Mechanistic Studies on Transcription Regulation by DksA and ppGpp

By

Albert Y. Chen

A dissertation submitted in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

(Microbiology)

at the

UNIVERSITY OF WISCONSIN-MADISON

2019

Date of final oral examination: 12/19/2018

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

Richard L. Gourse, Professor, Bacteriology
Karen M. Wassarman, Professor, Bacteriology
Robert C. Landick, Professor, Biochemistry
M. Thomas Record, Jr., Professor, Chemistry
Michael G. Thomas, Professor, Bacteriology

Mechanistic studies on transcription regulation by DksA and ppGpp

Albert Y. Chen

Under the supervision of Professor Richard L. Gourse

University of Wisconsin-Madison

Abstract

Bacteria need to alter their transcriptome quickly to respond to changes in their environment. In *E. coli*, the second messenger ppGpp and the transcription factor DksA regulate over 750 genes to help cells respond to various nutritional stresses. Neither ppGpp nor DksA bind to DNA. Instead, they act directly on RNAP, allosterically activating or inhibiting transcription of different genes. Understanding the conformational changes resulting from DksA and ppGpp interactions with RNAP is key to understanding the mechanism of regulation. I studied the ability of various RNAP variants to respond to DksA and ppGpp. Multiple mobile elements of RNAP are needed for activation and regulation by DksA and ppGpp, including the β' clamp, the trigger loop (TL), and β sequence insertion 1 (βSI1). In collaboration with structural biologists Andrey Feklistov and Seth Darst at Rockefeller, we determined that DksA and ppGpp result in significant movements of these elements that correlate with effects of mutations on RNAP function. Interaction of DksA with the TL favors clamp closure to favor melting of the transcription bubble. Movement of βSI1 by DksA and ppGpp widens the cleft and alters contacts to downstream DNA. The role of mobile elements has previously been difficult to characterize because they are hard to capture in crystal structures. These data provide a model for how DksA and ppGpp activate or inhibit transcription at different promoters. I also characterized a transcription intermediate in which the

transcription bubble is only partially melted. DNA footprinting and Bpa crosslinking data show that DksA/ppGpp can inhibit melting of the full transcription bubble at *rpsT* P2 but does not inhibit capture of the -11A base or melting of the upstream region. Furthermore, capture of the -11A base does not require DNA to be positioned near the active site within the cleft. Together, these data provide new insights into the mechanism of transcription initiation in bacteria and its regulation.

Acknowledgements

To Richard L. Gourse and Wilma Ross,

To my thesis committee members Robert C. Landick, M. Thomas Record, Jr.,

Michael G. Thomas, Jue D. Wang, and Karen M. Wassarman,

To members of the Gourse lab, past and present,

To my colleagues and collaborators,

To my fellow MDTP classmates,

To my parents,

To my friends and Journey family,

To the one who never gives up on me,

Thank you. I would not be here without you.

Table of Contents

List of Figures

Chapter 1 – Introduction

Chapter 2 – Stress regulation by reprogramming the conformational dynamics of RNA polymerase

Chapter 3 – DksA and TraR allow nucleation of promoter melting but inhibit melting post nucleation

List of Tables

Chapter 2 – Stress regulation by reprogramming the conformational dynamics of RNA polymerase

Chapter 3 – DksA and TraR allow nucleation of promoter melting but inhibit melting post nucleation

Appendix A – DksA weakens interactions between σ1.1 and the RNAP cleft

Appendix B – DksA tip residues D74 and A76 are required for a second step in transcription activation

Chapter 1

Introduction

This chapter provides an overview of the mechanism and regulation of transcription in bacteria and introduces the major topics discussed in this thesis.

Overview of transcription

Transcription is the essential cellular process by which DNA is transcribed into RNA. All organisms, from bacteria to archaea to eukaryotes, require this process for gene expression. The DNA-dependent RNA polymerase (RNAP) that carries out this process is conserved in bacteria, archaea, and in the three types of RNAP in eukaryotes (Lane and Darst, 2010). RNAPs in different species may contain variations, but all RNAPs have a conserved catalytic center and carry out a similar process.

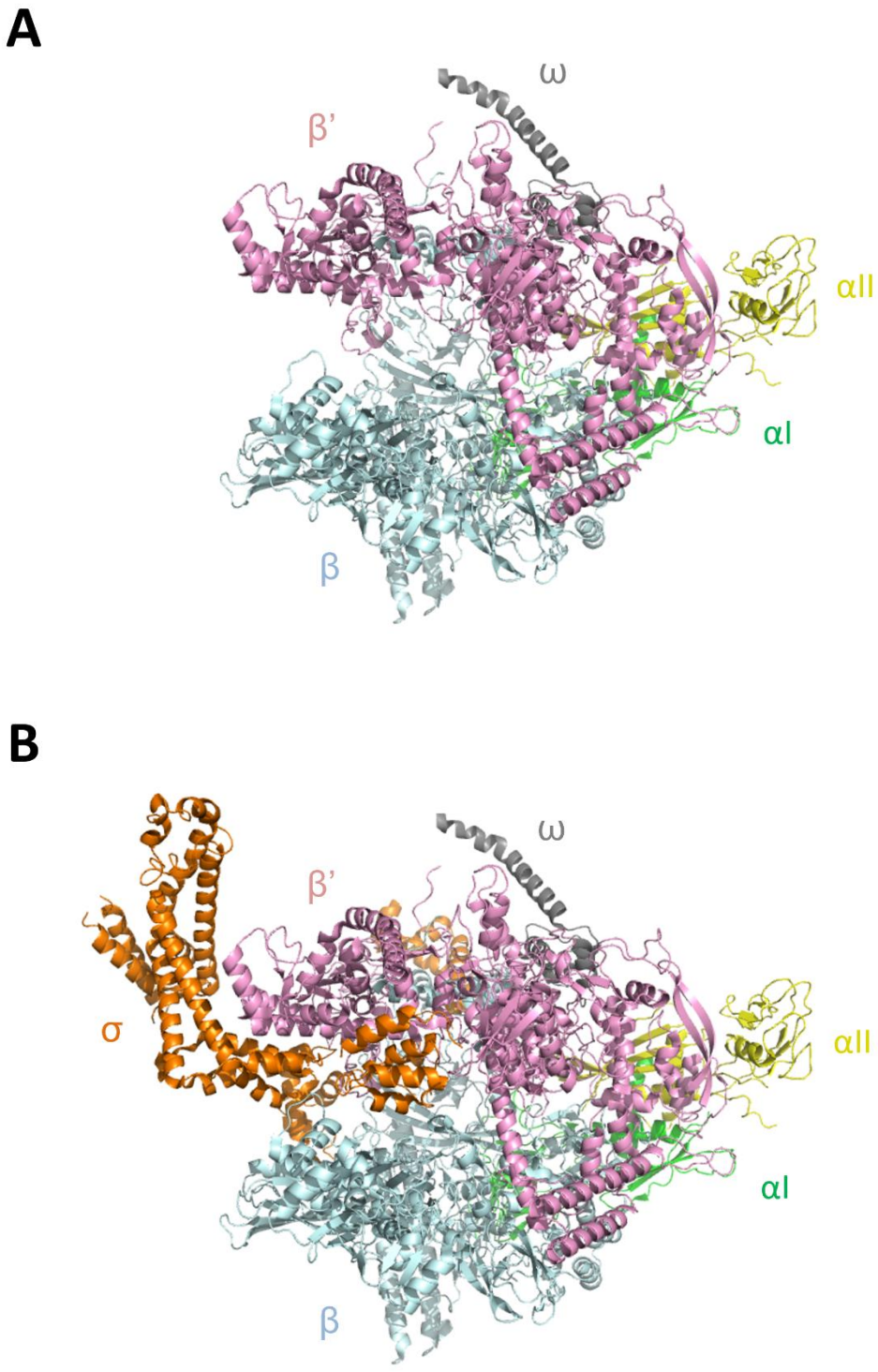
The bacterial RNAP core enzyme contains five subunits (Figure 1.1). Often described as a crab claw, the two large subunits, β and β', form the two pincers while the ω subunit and two subunits of α assemble at the base. The active site lies in the cleft, enclosed between these two pincers. Gram-positive bacteria have two additional subunits: the delta subunit involved in modulating gene expression and the epsilon subunit proposed to serve as protection from phage infection (Xue et al., 2010; Keller et al., 2014). While core RNAP is capable of RNA synthesis, transcription initiation from specific locations in DNA requires a dissociable subunit, σ. In *E. coli*, σ^{70} is the primary housekeeping sigma factor, directing transcription to genes during exponential growth. While *E. coli* has six other sigma factors, here I will focus mainly on transcription in *E. coli* from the primary sigma factor σ^{70}.

Transcription consists of three major steps: initiation, elongation, and termination. During transcription initiation, RNAP binds specifically to promoter DNA through DNA-binding domains in the σ subunit. The two strands of the transcription bubble must be separated, and the template strand DNA must be positioned in the active site in preparation for nucleotide addition. During transcription elongation, NTPs are added to

Figure 1.1. Structure of *E. coli* RNA Polymerase

Crystal structure of *E. coli* RNAP core **(A)** or holoenzyme **(B)** containing σ^{70} (PDB 4LK1; Bae et al., 2013). αI is colored in green, αII in yellow, β in cyan, β' in pink, ω in gray, and σ^{70} in orange.

Figure 1.1



the nascent RNA chain using the DNA as a template, until transcription is finally terminated.

Regulation of RNAP is critical for cell survival and adaptation to the cell's environment. An understanding of regulation requires an understanding of the mechanism of transcription. In this introduction, I will discuss the steps in transcription initiation and the different mobile regions in bacterial RNAP. I will also provide a general overview of transcription regulation and specifically its regulation by the second messenger ppGpp and its cofactor DksA.

Steps in transcription initiation

Transcription initiation begins with binding of RNAP to specific regions of DNA, called promoters. σ^{70} recognizes two major sequences in promoter DNA, the -35 element and the -10 element. Variation in sequence of these two hexamers, as well as the spacer length between the two hexamers, has large effects on promoter strength and activity (Siegele et al., 1989). Interactions between RNAP and other regions of the promoter also affect promoter activity, including the α subunits with an AT-rich UP element, σ region 3 with the extended -10 element, σ region 1.2 and the discriminator, and regions of the β subunit with the core recognition element (Ross et al., 1993; Barne et al., 1997; Haugen et al., 2006; Haugen et al., 2008; Zhang et al., 2012).

Upon binding of DNA to RNAP, RNAP forms a closed complex (RPc) in which DNA remains double-stranded. σ region 4.2 includes a helix-turn-helix motif, which inserts into the major groove of the -35 element, making base-specific contacts as well as contacts with the phosphate backbone (Campbell et al., 2002). These interactions

persist through transcription initiation. The bases in the A+T-rich sequences immediately upstream of the -35 element interact with the αCTD selectively, thereby increasing promoter binding and recruitment (Ross et al., 1993). These sequences can also make backbone contacts with RNAP, increasing promoter activity. Finally, sequences upstream of -60 can increase promoter activity sequence specifically and non-specifically (Davis et al., 2007; Ross et al., 2005). The interaction of σ region 2.4 with the -10 element does not appear to be sequence-specific in the closed complex, but as described below, base-specific interactions with the -10 element play a major role in DNA melting (Cook and deHaseth, 2007; Feklistov and Darst, 2011).

Once bound to DNA, RNAP undergoes a series of conformational changes to bend the DNA into the cleft, melt the transcription bubble, and stabilize DNA in the cleft to form a transcriptionally competent open complex (RPo). Bubble melting is initiated at the -11A base within the -10 element. The key tryptophan residue σ W433 invades the double-stranded helix, acting as a wedge to destabilize -11A from the helix (Tomsic et al., 2001). The unpaired -11A base is subsequently captured in a base-specific pocket in σ region 2.4 (Feklistov and Darst, 2011). Once DNA melting has initiated, other bases are captured in specific pockets to melt the transcription bubble from -11 to +2. Specifically, the highly conserved thymine base on the non-template strand at -7 is captured in a pocket in σ region 2.4 (Feklistov and Darst, 2011), a guanine at -6 or -5 at the upstream edge of the discriminator element is captured in σ region 1.2 (Haugen et al., 2006; Feklistov et al., 2006; Zhang et al., 2012), and a guanine at +2 in the core recognition element (CRE) is captured in a pocket in the β subunit (Zhang et al., 2012). Footprinting and FRET studies show that DNA upstream of the -35 element

wraps around RNAP early during transcription initiation (Davis et al., 2007; Sreenivasan et al., 2016), helping induce conformational changes to further facilitate DNA melting and open complex formation.

Though the mechanism responsible for inducing the 90° turn within the promoter DNA is unclear, this bend must be introduced in order for the template strand to be placed in the RNAP active site (Vassylyev et al., 2002). Perhaps -11A capture causes an increase in the flexibility of the downstream DNA, allowing it to bend into the cleft, or conversely DNA bending promotes unstacking of the -11A base. Footprinting of λP_R suggests that DNA melting does not occur until after DNA is bent into the active site cleft (Davis et al., 2007). The melted DNA bubble must then be stabilized in the main channel to form the final open complex. In that complex, mobile elements in RNAP, including the β' jaw, β' SI3, and the β lobe clamp down on the downstream DNA, burying ~2300 Å of amide surface area (Kontur et al., 2006; Ruff et al., 2015). These interactions will be discussed later in more detail.

Establishment of specific promoter DNA-RNAP interactions is critical for progression of RNAP through transcription initiation. As a result, variation in promoter DNA sequences leads to large differences in the rate of open complex formation and decay at each promoter. For instance, the open complex lifetime of λP_R is considerably longer (12.6 hours) than the open complex lifetime of *rrnB* P1 (<1s) (Henderson et al., 2017; Ruff et al., 2015). Converting the native T7A1 discriminator sequence to the λP_R sequence increases the open complex lifetime ~80-fold, similar to the lifetime of λP_R (Henderson et al., 2017). Similarly, a single mutation in the *rrnB* P1 discriminator sequence increases the open complex lifetime 37-fold (Haugen et al., 2006). Mutations

in other regions of the promoter affect different steps in transcription initiation (Hook-Barnard and Hinton, 2007; Cook and deHaseth, 2007). Furthermore, changes in the ability of RNAP to interact with DNA sequences also affects the rate of open complex formation (Artsimovitch et al., 2003; Drennan et al., 2012; Ruff et al., 2015). At the same promoter, RNAP from the proteobacterium *E. coli* and RNAP from the actinobacterium *Mycobacterium tuberculosis* have drastically different rates of open complex formation and decay due to differences in RNAP sequence (Davis et al., 2015; Hubin et al., 2017). Likewise, RNAP from the firmicute *Bacillus subtilis* also exhibits different transcription initiation kinetics compared to *E. coli* RNAP (Whipple and Sonenshein, 1992). A better understanding of how RNAP interacts with promoter DNA is critical to understand how transcription initiation occurs and how it is regulated.

X-ray crystal structures have provided significant information regarding the conformation of RNAP and the interactions made with DNA in the open complex (Murakami et al., 2002; Zhang et al., 2012; Bae et al., 2015). However, transcription initiation intermediates are transient and very unstable, making them difficult to characterize by x-ray crystallography. As a result, much of the information about the structures of the intermediates during the progression to the open complex has derived from footprinting data and fluorescence-based kinetic studies. The absence of crystal structures of the intermediates on the pathway to RPo formation has made it difficult to address many mechanistic questions. More recently, advancements in cryo-electron microscopy have provided new insights that will help to clarify the structures of these intermediates (Guo et al., 2018; Kang et al., 2018; Narayanan et al., 2018).

Mobile regions of RNAP

Except for σ^N-directed transcription, open complex formation occurs without any input of energy. Rather, it is driven solely by changes in binding free energy. Interfaces and conformations established in earlier intermediates trigger subsequent rearrangements to form later intermediates, ultimately leading to DNA opening and alignment of the start site base of the template strand with the catalytic site of the enzyme (Saecker et al., 2011). Flexible elements within RNAP play a large role in these conformational changes and thereby in transcription initiation (Figure 1.2).

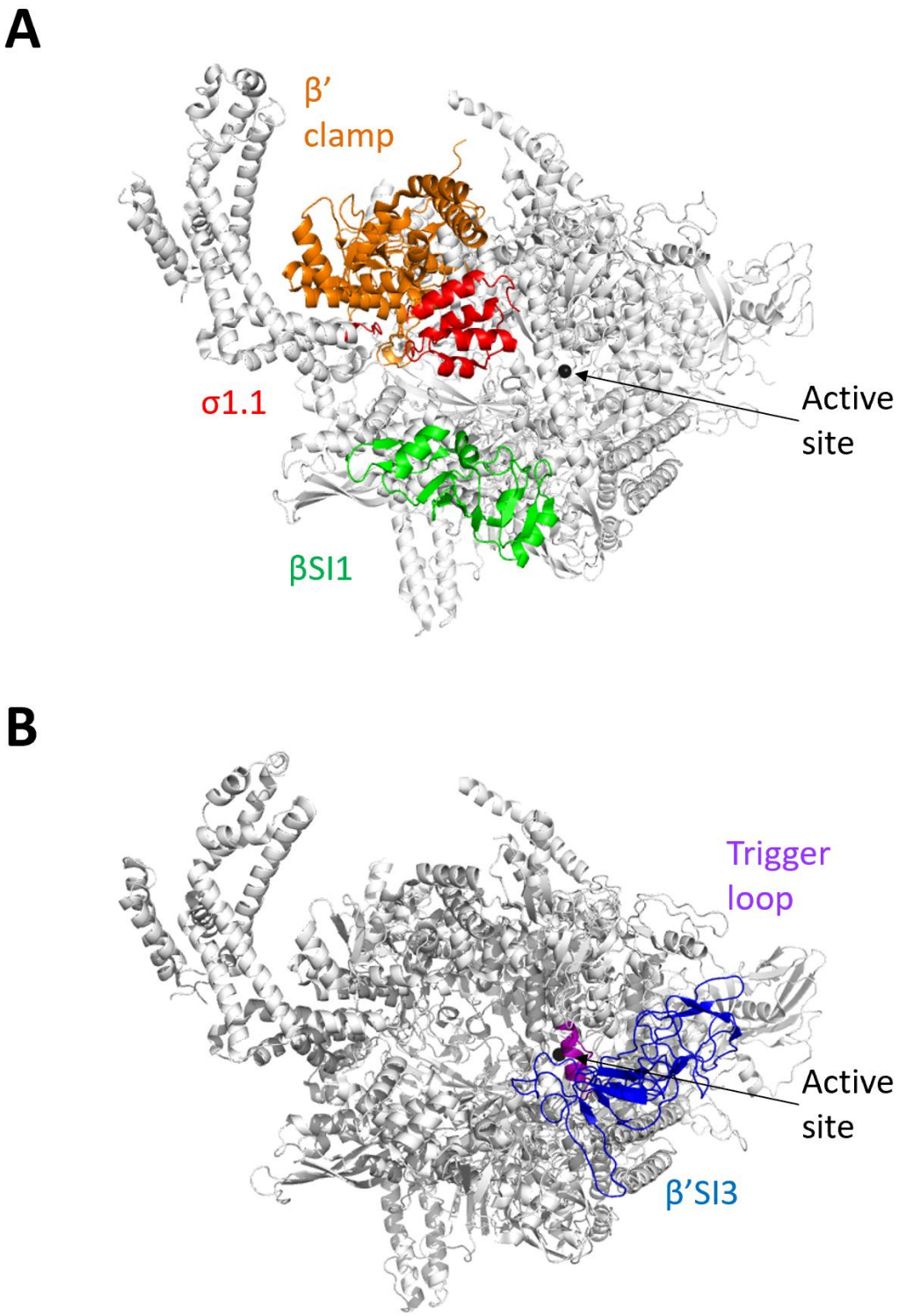
The β' clamp was one of the first mobile regions of RNAP discovered. Comparison of structures of yeast Pol II alone or in complex with DNA indicated that binding of DNA in the main channel resulted in closure of the β' clamp (Gnatt et al., 2001; Cramer et al., 2001). Upon binding to DNA, structures indicated that the clamp rotates by 30°, resulting in a 30 Å displacement of certain regions of RNAP. Clamp closure of bacterial RNAP was also detected in solution using FRET measurements (Chakraborty et al., 2012). During initiation, the clamp opens to allow DNA entry into the cleft, and subsequent closure is needed to stabilize DNA in the open complex and during processive transcription elongation (Chakraborty et al., 2012). More recent studies have also shown that clamp closure occurs early in initiation to allow -11A capture (Feklistov et al., 2017), and it was suggested that transient clamp opening and closure would allow DNA entry following -11A capture.

The clamp domain moves as a rigid body, but its movements are connected with conformational changes of the switch regions (Gnatt et al., 2001; Cramer et al., 2001). The switch regions consist of five linkers, many of which contact DNA. Binding of DNA

Figure 1.2. Mobile regions of RNAP

Mobile regions of interest are indicated on the *E. coli* RNAP structures. Model of RNAP shown from the same view as in Figure 1.1. **(A)** Crystal structure of RNAP holoenzyme showing the β' clamp in orange, σ1.1 in red, and βSI1 in green (PDB 4LK1; Bae et al., 2013). The trigger loop and β'SI3 are not resolved in this structure. **(B)** Crystal structure of RNAP initiation complex showing the trigger loop in purple and β'SI3 in blue (PDB 4YLN; Zuo and Steitz, 2015). DNA and RNA are hidden in this representation, and σ1.1 is not resolved in this structure.

Figure 1.2


A
β'
clamp
σ1.1
Active
site
βSI1
B
Trigger
loop
Active
site
β'SI3

alters the structures of these linkers, resulting in changes in the clamp state. The antibiotics myxopyronin, corallopyronin, and ripostatin bind to the switch regions to inhibit transcription (Mukhopadhyay et al., 2008), preventing DNA from binding to the active site and from interacting with the switch regions. By virtue of their location in the DNA-RNAP complex, these antibiotics also promote closure of the RNAP clamp (Chakraborty et al., 2012). In contrast, the antibiotic lipiarmycin has been reported to trap the clamp in the open state (Boyaci et al., 2018; Lin et al., 2018), preventing stabilization of DNA in the active site cleft.

The trigger loop (TL) is another key region of RNAP whose movements affect transcription. It alternates between a variety of unfolded states and a folded state, where it forms a three-helix bundle with the bridge helix to catalyze the nucleotide addition reaction (Vassylyev et al., 2007). Although its participation is key for all three phases of the transcription cycle, the precise role of the TL during transcription initiation is unclear (Fouqueau et al., 2013). However, its interaction with the bridge helix and its proximity to the switch regions suggests that the TL could affect clamp conformation.

σ region 1.1 goes through dramatic changes in position during transcription initiation. σ region 1.1 consists of a plug domain and an acidic linker, which acts like a DNA mimic (Schwartz et al., 2008), leading to auto-inhibition of Group 1 sigma factor activity, including *E. coli* σ^{70} (Gruber and Gross, 2003). That is, when σ is not associated with core RNAP, σ1.1 directly or indirectly masks σ region 4.2 and region 2.4 to prevent DNA binding (Dombroski et al., 1992; Schwartz et al., 2008). Upon association of σ with core RNAP, these regions are unmasked, and σ1.1 is placed into the cleft (Bae et al., 2013). During steps in transcription initiation, σ1.1 must be

displaced to allow DNA to enter the cleft. The mechanism behind σ1.1 displacement and the timing of this event is unclear, although FRET studies suggest that once σ1.1 is displaced, it is relocated to the tip of the β lobe (Mekler et al., 2002). This final state is likely dynamic, and the position of σ1.1 has not been visualized precisely by x-ray crystallography or cryo-EM in this state.

σ1.1 has large, promoter-dependent effects on different steps in transcription initiation. At λP_R, σ1.1 greatly increases the rate of open complex formation (Wilson and Dombroski, 1997). On the other hand, σ1.1 inhibits transcription and open complex formation from the weak promoter P_{minor} (Vuthoori et al., 2001). Since σ1.1 acts like a DNA mimic, some of its effects could be attributed to competition with DNA for the cleft. Additionally, σ1.1 is suggested to act through a complex network of interactions with downstream mobile elements including the β' jaw and β' sequence insertion 3, largely affecting the final stabilization of the open complex (Ruff et al., 2015). In summary, interactions of σ1.1 with the rest of the transcription complex are complex and dynamic, requiring further studies to understand how it affects transcription initiation.

E. coli RNAP contains several additional sequence insertions important for transcription initiation and its regulation, notably β' sequence insertion 3 (β'SI3) and β sequence insertion 1 (βSI1). β'SI3 (Artsimovitch et al., 2003), also known as β'i6 (Lane and Darst, 2010), is an insertion within the trigger loop that is present throughout the proteobacteria. Deletion of β'SI3 has large defects on open complex stability, abortive transcription, and cell growth (Artsimovitch et al., 2003). As a downstream mobile element, it participates in stabilization of DNA in the active site cleft during the final steps of open complex formation (Ruff et al., 2015). β'SI3 has also been shown to play

a key role in pausing and termination of elongating RNAPs (Windgassen et al., 2014; Artsimovitch et al., 2003).

βSI1 (Artsimovitch et al., 2003), also known as βi4 or β dispensable region 1 (Lane and Darst, 2010; Severinov et al., 1994), is an insertion within the β lobe domain of RNAP. Like β'SI3, it is distributed throughout the proteobacteria, although variations have been reported (Lane and Darst, 2010; Parshin et al., 2015). Unlike β'SI3, βSI1 is largely dispensable for basic functions in the transcription cycle, showing only minor effects on transcription initiation (Artsimovitch et al., 2003). It is physically unlikely to interact with DNA (Opalka et al., 2010). However, βSI1 is needed for growth under certain conditions, for example in minimal medium or after amino acid limitation, indicating that it may play an important role in regulation of transcription (Parshin et al., 2015). βSI1 has been found to be needed for the action of the bacteriophage T4 Alc protein, which redirects transcription from *E. coli* DNA to T4 phage DNA (Westblade et al., 2008; Snyder et al., 1976). As shown below, βSI1 also is required for function of ppGpp/DksA. In summary, interactions of RNAP mobile regions influence the progression of RNAP through transcription initiation.

Regulation of transcription initiation

Bacteria need to regulate gene expression in response to their environment. Although regulation undoubtedly occurs during all stages of gene expression, we understand regulation of transcription the best, and particularly at the level of transcription initiation. *E. coli* contains seven sigma factors, each recognizing a different set of genes. Although σ^{70} directs transcription of genes needed for products of general

metabolic pathways and genes needed for other housekeeping functions, the other six sigma factors primarily recognize sets of genes needed for responses to a variety of stresses, such as σ^E for extracytoplasmic stresses and σ^H for heat shock. The number of sigma factors in each bacterial species varies greatly. *Mycoplasma synoviae* only has one identified sigma factor (Vasconcelos et al., 2005) whereas *Streptomyces coelicolor* has more than 60 (Hahn et al., 2003), allowing it to direct transcription in response to a large number of inputs.

Generally, transcription factors can increase transcription of a gene by binding to DNA near the promoter to recruit RNAP to the promoter or to increase isomerization step(s) (reviewed in Browning and Busby, 2004). In contrast, if the binding site of the transcription factor overlaps the promoter, the transcription factor prevents RNAP from binding to the promoter and thereby inhibits transcription. There are also DNA-binding transcription factors that alter promoter DNA conformation in order to regulate transcription. MerR, for example, binds in the spacer region, inducing local unwinding of DNA to allow optimal RNAP binding (Ansari et al., 1992).

Some transcription factors, on the other hand, do not bind to DNA in a sequence-specific manner. CarD, for example, is an essential transcription factor in *Mycobacterium tuberculosis* that is found throughout the α-proteobacteria, actinobacteria, firmicutes, and other bacterial lineages (Stallings et al., 2009; Srivastava et al., 2013). It associates with β1 (sometimes called the protrusion) and stabilizes RPo through a conserved tryptophan residue inserted into the upstream edge of the bubble (Bae et al., 2015). Similarly, RbpA is an essential transcription factor in *Mtb* that increases transcription by interacting non-specifically with the DNA backbone and

specifically with the β' zinc-binding domain, β' zipper, and σ to prevent RPo collapse (Hubin et al., 2017). Specificity for the effects of these factors on transcription initiation derives from the kinetics of the promoter.

DksA, a secondary channel binding protein, and ppGpp, a small regulatory molecule, are two regulators that do not interact with DNA at all. These regulators are the focus of this thesis. As with CarD and RbpA, it has been difficult to correlate effects of DksA and ppGpp with specific DNA sequences. Instead, they exert their effects allosterically by changing the conformation of RNAP, affecting isomerization steps in transcription initiation.

Finally, small RNAs also serve as regulators of transcription. 6S RNA is a small RNA regulator that binds directly to RNAP, mimicking an open promoter complex (Wassarman and Storz, 2000; Chen et al., 2017; Wassarman, 2018). 6S RNA has been shown to regulate transcription of specific genes, contributing to the fitness and survival of bacteria in stationary phase (Cavanagh et al., 2008; Trotochaud and Wassarman, 2004). Together, protein, RNA, and small molecule regulators redirect transcription to allow bacteria to respond to specific internal and environmental signals.

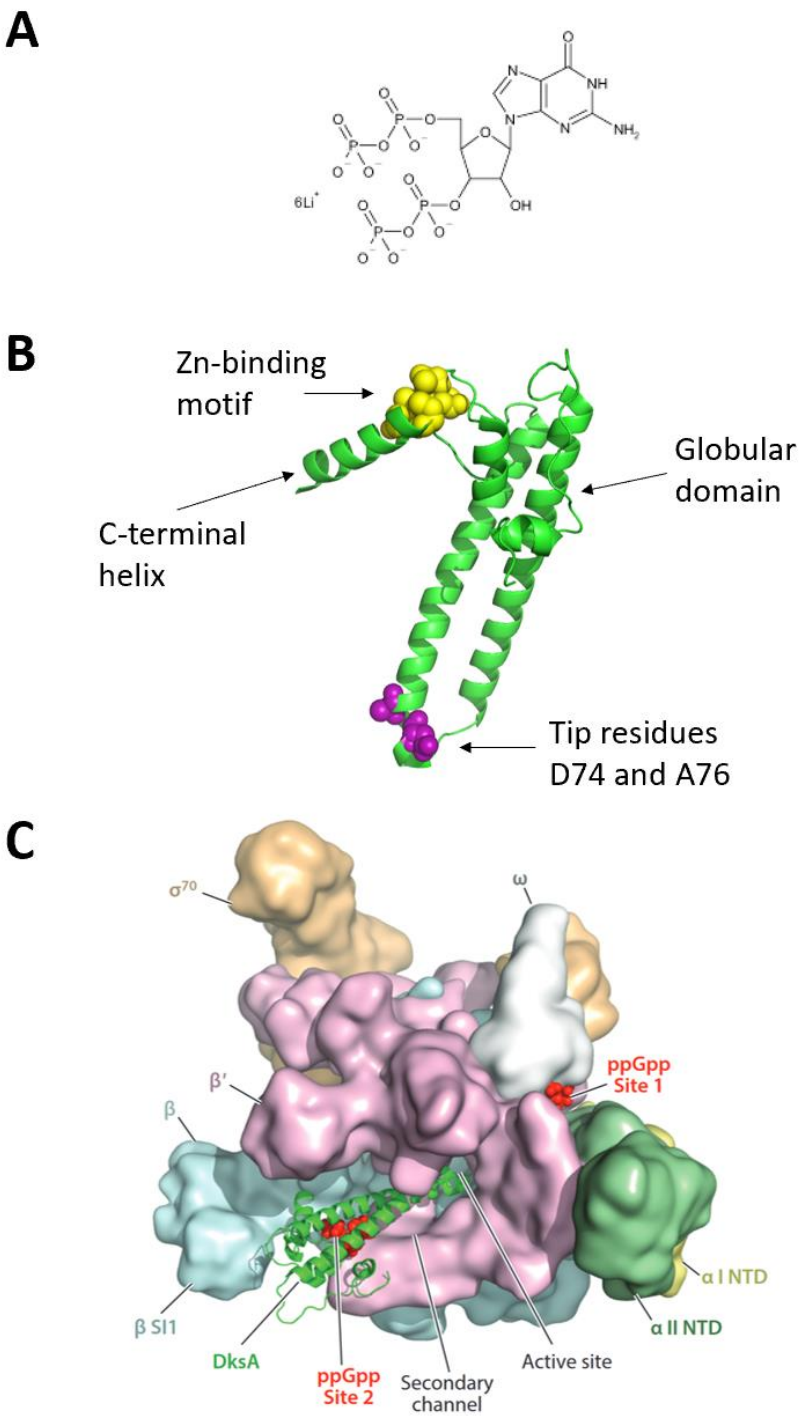
The global regulator ppGpp

The alarmone ppGpp is a global regulator that signals nutrient limitation in a process known as the stringent response (Cashel and Gallant, 1969; Figure 1.3A). Upon amino acid starvation, the major ppGpp synthase in *E. coli*, RelA, detects uncharged tRNAs in the ribosome A site and synthesizes pppGpp from ATP and GTP

Figure 1.3. ppGpp and DksA bind to RNAP

(A) Structure of ppGpp. **(B)** Crystal structure of DksA. Conserved tip residues D74 and A76 are shown in purple spheres. Cysteines in the Zn-binding motif are shown in yellow spheres (PDB: 1TJL; Perederina et al., 2004). **(C)** Location of DksA and ppGpp binding sites modeled onto RNAP. ppGpp is shown in red spheres. DksA is shown in green cartoon. Figure from Gourse et al., 2018.

Figure 1.3


A
B
Zn-binding motif
Globular domain
C-terminal helix
Tip residues D74 and A76
C
σ70
ω
β′
β
ppGpp Site 1
α I NTD
α II NTD
β SI1
DksA
ppGpp Site 2
Secondary channel
Active site

or ppGpp from ATP and GDP (Loveland et al., 2016; Haseltine and Block, 1973; Sy and Lipmann, 1973; Wendrich et al., 2002). The hydrolase GppA quickly converts pppGpp into ppGpp, the major form in *E. coli* (Hara and Sy, 1983; Mechold et al., 2013). For brevity, I will refer to the two molecules as ppGpp in this thesis. ppGpp levels can rise to near-millimolar concentrations (Varik et al., 2017), shutting down rRNA synthesis (Murray et al., 2003; Krasny and Gourse, 2004). Smaller increases in ppGpp lead to smaller decreases in rRNA synthesis, tuning ribosome synthesis of the bacterium to match nutrient availability (Ryals et al., 1982).

Effects of ppGpp go well beyond rRNA regulation. In *E. coli*, ppGpp affects transcription of over 700 genes, directly inhibiting genes involved in stable RNA synthesis, chemotaxis, and metabolism while activating genes involved in amino acid biosynthesis and stress response (Sanchez-Vazquez et al., submitted; Durfee et al., 2008; Traxler et al., 2008). ppGpp is also produced in response to iron limitation (Vinella et al., 2005), fatty acid limitation (Seyfzadeh et al., 1993), and phosphate limitation (Spira et al., 1995), adjusting transcription to adapt accordingly. *E. coli* contains an additional ppGpp synthase/hydrolase, SpoT, which produces ppGpp under these stresses, though its mechanism of action and regulation are much less clear.

relA is found in throughout the bacterial domain and even in plant chloroplasts (Potrykus and Cashel, 2008; Takahashi et al., 2004). In pathogens, ppGpp is used to control expression of virulence genes. In enterohaemorrhagic *E. coli* (EHEC), ppGpp is produced when EHEC enters the nutrient-poor large intestine, increasing expression of genes involved in attachment, colonization, and virulence (Nakanishi et al., 2006; Dalebroux et al., 2010). Similarly, when *Pseudomonas aeruginosa* enters the nutrient

poor, magnesium-depleted lungs of cystic fibrosis patients, ppGpp induction leads to increased production of quorum sensing molecules, biofilm formation, and colonization (Erickson et al., 2004; Dalebroux et al., 2010).

In *E. coli*, ppGpp regulates transcription by binding directly to two sites on RNAP (Figure 1.3C). The first one discovered, now called Site 1, is located at the interface between the ω subunit and the β' subunit of RNAP (Ross et al., 2013). Binding of ppGpp to this site was demonstrated genetically, biochemically, and structurally (Ross et al., 2013; Zuo et al., 2013; Molodtsov et al., 2018). Binding of ppGpp to this site, which is located near the hinge connecting the shelf and core domains of RNAP, reduced the lifetimes of open complexes formed by all promoters that were tested (Barker et al., 2001; Tagami et al., 2010; Ross et al., 2013). Substitutions in this site eliminated ppGpp-dependent inhibition *in vitro* and eliminated ppGpp binding as detected by Differential Radial Capillary Action of Ligand Assay (DRaCALA) (Ross et al., 2013). Furthermore, Site 1 mutants experienced a longer lag when cells were shifted from a richer to poorer medium (Ross et al., 2013). These results all supported the physiological relevance of binding of ppGpp to Site 1.

On the other hand, cells lacking Site 1 were not completely defective in shifting from rich to defined medium, like Δ*relAspoT* mutants (Ross et al., 2013). Additionally, ppGpp binding to this site was unable to activate transcription *in vitro* from a series of amino acid biosynthetic promoters unless the 151 amino acid protein DksA was also included in the reaction (Paul et al., 2005). Furthermore, elimination of Site 1 from RNAP was not sufficient to eliminate more than half the effect of ppGpp on transcription *in vitro* when DksA was included.

These results suggested that an additional binding site, now referred to as Site 2, is required for the full effects of ppGpp on transcription *in vitro* and *in vivo* (Ross et al., 2016). Using a combination of genetic, crosslinking, and other biochemical assays, it was determined that Site 2 was at an interface formed by the transcription factor DksA and the β' rim helices (Ross et al., 2016). Together, DksA and ppGpp fully inhibited and activated transcription to a similar extent as observed *in vivo* (Ross et al., 2016). Substitutions in both ppGpp binding sites on RNAP totally eliminated effects of ppGpp on transcription *in vivo*, and cells failed to recover from a downshift to minimal medium, the same phenotype as a *ΔrelAspoT* strain (Sanchez-Vazquez et al., submitted). The location of Site 2 was confirmed by stepwise soaking of DksA and then ppGpp into preformed RNAP crystals (Molodtsov et al., 2018). Although this approach confirmed the location of Site 2, RNAP packing within the crystals potentially constrained formation of contacts required for the full effects of ppGpp and DksA on transcription (discussed below).

ppGpp also binds directly to enzymes other than RNAP in *E. coli*. Purine salvage pathway enzymes including Gpt, Hpt, and GuaB are inhibited by ppGpp (Gallant et al., 1971; Hochstadt-Ozer and Cashel, 1972; Pao and Dyess, 1981). Translation factors including InfB, IF2, EF-G, RF3, and ObgE bind to ppGpp *in vitro*, and their activities may be inhibited by ppGpp *in vivo* (Milon et al., 2006; Mitkevich et al., 2010; Kihira et al., 2012; Persky et al., 2009). Recently, DRaCALA was performed on cell lysates, revealing ppGpp-binding to a total of twenty enzymes involved in nucleotide metabolism, ribosome biogenesis and translation, maturation of dehydrogenases, and metabolism of ppGpp (Zhang et al., 2018). Although ppGpp binds to multiple targets in

E. coli, its effects on transcription are completely dependent on its direct binding to RNAP: none of its other binding targets even indirectly accounted for the effects of ppGpp on the more than 750 genes that responded within 5 minutes to ppGpp induction *in vivo* (Sanchez-Vazquez et al., submitted).

ppGpp also regulates transcription in other bacteria, but its effects are indirect. Residues involved in ppGpp binding at Site 1 and Site 2 are not conserved in Firmicutes, including *Bacillus subtilis* (Ross et al., 2013; Ross et al., 2016). Instead, ppGpp directly inhibits enzymes like guanylate kinase (GMK) and HPRT (Kriel et al., 2012; Liu et al., 2015), decreasing the concentration of GTP, thereby indirectly inhibiting transcription from rRNA promoters (which all start with GTP in *B. subtilis*) (Krasny and Gourse, 2004). Thus, ppGpp regulates the concentration of GTP, thereby exerting a profound effect on *B. subtilis* physiology (Krasny and Gourse, 2004).

DksA and other secondary channel binding factors

As indicated above, DksA is a 151 amino acid transcription factor required for full effects of ppGpp on RNAP (Figure 1.3B). Like ppGpp, DksA destabilizes open complexes (Rutherford et al., 2009), but it is completely unable to activate transcription *in vitro,* and at the same concentrations where it is fully active when it works together with ppGpp, it has only partial effects on inhibition (Paul et al., 2004; Paul et al., 2005). Concentrations of DksA remain relatively constant throughout exponential growth and dip only ~ two-fold in stationary phase (Paul et al., 2004; Rutherford et al. 2007). Thus, the cell relies on changes in the ppGpp concentration not the DksA concentration to regulate transcription in response to nutritional stresses.

DksA's position in the secondary channel of RNAP was mapped by extensive genetic and crosslinking approaches (Lennon et al., 2012; Parshin et al., 2015). DksA is located at the interface between DksA and the rim helices (Ross et al., 2016), with the coiled-coil domain and globular domain docking on the β' rim helices located at the entrance of the secondary channel. These properties allow the DksA coiled-coil tip to approach the RNAP active site (Ross et al., 2016). An isoleucine substitution in DksA coiled-coil residue N88 or a phenylalanine substitution in globular domain residue L15 increase the affinity of DksA for the RNAP rim helices (Blankschien et al., 2009).

Unlike some other secondary channel factors that require β'SI3 for binding, DksA competes with β'SI3 for occupancy of the secondary channel. β'SI3 has been observed in multiple conformations depending on the presence of DNA/RNA in the main channel (Zuo and Steitz, 2015; Bae et al., 2015). Some conformations of β'SI3 are incompatible with DksA occupancy of the secondary channel. As a result, the affinity of DksA for RNAP is increased when RNAP is deleted for β'SI3 (Lennon et al., 2009), and DksA binding to RNAP is reduced when DNA is in the cleft, i.e. in an open complex (Lennon et al., 2009). Therefore β'SI3 may temporally regulate the binding of DksA during the transcription cycle.

The coiled-coil tip of DksA is cleaved by hydroxyl radicals generated by replacement of the active site magnesium in RNAP with an iron atom (Lennon et al., 2009), indicating that the coiled-coil tip is within ~10 Å of the active site. Substitutions for either D74 or A76, the most conserved residues in the DksA coiled-coil tip, abolish both the activation and inhibition functions of DksA (Lee et al., 2012). Modeling of DksA with RNAP suggested that DksA residue D74 could interact with active site residues βR1106

and/or βR678 (Parshin et al., 2015), although the role of this interaction in the DksA mechanism (and its modulation by ppGpp) is still unclear.

The globular domain of DksA contains a Zn-binding motif. Substitutions for any of its four cysteine residues abolish DksA activity (Paul et al., 2004; Gopalkrishnan et al., 2017). Although these residues are generally conserved in DksA homologs, there are examples that lack a Zn-binding motif. The DksA homolog from *Rhodobacter sphaeroides* lacks a Zn-binding motif, whereas the *Pseudomonas aeruginosa* contains two DksA homologs, one with and one without a Zn-binding motif (Lennon et al., 2014; Furman et al., 2013b). Both are structurally very similar to *E. coli* DksA. It has been proposed that the DksA without the Zn-binding motif performs DksA functions during Zn^{2+} starvation (Blaby-Haas et al., 2011; Furman et al., 2013b).

The C-terminal helix of DksA crosslinks to the βSI1 subdomain of RNAP, indicating they are in close proximity. Deletion of either βSI1 or the C-terminal helix of DksA reduces DksA binding to RNAP (Parshin et al., 2015; Furman et al., 2013a). Since ppGpp binds at the hinge of the C-terminal helix arm of DksA, making contacts with DksA residue K139, interactions between βSI1 and the C-terminal helix potentially affect ppGpp binding at Site 2 (Ross et al., 2016; Perederina et al., 2004). A co-crystal of DksA bound to RNAP showed a 4.5 Å shift of βSI1 toward the β' rim helices (Molodtsov et al., 2018), but this shift did not occur in the presence of ppGpp (Molodtsov et al., 2018). Either with or without ppGpp, the distance between DksA and βSI1 would be too large to account for the observed crosslink between (Parshin et al., 2015; also, see below). It is likely that crystal packing forces constrained movements in RNAP, thereby

preventing the DksA C-terminal helix interaction with βSI1 that occurs normally in solution and helps form ppGpp binding site 2.

DksA and ppGpp decrease open complex lifetime. At promoters with long-lived open complex lifetimes, there is little effect of ppGpp/DksA on transcriptional output, whereas transcription is inhibited from promoters like *rrnB* P1 with short-lived open complexes (Barker et al., 2001; Haugen et al., 2006). Suboptimal contacts between the cytosine at -7 of *rrnB* P1 (in the discriminator region) and σ region 1.2 correlate with regulation of this promoter *in vitro* and *in vivo* (Haugen et al., 2006).

Exactly how ppGpp and DksA inhibit promoter activity was unknown when I began my studies. ppGpp and DksA destabilize open complexes, and they clearly prevent RNAP from making contacts with DNA downstream of the transcription start site (Mekler et al., 2014). Suppressor mutations rescuing growth of a *ΔdksA* strain on minimal medium were identified in the switch regions and bridge helix, among other positions (Rutherford et al., 2009). These mutations could affect DNA binding directly and destabilize RPo or affect movements of the clamp domain. Like DksA and ppGpp, these mutations destabilized open complexes without preventing formation of earlier transcription intermediates (Rutherford et al., 2009). Interestingly, one of the substitutions in switch 3, βG1267V, made RNAP resistant to further effects of DksA (Rutherford et al., 2009), suggesting that switch 3 could be a key for transmitting effects of DksA and ppGpp to RNAP.

