Mechanistic studies of DksA mediated transcriptional regulation

By

Christopher William Lennon

A dissertation submitted in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

(Cellular and Molecular Biology)

at the

UNIVERSITY OF WISCONSIN-MADISON

2013

Date of final oral examination: 03/13/13

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

Richard L. Gourse, Professor, Bacteriology
Samuel E. Butcher, Professor, Biochemistry
Katrina T. Forest, Professor, Bacteriology
Robert C. Landick, Professor, Biochemistry
Karen M. Wassarman, Associate Professor, Bacteriology

Table of Contents

Chapter 1

Introduction

RNA Polymerase

Transcription is a process required for life in all free-living organisms. Because the degree and timing of gene expression is crucial to cell survival, transcription is an exquisitely regulated process. Transcription progresses through three distinct steps (initiation, elongation, and termination) and each of these steps is subject to extensive regulation. This chapter will focus primarily on transcription initiation and how the transcription factor DksA inhibits gene expression in *E. coli*.

All cellular life forms employ a multi-subunit DNA-dependent RNA Polymerase (RNAP) to synthesize RNA. Core RNAP structure is highly conserved in all domains of life, with the largest two subunits (β and β' in bacteria) forming a distinctive "crab-claw" that contains the enzyme active site cleft (Hirata and Murakami 2009). Bacterial RNAP core enzyme consists of five subunits ($\alpha_2\beta\beta'\omega$, ~400 kDa).

Bacterial sigma factors

Although core enzyme is transcriptionally competent at DNA ends and nicks, it requires binding of a σ factor (forming the RNAP holoenzyme or "Eσ") in order to specifically recognize promoter DNA and begin transcription initiation (Haugen et al. 2008). Seven σ factors are employed in *E. coli* to help regulate promoter binding and downstream gene expression under different environmental conditions. These include σ^{70} (primary "housekeeping"), σ^{S} (starvation/stationary phase), σ^{E} (extracellular stress), σ^{H} (heat shock), σ^{F} (flagellar synthesis), σ^{FecI} (iron transport) and σ^{N} (nitrogen limitation), with each factor targeting RNAP to

distinct promoter sequences upon holoenzyme formation. All *E. coli* σ factors except σ^N have a similar domain organization. Unlike σ^{70}, σ^N-dependent promoters require specific factors and external energy (ATP) during transcription initiation. σ^{70} has also been shown to influence certain aspects of transcription elongation, such as pausing, in addition to initiation (Mooney et al. 2005). For this thesis, I focus on the role of σ^{70} during transcription initiation.

σ^{70} has four conserved domains (1-4) connected by flexible linkers that help each domain recognize specific promoter DNA sequences. Group 1 σ factors (σ^{70} and σ^S in *E. coli*) also have a highly acidic amino-terminal extension in domain 1 (region 1.1) that inhibits promoter binding in the absence of core enzyme (Dombroski et al. 1993). Upon core RNAP binding, σ^{70} region 1.1 is thought to occupy the RNAP primary channel, acting as mimic of promoter DNA.

Promoter DNA recognition by Eσ^{70}

σ^{70} makes numerous contacts to specific promoter DNA sequences upon initial binding to promoter DNA and during its progression through the other steps in transcription initiation (Figure 1.1 summarizes these contacts on the consensus Eσ^{70} promoter) .The most highly conserved and arguably most important contact is formed by σ^{70} region 2.3-2.4, which binds directly to the -10 hexamer (Figure 1.1; promoter sequences are numbered relative to the transcription start site, +1). Recently, new structures of this interaction from *Thermus thermophilus* have given us an atomic level understanding of the role of some of these interactions during promoter recognition and dsDNA melting, as well as their likely

importance for the stability of initiation complexes (Feklistov and Darst 2011, Zhang et al. 2012). In addition, many other interactions have been characterized in detail between the -35 hexamer and σ^{70} region 4.2, extended -10 and σ^{70} region 3.0, and the discriminator (sequence between -10 hexamer and +1) and σ^{70} region 1.2 (Figure 1.1; Haugen et al. 2008).

Core RNAP itself can also aid in promoter-specific DNA recognition via the α subunits (UP elements, Ross et al. 1993). Although the two α subunits in RNAP are the same in sequence, one α subunit primarily associates with β, while the other primarily associates with β'. The α subunit is further divided into an amino-terminal domain (α-NTD), which is important in core enzyme assembly, and a carboxy-terminal domain (α-CTD), which can bind directly to DNA. The α-NTDs and α-CTDs are physically separated by an approximately 15 residue flexible linker, allowing each of the α-CTDs to move relatively freely with respect to the rest of RNAP. One or both α-CTDs can bind DNA, greatly enhancing holoenzyme recruitment to promoters. α-CTDs can recognize specific sequences directly upstream of the -35 hexamer (Ross et al. 1993) or associate non-specifically with backbone DNA further upstream of the -35 hexamer (Ross and Gourse, 2005). α-CTDs can also establish protein-protein interactions with transcription factors that bind to DNA upstream of the -35 hexamer to activate transcription (Browning and Busby 2004).

The precise DNA sequences of these elements are crucial for determining how strongly $E\sigma^{70}$ binds and transcribes a given promoter. Not all elements are identical or present at every promoter, and the precise sequences are critical not

only for determining the strength of binding in the initial complex but also for regulation during other steps in initiation. Sequence variation among promoters helps the cell achieve an unequal chromosomal distribution of RNAP that can match specific growth needs. Interestingly, a "full-consensus" promoter for $E\sigma^{70}$ ("perfect" recognition in each promoter element) is not present in wild-type *E. coli*, likely because tighter binding of $E\sigma^{70}$ to promoter DNA interferes with "escape" and progression into transcription elongation.

Mechanism of transcription initiation in bacteria

Transcription initiation is a stepwise process, characterized by distinct conformations of the DNA-RNAP complex that occur prior to elongation (Haugen et al. 2008, Saecker et al. 2011). Proceeding through these steps does not require the addition of external energy (at non σ^N-dependent promoters), but rather isomerization of the RNAP-DNA complex is needed to unwind and position the template DNA for RNA synthesis. To begin initiation, RNAP holoenzyme (R) binds to promoter DNA (P) to form an initial closed complex (RP_C, Figure 1.2). In RP_C, the DNA is double-stranded and outside the main DNA channel (primary channel). RP_C next passes through at least two intermediate complexes that are typically unstable and short-lived (denoted here as a composite term, RP_I, Figure 1.2) before forming a stable open complex (RP_O, Figure 1.2). In RP_O, the DNA is bent into the primary channel, and ~12 bp are unwound to form the "transcription bubble". In the transcription bubble, the non-template DNA strand of the -10 hexamer is bound to σ^{70} domain 2, and the template DNA strand is placed in the

RNAP active site to await NTP addition. DNAse I footprinting shows that RP_C, RP_I, and RP_O protect upstream DNA to the same extent (to about -60), but show clear differences in downstream DNA protection. Protection extends downstream stepwise from +1 in RP_C, to +12 RP_I, and finally to +20 when DNA forms the transcription bubble in RP_O. Unfortunately, because RP_C and RP_I are typically short-lived intermediates, the precise molecular nature of these complexes has been difficult to determine. However, recent high-resolution structures of "RP_O", (Feklistov and Darst 2011; Zhang et al. 2012) have significantly improved our understanding of initiation.

Although RNAP and DNA are thought to go through the same conformational changes at all promoters, different sequences determine the rates of complex formation and decay for individual steps in initiation, resulting in differential occupancy of specific intermediates at different promoters. For example, under the same conditions and in the absence NTPs at equilibrium, RNAP will primarily occupy RP_O at λP_R and RP_I at *rrnB* P1.

Transition from RP_O to a transcription elongation complex (TEC) requires that promoter contacts to $E\sigma^{70}$ be broken in a process known as promoter escape. In order to generate the energy required to overcome a thermodynamic barrier, RNAP forms a "scrunched" complex. During scrunching, downstream DNA is pulled into RNAP while the RNA strand is growing, which provides a high-energy state from which $E\sigma^{70}$ can escape contacts with promoter DNA to elongate. Alternatively, the scrunched complex may fail to escape and initiate

"abortively", releasing short RNA transcripts and "decaying" back to RP_O (Revyakin et al. 2006, Kapanidis et al. 2006).

rRNA initiation in *E. coli*

Ribosome production is precisely tuned to match the translation requirements of the cell through a number of complex regulatory mechanisms (Paul et al. 2004b). This makes energetic sense, as excess ribosome production would represent a substantial waste of cellular resources when the substrates of translation are scarce. Transcription from *E. coli* rRNA (*rrn*) promoters can account for more than half of all cellular RNAP activity at high cell growth rates (Schneider et al. 2003). The remarkable strength of *rrn* P1 promoters is primarily due to DNA sequences upstream of the core promoter region. Each *rrn* promoter has an UP element and 3-5 activator (Fis) binding sites, which can increase promoter activity ~300-fold (Hirvonen et al. 2001). RP_O at *rrn* promoters is very unstable and rapidly decays to RP_I in the absence of NTPs (rRNA promoters can progress to a stable RP_O-like state called RP_{AC} in the presence of the first two NTPs of the transcript) (Gourse,1988, Barker et al. 2001, Rutherford et al. 2009). This contrasts with the kinetics of initiation at most *E. coli* promoters, which form a very short-lived RP_I and a long-lived RP_O (Haugen et al. 2008). With the exception of a consensus -10 hexamer, the *rrn* P1 core promoter sequence is not ideal for $E\sigma^{70}$ recognition, with a nonconsensus -35 hexamer, a shorter than consensus spacer region (sequence between the -10 and -35 hexamers), no extended -10 element, and a suboptimal discriminator region. These

characteristics of *rrn* core promoters are essential, as mutations that stabilize RP_O through improved contacts with σ^{70} prevent *E. coli* from properly regulating rRNA synthesis in response to changing environmental conditions (Barker and Gourse, 2001, Haugen et al. 2006).

rRNA initiation control by small molecules in *E. coli*

rRNA transcription initiation is regulated by two small molecule effectors in *E. coli*, the transcripts initiating NTP (iNTP) and guanosine 5'-diphosphate 3'-diphosphate (ppGpp) (Murray et al. 2003). In contrast to most other promoters tested, *E. coli rrn* P1 promoters require especially high iNTP concentrations in order to maximize transcription *in vitro* (Gaal et al. 1997). Because ATP and GTP serve as vital cellular energy sources and are consumed during translation, and because 6 of the 7 *rrn* P1 promoters in *E. coli* initiate transcription with ATP and the seventh with GTP, it was reasoned that rRNA promoter-RNAP complexes "sense" NTPs directly to maximize transcription only when energy is abundant. To test this hypothesis, the relative cellular concentrations of ATP and GTP were manipulated via mutations in the purine biosynthetic pathway and shown to directly regulate rRNA promoter activity *in vivo* (Schneider et al. 2002, Murray et al. 2003).

In addition to the concentration of iNTP, the tightly controlled production of the unusual nucleotide ppGpp is required for proper rRNA control (Murray et al. 2003). ppGpp acts as a stress signaling "alarmone" in virtually all bacteria, and in *E. coli* ppGpp is rapidly synthesized when rapidly growing cells are suddenly

starved for amino acids. Synthesis occurs under these conditions from the activity of the ribosome-associated protein RelA, which increases the cellular concentration >10-20 fold after only 5 minutes. ppGpp binds to *E. coli* RNAP and this direct interaction is required for the drastic inhibition of rRNA production *in vivo* (Vrentas et al. 2005, Ross et al. in press). iNTP and ppGpp concentrations vary dramatically depending on nutrient availability,resulting in dynamic regulation of rRNA production during different phases of growth (Murray et al. 2003).

DksA is essential for rRNA regulation in *E. coli*

Within minutes of starvation for amino acids, new transcription from *rrnB* P1 is reduced >20-fold concordant with a dramatic increase in cellular ppGpp levels (Murray et al. 2003). However, because the inhibitory effect of ppGpp at *rrnB* P1 *in vitro* was fairly modest in comparison with the *in vivo* observations, it was correctly suspected that an important regulator of rRNA synthesis (which turned out to be the small protein DksA, ~17.5 kDa) had yet to be identified. When *dksA* is deleted from the *E. coli* chromosome, rRNA transcription is overproduced ~2-4-fold during log phase (Paul et al. 2004a). Δ*dksA* cells cannot stringently control rRNA synthesis in response to amino acid starvation, and in general the cell no longer regulates RNA to match the growth conditions (Paul et al. 2004a). These *in vivo* effects are explained by the fact that DksA directly binds to RNAP and amplifies the concentration-dependent effects of ppGpp and the iNTP on rRNA promoters *in vitro* (Paul et al. 2004a). Consistent with a concerted relationship

between DksA, ppGpp, and iNTP concentrations in the regulation of rRNA synthesis, mutations in the *rrnB* P1 discriminator region that stabilize RP_O are no longer inhibited by DksA, ppGpp, or low iNTP concentrations (Paul et al. 2004a, Barker et al. 2001, Schneider et al. 2002).

DksA is a non-classical transcription factor

The amount of transcript produced for a given gene is often regulated by transcription factors. To achieve promoter-specific effects, most transcription factors bind to precise DNA sequences near or within the promoter of the gene that is ultimately regulated (Browning and Busby 2004). In contrast, DksA is able to achieve promoter-specific regulation without a direct interaction with DNA. Rather, DksA directly binds via the secondary channel of RNAP, altering DNA contacts made during initiation (Perederina et al. 2004, Paul et al. 2004a, Haugen et al. 2008).

X-ray crystallography revealed that DksA has an overall structure similar to the GreA/B transcription elongation factors, having both a globular domain and a coiled-coil domain (Perederina et al. 2004, Figure 1.3). Generally speaking, Gre factors place their conserved acidic coiled-coil tip near the RNAP active site to rescue backtracked elongation complexes, reduce abortive initiation, and help remove misincorporated NTPs from the transcript (Opalka et al. 2003). The roles of DksA and Gre factors in the cell are distinct but partially overlapping. DksA and Gre factors share structural similarities, but little DNA sequence identity, and

the precise identity of the coiled-coil tip of each factor is absolutely vital for their specific functions (Rutherford et al. 2007, Lee et al. 2012).

DksA and ppGpp regulon

DksA and ppGpp inhibit much more than rRNA production in *E. coli*. Genome-wide studies have shown that DksA and ppGpp significantly inhibit, or even activate, the expression of ~20 % of all *E. coli* genes (Traxler et al. 2008, Durfee et al. 2008). Regulation by DksA and ppGpp can occur directly or indirectly. For our purposes, direct refers to cases in which DksA and ppGpp bind to RNAP complexes at a specific promoter to influence transcription, as in the case of *rrn* P1 promoters in *E. coli*. In contrast, we use indirect regulation to refer to secondary DksA and ppGpp effects, for example regulation of one gene product affecting the production of others. Biochemical analysis has shown that DksA and ppGpp directly inhibit transcription of promoters for tRNAs (Rutherford et al. unpublished.), many r-proteins (Lemke et al. 2011), the transcription activator Fis (Mallik et al. 2006), DksA itself (Chandrangsu et al. 2011) and others. DksA and ppGpp can also activate transcript production directly *in vitro* from some other promoters, including genes for amino acid biosynthesis and transport (Paul et al. 2005), virulence factors (Nakanishi et al. 2006), Hfq (Sharma and Payne 2006), and σ^E-dependent transcription (Constanzo et al. 2008). Thus, depending on the promoter sequence, DksA and ppGpp can have opposite effects on gene expression.

Mechanism of DksA mediated negative regulation

Recent biochemical studies from our laboratory have elucidated many of the molecular details resulting in direct negative regulation by DksA (Rutherford et al. 2009, Lennon et al. 2009, Lee et al. 2012, Lennon et al. 2012). DksA and ppGpp modify the kinetic properties of RNAP-DNA initiation complexes, favoring RNAP dissociation from promoter DNA and a decrease in the ½-life of open complexes at all promoters tested (in the presence of a competitor that prevents reformation of RP_O upon decay). However, transcriptional output is only inhibited from promoters that form intrinsically short-lived RP_O complexes (Paul et al. 2004, Paul et al. 2005, Rutherford et al. 2009). DNAse I footprinting revealed that DksA does not change the footprint of RP_O or RP_C but inhibits occupancy of RP_I and promotes formation of RP_C by disfavoring contacts between $E\sigma^{70}$ and downstream DNA.

RP_I is occupied at *rrnB* P1 under conditions where most other promoters favor RP_O, and DksA and ppGpp inhibit transcription from *rrnB* P1 by blocking RP_I occupancy. Most promoters form long-lived RP_O complexes, so the TEC forms when NTPs are present before DksA and ppGpp-mediated dissociation of RP_O significantly alters transcriptional output. In general, promoters where RP_O decays quickly to RP_I are inhibited by DksA and ppGpp whereas promoters that form long-lived RP_O complexes are not. This mechanism thereby results in selective promoter regulation (Barker et al. 2001, Paul et al. 2004a, Paul et al. 2005, Rutherford et al. 2009).

At a subset of promoters with long-lived RP_O complexes, DksA and ppGpp actually increase transcriptional output (direct activation). However, the precise mechanism of direct activation by DksA and ppGpp is not well understood and is currently being investigated.

Deleting *dksA* or the ability to synthesize ppGpp (Δ*relA/spoT*) in *E. coli* results in loss of survival in the absence of amino acids (Paul et al. 2004a, Paul et al. 2005, Rutherford et al. 2009). Suppressors that permit growth under this condition are often located in the genes for the β and β' RNAP subunits, mimicking the effects of DksA and ppGpp by reducing RP_O lifetime (Rutherford et al. 2009). The positions of suppressors in β and β' localized primarily to the trigger loop (TL), bridge helix (BH) and "switch regions" of RNAP, three extremely important and conserved regions of RNAP. The trigger loop (TL) of RNAP is a catalytically crucial domain that folds from an unstructured loop to a pair of α-helices (TH) during NTP addition (Vassylyev et al. 2007, Zhang et al. 2010). The TH form a three-helix bundle with the BH, a ~60Å long α-helix separating the DNA binding channel from the secondary channel. The BH makes direct contacts to the switch regions that are important for controlling RNAP "clamp" movements that occur during initiation (clamp "open" in holoenzyme and "closed" in RP_O) (Chakraborty et al. 2012).

Unfortunately, traditional techniques have yet to yield a high-resolution structure of DksA bound to RNAP. To improve our understanding of the mechanism of direct negative control by DksA, we constructed an evidence-based structural model of the DksA-RNAP complex using molecular docking and

distance constraints generated by crosslinking (Lennon et al. 2012 and Chapter 3). Our model predicted that the coiled-coil tip of DksA was in close proximity the TL domain of RNAP. Coiled-coil tip residues D74 and A76 are vital for DksA function, but not for DksA binding (Lee et al. 2012). We determined that these residues in DksA can be in very close proximity to RNAP residue β' F935 (within the mobile TL) using disulfide cross-linking (Lennon et al. 2012). This interaction is biologically significant, because like the coiled-coil tip (D74 and A76), β' F935 is also specifically required for DksA function (Lennon et al. 2012). We propose that to mediate transcription, the coiled-coil tip of DksA interacts directly with the TL domain (possibly altering its folded/unfolded state) at the base of the secondary channel, causing an allosteric change in the BH and switches, inducing opening of the RNAP clamp, ultimately leading to a loss downstream DNA contacts made during initiation (Rutherford et al. 2009, Lee et al. 2012, Lennon et al. 2012).

Conservation of the DksA/ppGpp system

Although ppGpp is thought to signal stress in virtually all bacteria, ppGpp doesn't bind directly to RNAP to regulate gene expression in all species. In one example, the model gram-positive bacterium *Bacillus subtilis*, ppGpp is produced and regulates rRNA expression, but not in the same way as in *E. coli* (gram-negative γ-proteobacterium). Rather than modulating *B. subtilis* RNAP directly, ppGpp synthesis acts to reduce the cellular concentration of GTP, which in turn inhibits rRNA production because all *B. subtilis* rRNA promoters use GTP as the iNTP

(Krásný and Gourse 2004). ppGpp not only reduces GTP levels by consumption for its own synthesis, but it also directly inhibits two GTP synthetases in *B. subtilis* (Kriell et al., 2012).

The ppGpp binding site on *E. coli* RNAP responsible for direct negative regulation was recently determined in two separate studies (Ross et al., Zuo et al., both manuscripts in press), and a *dksA*-like gene and most RNAP residues required for ppGpp binding are well-conserved among proteobacteria. In order to assess whether DksA and ppGpp work through a direct mechanism in α-proteobacteria as in *E. coli,* we tested whether RSP2654 (DksA-like) and ppGpp from the *Rhodobacter sphaeroides* reduced the ½-life of RP_O. We determined that purified RSP2654 and ppGpp, both alone and in concert, reduce the ½-life of two stable promoter complexes containing *R. sphaeroides* $E\sigma^{93}$ (Lennon, Lemmer et al. in preparation and Chapter 4), strongly suggesting that DksA and ppGpp work through very similar mechanisms to influence the initiation complexes formed by both γ- and α-proteobacterial RNAPs. (RSP2654 also serves as a "mega-mutant" (42% amino acid conservation), providing important clues as to amino acid requirements for function of *E. coli* DksA.) RSP2654 complements Δ*dksA E. coli* cells for growth on medium lacking amino acids and specifically represses *rrnB* P1 *in vivo*. *In vitro* using *E. coli* RNAP/promoters, we show that RSP2654 binds to RNAP, specifically represses *rrnB* P1 (ppGpp amplifies this effect), activates the amino acid biosynthesis promoter *hisG* (ppGpp is required for this effect), and requires residues corresponding to those in the DksA coiled-coil tip for function (Lennon, Lemmer et al. in preparation).

Other cellular roles of DksA

Most work to date has examined the effect of DksA on transcription initiation. However, DksA has also been shown to influence transcription elongation and termination *in vitro* (Perederina et al. 2004, Furman et al. 2012, Furman et al. 2013). In one study, WT DksA was shown to have very small inhibitory effects on elongation *in vitro*, even in the presence of ppGpp (Furman et al. 2012). To amplify this effect, the authors employed a tighter binding DksA variant (N88I, Blankschien et al. 2009) and concluded that DksA could have numerous effects on elongation. However, in another study published the following year (Furman et al. 2013), the same group found that DksA is restricted from binding (and thus influencing) most elongation complexes because it is prevented from accessing these complexes by the lineage-specific i6 domain of β'. Further, the authors showed that without exception (based on all sequenced bacterial genomes), organisms with a DksA homologue also have an i6 insertion in the β' subunit (Furman et al. 2013). Therefore, understanding the role of DksA on elongation will require further *in vitro* and *in vivo* studies.

Whereas DksA only influences transcriptional output from specific promoters, it has the potential to bind all RNAP molecules in the cell. Therefore, DksA could in theory influence RNAP-DNA complexes after the stage of initiation under certain conditions. DksA also competes with other secondary channel factors (i.e. Gre) for binding to RNAP in the cell, potentially decreasing their specific activities (Rutherford et al. 2007). Quite interestingly, DksA also plays an

important role in another critical cellular process, DNA replication. During amino acid starvation, DksA promotes replication fork progression in *E. coli* by reducing conflicts between the transcription and replication machinery (Tehranchi et al. 2010).

Thesis outline

In chapter 2, we present a quantitative assay for measuring DksA binding affinity, which has served as an important tool needed to understand the mechanism of DksA action as well as demonstrating that DksA and DNA compete for occupancy of open promoter complexes (Lennon et al. 2009). In chapter 3, we present a model of the DksA-RNAP complex based on cross-linking, as well as determine the precise contacts required for RNAP regulation by DksA (Lennon et al. 2012). Finally, in chapter 4 we genetically and biochemically examine the DksA and ppGpp system from the α-proteobacterium *R. sphaeroides*. Chapter 5 summarizes the major findings of my work and outlines possible future directions of investigation. Three appendices are also included. Appendices A and B describe specific contributions I made to publications on which I am a co-author (Blankschien et al. 2009, Lee et al. 2012). Appendix C provides preliminary results investigating the physical location of E. coli σ^{70} region 1.1 during transcription initiation via BPA crosslinking, a project that is being continued in the lab by doctoral student Albert Chen.

References

Barker M, Gaal T, Josaitis C, Gourse R. 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.

Browning DF, Busby SJ. 2004. The regulation of bacterial transcription initiation. *Nat Rev Microbiol* **2**: 57-65.

Blankschien M, Lee J, Grace E, Lennon C, Halliday J, Ross W, Gourse R, Herman C. 2009. Super DksAs: Substitutions In DksA enhancing its effects on transcription initiation. *EMBO J* **28:** 1720–1731.

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

Chakraborty A, Wang D, Ebright Y, Korlann Y, Kortkhonjia E, Kim T, Chowdhury S, Wigneshweraraj S, Irschik H, Jansen R, et al. 2012. Opening and closing of the bacterial RNA polymerase clamp. *Science* **337:** 591-595.

Constanzo A, Nicoloff H, Barchinger SE, Banta AB, Gourse RL, Ades SE. 2008. ppGpp and DksA likely regulate the activity of the extracytoplasmic stress factor sigma(E) in *Escherichia coli* by both direct and indirect mechanisms. *Mol Microbiol* **67**: 619-632.

Dombroski AJ, Walter WA, Gross. CA. 1993. Amino-terminal amino acids modulate sigma-factor DNA-binding activity. *Genes Dev* **12**: 2446-2455.

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

Feklistov A, Darst SA. 2011. Structural basis for promoter -10 element recognition by bacterial RNA polymerase sigma subunit. *Cell* **147**: 1257-1269.

Furman R, Sevostyanova A, Artsimovitch I. 2012. Transcription initiation factor DksA has diverse effects on RNA chain elongation. *Nucleic Acids Res* **40:** 3392-3402.

Furman R, Tsodikov OV, Wolf YI, Artsimovitch I. 2013. An insertion in the catalytic trigger loop gates the secondary channel of RNA polymerase. *J Mol Biol* **425:** 82-93.

Gaal T, Bartlett MS, Ross W, Turnbough Jr. CL, Gourse RL. 1997. Transcription regulation by initiating NTP concentrations: rRNA synthesis in bacteria. *Science* **278**: 2092-2097.

Gourse RL. 1988. Visualization and quantitative analysis of complex formation between *E. coli* RNA polymerase and an rRNA promoter *in vitro*. *Nucleic Acids Res* **16**: 9789-9809

Haugen SP, Berkmen MB, Ross W, Gaal T, Ward C, Gourse RL. 2006. rRNA promoter regulation by nonoptimal binding of sigma region 1.2: an additional recognition element for RNA polymerase. *Cell* **125**: 1069-1082.

Haugen SP, Ross W, Gourse RL. 2008. Advances in bacterial promoter recognition and its control by factors that do not bind DNA. *Nat Rev Microbiol* **6**: 507- 519

Hirata A, Murakami KS. 2009. Archael RNA polymerase Curr Opin Struct Biol **19**: 724-731.

Hirvonen CA, Ross W, Wozniak CE, Marasco E, Anthony JR, Aiyar SE, Newburn VH, Gourse RL. 2001. Contributions of UP elements and the transcription factor FIS to expression from the seven *rrn* P1 promoters in *Escherichia coli*. *J Bacteriol* **183**: 6305-6314

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

Kapanidis AN, Margeat E, Ho SO, Kortkhonjia E, Weiss S, Ebright RH. 2006. Initial transcription by RNA polymerase proceeds through a DNA-scrunching mechanism. *Science* **314**: 1144-1147.

Kriel A, Bittner AN, Kim SH, Liu K, Tehranchi AK, Zou WY, Rendon S, Chen Rui, Tu BP, Wang JD. 2012. Direct regulation of GTP homeostasis by (p)ppGpp: a critical component of viability and stress resistance. *Mol Cell* **48**: 231-241.

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

Lemke J, Durfee T, Gourse R. 2009. DksA and ppGpp directly regulate transcription of the *Escherichia coli* flagellar cascade. *Mol Microbiol* **74:** 1368-1379.

Lemke J, Sanchez-Vazquez P, Burgos H, Hedberg G, Ross W, Gourse R. 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, Gaal T, Ross W, Gourse R. 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 CW, Ross W, Martin-Tumasz S, Toulokhonov I, Vrentas CE, Rutherford ST, Lee JH, Butcher SE, Gourse RL. 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.

Mallik P, Paul BJ, Rutherford ST, Gourse RL, Osuna R. 2006. DksA is required for growth phase-dependent regulation, growth rate-dependent control, and stringent control of *fis* expression in *Escherichia coli. J Bacteriol* **188**: 5775-82.

Mooney RA, Darst SA, Landick R. 2005. Sigma and RNA polymerase: an on-again, off-again relationship? *Mol Cell* **20:** 335-345.

Murray H, Schneider D, Gourse, R. 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, 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.

Opalka N, Chlenov M, Chacon P, Rice W, Wriggers W, Darst, S. 2003. Structure and function of the transcription elongation factor GreB bound to bacterial RNA polymerase. *Cell* **114:** 335-345.

Paul B, Barker M, Ross W, Schneider D, Webb C, Foster J, Gourse R. 2004a. 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, Ross W, Gourse R. 2004b. rRNA transcription in *Escherichia coli. Annu Rev Genet* **38**: 749-770

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

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

Revyakin A, Liu C, Ebright RH, Strick TR. 2006. Abortive initiation and productive initiation by RNA polymerase involve DNA scrunching. *Science* **314**: 1139-1143.

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

Ross W, Gourse RL. 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 CE, Sanchez-Vazquez P, Gaal T, Gourse RL. 2013. The magic spot: a ppGpp binding site on *E. coli* RNA polymerase responsible for regulation of transcription initiation. *Mol Cell* in press.

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

Saecker R, Record M, deHaseth P. 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.

Schneider DA, Gaal T, Gourse RL. 2002. NTP-sensing by rRNA promoters in Escherichia coli is direct. *Proc Natl Acad Sci USA* **99**: 8602-8607.

Schneider DA, Ross W, Gourse RL. 2003. Control of rRNA expression in *Escherichia coli. Curr Opin Microbiol* **6**: 151-156.

Sharma AK, Payne SM. 2006. Induction of expression of *hfq* by DksA is essential for *Shigella flexneri* virulence. *Mol Microbiol* **62**: 469-479.

Tehranchi A, Blankschien M, Zhang Y, Halliday J, Srivatsan A, Peng J, Herman C, Wang J. 2010. The transcription factor DksA prevents conflicts between DNA replication and transcription machinery. *Cell* **141:** 595-605.

Traxler MF, Summers SM, Nguyen HT, Zacharia VM, Hightower GA, Smith JT, Conway T. 2008. The global, ppGpp-mediated stringent response to amino acid starvation in *Escherichia coli*. *Mol Microbiol* **68:** 1128-1148.

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

Vrentas C, Gaal T, Ross W, Ebright R, Gourse R. 2005. Response of RNA polymerase to ppGpp: requirement of the Ω subunit and relief of this requirement by DksA. *Genes Dev* **19:** 2378-2387.

Zhang J, Palangat M, 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 MX, Arnold E, Ebright RE. 2012. Structual basis of transcription initiation. *Science* **338**: 1076-1080.

Zuo Y, Wang Y, Steitz TA. 2013. The structure of magic spot bound to RNA polymerase suggests how it is regulated by ppGpp. *Mol Cell* in press.

Figures and legends

Figure 1.1. Eσ^{70} consensus core promoter non-template DNA strand and contacts made by specific σ^{70} regions. The DNA sequence is shown with optimal spacing between of the -10 and -35 hexamers relative to the transcription start site (+1).

Figure 1.1

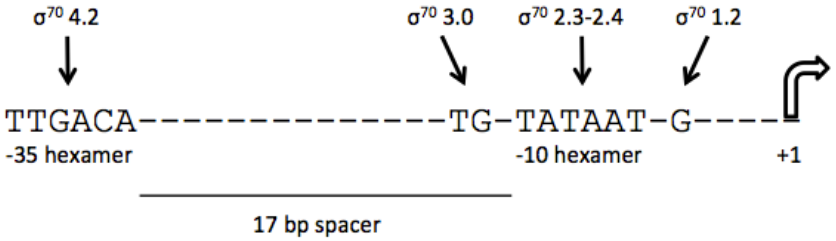
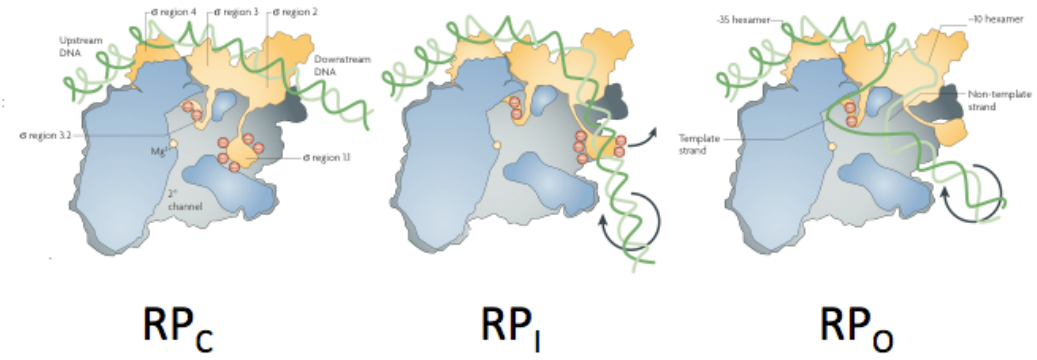

σ70 4.2
σ70 3.0
σ70 2.3-2.4
σ70 1.2
TTGACA----------TG-TATAAT-G----
-35 hexamer
-10 hexamer
+1
17 bp spacer

Figure 1.2. Step-wise progression of transcription initiation from RP_C, RP_I, and RP_O prior to elongation showing the predicted movements of the template DNA (dark green), non-template DNA (light green) relative to RNAP holoenzyme (σ^{70} orange, core blue). Figure modified from Haugen et al. 2008.

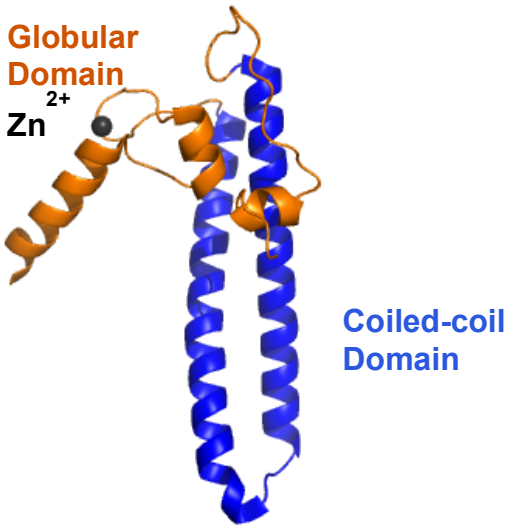
Figure 1.2



RP_C RP_I RP_O

Figure 1.3. Crystal structure of *E. coli* DksA (PDB: 1TJL; Perederina et al. 2004). The globular domain is colored orange and the coiled-coil domain colored blue. The globular domain contains a zinc ion (black).

Figure 1.3



Chapter 2

Escherichia coli DksA Binds to Free RNA Polymerase with Higher Affinity than to RNA Polymerase in an Open Complex

This chapter has been published in the Journal of Bacteriology (Christopher W. Lennon, Tamas Gaal, Wilma Ross and Richard L. Gourse. 2009. J Bacteriol 191: 5854-5858). I performed all experiments in the manuscript.

Abstract

The transcription factor DksA binds in the secondary channel of RNA polymerase (RNAP) and alters transcriptional output without interacting with DNA. Here we present a quantitative assay for measuring DksA binding affinity and illustrate its utility by determining the relative affinities of DksA for three different forms of RNAP. Whereas the apparent affinities of DksA for RNAP core and holoenzyme are the same, the apparent affinity of DksA for RNAP decreases almost 10-fold in an open complex. These results suggest that the conformation of RNAP present in an open complex is not optimal for DksA binding and that DNA directly or indirectly alters the interface between the two proteins.

Introduction

Transcription of rRNA accounts for the majority of RNA synthesis in *Escherichia coli* at high growth rates, but rRNA transcription decreases under less favorable nutritional conditions, primarily from regulation by two small molecules, ppGpp and the initial NTP in the transcript (Murray et al. 2003, Paul et al. 2004b). Current models suggest that the 151-amino-acid protein DksA binds directly to RNA polymerase (RNAP) and modifies the kinetic properties of the promoter complex, sensitizing it to changes in the concentrations of ppGpp and the initial NTP in the transcript (Paul et al. 2004a). Consistent with this model, mutants lacking *dksA* fail to reduce rRNA transcription at low steady-state growth rates, after amino acid starvation, and as cells enter stationary phase (Paul et al. 2004a). Although the concentration of DksA itself does not appear to vary, overproduction of DksA can further decrease rRNA transcription in vivo, and high concentrations of DksA by itself can inhibit promoter activity in the absence of ppGpp in vitro (Magnusson et al. 2007, Paul et al. 2004a, Potrykus et al. 2006, Rutherford et al. 2007). These results suggest that DksA concentrations are not saturating for control of rRNA promoters under normal conditions and that the requirement for *dksA* for regulation of rRNA synthesis in vivo is tied to its ability to act synergistically with the small molecules.

Although its role in regulation of rRNA transcription has been studied in the greatest detail, DksA affects many other promoters as well (Haugen et al. 2008, Magnusson et al. 2007, Paul et al. 2004b). Direct effects of DksA can be either negative or positive. DksA not only inhibits promoters responsible for

expressing rRNAs, many tRNAs, the transcription factor Fis, and many other gene products (Haugen et al. 2008, Magnusson et al. 2007, Mallik et al. 2006, Paul et al. 2004a, Potrykus et al. 2006) but also directly stimulates transcription from a variety of promoters. These include some promoters needed for amino acid biosynthesis and/or transport (Paul et al. 2005), the promoter responsible for synthesis of the mRNA binding protein Hfq (Sharma et al. 2006), and some promoters recognized by holoenzymes containing alternative sigma factors (Constanzo et al. 2008).

Unlike conventional transcription factors, DksA does not bind to DNA, and it interacts with promoter complexes by binding directly to RNAP (Paul et al. 2004a, Perederina et al. 2004). DksA resembles the transcription elongation factors GreA and GreB in overall size and in the structure of its coiled-coil domain (Perederina et al. 2004). Like these factors, DksA binds in the secondary channel of RNAP (Opalka et al. 2003, Perederina et al. 2004, Rutherford et al. 2007; I. Toulokhonov and R. L. Gourse, unpublished). DksA affects transcription initiation allosterically, disrupting RNAP- DNA interactions at least 25 Å away from its site of interaction with RNAP and shortening the half-life of one of the transcription initiation intermediates formed by RNAP on the pathway to open-complex formation (Rutherford et al. 2009).

In order to address the mechanism of DksA action further, we have developed a quantitative assay for measuring the affinity of DksA for RNAP. This assay facilitates genetic approaches to DksA function, e.g., it allows distinction between residues in DksA or RNAP that are needed specifically for binding

and/or for subsequent steps in the mechanism of DksA action. Furthermore, we show here that the assay allows comparison of the binding affinities of DksA for free RNAP and for promoter complexes.

The DksA binding assay (pictured schematically in Fig. 1) is an adaptation of a previously described method used to demonstrate the proximity of the transcription factors GreA and GreB to the RNAP active center (Laptenko et al. 2003). In this assay, an active-site Mg^{2+} is replaced by ferrous iron, generating hydroxyl radicals upon addition of dithiothreitol (DTT) (Zaychikov et al. 1996). The hydroxyl radicals cleave the coiled-coil tip of bound DksA, localizing it to within ~10 Å of the RNAP active site, the limit of hydoxyl radical diffusion (Perederina et al. 2004, Rutherford et al. 2009; Toulokhonov and Gourse, unpublished). The assay has not been adapted previously for measuring binding affinities of transcription factors for RNAP.

Standard binding assays, such as gel shifts, were unreliable in our preliminary attempts to measure the affinity of DksA for RNAP (data not shown), perhaps because the DksA-RNAP interaction is transient and most assays require that complexes persist over time (e.g., during electrophoresis). In contrast, the localized iron-mediated cleavage assay does not require formation of complexes with long lifetimes, it does not require separation of complexes from the free reactants, and it allows measurements to be made in solution. Finally, because it demands positioning of the DksA coiled coil at its functionally significant site (within ~10 Å of the active site), binding events in which DksA might interact only with the periphery of RNAP would not be scored as positive in

this assay.

HMK-His6 DksA binds to RNAP like wild-type DksA in vivo and in vitro

A vector expressing an N-terminally heart muscle kinase (HMK)-tagged, histidine (His6)-tagged DksA was constructed by inserting the *dksA* gene into pET33 to form pRLG8150 (Rutherford et al. 2009), and the protein was overexpressed in *E. coli* BL21(DE3) (Novagen). The resulting protein contains 25 extra amino acids at its N terminus compared to wild-type DksA (5 more residues at its N terminus than the His6 version described previously) (Paul et al. 2004a). HMK-His6 DksA was purified by metal affinity chromatography (Paul et al. 2004a).

His6 DksA is indistinguishable in function from wild-type DksA both in vitro and in vivo (Paul et al. 2004a). Tests were performed to ensure that HMK-His6 DksA was also indistinguishable from the native protein in function. First, we showed that HMK-His6 DksA encoded by a plasmid complemented a Δ*dksA* mutant strain, allowing growth on a medium lacking amino acids (Fig. 2A) (Paul et al. 2004a, Rutherford et al. 2009). Second, plasmid-encoded HMK-His6 DksA reduced the activity of an *rrnB* P1 promoter-*lacZ* fusion in a Δ*dksA* mutant strain, returning *rrnB* P1 activity to the same level as in wild-type cells and as in Δ*dksA* cells complemented with wild-type DksA (Fig. 2B). Third, purified HMK-His6 DksA inhibited transcription from the *rrnB* P1 promoter in vitro, with a concentration dependence similar to that for His6 DksA (Fig. 2C). The RNA-1 promoter was not affected by DksA, indicating that HMK-His6 DksA acts

promoter specifically. Taken together, the data shown in Fig. 2A to C indicate that HMK-His6 DksA functions like native DksA.

Purified HMK-His6-tagged DksA was ^{32}P-labeled as described previously (Rutherford et al. 2009). The buffer was then replaced with buffer A (20 mM Na-HEPES, pH 7.5, 20 mM NaCl) by two passes through a G-50 QuickSpin column (GE Healthcare). RNAP core and holoenzyme were purified by standard procedures (Burgess and Jendrisak 1975, Gaal et al. 2001), and the storage buffer was replaced with buffer A. The concentrations of both DksA and RNAP were then determined using a protein dye reagent assay (Bio-Rad).

For localized iron-mediated cleavage of DksA, ^{32}P-labeled DksA was added to RNAP for 5 min in buffer A at 30°C. Cleavage was initiated by addition of a mixture of 1 μl of freshly prepared 500 μM $(NH4)_2$-$Fe(SO_4)_2$ and 1 μl of 100 mM DTT to the 10-μl reaction mixture for 10 min. (Control time course experiments indicated that cleavage was maximal by this time [data not shown].) The reaction was stopped with 12 μl of 2X lithium dodecyl sulfate loading buffer (Invitrogen) containing glycerol, electrophoresis was performed on 4 to 12% NuPAGE Bis-Tris gels using MES [2-(N-morpholino) ethanesulfonic acid] buffer (Invitrogen), and the results were analyzed by phosphorimaging. Appearance of the cleaved product was dependent on the presence of Fe^{2+}, DTT, and RNAP (data not shown).

Competition experiments between ^{32}P-labeled HMK-His6 DksA and unlabeled His6 DksA (data not shown) or unlabeled HMK-His6 DksA (Fig. 2D)

confirmed that phosphorylation had no effect on the affinity of DksA for RNAP. When RNAP was limiting (10-fold excess of DksA over RNAP prior to addition of unlabeled DksA competitor), the fraction of ^{32}P-labeled DksA cleaved decreased as a function of increasing DksA competitor. The addition of equimolar concentrations of unlabeled DksA decreased cleavage by 50%. Thus, the decrease in the cleaved fraction of DksA fit the prediction for simple competition.

RPO binds DksA more weakly than free RNAP

To measure binding of DksA to RNAP holoenzyme in an open complex (RPO), we utilized a promoter that contains a perfect consensus DNA sequence for recognition by Eσ^{70}. This "full-con" promoter (Gaal et al. 2001) was chosen because it makes open complexes that are stable for many hours, much longer than the duration of the assay. The potential for contamination with free RNAP was reduced still further by removal of free RNAP with heparin beads (see below). Using a promoter that made a very stable complex and treatment of the complexes with heparin beads ensured that the estimate of the DksA binding constant for RPO would not be compromised by the presence of free RNAP in the reaction mixture.

The full-con promoter DNA fragment was generated by PCR from pRLG3749 (Gaal et al. 2001), a pRLG770 derivative (Ross et al. 1990) containing promoter positions -54 to +16 with respect to the transcription start site as well as 94 bp of flanking DNA. The DNA fragment was always in at least

twofold stoichiometric excess over RNAP at all RNAP concentrations. After addition of DNA, heparin beads (100 μg/ml final; Thermo Scientific) were added to remove any unbound RNAP, and then the beads were removed by microcentrifugation (1 min). The buffer was replaced with buffer A (see above) by two passes through a G-50 QuickSpin column, and the protein concentration was determined using a protein dye reagent assay (Bio-Rad).

The amount of DksA cleavage product generated by binding to RPO was ~10% of the amount of input DksA at the highest RNAP concentration tested (Fig. 3C, inset), sufficient to allow determination of an apparent Kd but lower than the ~30% plateau level typically obtained with free RNAP holoenzyme (Fig. 3B; also see the legend for Fig. 2D). When we employed the same analysis as for free RNAP core and holoenzyme, the binding constant obtained for DksA with RPO was 961 nM +/- 235 nM (Fig. 3C), almost 10-fold weaker than that obtained with free RNAP.

If our RPO preparation were contaminated with free RNAP, the difference in the apparent DksA binding constants for free RNAP versus RPO would have been even greater. To estimate the potential level of contamination with free RNAP, we analyzed the composition of our purified complex by native gel electrophoresis (data not shown). Under conditions in which contamination with 1% free RNAP would have been detected easily, we observed no free RNAP in our RPO preparation. We calculated that if there was as much as 1% free RNAP in our preparation, the real Kd of RPO would be 1,300 nM.

The low fraction of cleaved DksA with RPO likely resulted from at least two considerations. First, as indicated in Fig. 3C, the amount of RPO at the highest concentration tested (1,500 nM) was not quite saturating (technical considerations limited the use of concentrations of RPO high enough to reach full saturation). Our inability to achieve a plateau for binding experimentally resulted in a higher degree of error in the calculated value of the apparent Kd for DksA binding to RPO than to the value for binding to core and holoenzyme. Second, the level of cleavage of DksA in RPO was decreased by chelation of Fe^{2+} by DNA. Consistent with this interpretation, the addition of more Fe^{2+} increased the fraction of DksA cleaved but did not affect the concentration of RNAP at which half- maximal cleavage was observed (data not shown). We therefore added a gross excess of full-con promoter DNA (3 μM) to all reaction mixtures containing RPO to sequester the same amount of Fe^{2+} in each.

The increase in the apparent Kd of DksA for RPO relative to that for free RNAP indicates that promoter DNA changes the conformation of RNAP, directly or indirectly altering the RNAP-DksA interface. The apparent Kd for RPO binding to DksA reported here is similar to that obtained from estimates based on assays of DksA function, i.e., effects of DksA on transcription or on complex half-life (Rutherford et al. 2009). The apparent affinity of DksA for RPO reported here is also well within the range where binding of DksA to RPO would be expected to occur *in vivo*, considering the absolute concentrations of DksA estimated in cells (Paul et al. 2004a, Rutherford et al. 2007). The concentration of DksA was held constant in the assay described here, and the concentration of RNAP was varied.

In contrast, the RNAP concentration was fixed and the DksA concentration was varied in the measurements of binding based on DksA function (Rutherford et al. 2009). Thus, our results confirm that apparent Kd estimates can be made by varying either RNAP or DksA.

Future directions

DksA overexpression in vivo increases the effect of DksA on rRNA promoter activity. Therefore, the concentration of DksA is not saturating for effects on transcription initiation in cells (Potrykus et al. 2006, Rutherford et al. 2007), and differences in affinity of DksA for free RNAP versus RPO could differentially affect free versus promoter-bound RNAP in vivo. Further investigations will address the potential for the promoter sequence to affect DksA binding differentially and the nature of the DNA-induced RNAP conformational changes that modulate the binding of DksA. Although we have restricted our studies to the effects of DksA on transcription initiation, DksA also binds to RNAP during elongation (Perederina et al. 2004), and the presence of DNA could affect DksA occupancy of the RNAP core enzyme at this step in transcription. We note that this assay also could be adapted for the analysis of interactions between GreA or GreB and RNAP.

In summary, we have developed a localized Fe^{2+}-mediated, RNAP-dependent cleavage assay to determine the apparent affinity of DksA for RNAP. We illustrate the utility of the assay here by demonstrating that DksA binds with

equal affinities to RNAP core and holoenzyme and that promoter DNA significantly decreases the affinity of DksA for RNAP. We are using this assay for the analysis of DksA and RNAP variants in structure-function studies of the mechanism of DksA action. For example, DksA mutants (bearing substitution L15F or N88I) that inhibit rRNA promoter activity more strongly than wild-type DksA in vivo and in vitro were identified recently (Blankschien et al 2009). The assay described here allowed us to determine that the increased activities of these mutants resulted primarily from an increase in DksA binding affinity.

Acknowledgements

We thank I. Toulokhonov for construction of the HMK-His6 DksA expression plasmid, R. Saecker for help with data analysis, and S. Borukhov for comments on the manuscript. This work was supported by a predoctoral fellowship from the National Institutes of Health to C.W.L. and an NIH research grant (R37 GM37048) to R.L.G.

References

Blankschien MD, Lee JH, Grace EA, Lennon CW, Halliday JA, Ross W, Gourse RL, Herman C. 2009. Super DksAs: substitutions in DksA enhancing its effects on transcription initiation. *EMBO J* **28**:1720-1731.

Burgess RR, Jendrisak, JJ. 1975. Procedure for the rapid, large-scale purification of *Escherichia coli* DNA-dependent RNA polymerase involving polymin P precipitation and DNA-cellulose chromatography. *Biochemistry* **14**:4634-4638.

Costanzo A, Nicoloff H, Barchinger SE, Banta AB, Gourse RL, Ades, SE. 2008. ppGpp and DksA likely regulate the activity of the extra- cytoplasmic stress factor sigma E in *Escherichia coli* by both direct and indirect mechanisms. *Mol Microbiol* **67**:619-632.

Gaal T, Ross W, Estrem ST, Nguyen LH, Burgess RR, Gourse RL. 2001. Promoter recognition and discrimination by $E\sigma^S$ RNA polymerase. *Mol Microbiol* **42**:939-954.

Haugen SP, Ross W, Gourse RL. 2008. Advances in bacterial promoter recognition and its control by factors that do not bind DNA. *Nat Rev Microbiol* **6**:507-519.

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

Magnusson LU, Gummesson B, Jaksimović P, Farewell A, Nyström T. 2007. Identical, independent, and opposing roles of ppGpp and DksA in *Escherichia coli. J Bacteriol* **189**:5193-5202.

Mallik P, Paul BJ, Rutherford ST, Gourse RL, Osuna R. 2006. DksA is required for growth phase-dependent regulation, growth rate-de- pendent control, and stringent control of *fis* expression in *Escherichia coli. J Bacteriol* **188**:5775-5782.

Murakami KS, Darst SA. 2003. Bacterial RNA polymerases: the wholo story. *Curr Opin Struct Biol* **13**:31–39.

Murakami KS, Masuda S, Campbell EA, Muzzin O, Darst SA. 2002. Structural basis of transcription initiation: an RNA polymerase holoenzyme-DNA complex. *Science* **296**:1285-1290.

Murray HD, Schneider DA, Gourse RL. 2003. Control of rRNA expression by small molecules is dynamic and nonredundant. *Mol Cell* **12**:125-134.

Opalka N, Chlenov M, Chacon P, Rice WJ, Wriggers W, Darst SA. 2003.

Structure and function of the transcription elongation factor GreB bound to bacterial RNA polymerase. *Cell* **114**:335-345.

Paul BJ, Barker MM, Ross W, Schneider DA, Webb C, Foster JW, Gourse RL. 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 BJ, Berkmen MM, Gourse RL. 2005. DksA potentiates direct activation of amino acid promoters by ppGpp. *Proc Natl Acad Sci USA* **102**:7823-7828.

Paul BJ, Ross W, Gaal T, Gourse RL. 2004. rRNA transcription in *Escherichia coli. Annu Rev Genet* **38**:749-770.

Perederina A, Svetlov V, Vassylyeva MN, Tahirov TH, Yokoyama S, Artsimovitch I, Vassylyev DG. 2004. Regulation through the secondary channel-structural framework for ppGpp-DksA synergism during transcription. *Cell* **118**:297-309.

Potrykus D, Vinella D, Murphy H, Szalewska-Palasz A, D'Ari R, Cashel M. 2006. Antagonistic regulation of *Escherichia coli* ribosomal RNA *rrnB* P1 promoter activity by GreA and DksA. *J Biol Chem* **281**:15238-15248.

Ross W, Thompson JF, Newlands JT, Gourse RL. 1990. *E. coli* Fis protein activates ribosomal RNA transcription *in vitro* and *in vivo. EMBO J* **9**:3733-3742.

Rutherford ST, Lemke JJ, Vrentas CE, Gaal T, Ross W, Gourse RL. 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 ST, Villers CL, Lee JH, Ross W, Gourse RL. 2009. Allosteric control of *Escherichia coli* rRNA promoter complexes by DksA. *Genes Dev* **23**:236-248.

Sharma AK, Payne SM. 2006. Induction of expression of *hfq* by DksA is essential for *Shigella flexneri* virulence. *Mol Microbiol* **62**:469-479.

Zaychikov E, Martin E, Denissova L, Kozlov M, Markovtsov V, Kashlev M, Heumann H, Nikiforov V, Goldfarb A, Mustaev A. 1996. Mapping of catalytic residues in the RNA polymerase active center. *Science* **273**:107-109.

Figures and legends

Figure 2.1. Localized Fe^{2+}-mediated, RNAP-dependent cleavage of ^{32}P-labeled HMK-His6 DksA. (A) Free RNAP with Mg^{2+} (sphere) at the catalytic center. (B) ^{32}P-labeled HMK-His6 DksA binds in the secondary channel of RNAP holoenzyme. (C and D) The active-site Mg^{2+} in RNAP is replaced by Fe^{2+} (C), and DTT generates hydroxyl radicals that cleave the coiled-coil tip of DksA (D) (Fe^{2+} addition and DTT addition are pictured as sequential steps for puRPoses of illustration but were actually performed simultaneously). (E) Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) profile of DksA in the absence (-) and presence (+) of RNAP. For more detailed representations of the RNAP holoenzyme, see references 5, 9, and 10.

Figure 2.1

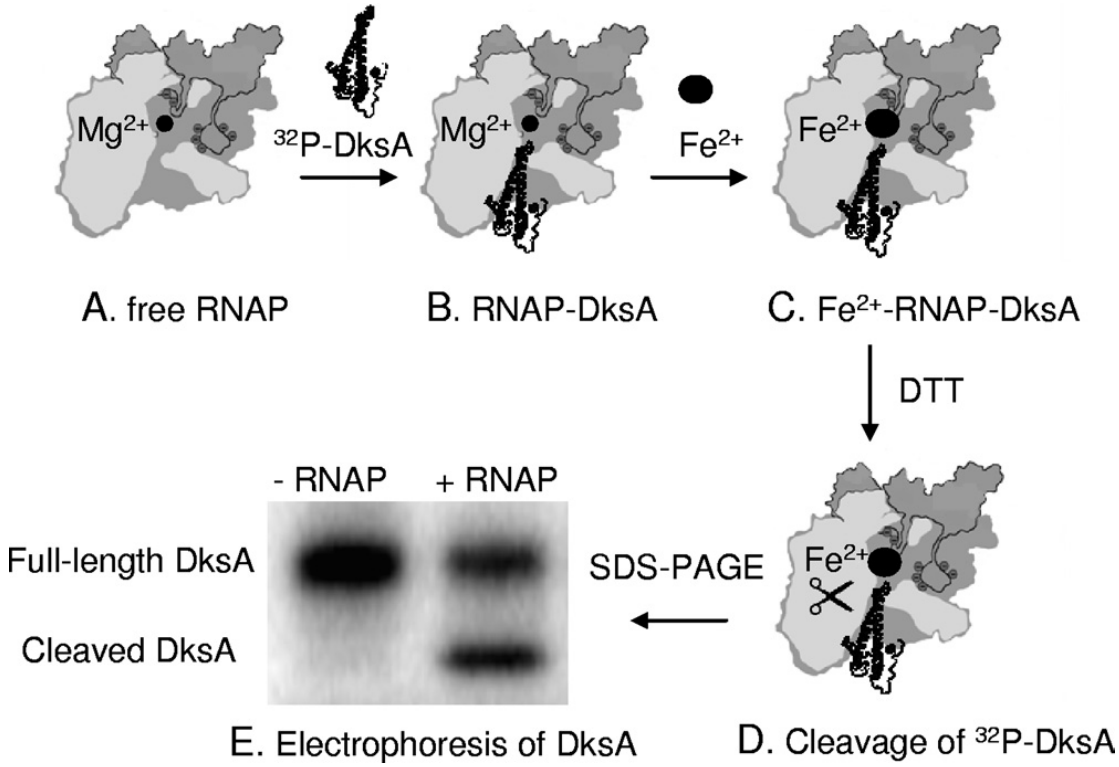

Mg2+
32P-DksA
Mg2+
Fe2+
Fe2+
A. free RNAP
B. RNAP-DksA
C. Fe2+-RNAP-DksA
DTT
- RNAP
+ RNAP
Full-length DksA
Cleaved DksA
SDS-PAGE
Fe2+
E. Electrophoresis of DksA
D. Cleavage of 32P-DksA

Figure. 2.2. HMK-His6 DksA binds to RNAP like wild-type DksA in vivo and in vitro. (A) HMK-His6 DksA complements a Δ*dksA* mutant strain for growth on M9 minimal medium with 0.4% glycerol and without amino acids. HMK-His6 DksA was produced from plasmid pRLG9511 (20). (B) HMK-His6 DksA reduces *rrnB* P1 promoter activity to the same extent as in cells containing native DksA. β-Galactosidase activities (in Miller units) in strains containing *rrnB* P1 promoter-*lacZ* fusions (promoter sequence endpoint, -61 to +1) in log phase (left) and stationary phase (right) were determined as described previously (19). Cells were grown at 30°C in LB medium. Error bars indicate standard deviations. Wild type (WT), RLG5950; Δ*dksA*, RLG7062. (C) HMK-His6 DksA inhibits rRNA promoter activity in vitro. The amount of transcript from the *rrnB* P1 promoter is indicated below each lane as a percentage of that without DksA. RNAP (10 nM) and DksA (at the indicated concentrations) were added to plasmid pRLG5944 containing the *rrnB* P1 promoter (sequence endpoints, -61 to +1) and the *RNA-1* promoter. Transcripts are indicated. Single-round transcription reactions were performed as described previously (Blankschien et al. 2009). The buffer contained 50 mM NaCl. The assay was performed multiple times with similar results, and results from a representative experiment are shown. (D) Unlabeled HMK-His6 DksA competes with ^{32}P-labeled HMK-His6 DksA for RNAP-dependent, Fe^{2+}-mediated cleavage of DksA. The RNAP holoenzyme concentration was 0.1 μM, the ^{32}P-DksA concentration was 1 nM, and the amounts of excess unlabeled DksA are indicated on the *x* axis. The curve was generated from five independent experiments in which the fraction of DksA cleaved in each experiment in the

absence of unlabeled DksA (typically~30% of input ^{32}P-DksA) was normalized to 1.0. Inset, representative gel from a single experiment in which 0, 2, 4, 8, or 16 μM unlabeled DksA was added to 1 nM ^{32}P-DksA and 0.1 μM RNAP holoenzyme for 5 min before addition of DTT to initiate DksA cleavage.

Figure 2.2

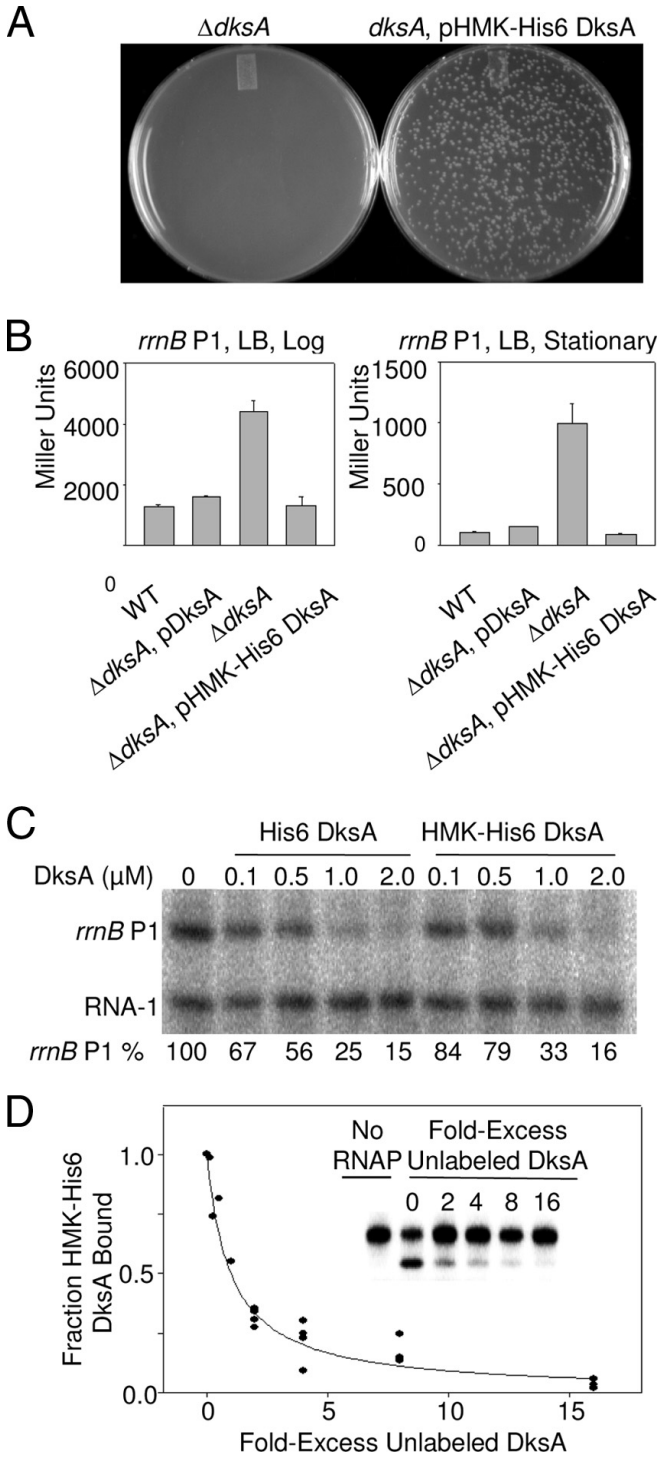
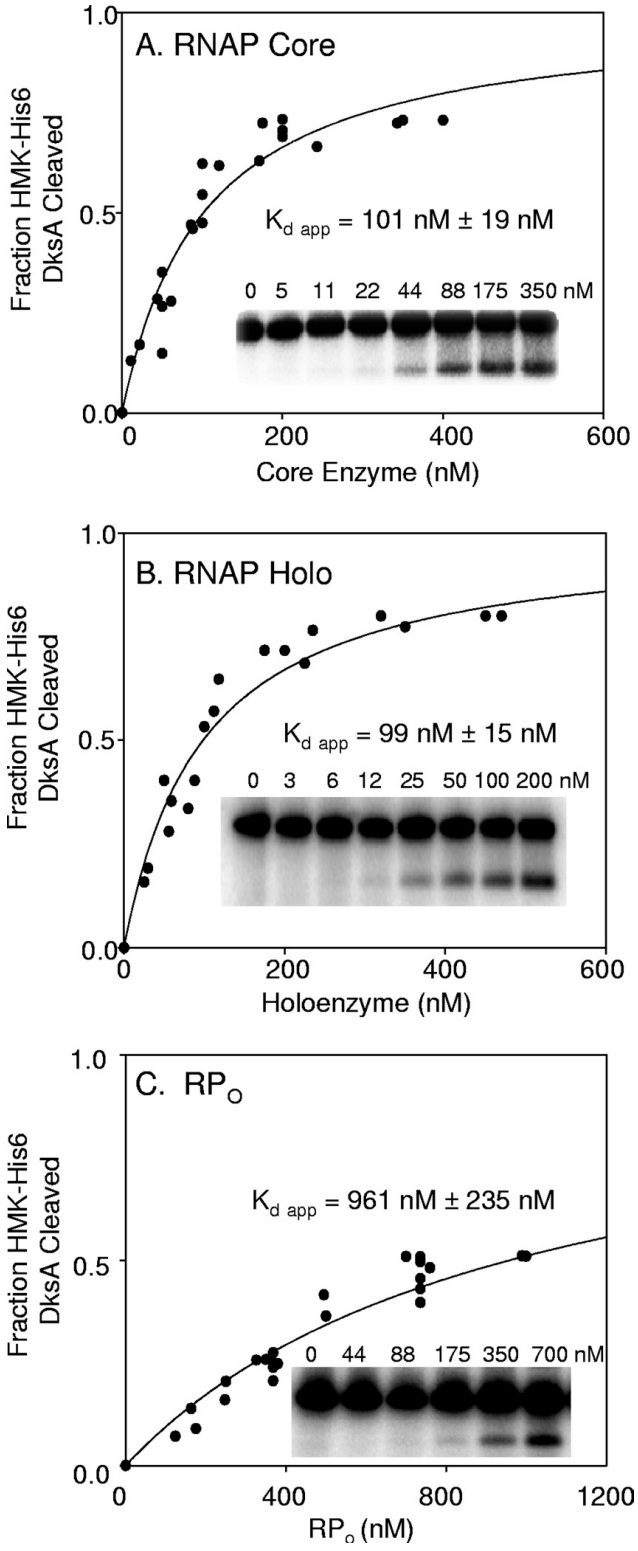

A
ΔdksA
dksA, pHMK-His6 DksA
B
rrnB P1, LB, Log
Miller Units
6000
4000
2000
0
WT
ΔdksA, pDksA
ΔdksA
ΔdksA, pHMK-His6 DksA
rrnB P1, LB, Stationary
Miller Units
1500
1000
500
0
WT
ΔdksA, pDksA
ΔdksA
ΔdksA, pHMK-His6 DksA
C
His6 DksA
HMK-His6 DksA
DksA (μM) 0 0.1 0.5 1.0 2.0 0.1 0.5 1.0 2.0
rrnB P1
RNA-1
rrnB P1 % 100 67 56 25 15 84 79 33 16
D
No RNAP
Fold-Excess Unlabeled DksA
0 2 4 8 16
Fraction HMK-His6 DksA Bound
1.0
0.5
0.0
0
5
10
15
Fold-Excess Unlabeled DksA

Figure 2.3. Binding curves for DksA with different forms of RNAP. Fe^{2+}-mediated cleavage assays were performed using ^{32}P-DksA and increasing amounts of RNAP. RNAP was in the form of core enzyme (A), holoenzyme (Holo) ($E\sigma^{70}$) (B), and heparin-resistant open complex (C). The DksA concentration was 1 nM for panels A and B or 10 nM for panel C. Each point represents the result for a single reaction from a single experiment. Results from at least six independent experiments were plotted together after normalization to the maximum fraction of DksA cleaved in each experiment. The apparent Kd (Kd app) represents the concentration of RNAP at which half-maximal cleavage of DksA was obtained (see the text for a discussion of the computational analysis). Insets, gel images from single representative experiments.