Similarly, not much was understood about the mechanism of activation by DksA and ppGpp when I began my work. Comparison of the promoter sequences from more than 30 promoters that were activated by ppGpp/DksA *in vitro* and *in vivo* did not

identify a highly-conserved motif for regulation (Sanchez-Vazquez et al., submitted). However, this analysis did show that promoters activated by DksA and ppGpp generally have nonconsensus -10 elements at the -9 and -8 positions (Barker et al., 2001; Sanchez-Vazquez et al., submitted), and mutations in the -10 element and/or in the discriminator region eliminated activation (Riggs et al., 1986; Da Costa and Artz, 1997). Promoters activated by DksA and ppGpp tend to have A+T-rich discriminator sequences (Barker et al., 2001; Gummesson et al., 2013). Multiple acceptable sequence combinations for activation and context effects complicate these sequence comparisons.

Kinetic studies of the arginine biosynthetic promoter *argI* showed that DksA and ppGpp have little if any effect on the initial binding step K_B, but rather increase transcription output by increasing the composite isomerization step k_f (Paul et al., 2005). Further studies are in progress to address how DksA and ppGpp increase the rate of isomerization and which specific isomerization step(s) is(are) rate-limiting at these promoters. Although it is now clear that both DksA and ppGpp are needed for activation because ppGpp binding site 2 is responsible, the role of Site 2 in inhibition remains unclear as does the mechanism of inhibition by ppGpp bound to Site 1.

Previous studies suggest that the trigger loop plays a role in DksA function. DksA did not destabilize open complexes formed with an RNAP lacking the trigger loop (Rutherford et al., 2009). Disulfide bonds were formed between cysteines engineered into both the trigger loop and the coiled-coil tip region of DksA (Lennon et al., 2012), suggesting that DksA is in close proximity to the trigger loop and interacts with it. DksA can promote partial folding of the trigger loop (Nayak et al., 2013; Lennon et al., 2012),

but the flexible nature of the trigger loop makes it difficult to visualize the structure of the trigger loop in the presence of DksA.

DksA homologs are found predominantly among the proteobacteria. Key residues are generally conserved, including the coiled-coil tip residues, residues involved in binding RNAP, and residues involved in binding ppGpp (Ross et al., 2016; Parshin et al., 2015), though DksA in certain species may lack the Zn-binding motif or other features (Furman et al., 2013b; Lennon et al., 2014). Proteobacterial DksA homologs that have been tested generally function similarly to regulate RNAP. *Rhodobacter sphareoides* DksA, for instance, is 42% identical to *E. coli* DksA, lacks the Zn-binding motif, and inhibits and activates transcription like *E. coli* DksA (Lennon et al., 2014). In non-proteobacterial cases that have been tested, ppGpp predominantly regulates transcription indirectly by affecting nucleotide pools rather than by directly interacting with RNAP (Liu et al., 2015). Although DksA-like proteins have been reported beyond the proteobacteria, many differ significantly in sequence from the proteobacterial DksA and have not been tested for function (Parshin et al., 2015).

TraR is an F plasmid-encoded protein 73 residues long with significant identity to the C-terminal half of DksA (Blankschien et al., 2009), and it binds to the RNAP secondary channel and regulates gene expression like DksA and ppGpp together (Gopalkrishnan et al., 2017). TraR contains residues corresponding to only one alpha helix of the DksA coiled-coil, but it does contain the two aspartates and the alanine corresponding to D71, D74, and A76 in DksA, as well as a Zn-binding motif. TraR lacks the residues that contribute to the ppGpp binding site in DksA, and RNAP-TraR complexes do not bind ppGpp. Instead, the residues in RNAP that interact with ppGpp

in the RNAP-DksA complex bind directly to TraR. TraR activates and inhibits transcription like DksA and ppGpp together (Gopalkrishnan et al., 2017).

Gre factors represent another secondary channel binding factor class that share no sequence homology with DksA (Perederina et al., 2004), although they share some functions with DksA. GreA and GreB are transcription factors that can function during promoter escape or during elongation. The acidic tip of the Gre factors participates in RNA cleavage and rescue of paused or stalled complexes (Borukhov et al., 2005; Laptenko et al., 2003). GreB and DksA are both able to inhibit transcription from *rrnB* P1, but GreB does not contain the residues analogous to those in DksA that help form Site 2. Thus, in retrospect it is apparent why GreB and ppGpp are not able to synergize to activate transcription (Rutherford et al., 2007; Lee et al., 2012; Ross et al., 2016). Interestingly, substitutions in the acidic tip of GreB analogous to those in DksA allow the chimeric GreB to activate transcription (Lee et al., 2012). Although this result indicates the importance of the coiled-coil tip in activation, its molecular basis remains unclear.

Outline of Thesis

The primary objective of this thesis research was to gain insights into the mechanism of how DksA and ppGpp regulate transcription. The two main approaches were to determine how DksA and ppGpp affect the conformation of mobile regions of RNAP and to determine how DksA and ppGpp affect interactions between RNAP and promoter DNA. I identified regions of RNAP important for regulation by DksA and ppGpp, allowing us to better understand how binding of DksA and ppGpp allosterically regulates RNAP activity at different promoters. I also characterized a transcription

intermediate enriched in the presence of DksA and ppGpp, providing insight into the mechanism of transcription initiation as well as the mechanism of regulation by DksA and ppGpp.

In Chapter 2, I identified mobile elements in RNAP important for regulation by DksA and ppGpp. I made substitutions and deletions in these elements to demonstrate their requirement in different steps of transcription. In conjunction with cryo-EM structures of RNAP bound to DksA or to both DksA and ppGpp obtained by our collaborators Andrey Feklistov and Seth Darst at Rockefeller University, our data show how DksA and ppGpp interact with the trigger loop and affect the clamp domain to increase strand opening. In addition, interaction of DksA with βSI1 repositions the β lobe, widening the active site cleft and altering contacts between RNAP and downstream DNA. These data provide an explanation for how DksA and ppGpp directly activate or inhibit different promoters.

In Chapter 3, I characterize a partially melted transcription intermediate formed in the presence of DksA at the S20 ribosomal protein promoter, *rpsT* P2. This work built on the discovery made by Jared Winkelman and Michael Maloney that DksA inhibits the formation of the open complex at *rpsT* P2, resulting in an intermediate in which only the first two positions of the -10 element are melted. This intermediate is distinct both from the initial bound closed complex as well as the final open complex. I determined the path of DNA on RNAP using Bpa crosslinking to understand the interactions between RNAP and DNA in this intermediate. My data are consistent with and inform the cryo-EM structures of the RNAP-TraR-*rpsT* P2 complex from James Chen, Elizabeth Campbell, and Seth Darst at Rockefeller University. Together, these data show that

DksA inhibits melting of downstream DNA at the *rpsT* P2 promoter but does not inhibit the initial nucleation. Furthermore, capture of base -11A within the -10 element to nucleate melting does not require downstream DNA to be bent into the active site cleft.

References

Ansari, A.Z., Chael, M.L., and O'Halloran, T.V. (1992). Allosteric underwinding of DNA is a critical step in positive control of transcription by Hg-MerR. *Nature* **355**:87-89.

Artsimovitch, I., Svetlov, V., Murakami, K.S., and Landick, R. (2003). Co-overexpression of Escherichia coli RNA polymerase subunits allows isolation and analysis of mutant enzymes lacking lineage-specific sequence insertions. *J. Biol. Chem.* **278**:12344-12355.

Bae, B., Davis, E., Brown, D., Campbell, E.A., Wigneshweraraj, S., and Darst, S.A. (2013). Phage T7 Gp2 inhibition of Escherichia coli RNA polymerase involves misappropriation of σ70 domain 1.1. *Proc. Natl. Acad. Sci. USA.* **110**:19772-19777.

Bae, B., Feklistov, A., Lass-Napiorkowska, A., Landick, R., and Darst, S.A. (2015). Structure of a bacterial RNA polymerase holoenzyme open promoter complex. *Elife* **4**:e08504.

Barker, M.M., Gaal, T., Josaitis, C.A., and Gourse, R.L. (2001). Mechanism of regulation of transcription initiation by ppGpp. I. Effects of ppGpp on transcription initiation in vivo and in vitro. *J. Mol. Biol.* **305**:673-688.

Barne, K.A., Brown, J.A., Busby, S.J., and S.D. Minchin. (1997). Region 2.5 of Escherichia coli RNA polymerase sigma70 subunit is responsible for the recognition of the 'extended-10' motif at promoters. *EMBO J*. **16**:4034-4040.

Blaby-Haas, C.E., Furman, R., Rodionov, D.A., Artsimovitch, I., and de Crécy-Lagard, V. (2011). Role of a Zn-independent DksA in Zn homeostasis and stringent response. *Mol. Microbiol.* **79**:700-715.

Blankschien, M.D., Lee, J.H., Grace, E.D., Lennon, C.W., Halliday, J.A., Ross, W., Gourse, R.L., and Herman, C. (2009). Super DksAs: substitutions in DksA enhancing its effects on transcription initiation. *EMBO J.* **28**:1720-1731.

Borukhov, S., Lee, J., and Laptenko, O. (2005). Bacterial transcription elongation factors: new insights into molecular mechanism of action. *Mol. Microbiol.* **55**:1315-1324.

Boyaci, H., Chen, J., Lilic, M., Palka, M., Mooney, R.A., Landick, R., Darst, S.A., and Campbell, E.A. (2018). Fidaxomicin jams *Mycobacterium tuberculosis* RNA polymerase motions needed for initiation via RbpA contacts. *Elife* **7**:e34823.

Browning, D.F. and Busby, S.F. (2004). The regulation of bacterial transcription initiation. *Nat. Rev. Microbiol.* **2**:57-65.

Campbell, E.A., Muzzin, O., Chlenov, M., Sun, J.L., Olson, C.A., Weinman, O., Trester-Zedlitz, M.L., and Darst, S.A. (2002). Structure of the bacterial RNA polymerase promoter specificity sigma subunit. *Mol. Cell* **9**:527-539.

Cashel, M. and Gallant, J. (1969). Two compounds implicated in the function of the RC gene of Escherichia coli. *Nature* **221**:838-841.

Cavanagh, A.T., Klocko, A.D., Liu, X., and Wassarman, K.M. (2008). Promoter specificity for 6S RNA regulation of transcription is determined by core promoter sequences and competition for region 4.2 of sigma70. *Mol. Microbiol.* **67**:1242-1256.

Chakraborty, A., Wang, D., Ebright, Y.W., Korlann, Y., Kortkhonjia, E., Kim, T., Chowdhury, S., Wigneshweraraj, S., Irschik, H., Jansen, R., Nixon, B.T., Knight, J., Weiss, S., and Ebright, R.H. (2012). Opening and closing of the bacterial RNA polymerase clamp. *Science* **337**:51-595.

Chen, J., Wassarman, K.M., Feng, S., Leon, K., Feklistov, A., Winkelman, J.T., Li, Z., Walz, T., Campbell, E.A., and Darst, S.A. (2017). 6S RNA Mimics B-Form DNA to Regulate Escherichia coli RNA Polymerase. *Mol. Cell* **68**:388-397.e6.

Cook, V.M. and deHaseth, P.L. (2007). Strand opening-deficient Escherichia coli RNA polymerase facilitates investigation of closed complexes with promoter DNA: effects of DNA sequence and temperature. *J. Biol. Chem.* **282**:21319-21326.

Cramer, P., Bushnell, D.A., and Kornberg, R.D. (2001). Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. *Science* **292**:1863-1876.

Da Costa, X.J. and Artz, S.W. (1997). Mutations that render the promoter of the histidine operon of Salmonella typhimurium insensitive to nutrient-rich medium repression and amino acid downshift. *J. Bacteriol.* **179**:5211-5217.

Dalebroux, Z.D., Svensson, S.L., Gaynor, E.C., and Swanson, M.S. (2010). ppGpp conjures bacterial virulence. *Microbiol. Mol. Biol. Rev.* **74**:171-199.

Davis, C.A., Bingman, C.A., Landick, R., Record, M.T. Jr., and Saecker, R.M. (2007). Real-time footprinting of DNA in the first kinetically significant intermediate in open complex formation by Escherichia coli RNA polymerase. *Proc. Natl. Acad. Sci. USA.* **104**:7833-7838.

Davis, E., Chen, J., Leon, K., Darst, S.A., and Campbell, E.A. (2015). Mycobacterial RNA polymerase forms unstable open promoter complexes that are stabilized by CarD. *Nucleic Acids Res.* **43**:433-445.

Dombroski, A.J., Walter, W.A., Record, M.T. Jr., Siegele, D.A., and Gross, C.A. (1992). Polypeptides containing highly conserved regions of transcription initiation factor sigma 70 exhibit specificity of binding to promoter DNA. *Cell* **70**:501-512.

Drennan, A., Kraemer, M., Capp, M., Gries, T., Ruff, E., Sheppard, C., Wigneshweraraj, S., Artismovitch, I., and Record, M.T. Jr. (2012). Key roles of the downstream mobile jaw of Escherichia coli RNA polymerase in transcription initiation. *Biochemistry* **51**:9447-9459.

Durfee, T., Hansen, A.M., Zhi, H., Blattner, F.R., and Jin, D.J. (2008). Transcription profiling of the stringent response in Escherichia coli. *J. Bacteriol.* **190**:1084-1096.

Erickson, D.L., Lines, J.L., Pesci, E.C., Venturi, V., and Storey, D.G. (2004). Pseudomonas aeruginosa relA contributes to virulence in Drosophila melanogaster. *Infect. Immun.* **72**:5638-5645.

Feklistov, A., Barinova, N., Sevostyanova, A., Heyduk, E., Bass, I., Vvedenskaya, I., Kuznedelov, K., Merkiene, E., Stavrovskaya, E., Klimasauskas, S., Nikiforov, V., Heyduk, T., Severinov, K., and Kulbachinskiy, A. (2006). A basal promoter element recognized by free RNA polymerase sigma subunit determines promoter recognition by RNA polymerase holoenzyme. *Mol. Cell* **23**:97-107.

Feklistov, A. and Darst, S.A. (2011). Structural basis for promoter-10 element recognition by the bacterial RNA polymerase σ subunit. *Cell* **147**:1257-1269.

Feklistov, A., Bae, B., Hauver, J., Lass-Napiorkowska, A., Kalesse, M., Glaus, F., Altmann, K.H., Heyduk, T., Landick, R., and Darst, S.A. (2017). RNA polymerase motions during promoter melting. *Science* **356**:863-866.

Fouqueau, T., Zeller, M.E., Cheung, A.C., Cramer, P., and Thomm, M. (2013). The RNA polymerase trigger loop functions in all three phases of the transcription cycle. *Nucleic Acids Res.* **41**:7048-7059.

Furman, R., Tsodikov, O.V., Wolf, Y.I., and Artsimovitch, I. (2013a). An insertion in the catalytic trigger loop gates the secondary channel of RNA polymerase. *J. Mol. Biol.* **425**:82-93.

Furman, R., Biswas, T., Danhart, E.M., Foster, M.P., Tsodikov, O.V., and Artsimovitch, I. (2013b). DksA2, a zinc-independent structural analog of the transcription factor DksA. *FEBS Lett.* **587**:614-619.

Gallant, J., Irr, J., and Cashel, M. (1971). The mechanism of amino acid control of guanylate and adenylate biosynthesis. *J. Biol. Chem.* **246**:5812-5816.

Gnatt, A.L., Cramer, P., Fu, J., Bushnell, D.A., and Kornberg, R.D. (2001). Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 A resolution. *Science* **292**:1876-1882.

Gopalkrishnan, S., Ross, W., Chen, A.Y., and Gourse, R.L. (2017). TraR directly regulates transcription initiation by mimicking the combined effects of the global regulators DksA and ppGpp. *Proc. Natl. Acad. Sci. USA.* **114**:E5539-E5548.

Gourse, R.L., Chen, A.Y., Gopalkrishnan, S., Sanchez-Vazquez, P., Myers, A., and Ross, W. (2018). Transcriptional Responses to ppGpp to DksA. *Annu. Rev. Microbiol.* **72**:163-184.

Gruber, T.M. and Gross, C.A. (2003). Multiple sigma subunits and the partitioning of bacterial transcription space. *Annu. Rev. Microbiol.* **57**:441-466.

Gummesson, B., Lovmar, M., and Nyström, T. (2013). A proximal promoter element required for positive transcriptional control by guanosine tetraphosphate and DksA protein during the stringent response. *J. Biol. Chem.* **288**:21055-21064.

Guo, X., Myasnikov, A.G., Chen, J., Crucifix, C., Papai, G., Takacs, M., Schultz, P., and Weixlbaumer, A. (2018). Structural Basis for NusA Stabilized Transcriptional Pausing. *Mol. Cell* **69**:816-827.

Hahn, M.Y., Bae, J.B., Park, J.H., and Roe, J.H. (2003). Isolation and characterization of Streptomyces coelicolor RNA polymerase, its sigma, and antisigma factors. *Methods Enzymol.* **370**:73-82.

Hara, A. and Sy, J. (1983). Guanosine 5'-triphosphate, 3'-diphosphate 5'-phosphohydrolase. Purification and substrate specificity. *J. Biol. Chem.* **258**:1678-1683.

Haseltine, W.A. and Block, R. (1973). Synthesis of guanosine tetra- and pentaphosphate requires the presence of a codon-specific, uncharged transfer ribonucleic acid in the acceptor site of ribosomes. *Proc. Natl. Acad. Sci. USA.* **70**:1564-1568.

Haugen, S.P., Berkmen, M.B., Ross, W., Gaal, T., Ward, C., and Gourse, R.L. (2006). rRNA promoter regulation by nonoptimal binding of sigma region 1.2: an additional recognition element for RNA polymerase. *Cell* **125**:1069-1082.

Haugen, S.P., Ross, W., Manrique, M., and Gourse, R.L. (2008). Fine structure of the promoter-sigma region 1.2 interaction. *Proc. Natl. Acad. Sci. USA.* **105**:3292-3297.

Henderson, K.L., Felth, L.C., Molzahn, C.M., Shkel, I., Wang, S., Chhabra, M., Ruff, E.F., Bieter, L., Kraft, J.E., and Record, M.T. Jr. (2017). Mechanism of transcription initiation and promoter escape by *E. coli* RNA polymerase. *Proc. Natl. Acad. Sci. USA.* **114**:E3032-E3040.

Hochstadt-Ozer, J., and Cashel, M. (1972). The regulation of purine utilization in bacteria. V. Inhibition of purine phosphoribosyltransferase activities and purine update in isolated membrane vesicles by guanosine tetraphosphate. *J. Biol. Chem.* **247**:7067-7072.

Hook-Barnard, I.G. and Hinton, D.M. (2007). Transcription initiation by mix and match elements: flexibility for polymerase binding to bacterial promoters. *Gene Regul. Syst. Bio.* **1**:275-293.

Hubin, E.A., Fay, A., Xu, C., Bean, J.M., Saecker, R.M., Glickman, M.S., Darst, S.A., and Campbell, E.A. (2017). Structure and function of the mycobacterial transcription initiation complex with the essential regulator RbpA. *Elife* **6**:e22520.

Kang, J.Y., Mishanina, T.V., Bellecourt, M.J., Mooney, R.A., Darst, S.A., and Landick, R. (2018). RNA Polymerase Accomodates a Pause RNA Hairpin by Global Conformational Rearrangements that Prolong Pausing. *Mol. Cell* **69**:802-815.E1.

Keller, A.N., Yang, X., Wiedermannová, J., Delumeau, O., Krásný, L., and Lewis, P.J. (2014). ε, a new subunit of RNA polymerase found in gram-positive bacteria. *J bacteriol.* **196**:3622-3632.

Kihira, K., Shimizu, Y., Shomura, Y., Shibata, N., Kitamura, M., Nakagawa, A., Ueda, T., Ochi, K., and Higuchi, Y. (2012). Crystal structure analysis of the translation factor RF3 (release factor 3). *FEBS Lett.* **586**:3705-3709.

Kontur, W.S., Saecker, R.M., Davis, C.A., Capp, M.W., and Record, M.T. Jr. (2006). Solute probes of conformational changes in open complex (RPo) formation by Escherichia coli RNA polymerase at the lambdaPR promoter: evidence for unmasking of the active site in the isomerization step and for large-scale coupled folding in the subsequent conversion to RPo. *Biochemistry* **45**:2161-2177.

Krásný, L. and Gourse, R.L. (2004). An alternative strategy for bacterial ribosome synthesis: Bacillus subtilis rRNA transcription regulation. *EMBO J.* **23**:4473-4483.

Kriel, A., Bittner, A.N., Kim, S.H., Liu, K., Tehranchi, A.K., Zou, W.Y., Rendon, S., Chen, R., Tu, B.P., and Wang, J.D. (2012). Direct regulation of GTP homeostasis by (p)ppGpp: a critical component of viability and stress resistance. *Mol. Cell* **48**:231-241.

Lane, W.J. and Darst, S.A. (2010). Molecular evolution of multisubunit RNA polymerases: sequence analysis. *J. Mol. Biol.* **395**:671-685.

Laptenko, O., Lee, J., Lomakin, I., and Borukhov, S. (2003). Transcript cleavage factors GreA and GreB act as transient catalytic components of RNA polymerase. *EMBO J.* **22**:6322-6334.

Lee, J.H., Lennon, C.W., Ross, W., and Gourse, R.L. (2012). Role of the coiled-coil tip of Escherichia coli DksA in promoter control. *J. Mol. Biol.* **416**:503-517.

Lennon, C.W., Gaal, T., Ross, W., and Gourse, R.L. (2009). Escherichia coli DksA binds to Free RNA polymerase with higher affinity than to RNA polymerase in an open complex. *J. Bacteriol.* **191**:5854-5858.

Lennon, C.W., Ross, W., Martin-Tumasz, S., Toulokhonov, I., Vrentas, C.E., Rutherford, S.T., Lee, J.H., Butcher, S.E., and Gourse, R.L. (2012). Direct interactions between the coiled-coil tip of DksA and the trigger loop of RNA polymerase mediates transcriptional regulation. *Genes Dev.* **26**:2634-2646.

Lennon, C.W., Lemmer, K.C., Irons, J.L., Sellman, M.I., Donohue, T.J., Gourse, R.L., and Ross, W. (2014). A Rhodobacter sphaeroides protein mechanistically similar to Escherichia coli DksA regulates photosynthetic growth. *MBio* **5**:e01105-e01114.

Lin, W., Das, K., Degen, D., Mazumder, A., Duchi, D., Wang, D., Ebright, Y.W., Ebright, R.Y., Sineva, E., Gigliotti, M., Srivastava, A., Mandal., S., Jiang, Y., Liu, Y., Yin, R., Zhang, Z., Eng, E.T., Thomas, D., Donadio, S., Zhang, H., Zhang, C., Kapanidis, A.N., and Ebright, R.H. (2018). Structural Basis of Transcription Inhibition by Fidaxomicin (Lipiarmycin A3). *Mol. Cell* **70**:60-71.e15.

Liu, K., Myers, A.R., Pisithkul, T., Claas, K.R., Satyshur, K.A., Amador-Noguez, D., Keck, J.L., and Wang, J.D. (2015). Molecular mechanism and evolution of guanylate kinase regulation by (p)ppGpp. *Mol. Cell* **57**:735-749.

Loveland, A.B., Bah, E., Madireddy, R., Zhang, Y., Brilot, A.F., Grigorieff, N., and Korostelev, A.A. (2016). Ribosome·RelA structures reveal the mechanism of stringent response activation. *Elife* **5**:e17029.

Mechold, U., Potrykus, K., Murphy, H., Murakami, K.S., and Cashel, M. (2013). Differential regulation by ppGpp versus pppGpp in Escherichia coli. *Nucleic Acids Res.* **41**:6175-6189.

Mekler, V., Kortkhonjia, E., Mukhopadhyay, J., Knight, J., Revyakin, A., Kapanidis, A.N., Niu, W., Ebright, Y.W., Levy, R., and Ebright, R.H. (2002). Structural organization of bacterial RNA polymerase holoenzyme and the RNA polymerase-promoter open complex. *Cell* **108**:599-614.

Mekler, V., Minakhin, L., Borukhov, S., Mustaev, A., and Severinov, K. (2014). Coupling of downstream RNA polymerase-promoter interactions with formation of catalytically competent transcription initiation complex. *J. Mol. Biol.* **426**:3973-3984.

Milon, P., Tischenko, E., Tomsic, J., Caserta, E., Folkers, G., La Teana, A., Rodnina, M.V., Pon, C.L., Boelens, R., and Gualerzi, C.O. (2006). The nucleotide-binding site of bacterial translation initiation factor 2 (IF2) as a metabolic sensor. *Proc. Natl. Acad. Sci. USA.* **103**:13962-13967.

Mitkevich, V.A., Ermakov, A., Kulikova, A.A., Tankov, S., Shyp, V., Soosaar, A., Tenson, T., Makarov, A.A., Ehrenberg, M., and Hauryliuk, V. (2010). Thermodynamic characterization of ppGpp binding to EF-G or IF2 and of initiator tRNA binding to free IF2 in the presence of GDP, GTP, or ppGpp. *J. Mol. Biol.* **402**:838-846.

Molodtsov, V., Sineva, E., Zhang, L., Huang, X., Cashel, M., Ades, S.E., and Murakami, K.S. (2018). Allosteric Effector ppGpp Potentiates the Inhibition of Transcript Initiation by DksA. *Mol. Cell* **69**:828-839.e5.

Mukhopadhyay, J., Das, K., Ismail, S., Koppstein, D., Jang, M., Hudson, B., Sarafianos, S., Tuske, S., Patel, J., Jansen, R., Irschik, H., Arnold, E., and Ebright, R.H. (2008). The RNA polymerase “switch region is a target for inhibitors. *Cell* **135**:295-307.

Murakami, K.S., Masuda, S., Campbell, E.A., Muzzin, O., and Darst, S.A. (2002). Structural basis of transcription initiation: an RNA polymerase holoenzyme-DNA complex. *Science* **296**:1285-1290.

Murray, H.D., Schneider, D.A., and Gourse, R.L. (2003). Control of rRNA expression by small molecules is dynamic and nonredundant. *Mol. Cell* **12**:125-134.

Nakanishi, N., Abe, H., Ogura, Y., Hayashi, T., Tashiro, K., Kuhara, S., Sugimoto, N., and Tobe, T. (2006). ppGpp with DksA controls gene expression in the locus of enterocyte effacement (LEE) pathogenicity island of enterohaemorrhagic Escherichia coli through activation of two virulence regulatory genes. *Mol Microbiol.* **61**:194-205.

Narayanan, A., Vago, F.S., Li, K., Qayyum, M.Z., Yernool, D., Jiang, W., and Murakami, K.S. (2018). Cryo-EM structure of *Escherichia coli* σ70 RNA polymerase and promoter DNA complex revealed a role of σ non-conserved region during the open complex formation. *J. Biol. Chem.* **293**:7367-7375.

Nayak, D., Voss, M., Windgassen, T., Mooney, R.A., and Landick, R. (2013). Cys-pair reporters detect a constrained trigger loop in a paused RNA polymerase. *Mol. Cell* **50**:882-893.

Opalka, N., Brown, J., Lane, W.J., Twist, K.A., Landick, R., Asturias, F.J., and Darst, S.A. (2010). Complete structural model of Escherichia coli RNA polymerase from a hybrid approach. *PLoS Biol.* **8**:e1000483.

Pao, C.C. and Dyess, B.T. (1981). Effect of unusual guanosine nucleotides on the activities of some Escherichia coli cellular enzymes. *Biochim. Biophys.Acta.* **677**:358-362.

Parshin, A., Shiver, A.L., Lee, J., Ozerova, M., Schneidman-Duhovny, D., Gross, C.A., and Borukhov, S. (2015). DksA regulates RNA polymerase in Escherichia coli

through a network of interactions in the secondary channel that includes Sequence Insertion 1. *Proc. Natl. Acad. Sci. USA.* **112**:E6862-E6871.

Paul, B.J., Barker, M.M., Ross, W., Schneider, D.A., Webb, C., Foster, J.W., and Gourse, R.L. (2004). DksA: a critical component of the transcription initiation machinery that potentiates the regulation of rRNA promoters by ppGpp and the initiating NTP. *Cell* **118**:311-322.

Paul, B.J., Berkmen, M.B., and Gourse, R.L. (2005). DksA potentiates direct activation of amino acid promoters by ppGpp. *Proc. Natl. Acad. Sci. USA.* **102**:7823-7828.

Perederina, A., Svetlov, V., Vassylyeva, M.N., Tahirov, T.H., Yokoyama, S., Artsimovitch, I., and Vassylyev, D.G. (2004). Regulation through the secondary channel—structural framework for ppGpp-DksA synergism during transcription. *Cell* **118**:297-309.

Persky, N.S., Ferullo, D.J., Cooper, D.L., Moore, H.R., and Lovett, S.T. (2009). The ObgE/CgtA GTPase influences the stringent response to amino acid starvation in Escherichia coli. *Mol. Microbiol.* **73**:253-266.

Potrykus, K. and Cashel, M. (2008). (p)ppGpp: still magical? *Annu. Rev. Microbiol.* **62**:35-51.

Riggs, D.L., Mueller, R.D., Kwan, H.S., and Artz, S.W. (1986). Promoter domain mediates guanosine tetraphosphate activation of the histidine operon. *Proc. Natl. Acad. Sci. USA.* **83**:9333-9337.

Ross, W., Gosink, K.K., Salomon, J., Igarashi, K., Zou, C., Ishihama, A., Severinov, K., and Gourse, R.L. (1993). A third recognition element in bacterial promoters: DNA binding by the alpha subunit of RNA polymerase. *Science* **262**:1407-1413.

Ross, W. and Gourse, R.L. (2005). Sequence-independent upstream DNA-alphaCTD interactions strongly stimulate Escherichia coli RNA polymerase-lacUV5 promoter association. *Proc. Natl. Acad. Sci. USA.* **102**:291-296.

Ross, W., Vrentas, C.E., Sanchez-Vazquez, P., Gaal, T., and Gourse, R.L. (2013). The magic spot: a ppGpp binding site on E. coli RNA polymerase responsible for regulation of transcription initiation. *Mol. Cell* **50**:420-429.

Ross, W., Sanchez-Vazquez, P., Chen, A.Y., Lee, J.H., Burgos, H.L., and Gourse, R.L. (2016). ppGpp Binding to a Site at the RNAP-DksA Interface Accounts for Its Dramatic Effects on Transcription Initiation during the Stringent Response. *Mol. Cell* **62**:811-823.

Ruff, E.F., Drennan, A.C., Capp, M.W., Poulos, M.A., Artsimovitch, I., and Record, M.T. Jr. (2015). E. coli RNA Polymerase Determinants of Open Complex Lifetime and Structure. *J. Mol. Biol.* **427**:2435-2450.

Rutherford, S.T., Lemke, J.J., Vrentas, C.E., Gaal, T., Ross, W., and Gourse, R.L. (2007). Effects of DksA, GreA, and GreB on transcription initiation: insights into the mechanisms of factors that bind in the secondary channel of RNA polymerase. *J. Mol. Biol.* **366**:1243-1257.

Rutherford, S.T., Villers, C.L., Lee, J.H., Ross, W., and Gourse, R.L. (2009). Allosteric control of Escherichia coli rRNA promoter complexes by DksA. *Genes Dev.* **23**:236-248.

Ryals, J., Little, R., and Bremer, H. (1982). Control of rRNA and tRNA syntheses in Escherichia coli by guanosine tetraphosphate. *J. Bacteriol*. **151**:1261-1268.

Saecker, R.M., Record, M.T. Jr., deHaseth, P.L. (2011). Mechanism of bacterial transcription initiation: RNA polymerase – promoter binding, isomerization to initiation-competent open complexes, and initiation of RNA synthesis. *J. Mol. Biol.* **412**:754-771.

Sanchez-Vazquez, P., Dewey, C.N., Kitten, N., Ross, W., and Gourse, R.L. Genome-wide effects of *Escherichia coli* transcription from ppGpp binding to its two sites on RNA polymerase. *Submitted.*

Schwartz, E.C., Shekhtman, A., Dutta, K., Pratt, M.R., Cowburn, D., Darst, S., and Muir, T.W. (2008). A full-length group 1 bacterial sigma factor adopts a compact structure incompatible with DNA binding. *Chem. Biol.* **15**:1091-1103.

Severinov, K., Kashlev, M., Severinova, E., Bass, I., McWilliams, K., Kutter, E., Nikiforov, V., Snyder, L., and Goldfarb, A. (1994). A non-essential domain of Escherichia coli RNA polymerase required for the action of the termination factor Alc. *J. Biol. Chem.* **269**:14254-14259.

Seyfzadeh, M., Keener, J., and Nomura, M. (1993). spoT-dependent accumulation of guanosine tetraphosphate in response to fatty acid starvation in Escherichia coli. *Proc. Natl. Acad. Sci. USA.* **90**:11004-11008.

Siegele, D.A., Hu, J.C., Walter, W.A., and Gross, C.A. (1989). Altered promoter recognition by mutant forms of the sigma 70 subunit of Escherichia coli RNA polymerase. *J. Mol. Biol.* **206**:591-603.

Snyder, L., Gold, L., and Kutter, E. (1976). A gene of bacteriophage T4 whose product prevents true late transcription on cytosine-containing T4 DNA. *Proc. Natl. Acad. Sci. USA.* **73**:3098-3102.

Spira, B., Silberstein, N., and Yagil, E. (1995). Guanosine 3’,5’-bispyrophosphate (ppGpp) synthesis in cells of Escherichia coli starved for Pi. *J. Bacteriol.* **177**:4053-4058.

Sreenivasan, R., Heitkamp, S., Chhabra, M., Saecker, R., Lingeman, E., Poulos, M., McCaslin, D., Capp, M.W., Artsimovitch, I., and Record, M.T. Jr. (2016).

Fluorescence Resonance Energy Transfer Characterization of DNA Wrapping in Closed and Open Escherichia coli RNA Polymerase-λP(R) Promoter Complexes. *Biochemistry* **55**:2174-1786.

Srivastava, D.B., Leon, K., Osmundson, J., Garner, A.L., Weiss, L.A., Westblade, L.F., Glickman, M.S., Landick, R., Darst, S.A., Stallings, C.L., and Campbell, E.A. (2013). Structure and function of CarD, an essential mycobacterial transcription factor. *Proc. Natl. Acad. Sci. USA.* **110**:12619-12624.

Stallings, C.L., Stephanou, N.C., Chu, L., Hochschild, A., Nickels, B.E., and Glickman, M.S. (2009). CarD is an essential regulator of rRNA transcription required for Mycobacterium tuberculosis persistence. *Cell* **138**:146-159.

Sy, J. and Lipmann, F. (1973). Identification of the synthesis of guanosine tetraphosphate (MS I) as insertion of a pyrophosphoryl group into the 3'-position in guanosine 5'-diphosphate. *Proc. Natl. Acad. Sci. USA.* **70**:306-309.

Tagami, S., Sekine, S., Kumarevel, T., Hino, N., Murayama, Y., Kamegamori, S., Yamamoto, M., Sakamoto, K., and Yokoyama, S. (2010). Crystal structure of bacterial RNA polymerase bound with a transcription inhibitor protein. *Nature* **468**:978-982.

Takahashi, K., Kasai, K., and Ochi, K. (2004). Identification of the bacterial alarmone guanosine 5'-diphosphate 3'-diphosphate (ppGpp) in plants. *Proc. Natl. Acad. Sci. USA.* **101**:4320-4324.

Tomsic, M., Tsujikawa, L., Panaghie, G., Wang, Y., Azok, J., and deHaseth, P.L. (2001). Different roles for basic and aromatic amino acids in conserved region 2 of Escherichia coli sigma(70) in the nucleation and maintenance of the single-stranded DNA bubble in open RNA polymerase-promoter complexes. *J. Biol. Chem.* **276**:31891-31896.

Traxler, M.F., Summers, S.M., Nguyen, H.T., Zacharia, V.M., Hightower, G.A., Smith, J.T., and Conway, T. (2008). The global, ppGpp-mediated stringent response to amino acid starvation in Escherichia coli. *Mol. Microbiol.* **68**:1128-1148.

Trotochaud, A.E. and Wassarman, K.M. (2004). 6S RNA Function Enhances Long-Term Cell Survival. *J. Bacteriol.* **186**:4978-4985.

Varik, V., Oliveira, S.R.A., Hauryliuk, V., and Tenson, T. (2017). HPLC-based quantification of bacterial housekeeping nucleotides and alarmone messengers ppGpp and pppGpp. *Sci. Rep.* **7**:11022.

Vasconcelos, A.T., Ferreira, H.B., Bizarro, C.V., Bonatto, S.L., Carvalho, M.O. et al. (2005). Swine and poultry pathogens: the complete genome sequences of two strains of Mycoplasma hyopneumoniae and a strain of Mycoplasma synoviae. *J. Bacteriol.* **187**:5568-5577.

Vassylyev, D.G., Sekine, S., Laptenko, O., Lee, J., Vassylyeva, M.N., Borukhov, S., and Yokoyama, S. (2002). Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 A resolution. *Nature* **417**:712-719.

Vassylyev, D.G., Vassylyeva, M.N., Zhang, J., Palangat, M., Artsimovtich, I., and Landick, R. (2007). Structural basis for substrate loading in bacterial RNA polymerase. *Nature* **448**:163-168.

Vinella, D., Albrecht, C., Cashel, M., and D'Ari, R. (2005). Iron limitation induces SpoT-dependent accumulation of ppGpp in Escherichia coli. *Mol Microbiol.* **56**:958-970.

Vuthoori, S., Bowers, C.W., McCracken, A., Dombroski, A.J., and Hinton, D.M. (2001). Domain 1.1 of the sigma(70) subunit of Escherichia coli RNA polymerase modulates the formation of stable polymerase/promoter complexes. *J. Mol. Biol.* **309**:561-572.

Wassarman, K.M. (2018). 6S RNA, A Global Regulator of Transcription. *Microbiol. Spectr.* **6**:10.1128/microbiol.spec.RWR-0019-2018.

Wassarman, K.M. and Storz, G. (2000). 6S RNA regulates E. coli RNA polymerase activity. *Cell* **101**:613-623.

Wendrich, T.M., Blaha, G., Wilson, D.N., Marahiel, M.A., and Nierhaus, K.H. (2002). Dissection of the mechanism for the stringent factor RelA. *Mol. Cell* **10**:779-788.

Westblade, L.F., Minakhin, L., Kuznedelov, K., Tackett, A.J., Chang, E.J., Mooney, R.A., Vvedenskaya, I., Wang, Q.J., Fenyö, D., Rout, M.P., Landick, R., Chait, B.T., Severinov, K., and Darst, S.A. (2008). Rapid isolation and identification of bacteriophage T4-encoded modifications of Escherichia coli RNA polymerase: a generic method to study bacteriophage/host interactions. *J. Proteome Res.* **7**:1244-1250.

Whipple, F.W. and Sonenshein, A.L. (1992). Mechanism of initiation of transcription by Bacillus subtilis RNA polymerase at several promoters. *J. Mol. Biol.* **223**:399-414.

Wilson, C. and Dombroski, A.J. (1997). Region 1 of sigma70 is required for efficient isomerization and initiation of transcription by Escherichia coli RNA polymerase. *J. Mol. Biol.* **267**:60-74.

Windgassen, T.A., Mooney, R.A., Nayak, D., Palangat, M., Zhang, J., and Landick, R. (2014). Trigger-helix folding pathway and SI3 mediate catalysis and hairpin-stabilized pausing by Escherichia coli RNA polymerase. *Nucleic Acids Res.* **42**:12707-12721.

Xue, X., Tomasch, J., Sztajer, H., and Wagner-Döbler, I. (2010). The delta subunit of RNA polymerase, RpoE, is a global modulator of Streptococcus mutans environmental adaptation. *J. Bacteriol.* **192**:5081-5092.

Zhang, Y., Feng, Y., Chatterjee, S., Tuske, S., Ho, M.X., Arnold, E., and Ebright, R.H. (2012). Structural basis of transcription initiation. *Science* **336**:1076-1080.

Zhang, Y., Zborníková, E., Rejman, D., and Gerdes, K. (2018). Novel (p)ppGpp Binding and Metabolizing Proteins of *Escherichia coli*. *MBio*. **9**:e02188-17.

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. *Mol. Cell* **50**:430-436.

Zuo, Y. and Steitz, T.A. (2015). Crystal structures of the E. coli transcription initiation complexes with a complete bubble. *Mol. Cell* **58**:534-540.

Chapter 2

Stress regulation by reprogramming the conformational dynamics of RNA polymerase

A version of this chapter will be submitted for publication (Albert Y. Chen, Wilma Ross, Saumya Gopalkrishnan, Seth A. Darst, Richard L. Gourse, and Andrey Feklistov). Andrey Feklistov performed and analyzed the cryo-EM and beacon assay experiments (Figures 2.1A-C, 2.2A, 2.3A, 2.4, 2.5A,B). Wilma Ross performed and analyzed the DRaCALA assay in Figure 2.3B. Saumya Gopalkrishnan performed the in vitro transcription in Figure 2.S6. I performed all other experiments (Figures 2.1D,E, 2.2B-D, 2.3C-E, 2.5C-D, 2.S1-2.S5). Andrey Feklistov, Wilma Ross, Richard Gourse, and I wrote this chapter.

Abstract

Bacteria constantly sense, and respond to, an ever-changing environment in part by controlling their gene expression with transcription factors (TFs). Many TFs recognize specific DNA sequences adjacent to the promoter, the RNA polymerase (RNAP) binding site, and thereby modulate RNAP's interactions with that promoter. Conversely, in proteobacteria, the stress regulator ppGpp and its co-factor DksA bind to RNAP without contacting DNA, and yet they specifically regulate nearly a fifth of the cell's promoters. Combining ensemble cryo-EM and biochemical assays, we show here that ppGpp and DksA reposition flexible modules in RNAP whose movements orchestrate promoter melting during transcription initiation. DksA and ppGpp alter the conformational dynamics of the initiating RNAP and the kinetics of transcription initiation thereby rapidly executing sweeping changes in gene expression to adapt to a stressful environment.

Introduction

Throughout the bacterial domain, the stringent response is responsible for reprogramming the transcriptome in response to changes in nutrition and the environment (Gourse et al., 2018; Hauryliuk et al., 2015; Potrykus and Cashel, 2008). When stressed, cells can increase production of the small molecule ppGpp as much as 100-fold (Varik et al., 2017; Traxler et al., 2011; Ryals et al., 1982), redirecting transcription from genes involved in growth and proliferation to those for survival and adaptation, as well as for persistence, antibiotic tolerance, and virulence in pathogenic organisms (Aberg et al., 2009; Greenway and England, 1999; Kim et al., 2005; Pomares

et al., 2008; Korch et al., 2003; Erickson et al., 2004; Dalebroux et al., 2010). The mechanism of regulation by ppGpp varies among bacterial species (Krasny and Gourse, 2004; Liu et al., 2015). In proteobacteria, this ppGpp-driven response requires the protein cofactor DksA (Paul et al., 2004; Perederina et al., 2004). Although DksA alone can affect transcription *in vitro*, as well as *in vivo* to a small extent even without overexpression, its concentrations vary little in cells (Paul et al., 2004; Rutherford et al., 2007; Chandrangsu et al., 2011). Together, both ppGpp and DksA are required for the large effects on transcription observed during the stringent response, and it is the changing levels of ppGpp that account for the timing of regulation. The effects of DksA and ppGpp together on a given gene depend on the kinetics of promoter utilization, which in turn is determined by the promoter sequence.

Previous genetic, biochemical, and crystallographic studies delineated the binding sites of DksA and ppGpp on RNAP (Ross et al., 2013; Zuo et al., 2013; Ross et al., 2016; Molodtsov et al., 2018), but the mechanism(s) accounting for the promoter-specific effects of these factors remain unclear. The highly mobile nature of RNAP (Chakraborty et al., 2012; Feklistov et al., 2017) suggests that DksA/ppGpp could affect the structural dynamics of RNAP during transcription initiation, and that these effects might not have been detected in earlier structural studies because of crystal packing constraints. Furthermore, past mechanistic studies have focused mostly on effects of DksA or ppGpp alone, so much is still unclear about how DksA and ppGpp regulate transcription together (Ross et al., 2013; Rutherford et al., 2009; Lennon et al., 2012; Parshin et al., 2015).

To overcome this limitation, we used single-particle cryo-electron microscopy (cryo-EM) and biochemical assays to examine how DksA and ppGpp together affect the conformation of RNAP. We obtained structures of RNAP bound to DksA, RNAP bound to both DksA and ppGpp, and RNAP bound to ppGpp and lipiarmycin. We find that DksA, properly positioned by ppGpp, affects the conformation of three mobile regions in RNAP: the swivel module (which includes the clamp), the trigger loop (TL), and the β lobe/β sequence insertion 1 (β SI1). Mutational analysis shows the importance of the DksA-trigger loop and DksA-β sequence insertion 1 interactions, as well as the role of ppGpp in repositioning DksA to dramatically increase its effects on transcription. Kinetic and biochemical analysis demonstrate the role of conformational changes in RNAP mobile elements in regulation by DksA and ppGpp.