Figure 2.3


A. RNAP Core
Fraction HMK-His6 DksA Cleaved
$K_{d\ app}$ = 101 nM ± 19 nM
0 5 11 22 44 88 175 350 nM
Core Enzyme (nM)
B. RNAP Holo
$K_{d\ app}$ = 99 nM ± 15 nM
0 3 6 12 25 50 100 200 nM
Holoenzyme (nM)
C. RP_O
$K_{d\ app}$ = 961 nM ± 235 nM
0 44 88 175 350 700 nM
RP_o (nM)

Chapter 3

Direct interactions between the coiled-coil tip of DksA and the trigger loop of RNA polymerase mediate transcriptional regulation

This chapter has been published in *Genes and Development* (Christopher W. Lennon, Wilma Ross, Steve Martin-Tumasz, Innokenti Toulokhonov, Catherine E. Vrentas, Steven T. Rutherford, Jeong-Hyun Lee, Samuel Butcher and Richard L. Gourse. 2012. Genes Dev 26: 2634-2636). Credit for experiments performed should go to SMT (HADDOCK modeling throughout MS), IT (Fig. 3.4 and Fig 3.S2), CEV (Fig 3.S4 panels B-D), STR (Fig. 3.5 panels A-B and Fig. 3.7 panels C-D), JHL (Fig. 3.6 panel F).

Abstract

Escherichia coli DksA is a transcription factor that binds to RNA polymerase (RNAP) without binding to DNA, destabilizing RNAP-promoter interactions, sensitizing RNAP to the global regulator ppGpp, and regulating transcription of several hundred target genes, including those encoding rRNA. Previously we described promoter sequences and kinetic properties that account for DksA's promoter specificity, but how DksA exerts its effects on RNAP has remained unclear. To better understand DksA's mechanism of action, we incorporated benzoyl-phenylalanine at specific positions in DksA and mapped its crosslinks to RNAP, constraining computational docking of the two proteins. The resulting evidence-based model of the DksA-RNAP complex, as well as additional genetic and biochemical approaches, confirmed that DksA binds to the RNAP secondary channel, defined the orientation of DksA in the channel, and predicted a network of DksA interactions with RNAP that includes the rim helices and the mobile trigger loop domain (TL). Engineered cysteine substitutions in the TL and DksA coiled-coil tip generated a disulfide bond between them, and the interacting residues were absolutely required for DksA function. We suggest that DksA traps the TL in a conformation that destabilizes promoter complexes, an interaction explaining the requirement for the DksA tip and its effects on transcription.

Introduction

Most bacterial transcription factors are DNA binding proteins that recognize a particular DNA sequence, thereby regulating transcription of nearby promoter(s). However, the 151 amino acid (17.5 kDa) DksA protein does not bind to DNA but rather has been proposed to bind to the RNAP secondary channel (Perederina et al. 2004; Rutherford et al. 2009; Lennon et al. 2009). DksA regulates transcription in conjunction with the small molecules guanosine tetraphosphate and pentaphosphate (referred to here as ppGpp) and the transcript's initiating nucleotide (iNTP) by modifying the kinetic properties of the promoter-RNAP complex (Barker et al. 2001; Paul et al. 2004). Although the concentration of DksA remains relatively constant during growth of *E. coli* (Paul et al. 2004; Rutherford et al. 2007; Chandrangsu et al. 2011), the concentrations of ppGpp and NTPs fluctuate, accounting for the dynamic effects of DksA on transcriptional responses to changing nutritional conditions (Murray et al. 2003; Paul et al. 2004).

DksA and ppGpp together inhibit transcription from rRNA promoters, many ribosomal protein and tRNA promoters, the promoter for FlhDC (the master regulator of flagella synthesis), the Fis promoter, DksA's own promoter, and as many as 300 others, and DksA/ppGpp directly activate promoters for amino acid biosynthesis and/or transport, virulence, the sRNA binding protein Hfq, σ^E-dependent transcription, and as many as 400 others (reviewed in Haugen et al. 2008; see also Potrykus et al. 2006; Magnusson et al. 2007; Durfee et al. 2008; Traxler et al. 2008; Lemke et al. 2009; Lemke et al. 2011; Chandrangsu et al.

2011; J. Lemke and R.L.G. ubpublished). However, the mechanism(s) responsible for inhibition and activation remain ill-defined.

Transcription initiation begins with the binding of RNAP holoenzyme (subunit composition $\alpha_2\beta\beta'\omega\sigma$) to promoter DNA to form an initial closed complex (RP_C) (Haugen et al. 2008; Saecker et al. 2011). This complex isomerizes to additional intermediates, abbreviated here as RP_I, before formation of a transcriptionally-competent open complex (RP_O) and RNA synthesis (Haugen et al. 2008; Saecker et al. 2011). DksA inhibits the transition from RP_C to RP_I (Rutherford et al. 2009). We have proposed that by binding in the RNAP secondary channel, DksA causes an allosteric change in the bridge helix (BH) and/or the trigger loop (TL) that ultimately affects the switch regions of RNAP. The switches interact with the promoter near the transcription start site, and they control RNAP clamp opening, thereby affecting DNA contacts further downstream in RP_O (Rutherford et al. 2009).

Rather than achieving promoter specificity by binding to specific DNA sites, DksA exploits promoter-specific variation in the kinetics of transcription initiation (Paul et al. 2004; Paul et al. 2005; Rutherford et al. 2009). At most promoters, the RNAP-promoter complex is long-lived. However, at rRNA promoters, this complex is intrinsically short-lived, such that dissociation of RNAP from rRNA promoters is in competition with NTP addition (Barker et al. 2001). DksA/ppGpp decrease the lifetime of the RNAP-promoter complex, inhibiting rRNA transcription (Barker et al. 2001; Haugen et al. 2008; Paul et al. 2004), whereas long-lived promoters are not inhibited by DksA/ppGpp because RNAP

escapes from the complex before dissociation significantly affects transcriptional output (Barker et al. 2001; Paul et al. 2004; Paul et al. 2005).

Although we have a kinetic framework for understanding DksA's effects on transcription initiation (Rutherford et al. 2009), and there is a high resolution structure of DksA (Perederina et al. 2004), a molecular understanding of its mechanism of action has been hampered by the absence of a structure of a DksA-RNAP complex. Structural information is available for the secondary channel binding factors GreB, TFIIS, and Gfh1 in complex with RNAP (Opalka et al. 2003; Kettenberger et al. 2003; Tagami et al. 2010). However, these factors primarily affect transcription elongation rather than initiation, and their amino acid sequences bear little resemblance to DksA. Furthermore, a DksA homolog has not been identified in the thermophilic organisms from which all reported bacterial RNAP structures to date have been obtained. An RNAP-DksA complex structure has not been reported, and *E. coli* DksA does not appear to bind to, or function on, *Thermus* RNAP (T. Gaal, C.E.V., and R.L.G., unpublished).

Here we present a model for the interaction of DksA with *E. coli* RNAP, developed by site-specific incorporation of the crosslinkable amino acid benzoyl-phenylalanine (Bpa; Chin et al. 2002) into DksA, mapping of crosslinked sites in RNAP, and then computational docking of DksA on RNAP. The model, in conjunction with the results of hydroxyl radical protein-protein footprinting and characterization of variants of RNAP that affect DksA binding and/or function, unambiguously positions DksA in the secondary channel and identifies interactions with both the rim helices (RH), located at the entrance to the

secondary channel, and the TL of the β' subunit. Engineered cysteine residues in the DksA coiled-coil tip (cc-tip) and the TL resulted in formation of a disulfide bond between them. We propose that a TL-DksA interaction provides an explanation for the sequence requirement of the cc-tip for regulating transcriptional output and plays a central role in the mechanism of DksA action.

Results

Benzoyl-phenylalanine (Bpa) incorporates site-specifically into DksA and crosslinks to the β' subunit of RNAP

To build an evidence-based model of DksA bound to RNAP, we utilized site-specific incorporation of the non-natural, UV-crosslinkable amino acid benzoyl-phenylalanine (Bpa) into DksA, a technique that relies on in vivo expression of an evolved tRNA and tRNA synthetase pair and incorporation of Bpa at an amber stop codon (Chin et al. 2002; Ryu and Schultz, 2006). After excitation with UV light, Bpa side chains form covalent crosslinks in vitro with carbons within ~3Å (Kauer et al. 1986).

DksA has two domains, a long coiled-coil consisting of residues from the central portion of the protein with two acidic residues at the tip, and a globular domain consisting of both N- and C-terminal segments of the protein including four cysteines that coordinate a zinc atom (Perederina et al. 2004; Fig. 1A). DksA variants containing single Bpa substitutions were created at 12 surface-accessible positions (codons S13, A18, R52, D64, F69, E79, E81, R87, N88, K94, E146, Q148), chosen because they were unlikely to participate in protein folding. Overexpressed DksA-Bpa proteins co-migrated with purified wild-type DksA (cf. Fig. 1B, lanes 1 and 4; in the example shown, Bpa is at residue 69) and with overexpressed wild-type DksA (lane 2). Little or no full-length DksA was obtained when the growth medium lacked Bpa (lane 3). The yield of DksA proteins containing Bpa was similar to that from expression of wild-type DksA (compare lanes 2 and 4 and data not shown). When Bpa was not added to the

growth medium, truncated DksA products were observed only from the constructs with stop codons at E146 or Q148 (data not shown), suggesting that the other truncated DksA products were unstable.

Specific crosslinks to RNAP were observed with ^{32}P-labeled HMK-tagged DksA proteins with a Bpa at six positions: S13, F69, E79, E81, N88 and E146 (Fig. 1C; E79-Bpa is shown as an example) and were used for modeling (below). [The other six Bpa substitutions did not form specific crosslinks (Supplemental Materials and Methods) and were not used further.] The crosslinks were UV-dependent (lanes 3,4), and the electrophoretic mobilities of the crosslinked complexes on SDS-PAGE corresponded to those expected for HMK-DksA (18 kDa) and one of the large RNAP subunits, β (151 kDa) or β' (155 kDa) (Fig. 1C). The DksA proteins that crosslinked specifically to RNAP retained function: each reduced transcription from *rrnB* P1, but not from the plasmid-encoded promoter, RNA-1 (Fig. S1).

To determine whether the DksA crosslink was to β or β', we repeated the reaction with each of the six DksA-Bpa proteins and RNAP containing a β' subunit with a C-terminal GFP fusion adding ~27 kDa to β' (Bratton et al. 2011). In all 6 cases, ~90-95% of the radiolabeled crosslinked species migrated slightly slower than with wild-type RNAP, to a position consistent with the MW of a complex containing β'-GFP (155 kDa plus 27 kDa) crosslinked to DksA (18 kDa) (Fig. 1C, lane 6; Fig. 1D, lanes 2-7). This is also consistent with the fact that the secondary channel of RNAP is composed almost entirely of β' (Vassylyev et al. 2002).

We demonstrated previously that core RNAP and holoenzyme bind DksA with the same affinity (Lennon et al. 2009), consistent with a location for DksA far from where σ binds to RNAP. Additionally, we did not observe any differences in DksA crosslinking in pilot experiments in the presence or absence of σ (data not shown). Therefore we used core RNAP to conserve material in the studies described below.

Mapping the DksA crosslinks within β'

DksA crosslinks were mapped within β' using thrombin, which does not cleave DksA under these conditions. Thrombin cut β' primarily at one site, corresponding to position ~900 (right lane for each DksA variant in Fig. 2A; see also Supplemental Experimental Procedures). The major radioactive digestion product of the crosslinked RNAP-DksA complex containing the S13, F69, E79, E81, or N88 DksA-Bpa variants migrated slightly above a 110 kDa size marker, consistent with the predicted size of a complex containing β' 1-900 (101 kDa) and HMK-DksA (18 kDa) (Fig. 2A, black dot). The crosslinked complex containing E146 DksA-Bpa migrated between the 60 kDa and 80 kDa size markers, consistent with the expected size of β' 901-1407 (54 kDa) crosslinked to HMK-DksA (18 kDa) (Fig. 2A; orange dot).

To confirm and further map the positions of the crosslinks, we used an RNAP containing β' with a thrombin site engineered at position 648 (β'-648^{Th}) in addition to the one at ~900 (Supplemental Materials and Methods), resulting in cleavage of β' 1-900 into β' 1-648 and β'649-900 fragments (Figs. 2B, 2C). When

the crosslinked complexes formed by β'-648Th RNAP and the S13, F69, E79, E81 or N88 DksA-Bpa variants were treated with thrombin, the predominant labeled digestion product in each case migrated as a smaller complex than that formed with wild-type RNAP, confirming that the crosslink occurred within the β' 1-900 fragment (Fig. 2B). In four cases (S13, F69, E81, and N88 DksA-Bpa), a product (cyan dots) migrated at approximately the position of the ~50 kDa size marker, consistent with a complex consisting of HMK-DksA (18 kDa) and β' 649-900 (28 kDa). An E79 DksA-Bpa cleavage product migrated between the 80 kDa and 110 kDa size markers, consistent with a crosslink of HMK-DksA (18 kDa) to β' 1-648 (73 kDa; green dot; Fig. 2B). Treatment of the E146 DksA-Bpa-β' 648Th RNAP complex with thrombin did not change the mobility of the crosslinked fragment (orange dot) compared to that observed with wild-type RNAP, consistent with the conclusion that E146 DksA-Bpa crosslinked to the C-terminal thrombin fragment of β' (901-1407; 54 kDa). This was confirmed by crosslinking β'-GFP RNAP with E146 DksA-Bpa, followed by digestion with thrombin, resulting in a complex that migrated slower than the same complex derived from wild-type β' (data not shown).

In summary, DksA residues S13, F69, E81, and N88 crosslinked to β' 649-900; E79 crosslinked to β' 1-648; and E146 crosslinked to β' 901-1407 (Fig. 2C). Fig. 2D illustrates the locations of these fragments on *E. coli* core RNAP (Opalka et al. 2010), color-coded as in Fig. 2C.

Construction of our model of the DksA-RNAP complex depended on one further constraint, namely that Fe^{2+} substitution for the active site Mg^{2+} in RNAP

results in cleavage of DksA upon binding to RNAP (Perederina et al. 2004; Lennon et al. 2009). Creation of genetically-truncated DksAs as size markers for comparison with the DksA cleavage product allowed us to refine the DksA-RNAP interaction. The ^{32}P-HMK DksA N-terminal Fe-$^{2+}$-cleavage product and a ^{32}P-HMK DksA genetically-truncated at residue 73 migrated similarly on gels (Fig. S2, compare lanes 2 and 3), whereas full-length DksA (lane 1) or a DksA genetically-truncated at residue 85 (lane 4) migrated more slowly. We conclude that the cc-tip of DksA (residues 66-76) is within 10Å of the catalytic Mg^{2+} in RNAP.

Model of DksA bound to RNAP

Modeling of the DksA-RNAP complex was performed using HADDOCK version 2.0 (Dominguez et al. 2003; de Vries et al. 2007), a molecular docking program that can incorporate biochemical and genetic information in addition to energetics and shape complementarity. We started with the structure of DksA (Perederina et al. 2004; PDB 1TJL) and the model of *E. coli* core RNAP (Opalka et al. 2010; PDB 3LU0; which in turn is based on the high resolution *T. thermophilus* RNAP structure; Vassylyev et al. 2002). The model was forced to satisfy the constraints imposed by the 6 crosslinks to β' and the proximity of the coiled-coli tip to the active site Mg^{2+} (see Supplemental Materials and Methods), but DksA was purposely not constrained to the secondary channel in order to build the model. The biochemical data thus limited the potential ways that DksA could dock on the surface of RNAP. The 24 best models were refined (see Fig. S3 for an overlay),

and the lowest energy model is shown in Fig. 3 (see Supplemental Materials and Methods for details).

As predicted from previous studies of DksA (Perederina et al. 2004; Rutherford et al. 2009) and from the location of the structurally similar Gre factors (Opalka et al. 2003), DksA binds in the secondary channel of RNAP in our model, with helix 2 of DksA (residues 77-107) rather than helix 1 (residues 34-65) closer to β'i6 (an *E. coli* lineage-specific sequence insertion previously called Sequence Insertion 3, SI3; pink in Fig. 3B), the TL (green), and the BH (cyan). Conversely, helix 1 is closer to the RH and to the back of the channel than helix 2. The amino terminal part of the globular domain of DksA (residues 1-33) appears to interact or even potentially clash with the RH, and the carboxy terminal section in the globular domain is between the RH and β'i6 (Fig. 3C, 3D). Segments in both helices of the coiled-coil domain of DksA interact with the RH (yellow in Fig. 3B, 3D).

The functionally-critical residues at the tip of the DksA coiled-coil (D74, A76; Lee et al. 2012; red in Fig. 3B) are oriented toward the TL. We showed previously that deletion of the TL eliminates DksA function (Rutherford et al. 2009). In the structure of RNAP on which our HADDOCK model is based, the TL is in the unfolded conformation (shown in green spacefill in Figs. 3B, 3D; Opalka et al. 2010), but as discussed below, the TL is mobile, and it oscillates between an unstructured loop conformation and folded or partially-folded α-helical hairpin conformations (trigger helices, TH; trigger loop/trigger helices, TL/TH) during the

catalytic cycle (Vassylyev et al. 2007; Zhang et al. 2010). There is little or no information about the state of the TL during transcription initiation.

Our HADDOCK model is physically-plausible and evidence-based. The crosslinks allowed us to fix the relative orientation of different regions of DksA relative to those of RNAP, but it is not a high-resolution structure and interactions can only be inferred. Therefore, additional approaches were taken to confirm and extend the predictions of the model.

Biochemical and genetic support for the HADDOCK model

Protein-protein footprinting provided an independent means to assess interactions between RNAP and DksA. Complexes formed with DksA (or GreB for comparison) and RNAP (^{32}P-labeled at an HMK site on the C-terminus of β') were cleaved by hydroxyl radicals generated in solution by Fe^{2+}-EDTA, and the cleaved fragments were separated by electrophoresis. The reactions shown in Fig. 4A were electrophoresed on a lower percentage gel to resolve additional regions in β' (Fig. 4D).

Sites protected by DksA and/or GreB are indicated by arrows. Graphs of individual lanes in the gel images from Figs. 4A and 4D are shown in Figs. 4B and 4E for DksA and for GreB in Figs. 4C and 4F. DksA reduced cleavage at several positions in and near the secondary channel, including the region flanking the cyanogen bromide-generated marker band at position 466 corresponding to the active site region (AS; peak 1 in Fig. 4E) and nearby residues in the secondary channel (peak 2 in Fig. 4E: β' ~485-510). DksA also

protected an extensive region corresponding to the RH and part of the segment of β' connecting the RH to the BH (peaks 3-5; β' ~648-699 and ~700-750). The reduced cleavage could result from direct protection by DksA or from conformational change(s) in RNAP induced by DksA binding.

The cleavage pattern in the presence of GreB (Figs. 4C and 4F) was distinctly different from that observed in the presence of DksA. Like DksA, GreB protected the active site region (peaks 1 and 2), but GreB did not protect peak 3 in the RH (β' 648-660), and it protected peaks 4 and 5 in the RH less well than DksA. Conversely, GreB affected multiple sites in β'i6 much more than DksA, protecting some residues (peak 8) and enhancing others (peaks 6 and 7). Both DksA and GreB protected the TL region near residue 932, although inefficient cleavage made the protected positions difficult to visualize.

We tested the importance of the proposed interaction of DksA with the RH by monitoring its effects on promoter complex lifetime. The half-life of a promoter with RNAP lacking the RH (β'Δ645-718, β'ΔRH) was similar to that containing wild-type RNAP (Fig. 5A). However, DksA decreased the lifetime of promoter complexes containing wild-type RNAP about 7-fold under the conditions tested, whereas it had little or no effect on complexes containing β'ΔRH RNAP (Fig. 5A). We conclude that DksA does not function on RNAP lacking the β' RH.

Using the Fe^{2+}-mediated cleavage assay (Lennon et al. 2009), we measured binding of DksA to β'ΔRH RNAP. No DksA cleavage was detected, even when the RNAP concentration was 1 μM (Fig. 5B), ~10-fold higher than the apparent binding constant for wild-type RNAP (100 nM; Lennon et al. 2009).

These results support the prediction of the HADDOCK model that DksA docks on the RH. The model also predicts that DksA residue N88 is in close proximity to, and points toward, the RH surface (Fig. 5C), consistent with the effect of an N88I substitution on DksA binding and function (Blankschein et al. 2009).

The HADDOCK model also suggested some residues in the secondary channel that might interact with DksA, including β' K598/K599 and β' R731 (Fig. S4A), positions at which we had previously constructed substitutions (Vrentas et al. 2008). β' K598A/K599A and β' R731A RNAPs exhibited 2 to 3-fold reduced affinity for DksA relative to wild-type RNAP in the Fe^{2+}-mediated cleavage assay (Fig. S4B). DksA inhibited transcription from *rrnB* P1 by the mutant RNAPs 2-fold or less in vitro, whereas it reduced transcription by wild-type RNAP > 30-fold (Fig. S4C). Therefore, the substitutions may affect steps in the DksA mechanism in addition to binding to RNAP. Although the responses of wild-type and β' R731A RNAP to GreB were similar, GreB affected β' K598A/K599A RNAP more strongly than wild-type RNAP (Fig. S4D), consistent with the overlapping but distinct modes of interaction of DksA and GreB with RNAP and with their overlapping but distinct functions (Fig. 4; Rutherford et al. 2007; Lee et al. 2012).

The coiled-coil tip of DksA interacts directly with the trigger loop domain of RNAP

Deletion of the TL domain in β' or substitutions at two positions in the DksA cc-tip, D74 or A76, eliminated DksA function (Rutherford et al. 2009; Lee et al. 2012). Even the conservative change of an aspartate to a glutamate at DksA

residue 74 eliminated the effects of DksA on transcription (Lee et al. 2012). Furthermore, the DksA-dependent protection of the TL observed in our protein-protein footprints (Fig. 4) suggested that DksA might interact with the TL. However, the TL domain did not appear to be in direct contact with the cc-tip of DksA in our HADDOCK model (Fig. 3).

Because the TL in this model is in the unfolded conformation, we speculated that perhaps a different conformation of the TL might interact directly with DksA, i.e. that a fully-folded (TH) or a partially-folded (TL/TH) conformation of the TL domain might happen during transcription initiation as in transcription elongation (Vassylvev et al. 2007; Zhang et al. 2010) and might interact with DksA. When the entire β' subunit with the fully-folded (TH) structure of the TL domain was used to replace β' in the HADDOCK model using the Align function of PyMol, residues in the TL approached DksA (Fig. 6A).

To address experimentally whether the TL is in close proximity to DksA, we engineered Bpa substitutions at three different positions in the TL, β' Q929, R933, and G1136, as well as at a position in β' outside of the TL as a control, R1148. RNAP with Bpa at β' R933 crosslinked most efficiently to DksA (Fig. 6B), and β' R933-Bpa crosslinking was greatly reduced with D74N-DksA compared to wild-type DksA, even at high RNAP concentrations (Fig. 6C). These results implicated the cc-tip as the potential target of the crosslink.

To assess the importance of residues in this section of the TL for DksA function, we used RNAP variants containing β' M932A, F935A, or H936A (β' R933A RNAP was not recovered for unknown reasons). β' M932A RNAP was

transcriptionally inactive, as expected from previous studies (Zhang et al. 2010). However, RNAPs containing β' F935A or H936A purified efficiently and retained transcription initiation activity. β' H936A RNAP responded to DksA similarly to wild-type RNAP, but β' F935A RNAP was severely defective in responding to DksA (Fig. 6D). We conclude that β' F935, an amino acid near the position of the R933-Bpa crosslink to DksA, plays an important role in DksA function.

To further determine whether the TL domain is in close proximity to the DksA coiled-coil tip, cysteines were introduced into β' at F935 and at the residues in the DksA tip most critical for function, D74 or A76. Under oxidizing conditions, D74C-DksA and β' F935C formed a complex that migrated to the position of β'-DksA, consistent with formation of a disulfide bond, indicating the two cysteines were within ~2Å of each other (Fig. 6E, lane 7). When β-mercaptoethanol (BME) was added to the SDS-loading solution, the complex was no longer observed (Fig. 6E, lane 8). D74C-DksA did not form a disulfide with wild-type RNAP (lane 5), and β' F935C did not form a disulfide with wild-type DksA (lane 2). As expected, R125C (a cysteine substitution in the globular domain of DksA) did not form a disulfide with β' F935C RNAP (lane 10) or with wild-type RNAP (lane 9). A disulfide bond was also obtained with β' F935C and A76C-DksA (Fig. S5). These data suggest that the cc-tip of DksA, at or very near residues D74 and A76, interacts directly with the β' TL at or near residue F935. β' F935C RNAP transcribed efficiently, but consistent with the importance of the DksA tip-TL interaction for DksA function and with the effect of the β' F935A substitution (Fig. 6D), it did not respond to DksA (Fig. 6F).

Role of the TL-DksA tip interaction in the mechanism of DksA action

RNAP containing a deletion of the β' TL does not respond to DksA (Rutherford et al. 2009). To address whether the TL-DksA tip interaction is important for DksA function because it contributes to the overall affinity of DksA for RNAP or rather because it affects a later step in the mechanism, we measured the effects of the TL deletion on DksA affinity using the Fe^{2+}-mediated cleavage assay (Lennon et al. 2009). *E. coli* RNAP contains the lineage-specific insertion β'i6 (residues 943-1130) in the TL (residues 926-942 and 1131-1143). Therefore, the TL deletion construct lacked β'i6 as well (ΔTL/Δi6). We compared the effects of a β'i6 deletion alone (Δi6) with the effects of the ΔTL/Δi6 deletion. The apparent binding constants for DksA were about the same for the Δi6 and ΔTL/Δi6 RNAPs (~100 nM; Fig. 7A, 7B) and for wild-type RNAP (Lennon et al. 2009; Lee et al. 2012), suggesting that the interaction of the TL domain with the DksA cc-tip is not a major determinant of the overall affinity of DksA for RNAP. Likewise, tip mutants D74A-DksA and A76T-DksA eliminate DksA function without affecting the overall affinity of DksA for RNAP (Lee et al. 2012).

It has been predicted that a double substitution in the TL, β'TL L930P/T931P/Δi6 (referred to below as β'TL LT→PP) blocks folding of the trigger helices (Vassylyev et al. 2007; Zhang et al. 2010). Because the TL is needed for nucleotide addition, we used a filter-binding assay to assess DksA function with this RNAP, because this assay measures open complex formation or decay without requiring RNA formation. Consistent with our previous results (Rutherford

et al. 2009), DksA reduced the half-life of the control Δi6 promoter complexes ~4.5-fold (from 18 min to 4 min; Fig. 7C). In contrast, β'TL LT→PP RNAP formed a 9-fold less stable promoter complex than the control, Δi6 RNAP (~2 min versus 18 min half-life), and the mutant complex was completely resistant to the effects of DksA (Figs. 7D).

RNAP with a different 2-residue substitution in the TL, β'G938A/G939A/Δi6 (referred to below as β'TL GG →AA) is hypersensitive to DksA (Rutherford et al. 2009). The β'TL GG→AA substitution is predicted to restrict the turn between the two trigger helices, locking the enzyme in a partially folded state (Vassylyev et al. 2007; Zhang et al. 2010). Thus, both the β'TL GG→AA and the β'TL LT→PP substitutions block full-folding of the trigger helices but are proposed to favor different TL conformations (Toulokhonov et al. 2007; Zhang et al. 2010). Although neither the β'TL GG →AA nor the β'TL LT→ PP substitutions affected the apparent affinity of DksA for RNAP, consistent with the lack of an effect of the β'ΔTL on overall DksA binding (Figs. 7A, 7B), our results suggest that the conformation of the TL has a major impact on DksA function. In our HADDOCK model, β' TL in the fully-folded (TH) conformation (from Vassylyev et al. 2007) appeared closer to the DksA tip than β' with an unfolded TL (models created using the Align function of PyMol; see above), and the model with a partially-folded (TL/TH) conformation was even closer to the DksA tip (Fig. 7E) (see Discussion).

Discussion

Conformational state of the TL and DksA function

To understand how DksA functions, we took biochemical and genetic approaches to probe its interactions with RNAP. Our results demonstrate that DksA binds in the RNAP secondary channel (as predicted previously), that residues in the N-terminal (and possibly C-terminal) segment of the globular domain of DksA and both helices of the DksA coiled-coil bind to the rim helices, and that the cc-tip of DksA interacts directly with the TL domain of β'. We prepose that this latter interaction explains the precise sequence requirement of the DksA tip residues for DksA function.

Our evidence-based model of the DksA-RNAP complex was made possible by introduction of the non-natural amino acid Bpa at precise positions in DksA. The model placed the DksA tip near the TL domain, but not close enough to explain the disulfide bond formed between D74C or A76C in the DksA tip and β' F935C in the TL. However, in the models containing the fully-folded conformation of the TL (trigger helices, TH) or the partially-folded conformation of the TL trapped by the antibiotic streptolydigin (TL/TH) (Vassylyev et al. 2007; see also Fig. 7E), it appears that β'TL residue F935 is considerably closer to the DksA coiled-coil tip, within a few angstroms of D74 and A76, than when the TL is in the unfolded conformation.

Although there is strong evidence that trigger loop-trigger helix transitions occur during transcription elongation (Vassylyev et al. 2007; Miropolskaya et al. 2010), there is little or no information about whether such transitions occur during

transcription initiation. DksA's effect on transcription initiation was eliminated either by deletion of the TL or by substitutions in the TL predicted to favor its unfolded form (β'TL LT→PP RNAP). In contrast, DksA function was enhanced by substitutions in the TL predicted to favor a partially-folded state (β'TL GG→AA RNAP) (Rutherford et al. 2009). Taken together, our data suggest that the favored conformation for DksA activity is a partially-folded state of the TL. One model for DksA function is that it traps the TL in this partially-folded state, that the β'TL GG →AA substitutions facilitate this TL conformation, and that the β'TL LT→PP substitutions preclude the interaction of the DksA cc-tip with the TL. Because the β'TL LT→PP RNAP forms a short-lived promoter complex in the absence of DksA (Fig. 7D), an alternative explanation for the observed resistance of the β'TL LT→PP complex to DksA is that the substitutions result in the same change in RNAP conformation that results from DksA interaction, precluding a further effect of DksA.

DksA disfavors contacts between RNAP and promoter DNA downstream of the transcription start site, inhibiting progression of initiation past the closed complex stage in negatively-regulated promoters (Rutherford et al. 2009). If DksA primarily occupies RP_C, and DksA functions by capturing a TL/TH form of the TL domain, then it seems reasonable that the TL/TH interaction with the tip of DksA takes place in RP_C. Substitutions in RNAP that bypass the DksA requirement were found predominantly in the TL, BH, and in 4 of the 5 switch regions of RNAP (Rutherford et al. 2009). These genetic results suggested a model in which the 60Å long BH (β' 769-804), which forms a three-helix bundle with the

TL, might link DksA to promoter DNA and the RNAP switch regions (Rutherford et al. 2009). By affecting the switch regions allosterically, DksA could affect DNA loading, DNA unwinding, and clamp closure following RP_C formation (see Chakraborty et al. 2012). Thus, we propose that the DksA cc-tip interaction with the TL is a critical step in DksA function.

Comparisons of DksA with other secondary channel binding factors

A number of small proteins, including DksA, GreA, GreB, Rnk, Gfh1, TraR, and TFIIS, bind in the RNAP secondary channel and affect transcription in different ways. Although there are high resolution structures of TFIIS with RNAP (Kettenberger et al. 2003) and of Gfh1 with RNAP (Tagami et al. 2010) and a lower resolution structure of GreB with RNAP (Opalka et al. 2003), details about interactions between RNAP and the other factors can only be inferred. Most of the different factors bear little sequence similarity to each other, and our data suggest that even very subtle changes in amino acid sequence can have dramatic effects on the outcomes of the interactions of the factors with RNAP (Lee et al. 2012). Thus, conclusions based only on structural analogy are of limited utility.

DksA has specific functions that are not shared by the other secondary channel binding proteins. Its most explored effects are on the control of transcription initiation, although it clearly also affects elongation, DNA repair, and RNAP-DNA polymerase collisions (Trautinger et al. 2005; Tehranchi et al. 2010). No DksA homologs are present in the thermophilic organisms from which we

have high resolution structural information of RNAP, and *E. coli* DksA does not affect *Thermus* RNAP like it does *E. coli* RNAP (T. Gaal, C.E.V., and R.L.G., data not shown).

Like DksA, GreB binds in the secondary channel of RNAP (Opalka et al. 2003), can inhibit *rrnB* P1 promoter activity (when overproduced to DksA-like levels in vivo or provided at the same concentration as DksA in vitro; Rutherford et al. 2007), and requires a conserved acidic tip residue for function (D74 for DksA, E44 for GreB). However, DksA and GreB do not share all functions. Gre factors can cleave RNA in backtracked transcription complexes (Laptenko et al. 2003), but DksA cannot (Furman et al. 2012), and DksA can positively regulate transcription from a number of promoters directly but Gre factors cannot (Rutherford et al. 2007). Recently, we reported that DksA residue D74 is required for both positive and negative transcription regulation and that changing the corresponding residue in either GreA or GreB (E44) to an aspartate resulted in acquisition of partial DksA function in positive control (Lee et al. 2012). Taken as a whole, it is likely that DksA and the Gre factors interact with RNAP in a similar manner overall, but that the network of interactions that DksA makes with RNAP, including its interaction with the TL, is finely-tuned for its specific role in transcription initiation.

An x-ray structure of the *T. thermophilus* secondary channel binding protein Gfh1 was reported in complex with *T. thermophilus* RNAP (Tagami et al. 2010). In the co-crystal, RNAP is in a so-called "ratcheted state", the TL is unstructured, and it appears that Gfh1 would clash with the TL if the TL were in

its fully-folded conformation. Tagami and coworkers proposed that Gfh1 alters the equilibrium between different conformations of RNAP (the ratcheted versus non-ratcheted states) in which the orientation between two rigid body modules (the core and shelf) is altered. The ratcheted conformation is proposed to affect RNAP function by widening the main and secondary channels, thereby weakening interactions with DNA. We recently found that ppGpp binds at the junction of the core and shelf modules (W.R. and R.L.G., in preparation), potentially explaining how ppGpp and DksA might work together. Thus, identification of the interactions between DksA and RNAP has implications not only for understanding the mechanism of DksA function, but also for how ppGpp and DksA might work together to control transcription.

In the ratcheted *T. thermophilus* RNAP-Gfh1 complex, Gfh1 appears to extend further into the secondary channel than DksA in our non-ratcheted *E. co*li RNAP model. To examine whether DksA might extend further into the secondary channel in the ratcheted state of RNAP, we created a HADDOCK model using *E. coli* DksA and the ratcheted form of *T. thermophilus* RNAP from the Gfh1 co-crystal (Tagami et al. 2010). The model of the ratcheted *T. thermophilus* RNAP - *E. coli* DksA complex is similar to the RNAP - DksA complex shown in Fig. 3. However, the tip region of DksA in this model does extend further into the channel than with the non-ratcheted form of the enzyme, albeit not as far as the tip of Gfh1 in the ratcheted RNAP complex (Fig. S6). In the ratcheted RNAP - DksA complex, the coiled-coil tip of DksA would be predicted to interact with a partially-folded TL, but there would be a greater potential for a clash with the

fully-folded form of the TL. DksA is slightly rotated and further from the RH compared to what was observed in our non-ratcheted RNAP-DksA model (Fig. S6). Whether the *E. coli* enzyme has a ratcheted state and whether DksA binds to it remain to be determined.

Prospect

In summary, our work provides the first evidence-based model for the DksA-RNAP complex. We suggest that a DksA cc-tip interaction with a partially-folded state of the TL occurs in RP_C and is a critical component of the network of interactions that is required for the specific effects of DksA on transcription initiation. Determining how the DksA tip-TL interaction contributes to inhibiting the progression from the closed to the open complex and whether this interaction is limited to specific steps in the transcription mechanism remain challenges for the future.

Materials and Methods

Bacterial strains and plasmids

E. coli strains and plasmids are listed in the supplement (Table S1). Plasmids were constructed using standard procedures for cloning and mutagenesis and were verified by DNA sequencing.

Expression and purification of Bpa-substituted proteins.

Plasmids containing genes encoding RNAP subunits or DksA were engineered with a stop codon (UAG) at the position chosen for benzoyl-phenylalanine (Bpa) incorporation (see Table S1 for plasmids and strains). Single UAG substitutions were constructed at multiple positions within each of the 5 sections of DksA (N-terminal helix, helix 1, coiled-coil tip, helix 2, and C-terminal helix). We chose positions based on the structure of DksA (Perederina et al. 2004) that were likely to be surface-exposed, unlikely to participate in protein folding, and unlikely to be required for DksA function based on previous information (Lee et al. 2012). DksA proteins containing Bpa at specific positions were over-expressed from the *araBAD* promoter on pRLG9502 and its derivatives in DH10B. pRLG9502 was constructed by inserting an NcoI-HindIII fragment coding for His_6-HMK-DksA from pRLG8150 into pRLG6640 (pBAD24/His_6-DksA; Webb et al. 1999) digested with NcoI and HindIII. We note that Bpa incorporation into DksA was unsuccessful when we used the T7 promoter in vector pET33 in BL21(DE3) cells.

Bpa-substituted RNAPs were over-expressed from the T7 promoter on a multi-subunit vector in BL21(DE3) cells induced with 1 mM IPTG. The

multisubunit RNAP vectors pIA299 and pIA423 do not contain the gene coding for ω, so strains used for overexpression of RNAP also carried pCDFω (Vrentas et al. 2005).

Strains for overexpression of Bpa-containing proteins also contained pSUPT/BpF which encodes a tRNA and tRNA-synthetase pair evolved for Bpa incorporation at UAG codons (Ryu and Schultz, 2006). Positive controls were always performed when establishing conditions for incorporation of Bpa into previously-untested positions or proteins.

For protein purification, co-transformation of the overexpression plasmid and the tRNA/tRNA synthetase plasmid was performed fresh for each experiment by electroporation, selecting for both antibiotic resistances. 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 37°C in LB with Bpa (1 mM), ampicillin (100 μg/ml), and chloramphenicol (25 μg/ml). Using a high 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, and an equal volume of 1 M HCl was added to produce a final pH of ~7.3. After 1 hr of growth with Bpa, expression was induced with 1 mM IPTG or 0.2% L-arabinose, depending on the expression system, and growth was continued for ~20 hr. Bpa incorporation was idiosyncratic: optimal temperature, inducer concentration, and time of expression varied with the protein of interest. Furthermore, UAG codons were not equally efficient at all positions, and it was critical to ensure that full-

length protein was made by stop codon readthrough and only when Bpa was present.

RNAPs with a His_6- or His_{10}-tag at the C-terminus of the β' subunit were purified using Ni-agarose and heparin affinity chromatography sequentially. For cell resuspension and Ni-agarose chromatography, Buffer A contained 40 mM Tris-Cl pH 7.9, 5 mM imidazole, and either 300 mM NaCl (for His_6 β' RNAPs) or 1M NaCl and 10% ethanol (for His_{10} β' RNAPs). Cell lysis buffer also included 10 µg/ml lysozyme and 23 µg/ml phenylmethylsullfonyl fluoride (PMSF). For culture volumes of 100 ml to 1 liter, cells were collected by centrifugation, resuspended in lysis buffer, and lysed by sonication. After centrifugation, the cleared lysate was added to pre-equilibrated Ni resin (~0.001 times the volume of the original culture), the column was washed with Buffer A, and the protein was eluted with Buffer A containing 150 mM imidazole. The eluate was diluted to 200 mM NaCl with TGED (Burgess and Jendrisak, 1975), bound to heparin resin (~0.001 times the volume of the original culture pre-equilibrated in TGED plus 200 mM NaCl), washed with the equilibration buffer, and RNAP was eluted with TGED plus 600 mM NaCl. RNAPs were concentrated using 0.5 ml Microcon centrifugal filtration units, diluted into storage buffer (50% glycerol, 10 mM Tris-Cl pH 7.9, 100 mM NaCl, 0.1 mM DTT, 0.1 mM EDTA) and stored at -20ºC. Protein concentrations were measured using the Bradford assay reagent (BioRad) using bovine serum albumen (BSA) as a standard.

Following purification of His-HMK-DksA from Ni-agarose as described above for RNAP, the His_6-tag was cleaved off using biotinylated-thrombin (EMD

Biosciences), leaving 8 residues (including the 5 residue HMK site) on the DksA N-terminus for labeling with ^{32}P. This N-terminal extension did not affect the activity of DksA (Lennon et al. 2009). Biotinylated thrombin was removed using streptavidin agarose (EMD Biosciences), and DksA was passed over Ni-NTA agarose again, collecting the flow-through to remove any impurities that co-eluted with DksA during the first purification step.

Expression and purification of mutant proteins not containing Bpa

Non-Bpa-containing DksAs and RNAPs were produced using overexpression vectors, as described previously (Paul et al. 2004; Rutherford et al. 2009; Lee et al. 2012) and in the Supplemental Materials and Methods.

Benzoyl-phenylalanine crosslink mapping

^{32}P-labeled HMK-DksA (0.01-1.0 μM) was incubated for 5-15 min at room temperature with RNAP (0.01-0.05 μM) in 20 mM HEPES pH 8, 20 mM NaCl, 10 mM $MgCl_2$ in a microfuge tube. The tube was placed directly on the filter of a 365 nm UV source for 3 min. The crosslinked region in β' was mapped by digestion with thrombin and SDS-PAGE. Where indicated, the β' subunit contained an engineered thrombin site at position 648. Procedures are described further in Results and Supplemental Materials and Methods.

Modeling the DksA-RNA Polymerase Complex

The structure of the complex between DksA and RNAP was modeled using HADDOCK version 2.0 (Dominguez et al. 2003; de Vries et al. 2007) starting

from PDB structures of DksA (1TJL; Perederina et al. 2004) and the model of *E. coli* core RNAP (3LU0; Opalka et al. 2010), which in turn is based on the high resolution *T. thermophilus* RNAP structure (Vassylyev et al. 2002). See Results and Supplemental Materials and Methods for further details.

Localized RNAP mediated Fe^{2+} cleavage of DksA

We employed a method in which hydroxyl radicals generated by replacement of Mg^{2+} with Fe^{2+} in the RNAP active center cleaved the coiled-coil tip of DksA (Lennon et al. 2009; Supplemental Materials and Methods).

Fe-EDTA protein-protein footprinting

Protein-protein footprinting was performed with hydroxyl radicals generated in solution by Fe-EDTA as described (Laptenko et al. 2003; see Supplemental Materials and Methods).

In vitro transcription

Single-round and multiple-round *in vitro* transcription reactions were performed essentially as described previously (Lee et al. 2012) with additional details in Supplemental Materials and Methods.

Disulfide formation

Cysteine residues were introduced into DksA and RNAP at positions predicted to be in close proximity in the complex. ^{32}P-labeled wild-type, D74C or R125C

DksA (5 nM) was incubated for 10 min at room temperature with wild-type or β' F935C RNAP (2.5 μM) in 20 mM Tris-HCl pH 7.9, 100 mM NaCl, 10 mM $MgCl_2$, 10 mM BME. Reactions were then exchanged into 20 mM Tris-HCl pH 7.9, 100 mM NaCl, 10 mM $MgCl_2$ with no BME using G-50 QuickSpin columns (GE Healthcare) and allowed to incubate at room temperature for 5 minutes. Reactions were stopped using 2-fold concentrated LDS loading buffer with or without BME as indicated in the figure. ^{32}P-labeled products were separated using 4-12% Bis-Tris NuPAGE gels and MOPS gel running buffer (Invitrogen).

RNAP-promoter complex lifetime assay

Competitor-resistant promoter complex lifetimes were measured by nitrocellulose filter retention in the presence of 10 μg/ml heparin with a DNA fragment containing the *lacUV5* promoter (endpoints -60 to +38 from pRLG4264) as described previously (Barker et al. 2001).

Acknowledgements

This work was supported by grants from the National Institutes of Health (R37 GM37048 to R.L.G. and a predoctoral fellowship to C.W.L.). Computational modeling resources were provided by the National Magnetic Resonance Facility at Madison, which is supported by NIH grants P41RR02301 (BRTP/ NCRR) and P41GM66326 (NIGMS), and a PyMol script for analysis and display of models was generously provided by L. Clos II. We also thank L. Li and D. Ma for mass-spectrometry of thrombin fragments, and J. Winkelman and other members of the Gourse lab, R. Mooney, J. Zhang, R. Landick, I. Artsimovitch, and P. Schultz for kindly providing strains, plasmids, and helpful comments.

References

Barker M, Gaal T, Josaitis C, Gourse R. 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.

Blankschien M, Lee J, Grace E, Lennon C, Halliday J, Ross W, Gourse R, Herman C. 2009. Super DksAs: Substitutions In DksA enhancing its effects on transcription initiation. *EMBO J* **28:** 1720–1731.

Bratton B, Mooney R, Weisshaar J. 2011. Spatial distribution and diffusive motion of RNA polymerase in live *Escherichia coli*. *J Bacteriol* **193:** 5138-5146.

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

Chakraborty A, Wang D, Ebright Y, Korlann Y, Kortkhonjia E, Kim T, Chowdhury S, Wigneshweraraj S, Irschik H, Jansen R, et al. 2012. Opening and closing of the bacterial RNA polymerase clamp. *Science* **337:** 591-595.

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

Chin J, Martin A, King D, Wang P, Schultz P. 2002. Addition of a photocrosslinking amino acid to the genetic code of *Escherichia coli. Proc Natl Acad Sci USA* **99:** 11020-11024.

de Vries S, van Dijk A, Krzeminski M, van Dijk M, Thureau A, Hsu V, Wassenaar T, Bonvin A. 2007. HADDOCK versus HADDOCK: New features and performance of HADDOCK 2.0 on the CAPRI targets. *Proteins* **69:** 726-733.

Dominguez C, Boelens R, Bonvin A. 2003. HADDOCK: A protein−protein docking approach based on biochemical or biophysical information. *J Am Chem Soc* **125:** 1731-1737.

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

Furman R, Sevostyanova A, Artsimovitch I. 2012. Transcription initiation factor DksA has diverse effects on RNA chain elongation. *Nucleic Acids Res* **40:** 3392-3402.

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

Kauer J, Erickson-Viitanen S, Wolfe H, DeGrado W. 1986. p-Benzoyl-L-phenylalanine, a new photoreactive amino acid. Photolabeling of calmodulin with a synthetic calmodulin binding peptide. *J Biol Chem* **261:** 10695-10700.

Kettenberger H, Armache K, Cramer P. 2003. Architecture of the RNA polymerase II TFIIS complex and implications for mRNA cleavage. *Cell* **114:** 347-357.

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

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

Lemke J, Durfee T, Gourse R. 2009. DksA and ppGpp directly regulate transcription of the *Escherichia coli* flagellar cascade. *Mol Microbiol* **74:** 1368-1379.

Lemke J, Sanchez-Vazquez P, Burgos H, Hedberg G, Ross W, Gourse R. 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, Gaal T, Ross W, Gourse R. 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.

Magnusson L, Gummesson B, Jaksimović P, Farewell A, Nyström T. 2007. Identical, independent, and opposing roles of ppGpp and DksA in *Escherichia coli. J Bacteriol* **189:** 5193-5202.

Miropolskaya N, Nikiforov V, Klimasauskas S, Artsimovitch I, Kulbachinskiy A. 2010. Modulation of RNA polymerase activity through the trigger loop folding. *Transcription* **1:**89-94.

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

Opalka N, Chlenov M, Chacon P, Rice W, Wriggers W, Darst, S. 2003. Structure and function of the transcription elongation factor GreB bound to bacterial RNA polymerase. *Cell* **114:** 335-345.

Opalka N, Brown J, Lane W, Twist K, Landick R, Asturias F, Darst S. 2010. Complete structural model of *Escherichia coli* RNA polymerase from a hybrid approach. *PLoS Biol* **8:** 9.

Paul B, Barker M, Ross W, Schneider D, Webb C, Foster J, Gourse R. 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, Barker M, Gourse R. 2005. DksA potentiates direct activation of amino acid promoters by ppGpp. *Proc Natl Acad Sci USA* **102:** 7823-7828.

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

Potrykus D, Vinella D, Murphy H, Szalewska-Palasz A, D'Ari R, Cashel M. 2006. Antagonistic regulation of *Escherichia coli* ribosomal RNA *rrnB* P1 promoter activity by GreA and DksA. *J Biol Chem* **281:** 15238-15248.

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

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

Saecker R, Record M, deHaseth P. 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.

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

Tehranchi A, Blankschien M, Zhang Y, Halliday J, Srivatsan A, Peng J, Herman C, Wang J. 2010. The transcription factor DksA prevents conflicts between DNA replication and transcription machinery. *Cell* **141:** 595-605.

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

Trautinger B, Jaktaji R, Ruskova E, Lloyd R. 2005. RNA polymerase modulators and DNA repair activities resolve conflicts between DNA replication and transcription. *Mol Cell* **19:** 247-258.

Traxler MF, Summers SM, Nguyen HT, Zacharia VM, Hightower GA, Smith JT, Conway T. 2008. The global, ppGpp-mediated stringent response to amino acid starvation in *Escherichia coli*. *Mol Microbiol* **68:** 1128-1148.

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

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

Vrentas C, Gaal T, Barker M, Rutherford S, Haugen S, Vassylyev D, Ross W, Gourse R. 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.

Webb C, Moreno M, Wilmes-Riesenberg M, Curtiss R, Foster J. 1999. Effects of DksA and ClpP protease on sigma S production and virulence in *Salmonella typhimurium*. *Mol Microbiol* **34:** 112-123.

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

Figures and legends

Figure 3.1. Benzoyl-phenylalanine (Bpa) incorporates site-specifically into DksA and crosslinks to the β' subunit of RNAP. (A) Structure of DksA (PDB 1TJL; Perederina et al. 2004). N- and C-termini and cc-tip region are indicated. Positions where Bpa substitutions were created are shown in spacefill. (B) Expression of DksA F69-Bpa. Protein lysates were examined on a 4-12% SDS gel stained with Coomassie Blue. Lane 1, purified H_6-HMK DksA (5 μg). Lane 2, DH10B containing pSUPT/BpF and pRLG9502 (DksA WT), no Bpa. Lane 3, DH10B containing pSUPT/BpF and pRLG9505 (DksA with UAG at codon 69), no Bpa. Lane 4, DH10B containing pSUPT/BpF and pRLG9505, 1 mM Bpa. Bpa was incorporated at UAG codons at S13, F69, E79, E81, N88, and E146 with similar efficiency. (C) ^{32}P-DksA E79-Bpa was incubated in the absence of RNAP (lanes 1,2), with WT RNAP (lanes 3,4) or with RNAP containing a β'-GFP fusion (lanes 5,6) and either exposed (lanes 2,4,6) or not exposed (lanes 1,3,5) to UV light. Samples were electrophoresed on a 4-12% SDS gel. Crosslinked β'-DksA (lane 4) and β'-GFP-DksA (lane 6) are indicated. (D) ^{32}P-DksA-Bpa proteins were crosslinked to β'-GFP RNAP (lanes 2-7), or ^{32}P DksA S13-Bpa was crosslinked to WT RNAP (lane 1), and samples were electrophoresed as in (C).

Figure 3.1

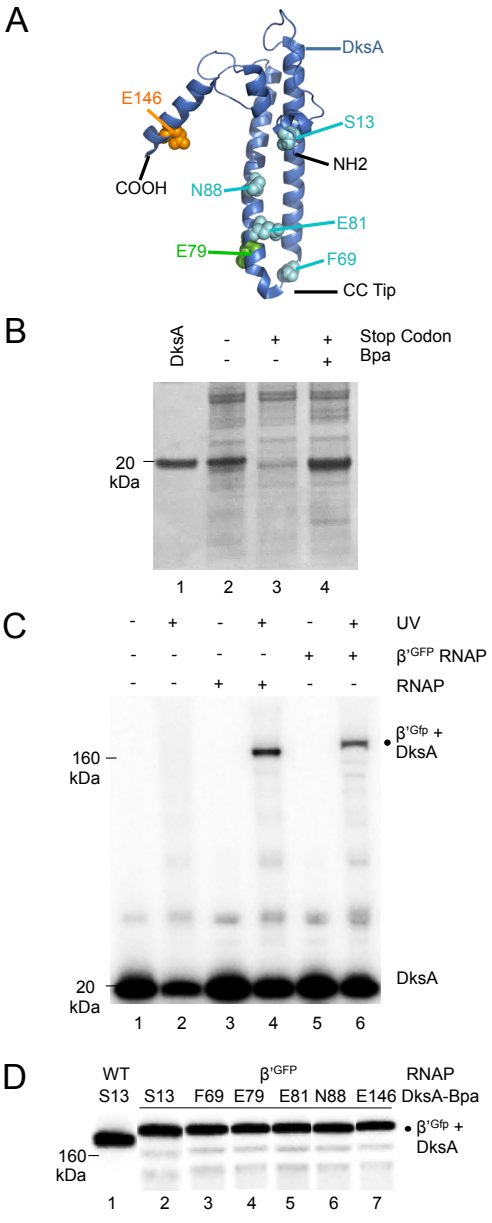

A
DksA
E146
S13
NH2
COOH
N88
E81
E79
F69
CC Tip
B
DksA
Stop Codon
Bpa
20 kDa
1 2 3 4
C
UV
β'GFP RNAP
RNAP
160 kDa
β'Gfp + DksA
20 kDa
DksA
1 2 3 4 5 6
D
WT
S13
β'GFP
S13 F69 E79 E81 N88 E146
RNAP DksA-Bpa
160 kDa
β'Gfp + DksA
1 2 3 4 5 6 7

Figure 3.2. Mapping DksA crosslinks within β'. ^{32}P HMK DksA proteins with S13-Bpa, F69-Bpa, E79-Bpa, E81-Bpa, N88-Bpa, or E146-Bpa were crosslinked to (A) wild-type RNAP (native thrombin site at residue ~900 of β') or to (B) RNAP with an additional thrombin site engineered at β' residue 648 (648^{Th}-RNAP), digested with thrombin, and electrophoresed on 4-12% SDS gels. The major thrombin cleavage product containing the crosslinked DksA in each reaction is indicated by a dot. Black dots, β' fragment 1-900. Orange dots, β' 901-1407. Green dot, β' 1-648. Cyan dots, β' 649-900. (C) Summary of DksA crosslinks with β' 648^{Th} thrombin fragments. DksA-Bpa residues are color-coded to match the β' fragments to which they crosslinked. (D) Model of *E. coli* core RNAP (PDB 3LU0; Opalka et al. 2010) with β' thrombin-fragment intervals containing crosslinks colored as in Figs. 2C and 1A.

Figure 3.2

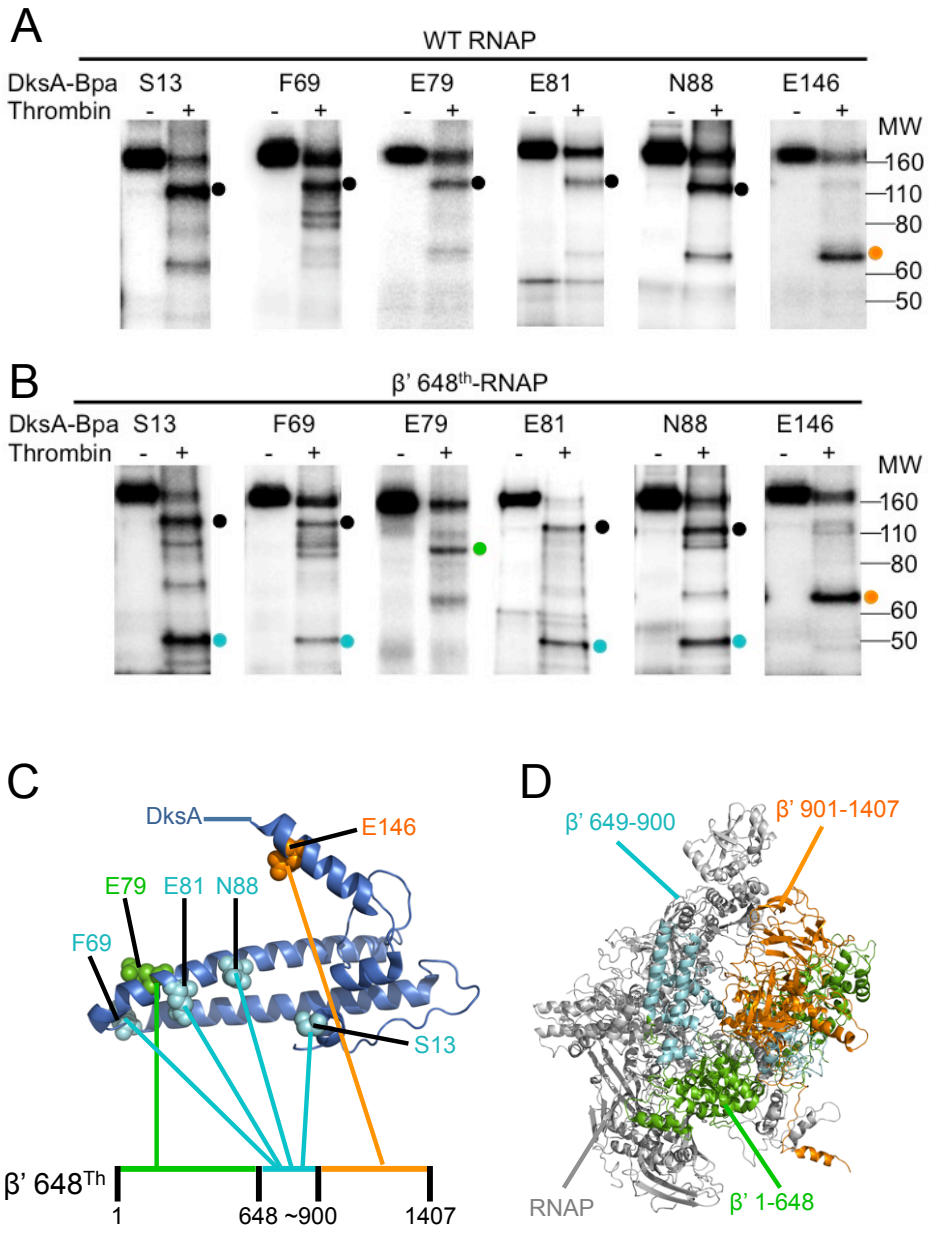

A
WT RNAP
DksA-Bpa S13 F69 E79 E81 N88 E146
Thrombin - + - + - + - + - + - +
MW
160
110
80
60
50
B
β' 648th-RNAP
DksA-Bpa S13 F69 E79 E81 N88 E146
Thrombin - + - + - + - + - + - +
MW
160
110
80
60
50
C
DksA
E146
E79 E81 N88
F69
S13
β' 648Th
1
648 ~900
1407
D
β' 649-900
β' 901-1407
RNAP
β' 1-648

Figure 3.3. HADDOCK model of *E. coli* DksA bound to *E. coli* core RNAP, with different segments of the complex colored for emphasis. Figs. 3A,B and 3C,D are different views of the RNAP-DksA complex. (A) Position of DksA relative to the RNAP subunits and the main and secondary channels. DksA, blue. αI subunit, green. αII, yellow. β, cyan. β', pink. ω, grey. (B) Position of DksA relative to specific features in RNAP. DksA, blue. DksA tip residues D74 and A76, red spacefill. Rim helices, yellow. Bridge helix, cyan.Trigger loop, green spacefill. Active site, orange spacefill. β'i6, pink. (C) Positions of segments of DksA relative to RNAP. The orientation has been rotated to look down the secondary channel. DksA residues 1-33 (N-terminal globular domain), green. DksA 34-65 (coiled-coil helix 1), blue. DksA 66-76 (tip region), red. DksA 77-107 (coiled-coil helix 2), cyan. DksA 108-151 (C-terminal globular domain), orange. RH, yellow. (D) Position of DksA relative to specific features of RNAP, colored as in (B), but from the angle shown in (C).