Results

DksA and ppGpp affect RNAP clamp positioning

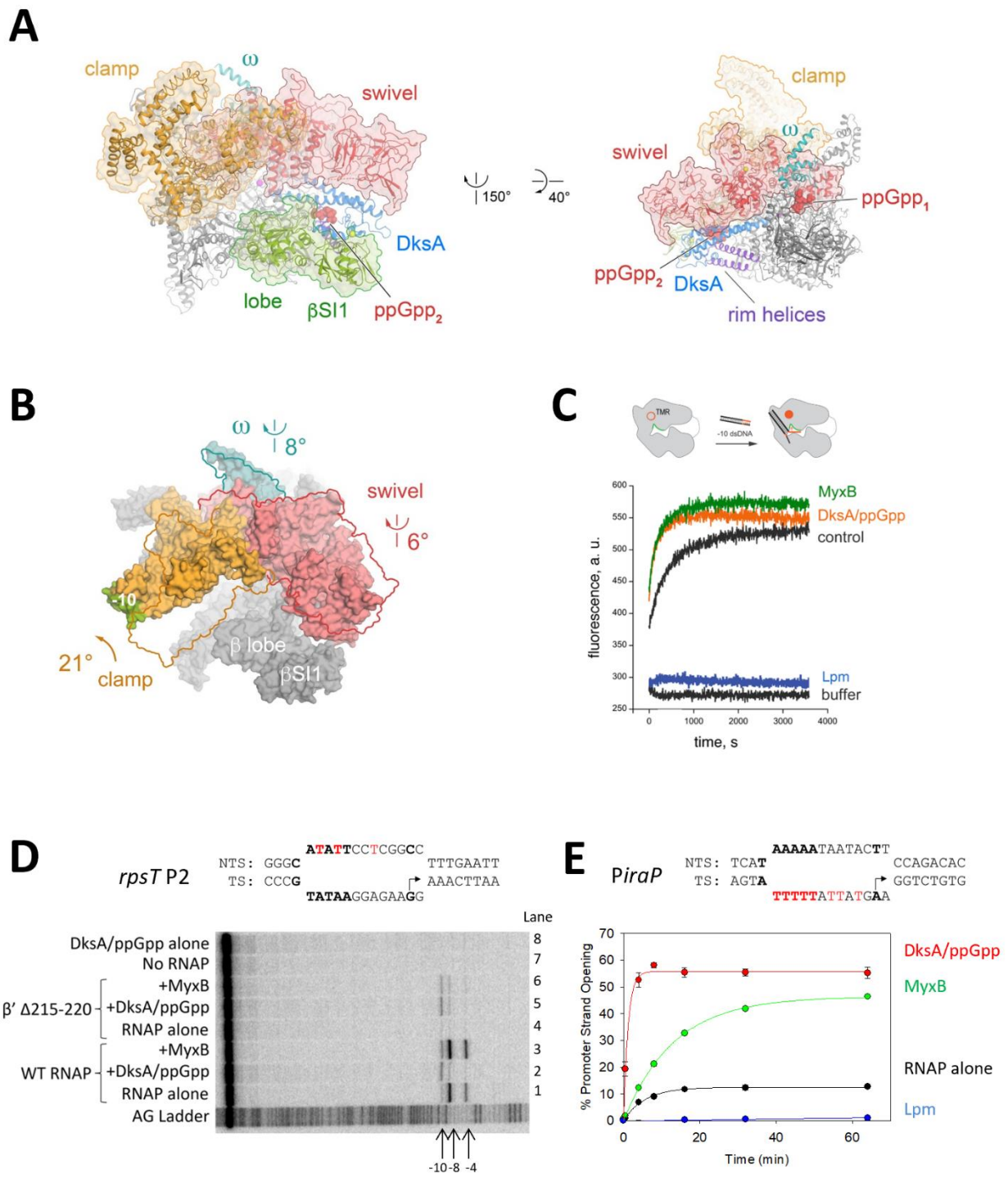
We first solved the cryo-EM structure of *E. coli* RNAP bound to both DksA and ppGpp. In this structure, DksA is bound in the RNAP secondary channel, while ppGpp is bound at two sites on RNAP, consistent with published genetic, biochemical, and structural data (Figure 2.1A; Ross et al., 2013; Ross et al., 2016; Parshin et al., 2015; Molodtsov et al., 2018).

In this structure, the RNAP clamp is closed. Since holoenzyme in the absence of factors is a mixed population with both open and closed clamp states (Chakraborty et al., 2012), for comparison we also solved the structure of *E. coli* RNAP holoenzyme with ppGpp and lipiarmycin (Lpm). The RNAP-Lpm-ppGpp complex was predicted to have a

Figure 2.1. DksA and ppGpp favor a closed RNAP clamp to increase strand opening.

(A) Structure of RNAP-DksA-ppGpp complex. Clamp, swivel, and β lobe modules shown in orange, red, and green, respectively. ω shown in light blue. DksA shown in dark blue cartoon. ppGpp at sites 1 and 2 shown in red spheres. **(B)** Comparison of clamp and swivel conformations in lipiarmycin-ppGpp-bound RNAP complex and in the RNAP-DksA-ppGpp complex. Lipiarmycin-ppGpp-RNAP complex shown in surface view. Position of clamp, swivel, and ω subunit in the RNAP-DksA-ppGpp complex shown in outline. **(C)** Time course of TMR-labeled RNAP (3 nM) fluorescence response to addition of 2 uM –10 element dsDNA. Unliganded RNAP (black trace) is compared to RNAP preincubated with MyxB (green), DksA/ppGpp (red) and Lpm (blue). (Inset) RNAP beacon assay schematic reporting on -10 element recognition. **(D)** Effects of DksA/ppGpp or MyxB on strand opening at *rpsT* P2 as measured by KMnO4 footprinting. *rpsT* P2 sequence indicated above, with reactive Ts indicated in red. The TSS and -10 element are indicated in bold, and the arrow indicates the direction of transcription. **(E)** Time course of strand opening as measured by KMnO4 footprinting at P*iraP*. Unliganded RNAP (black trace) is compared to RNAP preincubated with MyxB (green), DksA/ppGpp (red) or Lpm (blue). *iraP* promoter sequence indicated above, with reactive Ts indicated in red. The TSS and -10 element are indicated in bold, and the arrow indicates the direction of transcription.

Figure 2.1


A
clamp
ω
swivel
DksA
lobe
βSI1
ppGpp2
150°
40°
clamp
swivel
ω
ppGpp1
ppGpp2
DksA
rim helices
B
ω
8°
swivel
6°
-10
β lobe
βSI1
21°
clamp
C
TMR
-10 dsDNA
MyxB
DksA/ppGpp
control
Lpm
buffer
fluorescence, a. u.
time, s
D
rpsT P2
ATATTCCTCGGCC
NTS: GGGC TTTGAATT
TS: CCCG AAACTTAA
TATAAGGAGAAGG
Lane
DksA/ppGpp alone
No RNAP
β' Δ215-220
+MyxB
+DksA/ppGpp
RNAP alone
WT RNAP
+MyxB
+DksA/ppGpp
RNAP alone
AG Ladder
-10 -8 -4
E
PiraP
AAAAATAATACTT
NTS: TCAT CCAGACAC
TS: AGTA GGTCTGTG
TTTTTATTATGAA
% Promoter Strand Opening
Time (min)
DksA/ppGpp
MyxB
RNAP alone
Lpm

fully open clamp because Lpm binds at the base of the RNAP clamp, preventing it from closing (Lin et al., 2018; Boyaci et al., 2018). The clamp domain, along with the dock, shelf, β' SI3, and the β' C-terminal segment, form the swivel module, which rotates along an axis roughly parallel to the bridge helix (Kang et al., 2018). Rotation of the swivel module allows addition movement of the clamp domain in a plane orthogonal to swivel motion, called clamp opening (Kang et al., 2018). When the swivel module is not rotated, as observed in elongation complex structures (Kang et al., 2017), clamp opening is disfavored. In our RNAP-Lpm-ppGpp structure, the swivel module is rotated 6° parallel to the main cleft, and the clamp domain within the swivel module is additionally opened 21° away from the active site cleft (Figure 2.1B).

The swivel module also makes up one wall of the secondary channel. Since DksA binds in the secondary channel, its interaction with the swivel module likely prevents swiveling, thereby disfavoring the RNAP clamp from opening away from the active site cleft. Recent reports showed that clamp closure is required for a step after initial promoter DNA binding. Specifically, it was reported that clamp closure allows RNAP to nucleate transcription bubble melting through capture of the non-template -11A base (-10 element recognition) in a pocket in the σ subunit (Feklistov et al., 2017). We hypothesize that together DksA and ppGpp could facilitate -10 recognition during transcription initiation by preventing swiveling and increasing clamp closure.

To test this model, we used the "beacon" assay in which -10 element recognition specifically enhances the fluorescence intensity of a probe covalently linked to σ near the -10 recognition site (Mekler et al., 2011). By using a truncated promoter template lacking DNA downstream of the -10 element, we focused on early steps in the

mechanism, primarily -10 recognition, while avoiding effects on later steps in transcription initiation. In this assay, clamp-closing ligands like myxopyronin B (MyxB) favor -10 element recognition, resulting in a rapid increase in fluorescence. In contrast, clamp-opening ligands like lipiarmycin inhibit -10 recognition and do not increase fluorescence (Figure 2.1C; Feklistov et al., 2017). DksA/ppGpp caused a rate enhancement nearly identical to that observed with MyxB, consistent with the model that clamp closure resulting from DksA/ppGpp binding increases -10 recognition (Figure 2.1C).

Consistent with the results from the beacon assay, we were able to detect DksA/ppGpp-dependent nucleation of strand opening at the *rpsT* P2 promoter. WT RNAP formed a fully melted transcription bubble in this promoter complex in the absence of factors, as shown by the reactivity of single-stranded thymines at -8 and -4 on the non-template strand following $KMnO_4$ treatment (Figure 2.1D, lane 1). In the presence of DksA and ppGpp, transcription initiation stopped at an earlier intermediate in which the -10 thymine became hyper-reactive while downstream portions of the bubble were not yet melted (Figure 2.1D, lane 2). This shows that DksA/ppGpp inhibits formation of the fully open bubble at this promoter while allowing formation of an early intermediate.

To determine whether DksA and ppGpp actively promote formation of this early intermediate or simply does not prevent its formation, we used an RNAP variant lacking residues required for stabilizing RNAP interactions with downstream duplex DNA late in RPo formation (RNAP β'Δ215-220; Bartlett et al., 1998). In the absence of DksA and ppGpp, β'Δ215-220 RNAP formed a closed complex, showing little or no $KMnO_4$

reactivity at -10, -8, or -4 (Figure 2.1D, lane 4; Bartlett et al., 1998). Addition of DksA and ppGpp to β'Δ215-220 RNAP resulted in formation of a partially melted intermediate in which only the -10 thymine on the non-template strand was reactive while downstream residues were not (Figure 2.1D, lane 5). Use of the β'Δ215-220 RNAP allowed separation of effects on initial strand capture from stabilization of the fully melted transcription bubble, demonstrating that DksA/ppGpp directly promote nucleation of strand opening.

With β'Δ215-220 RNAP, MyxB was able to promote reactivity of the -10 base like DksA and ppGpp, suggesting that clamp closure alone is sufficient to promote nucleation of strand opening (Figure 2.1D, lane 6). With WT RNAP, however, promoter positions -10, -8, and -4 were all reactive in the presence of MyxB (Figure 2.1D, lane 3), showing that unlike DksA and ppGpp, MyxB does not inhibit melting of the full transcription bubble. These results are consistent with the effects of DksA and ppGpp on capture of -11A in the beacon assay and a role for DksA/ppGpp in nucleation of -10 recognition by clamp closure.

We note that the overall effect of DksA and ppGpp on the *rpsT* P2 promoter is negative (Lemke et al., 2011), even though we found that DksA and ppGpp promoted nucleation of strand opening. The overall negative effect of DksA and ppGpp results from their effects on later steps in the pathway to open complex formation, shifting the equilibrium from a fully open complex to an intermediate in which only the upstream end of the -10 hexamer was single-stranded. Thus, this promoter presented us with an opportunity to detect the intermediate complex and focus on the initial strand opening step. Our results are thus consistent with the model that DksA/ppGpp increases the

formation of an early step at all promoters, but DksA/ppGpp only increase transcriptional output from the class of promoters in which that step is rate-limiting (i.e. activated promoters; Paul et al., 2005).

To further test the model that an increased rate of clamp closure and the resulting -11 capture can account for activation of transcription by DksA/ppGpp, we examined effects of DksA/ppGpp on the rate of strand opening and transcription initiation by wild-type RNAP and the positively-regulated *iraP* promoter. As above, the progress of strand opening was monitored by $KMnO_4$ reactivity of single-stranded thymine residues. We found that DksA/ppGpp increased strand opening of the *iraP* promoter independent of the RNAP concentration (Figure 2.S1). These results are consistent with previous studies in which it was found that DksA/ppGpp activated transcription of a different promoter, P*argI*, even at saturating RNAP concentrations, indicating that they affect a step after initial binding of RNAP to form a closed complex (Paul et al., 2005).

As shown in Figure 2.1E, MyxB also increased DNA melting, but not as strongly as DksA/ppGpp (Figure 2.S2), suggesting there may be mechanism(s) in addition to clamp closure that contribute to full activation by DksA and ppGpp. Together, our results suggest that DksA/ppGpp-dependent clamp closure leads to an increase in -10 recognition, an increase in the rate of transcription bubble opening, and an increase in transcription at promoters where this step is rate-determining.

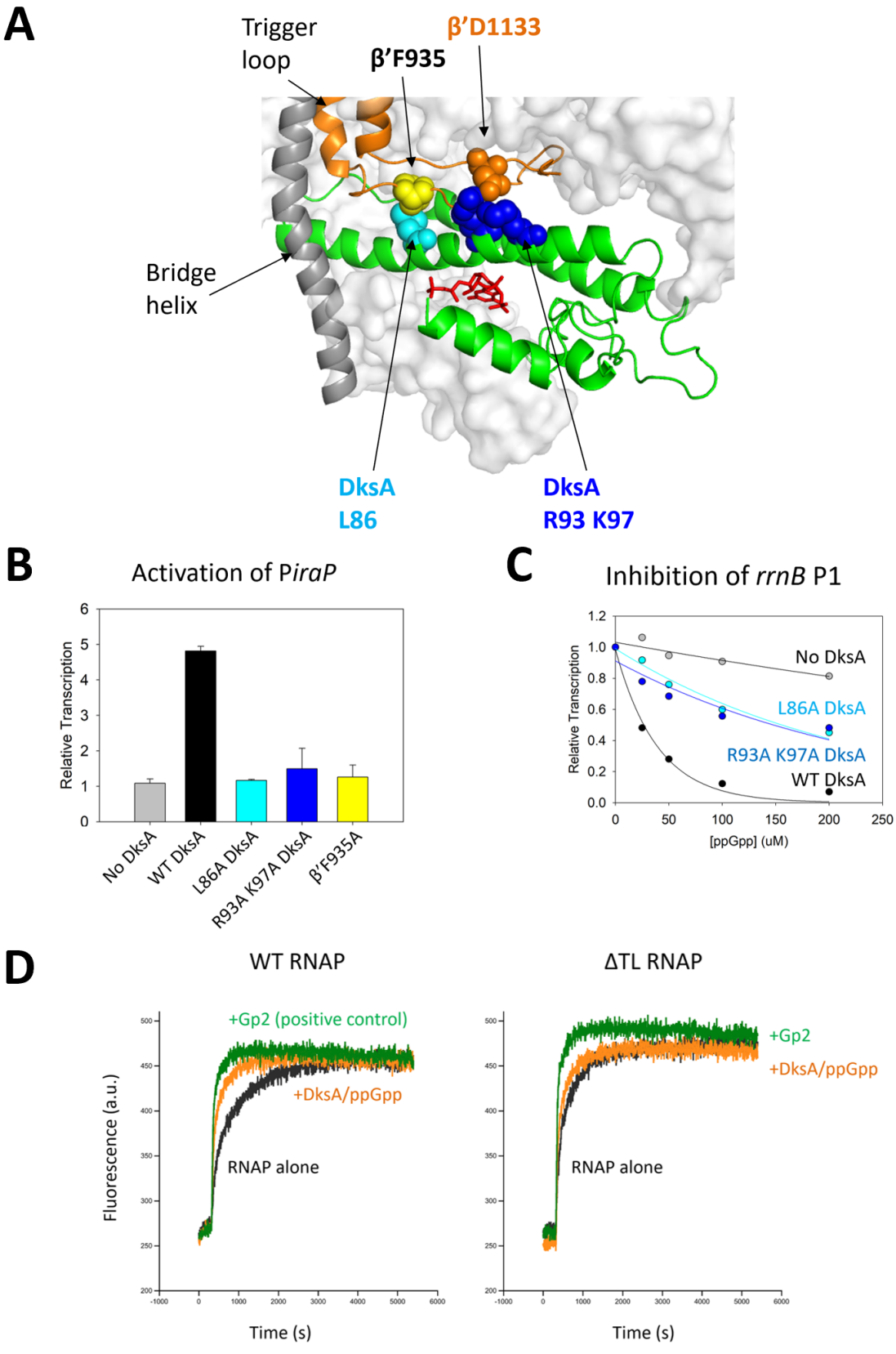
DksA interacts with the trigger loop

The cryo-EM structure of RNAP bound to DksA and ppGpp revealed a critical interaction of DksA with the trigger loop (TL). The TL is a highly mobile segment of β' that plays a critical role in the catalytic cycle of RNAP (Zhang et al., 2010). DksA was shown previously to favor a partially folded TL conformation, but there was no structural information about interactions between DksA and the TL and therefore little mechanistic understanding of the role of the TL in regulation by DksA/ppGpp (Rutherford et al., 2009; Lennon et al., 2012; Nayak et al., 2013). In our RNAP-DksA-ppGpp structure, the trigger loop is constrained, ordered, and lays across the coiled-coil of DksA, forming specific interactions between β'F935 and DksA residue L86, as well as between β'D1133 and DksA residues R93 and K97 (Figure 2.2A). Alanine substitutions disrupting this interaction abolish regulatory functions of DksA while having minimal impact on binding of DksA to RNAP (Figure 2.2B,C, 2.S3, 2.S4). The conformation of the trigger loop, in concert with movements of the bridge helix and switch regions, may also allosterically affect clamp movement and -10 recognition. Consistent with this model, an RNAP variant lacking the trigger loop no longer responded to DksA and ppGpp for -10 recognition in the beacon assay (Figure 2.2D) or in promoter complex stability assays (Rutherford et al., 2009).

β' sequence insertion 3 (β'SI3) is a lineage-specific insertion within the TL previously shown to affect TL folding and to regulate transcriptional pausing (Kang et al., 2018; Windgassen et al., 2014). In the *his* paused elongation complex, folding of the TL is inhibited by swiveling and swiveling-induced exclusion of β'SI3 from the secondary channel (Kang et al., 2018). In the RNAP-DksA-ppGpp structure, folding of the TL is

Figure 2.2. DksA interacts with the trigger loop to regulate transcription. (A) Interactions between trigger loop (orange) and DksA (green). **(B)** Multi-round *in vitro* transcription showing effects of substitutions in TL or DksA residues on activation of P*iraP*. **(C)** Multi-round *in vitro* transcription showing effects of substitutions in DksA residues on inhibition of *rrnB* P1. ω- RNAP used to focus on site 2-dependent effects. **(D)** Effect of TL deletion on -10 recognition in the Beacon assay. Gp2 in green (clamp closing agent) used as a positive control.

Figure 2.2


A
Trigger loop
β'D1133
β'F935
Bridge helix
DksA L86
DksA R93 K97
B
Activation of PiraP
Relative Transcription
No DksA
WT DksA
L86A DksA
R93A K97A DksA
β'F935A
C
Inhibition of rrnB P1
Relative Transcription
No DksA
L86A DksA
R93A K97A DksA
WT DksA
[ppGpp] (uM)
D
WT RNAP
ΔTL RNAP
+Gp2 (positive control)
+DksA/ppGpp
RNAP alone
+Gp2
+DksA/ppGpp
RNAP alone
Fluorescence (a.u.)
Time (s)
Time (s)

inhibited instead by interactions between the TL and the DksA coiled-coil. To determine if β'SI3 still played a role in activation by DksA/ppGpp, we tested whether RNAP lacking β'SI3 is still activated by DksA/ppGpp at the *iraP* promoter. We found that deletion of β'SI3 reduced the effect of DksA/ppGpp on transcription activation, and we suggest that β'SI3 increases the interaction of DksA and the TL (Figure 2.S5). Together, the data suggest that the interaction of DksA with the trigger loop, as modulated by the presence of β'SI3, favors clamp closure and DNA nucleation, potentially by inhibiting rotation of the swivel module, although we note that β'SI3 could also influence the affinity of RNAP for DksA.

The role of ppGpp in regulating DksA function

ppGpp concentrations change dramatically with nutritional conditions, whereas DksA concentrations change little in *E. coli* (Varik et al., 2017; Traxler et al., 2011; Paul et al., 2004; Rutherford et al., 2007; Chandrangsu et al., 2011). Therefore, the presence or absence of ppGpp in a DksA-ppGpp-RNAP complex must account for the timing of regulatory effects of DksA/ppGpp on transcription. ppGpp is absolutely required for activation of positively regulated promoters, and it greatly enhances the effects of DksA on inhibited promoters (Paul et al., 2004; Paul et al., 2005).

To gain further insights into the role of ppGpp in transcription regulation, we solved the cryo-EM structure of RNAP bound to DksA alone and compared it with the structure of RNAP bound to DksA and ppGpp together. In the RNAP-DksA-ppGpp structure, we found that ppGpp was bound at the two predicted sites on the RNAP surface. These were resolved to <3 Å in our structure, allowing us to determine the

molecular details of binding and compare them with those reported in previous lower resolution crystal structures. As reported previously, site 1 is in a cavity at the interface of the β' and ω subunits, and site 2 is at the interface of DksA and the β' rim helices (Ross et al., 2013; Zuo et al., 2013; Ross et al., 2016; Molodtsov et al., 2018). However, our structure showed that residue E146 in the DksA C-terminal helix (CTH) contributes to ppGpp binding through Mg^{2+}-coordinated interactions with the ppGpp phosphate moieties, and the neighboring E143 positions DksA K139 to interact with ppGpp (Figure 2.3A). Substitutions for these DksA CTH residues severely impaired ppGpp binding *in vitro* as measured by DRaCALA assays (Figure 2.3B), strongly reduced inhibition of the *rrnB* P1 promoter (Figure 2.3C, D), and virtually eliminated activation of the *iraP* promoter (Figure 2.3E).

The electron density of the DksA coiled-coil was weak in the absence of ppGpp (Figure 2.4A), suggesting that DksA may not bind stably in one conformation in the absence of ppGpp, only rarely or more transiently occupying the correct position required for function. While we do not know how ppGpp affects the binding affinity of DksA for the RNAP-DNA complex, we hypothesize that ppGpp acts as a "linchpin", securing DksA in the appropriate binding position, leading to the synergistic effects of ppGpp and DksA, explaining why DksA alone has a much smaller effect on transcriptional reprogramming than when ppGpp is also present (Paul et al., 2004; Paul et al., 2005; Sanchez-Vazquez et al., submitted).

In the absence of ppGpp, DksA interactions with the trigger loop were weak or absent, and the trigger loop was disordered. Notably, the positions of key residues R93 and K97 DksA were shifted as a result of ppGpp binding. Although it was reported that

Figure 2.3. DksA C-terminal helix residues participate in ppGpp binding at site 2. **(A)** ppGpp binding site 2 involves residues in DksA and β'. **(B)** Effects of substitutions in DksA residues on ppGpp binding as measured by DRaCALA. **(C)** Multi-round *in vitro* transcription showing effects of substitutions in DksA residues on inhibition of *rrnB* P1 in the presence of ppGpp. ω- RNAP used to focus on site 2-dependent effects. **(D)** Multi-round *in vitro* transcription showing effects of substitutions in DksA residues on inhibition of *rrnB* P1 in the absence of ppGpp. **(E)** Multi-round *in vitro* transcription showing effects of substitutions in DksA residues on activation of P*iraP*.

Figure 2.3

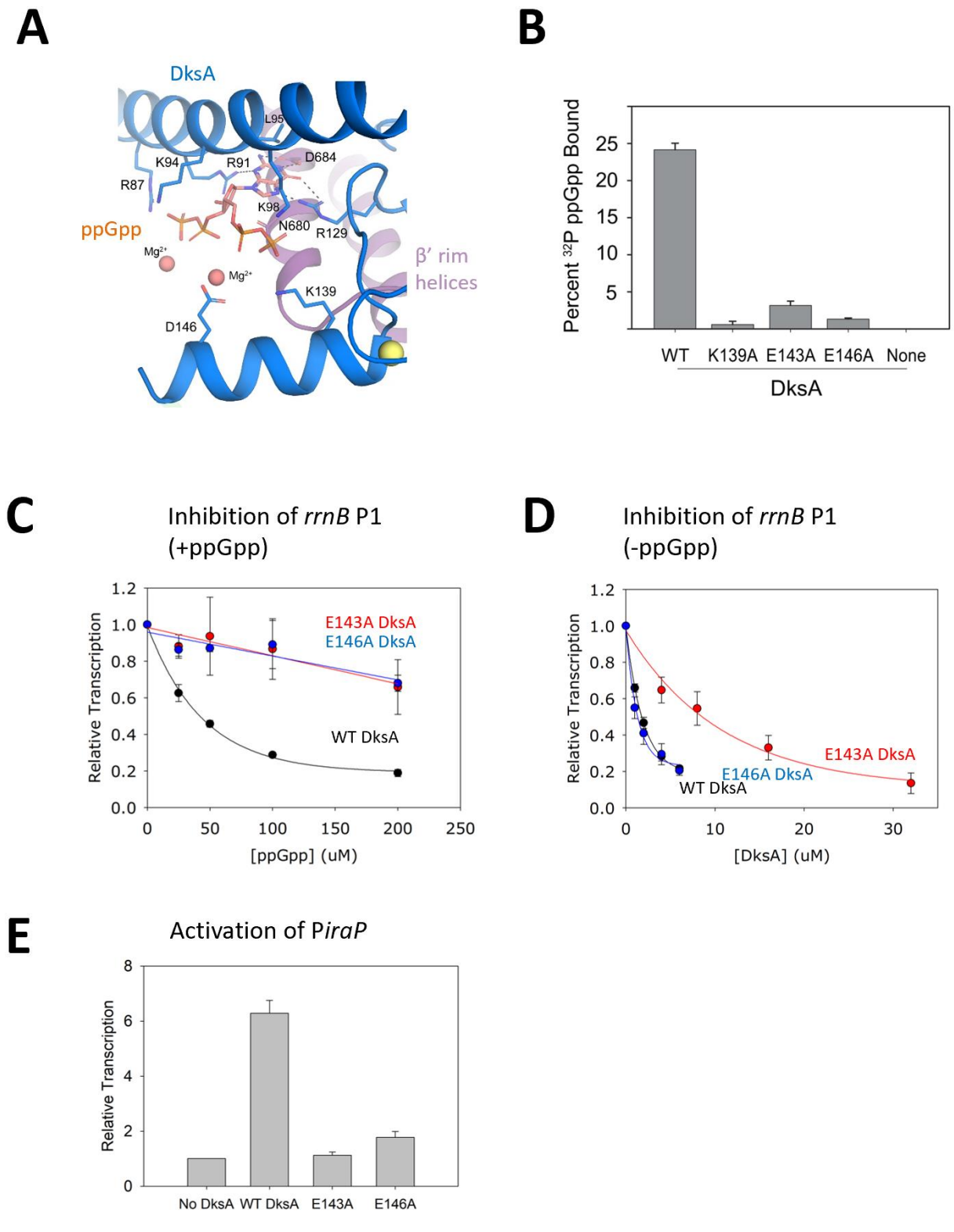
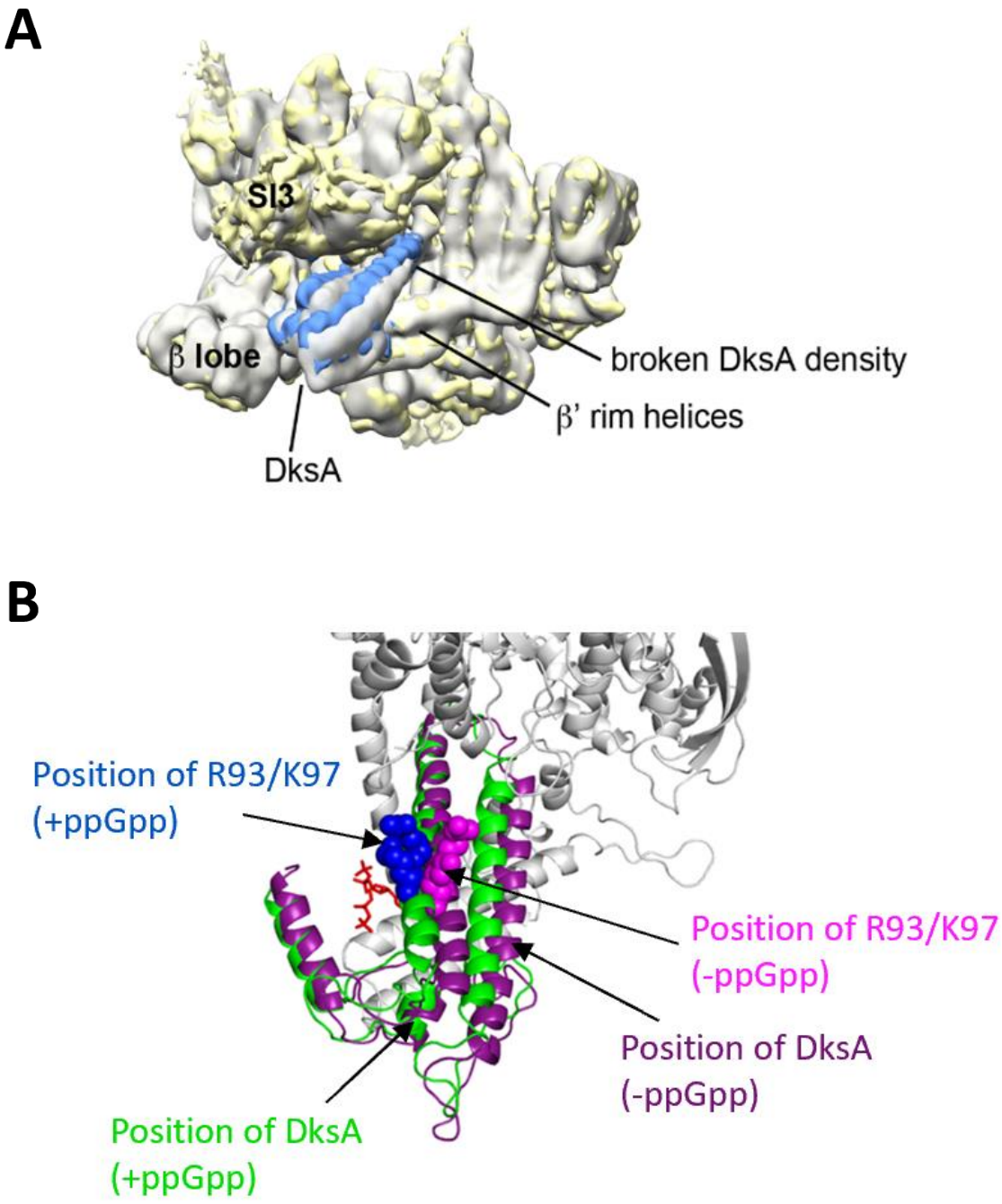

A
DksA
L95
K94
R91
D684
R87
K98
N680
R129
ppGpp
Mg2+
Mg2+
β′ rim
helices
K139
D146
B
Percent 32P ppGpp Bound
25
20
15
10
5
WT
K139A
E143A
E146A
None
DksA
C
Inhibition of rrnB P1
(+ppGpp)
Relative Transcription
1.2
1.0
0.8
0.6
0.4
0.2
0.0
E143A DksA
E146A DksA
WT DksA
0
50
100
150
200
250
[ppGpp] (uM)
D
Inhibition of rrnB P1
(-ppGpp)
Relative Transcription
E143A DksA
E146A DksA
WT DksA
0
10
20
30
[DksA] (uM)
E
Activation of PiraP
Relative Transcription
8
6
4
2
0
No DksA
WT DksA
E143A
E146A

Figure 2.4. ppGpp repositions DksA to interact stably with the trigger loop. (A) Density for RNAP/DksA complex (grey) is overlayed with the density for RNAP/DksA/ppGpp complex. Absence of ppGpp introduces disorder in DksA position. **(B)** Comparison of position of DksA in the secondary channel in the presence of ppGpp (in green) or in the absence (in purple). Positions of residues R93 and K97 in the presence of ppGpp shown in blue spheres or in the absence of ppGpp shown in magenta spheres.

Figure 2.4


A
SI3
β lobe
broken DksA density
β' rim helices
DksA
B
Position of R93/K97
(+ppGpp)
Position of R93/K97
(-ppGpp)
Position of DksA
(-ppGpp)
Position of DksA
(+ppGpp)

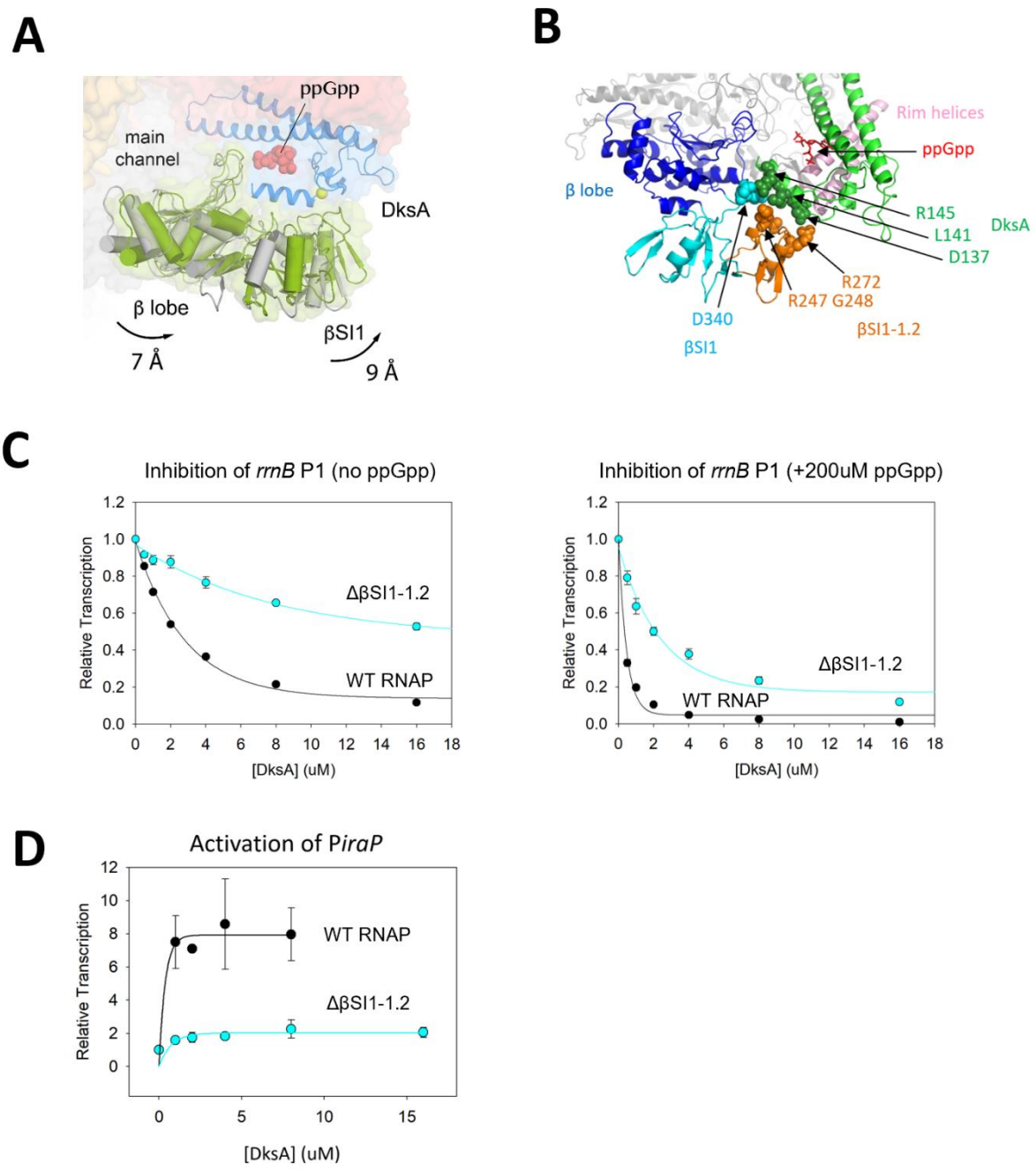
ppGpp slightly increases the binding affinity of DksA for RNAP (Molodtsov et al., 2018), we suggest that a major function of ppGpp is to reposition DksA in the secondary channel to interact with the trigger loop to allow DksA function (Figure 2.4B).

DksA and ppGpp reposition βSI1 and the β lobe

The cryo-EM structure also revealed a major DksA/ppGpp induced movement of RNAP involving the β lobe and the lineage-specific insertion within the lobe, β sequence insertion 1 (βSI1) parallel to the active site cleft towards the β' rim helices. The β lobe forms part of the active site cleft opposite the clamp domain and contacts the NT-strand in the open complex, and the DksA C-terminal α-helix forms an extensive interface with βSI1, consisting of DksA residues D137, L141, and R145 and βSI1 residues R247, G248, R272, and D340. This resulted in a 16 Å displacement of βSI1 and a 7 Å displacement of the β lobe (Figure 2.5A, B). This major displacement likely accounts for the loss of activation and inhibition of transcription by DksA/ppGpp observed with RNAPs lacking the βSI1 subregion that contacts DksA (β240-284; also known as βSI1-1.2) (Figure 2.5C, D; see also Parshin et al., 2015). Although it was reported previously that DksA crosslinks to βSI1 (Parshin et al., 2015), this interaction was not detected in the X-ray structure of the RNAP-DksA-ppGpp complex (Molodtsov et al., 2018). The lack of βSI1-DksA interactions in that structure may reflect crystal-packing constraints. Importantly, the major repositioning of the β lobe and βSI1 by DksA and ppGpp in the cryo-EM structure widens the active site cleft, most likely allowing DNA increased access to the active-site cleft, accounting for the transcription observed at unregulated and positively regulated promoters even when the clamp is closed.

Figure 2.5. DksA and ppGpp reposition the β lobe and βSI1. (A) Comparison of β lobe and βSI1 positions in RNAP-DksA-ppGpp complex (in green) and in the RNAP-Lpm-ppGpp complex (in gray). DksA shown in blue. ppGpp shown in red spheres. Movement of βSI1 as a result of DksA/ppGpp binding indicated by arrow. **(B)** Interactions between DksA C-terminal helix (green) and βSI1 (cyan; β225-343) or subregion βSI1-1.2 (orange; β240-284). Rim helices are shown in pink. ppGpp shown in red sticks. **(C)** Multi-round *in vitro* transcription showing effect of DksA on inhibition of WT or ΔβSI1-1.2 RNAP at *rrnB* P1 in the presence or absence of 200uM ppGpp. **(D)** Multi-round *in vitro* transcription showing effect of βSI1-1.2 deletion on activation of P*iraP*.

Figure 2.5. DksA and ppGpp reposition the β lobe and βSI1.



We propose that displacement of the β lobe and βSI1 has an additional consequence by disrupting important contacts to σ region 1.1. σ1.1 occupies the cleft early in transcription initiation, making contacts to the β lobe and blocking entry of DNA into the cleft (Mekler et al., 2002; Bae et al., 2013). Therefore, σ1.1 must be removed from the cleft to allow access of DNA to the active site. By shifting the position of the β lobe, DksA/ppGpp could weaken β lobe-σ1.1 contacts and favor removal of σ1.1 from the cleft.

To determine whether displacement of σ1.1 plays a role in activation of transcription by DksA/ppGpp, we tested whether RNAP lacking σ1.1 is still activated at the positively regulated *thrABC* promoter. We found that transcription by WT RNAP was activated by DksA/ppGpp 5-fold *in vitro*, but Δσ1.1 RNAP was insensitive to the effects of DksA and ppGpp at this promoter (Figure 2.S6A). Notably, even in the absence of DksA and ppGpp, transcription was increased dramatically in the absence of σ1.1. We infer that deletion of σ1.1 helps bypass a rate-limiting step at the *thrABC* promoter, whereas DksA and ppGpp are needed to activate transcription with WT RNAP (i.e. containing σ1.1), in part by promoting the removal of σ1.1 from the cleft. Consistent with this model, a triple alanine substitution in σ1.1 that weakens interactions between σ1.1 and the β lobe increases the effect of DksA/ppGpp on activation of transcription intiation (see Appendix A).

The β lobe also makes contacts to DNA in the open complex. To stabilize the open complex, multiple residues in the β lobe interact with the core recognition element of the promoter, and the β gate loop within the β lobe makes contacts to the non-template strand (Zhang et al., 2012; NandyMazumdar et al., 2016). Although deletion of

σ1.1 does not eliminate inhibition of the *rrnB* P1 promoter by DksA/ppGpp (Figure 2.S6C), it did increase the DksA concentration required. These results suggest that DksA/ppGpp-induced movement of the β lobe weakens interactions between the β lobe and σ1.1, as well as between the β lobe and DNA. In the absence of σ1.1, DksA/ppGpp still weakens β lobe-DNA interactions, but competition between σ1.1 and DNA no longer occurs, thereby reducing inhibition to some extent (see Appendix A).

DISCUSSION

Promoter melting by σ^{70} RNAP is driven solely by binding free energy, and the energetics and structural dynamics of the process are a function of promoter sequence. In the absence of regulators, the energy landscape for activated and inhibited promoters is uneven, weighted against early or late steps in the melting pathway, respectively. Reprogramming of RNAP motions reshapes the energy landscape of the initiation pathway, modifying read-out of the promoter sequences.

Expression of activated promoters is intrinsically weak in the absence of DksA/ppGpp, in part because of a sub-optimal -10 element, likely resulting in a high energy barrier for nucleation of melting (Figure 2.S1; Burgos, 2017). In contrast, the later stages of melting are facilitated by the presence of an A+T-rich discriminator and/or by σ1.2 contacts with guanines on the NT strand of the discriminator sequence (Haugen et al., 2006; Feklistov et al., 2006). Given sufficient time to overcome the energy barrier imposed by the first step, activated promoters are able to melt and form a stable RPo (Figure 2.S1). DksA and ppGpp lower the energy barrier for nucleation of melting through interactions of DksA with the trigger loop within the RNAP swivel

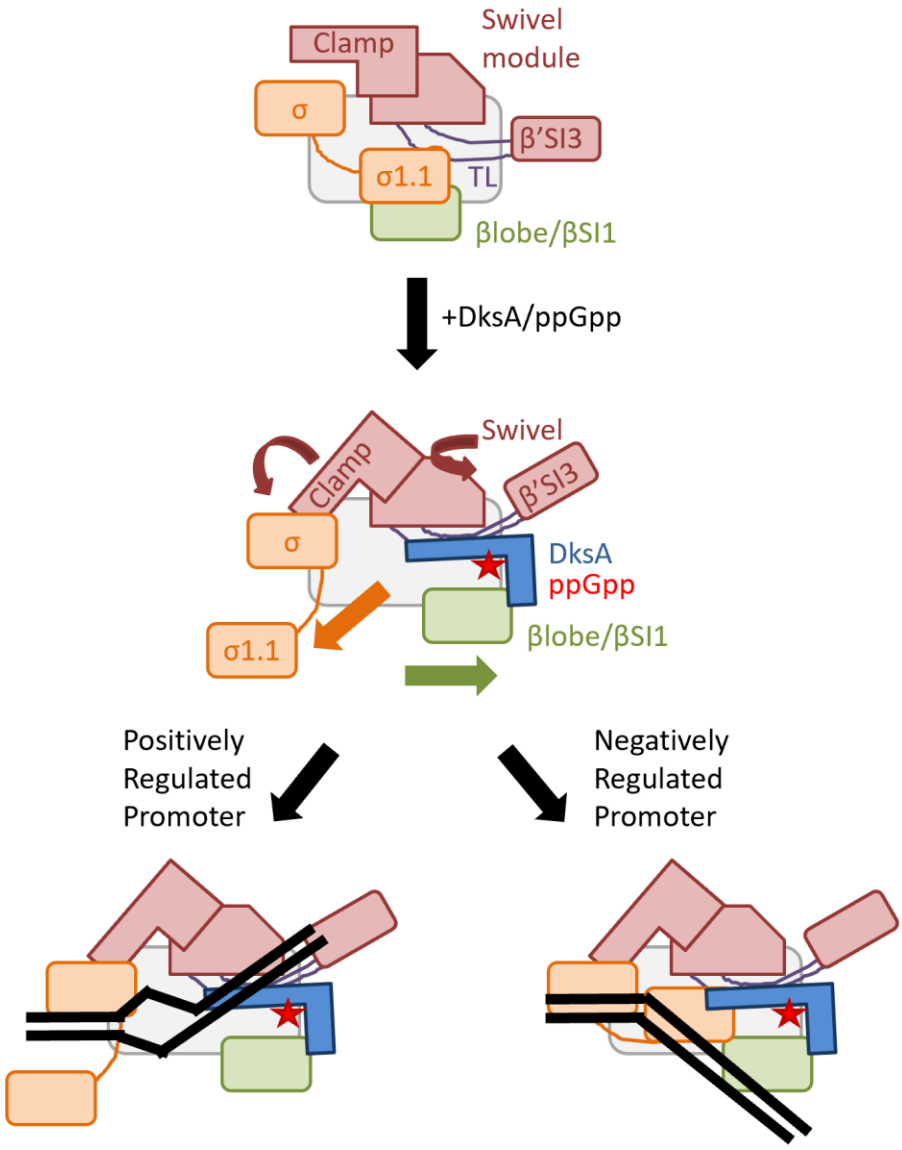
domain to induce a closed RNAP clamp (Figure 2.6). Furthermore, β lobe displacement by DksA widens the entrance for DNA into the active site cleft and weakens interactions with σ1.1, thereby favoring its removal from the cleft and increasing the rate of formation of RPo. Once formed, the stable open complex (RPo) is resistant to inhibitory effects of DksA/ppGpp (Figure 2.S1C; Rutherford et al., 2009) and reduces the binding affinity for DksA, allowing TL motion and catalysis (Lennon et al., 2009, Molodtsov et al., 2018).

In contrast, promoters inhibited by DksA/ppGpp tend to have an optimal -10 element, the absence of guanine residues on the NT strand in the upstream section of the discriminator element that interact with by σ1.2, and a G+C-rich downstream section of the discriminator element that disfavors melting and results in an unstable RPo (Sanchez-Vazquez et al., submitted; Haugen et al., 2006). These promoters can have high activity in the absence of regulators. DksA/ppGpp reprogramming further disfavors the melting step by destabilizing interactions with the T-strand due to β lobe displacement (Figure 2.6), leading to arrest of melting during early intermediates in the initiation pathway (Rutherford et al., 2009, Gopalkrishnan et al., 2017).

Transcriptional output from promoters regulated by transcription factors that interact directly with RNAP is determined by the interplay of many RNAP-promoter interactions and by the concentrations of the factors that contribute negatively or positively at different key steps along the promoter melting reaction coordinate (Haugen et al., 2008). In the case of DksA and ppGpp, the energy differences at play are seemingly subtle, caused by concerted small movements in flexible subdomains of RNAP involving a network of interactions of the clamp, βSI1, the trigger loop, and β'SI3, with the coiled-coil tip, the coiled-coil, and the C-terminal helix of DksA. These

Figure 2.6. Mechanism for transcription regulation by DksA/ppGpp. Schematic showing effects of DksA and ppGpp on RNAP regions and transcription initiation. RNAP swivel module (including the clamp and β'SI3), σ is shown in orange, the TL is shown in purple, and βlobe/βSI1 are shown in green. Binding of DksA (in blue) and ppGpp (red star) induce rotation of the swivel module, closing of the clamp, removal of σ1.1, and shifting of the βlobe/βSI1. At activated promoters, DksA/ppGpp-induced movements result in strand opening and DNA loading into the cleft. At negatively regulated promoters, DksA/ppGpp-induced movements result in destabilization of DNA in the cleft and inhibition of transcription.

Figure 2.6


Swivel
module
Clamp
σ
β'SI3
σ1.1
TL
βlobe/βSI1
+DksA/ppGpp
Swivel
Clamp
β'SI3
σ
DksA
ppGpp
σ1.1
βlobe/βSI1
Positively
Regulated
Promoter
Negatively
Regulated
Promoter

interactions are stabilized by ppGpp bound to site 2, accounting for its ability to regulate transcription in response to nutritional conditions. Even a single nucleotide change in the promoter can introduce a favorable protein contact able to deregulate transcription initiation (Haugen et al., 2006).

Despite the subtlety of the adjustments in RNAP, the stringent response is a robust and powerful regulatory mechanism allowing bacteria to employ a fast-diffusing small molecule inducer for rapid stress signaling likely across the entirety of the nucleoid, quickly reprograming nearly a fifth of the transcriptome (Sanchez-Vazquez et al., submitted). Allosteric control over RNAP motions, and hence promoter utilization mechanics, as well as the absence of site-specific interactions with DNA, grant DksA and ppGpp a regulatory scope that might be impossible to achieve so effectively by use of canonical DNA-binding TFs.

Materials and Methods

Construction of strains and plasmids

RNAP and DksA variants were generated by site-directed mutagenesis using the QuikChange Lightning Multi Site-Directed Mutagenesis Kit (Stratagene). Transformants were streaked and restreaked for single colonies, and the overexpression plasmids were purified and sequenced to verify the identity of the mutation. Plasmids for DNA footprinting were constructed by inserting EcoRI-HindIII digested fragments into pRLG1507. Endpoints are as noted. For pRLG9988, inserted fragment was additionally mutated using the QuikChange Lightning Multi Site-Directed Mutagenesis kit and primer 7714 to remove unwanted promoters downstream of *iraP*.

Purification of Proteins

Core RNAP containing ω was purified from BL21(DE3) containing the RNAP multi-subunit overexpression vector pIA900 or its variants (Table 1) using Ni-agarose and heparin affinity chromatography as previously described (Winkelman et al., 2015). Strains were grown in LB with 100μg/ml ampicillin at 37°C to OD_{600}=0.6. Expression was induced with 1mM IPTG. Cells were grown for 6hr after induction and then harvested by centrifugation. Cell pellets from 250ml cultures were resuspended in 5ml of BugBuster (Novagen), 23μg/ml phenylmethylsulfonyl fluoride (PMSF), and 5μl Lysonase (Novagen). Resuspended pellets were incubated at room temperature with gentle rocking for 30min before adding 15ml of RNAP resuspension buffer (1.4M NaCl, 40mM Tris-Cl pH 8.0, 30mM imidazole), followed by centrifugation at 14,000rpm for 40min at 4°C. The cleared lysate was added to 0.5ml pre-equilibrated Ni-NTA agarose (Qiagen), washed with RNAP wash buffer (300mM NaCl, 40mM Tris-Cl pH 8.0, 30mM imidazole), and eluted with wash buffer containing 300mM imidazole. The eluate was diluted to 200mM NaCl with TGED (10mM Tris-Cl pH 8.0, 5% glycerol, 0.1mM EDTA, 0.1mM DTT), and loaded onto 0.4ml of heparin sepharose (GE Healthcare) pre-equilibrated with TGED with 0.2M NaCl. The column was washed with 5ml of TGED containing 200mM NaCl, and RNAP was eluted with 1ml TGED containing 600mM NaCl. Eluate was concentrated using 5ml centrifugal filtration units (Corning) with a 100kDa molecular weight cutoff and exchanged with 2x storage buffer (20mM Tris-Cl pH 7.9, 200mM NaCl, 0.2mM DTT, 0.2mM EDTA). An equal volume of 100% glycerol was added to the remaining sample, and proteins were stored at -20°C or -80°C.

Protein concentration was measured using the Bradford assay reagent (Bio-Rad) using bovine serum albumen (BSA) as a standard.

Core RNAP lacking ω was purified from BL21(DE3) *ΔrpoZ::kan* containing the RNAP multi-subunit overexpression vector pIA299 or its variants (Table S1) by PEI precipitation, $AmmSO_4$ precipitation, Ni-NTA agarose chromatography, and heparin-sepharose chromatography as described previously (Ross et al., 2013; Vrentas et al., 2008). Cells were grown in 500mL LB with 100µg/ml ampicillin at 37°C to OD_{600}=0.6, induced with 1mM IPTG for 5-6 hours, and harvested by centrifugation. Cell pellets were suspended in 20ml lysis buffer (50mM Tris-Cl pH 8.0, 5% glycerol, 2mM EDTA, 233mM NaCl, 1mM BME, 0.1mM DTT), 23µg/ml PMSF, HALT Protease Inhibitor Cocktail (Thermo Scientific), and 0.15% sodium deoxycholate. Cells were lysed by sonication, diluted with TGED containing 0.2M NaCl, then centrifuged at 14,000rpm at 4°C for 20min. RNAP was purified from the cleared lysate by Polymin P precipitation followed by NH_4SO_4 precipitation overnight. Pellets were resuspended in RNAP wash buffer (20mM Tris-Cl pH 8.0, 500mM NaCl, 5% glycerol) containing 0.1% Tween20 and loaded onto a pre-equilibrated Ni-NTA agarose column (Qiagen). The column was washed with RNAP wash buffer containing 0.1% Tween20, RNAP wash buffer lacking 0.1% Tween20, and finally RNAP wash buffer containing 5mM imidazole. RNAP was eluted using RNAP wash buffer with 300mM imidazole. RNAP core enzyme was then purified using heparin chromatography, concentrated using centrifugal filtration units, and stored in 50% glycerol as above. Core RNAP lacking the TL was also purified using this method.

Core RNAP lacking β'SI3 was purified from BL21(DE3) *ΔrpoZ::kan* containing pIA331 as previously described (Artsimovitch et al., 2003). Strains were grown in 2L LB with 100μg/ml ampicillin at 30°C to OD_{600}=0.5, induced with 1mM IPTG for 3hr, then harvested by centrifugation. Pellets were resuspended in chitin buffer (20mM Tris-Cl pH 7.9, 5% glycerol, 500mM NaCl, 1mM EDTA) and HALT Protease Inhibitor Cocktail and lysed by sonication. Cleared lysate was added to 3ml pre-equilibrated chitin beads (New England Biolabs) and washed with 60ml buffer. Intein cleavage was induced by washing with 3 beds volumes of chitin buffer containing 50mM DTT, and then incubated at 4°C overnight. One column volume of chitin buffer was added to collect eluate. Core RNAP was then purified using heparin chromatography, concentrated using centrifugal filtration units, and stored in 50% glycerol as above.

DksA was purified from BL21(DE3) *ΔdksA::tet* as previously described (Paul et al., 2004; Ross et al., 2016). Strains were grown in 500ml LB with 30μg/ml kanamycin at 37°C to OD_{600}=0.4, induced with 1mM IPTG for 3hr at 30°C, and harvested by centrifugation. Pellets were resuspended in 20ml resuspension buffer (10mM Tris-Cl pH 8.0, 300mM NaCl, 5% glycerol, 2mM BME), 0.15% sodium deoxycholate, 0.25mg/ml lysozyme, 23μg/ml PMSF, and HALT protease inhibitor cocktail. Cells were lysed by sonication, and cleared lysate was loaded onto 0.4ml Ni-NTA agarose pre-equilibrated with wash buffer (300mM NaCl, 10mM Tris-Cl pH 8.0, 5mM imidazole, 2mM BME). The column was washed with 10ml wash buffer, 2ml wash buffer with 10mM imidazole, and eluted with 1.5ml wash buffer with 300mM imidazole. DksA was dialyzed into thrombin cleavage buffer overnight (20mM Tris-Cl pH 8.0, 150mM NaCl, 5mM $CaCl_2$, 2mM BME). The N-terminal His-tag was removed by cleavage with biotinylated thrombin (Novagen).

Thrombin and the cleaved His-tag were separated from DksA with Ni-NTA agarose and streptavidin agarose. Purified DksA was dialyzed against a storage buffer containing 10mM Tris-Cl pH 8.0, 100mM NaCl, 1mM DTT, and 50% glycerol.

Fluorescent assays and kinetic experiments

E. coli core RNAP and σ^{70} were obtained as described (Feklistov and Darst, 2011). Gp2 protein was expressed and purified as described (Bae et al., 2013). Myxopyronin and Ripostatin B were synthesized as described (Rentsch and Kalesse, 2012; Glaus and Altmann, 2012). Lipiarmycin (fidaxomicin) was purchased from Selleck Chemicals, USA. Promoter fragments were assembled as described (Feklistov and Darst, 2011) from DNA oligonucleotides (IDT, USA),

-10dsDNA

nt-strand: 5'-GTGGAAGCGCCGGCACGTATAATC-3',

t-strand: 5'-GATTATACGTGCCGGCGCTTCCAC-3',

For the beacon assay, RNAP holoenzyme containing σ^{70} harboring a single Cys residue at position 211 was labeled with TMR, and dsDNA promoter fragments were prepared as described (Feklistov and Darst, 2011; Mekler et al., 2011). All kinetic and equilibrium measurements with the RNAP beacon assay were carried out using a fluorescence plate reader (Infinite M1000, Tecan) at 25°C in binding buffer containing 40mM HEPES pH 8.0, 100mM KGlu, 5% glycerol, 1mM DTT, 10mM $MgCl_2$, 0.02% Tween-20. RNAP clamp modulators used in this work were added to the following final concentrations:

Gp2 protein (200nM), Ripostatin B (10μM), myxopyronin (5μM) and lipiarmycin (100μM). DksA was added to 2μM and ppGpp – to 100μM. Fluorescence intensities were recorded with an excitation wavelength of 550 nm and an emission wavelength of 578 nm with slit widths of 7 nm. For kinetic measurements of slow -10 dsDNA binding to RNAP on the plate reader the mixing dead time was ~20s. Typically, the final concentration of TMR-RNAP after mixing with DNA was 3nM.

Preparation of RNAP Complexes for Cryo-EM

E. coli holoenzyme (0.5 ml of 5 mg/ml) was purified from aggregates on a Superose 6 Increase column (GE Healthcare Life Sciences, Pittsburgh, PA) equilibrated with 20 mM Tris-HCl pH 8.0, 150 mM K-Glutamate, 5 mM MgCl2, 2.5 mM DTT. The peak fractions of the eluted protein were concentrated by centrifugal filtration (EMD-Millipore, Darmstadt, Germany) to 6 mg/mL protein concentration. The ligands were added in concentrations indicated above. The samples were incubated on ice for 15 min and then 3-([3-cholamidopropyl]dimethylammonio) 2-hydroxy-1-propane-sulfonate (CHAPSO) was added to the sample for a final concentration of 8 mM prior to grid preparation.

Cryo-EM grid preparation

C-flat CF-1.2/1.3-4Au 400 mesh gold grids (Protochips, Morrisville, NC) were glow-discharged for 10 s prior to the application of 3.5 μl of the sample (5 – 10 mg/ml protein concentration). After blotting for 3–4 s, the grids were plunge-frozen in liquid ethane using an FEI Vitrobot Mark IV (FEI, Hillsboro, OR) with 90% chamber humidity at 22 ℃.

Cryo-EM data acquisition and processing

Structural biology software was accessed through the SBGrid consortium (Morin et al., 2013). The grids were imaged using a 300 keV Titan Krios (FEI) equipped with a K2 Summit direct electron detector (Gatan, Warrendale, PA). Images were recorded with Leginon (Nicholson et al., 2010) in counting mode with a pixel size of 1.07 Å and a defocus range of 0.8 mm to 1.8 mm. Data were collected with a dose of 8 electrons/px/s. Images were recorded over a 10 s exposure with 0.2 s frames (50 total frames) to give a total dose of 66 electrons/$Å^2$. Dosefractionated subframes were aligned and summed using MotionCor2 (Zheng et al., 2017) and subsequent doseweighting was applied to each image. The contrast transfer function was estimated for each summed image using Gctf (Zhang, 2016). From the summed images, Gautomatch (developed by K. Zhang, MRC Laboratory of Molecular Biology, Cambridge, UK, http://www.mrc-lmb.cam.ac. uk/kzhang/Gautomatch) was used to pick particles with an auto-generated template. Autopicked particles were manually inspected, then subjected to 2D classification in RELION (Scheres, 2012) specifying 50 classes. Poorly populated classes were removed, and particles from ‘good’ 2D classes were 3D auto-refined in RELION using this ab-initio 3D template (low-pass filtered to 60 Å-resolution). RELION 3D classification into several classes was performed on the particles using the refined map and alignment angles. Among the 3D classes, the best-resolved classes, were 3D auto-refined and post-processed in RELION. Local resolution calculations were performed using blocres (Cardone et al., 2013).

Model building and refinement

To build initial models of the protein components of the complex model for E. coli RNAP holo was manually fit into the cryo-EM density maps using Chimera (Pettersen et al., 2004) and real-space refined using Phenix (Adams et al., 2010). In the real-space refinement, domains of RNAP were rigid-body refined. For the high-resolution structures, the rigid-body refined models were subsequently refined with secondary structure restraints. A model of Lpm was generated from a crystal structure (Serra et al., 2017), edited in Phenix REEL, and refined into the cryo-EM density. Refined models were inspected and modified in Coot (Emsley and Cowtan, 2004) according to cryo-EM maps, followed by further real-space refinement with PHENIX.

Multi-round In Vitro Transcription

Reactions with *rrnB* P1, *thrABC*, or *iraP* promoters were carried out at 30°C or 37°C with 50ng plasmid DNA template in buffer containing 10mM Tris-Cl pH8, 165mM NaCl, 10mM MgCl2, 1mM DTT, 100µg/ml BSA, 500µM ATP, 100µM GTP, 100µM CTP, 10µM UTP, and 1.5µCi α-32P UTP (Perkin Elmer). For reactions containing DksA and ppGpp, 1uM and 100uM were used, respectively, unless otherwise stated. Reactions were initiated with the addition of 20-30nM RNAP and allowed to proceed for 15min before quenching with an equal volume of transcription stop solution (7M urea, 2xTBE, 10mM EDTA, 1% SDS, and 0.05% bromophenol blue). Transcripts were visualized on a 5.5% polyacrylamide/7M urea gel and quantified using ImageQuant as described previously (Ross et al., 2013). Transcription with Δσ1.1 in Figure 2.S6 was performed at 23°C.

$KMnO_4$ Footprinting

$KMnO_4$ footprinting was carried out as previously described (Newlands et al., 1991). Plasmid pRLG9988 (*iraP*) or pRLG11272 (*rpsT* P2) were used to obtain radio-labeled template for footprinting. To label the non-template strand, 15ug plasmid was digested with NheI-HF (NEB), purified through phenol extraction and ethanol precipitation, end-labeled with α-32P dCTP (Perkin-Elmer) using Sequenase (Affymetrix), purified from unincorporated nucleotides through ethanol precipitation, and digested with NcoI-HF (NEB). To label the template strand, 15ug plasmid was digested with NcoI-HF, end-labeled with α-32P dCTP using Sequenase, and digested with NheI-HF. The resulting sample was run on a native 5% acrylamide gel. The promoter fragment was excised from the gel, diffused overnight into low salt elution buffer (200mM NaCl, 20mM Tris-Cl pH7.4, 1mM EDTA), purified using the Qiagen PCR purification kit, and stored in 10mM Tris-Cl pH8 buffer.

For footprinting at *rpsT* P2, 30nM RNAP was incubated with ~0.2nM template DNA in the presence or absence of 1μM DksA and 100μM ppGpp in 30mM KCl transcription buffer (10mM Tris-Cl pH8, 30mM KCl, 10mM MgCl2, 100μg/ml BSA) at 30°C for 10min. Reactions were treated with 2mM KMnO4 for 30 seconds before being quenched with β-mercaptoethanol (275mM final concentration). Reactions were ethanol precipitated with sodium acetate, resuspended in 2M ammonium acetate, and precipitated again with ethanol. The pellet was resuspended in 100μl 1M piperidine, incubated at 90°C for 30min, ethanol precipitated with sodium acetate, and resuspended in 4μl loading buffer (7M urea, 0.5x TBE, 0.05% bromophenol blue, 0.05% xylene cyanol).

A+G ladders were prepared using Maxam-Gilbert sequencing as previously described (Ross et al., 2001). 10µl DNA was incubated with 50µl formic acid at 23°C for 7min, ethanol precipitated, resuspended in 100µl 1M piperidine, incubated at 90°C for 30min, ethanol precipitated, and resuspended in 50µl loading buffer.

All samples were heated briefly to 90°C and run on a 9.5% acrylamide, 7M urea sequencing gel with 0.5X TBE. Gels were dried and exposed to a phosphorimager screen overnight.

Rate Analysis

Determination of rate constants at P*iraP* was performed based on the analysis from Drennan et al., 2012. Open complex formation rates were determined using $KMnO_4$ footprinting. Percent reactivity was calculated as the total the signal of the -11 to -2 bands divided by the total signal in the lane. Percent reactivity was plotted over time and fit to the equation: f = a*(1-exp(-β*x)), where β is the decay to equilibrium rate constant. β is related to rate constants k_d and α of dissociation and association, respectively, by $\beta = k_d + \alpha$, where $\alpha = K_1k_2[RNAP]/(1+K_1[RNAP])$ and $k_a = K_1k_2$. RNAP concentrations above 20nM were not included in the analysis due to artifacts at higher [RNAP].

k_d was determined by open complex stability assays using the same solution conditions on linear P*iraP* fragment DNA. Briefly, open complexes were allowed to form at 37°C before the addition of heparin at t=0. NTPs were added at various times after heparin addition to measure the fraction of open complexes remaining. α was plotted vs [RNAP] and fit to the hyperbolic equation f = ax/(b+x), where $a = k_2$ and $b = 1/K_1$.

Fe2+ Cleavage Assay

Binding of DksA to RNAP was measured using hydroxyl radical cleavage as described previously (Lennon et al., 2009). To 32-P label DksA, HMK-His6-DksA was incubated with γ-32P ATP and Protein Kinase A (Sigma) at 37°C for 30min in Buffer A (20mM HEPES pH7.5, 20mM NaCl, 10mM MgCl2). Free γ-32P ATP was removed using 2 G-50 columns (GE Healthcare) preequilibrated with Buffer A lacking MgCl2. MgCl2 and glycerol were similarly removed from RNAP preps using 2 G-50 columns.

RNAP at various concentrations was incubated with labeled DksA in buffer lacking MgCl2 for 10min at 23°C. Cleavage was induced by simultaneous addition of 10mM DTT and 10mM ferrous ammonium sulfate. Reactions proceeded for 10min and quenched with 2x loading buffer (125mM TrisCl pH8.0, 4% SDS, 20% glycerol, 1.4M β-mercaptoethanol, 0.1% bromophenol blue). Cleavage products were separated on 4-12% SDS gels and analyzed by phosphorimaging.

^{32}P-ppGpp Binding Assay (DRaCALA)

^{32}P-ppGpp binding to Site 2 at the RNAP-DksA interface was measured by using the DRaCALA assay (Differential Radial Capillary Action of Ligand; Roelofs et al., 2011) using RNAP lacking binding Site 1 at the ω-β' interface, as described previously (Ross et al., 2016). Briefly, ^{32}P-ppGpp was synthesized using γ-^{32}P-ATP and GDP (Amresco), and purified by using PEI cellulose thin layer chromatography with 1.5 M KH_2PO_4 pH 3.4. ^{32}P-ppGpp was eluted from the TLC plate in 4 M LiCl, precipitated, and stored at -80°C (Ross et al., 2016). Purity and identity of the resulting ^{32}P-labeled compound was verified by its co-migration on an analytical TLC plate with unlabeled ppGpp purchased

from TriLink Biotechnologies (imaged by UV shadowing). No other labeled compounds were present in the purified preparation.

Binding reactions (15µl) were carried out at 22ºC for 10 min. in 20mM Tris-Cl pH 7.5, 10mM $MgCl_2$, 100mM NaCl, 0.5mM β-mercaptoethanol, ~10nM ^{32}P-ppGpp, 2µM Δω RNAP (lacking ppGpp binding Site 1), and 20 µM WT or variant DksAs (a saturating concentration for DksA). Duplicate 4µl aliquots were spotted slowly onto dry nitrocellulose filter discs (Protran BA85; GE Healthcare) and dried filters were quantified by phosphorimaging. Protein-bound counts (in the central spot) were determined by correction of counts for background of unbound ^{32}P-ppGpp, and were expressed as the percent of total counts in the entire spot (adapted from Roelofs et al., 2011; Ross et al., 2016). Values shown (Fig. 3B) represent averages and ranges from two independent experiments.

Acknowledgements

We thank members of the Darst lab for the Δσ1.1 protein. We thank R. Saecker and M.T. Record for useful discussions on analyzing transcription initiation kinetics. We thank J.H. Lee for constructing pRLG13311 and pRLG13318. Work in the Gourse/Ross lab is supported by NIH R37 GM37048.

References

Aberg, A., Fernández-Vázquez, J., Cabrer-Panes, J.D., Sánchez, A., and Balsalobre, C. (2009). Similar and divergent effects of ppGpp and DksA deficiencies on transcription in Escherichia coli. *J. Bacteriol.* **191**:3226-3236.

Adams, P.D., Afonine, P.V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., McCoy, A.J., Moriarty, N.W., Oeffner, R., Read, R.J., Richardson, D.C., Richardson, J.S., Terwilliger,

T.C., and Zwart, P.H. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D. Biol. Crystallogr.* **66**:213-221.

Artsimovitch, I., Svetlov, V., Murakami, K.S., and Landick, R. (2003). Co-overexpression of Escherichia coli RNA polymerase subunits allows isolation and analysis of mutant enzymes lacking lineage-specific sequence insertions. *J. Biol. Chem.* **278**:12344-12355.

Bae, B., Davis, E., Brown, D., Campbell, E.A., Wigneshweraraj, S., and Darst, S.A. (2013). Phage T7 Gp2 inhibition of Escherichia coli RNA polymerase involves misappropriation of σ70 domain 1.1. *Proc. Natl. Acad. Sci. USA.* **110**:19772-19777.

Barker, M.M., Gaal, T., Josaitis, C.A., and Gourse, R.L. (2001). Mechanism of regulation of transcription initiation by ppGpp. I. Effects of ppGpp on transcription initiation in vivo and in vitro. *J. Mol. Biol.* **305**:673-688.

Bartlett, M.S., Gaal, T., Ross, W., and Gourse, R.L. (1998). RNA polymerase mutants that destabilize RNA polymerase-promoter complexes alter NTP-sensing by rrn P1 promoters. *J. Mol. Biol.* **279**:331-345.

Boyaci, H., Chen, J., Lilic, M., Palka, M., Mooney, R.A., Landick, R., Darst, S.A., and Campbell, E.A. (2018). Fidaxomicin jams *Mycobacterium tuberculosis* RNA polymerase motions needed for initiation via RbpA contacts. *Elife* **7**:e34823.

Burgos, H.L. (2017). Dynamic Regulation of Ribosomal Gene Expression During Adaptation to Fluctuating Nutritional Conditions. Ph.D. dissertation.

Cardone, G., Heymann, J.B., and Steven, A.C. (2013). One number does not fit all: mapping local variations in resolution in cryo-EM reconstructions. *J. Struct. Biol.* **184**:226-236.

Chakraborty, A., Wang, D., Ebright, Y.W., Korlann, Y., Kortkhonjia, E., Kim, T., Chowdhury, S., Wigneshweraraj, S., Irschik, H., Jansen, R., Nixon, B.T., Knight, J., Weiss, S., and Ebright, R.H. (2012). Opening and closing of the bacterial RNA polymerase clamp. *Science* **337**:51-595.

Chandrangsu, P., Lemke, J.J., and Gourse, R.L. (2011). The dksA promoter is negatively feedback regulated by DksA and ppGpp. *Mol. Microbiol.* **80**:1337-1348.

Dalebroux, Z.D., Svensson, S.L., Gaynor, E.C., and Swanson, M.S. (2010). ppGpp conjures bacterial virulence. *Microbiol. Mol. Biol. Rev.* **74**:171-199.

Drennen, A., Kraemer, M., Capp, M., Gries, T., Ruff, E., Sheppard, C., Wigneshweraraj, S., Artsimovitch, I., and Record, M.T. Jr. (2012). Key roles of the downstream

mobile jaw of Escherichia coli RNA polymerase in transcription initiation. *Biochemistry* **51**:9447-9459.

Emsley, P. and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta. Crystallogr. D. Biol. Crystallogr.* **60**:2126-2132.

Erickson, D.L., Lines, J.L., Pesci, E.C., Venturi, V., and Storey, D.G. (2004). Pseudomonas aeruginosa relA contributes to virulence in Drosophila melanogaster. *Infect. Immun.* **72**:5638-5645.

Feklistov, A., Barinova, N., Sevostyanova, A., Heyduk, E., Bass, I., Vvedenskaya, I., Kuznedelov, K., Merkiene, E., Stavrovskaya, E., Klimasauskas, S., Nikiforov, V., Heyduk, T., Severinov, K., and Kulbachinskiy, A. (2006). A basal promoter element recognized by free RNA polymerase sigma subunit determines promoter recognition by RNA polymerase holoenzyme. *Mol. Cell* **23**:97-107.

Feklistov, A. and Darst, S.A. (2011). Structural basis for promoter-10 element recognition by the bacterial RNA polymerase σ subunit. *Cell* **147**:1257-1269.

Feklistov, A., Bae, B., Hauver, J., Lass-Napiorkowska, A., Kalesse, M., Glaus, F., Altmann, K.H., Heyduk, T., Landick, R., and Darst, S.A. (2017). RNA polymerase motions during promoter melting. *Science* **356**:863-866.

Glaus, F. and Altmann, K.H. (2012). Total synthesis of the bacterial RNA polymerase inhibitor ripostatin B. *Angew. Chem. Int. Ed. Engl.* **51**:3405-3409.

Gopalkrishnan, S., Ross, W., Chen, A.Y., and Gourse, R.L. (2017). TraR directly regulates transcription initiation by mimicking the combined effects of the global regulators DksA and ppGpp. *Proc. Natl. Acad. Sci. USA.* **114**:E5539-E5548.

Gourse, R.L., Chen, A.Y., Gopalkrishnan, S., Sanchez-Vazquez, P., Myers, A., and Ross, W. (2018). Transcriptional Responses to ppGpp to DksA. *Annu. Rev. Microbiol.* **72**:163-184.

Greenway, D.L. and England, R.R. (1999). The intrinsic resistance of Escherichia coli to various antimicrobial agents requires ppGpp and sigma s. *Lett. Appl. Microbiol.* **29**:323-326.

Haugen, S.P., Berkmen, M.B., Ross, W., Gaal, T., Ward, C., and Gourse, R.L. (2006). rRNA promoter regulation by nonoptimal binding of sigma region 1.2: an additional recognition element for RNA polymerase. *Cell* **125**:1069-1082.

Haugen, S.P., Ross, W., and Gourse, R.L. (2008). Advances in bacterial promoter recognition and its control by factors that do not bind DNA. *Nat. Rev. Microbiol.* **6**:507-519.

Hauryliuk, V., Atkinson, G.C., Murakami, K.S., Tenson, T., and Gerdes, K. (2015). Recent functional insights into the role of (p)ppGpp in bacterial physiology. *Nat. Rev. Microbiol.* **13**:298-309.

Kang, J.Y., Olinares, P.D., Chen, J., Campbell, E.A., Mustaev, A., Chait, B.T., Gottesman, M.E., and Darst, S.A. (2017). Structural basis of transcription arrest by coliphage HK022 Nun in an *Escherichia coli* RNA polymerase elongation complex. *Elife* **6**:e25478.

Kang, J.Y., Mishanina, T.V., Bellecourt, M.J., Mooney, R.A., Darst, S.A., and Landick, R. (2018). RNA Polymerase Accomodates a Pause RNA Hairpin by Global Conformational Rearrangements that Prolong Pausing. *Mol. Cell* **69**:802-815.E1.

Kim, S., Watanabe, K., Suzuki, H., and Watarai, M. (2005). Roles of Brucella abortus SpoT in morphological differentiation and intramacrophagic replication. *Microbiology* **151**:1607-1617.

Korch, S.B., Henderson, T.A., and Hill, T.M. (2003). Characterization of the hipA7 allele of Escherichia coli and evidence that high persistence is governed by (p)ppGpp synthesis. *Mol. Microbiol.* **50**:1199-1213.

Krásný, L. and Gourse, R.L. (2004). An alternative strategy for bacterial ribosome synthesis: Bacillus subtilis rRNA transcription regulation. *EMBO J.* **23**:4473-4483.

Lemke, J.J., Sanchez-Vazquez, P., Burgos, H.L., Hedberg, G., Ross, W., and Gourse, R.L. (2011). Direct regulation of Escherichia coli ribosomal protein promoters by the transcription factors ppGpp and DksA. *Proc. Natl. Acad. Sci. USA.* **108**:5712-5717.

Lennon, C.W., Gaal, T., Ross, W., and Gourse, R.L. (2009). Escherichia coli DksA binds to Free RNA polymerase with higher affinity than to RNA polymerase in an open complex. *J. Bacteriol.* **191**:5854-5858.

Lennon, C.W., Ross, W., Martin-Tumasz, S., Toulokhonov, I., Vrentas, C.E., Rutherford, S.T., Lee, J.H., Butcher, S.E., and Gourse, R.L. (2012). Direct interactions between the coiled-coil tip of DksA and the trigger loop of RNA polymerase mediate transcriptional regulation. *Genes Dev.* **26**:2634-2646.

Lin, W., Das, K., Degen, D., Mazumder, A., Duchi, D., Wang, D., Ebright, Y.W., Ebright, R.Y., Sineva, E., Gigliotti, M., Srivastava, A., Mandal., S., Jiang, Y., Liu, Y., Yin, R., Zhang, Z., Eng, E.T., Thomas, D., Donadio, S., Zhang, H., Zhang, C., Kapanidis, A.N., and Ebright, R.H. (2018). Structural Basis of Transcription Inhibition by Fidaxomicin (Lipiarmycin A3). *Mol. Cell* **70**:60-71.e15.

Liu, K., Myers, A.R., Pisithkul, T., Claas, K.R., Satyshur, K.A., Amador-Noguez, D., Keck, J.L., and Wang, J.D. (2015). Molecular mechanism and evolution of guanylate kinase regulation by (p)ppGpp. *Mol. Cell* **57**:735-749.

Mekler, V., Kortkhonjia, E., Mukhopadhyay, J., Knight, J., Revyakin, A., Kapanidis, A.N., Niu, W., Ebright, Y.W., Levy, R., and Ebright, R.H. (2002). Structural organization of bacterial RNA polymerase holoenzyme and the RNA polymerase-promoter open complex. *Cell* **108**:599-614.

Mekler, V., Pavlova, O., and Severinov, K. (2011). Interaction of Escherichia coli RNA polymerase σ70, RNA polymerase holoenzyme, and the β'-σ70 complex. *J. Biol. Chem.* **286**:270-279.

Molodtsov, V., Sineva, E., Zhang, L., Huang, X., Cashel, M., Ades, S.E., and Murakami, K.S. (2018). Allosteric Effector ppGpp Potentiates the Inhibition of Transcript Initiation by DksA. *Mol. Cell* **69**:828-839.e5.

Morin, A., Eisenbraun, B., Key, J., Sanschagrin, P.C., Timony, M.A., Ottaviano, M., and Sliz, P. (2013). Collaboration gets the most out of software. *Elife* **2**:e01456.

NandyMazumdar, M., Nedialkov, Y., Svetlov, D., Sevostyanova, A., Belogurov, G.A., and Artsimovitch, I. (2016). RNA polymerase gate loop guides the nontemplate DNA strand in transcription complexes. *Proc. Natl. Acad. Sci. USA.* **113**:14994-14999.

Nayak, D., Voss, M., Windgassen, T., Mooney, R.A., and Landick, R. (2013). Cys-pair reporters detect a constrained trigger loop in a paused RNA polymerase. *Mol. Cell* **50**:882-893.

Newlands, J.T., Ross, W., Gosink, K.K., and Gourse, R.L. (1991). Factor-independent activation of Escherichia coli rRNA transcription. II. Characterization of complexes of rrnB P1 promoters containing or lacking the upstream activator region of Escherichia coli RNA polymerase. *J. Mol. Biol.* **220**:569-583.

Nicholson, W.V., White, H., and Trinick, J. (2010). An approach to automated acquisition of cryoEM images from lacey carbon grids. *J. Struct. Biol.* **172**:395-399.

Parshin, A., Shiver, A.L., Lee, J., Ozerova, M., Schneidman-Duhovny, D., Gross, C.A., and Borukhov, S. (2015). DksA regulates RNA polymerase in Escherichia coli through a network of interactions in the secondary channel that includes Sequence Insertion 1. *Proc. Natl. Acad. Sci. USA.* **112**:E6862-E6871.

Paul, B.J., Barker, M.M., Ross, W., Schneider, D.A., Webb, C., Foster, J.W., and Gourse, R.L. (2004). DksA: a critical component of the transcription initiation machinery that potentiates the regulation of rRNA promoters by ppGpp and the initiating NTP. *Cell* **118**:311-322.

Paul, B.J., Berkmen, M.B., and Gourse, R.L. (2005). DksA potentiates direct activation of amino acid promoters by ppGpp. *Proc. Natl. Acad. Sci. USA.* **102**:7823-7828.

Perederina, A., Svetlov, V., Vassylyeva, M.N., Tahirov, T.H., Yokoyama, S., Artsimovitch, I., and Vassylyev, D.G. (2004). Regulation through the secondary channel—structural framework for ppGpp-DksA synergism during transcription. *Cell* **118**:297-309.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera—a visualization system for exploration research and analysis. *J. Comput. Chem.* **25**:1605-1612.

Pomares, M.F., Vincent, P.A., Farías, R.N., and Salomón, R.A. (2008). Protective action of ppGpp in microcin J25-sensitive strains. *J. Bacteriol.* **190**:4328-4334.

Potrykus, K. and Cashel, M. (2008). (p)ppGpp: still magical? *Annu. Rev. Microbiol.* **62**:35-51.

Rentsch, A. and Kalesse, M. (2012). The total synthesis of corallopyronin A and myxopyronin B. *Angew. Chem. Int. Ed. Engl.* **51**:11381-11384.

Roelofs, K.G., Wang, J., Sintim, H.O., and Lee, V.T. (2011). Differential radial capillary action of ligand assay for high-throughput detection of protein-metabolite interactions. *Proc. Natl. Acad. Sci. USA.* **108**:15528-15533.

Ross, W., Thompson, J.F., Newlands, J.T., and Gourse, R.L. (1990). E. coli Fis protein activates ribosomal RNA transcription in vitro and in vivo. *EMBO J.* **9**:3733-3742.

Ross, W., Ernst, A., and Gourse, R.L. (2001). Fine structure of E. coli RNA polymerase-promoter interactions: alpha subunit binding to the UP element minor groove. *Genes Dev.* **15**:491-506.

Ross, W., Vrentas, C.E., Sanchez-Vazquez, P., Gaal, T., and Gourse, R.L. (2013). The magic spot: a ppGpp binding site on E. coli RNA polymerase responsible for regulation of transcription initiation. *Mol. Cell* **50**:420-429.

Ross, W., Sanchez-Vazquez, P., Chen, A.Y., Lee, J.H., Burgos, H.L., and Gourse, R.L. (2016). ppGpp Binding to a Site at the RNAP-DksA Interface Accounts for Its Dramatic Effects on Transcription Initiation during the Stringent Response. *Mol. Cell* **62**:811-823.

Rutherford, S.T., Lemke, J.J., Vrentas, C.E., Gaal, T., Ross, W., and Gourse, R.L. (2007). Effects of DksA, GreA, and GreB on transcription initiation: insights into the mechanisms of factors that bind in the secondary channel of RNA polymerase. *J. Mol. Biol.* **366**:1243-1257.

Rutherford, S.T., Villers, C.L., Lee, J.H., Ross, W., and Gourse, R.L. (2009). Allosteric control of Escherichia coli rRNA promoter complexes by DksA. *Genes Dev.* **23**:236-248.

Ryals, J., Little, R., and Bremer, H. (1982). Control of rRNA and tRNA syntheses in Escherichia coli by guanosine tetraphosphate. *J. Bacteriol.* **151**:1261-1268.

Sanchez-Vazquez, P., Dewey, C.N., Kitten, N., Ross, W., and Gourse, R.L. Genome-wide effects of *Escherichia coli* transcription from ppGpp binding to its two sites on RNA polymerase. *Submitted.*

Scheres, S.H. (2012). RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.* **180**:519-530.

Serra, S., Malpezzi, L., Bedeschi, A., Fuganti, C., and Fonte, P. (2017). Final Demonstration of the Co-Identity of Lipiarmycin A3 and Tiacumicin B (Fidaxomicin) through Single Crystal X-ray Analysis. *Antibiotics (Basel)* **6**:7.

Svetlov, V. and Artsimovitch, I. (2015). Purification of bacterial RNA polymerase: tools and protocols. *Methods Mol. Biol.* **1276**:13-29.

Traxler, M.F., Zacharia, V.M., Marquardt, S., Summers, S.M., Nguyen, H.T., Stark, S.E., and Conway, T. (2011). Discretely calibrated regulatory loops controlled by ppGpp partition gene induction across the 'feast to famine' gradient in Escherichia coli. *Mol. Microbiol.* **79**:830-845.

Toulokhonov, I., Zhang, J., Palangat, M., and Landick, R. (2007). A central role of the RNA polymerase trigger loop in active-site rearrangement during transcriptional pausing. *Mol. Cell* **27**:406-419.

Varik, V., Oliveira, S.R.A., Hauryliuk, V., and Tenson, T. (2017). HPLC-based quantification of bacterial housekeeping nucleotides and alarmone messengers ppGpp and pppGpp. *Sci. Rep.* **7**:11022.

Vrentas, C.E., Gaal, T., Berkmen, M.B., Rutherford, S.T., Haugen, S.P., Vassylyev, D.G., Ross, W., and Gourse, R.L. (2008). Still looking for the magic spot: the crystallographically defined binding site for ppGpp on RNA polymerase is unlikely to be responsible for rRNA transcription regulation. *J. Mol. Biol.* **377**:551-564.

Windgassen, T.A., Mooney, R.A., Nayak, D., Palangat, M., Zhang, J., and Landick, R. (2014). Trigger-helix folding pathway and SI3 mediate catalysis and hairpin-stabilized pausing by Escherichia coli RNA polymerase. *Nucleic Acids Res.* **42**:12707-12721.

Winkelman, J.T., Winkelman, B.T., Boyce, J., Maloney, M.F., Chen, A.Y., Ross, W., and Gourse, R.L. (2015). Crosslink Mapping at Amino Acid-Base Resolution Reveals the Path of Scrunched DNA in Initial Transcribing Complexes. *Mol. Cell* **59**:768-780.

Zhang, J., Palangat, M., and Landick, R. (2010). Role of the RNA polymerase trigger loop in catalysis and pausing. *Nat. Struct. Mol. Biol.* **17**:99-104.

Zhang, Y., Feng, Y., Chatterjee, S., Tuske, S., Ho, M.X., Arnold, E., and Ebright, R.H. (2012). Structural basis of transcription initiation. *Science* **338**:1076-1080.

Zhang, K. (2016). Gctf: Real-time CTF determination and correction. *J. Struct. Biol.* **193**:1-12.

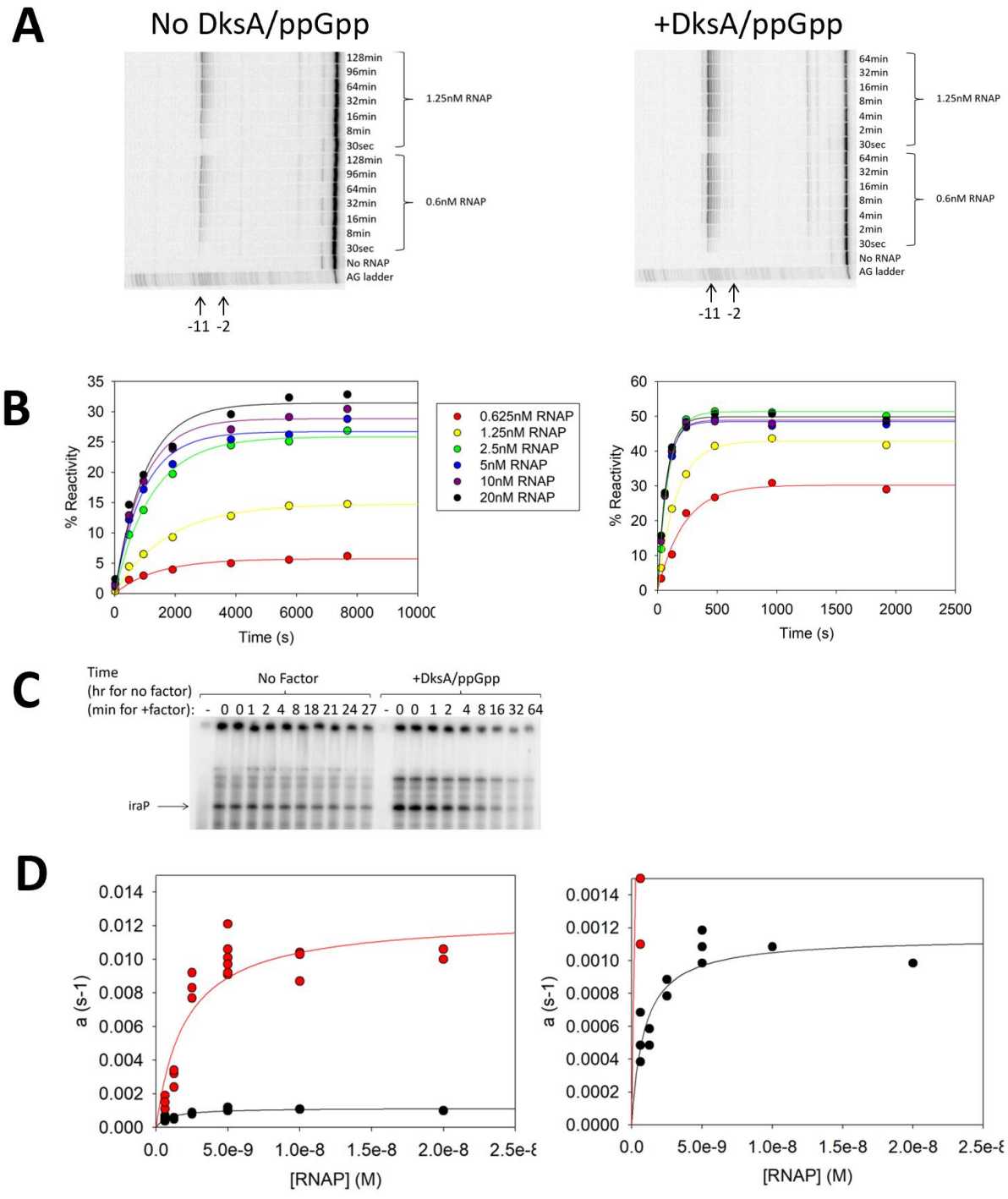
Zheng, S.Q., Palovcak, E., Armache, J.P., Verba, K.A., Cheng, Y., and Agard, D.A. (2017). MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**:331-332.