Figure 3.3

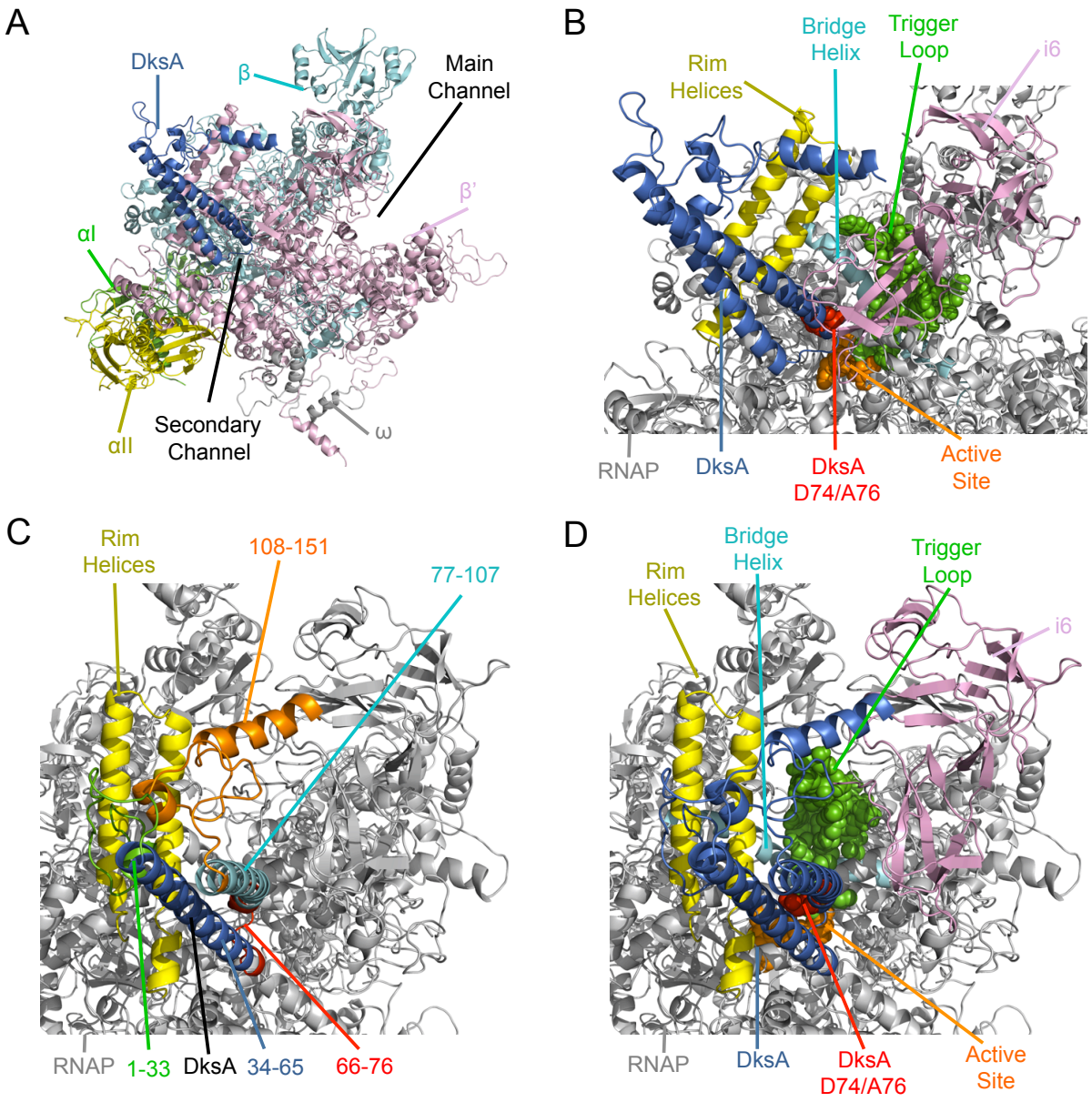

A
DksA
β
Main
Channel
β'
αI
αII
Secondary
Channel
ω
B
Rim
Helices
Bridge
Helix
Trigger
Loop
i6
RNAP
DksA
DksA
D74/A76
Active
Site
C
Rim
Helices
108-151
77-107
RNAP
1-33
DksA
34-65
66-76
D
Rim
Helices
Bridge
Helix
Trigger
Loop
i6
RNAP
DksA
DksA
D74/A76
Active
Site

Figure 3.4. Fe^{2+}-EDTA protein-protein footprints of RNAP with DksA and GreB. (A) RNAP was ^{32}P-labeled at a C-terminal HMK site on β'. Lane 1, CNBr-digested RNAP size markers. Lane 2, RNAP (140 nM), no Fe-EDTA. Lane 3, RNAP only (140 nM). Lane 4, RNAP (140 nM) plus DksA (3.5 μM). Lane 5, RNAP (140 nM) plus GreB (3.5 μM). Samples were electrophoresed on 8-16% acrylamide Tris-glycine SDS gels. Arrows indicate positions in/near the RH (3-5), TL (6-7), and β'i6 (8) regions of β' protected by DksA or GreB. Brackets represent regions with enhanced cleavages. (B) Graph of lane 3 (RNAP, grey line) vs 4 (RNAP plus DksA, red line) from gel shown in Fig. 4A. Arrows above panel indicate the positions of markers generated by CNBr cleavage at indicated positions in lane 1. (C) Graph of lanes 3 (RNAP, grey line) vs 5 (RNAP plus GreB, blue line). (D) Same samples as in Fig. 4A were electrophoresed on 8% gels to better resolve the active site region (AS; arrows 1,2) and the RH region (arrows 3-5). (E) Graph of lanes 3 (RNAP, grey) vs 4 (RNAP plus DksA, red). (F) Graph of lanes 3 (RNAP, grey) vs 4 (RNAP plus GreB, red).

Figure 3.4

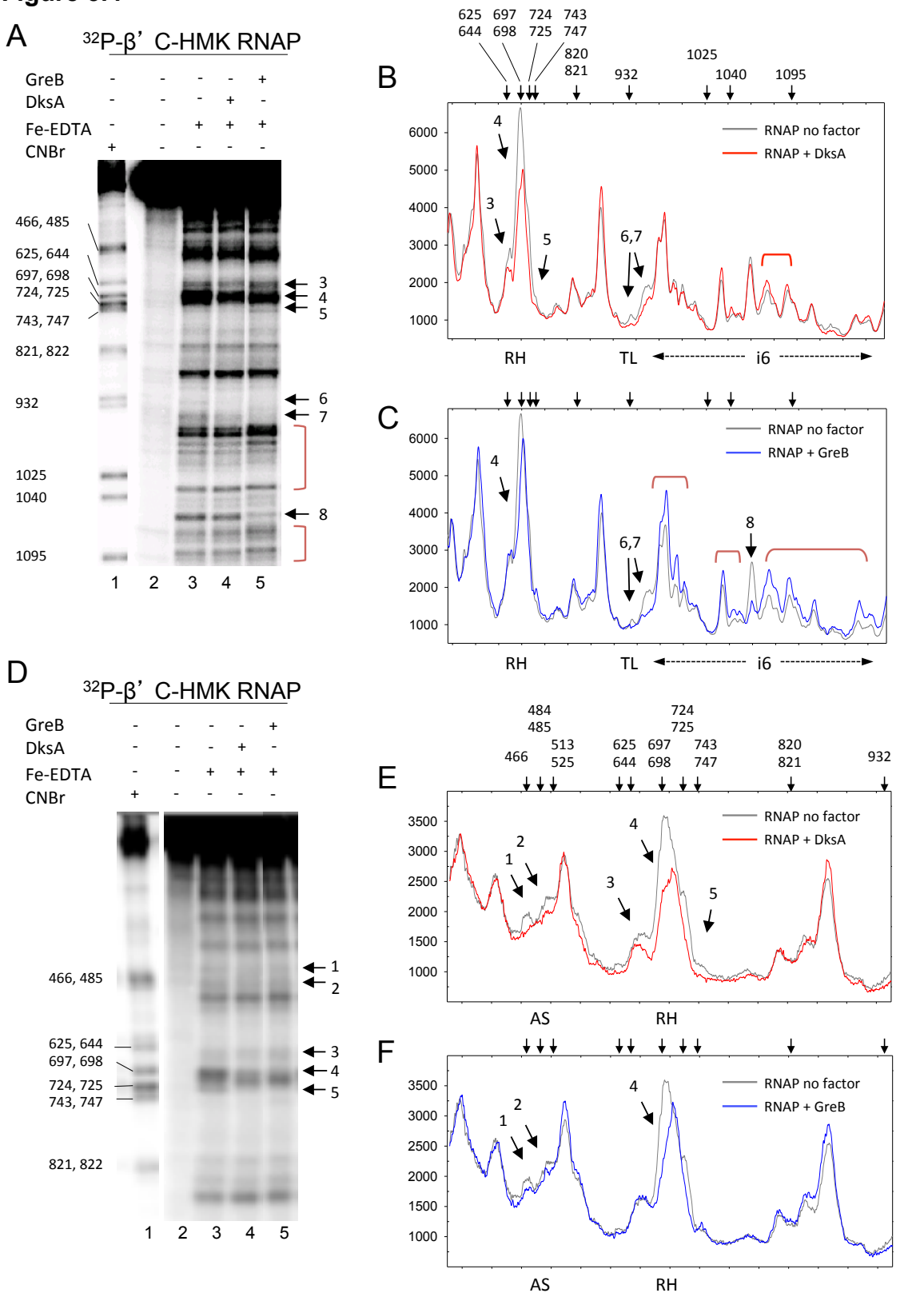

A
32P-β' C-HMK RNAP
GreB
DksA
Fe-EDTA
CNBr
466, 485
625, 644
697, 698
724, 725
743, 747
821, 822
932
1025
1040
1095
1 2 3 4 5
B
625 644
697 698
724 725
743 747
820 821
932
1025
1040
1095
RNAP no factor
RNAP + DksA
RH
TL
i6
C
RNAP no factor
RNAP + GreB
RH
TL
i6
D
32P-β' C-HMK RNAP
E
484 485
513 525
625 644
697 698
724 725
743 747
820 821
932
466
RNAP no factor
RNAP + DksA
AS
RH
F
RNAP no factor
RNAP + GreB
AS
RH

Figure 3.5. The β’ rim helices are required for DksA function and binding. (A) Lifetime of RNAP-*lacUV5* promoter complex, measured by filter retention in the presence or absence of 1.6 μM DksA. Plots show the fraction of promoter-RNAP complex remaining at times following heparin addition. WT, wild-type RNAP. ΔRH, RNAP lacking residues β' 645-718. (B) ΔRH RNAP does not bind DksA. Binding was measured by RNAP -localized Fe^{2+}-mediated cleavage of DksA (Lennon et al. 2009). (C) HADDOCK model illustrating the DksA interaction with RH. DksA, blue. RH, yellow. N88, red spacefill. Image was generated using PyMol with a surface transparency setting of 60%.

Figure 3.5

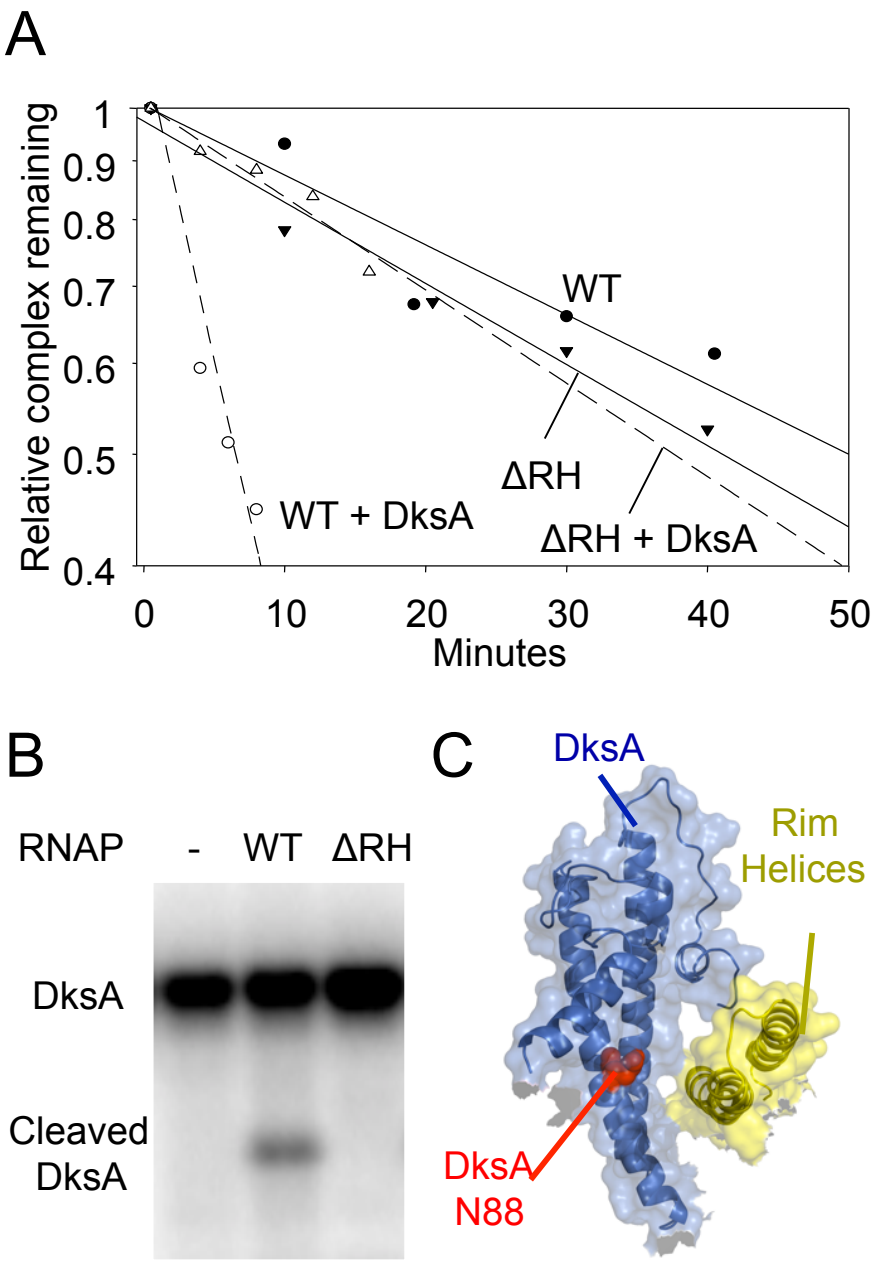

A
Relative complex remaining
1
0.9
0.8
0.7
0.6
0.5
0.4
0
10
20
30
40
50
Minutes
WT
ΔRH
ΔRH + DksA
WT + DksA
B
RNAP
-
WT
ΔRH
DksA
Cleaved DksA
C
DksA
Rim Helices
DksA N88

Figure 3.6. Interaction of the RNAP trigger loop (TL) with the DksA cc-tip is required for DksA function. (A) Residues where Bpa was inserted in the β' TL are shown in spacefill in HADDOCK model. The fully-folded trigger helices (TH) conformation of the TL was substituted for the unfolded TL domain (see Fig. 7). β'Q929, cyan. β' R933, orange. β'G1136, pink. DksA, blue. DksA cc-tip residue D74, red. (B) β' R933-Bpa crosslinks to ^{32}P-DksA. Crosslinked complexes were analyzed by SDS-PAGE as in Fig. 1. (C) D74N reduces the amount of DksA crosslinked to β'R933-Bpa. Crosslinking was measured at different RNAP concentrations. 2.5 and 5 μM were saturating. (D) β' TL substitution F935A RNAP reduces DksA function as measured by inhibition of *rrnB* P1 promoter activity. A plasmid containing *rrnB* P1 was transcribed with WT RNAP, β'F935A RNAP, or β'H936A RNAP at a range of DksA concentrations. Products were analyzed by gel electrophoresis and phosphorimaging, and the fraction of transcription relative to that without DksA is plotted for each enzyme. (E) DksA tip residue D74C forms a disulfide bond with TL β' F935C. ^{32}P-DksA (wild-type, D74C, or R125C) was incubated with RNAP (wild-type or β' F935C) under oxidizing (without BME) or reducing conditions (+ BME), and the products were examined by SDS-PAGE (see Supplemental Experimental Procedures). (F) β' F935C RNAP does not respond to DksA. Transcription was performed as in (C). Average transcription from *rrnB* P1 relative to that without DksA is indicated below the duplicate gel lanes.

Figure 3.6

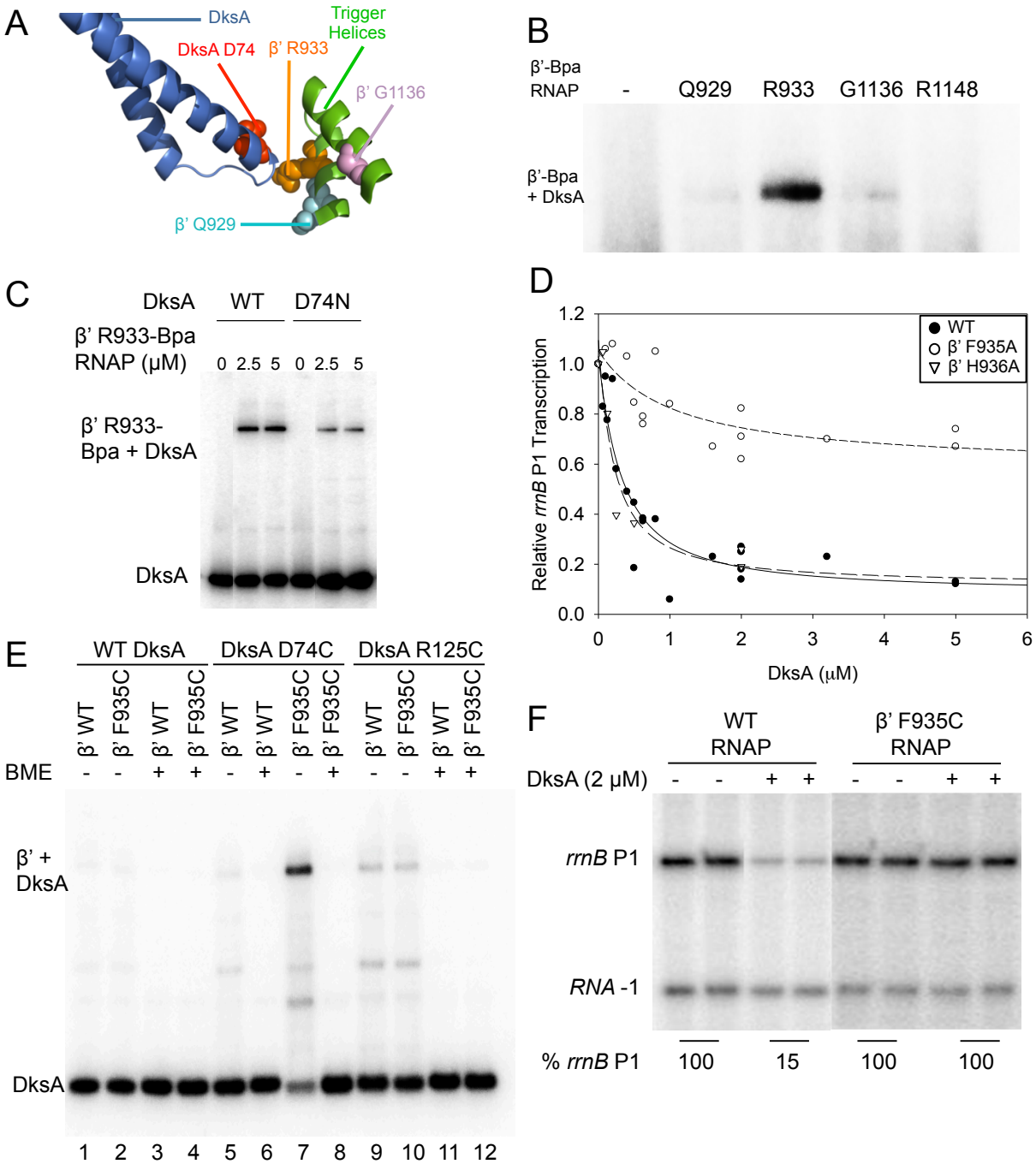
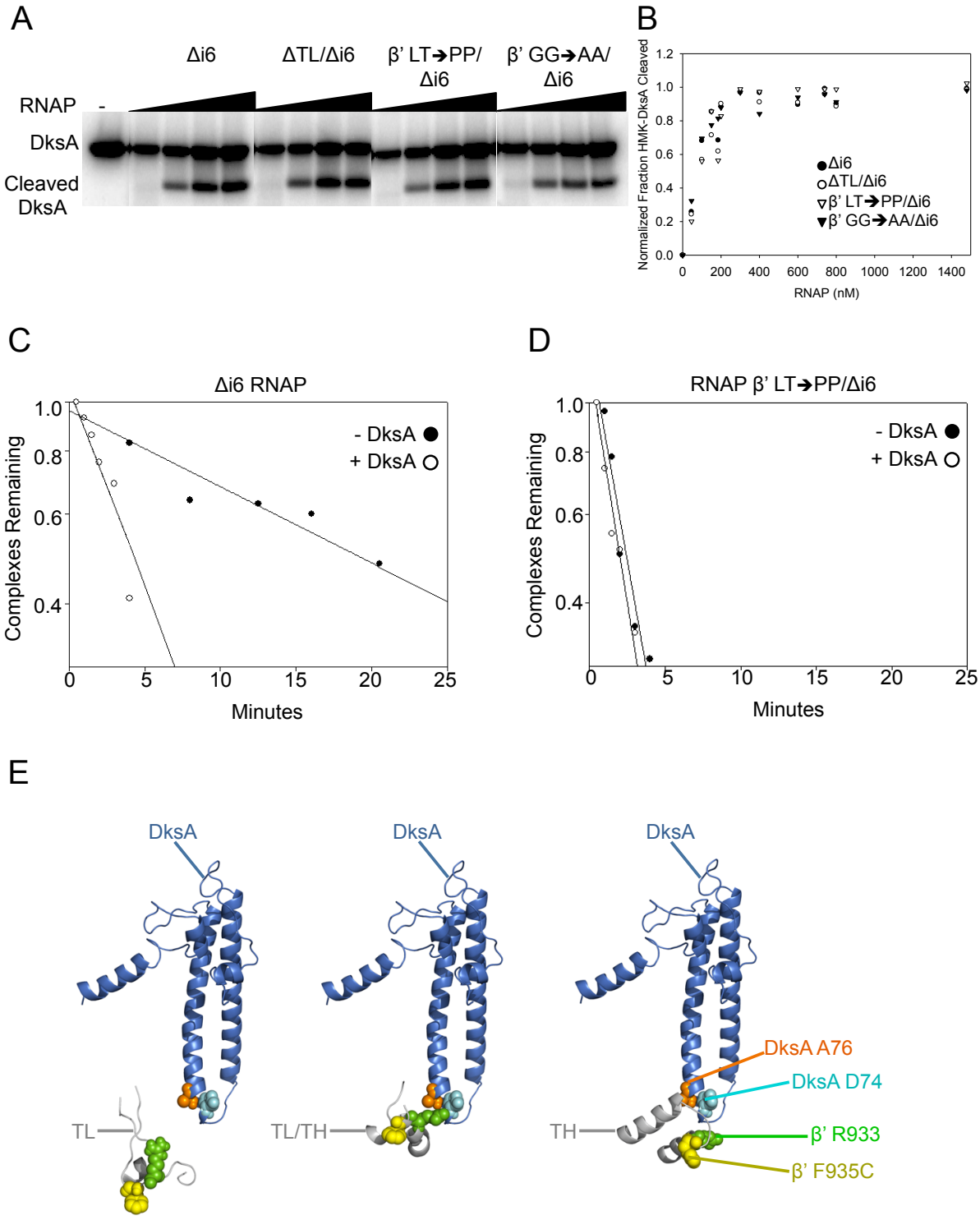

A
DksA
Trigger Helices
DksA D74
β' R933
β' G1136
β' Q929
B
β'-Bpa RNAP
-
Q929
R933
G1136
R1148
β'-Bpa + DksA
C
DksA
WT
D74N
β' R933-Bpa RNAP (μM)
0 2.5 5 0 2.5 5
β' R933-Bpa + DksA
DksA
D
Relative rrnB P1 Transcription
1.2
1.0
0.8
0.6
0.4
0.2
0.0
0 1 2 3 4 5 6
DksA (μM)
WT
β' F935A
β' H936A
E
WT DksA
DksA D74C
DksA R125C
β' WT
β' F935C
BME
β' +
DksA
DksA
1 2 3 4 5 6 7 8 9 10 11 12
F
WT RNAP
β' F935C RNAP
DksA (2 μM)
rrnB P1
RNA -1
% rrnB P1
100
15
100
100

Figure 3.7. The β' trigger loop (TL) is required for DksA function but not for DksA binding to RNAP. (A) Binding of DksA to Δi6 RNAP (deletion of β' residues 943-1130), ΔTL/Δi6 RNAP (deletion of β' residues 931-1137), β' 930,931 LT→PP/Δi6 RNAP (i.e. β'TL LT→PP RNAP), and β' 938,939 GG→AA/Δi6 RNAP (i.e. β'TL GG →AA) was measured by localized Fe^{2+}-mediated cleavage of DksA at different RNAP concentrations (Lennon et al. 2009). (B) Binding curves for TL variants shown in (A). Overall apparent binding affinities of mutant RNAPs for DksA were the same within error (K_d ~100 ± 15 nM for each RNAP). (C) Effect of DksA on Δi6 RNAP. Lifetimes of *lacUV5* promoter-RNAP complexes were determined in the presence/absence of 1.6 μM DksA. Each data point represents the fraction of ^{32}P-promoter DNA retained on nitrocellulose filters at times following competitor addition (Rutherford et al. 2009). (D) β'TL LT→PP RNAP is resistant to the effects of DksA. Half-life measurements were performed as in (C). (E) Models of DksA interactions with the TL domain in 3 different conformations: unfolded trigger loop (TL; see Fig. 3), partially-folded trigger loop (TL/TH; Vassylyev et al. 2007; PDB: 2PPB), and folded trigger helices (TH; Vassylyev et al. 2007; PDB: 2O5J). The entire β' subunit with the partially-folded (TL/TH) or fully-folded (TH) structures was used to replace β' from the HADDOCK model using the Align function of PyMol. Only the TL domain and DksA are shown. DksA, blue. DksA residue D74, cyan spacefill. DksA A76, orange spacefill. TL domain, grey. β' R933, green spacefill. β' F935, yellow spacefill.

Figure 3.7


A
Δi6
ΔTL/Δi6
β' LT➔PP/ Δi6
β' GG➔AA/ Δi6
RNAP
-
DksA
Cleaved DksA
B
Normalized Fraction HMK-DksA Cleaved
RNAP (nM)
●Δi6
○ΔTL/Δi6
▽β' LT➔PP/Δi6
▼β' GG➔AA/Δi6
C
Δi6 RNAP
Complexes Remaining
Minutes
- DksA ●
+ DksA ○
D
RNAP β' LT➔PP/Δi6
Complexes Remaining
Minutes
- DksA ●
+ DksA ○
E
DksA
TL
TL/TH
TH
DksA A76
DksA D74
β' R933
β' F935C

Chapter 3 Supplemental Materials

Direct interactions between the coiled-coil tip of DksA and the trigger loop of RNA polymerase mediate transcriptional regulation

Christopher W. Lennon, Wilma Ross, Stephen Martin-Tumasz, Innokenti Toulokhonov, Catherine E. Vrentas, Steven T. Rutherford, Jeong-Hyun Lee, Samuel E. Butcher, and Richard L. Gourse

I. Expanded Materials and Methods

Plasmid construction

β'Δ645-718 ΩAla4 (β'ΔRH) was constructed by site-directed PCR mutagenesis of plasmid pRL663 (Artsimovitch et al. 2003) with two fully complementary 66 nt oligonucleotides (5'-GGTGCATCTGTTGGTATCGATGACATGGCGGCAGCGGCGTTCAACAGCATCTACATGATGGCCGAC-3' and 5'-GTCGGCCATCATGTAGATGCTGTTGAACGCCGCTGCCGCCATGTCATCGATACCAACAGATGCACC-3') that flanked the deleted fragment. A linker encoding four amino acids (AlaAlaAlaAla) was introduced in place of the deleted residues. A fragment between unique Bsm1 and XhoI sites containing the deleted region was sequenced, excised, and cloned into the overexpression plasmid pIA423 (Toulokhonov et al. 2007) digested with BsmI and XhoI to obtain the β′ΔRH overexpressor (pRLG7844).

Benzoyl-phenylalanine crosslink mapping

The Bpa substitutions at residues S13, F69, E79, E81, N88, and E146 formed specific crosslinks with RNAP, whereas the Bpa substitutions at residues A18, R87, or K94 did not crosslink, the Bpa substitution at R52 severely reduced binding of DksA to RNAP, and the Bpa substitutions at positions D64 or Q148 resulted in non-specific crosslinking to RNAP (data not shown).

The crosslinked complexes were digested with thrombin. The crosslinked complex (10 µl) was mixed with 8 µl 8M urea and incubated for 10 min at 37°C. 17 µl H_20, 4 µl 10-fold concentrated thrombin cleavage buffer, and 1 µl (1U)

thrombin (Novagen) were added for 2 hr at room temperature. ^{32}P-labeled products were separated on 4-12% Bis-Tris NuPAGE gels and MOPS gel running buffer (Invitrogen). To eliminate the radioactive signal derived from non-crosslinked DksA or DksA dimers from overwhelming the signal of interest, the bottom section of the gel was cut off and discarded to remove products below ~40 kDa (by comparison with molecular weight markers). No crosslinks were observed without Bpa-substituted DksA (data not shown).

Where indicated, RNAP contained an additional thrombin site at amino acid residue 648 in β', created by introducing substitutions coding for E648R and K649G into the expression vector pIA299. Based on the wild-type β' amino acid sequence, thrombin was predicted to cleave at residues 860 and 1140 (http://web.expasy.org/peptide_cutter/). However, cleavage at the 1140 site was very inefficient, and the electrophoretic mobilities of fragments produced from thrombin treatment of RNAP with additional engineered thrombin sites at β' 525 or 566 suggested an actual cleavage site at ~900. Mutations constructed in the proline and arginine codons in the predicted thrombin site at position 860 had no effect on thrombin cleavage, further indicating that the site at 860 was not the site cleaved by thrombin (W.R., unpublished data). This conclusion was further strengthened by mass-spectrometry of gel-isolated thrombin fragments 525-900 and 566-900 which identified tryptic peptides derived from β' region 860-900 (D. Ma, W. R., R.L.G., and L. Li, unpublished results).

Localized RNAP-mediated Fe^{2+} cleavage of DksA

Ten µl reactions containing RNAP at the concentrations indicated, 1 to 5 nM ^{32}P-labeled HMK-DksA, 0.1 mg/ml BSA, 20 mM HEPES pH 7.9, and 20 mM NaCl were incubated for 10 min at 30°C and mixed with 1 µl each freshly-made 500 µM $(NH_4)_2Fe(SO_4)_2$ and 100 mM DTT (Laptenko et al. 2003, Lennon et al. 2009). After 10 min at 30°C, reactions were terminated with 2-fold concentrated LDS gel loading solution (Invitrogen). ^{32}P-labeled products were separated using 4-12% Bis-Tris NuPAGE gels and MES gel running buffer (Invitrogen).

Modeling the DksA-RNA Polymerase Complex

The structure of the complex between DksA and RNA Polymerase was modeled using HADDOCK version 2.0 (Dominguez et al. 2003; de Vries et al. 2007) starting from PDB structures of DksA (1TJL; Perederina et al. 2004) and the model of *E. coli* core RNAP (3LU0; Opalka et al. 2010), which in turn is based on the high resolution *T. thermophilus* RNAP structure (Vassylyev et al. 2002). HADDOCK is a molecular docking program that incorporates biochemical data, energetic data, and shape complementarity. In an initial round of rigid body docking, 240 structures of the DksA/RNAP complex were calculated. Based on their overall energies and the number of distance restraint violations, the best 24 structures underwent semi-flexible simulated annealing, followed by refinement in an explicit water solvent.

During docking, HADDOCK can be configured for varying degrees of flexibility within the molecules to allow for an energetically favorable intermolecular interface while maintaining the overall structures. DksA residues not in secondary structural elements (residues 19-31, 68-73, 109-122, and 129-

135) were classified as fully flexible (allowing movement at all stages of the docking), whereas all other residues in DksA and RNAP were classified as semi-flexible (allowing movement only during final refinement). To maintain the structure of DksA, inter-residue N-N, CA-N, and CA-CA distances of less than 10 Å in α-helical regions of the crystal structure were restrained to be that distance ± 0.1 Å (2,619 restraints, generated using an in-house script). These intra-DksA restraints had 20% of the weight of the other restraints. In addition, artificial restraints were included to maintain the structure of the RNAP active site.

For intermolecular distance restraints, HADDOCK ambiguous interaction restraints (AIRs) were used, derived from the crosslinking data. AIRs with an upper distance bound of 6 Å were assigned for DksA residues that crosslinked to β'. These AIRs effectively require any atom of the specific residue in DksA to be contacting any atom of any solvent accessible residue in the region of β'. (Solvent accessible residues are defined as having an accessibility score > 10 when the entire RNAP structure was analyzed in the "Accessible Molecular Surface" module of WHAT IF (Rodriguez et al. 1998). Specifically, AIRs were created between residues S13, F69, E81, and N88 in DksA and β' residues 649-900; between DksA E79 and β' 1-648; and between DksA E146 and β' 901-1407. In addition, hydroxyl radicals generated by Fe^{2+} at the position of the catalytic Mg^{2+} in RNAP cleave the loop between the two α-helices of the DksA coiled-coil (Fig. S2), indicating the coiled-coil is close to the active site of β'. Therefore, an AIR with an upper distance bound of 10 Å was assigned between DksA residues 71-75 and the active site magnesium.

An ensemble of all 24 models is shown in Fig. S3. All 24 models placed DksA in the secondary channel, and 22 of the 24 models were almost identical. The lowest energy structure with no restraint violations was selected as the final model of the DksA/RNAP complex. The lowest energy structure was in the class of very similar 22 models.

To generate the HADDOCK model of DksA bound to the ratcheted state of RNAP (Fig. S6), we followed the same procedure described above, except that the structure of *T. thermophilus* RNAP from the Gfh1 co-crystal was used (3AOI; Tagami et al. 2010) instead of the model of *E. coli* core RNAP (3LU0; Opalka et al. 2010), and Gfh1 and nucleic acid were removed before docking DksA on RNAP.