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. *Mol. Cell* **50**:430-436.

Supplemental Materials

Figure 2.S1. DksA and ppGpp do not affect K_B at P*iraP*. (A) Representative $KMnO_4$ footprinting gels at P*iraP* in the presence or absence of 1uM DksA and 100uM ppGpp. **(B)** Representative graph of $KMnO_4$ data from (A) plotting reactivity of transcription bubble as a function of time. **(C)** Representative open complex decay assay where open complexes at P*iraP* are challenged with heparin at t=0 and fraction of open complexes remaining is measured by *in vitro* transcription. Reactions were performed in the presence or absence of 1 uM DksA and 100 uM ppGpp. **(D)** Plot of the rate of association α vs [RNAP] obtained from $KMnO_4$ and open complex decay experiments (left). RNAP alone shown in black and RNAP+DksA/ppGpp shown in red. Plot on the right shows the same plot with a different y-axis scale. **(E)** Using a simplified two step mechanism, where RNAP binds DNA to form RP_C in step 1 and RP_C isomerizes to form RP_O in step 2, K_1 and k_2 calculated from α vs [RNAP], where K_1 represents the equilibrium constant for forming RP_C and k_2 is the forward isomerization rate constant where RP_C isomerizes to RP_O. See methods for more details.

Figure 2.S1



	K_1 (M^{-1})	k_2 (s^{-1})
RNAP alone	$(1.5 \pm 0.4) \times 10^{9}$	$(1.00 \pm 0.05) \times 10^{-3}$
+DksA/ppGpp	$(7 \pm 2) \times 10^{8}$	$(1.09 \pm 0.08) \times 10^{-2}$

Figure 2.S2. Myx and Lpm affect strand opening at P*iraP*. Representative gel showing raw data for Figure 2.1E.

Figure 2.S2

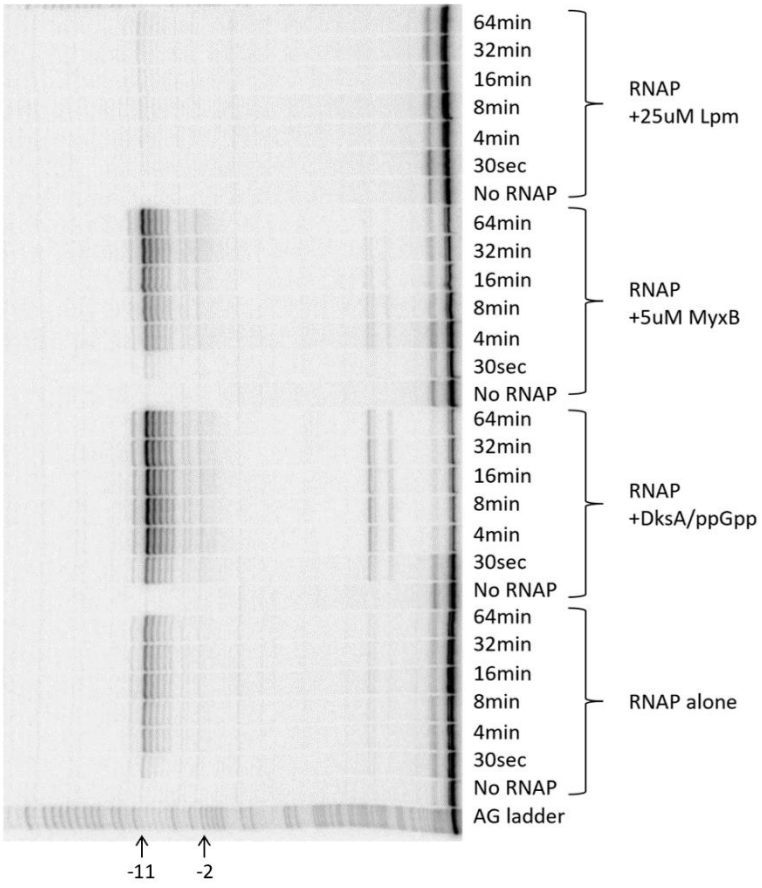

64min
32min
16min
8min
4min
30sec
No RNAP
RNAP
+25uM Lpm
64min
32min
16min
8min
4min
30sec
No RNAP
RNAP
+5uM MyxB
64min
32min
16min
8min
4min
30sec
No RNAP
RNAP
+DksA/ppGpp
64min
32min
16min
8min
4min
30sec
No RNAP
RNAP alone
AG ladder
-11
-2

Figure 2.S3. L86A DksA and R93A K97A DksA are defective for function. (A) Representative gel showing activation of P*iraP* by DksA variants in the presence or absence of ppGpp. Gel quantified and graphed in Figure 2.2B. **(B)** Representative gel showing inhibition of *rrnB* P1 by DksA variants (15uM) with increasing concentrations of ppGpp. Gel quantified and graphed in Figure 2.2C.

Figure 2.S3

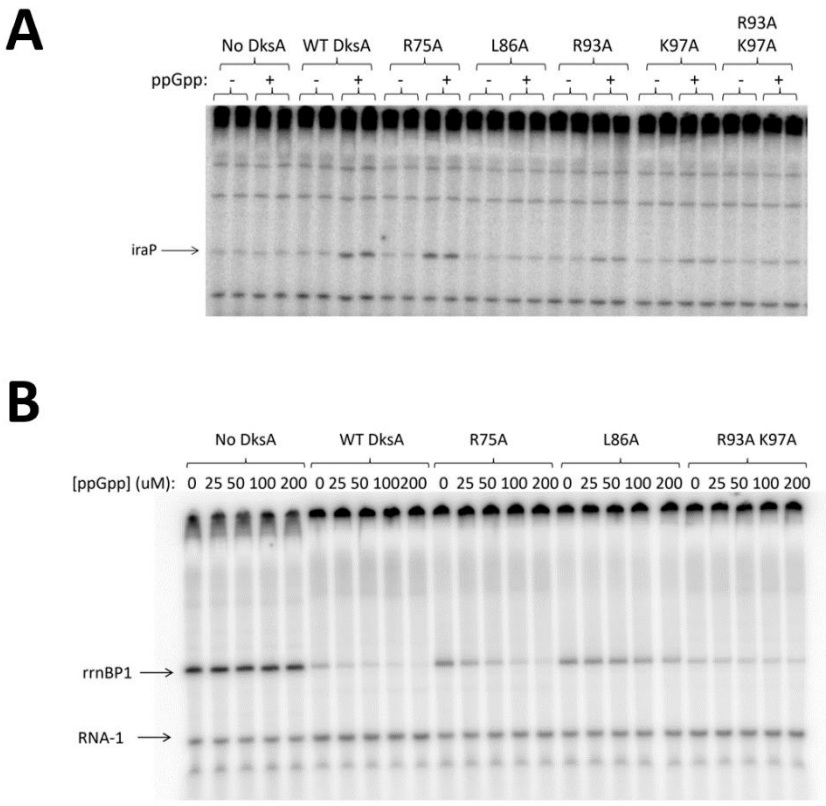

A
No DksA
WT DksA
R75A
L86A
R93A
K97A
R93A K97A
ppGpp: - + - + - + - + - + - + - +
iraP
B
No DksA
WT DksA
R75A
L86A
R93A K97A
[ppGpp] (uM): 0 25 50 100 200 0 25 50 100200 0 25 50 100 200 0 25 50 100 200 0 25 50 100 200
rrnBP1
RNA-1

Figure 2.S4. L86A DksA and R93A K97A DksA are not defective for binding RNAP. **(A)** Representative gel measuring the binding affinity of DksA variants to RNAP using the iron cleavage assay. **(B)** Quantification of iron cleavage data. n=3.

Figure 2.S4

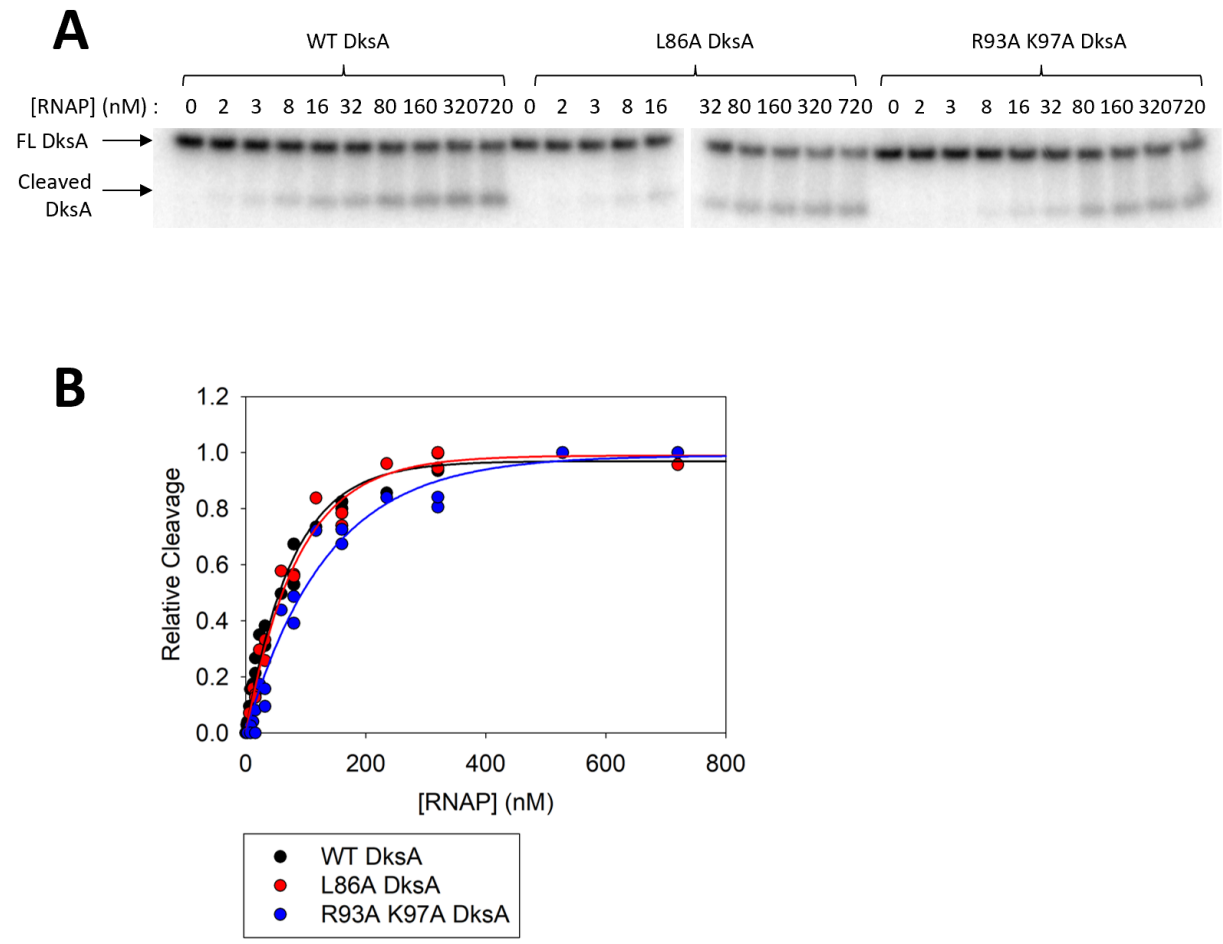

A
WT DksA
L86A DksA
R93A K97A DksA
[RNAP] (nM) : 0 2 3 8 16 32 80 160 320 720 0 2 3 8 16 32 80 160 320 720 0 2 3 8 16 32 80 160 320 720
FL DksA
Cleaved DksA
B
Relative Cleavage
1.2
1.0
0.8
0.6
0.4
0.2
0.0
0
200
400
600
800
[RNAP] (nM)
WT DksA
L86A DksA
R93A K97A DksA

Figure 2.S5. Δβ'SI3 is partially defective for activation at P*iraP*. (A) Representative gel showing multi-round *in vitro* transcription of P*iraP* by WT RNAP or Δβ'SI3. All reactions have 100uM ppGpp. **(B)** Quantification of transcription data from (A). WT RNAP shown in black. Δβ'SI3 shown in red. n=3.

Figure 2.S5

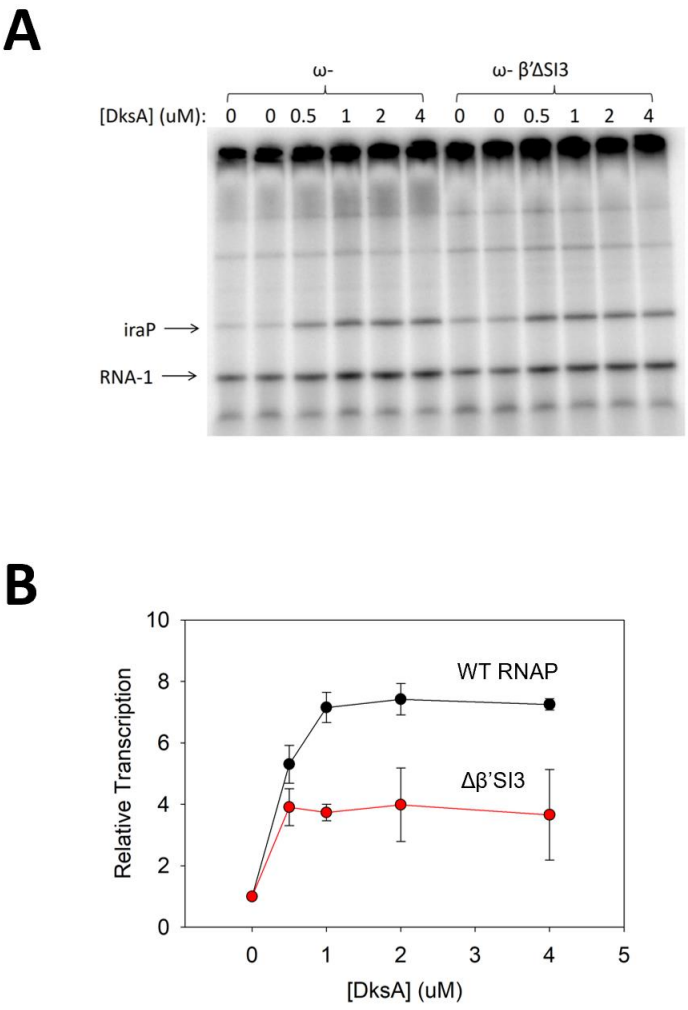

A
ω-
ω- β'ΔSI3
[DksA] (uM): 0 0 0.5 1 2 4 0 0 0.5 1 2 4
iraP
RNA-1
B
Relative Transcription
[DksA] (uM)
WT RNAP
Δβ'SI3
0
2
4
6
8
10
0
1
2
3
4
5

Figure 2.S6. σ1.1 is required for activation by DksA and ppGpp, but not for inhibition. (A) Representative gel showing multi-round *in vitro* transcription from *rrnB P1* at varying concentrations of DksA. **(B)** Quantification of transcription data from (A), normalized to transcription levels in the absence of DksA (n=3). Transcription by WT RNAP is in black, and transcription by holoenzyme with Δσ1.1 is in blue (final figure will only have one blue line). **(C)** Representative gel showing multi-round *in vitro* transcription from P*thrABC,* quantified in **(D)**.

Figure 2.S6

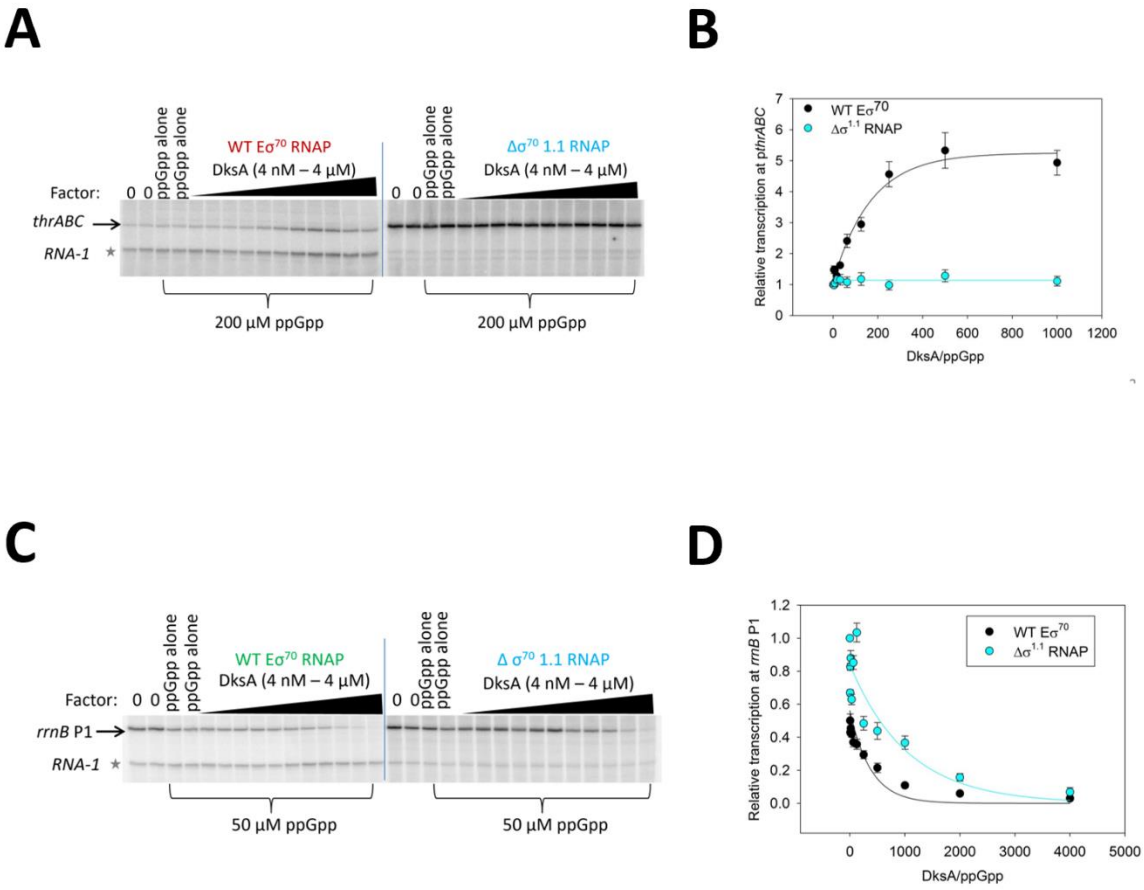

A
B
C
D
WT Eσ70 RNAP
Δσ70 1.1 RNAP
DksA (4 nM – 4 μM)
ppGpp alone
Factor:
thrABC
RNA-1
200 μM ppGpp
Relative transcription at pthrABC
DksA/ppGpp
rrnB P1
50 μM ppGpp
Relative transcription at rrnB P1

Table 2.S1. Strains and Plasmids

Strain	**Description**	**Source**
RLG4677	BL21(DE3)	Novagen
RLG12115	BL21(DE3) *rpoZ::kan*	Ross et al., 2013
RLG7075	BL21(DE3) *dksA::tet*	Paul et al., 2004
Plasmid	**Description**	**Source**
pIA299	pT7 αββ'-H6 overexpression vector	Artsimovitch et al., 2003
pRLG9962	pIA299 *rpoB* ΔSI1-1.2 (ΔE240-L284ΩAAA)	Gopalkrishnan et al., 2017
pIA900	pT7 αββ'-H10ω overexpression vector	Svetlov and Artsimovitch, 2015
pRLG15410	pIA900 *rpoC* F935A	This work
pRLG10030	pIA900 *rpoC* Δ215-220	This work
pJZ9	pT7 αββ'(Δ931-1137)-His6ω overexpression vector	Toulokhonov et al., 2007
pIA331	pT7 αββ'(Δ943-1130)-CBP/intein overexpression vector	Artsimovitch et al., 2003
pIA389	pET28 *rpoD* Δ1.1	Haugen et al., 2006
pRLG13300	pET33 His10 HMK DksA overexpression vector	Ross et al., 2016
pRLG13311	pRLG13300 DksA L86A	This work (Jeong Hyun)
pRLG13318	pRLG13300 DksA R93A	This work (Jeong Hyun)
pRLG15426	pRLG13300 DksA R93A K97A	This work
pRLG10138	pRLG13300 DksA E143A	This work
pRLG10141	pRLG13300 DksA E146A	This work
pRLG770	*In vitro* transcription vector	Ross et al., 1990
pRLG1616	pRLG770 with *rrnB P1* (-88 to +50)	Ross et al., 1990
pRLG5073	pRLG770 with P*thrABC* (-72 to +16)	Barker et al., 2001
pRLG11350	pRLG770 with *iraP* (-207 to +21)	Ross et al., 2016
pRLG9988	pRLG1507 with *iraP* (-64 to +50 with mutations)	This work
pRLG11272	pRLG1507 with *rpsT P2* (-68 to +50)	Chapter 3

Table 2.S2. Oligonucleotides

Primer	Sequence	Purpose
8354	GACCATGCGTACGGCCCACATCGG	*rpoC* F935A
7358	CGAAACCAACTCCGAAACCAAGCGTATCAAAC TGCTGGAAGCGTTCGTTCAG	*rpoC* ΔK215-R220
6345	AGCCTCGAAGCGCGTAACCGCGAT	DksA L86A
6359	GATCGCGAGGCTAAGCTGATCAAA	DksA R93A
6367	AAGCTGATCGCAAAGATCGAGAAG	DksA K97A
7240	CAAAACGCTGGCTGCAATTCGCGAAAAAC	DksA E143A
7243	GAAATTCGCGCAAAACAGATGGC	DksA E146A
7714	CATAAAAATAATACTTCCAGACACGAAGAAGTT GTGAAACCTCATGTTGACTGCTCCATACTTTGA AGCTTGGGTCGACACGCGTAGATC	Remove downstream promoters from iraP

Chapter 3

DksA and TraR allow nucleation of promoter melting but inhibit melting post nucleation

Jared Winkelman performed the DNaseI and $KMnO_4$ footprints in Figures 3.1A,B,C and 3.3. Michael Maloney performed the DMS footprints in Figure 3.S1 and performed other footprints that are not shown. Michael Maloney, Jared Winkelman, and I purified all the Bpa-containing RNAPs. Saumya Gopalkrishnan performed the *in vitro* transcription experiments in Figure 3.5A and Figure 3.S5. James Chen from the Darst lab performed the cryo-EM experiments. I performed all the other experiments (Figures 3.1D, 3.2, 3.4, 3.5B,C, 3.S2, 3.S3, 3.S5, and 3.S6). Sections in the introduction and relevant to Figures 3.1, 3.3, and 3.4 were modified from Chapter 5 of Jared Winkelman's thesis. This work will be submitted as part of a paper with cryo-EM structures from James Chen.

Abstract

DksA and TraR are two transcription factors that directly inhibit transcription from some promoters while directly activating transcription at others by affecting the progression of RNAP through transcription initiation. To gain insight into how DksA and TraR affect RNAP-promoter complexes, we performed footprinting and crosslinking studies on the negatively regulated promoter *rpsT* P2 and solved cryo-EM structures of initiation intermediates at this promoter in the presence of TraR. Our data suggest a model in which DksA together with ppGpp and TraR by itself activate promoters that are limited at the nucleation step and inhibit promoters that are limited at steps after nucleation. Our results not only suggest that DksA and TraR facilitate nucleation of promoter melting and inhibit melting post-nucleation, but they allow structural analyses of transcription intermediates early on the pathway to open complex formation for the first time, thereby providing information about the general mechanism of open complex formation.

Introduction

The first step in promoter recognition is sequence-specific binding of RNAP to double-stranded promoter DNA to form a closed complex (RP_C). The RNAP-promoter complex then proceeds through a series of isomerization steps that result in melting of an ~13 bp region of the promoter and placement of double stranded DNA downstream of the bubble within a channel formed by the β and β′ pincers to form an open complex (RP_O).

Promoter melting by $E\sigma^{70}$ holoenzyme has been proposed to be nucleated by the flipping out and capture of the second nucleotide of the -10 hexamer (TATAAT) within a highly-conserved hydrophobic pocket in σ region 2 (Chen and Helmann, 1997; Felklistov and Darst, 2011; Heyduk et. al., 2006; Lim et. al., 2001; Schroeder et. al., 2009; Tsujikawa et. al., 2002). Melting of the rest of the ~13 nucleotide (nt) bubble requires a 90° bend in order to establish interactions between the highly conserved final thymine (TATAAT) in a second pocket in σ region 2. Bending is also required to place downstream DNA in the cleft formed by the β and β' pincers (Feklistov and Darst, 2011). Although there is evidence that nucleation of melting at -11A precedes melting of the rest of the bubble, this nucleated state has been difficult to detect in kinetic experiments and has not been observed at equilibrium. Therefore, there is little information about the structure of any intermediates that have nucleated, but not completed, promoter melting.

The secondary channel-binding protein DksA affects the rate of transcription from some promoters by affecting the kinetics of open complex formation. The best-characterized function of DksA is its inhibition of transcription from stable RNA promoters (e.g. *rrnB* P1) by inhibiting open complex formation (Paul et al., 2004). Footprinting studies on *rrnB* P1 suggested that DksA and ppGpp together prevent progression from the initial bound closed complex (Rutherford et al., 2009). Intriguingly, DksA (together with ppGpp) is also able to activate transcription from some promoters directly by increasing the rate of open complex formation (Paul et al., 2005). Studies on the DksA-like transcription factor TraR revealed a similar mechanism of regulation (Gopalkrishnan et al., 2017). However, it remains unclear how a factor that inhibits open

complex formation on some promoters can also increase the rate of open complex formation on others.

Here we analyze complexes formed on the DksA-inhibited promoter *rpsT* P2 and show that complexes formed at low temperature (RP_C), complexes formed at high temperature in the presence of DksA (RP_{DksA}), and complexes formed at high temperature in the absence of DksA (RP_{CTP}) have different structural features. The most important observation is that DksA favors a complex that appears to have nucleated melting of the -11 and -10 positions but has not melted DNA downstream of these positions and has not placed the downstream duplex DNA in the cleft. Studies on this intermediate reveal information regarding the mechanism of regulation by secondary channel binding factors as well as the mechanism of transcription initiation in general.

Results

Secondary channel-binding factors allow nucleation of promoter melting

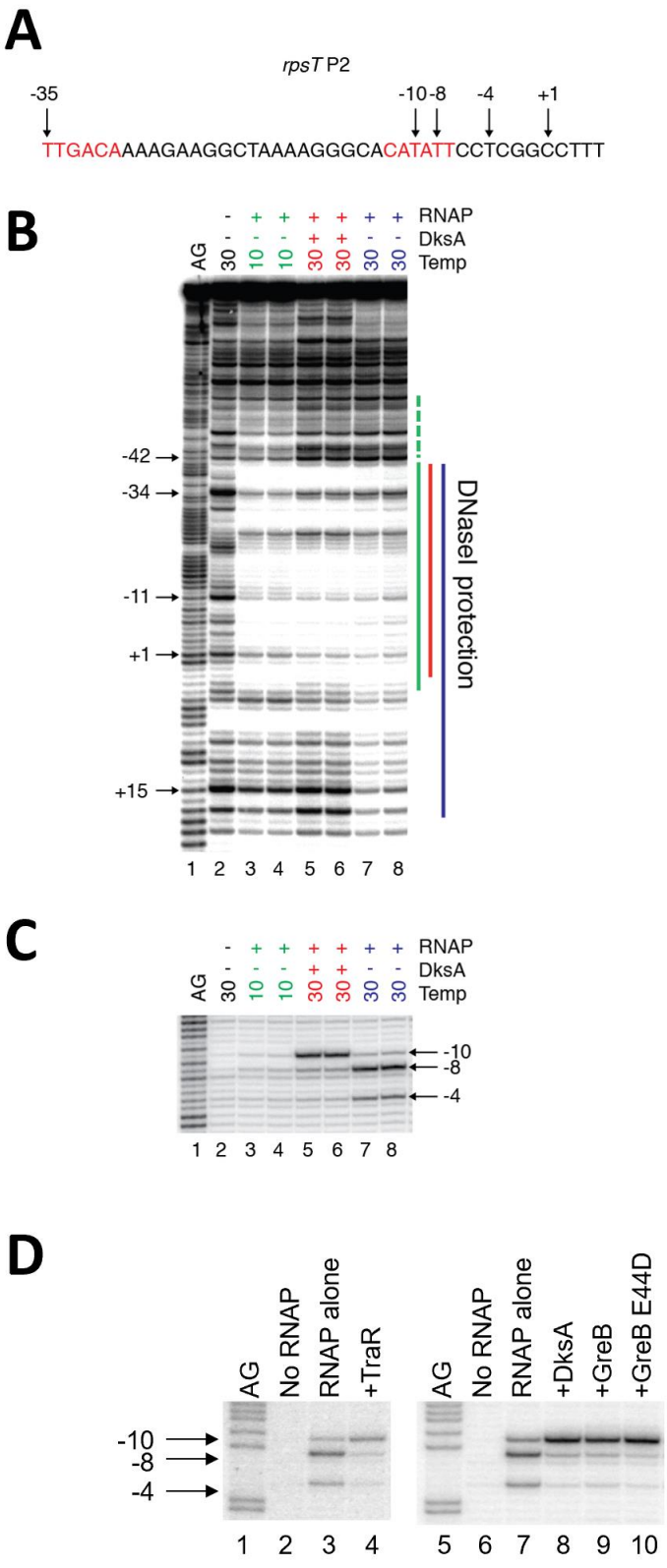
DksA-dependent transcription inhibition has been most thoroughly studied on the model ribosomal RNA promoter, *rrnB* P1. On *rrnB* P1, complexes formed in the presence of DksA exhibit properties in footprints similar to complexes formed at low temperature: the bubble is not $KMnO_4$-reactive and DNA downstream of +1 is not protected from DNaseI cleavage by RNAP (Rutherford et al., 2009). To determine if this behavior is general for DksA-inhibited promoter complexes, we performed footprinting experiments on the DksA and ppGpp-sensitive *rpsT* P2 promoter (Lemke et al., 2011), the promoter for ribosomal protein S20. Complexes formed on *rpsT* P2 at 10°C resembled previously-studied closed complexes: the DNaseI footprint extended from

-41 to +4, and no thymines were reactive to $KMnO_4$ (Figures 3.1B and 3.1C, lanes 3,4). RNAP-*rpsT* P2 complexes formed at 30°C also exhibited footprints characteristic of standard RP_O promoter complexes: RNAP protected DNA from -41 to +20 from DNaseI, and non-template strand thymines at -10, -8, and -4 were reactive to $KMnO_4$ (Figures 3.1B and 3.1C, lanes 7,8). DNaseI footprints in the presence of DksA were similar to those formed at low temperatures, with a downstream boundary of protection at +3. However, the $KMnO_4$ footprint showed an intriguing pattern in which the thymine at -10 was more reactive than in RP_O, whereas the thymines at -8 and -4 were almost completely unreactive to $KMnO_4$. Analysis of $KMnO_4$ reactivity of the template strand showed a similar pattern: the thymines at -11 and -9 were reactive at 30°, whereas in the presence of DksA reactivity of the -9 thymine was reduced while -11 was still reactive (Figure 3.S1A). We also monitored the melted state of cytosines on the non-template strand using dimethyl sulfate. Cytosines at -6, -5, -3, +1, and +2 were all reactive to DMS at 37°C but none of these Cs were reactive in the presence of DksA (Figure 3.S1B,C). Together, these data suggested that RNAP was bound to *rpsT* P2 from -40 to around +1 and had nucleated melting at -11 and -10, but that DNA downstream of -10 was double-stranded and not bent into the cleft.

To determine how other secondary channel-binding factors affect RNAP at *rpsT* P2, we performed $KMnO_4$ footprinting at *rpsT* P2 in the presence of TraR or GreB. TraR is another secondary channel binding factor that regulates transcription similarly to DksA and ppGpp (Gopalkrishnan et al., 2017). At *rpsT* P2, TraR has been shown to inhibit transcription and shorten the downstream protection of the DNaseI footprint to +3

Figure 3.1. The transcription bubble is partially melted at *rpsT* P2 in the presence of DksA, TraR, or GreB. (A) *rpsT* P2 promoter sequence. -35 hexamer and -10 hexamer sequence are in red. DNaseI **(B)** or $KMnO_4$ **(C)** footprints of RNAP-*rpsT* P2 complexes formed at low temperature (lanes 3,4), at 30°C in the presence of DksA (lanes 5,6), or at 30° in the absence of DksA (lanes 7,8). Lane 1, A+G sequence ladder. Lane 2, no RNAP control. The DNA fragment was 3' end-radiolabeled on the non-template strand. **(D)** $KMnO_4$ footprints of RNAP-*rpsT* P2 complexes in the presence of TraR (lane 4), DksA (lane 8), WT GreB (lane 9) or E44D GreB (lane 10). Lane 2 and 6, no RNAP control. Lane 1 and 5, A+G sequence ladder.

Figure 3.1


A
rpsT P2
-35
-10
-8
-4
+1
TTGACAAAAGAAGGCTAAAAGGGCACATATTCCTCGGCCTTT
B
RNAP
DksA
Temp
AG
30 - 10 - 10 - 30 + 30 + 30 - 30 -
-42
-34
-11
+1
+15
DNaseI protection
1 2 3 4 5 6 7 8
C
RNAP
DksA
Temp
AG
30 - 10 - 10 - 30 + 30 + 30 - 30 -
-10
-8
-4
1 2 3 4 5 6 7 8
D
AG
No RNAP
RNAP alone
+TraR
AG
No RNAP
RNAP alone
+DksA
+GreB
+GreB E44D
-10
-8
-4
1 2 3 4 5 6 7 8 9 10

(Gopalkrishnan et al., 2017). We found that like with DksA, the -10 thymine on the non-template strand is hyper-reactive in the presence of TraR while downstream thymines -8 and -4 are not reactive (Figure 3.1D).

GreB is another secondary channel binding factor that also destabilizes open complexes (Rutherford et al., 2007; Lee et al., 2012). Unlike DksA and TraR, GreB does not activate transcription, but instead functions mainly during transcription elongation to rescue stalled or paused complexes (Laptenko et al., 2003; Furman et al., 2013; Tetone et al., 2017). At *rpsT* P2, GreB also caused RNAP to form a partially melted intermediate where the -10 thymine is hyperreactive (Figure 3.1D, lane 9). An E44D substitution in GreB, which allows GreB to inhibit and activate transcription like DksA (Lee et al., 2012), was also able to favor the same intermediate (Figure 3.1D, lane 10). Together, these data indicate that secondary channel binding factors that can destabilize the open complex can also form the partially melted complex at *rpsT* P2 (see also Discussion).

ΔdksA suppressor mutants cannot mimic the DksA inhibited complex on *rpsT* P2 in the absence of DksA

To determine whether destabilization of *rpsT* P2 was sufficient to form the partially melted intermediate, we tested RNAP variants that form open complexes with intrinsically short lifetimes in the absence of secondary channel binding factors. DksA is required for growth of *E. coli* on media lacking amino acids, and suppressor mutations that rescue growth were mapped to *rpoB* and *rpoC* (Rutherford et al., 2009). Of the 28 unique mutations examined, 18 coded for substitutions in or near DNA binding residues

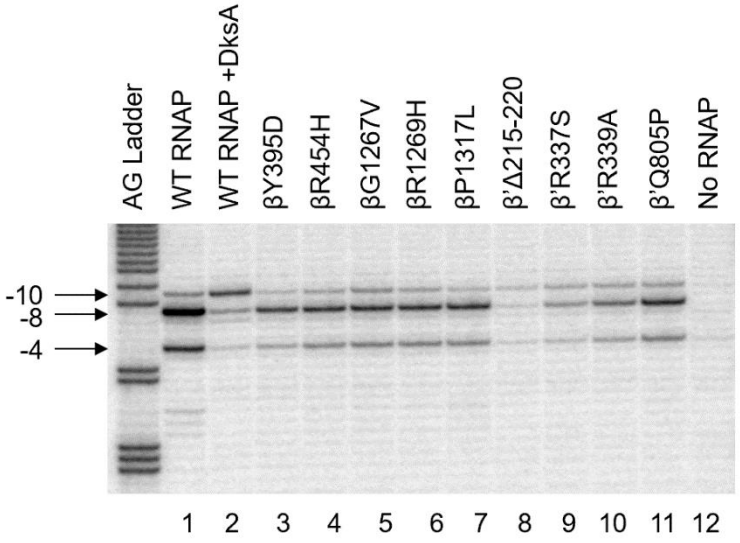
while 21 coded for substitutions in or around the switch regions (Rutherford et al., 2009). The switch regions control the RNAP clamp, which is needed for open complex formation and for stabilizing DNA in the cleft (Feklistov et al., 2017; Chakraborty et al., 2012; see also Chapter 2). By directly or allosterically affecting DNA contacts, these suppressor mutations encoded changes that mimicked the effect of DksA and ppGpp on transcription *in vivo*. Seven representative RNAP variants were purified and tested *in vitro* and all resulted in reduced open complex lifetimes and decreased transcription from *rrnB* P1, similar to the effects of DksA and ppGpp (Rutherford et al., 2009).

To determine whether these RNAP variants were also able to mimic the DksA/ppGpp inhibited complex containing *rpsT* P2, we purified ten variants and performed $KMnO_4$ footprinting to detect promoter melting. Nine out of ten variants formed open complexes, where the non-template -10, -8, and -4 thymines were all reactive (Figure 3.2). While these RNAP variants formed shorter-lived open complexes than WT RNAP, RPo was nevertheless able to form with most of the mutant RNAPs, and $KMnO_4$ was able to detect a fully melted bubble in a fraction of the promoter population. However, one of the variants, β'Δ215-220, formed a closed complex in which little reactivity was observed at any position (Figure 3.2). Notably, no variants resulted in formation of the partially-melted intermediate observed with this promoter and DksA. These data suggest that simply destabilizing an open complex is not sufficient to observe a partially melted intermediate at this promoter, and that DksA alters RNAP in a specific way to capture this intermediate.

Additionally, it was shown in Chapter 2 that while Δβ'215-220 forms a closed complex, DksA-promoted clamp closure or clamp closure promoted by the antibiotic

Figure 3.2. *ΔdksA* suppressor substitutions in RNAP do not form the partially melted intermediate in the absence of DksA. $KMnO_4$ footprints of RNAP variants on WT *rpsT* P2 promoter DNA. The DNA fragment was 3' end-radiolabeled on the non-template strand. Reactions were performed at 30°C.

Figure 3.2


AG Ladder
WT RNAP
WT RNAP +DksA
βY395D
βR454H
βG1267V
βR1269H
βP1317L
β'Δ215-220
β'R337S
β'R339A
β'Q805P
No RNAP
-10
-8
-4
1 2 3 4 5 6 7 8 9 10 11 12

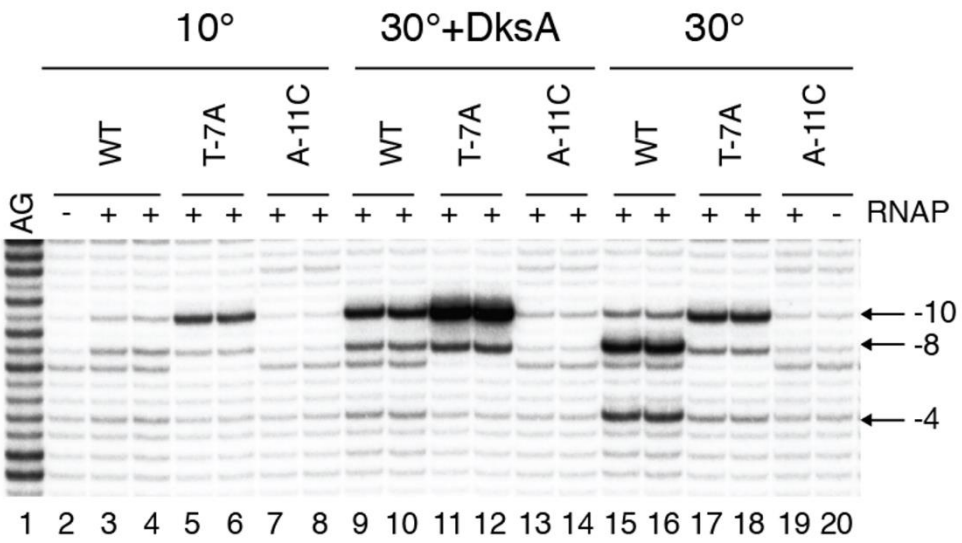
myxopyronin B (MyxB) were sufficient to induce partial transcription bubble melting (Figure 2.1C). And while *ΔdksA* suppressors were able to mimic effects of DksA in that they produced shorter-lived open complexes, they were unable to promote isomerization and strand opening. These data support the model that RP_{DksA} is an intermediate subsequent to formation of the initial closed complex (RP_C) that has not yet progressed to a fully formed RP_O.

Complexes formed with *rpsT* P2 -11A mutated promoters mimic the complexes formed at low temperature whereas complexes formed with -7T mutated promoters mimic the complexes inhibited by DksA

Recognition of the nontemplate strand of the -10 hexamer is critical to formation of RP_O. The most highly conserved and functionally important bases are -11A (T<u>A</u>TAAT) and -7T (TATAA<u>T</u>), which are flipped out and buried in pockets in σ region 2 (Feklistov and Darst, 2011). Therefore, mutation of -11A or -7T would be predicted to impair open complex formation. As shown in Figure 3.1, footprints of RNAP in the presence of DksA at *rpsT* P2 (RP_{DksA}) suggested that melting had been nucleated but that DNA downstream of -10 was still double-stranded. Thus -11A could potentially be recognized and be important for formation of RP_{DksA}, but -7T would not be recognized. Footprinting data for RP_C suggested that -11A and -7T were not melted and therefore were not significant for formation of RP_C (Figure 3.1). To determine the role of these bases in formation of RP_C, RP_{DksA}, and RP_O, we changed these positions individually and performed footprinting experiments at 10°, at 30°C in the presence of DksA, and at 30°C in the absence of DksA (Figure 3.3).

Figure 3.3. Promoter mutations mimic closed and DksA-inhibited complexes. $KMnO_4$ footprints were performed using nontemplate strand-labeled *rpsT* P2, T-7A *rpsT* P2, or A-11C *rpsT* P2 promoter DNA. A-11C promoter DNA is not reactive to $KMnO_4$, T-7A promoter DNA exhibits a pattern of reactivity similar to that seen in the presence of DksA, with -10 more reactive than -8 or -4.

Figure 3.3


10°
30°+DksA
30°
WT
T-7A
A-11C
WT
T-7A
A-11C
WT
T-7A
A-11C
AG
- + + + + + + + + + + + + + + + + + + -
RNAP
-10
-8
-4
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

Consistent with the proposed role of -11A in nucleation of melting, we found that the A-11C mutant *rpsT* P2 promoter was not $KMnO_4$-reactive under any of the conditions tested (Figure 3.3 lanes 7,8,13,14,19, 20). However, this promoter was able to form a specific promoter complex, so -11A recognition is not required for closed complex formation. We conclude that the A-11C *rpsT* P2 mutant promoter mimics the low temperature complex and can be used as a model for RP_C at higher temperatures.

Interestingly, we found that T-7A *rpsT* P2 mimicked RP_{DksA} in the absence of DksA. The -10T was hyper-reactive to $KMnO_4$, whereas reactivity of -8T and -4T was diminished at 30°C in the presence or absence of DksA (Figure 3.3 lanes 11,12, 17,18). These data support the hypothesis that RP_{DksA} represents a complex in which nucleation of melting at -11A has happened, but melting has not continued downstream to include -7T. We also note that combination of the T-7A mutation and DksA led to enrichment of a complex in which the -10 position was single-stranded, but positions downstream of -10 were not melted (compare lanes 9,10 to 11,12). Therefore, in the next sections we use the T-7A *rpsT* P2 promoter in addition to DksA to enrich for RP_{DksA}.

The interface between RNAP and promoter DNA is different in RP_C, RP_{DksA}, and RP_{CTP}

To directly monitor how the path of DNA differs in RP_C, RP_{DksA}, and RP_O we monitored the proximity of specific amino acids in RNAP to specific nucleotides in *rpsT* P2 using crosslinking with *p*-benzoyl-L-phenylalanine (Bpa). Bpa is a non-natural amino acid that can be incorporated into a polypeptide, and upon UV exposure, it forms

crosslinks with a carbon atom located 2-3 Å away (Ryu and Schultz, 2006; Winkelman et al., 2015). We utilized our library of Bpa-containing RNAPs to identify surfaces of RNAP that crosslinked to *rpsT* P2 in different complexes and mapped the location of each crosslink at single-nucleotide resolution on DNA using primer extension. For RP_O, promoters like *rrnB* P1 and *rpsT* P2 are intrinsically unstable (Rutherford et al., 2009), so we enriched for an open bubble by adding the initiating nucleotide CTP. At this promoter, addition of CTP likely produced a two-mer product, but we reasoned that formation of one phosphodiester bond should not drastically affect the position of the crosslinks to the complex. The complexes examined were A-11C *rpsT* P2 (to form RP_C), T-7A *rpsT* P2 in the presence of N88I DksA (to form RP_{DksA}), and wild-type *rpsT* P2 in the presence of CTP (to form RP_{CTP}). All reactions were performed at 37°.

To ensure that the Bpa-containing RNAPs formed RP_C, RP_{DksA}, and RP_{CTP}, we performed control footprinting experiments. Only Bpa-containing RNAPs that protected DNA from -40 to +4 on the A-11C promoter were used to analyze RP_C (Figure 3.S2). For RP_{DksA}, many mutations in DNA-binding residues eliminated the capture of the partially melted intermediate, instead resulting in reactivity of all three positions (-10, -8, and -4) or resulting in no reactivity at all (Figure 3.S3). Use of the high-affinity variant N88I DksA on T-7A *rpsT* P2 enriched for the intermediate of interest, however, which allowed capture of the partially melted intermediate in which the -10 base was more reactive than the -8 base. As stated above, RP_{CTP} was formed on WT *rpsT* P2 in the presence of CTP to enrich for an open bubble (reactive -10T, -8T, and -4T).

We identified several different patterns of crosslinking, summarized in Table 3.1 and Figure 3.4 (Figure 3.S4). First, we identified two positions, β'T48Bpa and βS63Bpa,

Figure 3.4. RP_C, RP_{DksA}, and RP_{CTP} have distinct crosslinking patterns. Schematic summarizing Bpa crosslinking to *rpsT* P2 in RP_C, RP_{DksA}, or RP_{CTP}. Line indicates crosslink between the Bpa substitution and the corresponding DNA base. Residues without a line indicate that crosslinking was performed but no crosslinks were detected. -10 hexamer and the +1 start site are highlighted in red. Residues that do not change crosslinking between RP_{DksA} and RP_{CTP} are shown in orange. Other residues are shown in blue. See Figure 3.S4 for gels and Table 3.1 for summary of data.

Figure 3.4

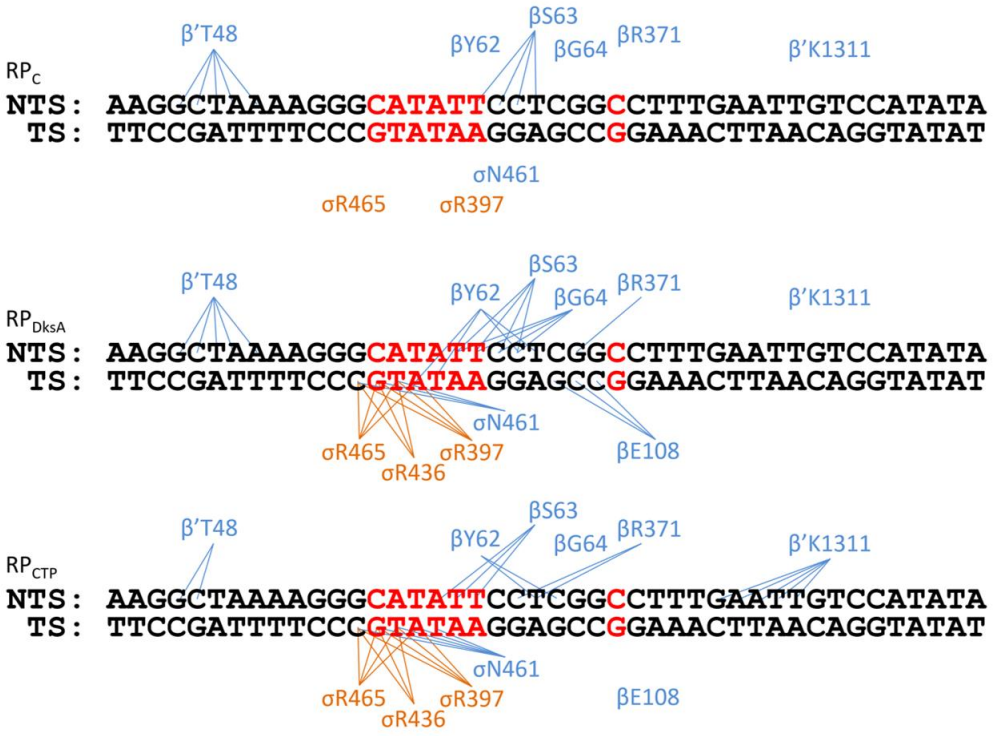

RP$_C$
β'T48
βS63
βY62
βG64
βR371
β'K1311
NTS: AAGGCTAAAAGGGCATATTCCTCGGCCTTTGAATTGTCCATATA
TS: TTCCGATTTTCCCGTATAAGGAGCCGGAAACTTAACAGGTATAT
σN461
σR465
σR397
RP$_{DksA}$
β'T48
βS63
βY62
βG64
βR371
β'K1311
NTS: AAGGCTAAAAGGGCATATTCCTCGGCCTTTGAATTGTCCATATA
TS: TTCCGATTTTCCCGTATAAGGAGCCGGAAACTTAACAGGTATAT
σN461
σR465
σR397
σR436
βE108
RP$_{CTP}$
β'T48
βS63
βY62
βG64
βR371
β'K1311
NTS: AAGGCTAAAAGGGCATATTCCTCGGCCTTTGAATTGTCCATATA
TS: TTCCGATTTTCCCGTATAAGGAGCCGGAAACTTAACAGGTATAT
σN461
σR465
σR397
σR436
βE108

Table 3.1. Summary of *rpsT* P2 crosslinking. The position of Bpa incorporation is in the left column. "--" means that crosslinks were not detected. "N/A" means Bpa variant does not form the proper footprint under these conditions. Bolded numbers refer to strong crosslinks within a group of crosslinks. See Figure 3.S4.

RNAP	RP_C	RP_{DksA}	RP_{CTP}
βY62	--	TS -10, -9// NTS -5, -4	NTS -4, -3
βS63	(weak) NTS -7, -6, -5, -4	NTS -8, -7, -6, -5	NTS -9, -8, **-7**
βG64	--	NTS -8, -7, -6, -5	--
βE108	N/A	TS -3, -2, -1	--
βR371	--	NTS -2	NTS -5, -4
β'T48	TS -22, -21, -20, -19, -18	TS -22, -21, -20, -19, -18	TS -22, -21
β'K1311	--	--	NTS +5, +6, +7, +8, +9
σR397	--	TS -13, **-12**, -11, -10	TS -13, -12, **-11**, -10
σR436	N/A	TS -13, -12, -11	TS -13, -12, -11
σN461	--	TS -13, -12, -11	TS -13, -12, **-11**, -10, -9
σR465	--	TS -13, -12, **-11**, -10	TS -13, -12, -**11**, -10

that crosslinked in RP_C and in RP_{DksA}, suggesting that these residues are in the proximity of promoter DNA in both states. The crosslinking and footprinting signals were generally weaker in these complexes, consistent with the less stable binding of RNAP to DNA in RP_C. Additionally, the crosslinking pattern changed for the βS63Bpa complex between RP_C and RP_{DksA}, indicating that there were changes in the RNAP-DNA interactions upon isomerization.

We also identified positions that crosslinked similarly in RP_{DksA} and RP_{CTP}, but not in RP_C. Several Bpa-containing RNAPs that crosslinked to the upstream part of the bubble (template strand -13 to -9) showed this pattern (σR397Bpa, σR436Bpa, σE458Bpa, σR465Bpa) (orange residues in Figure 3.4). This is the region where melting is nucleated, suggesting that upon nucleation, DNA nucleotides from -13 to -9 are positioned differently than in RP_C, and that once these contacts have been established in RP_{DksA}, they are maintained through RP_{CTP}.

We also identified positions that crosslinked to DNA in both RP_{DksA} and in RP_{CTP} but to different promoter positions (βY62Bpa, βS63Bpa, βR371Bpa, β'T48Bpa, σN461Bpa). A representative example is βR371Bpa, which crosslinked to non-template strand nucleotide -2 in RP_{DksA} but to -5 and -4 in RP_{CTP}. Most of these residues lie downstream of the -10 element, consistent with a dramatic rearrangement that occurs as the DNA from -9 to +2 goes from double-stranded in RP_{DksA} to single-stranded in RP_{CTP}. There were also slight changes in the crosslink pattern of the complex containing β'T48Bpa, which crosslinks to the -10/-35 spacer region of the promoter, suggesting that there are changes in the interactions with sigma upon open complex formation.

βG64Bpa and βE108Bpa are unique in that they crosslinked to DNA in RP_{DksA} but not in RP_{CTP}. These residues are positioned near where DNA bends into the cleft. Therefore, when DNA is not bent into the cleft, as in RP_{DksA}, these residues are in close proximity with DNA, but when DNA bends into the cleft, DNA moves further away from these residues, so crosslinks are not formed.

Finally, crosslinking from the cleft by certain Bpa-containing RNAP variants only occurred in the open complex, an example of which was β'K1311Bpa. Since there is no protection of promoter DNA by RNAP downstream of +4 in either RP_C or RP_{DksA}, the crosslinking pattern was consistent with the positions of Bpa in these complexes.

Cryo-EM reveals a pocket for the template strand -9T

We collaborated with the Darst lab to solve the cryo-EM structure of the partially melted complex formed at *rpsT* P2. Since TraR traps the same intermediate as DksA and ppGpp together, TraR was used in these cryo-EM experiments to reduce the complexity of having populations containing zero, one, or two ppGpp-bound complexes. Two sets of cryo-EM data were obtained. The first set contained WT RNAP in complex with TraR at the WT *rpsT* P2 promoter. This data set produced multiple transcription initiation intermediates, including RP_C, multiple partially melted intermediates, and RP_O. The intermediate most similar to RP_{DksA} showed flipping of the -11A base out of the double helix, but DNA downstream of -11 was unresolved, so the state of downstream DNA is unclear.

The second set of cryo-EM data contained WT RNAP in complex with TraR on *rpsT* P2 promoter DNA engineered to have a mismatch at -11 and -10 to stabilize the

partially melted state. An extra mismatch was also introduced at the -7 position during template generation. This cryo-EM dataset resulted in a partially melted transcription bubble, single-stranded between -11 and -7. Downstream DNA was outside the cleft and continued as a double helix past the β lobe, consistent with a DNaseI footprinting pattern that showed protection to ~+3. The position of DNA relative to the RNAP was consistent with the crosslinking pattern obtained with DksA on the T-7A *rpsT* P2 promoter.

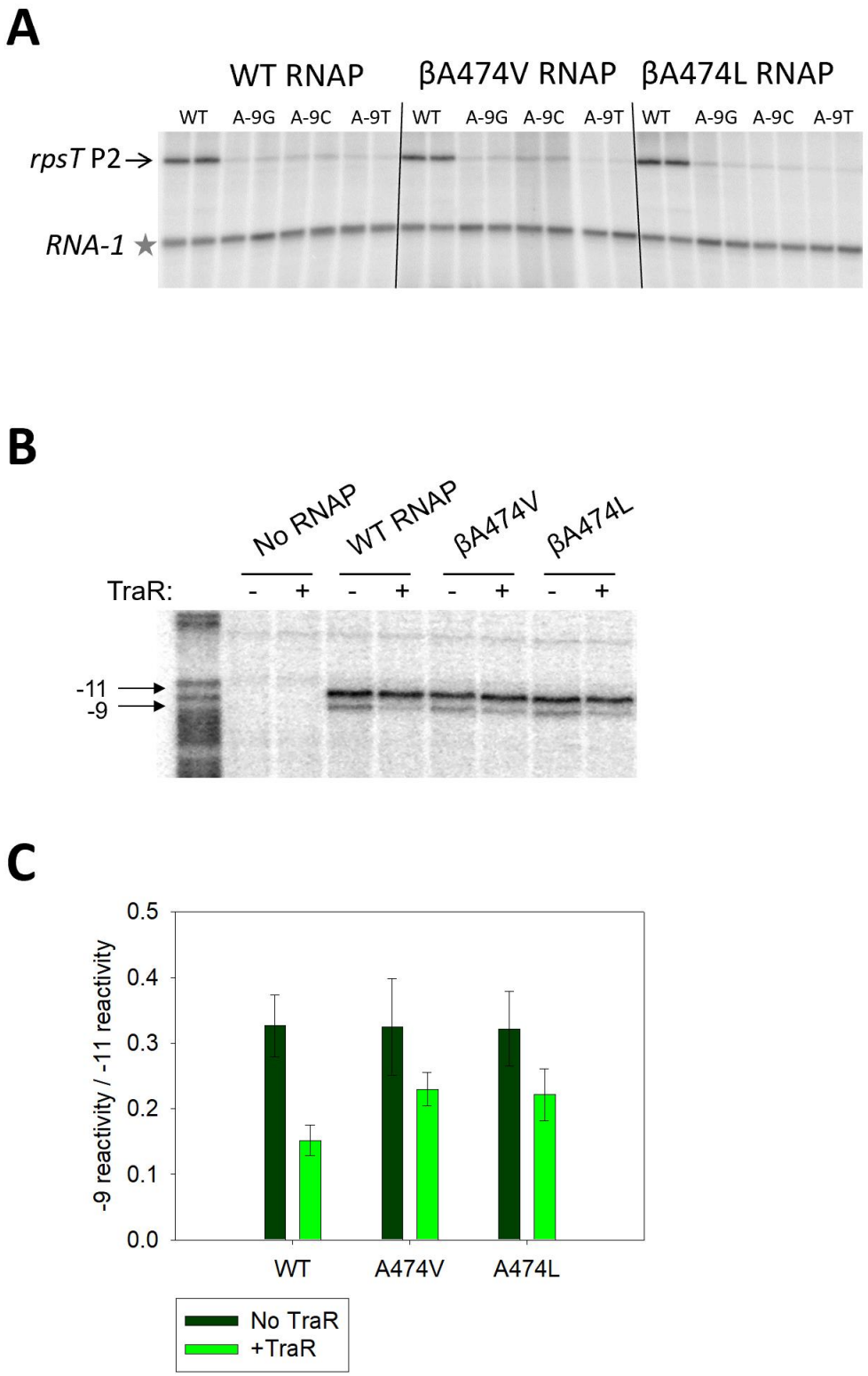
The cryo-EM structure of the partially melted intermediate also showed that the template strand -9T bound in a pocket in the β protrusion. Binding of DNA in this pocket has not been observed previously. -9T was not bound in this pocket in RP_O, suggesting that interaction of -9T with this pocket may occur during initiation intermediates to aid in open complex formation.

The pocket appears large enough to fit a thymine base but too small to fit larger bases. The consensus base at this position is a thymine, and *rpsT* P2 has the thymine at this position. To test the functional importance of the -9T in *rpsT* P2, we mutated this position and used *in vitro* transcription to measure its effect on output and on regulation by TraR. Mutation of this basepair to any other basepair resulted in a dramatic decrease in transcription, consistent with previous observations on certain other promoters (Figure 3.5A; Xu et al., 2001). However, substitutions in this position had no effect on regulation by TraR, suggesting it is important for transcription initiation in general, but not for regulation by TraR (Figure 3.S5A).

Looking at the conservation of the -9T pocket, we found that βA474 is present in almost 50% of bacterial species. Another 25% have a valine at this position, which

Figure 3.5. -9 position in the -10 hexamer is important for *rpsT* P2 transcription. **(A)** Multi-round *in vitro* transcription showing effect of A-9G, A-9C, and A-9T mutations in *rpsT* P2 on transcription output. Reactions carried out at room temperature with WT RNAP or with variants βA474A or βA474L. **(B)** $KMnO_4$ footprints with WT RNAP, βA474A, or βA474L on template strand-labeled *rpsT* P2. Reactions are performed in the presence or absence of TraR. Relative reactivity of -9 to -11 bands plotted in **(C)**.

Figure 3.5


A
WT RNAP
βA474V RNAP
βA474L RNAP
WT A-9G A-9C A-9T WT A-9G A-9C A-9T WT A-9G A-9C A-9T
rpsT P2
RNA-1
B
No RNAP
WT RNAP
βA474V
βA474L
TraR: - + - + - + - +
-11
-9
C
-9 reactivity / -11 reactivity
0.5
0.4
0.3
0.2
0.1
0.0
WT
A474V
A474L
No TraR
+TraR

would be expected to increase the fit of -9T in the pocket. Leucine does not appear to occur at this position, and modeling suggests that binding of -9T is incompatible with βA474L. Thus, to test the role of the -9 pocket further, we made a βA474V or βA474L substitution and determined its effect on transcription. Surprisingly, there was little effect of these substitutions on overall transcription output and only a minor defect in regulation by TraR (Figures 3.5A, 3.S5B).

We used $KMnO_4$ footprinting on the template strand of *rpsT* P2 to determine if the βA474V and βA474L substitutions had effects on formation of the partially melted intermediate (Figures 3.5B, 3.5C, 3.S6). The -11T on the T strand was reactive in all of the footprints. However, -9T on the T strand was reactive in RP_O, but became unreactive in the presence of TraR, likely because the -9T was bound in the pocket and inaccessible to $KMnO_4$. Both the βA474V and βA474L substituted RNAPs resulted in a slight increase in the reactivity at -9 relative to the reactivity observed at -11, suggesting that these mutations both decreased binding of -9T in the pocket. However, the change was subtle. More mutations in the pocket may be necessary to eliminate binding of -9T in the proposed pocket.

Discussion

Footprints of *rpsT* P2 at 10°C, at 30°C in the presence of DksA, and at 30°C in the absence of DksA suggested that there are three different complexes that can be distinguished from each other under these conditions. $KMnO_4$ footprinting indicated that DNA in the 10°C complex is completely double-stranded, that DNA melting has been nucleated at 30°C in the complex containing DksA, and that the bubble has been melted

fully in the complex formed at 30°C in the absence of DksA. Analysis of the footprints formed on promoters with the A-11C or T-7A mutations indicated that -11A is important for formation of RP_{DksA} and RP_O, but -11A is not required for formation of RP_C at 10°C or 30°C. The T at -7 was required for formation of a fully-melted bubble, but the T-7A promoter could form RP_{DksA}. Our crosslinking data and cryo-EM structures also showed that three different complexes with different paths of DNA can be distinguished using the three solution conditions.

The footprints and structure of the partially melted intermediate showed that transcription bubble melting can be initiated without DNA being bent into the cleft. However, bubble melting does not progress very far. Capture of the -11A base in σ may provide sufficient flexibility for downstream DNA to bend 90° into the cleft and allow further melting to occur. In the cleft, further contacts with RNAP would stabilize single stranded DNA (Zhang et al., 2012). One of these interactions may be with the -9 pocket in the β protrusion. The cryo-EM structure revealed an interaction between the template strand -9T and the protrusion pocket, although substitutions in this pocket did not appear to have significant effects. It is possible that a single substitution in the pocket is not sufficient to abolish the interaction of -9T with the pocket.

If the protrusion pocket is not physiologically relevant, what is the basis for the base specificity at position -9? Mutations in this residue result in significant defects in transcription, although these experiments did not distinguish between effects of the -9T on the template strand vs. the -9A on the non-template strand. While RNAP makes multiple contacts to the phosphate backbone of the non-template -9 position, few base-specific interactions have been observed to the non-template -9 base (Zhang et al.,

2012; Feklistov et al., 2011). However, mutations of the non-template -9A away from consensus impairs the ability of single-stranded promoter DNA to bind to RNAP holoenzyme, suggesting that RNAP can make base-specific contacts to the non-template -9A (Qiu et al., 1999). It is also possible that both the template strand and non-template strand bases play a role in transcription initiation, but a more thorough investigation of the -9 pocket is needed to determine its role in transcription.

DksA and ppGpp together directly inhibit transcription from some promoters, while they activate transcription from other promoters. Our observation that DksA allows nucleation of melting but inhibits melting downstream of the point of nucleation suggests a potential explanation for these contrasting regulatory effects. We propose that DksA facilitates, rather than just allows, nucleation of melting, and it inhibits steps after nucleation. According to this model, promoters that are rate-limited at steps after nucleation of melting, such as the *rpsT* P2 and ribosomal RNA promoters, would be inhibited by DksA and ppGpp, while promoters that are rate-limited at the nucleation step would be activated (see also Chapter 2). This model is consistent with data indicating that activated promoters form RP_O slowly but the complex formed is stable. This slow rate of open complex formation could result from a slow rate of nucleation. In contrast, promoters that are inhibited form transcriptionally active complexes quickly but do not form stable open complexes.

Interestingly, GreB is also able to induce formation of the partially melted intermediate even though it cannot activate transcription. One explanation could be that formation of this partially melted intermediate is necessary but not sufficient for activation. While both DksA and GreB bind in the secondary channel, their interactions

with RNAP differ. β'SI3 is required for GreB to bind to RNAP, but β'SI3 inhibits binding of DksA (Furman et al., 2013; Parshin et al., 2015). Both DksA and GreB influence the conformation of the trigger loop, but GreB inhibits partial folding of the trigger loop while DksA permits certain other conformations (Nayak et al., 2013; Chapter 2). Finally, DksA interacts with βSI1 while GreB has not been shown to do so (Parshin et al., 2015; Chapter 2). The presence of GreB in the secondary channel may be sufficient to favor clamp closure and strand opening, but it may be unable to make a necessary interaction with βSI1 to achieve transcription activation. DksA and ppGpp, on the other hand, are able to interact with RNAP in a way that allows for transcription activation. Further studies will be needed to reveal whether this partially melted intermediate occurs at all promoters and how it plays into the mechanism of regulation by DksA and ppGpp.

Materials and Methods*

Modified from J. Winkelman's thesis, Chapter 5.

Plasmid Mutagenesis

RNAP and promoter variants were generated by site-directed mutagenesis using the QuikChange Lightning Multi Site-Directed Mutagenesis Kit (Stratagene). Transformants were streaked and restreaked for single colonies, and the overexpression plasmids were purified and sequenced to verify the identity of the mutation.

Preparation of DNA templates for crosslinking

DNA templates for crosslinking experiments were prepared by PCR amplification using Taq DNA polymerase (New England Biolabs) and primers 1620 and 6625 on p11272 for WT *rpsT* P2, p12829 for A-11C *rpsT* P2, or p12844 for T-7A *rpsT* P2. PCR was performed with an annealing temperature of 53°C and run for 26 cycles. PCR products were purified using the Qiagen PCR Purification kit.

Plasmids for overexpression of WT or Bpa-containing RNAPs

Plasmids used to overexpress WT or Bpa containing core RNAP were derived from pIA900 (encoding α,β,β',and ω with His_{10} on the C-terminus of β' (Svetlov and Artsimovitch, 2015). Plasmids used to overexpress WT or Bpa-containing σ^{70} proteins were derived from pRLG13105 (encoding σ^{70} with and His_{10} at the N-terminus and a PreScission protease cleavage site between the His tag and σ^{70} sequence; Winkelman et al., 2015).

Stop codons (TAG) were introduced into *rpoB*, *rpoC*, or *rpoD* at the position chosen for Bpa incorporation. Mutagenesis was performed using the QuikChange Lightning Multi Site-Directed Mutagenesis Kit (Agilent). Transformants were streaked and restreaked for single colonies, and the overexpression plasmids were purified and sequenced to verify the identity of the mutation.

Incorporation of Bpa and protein expression

Protein expression was performed as described in Winkelman et al., 2015. For protein purification, cotransformation of the overexpression plasmid and the tRNA/tRNA

synthetase plasmid into the host strain (BL21(DE3) for pIA900; DH10B for pRLG13105) was performed fresh for each experiment. Plasmids were cotransformed by electroporation, selecting for both ampicillin and chloramphenicol resistance. Fresh transformants were scraped from plates for use as an inoculum, generating a relatively high starting culture density (OD_{600} ~0.3), and grown at 30°(for σ^{70} overexpression) or 37°C (for core RNAP overexpression) in LB with Bpa (1 mM), ampicillin (100 μg/ml), and chloramphenicol (25 μg/ml). Using a large inoculum from plates avoided suppressor accumulation from extended growth in liquid culture. The culture medium was prepared by addition of Bpa to LB medium dropwise from a freshly made 100 mM Bpa stock in 1 M NaOH. 1 M HCl was added to produce a final pH of ~7.2. After 1 hr of growth with Bpa, expression was induced with 1 mM IPTG or 0.2% L-arabinose, depending on the expression system. Core RNAP overexpressing cells were grown in the dark at 37°C for 6-20 hr after induction, and σ^{70} -overexpressing cells were grown in the dark at 30° for 1-1.5 hr after induction. Cells were harvested by centrifugation and stored at -20° for up to one week. For non-Bpa containing proteins, cells were grown in the absence of Bpa and without the pSupT/BpF plasmid.

Purification of overexpressed core RNAP and σ^{70}

Core RNAPs with a His_{10} tag at the C-terminus of the β′ subunit were purified using Ni-agarose and heparin affinity chromatography sequentially, as described in Winkelman et al., 2015. Cell pellets harvested from 250 ml cultures were suspended in 5 ml of BugBuster (Novagen), phenylmethylsulfonyl fluoride (PMSF) to a final concentration of 23 μg/ml, and 5 μl Lysonase (Novagen). Resuspended pellets were

incubated at room temperature with gentle rocking for 30 min before adding 15 ml of RNAP resuspension buffer (1.4 M NaCl, 40 mM Tris-Cl pH 8.0, 30 mM imidazole), followed by centrifugation at 14,000 rpm for 40 min at 4°C. The cleared lysate was then added to 0.5 ml pre-equilibrated Ni resin, the column was washed with RNAP wash buffer (300 mM NaCl, 40 mM Tris-Cl pH 8.0, 30 mM imidazole), and the protein was eluted with wash buffer containing 300 mM imidazole. The eluate was diluted to 200 mM NaCl with TGED (10 mM Tris-Cl pH 8.0, 5% glycerol, 0.1 mM EDTA, 0.1 mM DTT; (Burgess and Jendrisak, 1975) and 100 mM NaCl, loaded onto 0.4 ml of heparin resin column that had been pre-equilibrated in TGED plus 200 mM NaCl. The column was washed with 5 ml of TGED plus 200 mM NaCl, and RNAP was eluted with 1 ml of TGED plus 600 mM NaCl. RNAPs were concentrated using 5 ml Microcon centrifugal filtration units with a 100 kDa molecular weight cutoff, and exchanged with 2X storage buffer without glycerol (20 mM Tris-Cl at pH 7.9, 200 mM NaCl, 0.2 mM DTT, 0.2 mM EDTA). The final volume was measured and an equal volume of 100% glycerol was added. Proteins were stored at −20°C. Protein concentrations were measured using the Bradford assay reagent (Bio-Rad) using bovine serum albumen (BSA) as a standard.

σ^{70} RNAPs with a His_{10} tag at the N-terminus were purified using Ni-agarose. Cell pellets harvested from 250 ml cultures were suspended in 15 ml buffer A (40 mM Tris-Cl pH 7.9, 200 mM NaCl, 10 mM imidazole), 1X HALT protease inhibitor (Pierce), and PMSF to a final concentration of 23 μg/ml. Resuspended pellets were lysed by sonication before being centrifugation at 14,000 rpm for 20 min at 4°C. The cleared lysate was then added to 0.5 ml Ni resin pre-equilibrated in buffer A, the column was washed sequentially with 10 ml buffer A, 5 ml of buffer B (40 mM Tris-Cl pH 7.9, 1 M

NaCl) with 10 mM imidazole, 5 ml buffer B with 20 mM imidazole, and 5 ml buffer B with 50 mM imidazole, before elution with 1.5 ml buffer B with 300 mM imidazole. The eluate was dialyzed into PPX buffer (50 mM Tris-Cl pH 6.5, 150 mM NaCl, 1 mM EDTA, 1mM DTT) for 12 hr at 4°. PreScission protease (4 u; GE Healthcare) was added to the dialyzed protein and incubated for another 12 hr at 4°C before the sample was applied to a 0.5 ml of Ni-NTA column equilibrated in PPX buffer. The flowthrough was collected and concentrated using 5 ml Microcon centrifugal filtration units with a 10 kDa molecular weight cutoff, exchanged with storage buffer (50% glycerol, 10 mM Tris-Cl at pH 7.9, 100 mM NaCl, 0.1 mM DTT, 0.1 mM EDTA), and stored at −20°C. Protein concentrations were measured using the Bradford assay reagent (Bio-Rad) using bovine serum albumen (BSA) as a standard. Holoenzymes were formed with 4-10 fold excess Bpa-containing σ^{70}.

DksA was purified as described in Paul et al., 2004 and Ross et al., 2016. TraR was purified as described in Gopalkrishnan et al., 2017. GreB was purified as described in Lee et al., 2012.

Crosslinking and primer extension mapping

10 μl crosslinking reactions were performed by incubating 40 nM Bpa-containing RNAP with 2 nM linear PCR DNA containing the wild-type *rpsT* P2 or mutant *rpsT* P2 promoters in transcription buffer [10 mM Tris-Cl pH 8.0, 30 mM KCl, 10 mM $MgCl_2$, 0.2 mg/ml BSA (NEB) and 1 mM DTT] for 5 min in a 37° water bath. Reactions in 1.5 ml microfuge tubes were placed directly onto the surface of a UV transilluminator with two 15 watt bulbs and irradiated with 365 nm UV light for 1 min. Samples were then

returned to a water bath for 1 min while the lamp was turned off to prevent UV-bulb overheating. Irradiation and water bath incubation were repeated for a total of 10 min of UV-exposure.

2 μl of each crosslinking reaction was used as a template in 12.5 μl primer extension reactions. Reactions also contained 1.25 units of *Taq* DNA polymerase (NEB), 1X Taq buffer (NEB), 250 μM of each dNTP, 2M betaine, 5% DMSO, and ~1 pmol of radiolabeled primer (primer 5910 to monitor crosslinks to the non-template strand and primer 7176 to monitor crosslinks to the template strand). Primer 5910 annealed to the non-template strand of the plasmid backbone, 51-76 nt downstream from the transcription start site. Primer 7176 annealed to the template strand from –77 to -55 relative to the start site (-77 to -69 was from the plasmid backbone sequence and -68 to -55 was from the promoter sequence). Extension products were amplified by 18 cycles of PCR (30 s at 95°C, 30 s at 53°, and 30 s at 72°). An equal volume of primer extension reaction and loading solution (8 M urea, 0.5X TBE, 0.05% bromophenol blue, 0.05% xylene cyanol) were mixed, loaded onto a 40 cm, 9.5% acrylamide, 0.5X TBE, 7M urea gel, and electrophoresed for ~2.5 hr at 2000 V. GATC sequencing ladders were generated with primer 5910 or 5853 and the same template DNAs used for crosslinking, using the Thermo Sequenase Cycle Sequencing Kit (Affymetrix).

DNaseI and $KMnO_4$ Footprinting

For DNaseI footprinting (Bartlett et al., 1998), the template strand from either plasmid pRLG11272 (containing an *rpsT* P2 promoter fragment with endpoints -68 to +50), pRLG12844 (containing a T-7A *rpsT* P2 promoter fragment with endpoints -68 to

+50), or pRLG12829 (containing an A-11C *rpsT* P2 promoter fragment with endpoints -68 to +50) was digested with NcoI (NEB), end-labeled by filling-in with [α-^{32}P] dCTP (Perkin-Elmer) using Sequenase (USB), and digested with NheI (NEB). To label the non-template strand, the above plasmids were first digested with NheI, labeled by filling-in the end of the promoter fragment with [α-^{32}P] dCTP, then digested with NcoI. The DNA was concentrated after each step by ethanol precipitation, and electrophoresed on a 5% acrylamide gel. The promoter fragments were then excised from the gel, diffused overnight into low salt elution buffer (0.2 M NaCl, 20 mM Tris-Cl pH 7.4, 1 mM EDTA), purified using a Qiagen PCR purification kit, ethanol precipitated, and resuspended in 100 µl 10 mM Tris-HCl pH 8.0.

For DNaseI footprinting, 40-80 nM RNAP was added to ~0.2 nM template DNA in 30 mM KCl transcription buffer (10 mM Tris-HCl, pH 8.0, 30 mM KCl, 10 mM $MgCl_2$, 1 mM DTT, 0.1 µg/µl BSA).1 µl of DNaseI (to a final concentration of 0.18-0.5 µg/ml was added for 30 sec before the reaction was stopped by addition of 10 mM EDTA, 0.3 M sodium acetate, and phenol. Glycogen was added to the aqueous fraction, and the DNA was precipitated with ethanol, washed with 100% ethanol, dried, and suspended in 4 µl loading buffer (7 M urea, 0.5x TBE 0.05% bromophenol blue, 0.05% xylene cyanole). Control reactions were performed without DNaseI and RNAP.

For $KMnO_4$ footprinting (Newlands et al., 1991), RNAP-promoter complexes were formed with promoter fragments radiolabeled on the template or non-template strand and treated with $KMnO_4$ (2.5 mM) for 2 min at 37°C. Reactions were terminated with β-mercaptoethanol (0.34 M final concentration) and precipitated with 0.5 M sodium acetate, glycogen, and 2 vol of ethanol. DNA was resuspended in 2 M ammonium

acetate and precipitated again in ethanol. The pellet was washed with ethanol, dried, and suspended in 100 µl 1 M piperidine, heated at 90°C for 30 min, precipitated with 0.3 M sodium acetate and ethanol, washed extensively with ethanol, air-dried, and resuspended in 4 µl loading buffer.

A+G ladders used as markers were made using the same templates as those used for footprinting. 12 µl of template DNA (in H_2O) was depurinated with 50 µl formic acid at room temperature for 7 min (Ross et al., 2001). The DNA was precipitated with ethanol, washed, dried, and suspended in 100 µl 1 M piperidine, heated at 90°C for 30 min, precipitated with ethanol, washed, air-dried, and suspended in 15 µl loading solution.

All samples were heated briefly to 90°C before loading on a 9.5% acrylamide, 7 M urea sequencing gel. Gels were dried under vacuum at 80°C and exposed to a phosphorimager screen overnight.

In vitro Transcription

Multi-round *in vitro* transcription were carried out as in Gopalkrishnan et al., 2017. Briefly, transcription was performed at 23°C in buffer containing 10mM Tris-Cl pH7.9, 170mM NaCl, 10mM $MgCl_2$, 1mM DTT, 100µg/ml BSA, 500µM ATP, 100µM GTP, 100µM CTP, 10µM UTP, and 1.5µCi α-32P UTP (Perkin Elmer). 50ng supercoiled plasmid DNA template was used for each reaction. Reactions were initiated by addition of 20nM RNAP and incubated for 15min before quenching with an equal volume of stop solution (95% formamide, 20mM EDTA, 0.05% bromophenol blue, 0.05% xylene

cyanol). Transcripts were separated on a 6% acrylamide-7M urea denaturing gel and analyze by phosphorimaging.

References

Artsimovitch, I., Svetlov, V., Murakami, K.S., and Landick, R. (2003). Co-overexpression of Escherichia coli RNA polymerase subunits allows isolation and analysis of mutant enzymes lacking lineage-specific sequence insertions. *J. Biol. Chem.* **278**:12344-12355.

Bartlett, M.S., Gaal, T., Ross, W., and Gourse, R.L. (1998). RNA polymerase mutants that destabilize RNA polymerase-promoter complexes alter NTP-sensing by rrn P1 promoters. *J. Mol. Biol.* **279**:331-345.

Burgess, R.R. and Jendrisak, J.J. (1975). A procedure for the rapid, large-scall purification of Escherichia coli DNA-dependent RNA polymerase involving Polymin P precipitation and DNA-cellulose chromatography. *Biochemistry* **14**:4634-4638.

Chakraborty, A., Wang, D., Ebright, Y.W., Korlann, Y., Kortkhonjia, E., Kim, T., Chowdhury, S., Wigneshweraraj, S., Irschik, H., Jansen, R., Nixon, B.T., Knight, J., Weiss, S., and Ebright, R.H. (2012). Opening and closing of the bacterial RNA polymerase clamp. *Science* **337**:51-595.

Chen, Y.F. and Helmann, J.D. (1997). DNA-melting at the Bacillus subtilis flagellin promoter nucleates near -10 and expands unidirectionally. *J. Mol. Biol.* **267**:47-59.

Feklistov, A. and Darst, S.A. (2011). Structural basis for promoter-10 element recognition by the bacterial RNA polymerase σ subunit. *Cell* **147**:1257-1269.

Feklistov, A., Bae, B., Hauver, J., Lass-Napiorkowska, A., Kalesse, M., Glaus, F., Altmann, K.H., Heyduk, T., Landick, R., and Darst, S.A. (2017). RNA polymerase motions during promoter melting. *Science* **356**:863-866.

Furman, R., Tsodikov, O.V., Wolf, Y.I., and Artsimovitch, I. (2013). An insertion in the catalytic trigger loop gates the secondary channel of RNA polymerase. *J. Mol. Biol.* **425**:82-93.

Gopalkrishnan, S., Ross, W., Chen, A.Y., and Gourse, R.L. (2017). TraR directly regulates transcription initiation by mimicking the combined effects of the global regulators DksA and ppGpp. *Proc. Natl. Acad. Sci. USA.* **114**:E5539-E5548.

Heyduk, E., Kuznedelov, K., Severinov, K., and Heyduk, T. (2006). A consensus adenine at position -11 of the nontemplate strand of bacterial promoter is important for nucleation of promoter melting. *J. Biol. Chem.* **281**:12362-12369.

Laptenko, O., Lee, J., Lomakin, I., and Borukhov, S. (2003). Transcript cleavage factors GreA and GreB act as transient catalytic components of RNA polymerase. *EMBO J.* **22**:6322-6334.

Lee, J.H., Lennon, C.W., Ross, W., and Gourse, R.L. (2012). Role of the coiled-coil tip of Escherichia coli DksA in promoter control. *J. Mol. Biol.* **416**:503-517.

Lemke, J.J., Sanchez-Vazquez, P., Burgos, H.L., Hedberg, G., Ross, W., and Gourse, R.L. (2011). Direct regulation of Escherichia coli ribosomal protein promoters by the transcription factors ppGpp and DksA. *Proc. Natl. Acad. Sci. USA.* **108**:5712-5717.

Lim, H.M., Lee, H.J., Roy, S., and Adhya, S. (2001). A “master” in base unpairing during isomerization of a promoter upon RNA polymerase binding. *Proc. Natl. Acad. Sci. USA.* **98**:14849-14852.

Nayak, D., Voss, M., Windgassen, T., Mooney, R.A., and Landick, R. (2013). Cys-pair reporters detect a constrained trigger loop in a paused RNA polymerase. *Mol. Cell* **50**:882-893.

Newlands, J.T., Ross, W., Gosink, K.K., and Gourse, R.L. (1991). Factor-independent activation of Escherichia coli rRNA transcription. II. Characterization of complexes of rrnB P1 promoters containing or lacking the upstream activator region with Escherichia coli RNA polymerase. *J. Mol. Biol.* **220**:569-583.

Parshin, A., Shiver, A.L., Lee, J., Ozerova, M., Schneidman-Duhovny, D., Gross, C.A., and Borukhov, S. (2015). DksA regulates RNA polymerase in Escherichia coli through a network of interactions in the secondary channel that includes Sequence Insertion 1. *Proc. Natl. Acad. Sci. USA.* **112**:E6862-E6871.

Paul, B.J., Barker, M.M., Ross, W., Schneider, D.A., Webb, C., Foster, J.W., and Gourse, R.L. (2004). DksA: a critical component of the transcription initiation machinery that potentiates the regulation of rRNA promoters by ppGpp and the initiating NTP. *Cell* **118**:311-322.

Paul, B.J., Berkmen, M.B., and Gourse, R.L. (2005). DksA potentiates direct activation of amino acid promoters by ppGpp. *Proc. Natl. Acad. Sci. USA.* **102**:7823-7828.

Qiu, J. and Helmann, J.D. (1999). Adenines at -11, -9, and -8 play a key role in the binding of Bacillus subtilis Esigma(A) RNA polymerase to -10 region single-stranded DNA. *Nucleic Acids Res.* **27**:4541-4546.

Ross, W., Ernst, A., and Gourse, R.L. (2001). Fine structure of E. coli RNA polymerase-promoter interactions: alpha subunit binding to the UP element minor groove. *Genes Dev.* **15**:491-506.

Ross, W., Sanchez-Vazquez, P., Chen, A.Y., Lee, J.H., Burgos, H.L., and Gourse, R.L. (2016). ppGpp Binding to a Site at the RNAP-DksA Interface Accounts for Its Dramatic Effects on Transcription Initiation during the Stringent Response. *Mol. Cell* **62**:811-823.

Rutherford, S.T., Lemke, J.J., Vrentas, C.E., Gaal, T., Ross, W., and Gourse, R.L. (2007). Effects of DksA, GreA, and GreB on transcription initiation: insights into the mechanisms of factors that bind in the secondary channel of RNA polymerase. *J. Mol. Biol.* **366**:1243-1257.

Rutherford, S.T., Villers, C.L., Lee, J.H., Ross, W., and Gourse, R.L. (2009). Allosteric control of Escherichia coli rRNA promoter complexes by DksA. *Genes Dev.* **23**:236-248.

Ryu, Y. and Schultz, P.G. (2006). Efficient incorporation of unnatural amino acids into proteins in Escherichia coli. *Nat. Methods* **3**:263-265.

Schroeder, L.A., Gries, T.J., Saecker, R.M., Record, M.T. Jr., Harris, M.E., and DeHaseth, P.L. (2009). Evidence for a tyrosine-adenine stacking interaction and for a short-lived open intermediate subsequent to initial binding of Escherichia coli RNA polymerase to promoter DNA. *J. Mol. Biol.* **385**:339-349.

Svetlov, V. and Artsimovitch, I. (2015). Purification of bacterial RNA polymerase: tools and protocols. *Methods Mol. Biol.* **1276**:13-29.

Tetone, L.E., Friedman, L.J., Osborne, M.L., Ravi, H., Kyzer, S., Stumper, S.K., Mooney, R.A., Landick, R., and Gelles, J. (2017). Dynamics of GreB-RNA polymerase interaction allow a proofreading accessory protein to patrol for transcription complexes needing rescue. *Proc. Natl. Acad. Sci. USA.* **114**:E1081-E1090.

Tsujikawa, L., Strainic, M.G., Watrob, H., Barkley, M.D., and DeHaseth, P.L. (2002). RNA polymerase alters the mobility of an A-residue crucial to polymerase-induced melting of promoter DNA. *Biochemistry* **41**:15334-15341.

Winkelman, J.T., Winkelman, B.T., Boyce, J., Maloney, M.F., Chen, A.Y., Ross, W., and Gourse, R.L. (2015). Crosslink Mapping at the Amino Acid-Base Resolution Reveals the Path of Scrunched DNA in Initial Transcribing Complexes. *Mol. Cell* **59**:768-780.

Xu, J., McCabe, B.C., and Koudelka, G.B. (2001). Function-based selection and characterization of base-pair polymorphisms in a promoter of Escherichia coli RNA polymerase-sigma(70). *J. Bacteriol.* **183**:2866-2873.

Zhang, Y., Feng, Y., Chatterjee, S., Tuske, S., Ho, M.X., Arnold, E., and Ebright, R.H. (2012). Structural basis of transcription initiation. *Science* **336**:1076-1080.

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. *Mol. Cell* **50**:430-436.

Supplemental Materials

Figure 3.S1. DksA favors a melting-nucleated state. (A) $KMnO_4$ footprints were performed on template strand labeled DNA without RNAP (lane 1), with RNAP at 37°C (lane 2), with RNAP at 28°+DksA (lane 3), with RNAP at 12° (lane 4). **(B)** Dimethyl sulfate (DMS) footprint. DMS was reacted with nontemplate strand labeled *rpsT* P2 DNA. Hydrazine cleavage was used to detect methylated cytosines. **(C)** Gel traces corresponding to DMS footprints shown in **(B)**.

Figure 3.S1

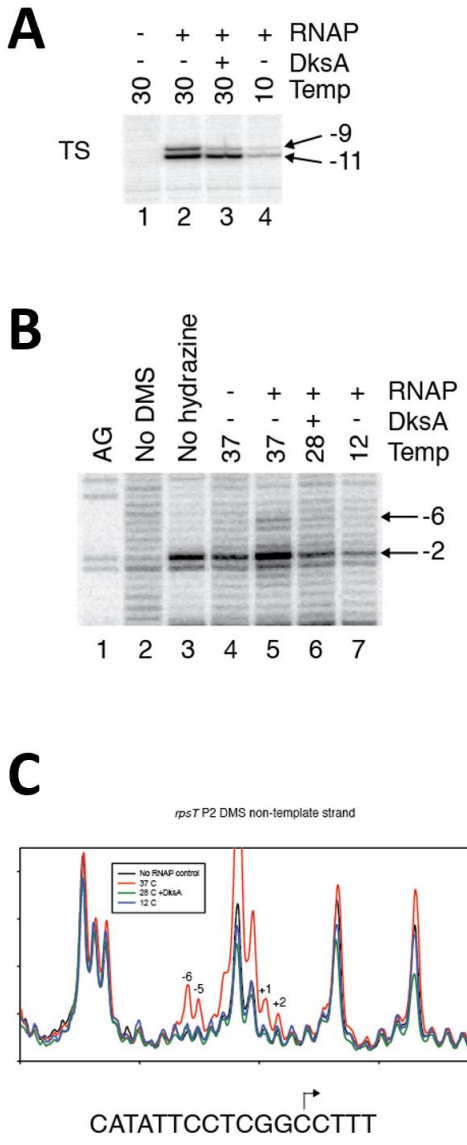

A
RNAP
DksA
Temp
30 - 30 + 30 + 10 -
TS
-9
-11
1 2 3 4
B
AG
No DMS
No hydrazine
RNAP
DksA
Temp
37 - 37 + 28 + 12 -
-6
-2
1 2 3 4 5 6 7
C
rpsT P2 DMS non-template strand
No RNAP control
37 C
28 C +DksA
12 C
-6
-5
+1
+2
CATATTCCTCGGCCTTT

Figure 3.S2. Bpa-containing RNAPs form RP_C. Representative example of footprinting experiments on A-11C *rpsT* P2 at 37° to ensure that Bpa-containing RNAPs bind DNA to form RP_C. Region of DNaseI protection is indicated on the right. Bpa-containing RNAPs that protect DNA in the correct region are labeled in blue. Bpa-containing RNAPs that do not bind DNA are labeled in black.