Fe-EDTA protein-protein footprinting

HMK-tagged β' core RNAP (10 µg) was ^{32}P-labeled as described (Loizos et al. 1999) and purified with a G-50 QuickSpin column (GE Healthcare) into 20 mM HEPES pH 7.5, 50 mM NaCl, 1 mM $MgCl_2$. Core RNAP (0.6 µg) containing ^{32}P-labeled-β' was mixed with DksA (0.6 µg) and BSA (7 µg) in 7 µl. One µl each of fresh 10-fold concentrated Fe–EDTA (10 mM $(NH_4)_2Fe(SO_4)_2$ and 20 mM EDTA), 200 mM ascorbate and 10 mM H_2O_2 were added, the reaction was incubated for 15 min at 25°C, and then the reaction was terminated with 3.4 µl of 4-fold concentrated SDS gel loading buffer containing 50 mM BME. ^{32}P-β' molecular weight markers were prepared by subjecting radiolabeled HMK-β' to chemical cleavage at methionine residues by CNBr (Mustaev et al. 1996; Markovtsov et al.

1996; Laptenko et al. 2003). The samples were separated on 20 × 40 cm long 8% or 8-16% acrylamide, Tris–Glycine, SDS–PAGE gels.

In vitro transcription

Single-round transcription reactions were performed essentially as described previously (Lee et al. 2012) at 30°C in transcription buffer (40 mM Tris-HCl pH 7.9, 50 mM NaCl, 10 mM $MgCl_2$, 1 mM DTT, 0.1 mg/ml BSA) using a supercoiled plasmid containing *rrnB* P1 promoter sequences -88 to +50 with respect to the transcription start site (pRLG1616). RNAP holoenzyme (10 nM) and plasmid DNA (50 ng) were mixed, and transcription was initiated by addition of NTPs (including ^{32}P-UTP). Heparin (to 10 µg/ml) was added with the NTPs to prevent rebinding of RNAP once it left the promoter. Multiple-round transcription was performed at 30°C in transcription buffer (40 mM Tris-HCl pH 7.9, 150 mM NaCl, 10 mM $MgCl_2$, 1 mM DTT, 0.1 mg/mL BSA), pRLG1616, RNAP holoenzyme, and NTPs (^{32}P-UTP). Reactions were initiated by addition of RNAP. For single or multiple round experiments, transcription was allowed to proceed for 10 min before addition of 2-fold concentrated stop solution. DksA was added at the concentrations indicated in the figure legends. RNA transcripts were separated on 5.5% acrylamide-7M urea gels.

II. Supplemental Table 1

Strain	Description	Source
DH10B	For expression of genes under control of *araBAD* promoter	Novagen
BL21(DE3)	For expression of genes under control of T7 promoter	Invitrogen
RL1314	GFP fused to β' C-terminus	Bratton et al. 2011

Plasmid	Description	Source
pRLG1616	pRLG770 containing *rrnB* P1 (-88 to +50)	Rutherford et al. 2007
pRLG4264	pSL6 containing *lacUV5* (-60 to +38)	Barker et al. 2001
pRLG7067	His_6-DksA from T7 promoter	Paul et al. 2004
pRLG8150	His_6-HMK-DksA from T7 promoter	Lennon et al. 2009
pRLG12050	pRLG8150 with DksA-D74C	This work
pRLG9139	pRLG8150 with DksA-R125C	This work
pRLG6640	pBAD24 with His_6-DksA from *araBAD* promoter	Webb et al. 1999
pRLG9502	pRLG6640 with His_6-HMK-DksA	This work
pRLG9532	pRLG9502 with His_6-HMK-DksA S13-UAG	This work
pRLG9505	pRLG9502 with His_6-HMK-DksA F69-UAG	This work
pRLG11413	pRLG9502 with His_6-HMK-DksA E79-UAG	This work
pRLG10832	pRLG9502 with His_6-HMK-DksA E81-UAG	This work
pRLG10835	pRLG9502 with His_6-HMK-DksA N88-UAG	This work
pRLG9543	pRLG9502 with His_6-HMK-DksA E146-UAG	This work
pRLG9535	pRLG9502 with His_6-HMK-DksA A18-UAG	This work
pRLG9504	pRLG9502 with His_6-HMK-DksA R52-UAG	This work
pRLG10830	pRLG9502 with His_6-HMK-DksA D64-UAG	This work
pRLG9541	pRLG9502 with His_6-HMK-DksA R87-UAG	This work
pRLG9542	pRLG9502 with His_6-HMK-DksA K94-UAG	This work
pRLG9544	pRLG9502 with His_6-HMK-DksA Q148-UAG	This work
pSUPT/BpF	Expresses BPA synthetase/tRNA	Ryu and Schultz, 2006.
pIA299	Expresses α_2, β, β'-C-His_6 from T7 promoter	Artsimovitch et al. 2003
pCDFω	Expresses ω from T7 promoter	Vrentas et al. 2005
pIA423	Expresses α_2, β, β'-C- CBP from T7 promoter	Artsimovitch et al. 2003
pRLG7844	pIA423 with β' ΔRH (β'Δ645-718 Ω ala4)	This work
pRLG7668	pIA299 with β' K598A/K599A	Vrentas et al. 2008
pRLG7673	pIA299 with β' R731A	Vrentas et al. 2008
pRLG10314	pIA299 with thrombin site at β' 648 (E648R, K649G)	This work
pRLG11415	pIA299 with β' Q929-UAG	This work
pRLG11416	pIA299 with β' R933-UAG	This work
pRLG11418	pIA299 with β' G1136-UAG	This work
pRLG12773	pIA900 with β' G1148-UAG	J. Winkelman, unpublished
pIA900	α_2, β, β'-C-His_{10}, ω from T7 promoter	I. Artsimovitch, unpublished

pRLG11447	pIA900 with β' F935C	This work
pIA331	α_2, β, β'-C-CPB Δi6 (β'Δ943-1130) from T7 promoter	Artsimovitch et al. 2003
pIT121	α_2, β, β'-C-CPB with β' M932A from T7 promoter	Toulokhonov et al. 2007
pJZ14	α_2, β, β'-C-CPB with β' R933A from T7 promoter	Toulokhonov et al. 2007
pJZ15	α_2, β, β'-C-CPB with β' F935A from T7 promoter	Toulokhonov et al. 2007
pJZ16	α_2, β, β'-C-CPB with β' H936A from T7 promoter	Toulokhonov et al. 2007
pJZ9	α_2, β, β'-C-$His_{6,}$ ω, ΔTL/Δi6 (β' Δ931-1137) from T7 promoter	Toulokhonov et al. 2007
pRLG8155	α_2, β, β'-C-CPB HMK from T7 promoter	This work

III. Supplemental References

Artsimovitch I, Svetlov V, Murakami KS, 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.

Barker M, Gaal T, Josaitis C, Gourse R. 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.

Bratton B, Mooney R, Weisshaar J. 2011. Spatial distribution and diffusive motion of RNA polymerase in live *Escherichia coli*. *J Bacteriol* **193:** 5138-5146.

de Vries S, van Dijk A, Krzeminski M, van Dijk M, Thureau A, Hsu V, Wassenaar T, Bonvin A. 2007. HADDOCK versus HADDOCK: New features and performance of HADDOCK 2.0 on the CAPRI targets. *Proteins* **69:** 726-733.

Dominguez C, Boelens R, Bonvin A. 2003. HADDOCK: A protein−protein docking approach based on biochemical or biophysical information. *J Am Chem Soc* **125:** 1731-1737.

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

Lee J, Lennon C, Ross W, Gourse R. 2012. Role of the coiled-coil tip of *Escherichia coli* DksA in promoter control. *J Mol Biol* **416:** 503-17.

Lennon C, Gaal T, Ross W, Gourse R. 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.

Markovtsov V, Mustaev A, Goldfarb A. 1996. Protein-RNA interactions in the active center of transcription elongation complex. *Proc Natl Acad Sci USA.* **93:** 3221-3226.

Mustaev A, Kozlov M, Markovtsov V, Zaychikov E, Denissova L, Goldfarb A. 1997. Modular organization of the catalytic center of RNA polymerase. *Proc Natl Acad Sci USA* **94:** 6641-6645.

Opalka N, Brown J, Lane W, Twist K, Landick R, Asturias F, Darst S. 2010. Complete structural model of *Escherichia coli* RNA polymerase from a hybrid approach. *PLoS Biol* **8:** 9.

Paul B, Barker M, Ross W, Schneider D, Webb C, Foster J, Gourse R. 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.

Rodriguez R, Chinea G, Lopez N, Pons T, Vriend G. 1998. Homology modeling, model and software evaluation: three related resources. *Bioinformatics* **14:** 523-528.

Rutherford S, Lemke J, Vrentas C, Gaal T, Ross W, Gourse R. 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.

Ryu Y, Schultz PG. 2006. Efficient incorporation of unnatural amino acids into proteins of *Escherichia coli*. *Nat Methods* **3:** 263-265.

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

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

Vassylyev DG, Vassylyeva M, Perederina A, Tahirov TH, Artsimovitch I. 2007. Structural basis for transcription elongation by bacterial RNA polymerase. *Nature* **448:** 157-162.

Vrentas C, Gaal T, Ross W, Ebright R, Gourse R. 2005. Response of RNA polymerase to ppGpp: requirement of the Ω subunit and relief of this requirement by DksA. *Genes Dev* **19:** 2378-2387.

Vrentas C, Gaal T, Barker M, Rutherford S, Haugen S, Vassylyev D, Ross W, Gourse 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.

Webb C, Moreno M, Wilmes-Riesenberg M, Curtiss R, Foster J. 1999. Effects of DksA and ClpP protease on sigma S production and virulence in *Salmonella typhimurium*. *Mol Microbiol* **34:** 112-123.

IV. Supplemental Figures and Legends

Figure 3.S1. DksA proteins containing Bpa substitutions are functional. Transcription from *rrnB* P1 promoter (-88 to +50) on pRLG1616 was measured in vitro (Lee et al. 2012). (A) Effects of 2 μM wild-type DksA, E81 DksA-Bpa, and N88 DksA-Bpa. Percent transcription from *rrnB* P1 ± DksA (single-round assay) is indicated below the gel lanes (average of duplicate reactions). (B) Effects of 10 μM WT DksA and 5 μM E79 DksA-Bpa in multiple-round transcription (duplicate lanes). (C) Effects of wild-type DksA and S13 DksA-Bpa, F69 DksA-Bpa, and E146 DksA-Bpa (single-round assay) at the DksA concentrations indicated. Effects of the DksA-Bpa variants on transcription should be compared only to that of the wild-type DksA in the same experiment, not between experiments. Although we conclude that all the variants are active in promoter-specific inhibition of transcription, different DksA preparations could in theory have significantly different specific activities.

Figure 3.S1

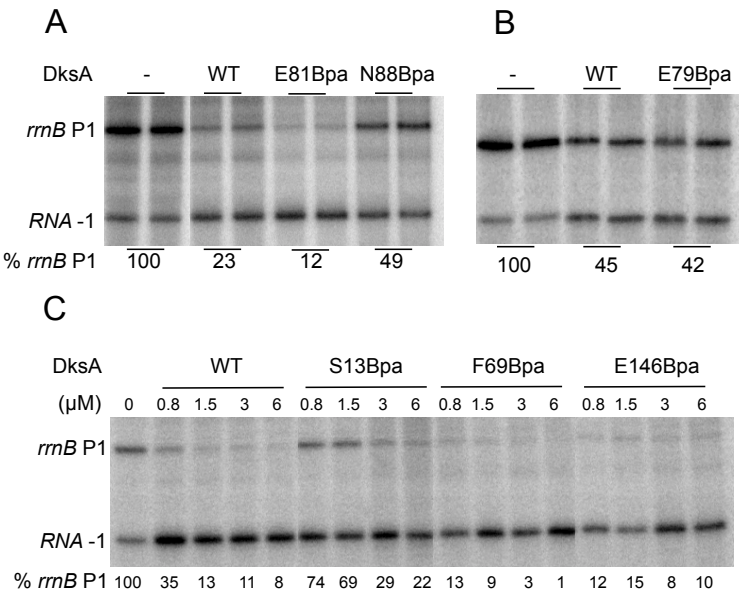

A
DksA
-
WT
E81Bpa
N88Bpa
rrnB P1
RNA -1
% rrnB P1
100
23
12
49
B
-
WT
E79Bpa
100
45
42
C
DksA
WT
S13Bpa
F69Bpa
E146Bpa
(μM)
0
0.8
1.5
3
6
rrnB P1
RNA -1
% rrnB P1
100
35
13
11
8
74
69
29
22
13
9
3
1
12
15
8
10

Figure 3.S2. The DksA cc-tip region is the target of localized RNAP-mediated Fe^{2+} cleavage (see Supplemental Experimental Procedures; Lennon et al. 2009). SDS-PAGE of ^{32}P-HMK-DksA variants. The N-terminal ^{32}P- HMK-DksA product generated by Fe^{2+}-mediated cleavage migrates similarly to the DksA variant terminated at residue 73. Lane 1, ^{32}P-HMK-DksA and RNAP without Fe^{2+}. Lane 2, ^{32}P-HMK-DksA and RNAP with Fe^{2+}. Lane 3, purified ^{32}P-HMK 1-73 DksA (i.e. genetically truncated at position 73). Lane 4, purified ^{32}P-HMK 1-85 DksA (i.e. genetically truncated at position 85**).**

Figure 3.S2

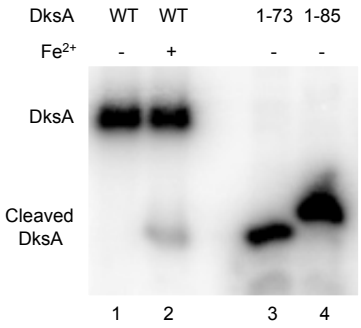


Figure 3.S3. Overlay of the 24 refined HADDOCK models. Figs. S3A and B and Figs. S3C and D are two different views of the RNAP-DksA complex. DksA is shown in blue, and the rim helices are shown in yellow. The 24 refined models are shown in Fig. S3A and S3C, and the lowest energy model is shown in Figs. S3B and D. Twenty-two of the 24 models were very similar to the lowest energy model used in Fig. 3 and consistent with the biochemical and genetic experiments. The 2 other models still placed DksA in the secondary channel, but were inconsistent with the data supporting binding of DksA to the RH.

Figure 3.S3

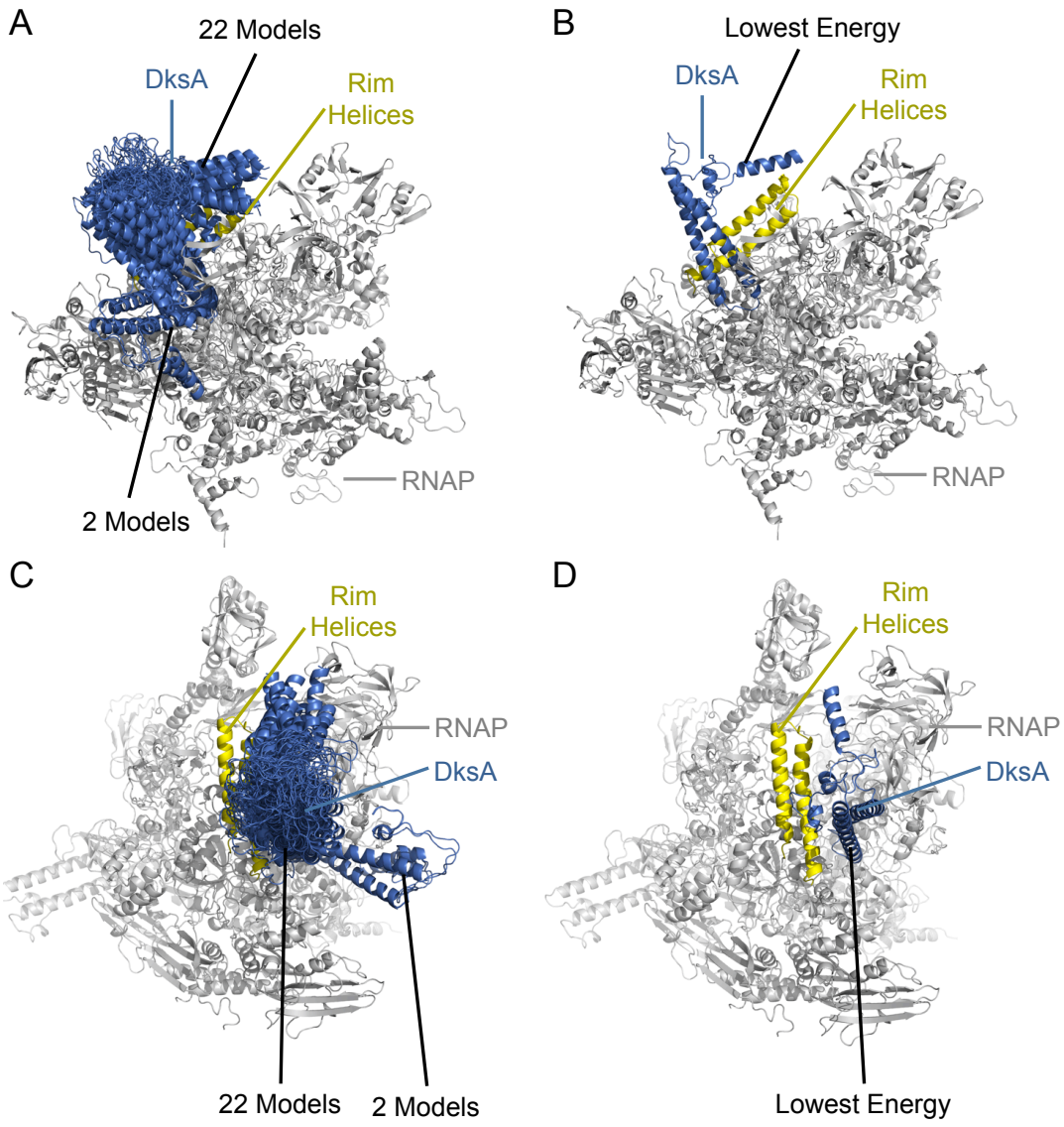

A
22 Models
DksA
Rim
Helices
RNAP
2 Models
B
Lowest Energy
DksA
Rim
Helices
RNAP
C
Rim
Helices
RNAP
DksA
22 Models
2 Models
D
Rim
Helices
RNAP
DksA
Lowest Energy

Figure 3.S4. β' residues K598/K599, and R731 play a role in binding of DksA to RNAP and DksA function. (A) HADDOCK model of DksA-RNAP complex illustrating the positions of the 3 substitutions in spacefill. DksA, blue. Rim helices, yellow. β' K598 and K599, red spacefill. β' R731, green spacefill. (B) Representative binding curves for DksA and wild-type RNAP, β' K598A/K599A RNAP, and R731A RNAP, determined by localized Fe^{2+}-mediated cleavage (Lennon et al. 2009). β' K598A/K599A and R731A reduce the affinity of DksA for RNAP from ~100 nM to 200-300 nM. (C) Effect of DksA and (D) effect of GreB on single-round transcription with wild-type, β' K598A/K599A, or R731A RNAP and a plasmid containing the *rrnB* P1 promoter (-88 to +50) and the RNA-1 promoter (Lee et al. 2012). Percent transcription ± DksA is indicated below gel lanes.

Figure 3.S4

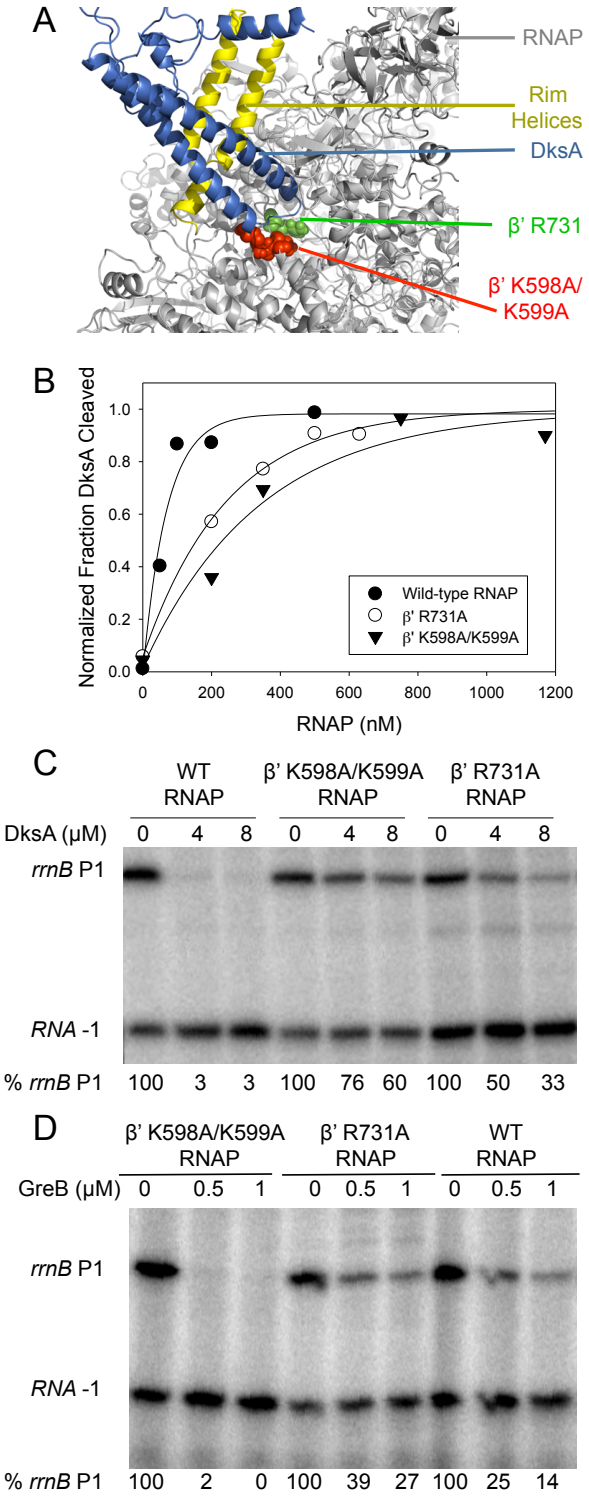

A
RNAP
Rim Helices
DksA
β' R731
β' K598A/ K599A
B
Normalized Fraction DksA Cleaved
1.0
0.8
0.6
0.4
0.2
0.0
0
200
400
600
800
1000
1200
RNAP (nM)
Wild-type RNAP
β' R731A
β' K598A/K599A
C
WT RNAP
β' K598A/K599A RNAP
β' R731A RNAP
DksA (μM) 0 4 8 0 4 8 0 4 8
rrnB P1
RNA -1
% rrnB P1 100 3 3 100 76 60 100 50 33
D
β' K598A/K599A RNAP
β' R731A RNAP
WT RNAP
GreB (μM) 0 0.5 1 0 0.5 1 0 0.5 1
rrnB P1
RNA -1
% rrnB P1 100 2 0 100 39 27 100 25 14

Figure 3.S5. DksA-A76C forms a disulfide bond with trigger loop residue β' F935C. Wild-type, A76C, or R125C variants of DksA were labeled on their N-termini with ^{32}P, incubated with β' F935C RNAP under oxidizing conditions, and the products were examined by SDS-PAGE (See Supplemental Experimental Procedures and Fig. 6D). Lane 1, Wild-type DksA. Lane 2, DksA-A76C. Lane 3, R125C-DksA. Positions of β'-DksA and DksA-DksA complexes and free DksA are indicated.

Figure 3.S5

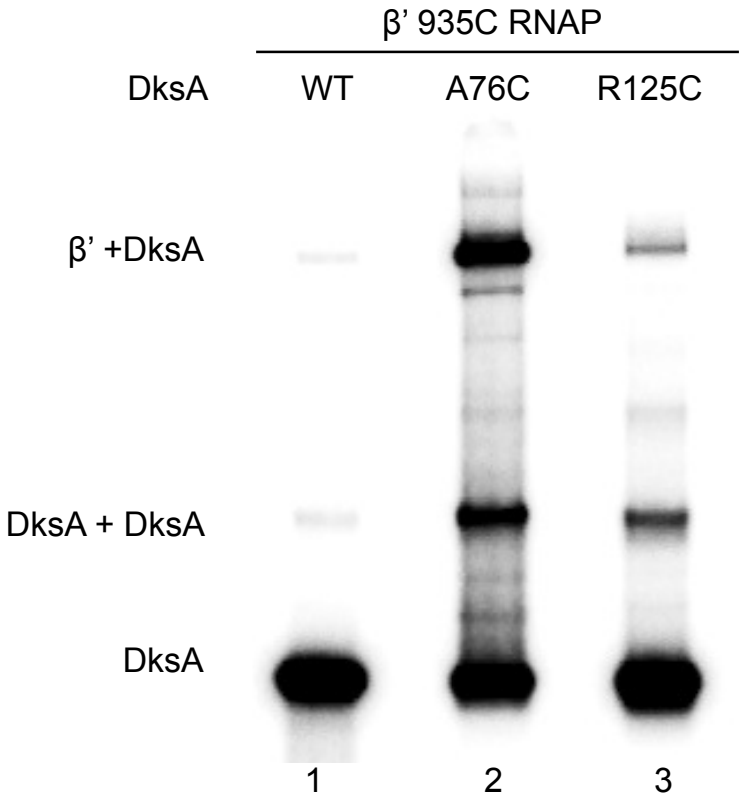
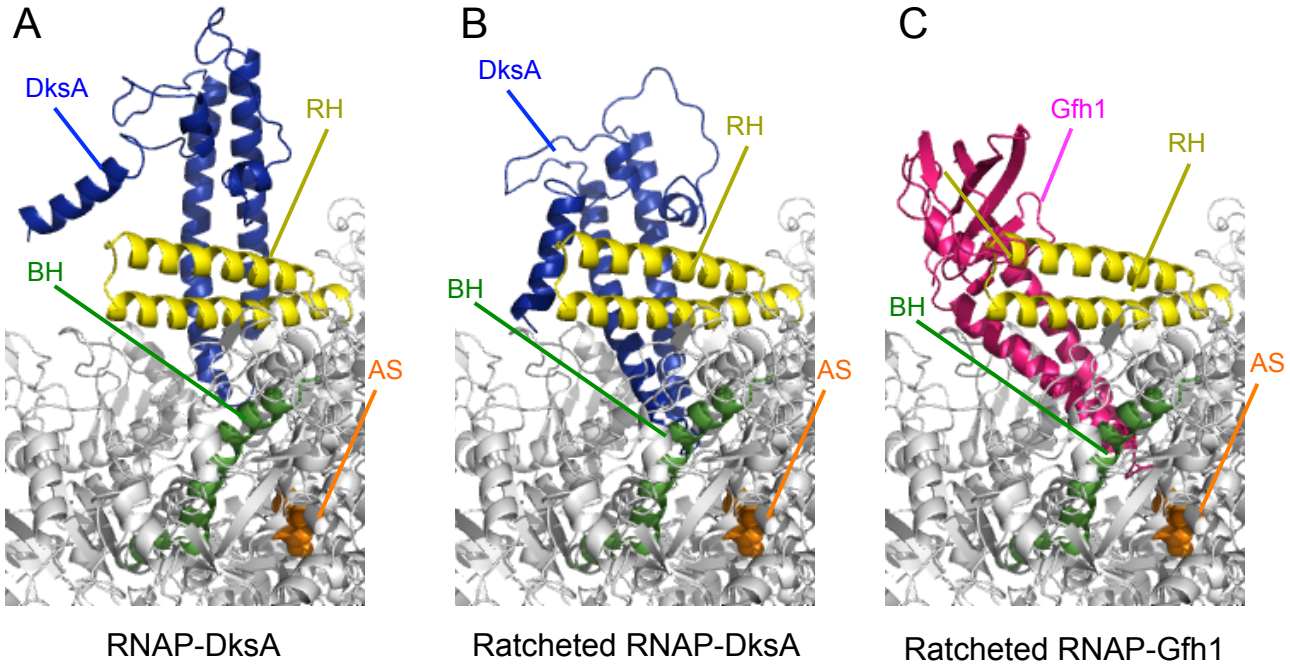


Figure 3.S6: Docking DksA on the ratcheted conformation of RNAP. (A) Non-ratcheted RNAP-DksA HADDOCK model, as in Fig. 3. (B) Ratcheted RNAP-DksA HADDOCK model. (C) Ratcheted RNAP-Gfh1 complex (Tagami et al. 2010). In the ratcheted RNAP-DksA model (B), the coiled-coil tip of DksA extends further into the secondary channel, DksA is rotated slightly, and the distance of DksA from the rim helices is increased slightly compared to that in the RNAP-DksA model shown in (A). In the RNAP-Gfh1 structure (C), the coiled-coil tip of Gfh1 extends even further into the secondary channel than the tip of DksA as shown in (B). The RNAP-DksA model in (A) was generated using *E. coli* DksA and *E. coli* RNAP (Opalka et al. 2010). The ratcheted RNAP-DksA model in (B) was generated using *E. coli* DksA and ratcheted *T. thermophilus* RNAP (Tagami et al. 2010). Rim helices (RH, yellow), bridge helix (BH, green), DksA (blue), Gfh1 (magenta), active site aspartate residues (AS, β' D460, D462, D464; orange).

Figure 3.S6


A
DksA
RH
BH
AS
RNAP-DksA
B
DksA
RH
BH
AS
Ratcheted RNAP-DksA
C
Gfh1
RH
BH
AS
Ratcheted RNAP-Gfh1

Chapter 4

Genetic and biochemical dissection of the *Rhodobacter sphaeroides* DksA/ppGpp system

Here I present work examining *R. sphaeroides* RSP2654 (*E. coli* DksA homologue) both *in vivo* and *in vitro*. This project is a collaboration with the Donohue lab (UW-Madison) and will be submitted for publication in 2013 with the following author list: Christopher W. Lennon, Kimberly C. Lemmer, Jessica L. Irons, Max I. Sellman, Timothy J. Donohue, Richard L. Gourse and Wilma Ross. CWL and KCL will be co-first authors. Experiments used for figures 4.1, 4.2A, 4.3D and 4.3E were performed by KCL and MIS; Figures 4.3A and Figure 4.4A were performed by JLI. I performed all other experiments used in the figures for this chapter. Portions of the text relating to experiments performed by KCL and MIS were written by KCL.

Abstract

$DksA_{Ec}$ regulates transcription from hundreds of promoters in the gammaproteobacterium *E. coli*, alone and in concert with the alarmone ppGpp. Here, we analyze two genes annotated as DksA-like, RSP2654 and RSP0166, in the highly divergent alphaproteobacterium *R. sphaeroides*. RSP2654, which has sequence similarity to the coiled-coil region of $DksA_{Ec}$ but not the zinc-finger domain, is required for proper photosynthetic growth in *R. sphaeroides*, functions similarly to $DksA_{Ec}$ *in vivo* in *E. coli*, functions on *E. coli* RNAP and promoters, and directly destabilizes *R. sphaeroides* RNAP-promoter complexes alone and synergistically with ppGpp. Our work provides a basis for understanding the mechanism of DksA/ppGpp action in *R. sphaeroides* and other alphaprotobacteria, and it provides clues about $DksA_{Ec}$ structure/function relationships.

Introduction

Escherichia coli (gammaproteobacteria) and *Rhodobacter sphaeroides* (alphaproteobacteria) are distantly related bacteria that live in very different environmental and nutritional conditions. *E. coli* DksA ($DksA_{EC}$) is a crucial transcription factor regulating gene expression under non-optimal conditions. $DksA_{EC}$ is a non-classical transcription factor in that it binds only to RNAP and not to promoter DNA (Haugen et al. 2008, Lennon et al. 2012). Thus, $DksA_{Ec}$ has the potential to bind and influence RNA Polymerase (RNAP) at any promoter. $DksA_{Ec}$ destabilizes all transcription initiation complexes in concert with the alarmone ppGpp, but it only inhibits transcriptional output from intrinsically unstable RNAP-promoter DNA complexes. (Paul et al. 2004, Paul et al. 2005, Rutherford et al. 2009, Lee et al. 2012). In *E. coli*, RNAP-rRNA promoter complexes are especially unstable, such that there is a direct competition between addition of the transcript's initiating NTP and complex decay (Barker et al. 2001), making $DksA_{Ec}$ critical for the regulation rRNA expression (Paul et al. 2004). $DksA_{Ec}$ has two domains; a coiled-coil domain containing the residues D74 and A76 that are used for its promoter specific effects and a globular domain with a 4 cysteine zinc-finger motif (Perederina et al. 2004, Lee et al. 2012). *R. sphaeroides* encodes two potential DksA homologues, RSP0166 and RSP2654). One contains a domain homologous to the coiled-coil (RSP2654) and one a domain homologous to the zinc-finger (RSP0166). Here we seek to understand whether these *R. sphaeroides* protein(s) are functionally similar to $DksA_{Ec}$, if

these protein(s) employ a similar mechanism of transcriptional regulation, and if these protein(s) potentiate the effects of the alarmone ppGpp.