Figure 3.S2

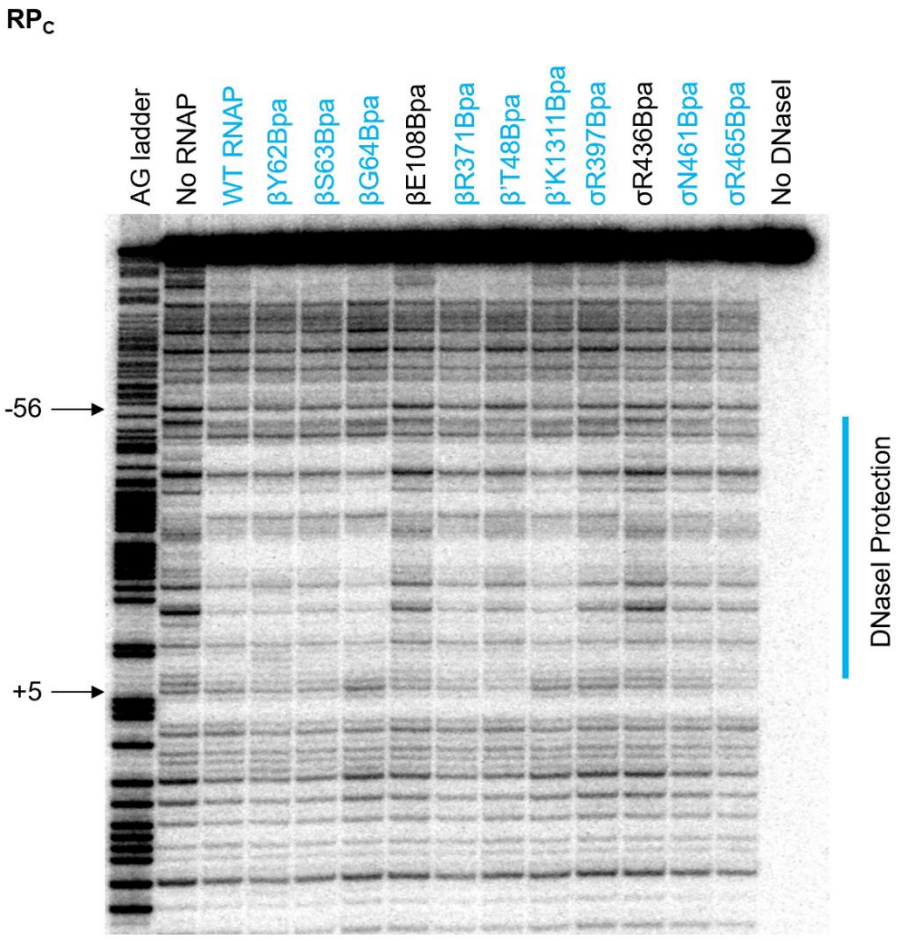

RPc
AG ladder
No RNAP
WT RNAP
βY62Bpa
βS63Bpa
βG64Bpa
βE108Bpa
βR371Bpa
β'T48Bpa
β'K1311Bpa
σR397Bpa
σR436Bpa
σN461Bpa
σR465Bpa
No DNaseI
-56
+5
DNaseI Protection

Figure 3.S3. Bpa-containing RNAPs form RP_{DksA} and RP_{CTP}. (A) $KMnO_4$ footprints were performed using nontemplate strand labeled *rpsT* P2 in the presence of DksA to determine if Bpa-containing RNAPs formed the correct complex. Bpa substitutions in the σ subunit resulted in reactivity of -8 and -4 residues, unlike in the DksA-complex with WT RNAP. **(B)** $KMnO_4$ footprints were performed using nontemplate strand labeled T-7A *rpsT* P2 in the presence of N88I. T-7A *rpsT* P2 and N88I allow formation of the partially melted intermediate with most Bpa substitutions. **(C)** $KMnO_4$ footprints were performed using nontemplate strand labeled *rpsT* P2 in the presence of CTP to determine if Bpa-containing RNAPs form RP_{CTP}.

Figure 3.S3

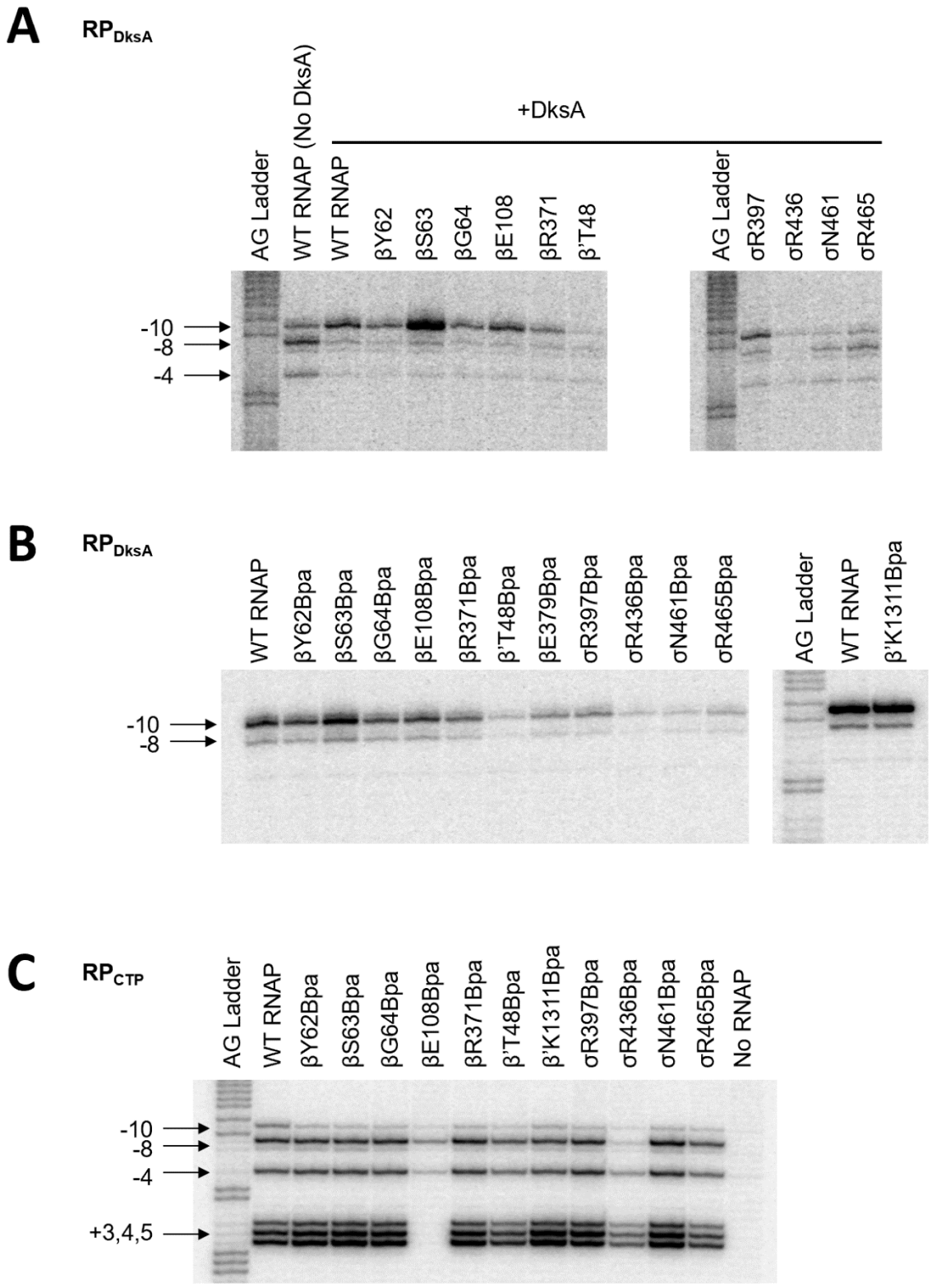

A
RPDksA
+DksA
AG Ladder
WT RNAP (No DksA)
WT RNAP
βY62
βS63
βG64
βE108
βR371
β'T48
AG Ladder
σR397
σR436
σN461
σR465
-10
-8
-4
B
RPDksA
WT RNAP
βY62Bpa
βS63Bpa
βG64Bpa
βE108Bpa
βR371Bpa
β'T48Bpa
βE379Bpa
σR397Bpa
σR436Bpa
σN461Bpa
σR465Bpa
AG Ladder
WT RNAP
β'K1311Bpa
-10
-8
C
RPCTP
AG Ladder
WT RNAP
βY62Bpa
βS63Bpa
βG64Bpa
βE108Bpa
βR371Bpa
β'T48Bpa
β'K1311Bpa
σR397Bpa
σR436Bpa
σN461Bpa
σR465Bpa
No RNAP
-10
-8
-4
+3,4,5

Figure 3.S4. Crosslink mapping of RP_C, RP_{DksA}, and RP_{CTP} complexes by Bpa-RNAPs at the *rpsT* P2 promoter. Primer extension mapping of the position of the crosslink formed in *rpsT* P2 complexes by Bpa-RNAPs. Crosslinks grouped by (1) crosslinks that form in RP_C, (2) crosslinks that occur in both RP_{DksA} and RP_{CTP} and crosslink to the same bases, and (3) crosslinks that occur in both RP_{DksA} and RP_{CTP} but crosslink to different bases. NTS indicates that crosslinks occurred to bases in the nontemplate strand. TS indicates that crosslinks occurred to bases in the template strand.

Figure 3.S4

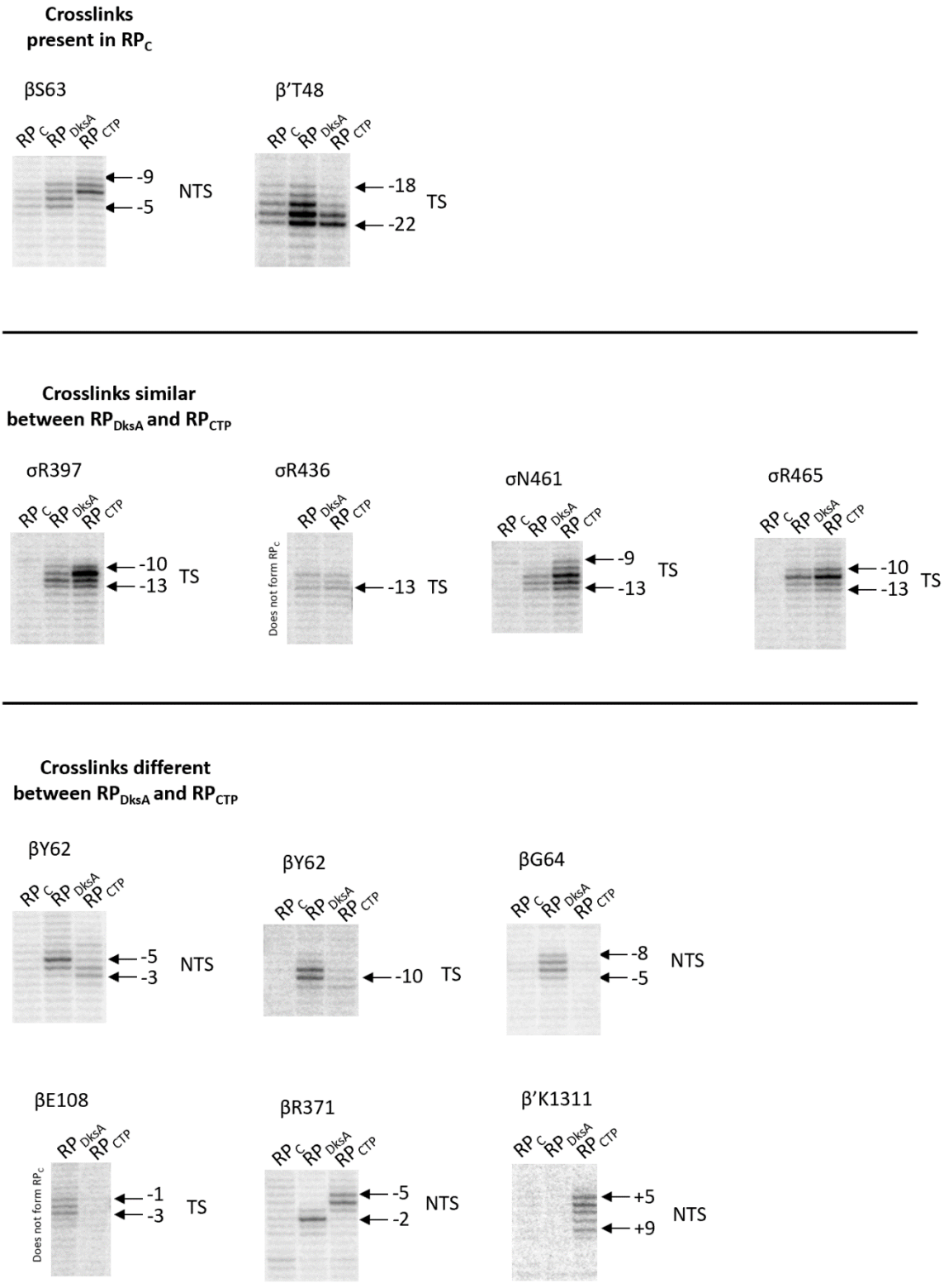

Crosslinks present in RPC
βS63
RPC
RPDksA
RPCTP
-9
-5
NTS
β'T48
-18
-22
TS
Crosslinks similar between RPDksA and RPCTP
σR397
-10
-13
TS
σR436
Does not form RPC
-13
TS
σN461
-9
-13
TS
σR465
-10
-13
TS
Crosslinks different between RPDksA and RPCTP
βY62
-5
-3
NTS
βY62
-10
TS
βG64
-8
-5
NTS
βE108
Does not form RPC
-1
-3
TS
βR371
-5
-2
NTS
β'K1311
+5
+9
NTS

Figure 3.S5. Mutations in A-9 or in the -9 pocket are still regulated by TraR. (A) Multi-round *in vitro* transcription showing the effect of TraR on WT *rpsT* P2 or *rpsT* P2 variants. Graphs show quantification from two independent experiments. **(B)** Multi-round *in vitro* transcription showing the effect of TraR on -9 pocket variants βA474V and βA474L at WT *rpsT* P2. Graphs show quantification from two independent experiments.

Figure 3.S5

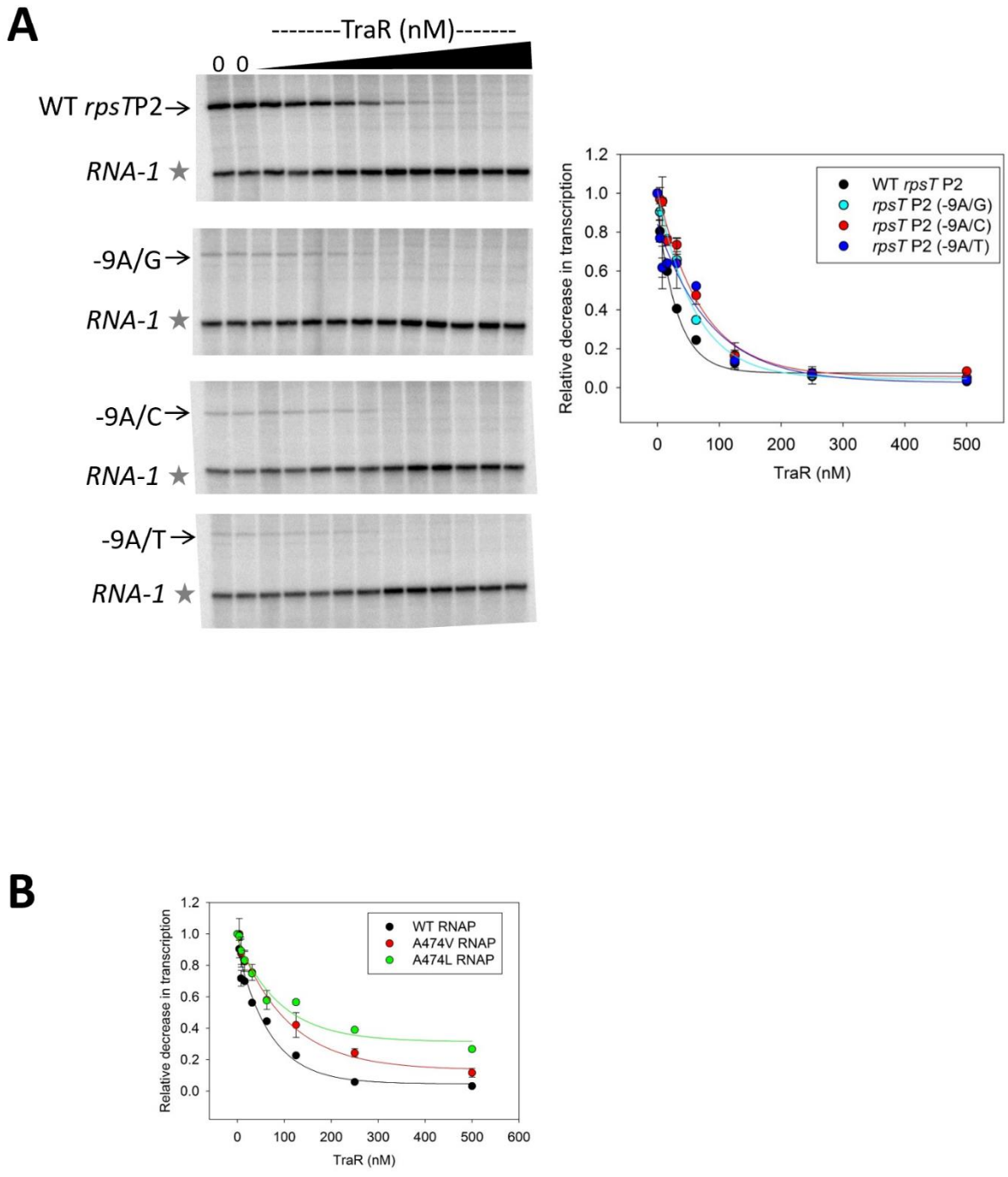

A
--------TraR (nM)-------
0 0
WT rpsTP2→
RNA-1 ★
-9A/G→
RNA-1 ★
-9A/C→
RNA-1 ★
-9A/T→
RNA-1 ★
Relative decrease in transcription
TraR (nM)
WT rpsT P2
rpsT P2 (-9A/G)
rpsT P2 (-9A/C)
rpsT P2 (-9A/T)
B
Relative decrease in transcription
TraR (nM)
WT RNAP
A474V RNAP
A474L RNAP

Figure 3.S6. -9 pocket mutants still form the partially melted intermediate. $KMnO_4$ footprints of βA474V or βA474L on non-template strand-labeled *rpsT* P2 in the presence or absence of TraR. Both βA474V and βA474L respond to TraR and form the same complexes as WT RNAP.

Figure 3.S6

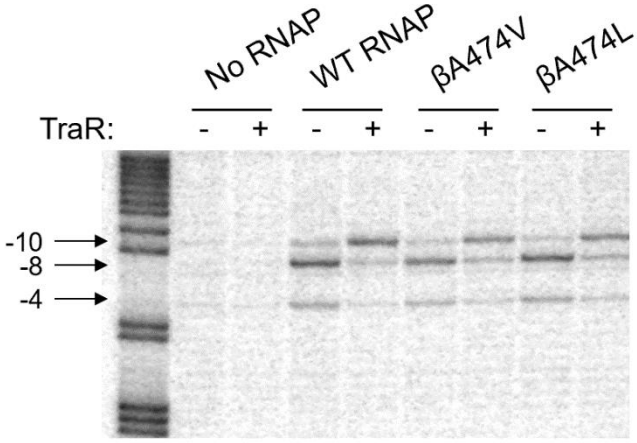


Table 3.S1. Strains and Plasmids

Strain	**Description**	**Source**
RLG4677	BL21(DE3)	Novagen
RLG7075	BL21(DE3) *dksA::tet*	Paul et al., 2004
RLG10776	DH10B	Invitrogen
Plasmid	**Description**	**Source**
pIA900	pT7 αββ'-H10ω overexpression vector	Svetlov and Artsimovitch, 2015
pRLG10148	pIA900 *rpoB* Y395D	This work
pRLG10149	pIA900 *rpoB* R454H	This work
pRLG10144	pIA900 *rpoB* G1267V	This work
pRLG10145	pIA900 *rpoB* R1269H	This work
pRLG10029	pIA900 *rpoB* P1317L	This work
pRLG10030	pIA900 *rpoC* Δ215-220	This work
pRLG10146	pIA900 *rpoC* R337S	This work
pRLG10031	pIA900 *rpoC* R339A	This work
pRLG10147	pIA900 *rpoC* Q805P	This work
pRLG15444	pIA900 *rpoB* A474V	This work
pRLG15445	pIA900 *rpoB* A474L	This work
pRLG11839	pIA900 *rpoB* Y62TAG for Bpa incorporation	Winkelman et al., 2015
pRLG11831	pIA900 *rpoB* S63TAG for Bpa incorporation	This work
pRLG13336	pIA900 *rpoB* G64TAG for Bpa incorporation	This work
pRLG13283	pIA900 *rpoB* E108TAG for Bpa incorporation	This work
pRLG11830	pIA900 *rpoB* R371TAG for Bpa incorporation	This work

pRLG13581	pIA900 *rpoC* T48TAG for Bpa incorporation	Winkelman et al., 2015
pRLG12778	pIA900 *rpoC* K1311TAG for Bpa incorporation	Winkelman et al., 2015
pRLG13105	Expresses His10-σ70 from P_{BAD} promoter	Winkelman et al., 2015
pRLG13252	p13105 *rpoD* R397TAG for Bpa incorporation	Winkelman et al., 2015
pRLG13107	p13105 *rpoD* R436TAG for Bpa incorporation	Winkelman et al., 2015
pRLG14147	p13105 *rpoD* N461TAG for Bpa incorporation	Winkelman et al., 2015
pRLG14146	p13105 *rpoD* R465TAG for Bpa incorporation	Winkelman et al., 2015
pSupT/BpF	Expresses Bpa synthetase/tRNA	Ryu and Schultz, 2006
pRLG13300	pET33 His10 HMK DksA	Ross et al., 2016
pRLG9066	pET33 His6 HMK GreB	Lee et al., 2012
pRLG10843	pET33 His6 HMK GreB-E44D	Lee et al., 2012
pRLG13084	pET28a TraR His6	Gopalkrishnan et al., 2017
pRLG11272	pRLG1507 with *rpsT P2* (-68 to +50)	This work
pRLG12829	pRLG1507 with *rpsT* P2 A-11C (-68 to +50)	This work
pRLG12844	pRLG1507 with *rpsT* P2 T-7A (-68 to +50)	This work
pRLG9237	p770 containing rpsT P2 (-89 to +50)	Lemke et al., 2011
pRLG15418	pRLG9237 with rpsT P2 A-9G	This work
pRLG15419	pRLG9237 with rpsT P2 A-9C	This work
pRLG15420	pRLG9237 with rpsT P2 A-9T	This work

Table 3.S2. Oligonucleotides

Primer	Sequence	Purpose
5637	GAAATCTACCGCATGATGTAGCCTGGC GAGCCGCCGAC	*rpoB* R371TAG
5670	CCCGATTCAGAGCTACTAGGGTAATTC CGAGCTGC	*rpoB* S63TAG
7059	GATCTATGAGCGCTAGGCGCCGGAAGGC	*rpoB* E108TAG
7168	GAGCTACAGCTAGAATTCCGAGC	*rpoB* G64TAG
7318	ACAGTTCGGTGTTCAGCGTTTCGGG	*rpoB* G1267V
7319	GGTGGTCAGCATTTCGGGGAGATGG	*rpoB* R1269H
7320	GTAAACAGGGTTCTTTCCGTCAG	*rpoB* R337S
7321	ACGTGGCGCCGGACCTGGTGGTTACC	*rpoB* Q805P
7350	GAAGACCGTGATGACTTGTCTGCGG	*rpoB* Y395D
7351	GTCGTATCCATTCCGTTGGCGAAATG	*rpoB* R454H
7353	GATGGAGCTGGGCATGCCAGAATCC	*rpoB* P1317L
7358	CGAAACCAACTCCGAAACCAAGCGTAT CAAACTGCTGGAAGCGTTCGTTCAG	*rpoC* Δ215-220
7360	GGGTCGTTTCGCTCAGAACCTGCTCGG	*rpoC* R339A
8369	GTGTAGAGCGTGTGGTGAAAGAGCGTC	*rpoB* A474V
8370	GTGTAGAGCGTCTGGTGAAAGAGCGTC	*rpoB* A474L
8388	GAAGGCTAAAAGGGCATGTTCCTCGGCCT TTGAAT	*rpsT* P2 A-9G
8389	GAAGGCTAAAAGGGCATCTTCCTCGGCCT TTGAAT	*rpsT* P2 A-9C
8390	GAAGGCTAAAAGGGCATTTTCCTCGGCCT TTGAAT	*rpsT* P2 A-9T
6519	GGCTAAAAGGGCCTATTCCTCGGCC	*rpsT* P2 A-11C

6564	GGCTAAAAGGGCATATACCTCGGCCTTTG	*rpsT* P2 T-7A
5910	GATCTACGCGTGTCGACCCAAGC	Map crosslinks to NTS DNA
7176	GGAATTCCGGCACATTAACGGC	Map crosslinks to TS DNA
1620	GCGCTACGGCGTTTCACTTC	For PCR amplification of crosslinking template DNA
6625	CGTATCACGAGGCCCTTTCGTC	For PCR amplification of crosslinking template DNA

Chapter 4

Conclusion and Future Directions

Conclusions

The objective of my thesis research was to understand how DksA and ppGpp regulate transcription in *E. coli*. DksA and ppGpp must regulate transcription initiation allosterically to activate or inhibit transcription, but it was unclear what regions of RNAP were affected and what conformational changes occurred upon DksA and ppGpp binding. Additionally, it was unclear how DksA and ppGpp activate transcription. One study showed that DksA and ppGpp increased the rate of isomerization between a closed complex and an open complex at the positively regulated promoter P*argI* (Paul et al., 2005), but isomerization during transcription initiation is a composite of multiple steps and the specific step(s) involved were unknown. The effects of DksA and ppGpp on RNAP-DNA interactions during transcription initiation were also unknown.

In Chapter 2, I addressed how DksA and ppGpp affect the conformation of RNAP. By testing mutations and deletions of various mobile regions of RNAP *in vitro*, paired with cryo-EM structures by Andrey Feklistov, I found that DksA makes direct contacts to the trigger loop, preventing rotation of the swivel module and causing closure of the clamp domain to increase promoter melting. Increasing promoter melting could explain effects of DksA and ppGpp on activation, but additional steps are required, since promoting clamp closure alone using the antibiotic myxopyronin B does not fully replicate effects of DksA and ppGpp. ppGpp binding at site 2 repositions the binding of DksA in the secondary channel and stabilizes the DksA-trigger loop interaction. Since the DksA-trigger loop interaction is needed for clamp closure, this provides an explanation for why ppGpp is required to see effects of DksA on activation.

I also found that DksA makes contacts to β sequence insertion 1, causing a 7 Å shift of the β lobe. The movement of the β lobe increases the width of the cleft, allowing increased access of DNA into the cleft, even in the presence of a closed clamp. Together with the effect of DksA and ppGpp on clamp closure, DksA and ppGpp activate transcription by increasing DNA melting and DNA loading to form the transcriptionally competent open complex. Furthermore, the β lobe participates in binding of DNA in the open complex, and repositioning of the β lobe likely disrupts RNAP-DNA contacts. Destabilization of DNA in the open complex contributes to inhibition of promoter complexes by DksA and ppGpp.

In Chapter 3, I focused on effects of DksA and ppGpp on RNAP-DNA interactions. Using DNaseI and $KMnO_4$ footprinting as well as Bpa-crosslinking, I found that DksA or its homolog TraR trap a transcription initiation intermediate at the negatively regulated promoter *rpsT* P2 where promoter melting has been nucleated but melting does not progress beyond a 3-nucleotide bubble. This intermediate is further along the transcription initiation pathway than the closed complex, where the -11A of the nontemplate strand is still in the DNA duplex and DNA is double-stranded. Once -11A has been captured in its pocket in σ region 2, as it is in the DksA-inhibited complex, these interactions remain through formation of the open complex. Furthermore, DNA downstream of the transcription start site remains sensitive to DNaseI cleavage in this partially melted intermediate, showing that DNA is not loaded in the cleft. These data are consistent with a cryo-EM structure by James Chen from the Darst lab at Rockefeller University.

Taken together, we conclude that DksA and ppGpp promote A-11 capture, nucleating transcription bubble melting. Melting of the upstream portion of the bubble does not require DNA to be loaded in the cleft. At positively regulated promoters, where DNA melting and loading are rate-limited, the increase in promoter melting in conjunction with widening of the cleft through DksA-induced β lobe movement allows increased open complex formation and increased transcription output. Since positively regulated promoters have stable open complexes, open complex destabilization by DksA and ppGpp have little effect. At negatively regulated promoters, which have intrinsically unstable open complexes, further destabilization of downstream DNA interactions with the β lobe and the cleft results in a shift towards the partially melted intermediate at *rpsT* P2 or to a closed complex at *rrnB* P1 (Rutherford et al., 2009) and decreasing transcription output.

In a broader context, these studies on regulation by DksA and ppGpp reveal the intricacies of transcription initiation. I showed that mobile regions of RNAP play a key role in transcription initiation and its regulation. Many of these movements have not been previously seen in published crystal structures, and the advancement of cryo-EM will greatly improve our understanding of the mechanism of transcription initiation and the conformational changes that occur. Furthermore, the impact of non-DNA binding transcription factors is becoming increasingly relevant (Hubin et al., 2017; Bae et al., 2015; Gopalkrishnan et al., 2017). Our studies with the non-DNA binding transcription factors DksA and ppGpp provide insights into possible RNAP regulatory mechanisms.

Future Directions

What is the dwell time of DksA on RNAP?

At positively regulated promoters, DksA and ppGpp are required for formation of an open complex. Our studies show that DksA and ppGpp bind to the RNAP holoenzyme and remain bound in the DksA-inhibited complex. However, it is unclear whether DksA and ppGpp remain bound through other intermediates of transcription initiation and when DksA and ppGpp dissociate from RNAP. It has been shown that DksA binds with decreased affinity to an open complex (Lennon et al., 2009). Furthermore, if DksA interacts with the trigger loop, it must dissociate before formation of the first nucleotide addition reaction.

At negatively regulated promoters, it is unclear whether DksA must be bound to RNAP before binding to promoter DNA to prevent open complex formation or if DksA can bind to the open complex to destabilize it. Since DksA binds poorly to an open complex, this suggests that DksA must bind to RNAP before association of RNAP with the promoter, although this must be tested.

Multiwavelength single-molecule fluorescence microscopy has previously been used to study the dynamics of GreB interactions with elongation complexes (Tetone et al., 2017). A similar study could be done to monitor DksA occupancy in the secondary channel through transcription initiation at different promoters.

How does DksA promote transcription elongation?

DksA was shown to associate with RNAP throughout transcription elongation and prevent stalling *in vivo* (Zhang et al., 2014). Since DksA binds with low affinity to an

open complex, it is unclear how it binds to elongation complexes. Furthermore, if DksA interacts with the trigger loop during transcription initiation, it is unclear what its effect would be during transcription elongation, when folding of the trigger loop is required for nucleotide addition. Since elongation pausing often involves opening of the RNAP clamp and swiveling (Weixlbaumer et al., 2013; Kang et al., 2018), it is possible that brief interactions with the trigger loop help close the clamp and subsequent dissociation of DksA allows restoration of transcription elongation. It would be helpful to determine whether interactions between DksA and the trigger loop are required for promoting transcription elongation *in vivo*.

How do DksA and ppGpp destabilize open complexes?

Movement of βSI1 by DksA and ppGpp repositions the β lobe, which contains residues that interact with downstream DNA (Zhang et al., 2012). When a region of βSI1 is deleted, activation does not occur, but increasing concentrations of DksA still allow inhibition. If β lobe movement is not the only mechanism for destabilization of open complexes, what other interactions contribute? Studies showed that deletion of the trigger loop prevents DksA from destabilizing open complexes (Rutherford et al., 2009). More studies are required to understand how this interaction allows DksA to destabilize open complexes.

How do movements of RNAP mobile regions change during the different steps of transcription initiation?

Our cryo-EM studies showed DksA and ppGpp bound to RNAP in the absence of DNA. Are these RNAP rearrangements maintained throughout the transcription cycle? Do these rearrangements occur in the absence of regulatory factors? Cryo-EM studies are currently being done by James Chen in the Darst lab at Rockefeller University that will begin to address this question at the negatively regulated promoter *rpsT* P2 in the presence of TraR. Cryo-EM has become a very useful tool for monitoring conformational changes of complex enzymes like RNAP. Studies on other promoters will help us understand whether the RNAP movements and RNAP intermediates we observed occur at all promoters or are specific to regulation by secondary channel binding factors.

References

Bae, B., Feklistov, A., Lass-Napiorkowska, A., Landick, R., and Darst, S.A. (2015). Structure of a bacterial RNA polymerase holoenzyme open promoter complex. *Elife* **4**:e08504.

Gopalkrishnan, S., Ross, W., Chen, A.Y., and Gourse, R.L. (2017). TraR directly regulates transcription initiation by mimicking the combined effects of the global regulators DksA and ppGpp. *Proc. Natl. Acad. Sci. USA.* **114**:E5539-E5548.

Hubin, E.A., Fay, A., Xu, C., Bean, J.M., Saecker, R.M., Glickman, M.S., Darst, S.A., and Campbell, E.A. (2017). Structure and function of the mycobacterial transcription initiation complex with the essential regulator RbpA. *Elife* **6**:e22520.

Kang, J.Y., Mishanina, T.V., Bellecourt, M.J., Mooney, R.A., Darst, S.A., and Landick, R. (2018). RNA Polymerase Accomodates a Pause RNA Hairpin by Global Conformational Rearrangements that Prolong Pausing. *Mol. Cell* **69**:802-815.E1.

Lennon, C.W., Gaal, T., Ross, W., and Gourse, R.L. (2009). Escherichia coli DksA binds to Free RNA polymerase with higher affinity than to RNA polymerase in an open complex. *J. Bacteriol.* **191**:5854-5858.

Paul, B.J., Berkmen, M.B., and Gourse, R.L. (2005). DksA potentiates direct activation of amino acid promoters by ppGpp. *Proc. Natl. Acad. Sci. USA.* **102**:7823-7828.

Rutherford, S.T., Villers, C.L., Lee, J.H., Ross, W., and Gourse, R.L. (2009). Allosteric control of Escherichia coli rRNA promoter complexes by DksA. *Genes Dev.* **23**:236-248.

Tetone, L.E., Friedman, L.J., Osborne, M.L., Ravi, H., Kyzer, S., Stumper, S.K., Mooney, R.A., Landick, R., and Gelles, J. (2017). Dynamics of GreB-RNA polymerase interaction allow a proofreading accessory protein to patrol for transcription complexes needing rescue. *Proc. Natl. Acad. Sci. USA.* **114**:E1081-E1090.

Weixlbaumer, A., Leon, K., Landick, R., and Darst, S.A. (2013). Structural basis of transcriptional pausing in bacteria. *Cell* **152**:431-441.

Zhang, Y., Feng, Y., Chatterjee, S., Tuske, S., Ho, M.X., Arnold, E., and Ebright, R.H. (2012). Structural basis of transcription initiation. *Science* **336**:1076-1080.

Zhang, Y., Mooney, R.A., Grass, J.A., Sivaramakrishnan, P., Herman, C., Landick, R., and Wang, J.D. (2014). DksA guards elongating RNA polymerase against ribosome-stalling-induced arrest. *Mol. Cell* **53**:766-778.

Appendix A

DksA weakens interactions between σ1.1 and the RNAP cleft

I performed all experiments in this appendix.

Introduction

σ region 1.1 is an autoinhibitory domain present in all housekeeping sigma factors, including the housekeeping sigma factor σ^{70} in E. coli (Gruber and Gross, 2003). When σ^{70} is free in solution, σ1.1 interacts with σ regions 2.4 and 4.2, preventing free sigma from binding promoter DNA (Dombroski et al., 1992; Schwartz et al., 2008). Upon association of σ^{70} with core RNA polymerase, σ1.1 relocates to the cleft of RNAP, unmasking σ regions 2.4 and 4.2 to allow binding of the -10 element and -35 element respectively. σ1.1 must be subsequently removed from the cleft to allow DNA to enter and bind near the active site (Mekler et al., 2002).

σ1.1 (σ N-terminal residues 1-98) consists of a plug domain (σ residues 1-50) and an acidic linker (σ residues 51-98). Deletion of the plug domain has mild effects on transcription initiation while deletion of the linker results in a large decrease in forward isomerization (Wilson et al., 1997), as well as effects on open complex stabilization (Ruff et al., 2015). However, its effects vary at different promoters, increasing transcription at some while decreasing transcription at others (Vuthoori et al., 2001).

Both σ1.1 and DksA affect transcription initiation. In Chapter 2, I showed that σ1.1 is required for activation by DksA and ppGpp, but not for inhibition. One model is that DksA and ppGpp activate transcription by favoring the removal of σ1.1 from the cleft by disrupting contacts between the β lobe and σ1.1. Here we test whether DksA and ppGpp decrease σ1.1 occupancy in the cleft and test if weakening σ1.1 contacts to the cleft affects regulation by DksA and ppGpp.

Results

DksA decreases crosslinking between σ1.1 and core RNAP

To test whether DksA favors removal of σ1.1 from the cleft, we used Bpa-crosslinking to measure σ1.1 occupancy in the presence or absence of DksA and ppGpp. The photoactivatable, non-natural amino acid *p*-benzoyl-L-phenylalanine (Bpa) can be incorporated at specific residues in the polypeptide (Winkelman et al., 2015; Ryu and Schultz, 2006). Residue 163 in the β subunit is in the cleft, and when Bpa is incorporated at this position, Bpa can crosslink to σ in a UV-dependent manner (Figure A.1A). Upon addition of dsDNA, this crosslink does not occur (Figure A.1B), consistent with the idea that σ1.1 must leave the cleft to allow entry of DNA.

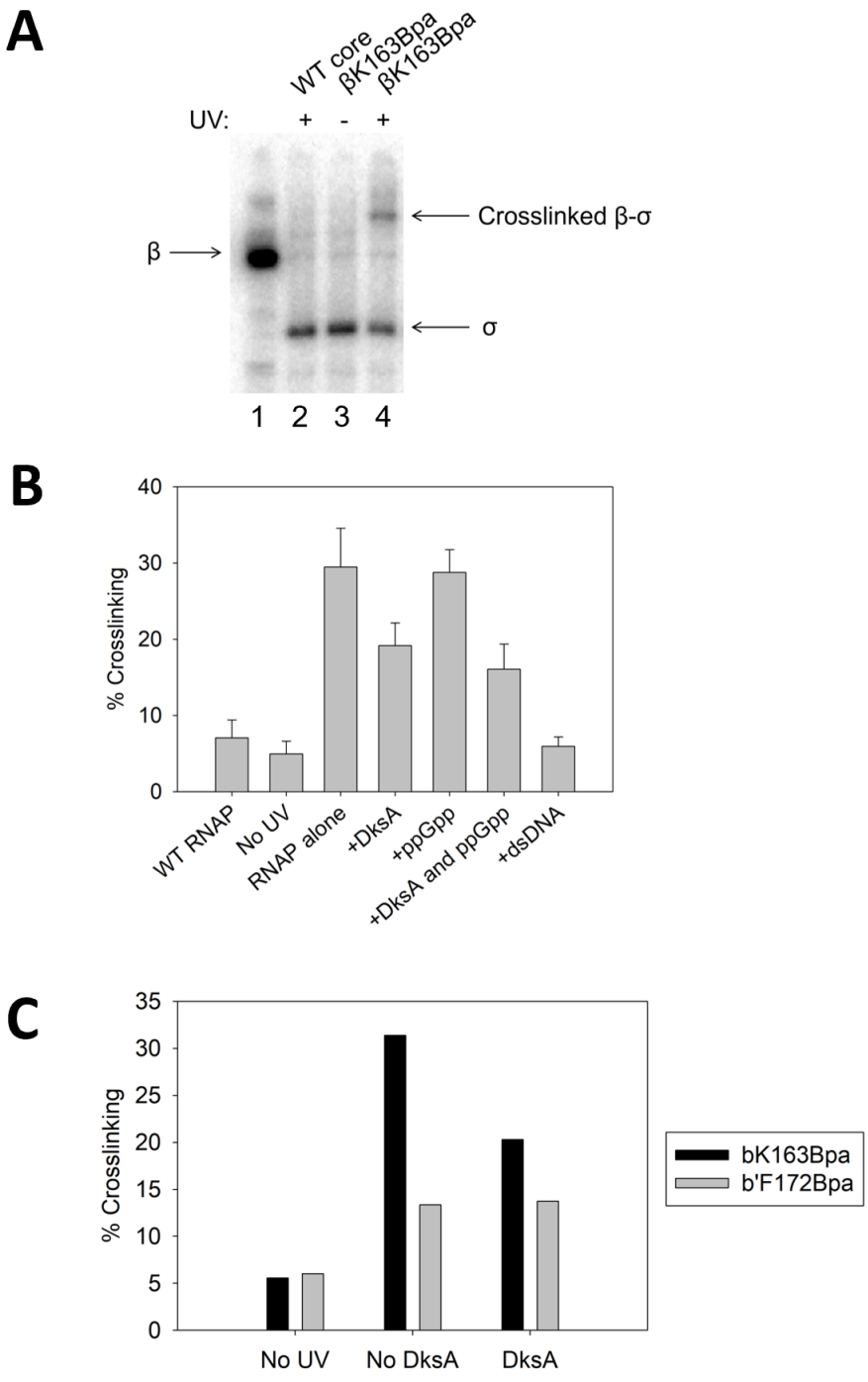
When DksA is present, crosslinking efficiency decreases by nearly 50% (Figure A.1B). ppGpp slightly enhances the effect of DksA, while ppGpp alone has no effect on crosslinking efficiency. As a control, Bpa was incorporated at position 172 in the β' subunit, which crosslinks to the non-conserved region of σ (Banta, 2013). DksA does not affect crosslinking efficiency at this position, indicating that DksA specifically affects crosslinking to σ1.1 (Figure A.1C). Together, this suggests that DksA repositions σ1.1 or favors its removal from the cleft to promote activation of transcription.

Mutations in σ1.1 affect activation by DksA and ppGpp

The interaction between σ1.1 and the cleft involves many residues. We made single or triple alanine substitutions in various parts of σ1.1 to weaken the stability of σ1.1 in the cleft (Figure A.2A). These substitutions were designed to disrupt

Figure A.1. DksA decreases crosslinking from βK163Bpa to σ, but not from β'F172Bpa. (A) SDS-PAGE gel showing crosslinking between ^{32}P-labeled σ and βK163Bpa with or without UV exposure. ^{32}P-labeled β is used as a marker in lane 1. WT RNAP is used as a negative control in lane 2. **(B)** Quantification of crosslinking between ^{32}P-labeled σ and βK163Bpa in the presence of various factors (n=3). **(C)** Quantification of crosslinking between ^{32}P-labeled σ and βK163Bpa (black bars) or β'F172Bpa (gray bars). n=1.

Figure A.1


A
WT core
βK163Bpa
βK163Bpa
UV: + - +
β
Crosslinked β-σ
σ
1 2 3 4
B
% Crosslinking
0
10
20
30
40
WT RNAP
No UV
RNAP alone
+DksA
+ppGpp
+DksA and ppGpp
+dsDNA
C
% Crosslinking
0
5
10
15
20
25
30
35
No UV
No DksA
DksA
bK163Bpa
b'F172Bpa

the interface between σ1.1 and the β lobe (σ E18A Q19A Y21A and σ D26A N28A D29A), switch 1 (σ N70A T71A D73A), or the β' clamp (σ E88A).

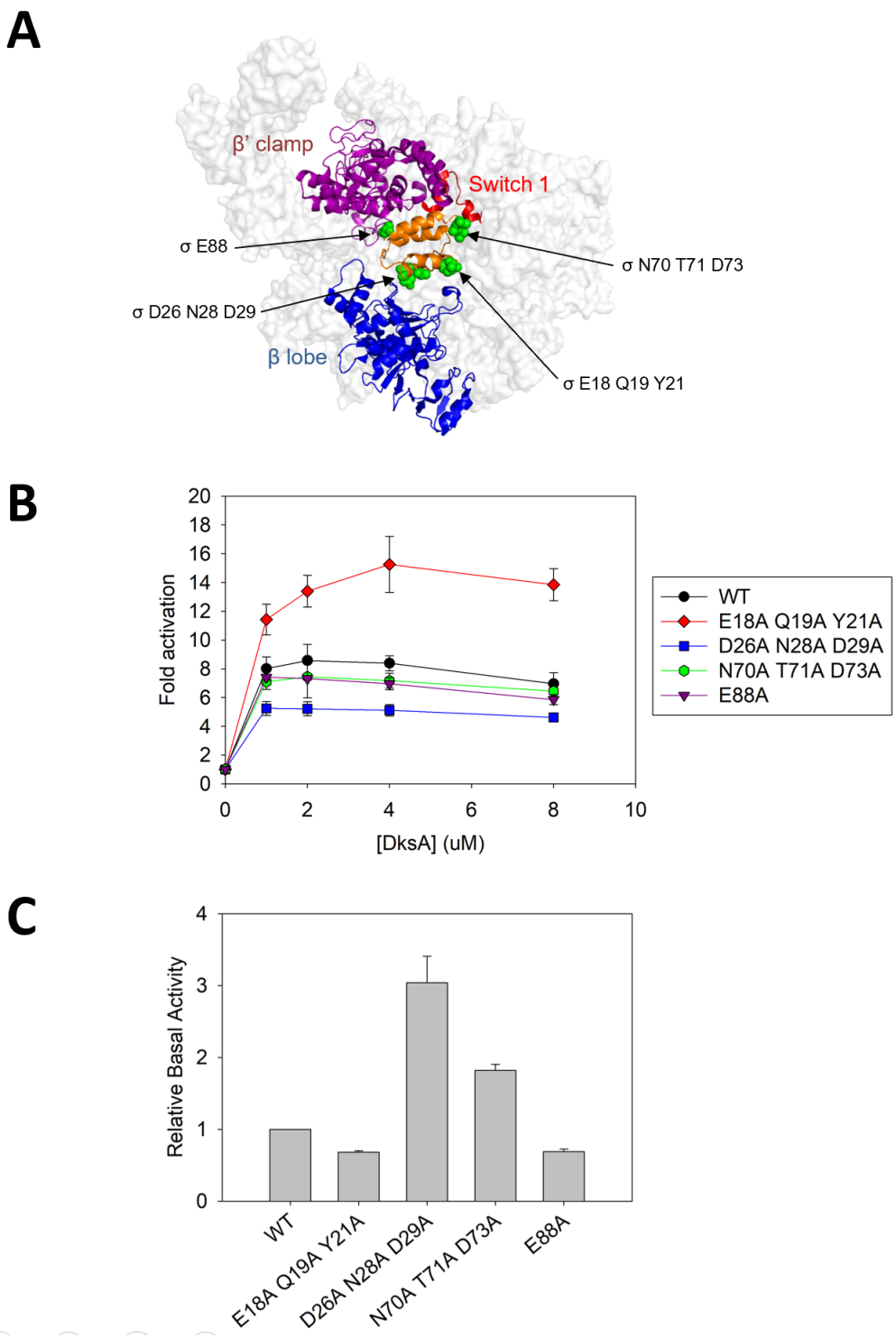
To test whether stability of σ1.1 in the cleft affects activation by DksA and ppGpp, we measured activation using *in vitro* transcription at the positively regulated *iraP* promoter. WT RNAP is activated ~8-fold under these conditions (Figure A.2B). A triple alanine substitution of σ residues E18, Q19, and Y21 increased activation to 14-fold. By weakening the interaction between σ1.1 and the cleft through alanine substitutions, DksA and ppGpp more easily promotes RPo formation, resulting in a greater activation effect.

Surprisingly, holoenzyme with other σ variants D26A N28A D29A, N70A T71A D73A, and E88A all acted like WT, being activated ~8-fold by DksA and ppGpp. While basal activity of these σ variants varies slightly, it does not explain the difference in response to DksA and ppGpp. Both σ E18A Q19A Y21A and σ E88A variants had similar decreased basal activity, but σ E88A acted like WT for activation while the triple mutant did not (Figure A.2C).

All substitutions were intended to destabilize σ1.1 binding in the cleft, but it is unclear if all were successful. It is possible that only σ E18A Q19A Y21A impairs binding while others are unaffected, explaining why only σ E18A Q19A Y21A had a unique response to DksA and ppGpp. Alternatively, if all substitutions affect binding of σ1.1 in the cleft, the results would indicate that DksA and ppGpp specifically affects interactions with E18 Q19 Y21 as opposed to general occupancy of σ1.1 in the cleft. Consistent with this idea, βK163 is positioned to interact with these

Figure A.2. Effect of alanine substitutions in σ1.1 on activation by DksA and ppGpp. (A) Crystal structure of RNAP holoenzyme (PDB: 4LK1; Bae et al., 2013) showing location of alanine substitutions in green spheres. σ1.1 is depicted in orange, β lobe in blue, switch 1 in red, and the β' clamp in purple. **(B)** Quantification of multi-round *in vitro* transcription from P*iraP* with different σ variants. Transcription was performed with 0-8uM DksA and 100uM ppGpp. Data are normalized to transcript levels in the absence of DksA (n=3). **(C)** Basal level of transcription in the absence of DksA from (B), normalized to WT RNAP.

Figure A.2


A
β' clamp
Switch 1
σ E88
σ N70 T71 D73
σ D26 N28 D29
β lobe
σ E18 Q19 Y21
B
Fold activation
[DksA] (uM)
WT
E18A Q19A Y21A
D26A N28A D29A
N70A T71A D73A
E88A
C
Relative Basal Activity
WT
E18A Q19A Y21A
D26A N28A D29A
N70A T71A D73A
E88A

residues, and the difference in crosslinking from this position may reflect a conformational change in this specific region of σ1.1.

Discussion

Preliminary data here suggest that DksA and ppGpp favor removal of σ1.1 from the cleft, and that this plays a role in the mechanism of transcription activation. As shown in chapter 2, when σ1.1 is deleted, transcription output from the positively regulated promoter *thrABC* is greatly increased, and DksA and ppGpp have no further activation effect. Bpa crosslinking shows that DksA decreases an interaction between the cleft and σ1.1. *In vitro* transcription data with a triple alanine substitution intended to destabilize binding of σ1.1 in the cleft are consistent with this model.

An alternative explanation is that DksA and ppGpp affect a specific interaction between σ1.1 and the β lobe. The Bpa crosslink affected by DksA and ppGpp (βK163Bpa) is positioned to interact with σ residues of interest E18, Q19, and Y21. It is unclear how these residues affect transcription initiation, but interaction with these residues may be needed for activation by DksA and ppGpp. More work is needed to distinguish between these models. Specifically, it would be helpful to test Bpa crosslinking between σ1.1 and other residues in the cleft to determine whether σ1.1 is fully displaced from the cleft or if it is simply repositioned such that crosslinking decreases from specific positions. The cryo-EM structure of RNAP bound to TraR or DksA do not show major changes in the conformation of σ1.1, although a shift of a few angstroms could be sufficient to prevent crosslinking.

Previous data show that σ1.1 has different effects on different promoters (Ruff et al., 2015). Deletion of σ1.1 resulted in decreased transcription from a strong promoter, Ptac, while it increased transcription from a weak promoter, Pminor (Vuthoori et al., 2001). Promoters activated by DksA and ppGpp have weak basal activity and likely have kinetic properties more like Pminor than well-studied promoters like Ptac, λPr, and T7A1. A thorough study of effects of σ1.1 on transcription initiation kinetics at positively regulated promoters would be valuable in understanding the role of σ1.1 and DksA/ppGpp.

Finally, σ1.1 is present only in Group 1 sigma factors. If σ1.1 is required for activation by DksA and ppGpp, how does DksA and ppGpp activate transcription from holoenzyme containing σ factors that naturally lack σ1.1, such as σE? DksA and ppGpp may favor open complex formation through multiple interactions, and the role of σ1.1 may be partly to tune the kinetics of transcription initiation such that certain promoters become sensitive to effects of DksA and ppGpp. At σE-dependent promoters, open complex formation may be poised at certain promoters to be regulated by DksA and ppGpp in the absence of σ1.1 solely through interactions with the trigger loop, β sequence insertion 1, or other regions of RNAP.

Materials and Methods

Protein Mutagenesis and Purification

RNAP and promoter variants were generated by site-directed mutagenesis using the QuikChange Lightning Multi Site-Directed Mutagenesis Kit (Stratagene). Transformants were streaked and restreaked for single colonies, and the

overexpression plasmids were purified and sequenced to verify the identity of the mutation. Proteins were purified as described in Chapter 3.

Bpa crosslinking

WT σ was radiolabeled using Protein Kinase A (Sigma). 400nM WT σ was incubated with 4 units of Protein Kinase A and 10μCi γ-^{32}P-ATP (Perkin Elmer) for 30min at 37°C. Free ATP was removed using a G-50 illustra MicroSpin column (GE Healthcare).

Labeled σ was incubated at 37°C for 10min in 200ul PCR strip tubes with Bpa-containing core RNAP in the presence or absence of factors in buffer containing 30mM KCl, 10mM Tris-Cl pH8, and 10mM $MgCl_2$. Reactions were exposed to 365nm UV light for 10min at 23°C. Loading dye solution was added (125mM Tris-Cl pH8.0, 4% SDS, 20% glycerol, 1.4M β-mercaptoethanol, 0.1% bromophenol blue) and reactions were run on a NuPAGE 4-12% Bis-Tris gel (Thermofisher) using MES buffer. Gels were dried, visualized using phosphorimaging, and quantified using ImageQuant.

***In vitro* Transcription**

Multi-round transcription was carried out essentially as described in Chapter 2. Reactions were carried out at 30°C in buffer containing 165mM NaCl, 10mM Tris-Cl pH8, 10mM $MgCl_2$, 1mM DTT, 100μg/ml BSA, 200μM ATP, 200μM GTP, 200μM CTP, 10μM UTP, and 1.5μCi α-32P UTP (Perkin Elmer). 20nM RNAP was incubated with 50ng supercoiled plasmid DNA template, 100uM ppGpp, and various concentrations of DksA for 15min before quenching with an equal volume of transcription stop solution

(7M urea, 2xTBE, 10mM EDTA, 1% SDS, and 0.05% bromophenol blue). Transcripts were separated on a 5.5% acrylamide-7M urea gel, visualized using phosphorimaging, and quantified using ImageQuant.

References

Banta, A.B. (2013). Molecular Interactions Between the Transcription Factor Crl and Sigma S RNA Polymerase Holoenzyme in *Escherichia coli.* Ph.D. dissertation.

Bae, B., Davis, E., Brown, D., Campbell, E.A., Wigneshweraraj, S., and Darst, S.A. (2013). Phage T7 Gp2 inhibition of Escherichia coli RNA polymerase involves misappropriation of σ70 domain 1.1. *Proc. Natl. Acad. Sci. USA.* **110**:19772-19777.

Dombroski, A.J., Walter, W.A., Record, M.T. Jr., Siegele, D.A., and Gross, C.A. (1992). Polypeptides containing highly conserved regions of transcription initiation factor sigma 70 exhibit specificity of binding to promoter DNA. *Cell* **70**:501-512.

Gruber, T.M. and Gross, C.A. (2003). Multiple sigma subunits and the partitioning of bacterial transcription space. *Annu. Rev. Microbiol.* **57**:441-466.

Mekler, V., Kortkhonjia, E., Mukhopadhyay, J., Knight, J., Revyakin, A., Kapanidis, A.N., Niu, W., Ebright, Y.W., Levy, R., and Ebright, R.H. (2002). Structural organization of bacterial RNA polymerase holoenzyme and the RNA polymerase-promoter open complex. *Cell* **108**:599-614.

Paul, B.J., Barker, M.M., Ross, W., Schneider, D.A., Webb, C., Foster, J.W., and Gourse, R.L. (2004). DksA: a critical component of the transcription initiation machinery that potentiates the regulation of rRNA promoters by ppGpp and the initiating NTP. *Cell* **118**:311-322.

Ross, W., Sanchez-Vazquez, P., Chen, A.Y., Lee, J.H., Burgos, H.L., and Gourse, R.L. (2016). ppGpp Binding to a Site at the RNAP-DksA Interface Accounts for Its Dramatic Effects on Transcription Initiation during the Stringent Response. *Mol. Cell* **62**:811-823.

Ruff, E.F., Drennan, A.C., Capp, M.W., Poulos, M.A., Artsimovitch, I., and Record, M.T. Jr. (2015). E. coli RNA Polymerase Determinants of Open Complex Lifetime and Structure. *J. Mol. Biol.* **427**:2435-2450.

Ryu, Y. and Schultz, P.G. (2006). Efficient incorporation of unnatural amino acids into proteins in Escherichia coli. *Nat. Methods* **3**:263-265.

Schwartz, E.C., Shekhtman, A., Dutta, K., Pratt, M.R., Cowburn, D., Darst, S., and Muir, T.W. (2008). A full-length group 1 bacterial sigma factor adopts a compact structure incompatible with DNA binding. *Chem. Biol.* **15**:1091-1103.

Svetlov, V. and Artsimovitch, I. (2015). Purification of bacterial RNA polymerase: tools and protocols. *Methods Mol. Biol.* **1276**:13-29.

Vuthoori, S., Bowers, C.W., McCracken, A., Dombroski, A.J., and Hinton, D.M. (2001). Domain 1.1 of the sigma(70) subunit of Escherichia coli RNA polymerase modulates the formation of stable polymerase/promoter complexes. *J. Mol. Biol.* **309**:561-572.

Wilson, C. and Dombroski, A.J. (1997). Region 1 of sigma70 is required for efficient isomerization and initiation of transcription by Escherichia coli RNA polymerase. *J. Mol. Biol.* **267**:60-74.

Winkelman, J.T., Winkelman, B.T., Boyce, J., Maloney, M.F., Chen, A.Y., Ross, W., and Gourse, R.L. (2015). Crosslink Mapping at Amino Acid-Base Resolution Reveals the Path of Scrunched DNA in Initial Transcribing Complexes. *Mol. Cell* **59**:768-780.

Supplemental Materials

Table A.S1. Strains and Plasmids

Strain	**Description**	**Source**
RLG4677	BL21(DE3)	Novagen
RLG7075	BL21(DE3) *dksA::tet*	Paul et al., 2004
RLG10776	DH10B	Invitrogen
Plasmid	**Description**	**Source**
pRLG13300	pET33 His10 HMK DksA	Ross et al., 2016
pIA900	pT7 αββ'-H10ω overexpression vector	Svetlov and Artsimovitch, 2015
pRLG10039	pIA900 *rpoB* K163TAG for Bpa incorporation	This work
pRLG13515	pIA900 *rpoC* F172TAG for Bpa incorporation	Banta, 2013
pRLG13105	Expresses His10-σ70 from P_{BAD} promoter	Winkelman et al., 2015
pRLG9982	pRLG13105 *rpoD* E18A Q19A Y21A	This work
pRLG9986	pRLG13105 *rpoD* D26A N28A D29A	This work
pRLG9983	pRLG13105 *rpoD* N70A T71A D73A	This work
pRLG9984	pRLG13105 *rpoD* E88A	This work
pSupT/BpF	Expresses Bpa synthetase/tRNA	Ryu and Schultz, 2006
pRLG11350	pRLG770 with *iraP* (-207 to +21)	Ross et al., 2016

Table A.S2. Oligonucleotides

Primer	Sequence	Purpose
3305	GGCTCTTGACAAAAGTGTTAAATTGTGCTA TACTGTATTGGTATGGATGACAGAATTCGG	dsDNA top strand
3306	CCGAATTCTGTCATCCATACCAATACAGTA TAGCACAATTTAACACTTTTGTCAAGAGCC	dsDNA bottom strand
7362	CGACAAAGGTTAGACCCACTCTTCGG	*rpoB* K163TAG
7707	CCGTGGTAAGGCAGCAGGCGCTCTGACCT ATGC	*rpoD* E18A Q19A Y21A
7708	CTGACCTATGCCGCGGTCGCTGCCCATCT GCCGGAAG	*rpoD* D26A N28A D29A
7710	GATGCTGGCTGAAGCCGCCGCGGCCGAA GATGC	*rpoD* N70A T71A D73A
7711	CTTTCCAGCGTGGCATCTGAAATCGG	*rpoD* E88A

Appendix B

DksA tip residues D74 and A76 are required for a second step in transcription activation

Wilma Ross performed the experiments in Figure B.3. I performed all other experiments.

Introduction

DksA is a transcription factor that binds in the RNAP secondary channel (Paul et al., 2004; Perederina et al., 2004). Its coiled-coil tip extends into the RNAP active site. Substitutions in the conserved DksA tip residues D74 and A76 have large defects in function, but do not affect the binding affinity of DksA to RNAP (Lee et al., 2012). D74N DksA does not destabilize open complexes and is defective for inhibiting transcription from negatively regulated promoter *rrnB* P1, either in the presence or absence of ppGpp. Additionally, D74N DksA does not activate transcription. Tip residues have been shown to crosslink to the trigger loop, as well as to interact with residues near the active site (Lennon et al., 2012; Parshin et al., 2015). Here we find that despite being defective for transcription activation, D74N and A76T DksA are still capable of increasing strand opening and are required for a proposed second step in activation.

Results

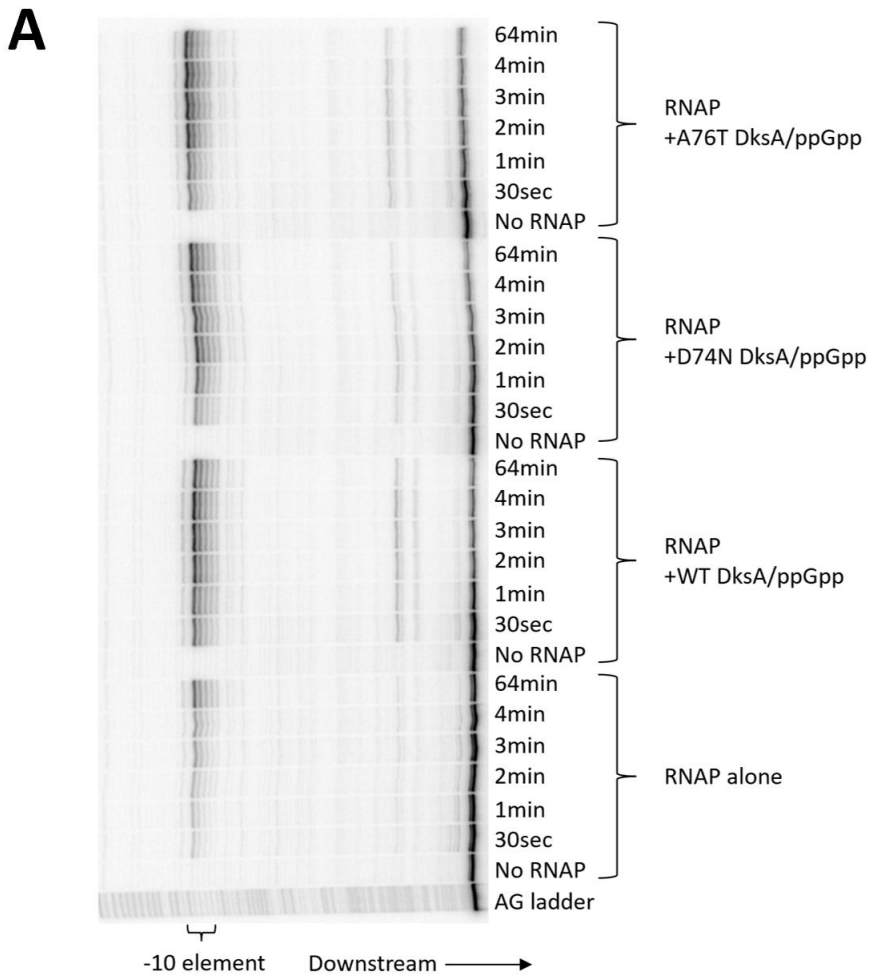
DksA tip mutants increase strand opening at P*iraP*

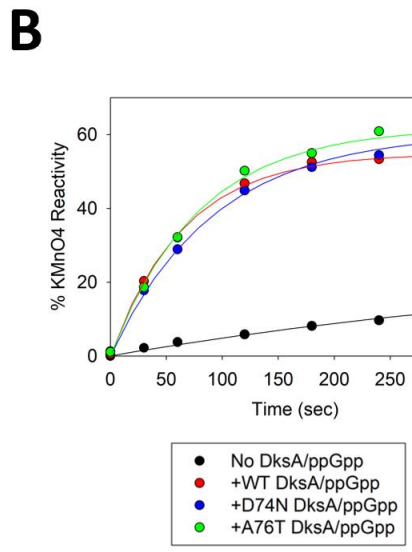
DksA and ppGpp increase the rate of open complex formation to activate transcription at certain promoters. Open complex formation requires melting of the transcription bubble, which can be detected using $KMnO_4$ footprinting. By measuring $KMnO_4$ reactivity of the transcription bubble over time, the rate of strand opening can be determined. At the positively regulated promoter P*iraP*, we find that WT DksA and ppGpp increase the observed rate of strand opening, k_{obs}, ~75-fold, from 0.002 s^{-1} in the absence of factors to 0.0153 s^{-1} in the presence of factors (Figure B.1). Surprisingly, tip mutants D74N and A76T increase the rate of strand opening to a similar magnitude,

Figure B.1. D74N and A76T DksA increase strand opening at P*iraP*.

(A) Representative gel showing $KMnO_4$ footprinting of RNAP at P*iraP* in the presence or absence of DksA/ppGpp. Template strand is labeled. **(B)** Quantification of $KMnO_4$ footprinting data from (A). Reactivity of the transcription bubble is plotted over time and fit to a single exponential curve to obtain k_{obs}, shown in the table on the right (n=3).

Figure B.1





DksA variant	k_{obs} (s^{-1})
No DksA/ppGpp	0.002 +/- 9.3664E-5
WT DksA/ppGpp	0.0153 +/- 0.0005
D74N DksA/ppGpp	0.0108 +/- 0.0005
A76T DksA/ppGpp	0.0127 +/- 0.0008

0.0108 s^{-1} and 0.0127 s^{-1} respectively. Considering the significant defect in transcription activation, the observation that DksA tip residues D74 and A76 are not required for strand opening suggests the tip residues are required for an additional function needed to achieve transcription activation.

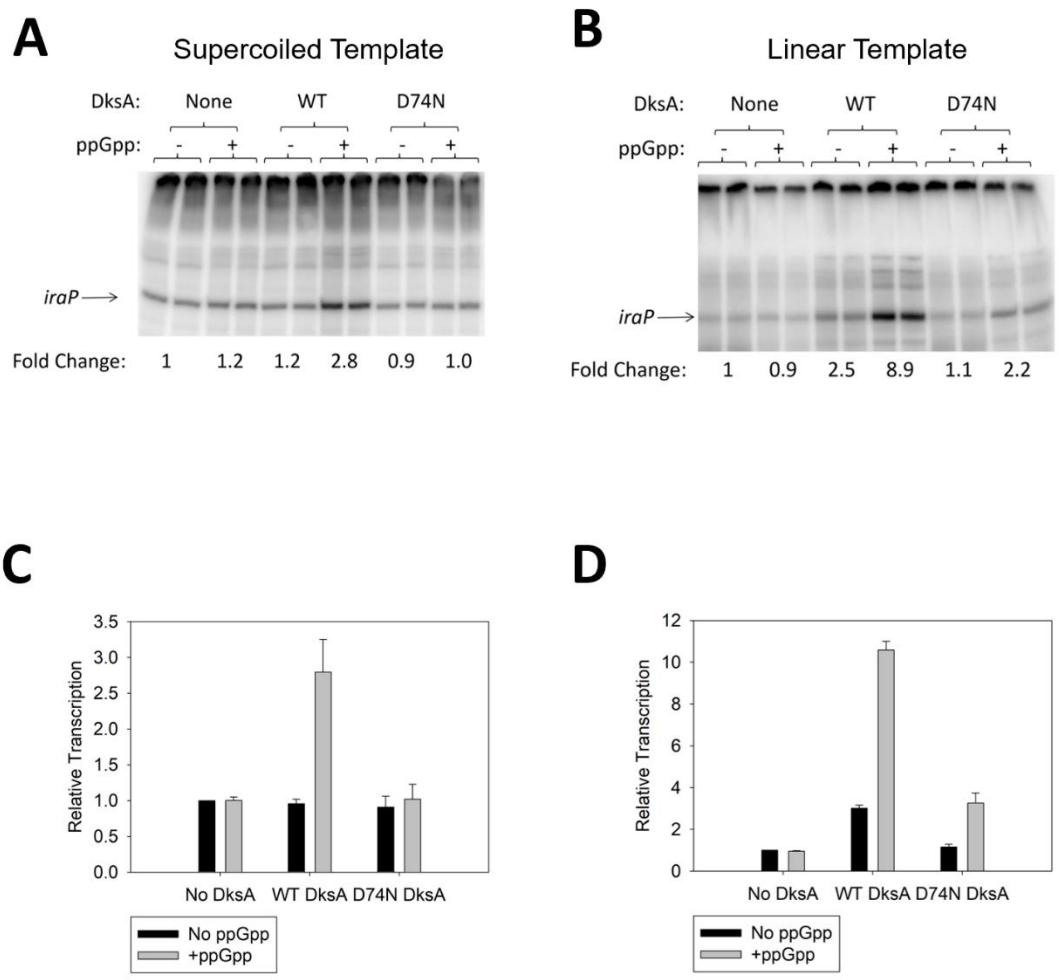
D74N DksA is defective for activation on both linear and supercoiled templates

DNA supercoiling state, temperature, and differing salt concentrations can all affect open complex formation. Activation at P*iraP* was previously studied on supercoiled templates (Ross et al., 2016), while current $KMnO_4$ experiments were performed on linear templates. To determine whether these differences allow for activation by D74N, we performed multi-round *in vitro* transcription on both supercoiled and linear templates. Consistent with previous observations, we find that WT DksA and ppGpp activate transcription ~3-fold on supercoiled templates while D74N has no activation effect (Figure B.2).

On linear templates, under the conditions used for the $KMnO_4$ footprinting, WT DksA and ppGpp activate transcription ~10-fold while D74N and ppGpp activate transcription ~2-fold. Melting of the transcription bubble is known to be less favorable if the DNA template is not supercoiled. If D74N is still able to increase strand opening, as seen in Figure B.1, the partial activation effect of D74N on linear templates may be attributed to an increase in strand opening. On supercoiled templates, strand opening may not be rate-limiting, and therefore the partial effect of D74N is not seen. Most importantly, D74N is not able to fully activate transcription on supercoiled or linear templates despite its ability to increase strand opening like WT DksA.

Figure B.2. D74N DksA is unable to activate transcription on linear or supercoiled templates. Representative gel showing multi-round *in vitro* transcription of P*iraP* from a supercoiled **(A)** or linear **(B)** template. Transcription from supercoiled templates were performed at 37°C in 165mM NaCl buffer while transcription from linear templates were performed at 37°C in 30mM KCl buffer. **(C)** Quantification of transcription on supercoiled templates. Note: This panel was were performed on a p770 based vector while all others were done on pSL6. I will repeat this for final draft. (n=3). **(D)** Quantification of transcription on linear templates. (n=3).

Figure B.2


A
Supercoiled Template
DksA: None WT D74N
ppGpp: - + - + - +
iraP
Fold Change: 1 1.2 1.2 2.8 0.9 1.0
B
Linear Template
DksA: None WT D74N
ppGpp: - + - + - +
iraP
Fold Change: 1 0.9 2.5 8.9 1.1 2.2
C
Relative Transcription
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
No DksA WT DksA D74N DksA
No ppGpp
+ppGpp
D
Relative Transcription
12
10
8
6
4
2
0
No DksA WT DksA D74N DksA
No ppGpp
+ppGpp

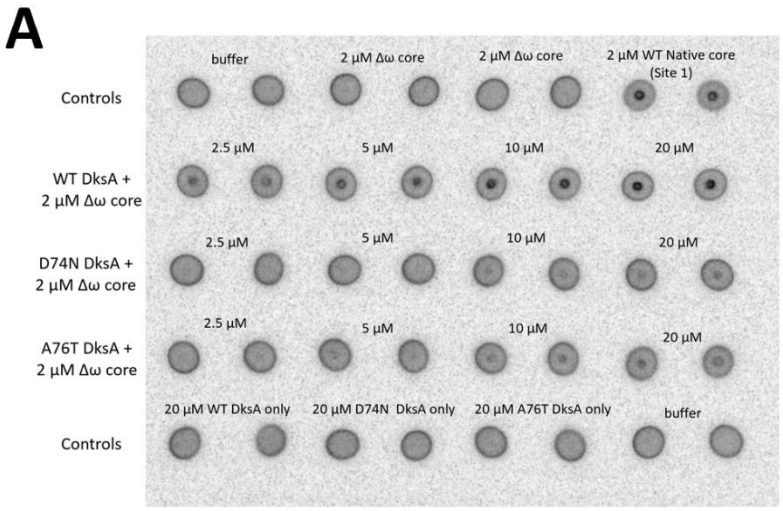
DksA tip mutants are defective for binding ppGpp

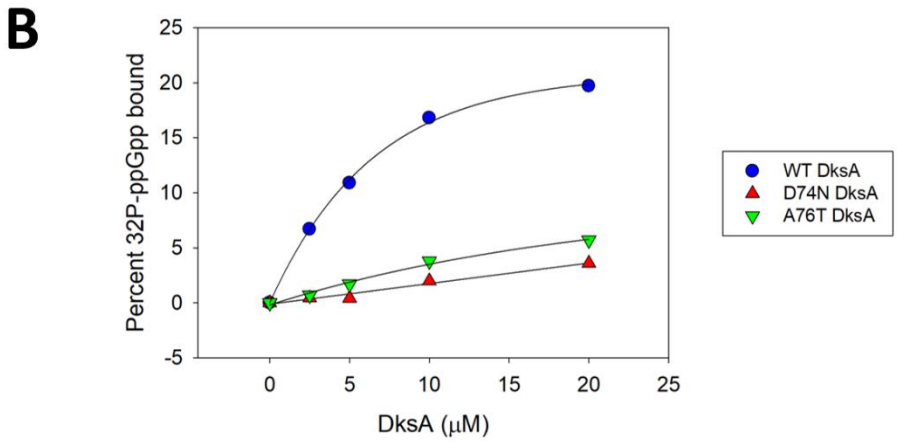
The 2-fold effect may also be attributed to a difference in salt concentrations. Linear templates require lower salt for transcription to occur, which may favor binding of DksA and ppGpp. ppGpp has previously been shown to bind at the interface between DksA and the β' rim helices, called site 2 (Ross et al., 2016). Since binding of ppGpp at site 2 is required for activation, we tested whether D74N or A76T affected ppGpp binding at site 2. We measured binding of ^{32}P-labeled ppGpp using the Differential Radial Capillary Action of Ligand Assay (DRaCALA). In the presence of WT DksA, ppGpp bound to RNAP with high affinity (Figure B.3A,B). With either D74N or A76T DksA, binding of ppGpp was dramatically weakened. The same result was obtained using multi-round *in vitro* transcription (Figure B.3C). While WT DksA required ~100uM ppGpp to reach maximal activation, D74N or A76T DksA required much higher concentrations of ppGpp to see activation. Furthermore, even at 800uM ppGpp, D74N and A76T DksA could not reach the same level of activation as WT DksA. Since the tip of DksA is located ~30 Å from the ppGpp binding site, it cannot directly interact with ppGpp. However, by anchoring the tip of DksA near the active site (Parshin et al., 2015; Molodtsov et al., 2018), it may affect the ability of DksA to properly form the ppGpp binding site.

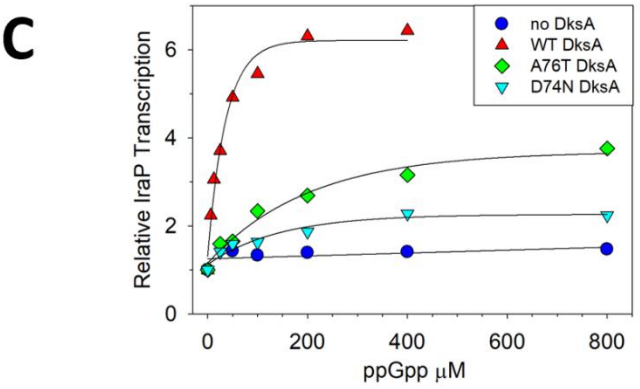
ppGpp has been shown to increase the affinity of DksA to RNAP (Molodtsov et al., 2018). Additionally, cryo-EM structures of RNAP bound to DksA in the presence and absence of ppGpp also shows that ppGpp stabilizes the binding of DksA in the secondary channel to correctly interact with the trigger loop (Chen, chapter 2). By

Figure B.3. D74N and A76T DksA are defective for ppGpp binding at site 2. (A) DRaCALA spots measuring binding of ^{32}P-labeled ppGpp to RNAP in the presence of different DksA variants. Δω RNAP is used to prevent binding of ppGpp to site 1. **(B)** Quantification of DRaCALA data from (A). **(C)** Quantification of multi-round *in vitro* transcription data at P*iraP* at various concentrations of ppGpp.

Figure B.3


A
buffer
2 μM Δω core
2 μM Δω core
2 μM WT Native core (Site 1)
Controls
2.5 μM
5 μM
10 μM
20 μM
WT DksA + 2 μM Δω core
2.5 μM
5 μM
10 μM
20 μM
D74N DksA + 2 μM Δω core
2.5 μM
5 μM
10 μM
20 μM
A76T DksA + 2 μM Δω core
20 μM WT DksA only
20 μM D74N DksA only
20 μM A76T DksA only
buffer
Controls


B
Percent 32P-ppGpp bound
25
20
15
10
5
0
-5
0
5
10
15
20
25
DksA (μM)
WT DksA
D74N DksA
A76T DksA


C
Relative IraP Transcription
6
4
2
0
0
200
400
600
800
ppGpp μM
no DksA
WT DksA
A76T DksA
D74N DksA

preventing ppGpp binding, substitutions in the DksA tip could prevent interactions with the trigger loop and prevent activation.

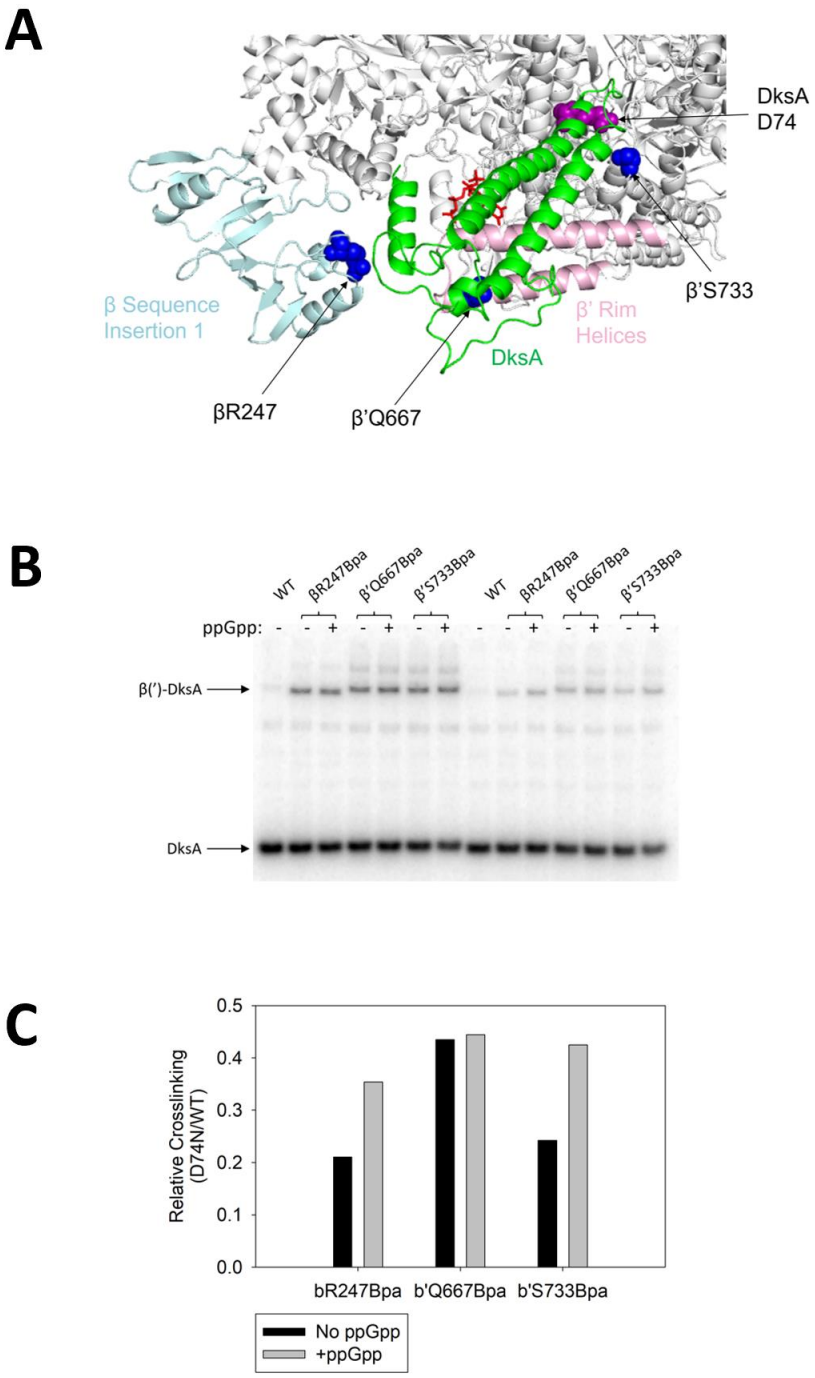
DksA tip residues anchor DksA in the secondary channel

To determine whether D74N affects the positioning of DksA in the secondary channel, we used Bpa crosslinking to monitor DksA placement in the presence and absence of ppGpp. We tested RNAP variants with Bpa substituted at three different sites. βR247Bpa is in β sequence insertion 1, a lineage-specific sequence insertion necessary for activation as well as for binding of DksA to RNAP (Parshin et al., 2015; Chen, chapter 2). β'Q667Bpa is in the β' rim helices, where DksA docks onto RNAP (Lennon et al., 2012). β'S733Bpa is in the secondary channel. All three positions crosslink to DksA in the presence or absence of ppGpp (Figure B.4B). Crosslinking from all three positions to D74N is more than 50% weaker than to WT DksA. Previous iron cleavage data showed that D74N binds with the same affinity as WT DksA to RNAP (Lee et al., 2012). Therefore, a general decrease in crosslinking may suggest greater mobility of DksA, preventing efficient crosslinking.

Looking at specific positions, ppGpp increases the crosslink efficiency from βR247Bpa and β'S733Bpa to D74N more than it affects crosslinking to WT DksA (Figure B.4C). In contrast, crosslinking from β'Q667Bpa is not affected by ppGpp. This shows that in the absence of ppGpp, D74N DksA may have more motion, pivoting around its binding site on the rim helices. Since β'Q667Bpa is on the rim helices where DksA docks, Bpa is able to crosslink to DksA regardless of minor changes in binding position. However, D74N DksA may move further away from βR247Bpa and

Figure B.4. Crosslinking to D74N DksA is decreased compared to WT DksA. (A) Cryo-EM structure showing DksA and ppGpp bound to RNAP. Locations of Bpa substitutions βR247, β'Q667, and β'S733 are shown in blue spheres. DksA is shown in green cartoon, β sequence insertion 1 is shown in cyan, the β' rim helices are shown in pink. D74 DksA is shown in purple spheres. ppGpp at site 2 is shown in red sticks. **(B)** SDS-PAGE gel showing crosslinking from Bpa substitutions to ^{32}P-labeled DksA. Crosslinking results in a super-shift of the ^{32}P-labeled DksA. **(C)** Quantification of crosslinking data, showing relative crosslinking to D74N DksA vs WT DksA from different Bpa substitutions in the presence or absence of ppGpp. n=1.

Figure B.4


A
DksA
D74
β Sequence
Insertion 1
β' Rim
Helices
β'S733
DksA
βR247
β'Q667
B
WT
βR247Bpa
β'Q667Bpa
β'S733Bpa
WT
βR247Bpa
β'Q667Bpa
β'S733Bpa
ppGpp: - - + - + - + - - + - + - +
β(')-DksA
DksA
C
Relative Crosslinking
(D74N/WT)
0.5
0.4
0.3
0.2
0.1
0.0
bR247Bpa
b'Q667Bpa
b'S733Bpa
No ppGpp
+ppGpp

β’S733Bpa, decreasing the crosslinking efficiency. When ppGpp binds, it stabilizes DksA in a more optimal conformation, increasing crosslinking efficiency. Since D74N impairs ppGpp binding, the overall levels of crosslinking never reach WT levels. The consequence of “wobbly binding” is that DksA is unable to properly interact with RNAP to activate transcription.

Discussion

We show here that DksA tip mutants D74N and A76T increase strand opening at positively regulated promoter P*iraP* but are still defective for activation. This occurs because tip mutants are impaired for binding ppGpp at site 2, resulting in dynamic binding of DksA in the secondary channel and suboptimal contacts between DksA and RNAP. These data suggest that P*iraP* is still rate-limited for a step after strand opening that D74N is unable to increase and that this step is ppGpp-dependent. Preliminary studies suggest that promoter escape is not a likely explanation.

There are alternative models that need to be tested. DksA may be required for proper template strand placement for positively regulated promoters. $KMnO_4$ footprinting only detects strand opening, but not DNA positioning within the cleft. Bpa crosslinking to map the path of template strand DNA in the presence and absence of DksA and ppGpp would address this issue. Additionally, DksA has been shown to interact with the trigger loop. Substitutions in the DksA tip may prevent this interaction from occurring properly. Finally, DksA must bind to RNAP and then dissociate at the proper time to allow RNAP promoter escape and elongation. Effects of tip substitutions on DksA dwell time may have important effects on DksA function and activation.

What is the contribution of strand opening to activation by DksA and ppGpp? The difference in activation by D74N on linear vs supercoiled templates suggest that an increase in strand opening will only result in a 2-fold effect, compared to the 10-fold effect of WT DksA. This difference could also be attributed to the salt differences in the two experiments, and a more thorough study modulating supercoiling state through topoisomerases may properly address this question.

References

Artsimovitch, I., Svetlov, V., Murakami, K.S., and Landick, R. (2003). Co-overexpression of Escherichia coli RNA polymerase subunits allows isolation and analysis of mutant enzymes lacking lineage-specific sequence insertions. *J. Biol. Chem.* **278**:12344-12355.

Lee, J.H., Lennon, C.W., Ross, W., and Gourse, R.L. (2012). Role of the coiled-coil tip of Escherichia coli DksA in promoter control. *J. Mol. Biol.* **416**:503-517.

Lennon, C.W., Ross, W., Martin-Tumasz, S., Toulokhonov, I., Vrentas, C.E., Rutherford, S.T., Lee, J.H., Butcher, S.E., and Gourse, R.L. (2012). Direct interactions between the coiled-coil tip of DksA and the trigger loop of RNA polymerase mediate transcriptional regulation. *Genes Dev.* **26**:2634-2646.

Molodtsov, V., Sineva, E., Zhang, L., Huang, X., Cashel, M., Ades, S.E., and Murakami, K.S. (2018). Allosteric Effector ppGpp Potentiates the Inhibition of Transcript Initiation by DksA. *Mol. Cell* **69**:828-839.e5.

Parshin, A., Shiver, A.L., Lee, J., Ozerova, M., Schneidman-Duhovny, D., Gross, C.A., and Borukhov, S. (2015). DksA regulates RNA polymerase in Escherichia coli through a network of interactions in the secondary channel that includes Sequence Insertion 1. *Proc. Natl. Acad. Sci. USA.* **112**:E6862-E6871.

Paul, B.J., Barker, M.M., Ross, W., Schneider, D.A., Webb, C., Foster, J.W., and Gourse, R.L. (2004). DksA: a critical component of the transcription initiation machinery that potentiates the regulation of rRNA promoters by ppGpp and the initiating NTP. *Cell* **118**:311-322.

Perederina, A., Svetlov, V., Vassylyeva, M.N., Tahirov, T.H., Yokoyama, S., Artsimovitch, I., and Vassylyev, D.G. (2004). Regulation through the secondary channel—structural framework for ppGpp-DksA synergism during transcription. *Cell* **118**:297-309.

Ross, W., Vrentas, C.E., Sanchez-Vazquez, P., Gaal, T., and Gourse, R.L. (2013). The magic spot: a ppGpp binding site on E. coli RNA polymerase responsible for regulation of transcription initiation. *Mol. Cell* **50**:420-429.

Ross, W., Sanchez-Vazquez, P., Chen, A.Y., Lee, J.H., Burgos, H.L., and Gourse, R.L. (2016). ppGpp Binding to a Site at the RNAP-DksA Interface Accounts for Its Dramatic Effects on Transcription Initiation during the Stringent Response. *Mol. Cell* **62**:811-823.

Ryu, Y. and Schultz, P.G. (2006). Efficient incorporation of unnatural amino acids into proteins in Escherichia coli. *Nat. Methods* **3**:263-265.

Svetlov, V. and Artsimovitch, I. (2015). Purification of bacterial RNA polymerase: tools and protocols. *Methods Mol. Biol.* **1276**:13-29.

Supplemental Materials

Table B.S1. Strains and Plasmids

Strain	**Description**	**Source**
RLG4677	BL21(DE3)	Novagen
RLG12115	BL21(DE3) *rpoZ::kan*	Ross et al., 2013
RLG7075	BL21(DE3) *dksA::tet*	Paul et al., 2004
Plasmid	**Description**	**Source**
pIA299	pT7 αββ'-H6 overexpression vector	Artsimovitch et al., 2003
pIA900	pT7 αββ'-H10ω overexpression vector	Svetlov and Artsimovitch, 2015
pRLG9997	pIA900 *rpoC* S733TAG for Bpa incorporation	This work
pRLG9998	pIA900 *rpoC* Q667TAG for Bpa incorporation	This work
pRLG12214	pIA900 *rpoB* R247TAG for Bpa incorporation	This work
pSupT/BpF	Expresses Bpa synthetase/tRNA	Ryu and Schultz, 2006
pRLG13300	pET33 His10 HMK DksA	Ross et al., 2016
pRLG9059	pET33 His6 HMK DksA-D74N	Lee et al., 2012
pRLG9060	pET33 His6 HMK DksA-A76T	Lee et al., 2012
pRLG9988	pRLG1507 with *iraP* (-64 to +50 with mutations)	Ch. 2
pRLG11350	pRLG770 with *iraP* (-207 to +21)	Ross et al., 2016

Table B.S2. Oligonucleotides

Primer	**Sequence**	**Purpose**
5978	CCGGAACGCCTGTAGGGTGAAACCGCATC	*rpoB* R247TAG
7849	GCGTGGTTAGGCGGCACAGATTCGTCAGC	*rpoC* S733TAG
7851	CTGAAATTCAGGAGTAGTTCCAGTCTG	*rpoC* Q667TAG