Results

R. sphaeroides encodes a DksA-like protein that is required for normal photosynthetic growth

R. sphaeroides contains two genes that are annotated to encode DksA-like proteins, RSP2654 and RSP0166. RSP2654 encodes a 158 amino acid protein that is 42% identical to $DksA_{Ec}$. The residues located at the tip of the coiled-coil fold, an aspartic acid and an alanine, that are critical to $DksA_{Ec}$ function (Lee et al., 2012) are conserved in RSP2654, but not the Cys-4 zinc finger. In contrast, the coiled-coil domain homology is not conserved in RSP0166, a 105 amino acid protein that is 35% identical to $DksA_{Ec}$, but this protein does contain the characteristic Cys-4 zinc finger motif found in DksA proteins (Perederina et al., 2004). In $DksA_{Ec}$, mutation of any of these 4 cysteine residues in $DksA_{Ec}$ abolishes function (Henard and Vazquez-Torres 2012, Paul et al., 2005). RSP0166 contains all 4 cysteines of this motif, but RSP2654 only contains one of these cysteines. Instead of a zinc-finger, RSP2654 contains a Cxx[S/T]-(x17)-[C/S/T]xxA motif characteristic of DksA2 proteins that do not bind to zinc (Blaby-Haas et al. 2011). DksA2 is a protein found in *P. aeruginosa* that retains DksA function in the absence of a bound zinc ion. *P. aeruginosa* also contains a typical $DksA_{Ec}$ homologue. Thus, *R. sphaeroides* appears to encode 2 DksA-like proteins, each with some but not all of the characteristics of the well characterized $DksA_{Ec}$.

In order to test the function of RSP2654 and RSP0166 in *R. sphaeroides*, we constructed mutant strains with markerless deletions of each open reading

frame (strains Δ*2654* and Δ*0166*). In response to decreasing oxygen tensions, *R. sphaeroides* synthesizes components of the photosynthetic apparatus including the light-harvesting pigment-protein complexes and the photosynthetic electron carriers (Zeilstra-Ryalls and Kaplan, 2004). A pale colony color phenotype is typically associated with a defect in photosynthetic pigment production or an inability to assemble functional photosynthetic pigment–protein complexes under inducing conditions of decreased oxygen tension (Davis et al. 1988, Eraso and Kaplan, 1994, Gomelsky and Kaplan, 1998). WT *R. sphaeroides* forms red-colored colonies when grown aerobically, as was also observed for the Δ*0166* mutant (Fig. 4.1A). In contrast the Δ*2654* mutant formed much paler colonies (Fig. 4.1A), suggesting it was defective in pigment production.

In order to test if photosynthetic-pigment protein complex assembly was impaired in the Δ*2654* mutant, we grew WT and Δ*2654* cells under low oxygen tension (0.5%) where pigment-protein complexes are normally assembled and inserted into the membrane. The visible spectra of intact cells show that, despite its altered colony pigmentation, Δ*2654* contains pigment-protein complexes (Fig. 4.1B). Quantitation of the levels of the pigment-protein complexes show that these are present at equivalent levels in Δ*2654* and WT cells (data not shown).

We also tested the function of the photosynthetic apparatus by assessing the ability of Δ*2654* cells to grow via photosynthesis. Surprisingly, when Δ*2654* was incubated under photosynthetic conditions (anaerobic with light) on agar plates, poor growth was observed (Fig. 4.1C) despite its ability to make pigment-protein complexes at low oxygen tensions. Photosynthetic growth of Δ*2654* was

rescued by complementation with an IPTG-inducible plasmid expressing RSP2654 (pIND-2654; see below), confirming that this growth defect was associated with the loss of this protein. Plating efficiency experiments demonstrated that Δ*2654* plated with 93-99% efficiency under photosynthetic conditions (data not shown), but it formed extremely small colonies (Fig. 4.1D), suggesting that loss of this protein results in a defect in photosynthetic growth rate. Thus, although *R. sphaeroides* RSP2654 is not required for the making of photosynthetic pigment-protein complexes, it is required for WT growth under anaerobic conditions in the light. Further studies will be needed to define the precise cause of this photosynthetic defect.

RSP2654 is necessary for normal utilization of exogenous amino acids

In *E. coli*, DksA acts with ppGpp to activate promoters for amino acid biosynthesis and transport (Paul et al., 2005), and cells deleted for *dksA* are unable to grow on media lacking amino acids (Paul et al. 2004, Lennon et al., 2009). Therefore, we asked if deletion of RSP2654 had a similar effect in *R. sphaeroides. R. sphaeroides* is typically grown in Sistrom's media (SIS), a minimal media containing aspartic acid and glutamic acid (Sistrom, 1960). When WT 2.4.1 and Δ*2654* were grown aerobically in SIS lacking all amino acids, both strains had equivalent levels of total growth, but the Δ*2654* mutant had a reproducibly slower generation time than wild type cells (6.5 versus 5.5 hours, respectively). In addition, the Δ*2654* mutant exhibited an extended lag phase (Fig. 4.1E). To test if addition of amino acids could rescue the increased lag time

observed for Δ*2654*, both strains were grown in SIS supplemented with 0.4% casamino acids. The Δ*2654* mutant showed a growth defect relative to WT cells even in the presence of casamino acids (Fig 4.1F). In casamino acids, WT cells showed a decrease in generation time relative to a media without amino acids (4.4 hours vs. 5.7 hours, respectively), whereas Δ*2654* cells had approximately the same generation time independent of the presence of casamino acids (6.5 hours vs. 6.8 hours) (Fig. 4.1F). These results suggest that WT cells are able to utilize exogenous amino acids to increase growth rate, but the Δ*2654* mutant cannot. This phenotype is consistent with a potential role for RSP2654 in regulating amino acid metabolism and/or transport.

E. coli DksA complements the photosynthetic growth defect of the R. sphaeroides Δ*2654* mutant

To test whether RSP2654 is functionally related to $DksA_{Ec}$, we asked if $DksA_{Ec}$ expressed in trans from an inducible plasmid (pIND-DksA) could rescue photosynthetic growth of the *R. sphaeroides* Δ*2654* mutant. As shown in Figure 4.2A, $DksA_{Ec}$ is able to complement the photosynthetic growth defect of this strain, albeit not quite as well as cells containing a copy of RSP2654 in trans.

In *E. coli*, two resides in the coiled-coil tip of the protein (D74 and A76, but not D71), are critical for $DksA_{Ec}$ function. Single substitutions at these positions, including a D74N substitution, significantly decreased the activity of $DksA_{Ec}$ *in vitro* and *in vivo* (Lee et al. 2012). As shown in Figure 4.2A, $DksA_{Ec}$ D74N expressed in trans (pIND- $DksA_{Ec}$ D74N), is unable to complement the

photosynthetic growth defect of *R. sphaeroides* Δ*2654*. These results indicate that $DksA_{Ec}$ can functionally substitute for RSP2654 and that the same aspartic acid at the coiled-coil tip that is critical for function in *E. coli* is critical for its ability to compensate for the loss of RSP2654 in *R. sphaeroides.*

RSP2654 complements Δ*dksA E. coli* cells for growth in the absence of amino acids and inhibits rRNA promoter activity *in vivo*

E. coli $dksA_{Ec}$ mutants are amino acid auxotrophs (Webb et al. 1999). Plasmid encoded $DksA_{Ec}$ (pINIIA-$dksA_{Ec}$) restored the ability of cells to grow on plates without amino acids (Lennon et al. 2009, Lee et al. 2012). Since $dksA_{Ec}$ can complement *R. sphaeroides Δ2654* (for the phenotypes described above), we asked whether the *R. sphaeroides* DksA-like proteins could complement *E. coli ΔdksA* cells for growth on medium lacking amino acids. *E. coli* strains carried either the pINIIIA vector or pINIIIA constitutively expressing $dksA_{Ec}$, RSP2654, or RSP0166. WT cells with the pIN/IIIA vector or Δ*dksA* cells with plasmid generated $dksA_{Ec}$ were able to grow in minimal media lacking amino acids (Fig. 4.2B), while Δ*dksA* cells with the pIN/IIIA vector did not, consistent with previous observations (Lennon et al. 2009, Lee et al. 2012). Plasmid encoded RSP2654, but not RSP0166, restored the ability of Δ*dksA* cells to grow without amino acids (Fig. 4.2B).

DksA reduces *E. coli* rRNA promoter activity *in vivo* (Paul et al. 2004). To determine whether *R. sphaeroides* DksA-like proteins could inhibit rRNA production in *E. coli*, we tested the effect of these genes on *rrnB* P1 promoter

activity using an *rrnB* P1-*lacZ* fusion as a reporter. Compared to WT, *rrnB* P1 activity in Δ*dksA* cells is elevated approximately 3.5-fold in log phase when both strains contain an empty pIN/IIIA vector (Fig. 4.2C), consistent with previous studies (Paul et al. 2004, Rutherford et al. 2009, Lee et al. 2012). When $dksA_{Ec}$ is supplied on the plasmid, *rrnB* P1 promoter activity resembles that in WT cells (Fig. 4.2C). Under the same conditions, RSP2654 reduced *rrnB* P1 promoter activity, nearly as well as *E. coli dksA* (Fig. 4.2C). Consistent with the lack of complementation for growth of Δ*dksA* cells on a medium lacking amino acids (Fig. 4.2B), plasmid-encoded RSP0166 did not reduce *rrnB* P1 promoter activity to WT levels (Fig. 4.2C). Because RSP0166 did not complement *E. coli* for the above Δ*dksA* phenotypes, and because the *R. sphaeriodes* deletion mutant did not have any other apparent phenotypes, we did not examine the function of this gene further.

RSP2654 specifically reduces *E. coli rrnB* P1 activity *in vitro*

We next asked whether RSP2654 possessed specific $DksA_{Ec}$ functions *in vitro* with purified *E. coli* components. We reasoned that this could not only inform us about possible roles of RSP2654 in *R. sphaeroides*, but also provide information about the structure/function relationships of $DksA_{Ec}$. In effect, RSP2654 serves as a mutant of $DksA_{Ec}$ in which there is only 42% amino acid identity conservation.

DksA inhibits *E. coli* rRNA promoter activity *in vitro* (Paul et al. 2004). To determine whether the effect of RSP2654 on rRNA promoter activity *in vivo* (Fig.

4.2C) was direct, single round transcription of the *E. coli rrnB* P1 promoter was performed with *E. coli* RNAP at a range of concentrations of His_6-HMK-RSP2654 or His_6-HMK-$DksA_{Ec}$ (His_6-HMK- $DksA_{Ec}$ retains $DksA_{Ec}$ function *in vivo* and *in vitro*, Lennon et al. 2009). The *RNA-1* promoter, which is not affected by $DksA_{Ec}$, was also present on the template. Both $DksA_{Ec}$ and RSP2654 specifically inhibited *rrnB* P1 activity in a concentration-dependent manner (Fig. 4.3A and Fig. 4.3B). Given the evolutionary distance between *E. coli* and *R. sphaeroides*, it was not surprising that higher concentrations of RSP2654 than $DksA_{Ec}$ were required to achieve the same reduction in *E. coli rrnB* P1 promoter activity with *E. coli* RNAP. We estimate the IC_{50} for reduction of *E. coli rrnB* P1 promoter activity to be approximately 5-fold lower for the recombinant RSP2654 protein when compared to $DksA_{Ec}$. However, we cannot be certain that the specific activities of our RSP2654 and $DksA_{Ec}$ preparations are equivalent.

We reasoned that RSP2654 D80 and A82, which correspond to $DksA_{Ec}$ D74 and A76, might be required for inhibition of *E. coli* rRNA. WT His_6-HMK-$DksA_{Ec}$ was compared to WT, D80E, D80V, A82T and D80V/A82T His_6-HMK-RSP2654 (Fig. 4.3C) for inhibition of *E. coli rrnB* P1 promoter activity using *E. coli* RNAP $E\sigma^{70}$. Significant defects in the ability of RSP2654 to inhibit the *E. coli rrnB* P1 promoter were observed with the D80V, A82T and D80V/A82T substitutions (Fig. 4.3C). Therefore, as was the case for $DksA_{Ec}$ D74 and A76, RSP2654 D80 and A82 were critical for transcriptional regulation. Interestingly, in contrast to the severe loss of activity previously observed with $DksA_{Ec}$ D74E (Lee

et al. 2012), RSP2654 D80E inhibited *E. coli rrnB* P1 nearly as well as WT 2654 (Fig. 4.2C), indicating a less strict requirement for aspartate at this position.

The putative coiled-coil tip residues of RSP2654 are important for *in vivo* function in *R. sphaeroides*

We also tested the importance of D80 and A82 *in vivo* in *R. sphaeroides,* using photosynthetic growth on solid media as a reporter of RSP2654 function. Single amino acid substitutions were made in the chromosomally encoded copy of RSP2654 in WT 2.4.1 background by marker-less homologous recombination. The resulting strains (2654-A82T and 2654-D80N) were compared to WT and Δ*2654* strains for photosynthetic growth. Both 2654-A82T and 2654-D80N showed decreased photosynthetic growth relative to WT 2.4.1, but growth was not as severely impaired as in the Δ*2654* mutant, which has an in-frame deletion in the RSP2654 gene (Fig 4.3D). Western blot analysis indicated that RSP2654 protein levels were at least as high in the strains 2654-A82T and 2654-D80N (Fig 4.3E), so the observed phenotype was not a result of protein instability caused by the amino acid substitutions. These results suggest the D80 and A82 residues that are conserved relative to the *E. coli* coiled-coil tip residues are also important for the function of RSP2654 in *R. sphaeroides*. Consistent with this, RSP2654 A82T was defective for inhibition of *E. coli rrnB* P1 using *E. coli* RNAP *in vitro* (Fig. 4.3C). Under the same conditions, RSP2654 D80N inhibited transcription in a non-promoter-specific manner (data not shown).

RSP2654 potentiates the effects of ppGpp on *E. coli* promoters *in vitro*

Along with the unusual nucleotide ppGpp, $DksA_{Ec}$ can either inhibit or activate transcription of different promoters (Paul et al. 2005) in *E. coli*. As displayed above (Fig. 4.3), $DksA_{Ec}$ can inhibit *rrnB* P1 transcription without ppGpp, but it is known that the effects of this protein at this promoter are greatly amplified when ppGpp is also present (Paul et al. 2004). To test whether RSP2654 can also potentiate the inhibitory effects of ppGpp, single round transcription assays were performed using *E. coli rrnB* P1 and *E coli* RNAP. Increasing concentrations of ppGpp were added in the absence of factor, with His_6-HMK-$DksA_{Ec}$, or with His_6-HMK-RSP2654. *rrnB* P1 transcription was reduced modestly in the presence of ppGpp alone (Fig. 4.4A, 3-fold). Both together inhibited transcription drastically (Fig. 4.4A, 10-fold), even at a concentration of $DksA_{Ec}$ or RSP2654 (0.5 μM) in which transcription was only slightly inhibited in the absence of ppGpp (Fig. 4.2A; 50% reduction with $DksA_{Ec}$ alone, 20% reduction with RSP2654).

Direct transcriptional activation can occur at certain *E. coli* promoters in the presence of ppGpp/ $DksA_{Ec}$ (Paul et al. 2005). Unlike transcriptional inhibition by ppGpp/$DksA_{Ec}$, in which either ppGpp or $DksA_{Ec}$ can affect transcription in the absence of the other, transcriptional activation by ppGpp/$DksA_{Ec}$ is dependent on both factors being present, i.e., whereas no effect is observed with either ppGpp or $DksA_{Ec}$ alone (Paul et al. 2005). To determine if RSP2654 and ppGpp together could activate the *hisG* promoter, multiple round transcription assays were performed with purified *E. coli* RNA polymerase (RNAP) $E\sigma^{70}$, His_6-HMK-$DksA_{Ec}$, or His_6-HMK-RSP2654. *hisG* transcriptional activation was achieved by ppGpp

and $DksA_{Ec}$ (4.2 fold) or by ppGpp and RSP2654 (2.3 fold), but not by ppGpp, $DksA_{Ec}$, or RSP2654 alone (Fig. 4.4B). Therefore, RSP2654, along with ppGpp can directly activate transcription from the *E. coli hisG* promoter, albeit to a lesser degree than observed for ppGpp and $DksA_{Ec}$.

RSP2654 binds directly to *E. coli* RNAP

DksA binds directly to *E. coli* RNAP via the secondary channel (Lennon et al. 2012). To determine whether RSP2654 functions from the same site as $DksA_{Ec}$, we employed an Fe^{2+}-mediated cleavage assay in which hydroxyl radicals are generated from the RNAP active site, cleaving proteins non-specifically within ~10Å of the Fe^{2+} (Lennon et al. 2009). Previously, we showed that the coiled-coil tip of $DksA_{Ec}$, near residue 73, is a major target of hydroxyl radical cleavage (Lennon et al. 2012, Fig. 4.5A and Fig. 4.5B). His_6-HMK-RSP2654 and His_6-HMK-$DksA_{Ec}$ were ^{32}P-labeled at the amino-terminus and incubated with RNAP. Following hydroxyl radical generation, products were separated using SDS-PAGE. ^{32}P-His_6-HMK-RSP2654, like ^{32}P-His_6-HMK-$DksA_{Ec}$, was cleaved only in the presence of *E. coli* core RNAP (Fig. 4.5C), indicating a direct interaction *in vitro*. Interestingly, the ^{32}P-His_6-HMK-RSP2654 amino-terminal cleavage product was approximately the same size as the ^{32}P-His_6-HMK-$DksA_{Ec}$ amino-terminal cleavage product (Fig. 4.4C). Although we did not determine the precise site of RSP2654 cleavage, the size of the RSP2654 amino-terminal product suggested that cleavage occurred near the middle of the polypeptide, as with $DksA_{Ec}$ (Lennon et al. 2012). Therefore, D80 and A82 of RSP2654 likely are positioned

similarly to $DksA_{Ec}$ D74 and A76 in *E. coli* RNAP, consistent with the requirement of these residues for RSP2654 inhibition of *E. coli rrnB* P1 (Figure 4.3C).

RSP2654 and ppGpp decrease the lifetime of *R. sphaeroides* RNAP-promoter DNA complexes directly

$DksA_{Ec}$ and ppGpp directly reduce the lifetime of all *E. coli* RNAP-promoter DNA complexes tested to date, inhibiting transcriptional output from a subset of promoters that form intrinsically unstable complexes with RNAP (Paul et al. 2004, Rutherford et al. 2009, Lee et al. 2012). Because we had no specific information about which promoters are targeted by RSP2654 and ppGpp in *R. sphaeroides*, we examined RNAP-promoter DNA complex lifetime as an assay for function. Using the *lacUV5* and *RNA-1* promoters, two promoters that form long-lived complexes with *E. coli* RNAP ($E\sigma^{70}$), we tested whether RSP2654 and ppGpp could reduce *R. sphaeroides* holoenzyme ($E\sigma^{93}$) promoter DNA complex lifetime directly. To determine half-life, RNAP was pre-bound to promoter DNA (on a supercoiled plasmid template) and allowed to equilibrate. Transcription was initiated by the addition of NTPs at different time points after addition of heparin as a competitor for free RNAP as described (Lee et al. 2012). Using *E. coli* holoenzyme ($E\sigma^{70}$), both DksA and RSP2654 reduced complex lifetimes to a similar degree (data not shown). Using *R. sphaeroides* holoenzyme ($E\sigma^{93}$), RSP2654 and ppGpp individually and cooperatively reduced the half-life of *lacUV5* and *RNA-1* promoter complexes (Fig. 4.6). Individually, RSP2654 and ppGpp had modest effects (2-4 fold reduction in ½ life) on the *lacUV5* and *RNA-1*

promoter-complexes (Fig. 4.6). However, when combined, RSP2654 and ppGpp decreased complex stability dramatically (~50 reduction in ½ life) (Fig. 4.6), suggesting that the two factors may function even more synergistically on *R. sphaeroides* RNAP-promoter complexes than on *E. coli* RNAP-promoter complexes.

Conclusions

$DksA_{Ec}$ is an RNAP binding transcription factor in *E. coli*, influencing the expression of hundreds of genes alone and in concert with the global stress regulator ppGpp. Using a variety of *in vivo* and *in vitro* assays, we showed that RSP2654 acts by a similar mechanism as $DksA_{Ec}$. Further, our analysis of RSP2654 *in vitro* with *E. coli* RNAP and promoters provides clues about the structure and function of $DksA_{Ec}$, for which considerable work has been done to define the detailed mechanism of action. Finally, to our knowledge, this is the first biochemical analysis of DksA homologues from an alphaproteobacterium.

It has been reported that *R. sphaeroides* does not utilize ppGpp to regulate gene expression in response to amino acid starvation, as in *E. coli*. In response to leucine deprivation, *R. sphaeroides* halts stable RNA synthesis without ppGpp accumulation (Acosta and Lueking, 1987). A similar result was reported in two other oligotrophic alphaproteobacteria that contain *dksA* homologues more closely related to *R. sphaeroides* than to *E. coli*. In *C. crescentus*, carbon and ammonia starvation, but not amino acid starvation, led to ppGpp accumulation (Boutte and Crosson, 2011). In *S. meliloti*, the same general trend as in *C. crescentus* is observed. Amino acid starvation doesn't induce ppGpp synthesis, but carbon and ammonia starvation does. *S. meliloti* transcriptome analysis comparing WT, Δ*relA* (ppGpp°), and Δ*dksA* cells following nitrogen starvation showed that many of the same genes regulated by ppGpp were also regulated by DksA (Krol and Becker, 2011). In agreement with this report, our *in vitro* results (Fig. 4.6) argue that ppGpp and DksA regulate

promoter targets together in *R. sphaeroides.* It was reported previously, consistent with our observed *R. sphaeroiedes* Δ*2654* phototrophic growth defect (Fig. 4.1), that ppGpp is rapidly synthesized in *R. sphaeroides* when phototrophically growing cells are shifted to decreased light intensity (Eccleston Jr. and Gray, 1973). These new data suggest RSP2654 and ppGpp work in concert to regulate critical transcriptional changes required for photosynthetic growth, explaining the cellular requirement of RSP2654.

RSP2654 and ppGpp destabilized promoter complexes formed by *R. sphaeroiedes* RNAP ($E\sigma^{93}$) directly, a critical mechanistic characteristic conserved with $DksA_{EC}$ and ppGpp. While RSP2654 and ppGpp both decreased *R. sphaeroiedes* $E\sigma^{93}$-*lacUV5* and *RNA-1* promoter complex lifetimes individually (~2-4 fold reduction in half life, Fig. 4.6), RSP2654 and ppGpp together reduced complex half lives ~50 fold (Fig. 4.6), suggesting the two factors may act even more synergistically on *R. sphaeroiedes* RNAP than on *E. coli* RNAP. Recently, the binding site of ppGpp on *E. coli* RNAP was determined (WER and RLG, in prep.). This site is conserved quite well in *R. sphaeroides*, as well as in the alphaproteobacteria *C. crescentus* and *S. meliloti* (Table 4.1). RSP2654 is also highly conserved with putative *C. crescentus* and *S. meliloti dksA* (Fig. 4.7 shows amino acid sequence alignment; 65-70% amino acid identity conservation). RSP2654 residues D80 and A82, which are critical for regulating *E. coli rrnB* P1 promoter activity (Fig. 4.3C), are conserved in *C. crescentus* (D64/A66) and *S. meliloti* (D61/A63). In addition, *R. sphaeroides, C. crescentus* and *S. meliloti* only contain a single cysteine in their globular domain (Fig. 4.7), in contrast to *E. coli,*

which has a four-cysteine zinc finger (Perederina et al. 2004 and Fig. 4.7). DksA2 from *P. aeruginosa*, which only has two cysteine residues and cannot coordinate Zn^{2+}, has been shown to be homologous to $DksA_{Ec}$, both functionally (Blaby-Haas et al. 2011) and structurally (Furman et al. 2013). Therefore, the presence of a zinc-finger does not appear necessary to retain DksA-like function. Furthermoure, our results showing that RSP0166 does not retain DksA-like activity *in vivo* in *E. coli*, argues that a zinc-finger alone is insufficient to provide DksA-like function. We suggest that genes annotated as DksA based on zinc-finger homology alone should be evaluated with caution.

Materials and methods

Bacterial Strains and growth conditions

E. coli DH5α was used as a plasmid host, and *E. coli* S17-1 was used as a donor for plasmid conjugation into *R. sphaeroides*. *E. coli* strains were grown at 37°C in Luria-Bertani medium. *R. sphaeroides* 2.4.1 strains were grown at 30°C in Sistrom's succinate-based medium (Sistrom, 1960), unless otherwise noted. When necessary the media were supplemented with kanamycin (25 µg/ml for *R. sphaeroides*, 50 µg/ml for *E. coli*). For growth at 0.5% oxygen conditions, 500 ml *R. sphaeroides* cultures were bubbled with 98.5% N_2, 0.5% O_2, and 1% CO_2. For photosynthetic growth, agar plates were incubated in sealed canisters containing a GasPak EZ Anaerobe Container System (BD biosciences) and incubated at room temperature in front of an incandescent light with a light intensity of 10 W/m^2.

For aerobic growth curves, *R. sphaeroides* strains were grown in Sistom's media was made without aspartic acid and glutamic acid, or in Sistrom's media supplemented with 0.4% casamino acids and 0.004% tryptophan. 200 ml cultures were incubated at 30^{o}C in clear 96-well plates in an Infinite® F500 plate reader (Tecan, Männedorf, Switzerland) shaking at 33.2 rpm orbitally. Absorbance was measured every ~10 minutes at 595 nm after 10 seconds of linear shaking.

Construction of R. sphaeroides mutants

Site-specific gene deletion of RSP2654 and RSP0166 was carried out to create strains D2654 and D0166 using the nonreplicable integration vector pK19mobsacB, which allows for the marker-free deletion by two-step homologous recombination (Schafer et al., 1994). For each gene, 2.3 kb fragments were amplified from genomic DNA of *R. sphaeroides* containing each open reading frame flanked by 0.8-1 kb of sequence on each side with primers containing XbaI and EcoRI (RSP2654) or HindIII (RSP0166). These PCR products were digested with the relevant restriction enzymes and inserted into pK18mobsacB cut with the same enzymes to create plasmids pKC09 and pKCL07. The entire coding region of RSP2654 or RSP0166 was deleted from the respective plasmids by performing PCR with primers facing outward from each end of the ORF and ligation of the resulting fragment with T4 DNA ligase (Promega, Madison, WI) to create pKCL08 and pKCL10. These plasmids were mated into *R. sphaeroides* from *E. coli* S17-1. Single crossovers were selected by kanamycin resistance, and double crossovers by loss of sucrose sensitivity. Sucrose-resistant strains in which the RSP2654 or RSP0166 coding regions had been excised were identified by PCR.

To create strains 2654-D80N and 2654-A82T site-specific mutagenesis was performed using pK18mobsacB-derived plasmids. Two-step PCR mutagenesis was performed to create 2.3 kb genomic fragments containing RSP2654 (as described above) in which RSP2654 had mutations for each desired amino acid substitution. One base pair change in each PCR product resulted in a codon change of the desired amino acid substitution (D80 to N, or

A82 to T), and a second base pair change mutated an EarI restriction site, to aid in screening (see below). These PCR products were inserted into the XbaI and EcoRI sites of pK18mobsacB to make plasmids pKCL11 and pKCL12. The plasmids were mobilized into *R. sphaeroides* and selected as described above. The resulting sucrose-resistant strains were screened for a copy of RSP2654 containing the 2 mutated nucleotides by PCR of the gene and digestion with EarI. Strains containing a copy of RSP2654 that was not digested by EarI were sequenced for verification of the codon change.

Spectroscopy

To assess photosynthetic pigment-protein complex assembly, aliquots of exponential-phase cell culture were assayed by visible spectroscopy on an Olis DW-2/2000 spectrophotometer. To normalize for cell density, all spectra were scaled to an absorbance of 1 at 680 nm.

Construction of plasmids for expression of RSP2654 or E. coli DksA in R. sphaeroides

The coding sequences for RSP2654 was PCR amplified from genomic DNA and inserted into the NdeI and HindIII sites of pIND5 downstream of the IPTG-inducible promoter. The coding sequences for *E. coli* DksA and DksA-D74N was PCR amplified from plasmids pRLG6333 and pRLG8873 (Paul et al., 2005), respectively, with primers containing AseI and BglII restriction sites. These

PCR products were inserted into the NdeI and BglII sites of pIND5, adding a hexa-histidine tag onto the C-terminus of the expressed protein.

Western blot analysis

Exponentially growing cultures were harvested, resuspended in urea buffer (8M urea, 100 mM NaH_2PO4, 10 mM Tris pH 8.0) supplemented with 50 mM phenylmethylsulfonyl fluoride, and then heated at 95^o for 10 minutes. Samples were centrifuged to remove debris and total protein concentration of the samples was determined with using Bio-Rad Protein Assay Reagent (Bio-Rad, Hercules, CA) following the manufacturer's protocol. Western blotting was performed as previously described (Dufour et al., 2012) using a rabbit polyclonal antibody raised against recombinant RSP2654, purified as described above (Harlan Laboratories, Madison, WI). Detection was performed with Pierce ECL Western Blotting Substrate (Pierce, Rockford, IL).

Construction of plasmids for expression of RSP2654 and RSP0166 in E. coli

R. sphaeroides 0166 and *2654* DNA was synthesized (GeneArt) with codons optimized for expression in *E. coli* and cloned into the pINIIIA vector at the XbaI and HindIII sites and the pET33 vector at the NheI and HindIII sites. Mutagenesis of *R. sphaeroides 2654* was performed using a QuikChange Lighting Multi Site-Directed Mutagenesis Kit (Stratagene) by standard procedures.

E. coli growth without amino acids

WT *E. coli* cells were transformed with the empty pINIIIA vector and *ΔdksA E. coli* cells were transformed with either the empty pINIIIA vector or the pINIIIA vector constitutively expressing one of the following: *E. coli* DksA, *R. sphaeroides* 2654, or *R. sphaeroides* 0166. Following 1 hour outgrowth in LB at 30°C, cells were plated on M9 agar containing 0.4% glycerol, 100 μg/mL ampicillin and grown an additional ~16 hours at 30°C.

β-galatosidase assay

RLG5950 (WT *E. coli* containing *rrnB* P1 promoter -61 to +1 fused to a *lacZ* reporter) was transformed with the empty pINIIIA vector and RLG7238 (Δ*dksA*::*tet E. coli* containing an *rrnB* P1 promoter -61 to +1 fused to a *lacZ* reporter) was transformed with the empty pINIIIA vector or the pINIIIA vector constitutively expressing either *E. coli dksA*, *R. sphaeroides 2654*, or *R. sphaeroides 0166*. Cells were grown in M9 media containing 0.4% glycerol, casamino acids and 100 μg/mL ampicillin to an optical density of ~0.4 at 600 nm (~4 generations) for log phase measurements. Cells were chilled on ice ~20 minutes, sonicated and β-galatosidase activity was measured by standard procedures as described (Barker et al. 2001).

Protein purification

His_6-HMK-DksA and His_6-HMK-2654 were purified by Ni^{2+}-affinity chromatography using conditions previously described for His_6-DksA (Paul et al. 2004). Native *E. coli* RNAP holoenzyme ($E\sigma^{70}$) was purified as described

(Burgess and Jendrisak 1975). Native *R. sphaeroides* core RNAP was purified as described (Anthony et al. 2004), except that heparin resin was substituted for DNA cellulose. For the heparin purification step, partially pure RNAP in TGE (10 mM Tris-HCl, 0.1 mM EDTA, 5% glycerol) + 200 mM NaCl was bound to heparin resin equilibrated in the same buffer, washed with one column volume of each TGE + 300, 400, 500, 600 or 700 mM NaCl. RNAP eluted during the 500-700 mM NaCl wash steps and was subsequently concentrated into storage buffer (20 mM Tris-HCl pH 7.9, 100 mM NaCl, 0.1 mM DTT, 0.1 mM EDTA, 50% glycerol). His_6-σ^{93} was purified from the soluble fraction using Ni^{2+}-affinity chromatography. Briefly, cells were resuspended in buffer A (40 mM Tris-HCl pH 7.9, 10 mM imidazole) plus 300 mM NaCl, lysed via sonication, centrifuged, and the cleared lysate was passed over Ni-NTA resin (Qiagen) equilibrated with buffer A plus 300 mM NaCl. The column was subsequently washed with buffer A plus 600 mM NaCl, buffer A plus 900 mM NaCl, eluted with buffer A plus 900 mM NaCl with 300 mM imidazole, and finally dialyzed for storage against 20 mM Tris-HCl pH 7.9, 200 mM NaCl, 0.1 mM DTT, 0.1 mM EDTA, 50% glycerol.

In vitro transcription

Supercoiled template DNA (150 ng) containing the *rrnB* P1 and *RNA-1* promoters (pRLG1616) or *hisG* and *RNA-1* promoters (pRLG4413) was incubated with His_6-HMK-DksA, His_6-HMK-2654 (WT or variant), or no factor (storage buffer) in transcription buffer (20 mM Tris-HCl pH 7.9, 10 mM $MgCl_2$, 1 mM DTT, 0.1 mg/mL BSA) at room temperature (~20°C) for 10 minutes. Additionally,

transcription buffer contained either 50 mM NaCl (single round) or 165 mM NaCl (multiple round). ppGpp was present at 100 μM where indicated. *E. coli* RNAP ($E\sigma^{70}$) was added to a final concentration of 10 nM and NTPs were added at a final concentration of 500 μM ATP, 200 μM GTP, 200 μM CTP, 10 μM UTP, and 1.0 μCi of α-^{32}P-UTP. Depending on whether single or multiple round transcription was performed, RNAP or NTPs were added to initiate transcription. For single round transcription, reactions were initiated by the addition of NTPs with heparin (100 μg/mL final). For multiple round transcription, reactions were initiated by the addition of RNAP. Reactions were allowed to proceed for 10 minutes at room temperature (~20°C) or 30°C and halted by the addition of 2X stop buffer (95% formamide, 20 mM EDTA, 0.05% bromophenol blue, and 0.05% xylene cyanol). Transcripts were separated using 5.5% polyacrylamide, 7M urea gels. RNA was quantified using phosporimaging and ImageQuant software.

Fe^{2+}-mediated cleavage assay

E. coli His_6-HMK-DksA or *R. sphaeroides* His_6-HMK-2654 were ^{32}P-labeled as described (Lennon et al. 2009). Excess ^{32}P-γ-ATP was removed and proteins were exchanged into cleavage buffer (20 mM NaCl, 20 mM HEPES pH 7.9) using G-50 size exclusion spin columns (GE Healthcare). *E. coli* core RNAP was also exchanged into cleavage buffer. 1.8 μM core RNAP was incubated at 30°C for 10 minutes with ~20 nM ^{32}P-labeled His_6-HMK-DksA or *R. sphaeroides* His_6-HMK-2654 in a 10 μl reation. Hydroxyl radicals were generated by from the active site of RNAP by the concurrent addition of 1 μl 100 mM DTT and 1 μl 500

μM $(NH4)_2$-$Fe(SO_4)_2$. Reactions were incubated at 30°C for 10 minutes and stopped by the addition of an equal volume of 2X LDS (Invitrogen). Reactions were electrophoresed using 4-12% NuPAGE gels with MES buffer (Invitrogen). Phosphorimaging was used to visualize ^{32}P-labeled products.

RNAP-promoter complex lifetime assay

Promoter (*lacUV5*) complex lifetime with *R. sphaeroides* $E\sigma^{93}$ was determined by measuring transcription from a supercoiled plasmid template at different times following heparin addition. 150 ng DNA (pRLG3422) was incubated with ~10 nM RNAP in transcription buffer (200 mM NaCl, 20 mM Tris-HCl pH 7.9, 10 mM $MgCl_2$, 1 mM DTT, 0.1 mg/mL BSA) at room temperature (~20°C) for 10 minutes. 4 μM His_6-HMK-2654, 333 μM ppGpp, 4 μM His_6-HMK-2654 and 333 μM ppGpp, or no factor (storage buffer) were next added, followed by heparin (to a final concentration of 100 μg/mL). Transcription was initiated by NTP addition at the indicated times following heparin addition. Reactions were allowed to proceed for 10 minutes. Reactions were stopped and transcripts were quantified as described above (*in vitro transcription*).

References

Acosta R, Lueking DR. 1987. Stringency in the absence of ppGpp accumulation in *Rhodobacter sphaeroides*. *J Bacteriol* **169**: 908-912.

Anthony JR, Green HA, Donohue TJ. 2004. Purification of Rhodobacter sphaeroides RNA polymerase and its sigma factors. *Meth Enzymol* **370**:54-65.

Barker M, Gaal T, Josaitis C, Gourse R. 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.

Blaby-Haas CE, Furman R, Rodionov DA, Artsimovitch I, de Crecy-Lagard V. 2011. Role of a Zn-independent DksA in Zn homeostasis and stringent response. *Mol Microbiol* **79**: 700-715.

Boutte CC, Crosson S. 2011. The complex logic of stringent response regulation in *Caulobacter crescentus*: starvation signaling in an oligotrophic environment. *Mol Microbiol* **80**: 695-714.

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

Davis J, Donohue TJ, Kaplan S. 1988. Construction, characterization, and complementation of a Puf- mutant of *Rhodobacter sphaeroides*. *J Bacteriol* **170**: 320-329.

Dufour, YS, Imam S, Koo BM, Green HA, Donohue TJ.2012. Convergence of the transcriptional responses to heat shock and singlet oxygen stresses. *PLoS Genet* **8**: e1002929.

Eccleston Jr ED, Gray ED. 1973. Variations in ppGpp levels in *Rhodopseudomonas spheroides* during adaptation to decreased light intensity. *Biochem Biophys Res Commun*. **54**:1370–1376.

Eraso, JM, Kaplan S. 1994. prrA, a putative response regulator involved in oxygen regulation of photosynthesis gene expression in *Rhodobacter sphaeroides. J Bacteriol* **176**: 32-43.

Furman R, Biwas T, Danhart EM, Foster MP, Tsodikov OV, Artsimovitch I. 2013. DksA2, a zinc-independent structural analog of the transcription factor DksA. *FEBS letters*, in press.

Gomelsky M, Kaplan S. 1998. AppA, a redox regulator of photosystem formation in Rhodobacter sphaeroides 2.4.1, is a flavoprotein. Identification of a novel fad binding domain. *J Biol Chem* **273**: 35319-35325.

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

Henard CA, Vazquez-Torres A. 2012. DksA-dependent resistance of Salmonella enterica serovar Typhimurium against the antimicrobial activity of inducible nitric oxide synthase. *Infect Immun* **80**: 1373-1380.

Karls RK, Jin DJ, Donohue TJ. 1993. Transcription properties of RNA polymerase holoenzymes isolated from the purple nonsulfur bacterium *Rhodobacter sphaeroides. J Bacteriol* **175**:7629-7638.

Krol E, Becker A. 2011. ppGpp in *Sinorhizobium meliloti*: biosynthesis in response to sudden nutritional downshifts and modulation of the transcriptome. *Mol Microbiol* **81** 1233-1254.

Lee JH, Lennon CW, Ross W, Gourse RL. 2012. Role of the coiled-coil tip of Escherichia coli DksA in promoter control. *J Mol Biol* **416**: 503-517.

Lennon CW, Gaal T, Ross W, Gourse RL. 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 CW, Ross W, Martin-Tumasz S, Toulokhonov I, Vrentas CE, Rutherford ST, Lee JH, Butcher SE, Gourse RL. 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.

Paul B, Barker M, Ross W, Schneider D, Webb C, Foster J, Gourse R. 2004a. 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, Barker M, Gourse R. 2005. DksA potentiates direct activation of amino acid promoters by ppGpp. *Proc Natl Acad Sci USA* **102:** 7823-7828.

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

Rutherford S, Villers C, Lee J, Ross W, Gourse RL. 2009. Allosteric control of *Escherichia coli* rRNA promoter complexes by DksA. *Genes Dev* **23:** 236-248.

Schafer A, Tauch A, Jager W, Kalinowski J, Thierbach G, Puhler A. 1994. Small mobilizable multi-purpose cloning vectors derived from the Escherichia coli plasmids pK18 and pK19: selection of defined deletions in the chromosome of Corynebacterium glutamicum. *Gene* **145**: 69-73.

Sistrom WR. 1960. A requirement for sodium in the growth of *Rhodopseudomonas spheroides. J Gen Microbiol* **22**: 778-785.

Webb C, Moreno M, Wilmes-Riesenberg M, Curtiss R, Foster J. 1999. Effects of DksA and ClpP protease on sigma S production and virulence in *Salmonella typhimurium*. *Mol Microbiol* **34:** 112-123.

Zeilstra-Ryalls JH, Kaplan S. 2004. Oxygen intervention in the regulation of gene expression: the photosynthetic bacterial paradigm. *Cell Mol Life Sci* **61**: 417-436.

Table 4.1

Residues required for ppGpp binding in *E. coli*							
Subunit	**ω**	**β'**	**β'**	**β'**	**β'**	**β'**	**β'**
Residue (*E. coli* numbers)	**2-5**	**362**	**417**	**615**	**619**	**622**	**626**
E. coli	**ARVT**	**R**	**R**	**K**	**I**	**D**	**Y**
R. sphaeroides	**ARVT**	**K**	**R**	**K**	**I**	**D**	**G**
C. crescentus	**ARVT**	**K**	**R**	**K**	**I**	**D**	**G**
S. meliloti	**ARVT**	**K**	**R**	**K**	**I**	**D**	**Q**

Table 4.2

Strain	Description
BL21(DE3)	*E. coli* expressing T7 RNAP for pET vector expression
RLG5950	*rrnB* P1-*lacZ* fusion, *E. coli*
RLG7238	Δ*dksA*, *rrnB* P1-*lacZ* fusion, *E. coli*
R. sph. 2.4.1	WT lab strain
D2654	In-frame deletion of RSP2654 coding sequence in 2.4.1
D0166	In-frame deletion of RSP0166 coding sequence in 2.4.1
2654-D80N	2.4.1 derivative encoding a D80N substitution in RSP2654
2654-A82T	2.4.1 derivative encoding an A82T substitution in RSP2654

Table 4.3

Plasmid	Description
pRLG8150	pET33-DksA
pRLG10347	pET33-2654
pRLG12055	pET33-2654 D80V
pRLG12056	pET33-2654 D80E
pRLG12057	pET33-2654 A82T
pRLG12058	pET33-2654 D80V/A82T
pRLG6332	pINIIIA
pRLG6333	pINIIIA-DksA
pRLG10841	pINIIIA-0166
pRLG11003	pINIIIA-2654
pRLG1616	pRLG770 containing *rrnB* P1 and *RNA-1* promoters
pRLG3422	pRLG770 containing *lacUV5* and *RNA-1* promoters
pRLG11367	pRLG770 containing *hisG* and *RNA-1* promoters
pRLG13055	His_6-σ^{93}
pKCL07	2286-bp *R. sphaeroides* genomic region containing RSP0166 in pK18mobsacB
pKCL08	pKC07 with the entire RSP0166 ORF deleted
pKCL09	2281-bp *R. sphaeroides* genomic region containing RSP2654 in pK18mobsacB
pKCL10	pKCL09 with the entire RSP2654 ORF deleted
pKCL11	As pKCL09 but encoding RSP2654-D80N
pKCL12	As pKCL09 but encoding RSP2654-A82T
pKCL15	pIND5-RSP2654
pKCL17	pIND5-$DksA_{Ec}$
pKCL18	pIND5-$DksA_{Ec}$ D74N

Figures and legends

Figure 4.1. Phenotype of *R. sphaeroides* strains D2654 and D0166. (A) Colony pigmentation for WT and mutant strains grown aerobically on SIS agar plates. (B) Absorbance spectra of WT and D2654 liquid cultures grown at 0.5% oxygen in the dark. Spectra were obtained from intact cells, normalized to equal absorbance at 680 nm, and staggered for presentation on one vertical axis. (C) Photosynthetic growth of WT and mutant strains on SIS agar plates. (D) Close up image shows colony size of WT and D2654 grown photosynthetically. A 10^{-6} dilution of culture was plated on SIS agar plates and incubated anaerobically for 7 days in the light until maximum colony size was achieved. (E and F) Growth curves of WT and mutant strains grown aerobically in 96-well plates in either SIS lacking all amino acids (E) or SIS supplemented with 0.4% casamino acids (F). The generation times represent the average ± standard deviation of five replicates in one representative experiment. One representative curve for each strain is shown.

Figure 4.1

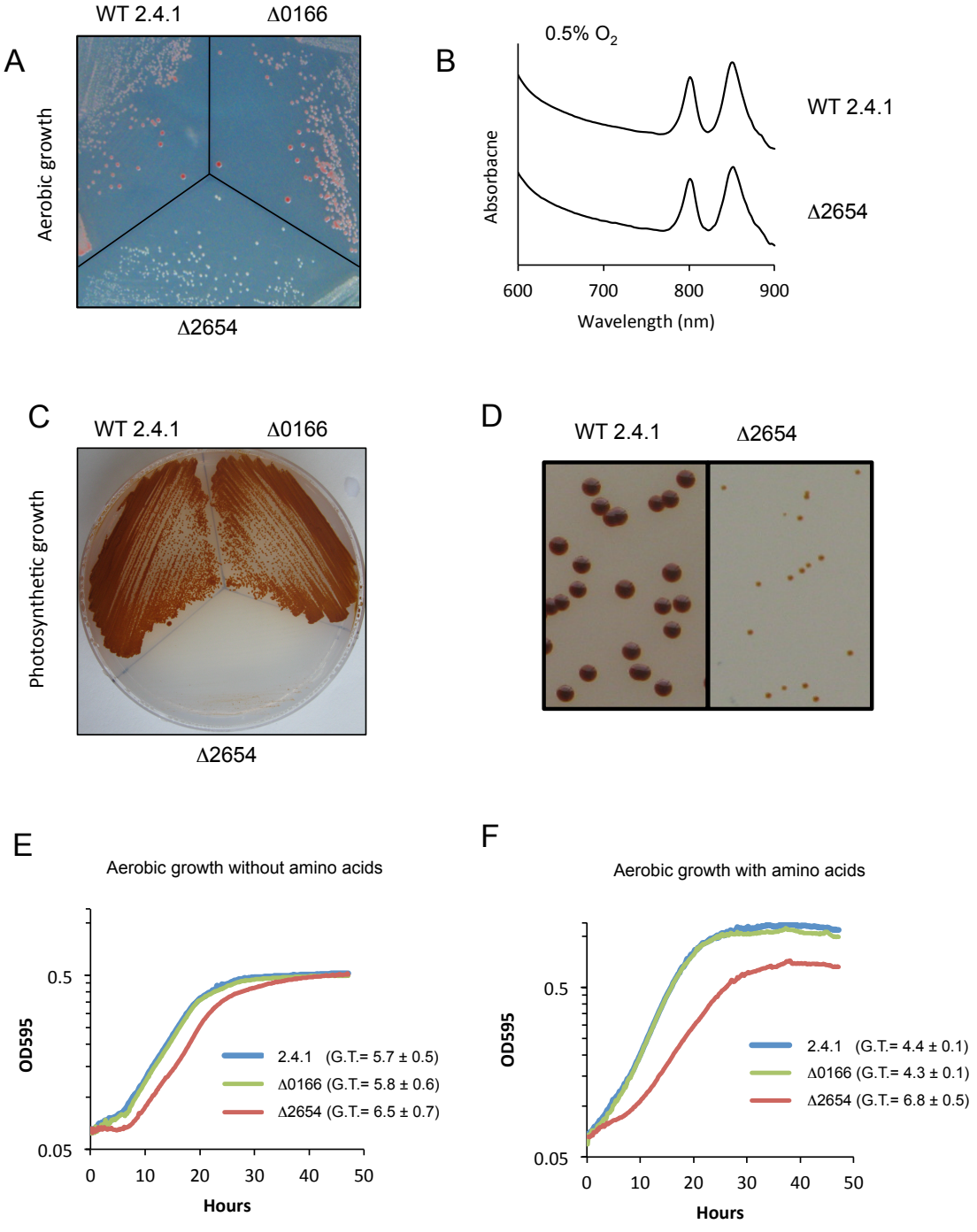

A
WT 2.4.1
Δ0166
Aerobic growth
Δ2654
B
0.5% O2
WT 2.4.1
Δ2654
Absorbacne
600
700
800
900
Wavelength (nm)
C
WT 2.4.1
Δ0166
Photosynthetic growth
Δ2654
D
WT 2.4.1
Δ2654
E
Aerobic growth without amino acids
OD595
0.5
0.05
0
10
20
30
40
50
Hours
2.4.1 (G.T.= 5.7 ± 0.5)
Δ0166 (G.T.= 5.8 ± 0.6)
Δ2654 (G.T.= 6.5 ± 0.7)
F
Aerobic growth with amino acids
OD595
0.5
0.05
0
10
20
30
40
50
Hours
2.4.1 (G.T.= 4.4 ± 0.1)
Δ0166 (G.T.= 4.3 ± 0.1)
Δ2654 (G.T.= 6.8 ± 0.5)

Figure 4.2. (A) Photosynthetic growth of *R. sphaeroides* WT, Δ2654, and Δ2654 derivatives carrying the indicated plasmids. Strains were streaked onto SIS agar plates containing 100 mM IPTG and incubated anaerobically in the light. (B) *E. coli* Δ*dksA* cells cannot grow in the absence of amino acids. Growth is restored by plasmid encoded $DksA_{Ec}$ and RSP2654, but not RSP0166. (C) *rrnB* P1 promoter activity is de-repressed ~3.5-fold during log phase in Δ*dksA* strain when compared to WT as measured by β-galatosidase activity. *rrnB* P1 promoter activity is repressed to WT levels in the presence of plasmid bourn $DksA_{Ec}$ and RSP2654, but not RSP0166. In B-C, $DksA_{Ec}$, RSP2654 and RSP0166 were expressed constitutively from the pINIIIA vector.

Figure 4.2

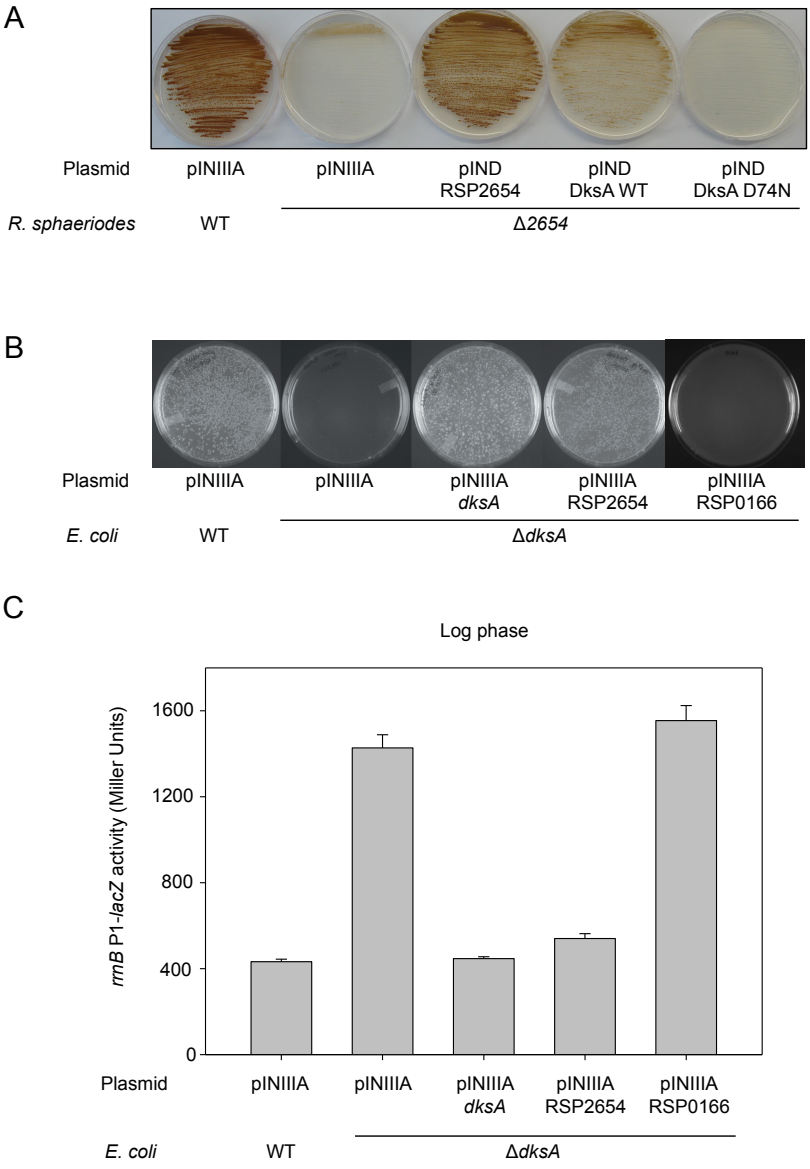

A
Plasmid
pINIIIA
pINIIIA
pIND RSP2654
pIND DksA WT
pIND DksA D74N
R. sphaeriodes
WT
Δ2654
B
Plasmid
pINIIIA
pINIIIA
pINIIIA dksA
pINIIIA RSP2654
pINIIIA RSP0166
E. coli
WT
ΔdksA
C
Log phase
rrnB P1-lacZ activity (Miller Units)
1600
1200
800
400
0
Plasmid
pINIIIA
pINIIIA
pINIIIA dksA
pINIIIA RSP2654
pINIIIA RSP0166
E. coli
WT
ΔdksA

Figure 4.3. *R. sphaeroides* 2654 specifically inhibits *E. coli rrnB* P1 and requires residues D80 and A82 for the funciton. (A) Transcription of *E. coli rrnB* P1 is inhibited in a concentration dependent manner by His_6-HMK-$DksA_{Ec}$ or His_6-HMK-RSP2654. Single round transcription from the *E. coli rrnB* P1 and *RNA-1* promoters by *E. coli* RNAP $E\sigma^{70}$ with increasing concentrations (0.5-4 μM) of His_6-HMK-$DksA_{Ec}$ or His_6-HMK-RSP2654. Reactions are in duplicate. (B) Quantification of (A) and two replicate experiments. (C) Residues D80 and A82 of RSP2654 are required for inhibition of *E. coli rrnB* P1. Single round transcription as in (A) with 2 μM His_6-HMK-$DksA_{Ec}$ (WT) or 10 μM His_6-HMK-RSP2654 (WT and mutants). Reactions are in triplicate. (D) Photosynthetic growth of *R. sphaeroides* WT and D2654 strains encoding the indicated single amino acid substitution in their chromosomal copy of RSP2654 (2654-D80N and 2654-A82T). (E) Western blot analysis showing the levels of RSP2654 protein in the strains shown in (D). Equal amount of total protein was loaded in each lane.

Figure 4.3

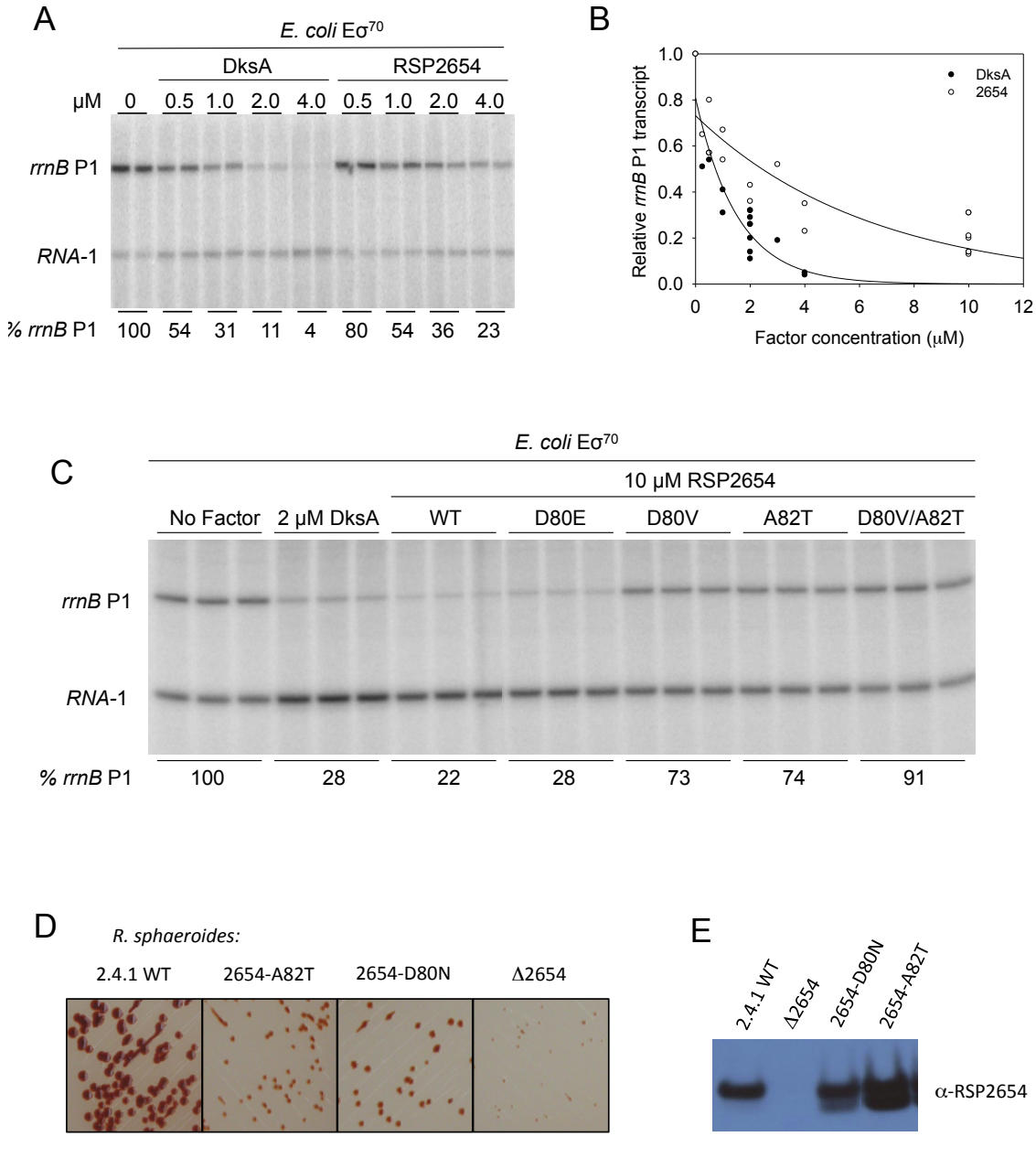

A
E. coli Eσ70
DksA
RSP2654
μM 0 0.5 1.0 2.0 4.0 0.5 1.0 2.0 4.0
rrnB P1
RNA-1
% rrnB P1 100 54 31 11 4 80 54 36 23
B
Relative rrnB P1 transcript
DksA
2654
Factor concentration (μM)
C
E. coli Eσ70
10 μM RSP2654
No Factor
2 μM DksA
WT
D80E
D80V
A82T
D80V/A82T
rrnB P1
RNA-1
% rrnB P1 100 28 22 28 73 74 91
D
R. sphaeroides:
2.4.1 WT
2654-A82T
2654-D80N
Δ2654
E
2.4.1 WT
Δ2654
2654-D80N
2654-A82T
α-RSP2654

Figure 4.4. RSP2654 potentiates the negative and positive effects of ppGpp on *E. coli* promoters. (A) His_6-HMK-$DksA_{Ec}$ or His_6-HMK-RSP2654 (0.5 μM) amplify the inhibitory effects of ppGpp on *E. coli rrnB* P1 using *E. coli* RNAP as measured by single round transcription. The concentration of ppGpp ranges from 12.5-200 μM. (B) His_6-HMK-$DksA_{Ec}$ (2 μM) or His_6-HMK-RSP2654 (10 μM) and 100 μM ppGpp synergistically activate transcription from the *E. coli hisG* promoter using *E. coli* RNAP as measured by multiple round transcription. Activation only occurs in the presence of both ppGpp and His_6-HMK-$DksA_{Ec}$ or His_6-HMK-RSP2654. Reactions are in duplicate.

Figure 4.4

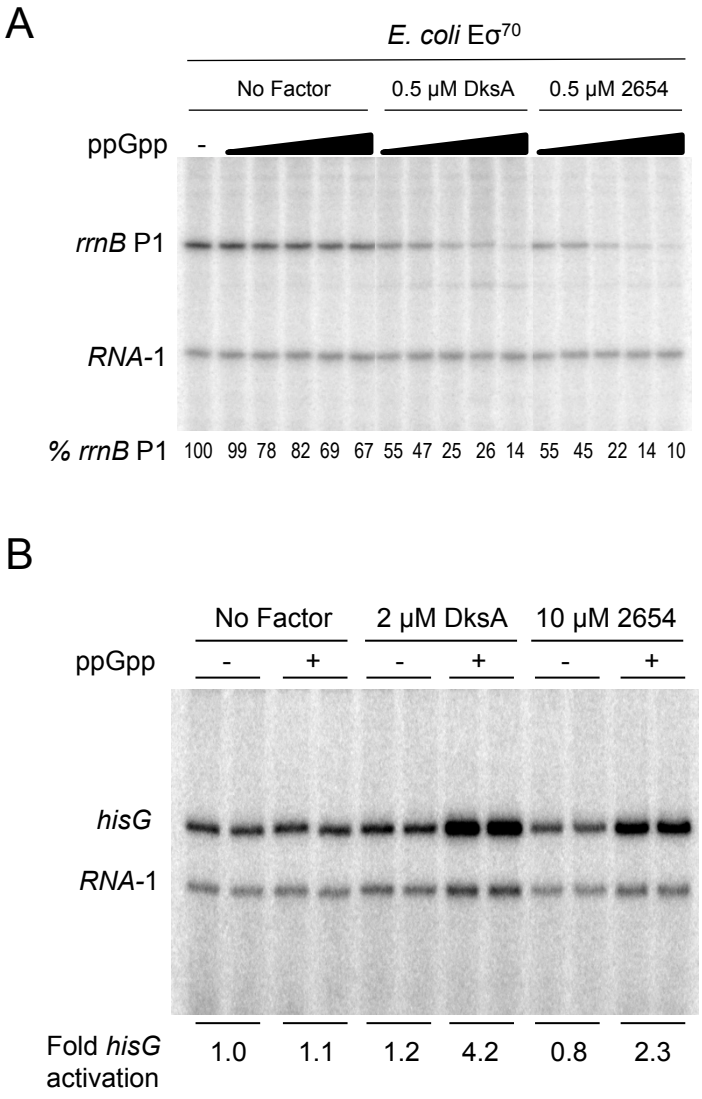

A
E. coli Eσ70
No Factor
0.5 μM DksA
0.5 μM 2654
ppGpp
-
rrnB P1
RNA-1
% rrnB P1
100 99 78 82 69 67 55 47 25 26 14 55 45 22 14 10
B
No Factor
2 μM DksA
10 μM 2654
ppGpp
- + - + - +
hisG
RNA-1
Fold hisG activation
1.0 1.1 1.2 4.2 0.8 2.3

Figure 4.5. RSP2654, like $DksA_{Ec}$, binds directly to *E. coli* RNAP near the active site. (A) Schematic representation of RNAP Fe^{2+}-mediated cleavage of $DksA_{Ec}$ and approximate position of $DksA_{Ec}$ cleavage (Lennon et al. 2012). (B) RNAP Fe^{2+}-mediated cleavage of DksA (grey) occurs at the tip, near residue 73 (black) (PDB: 1TJL). (C) Cleavage of His_6-HMK-$DksA_{Ec}$ and His_6-HMK-RSP2654 occurs in the presence of *E. coli* core RNAP.

Figure 4.5

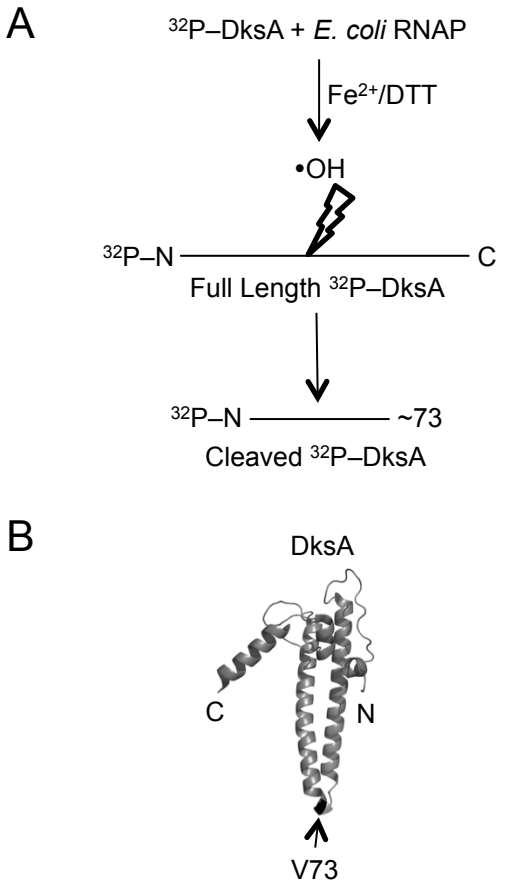
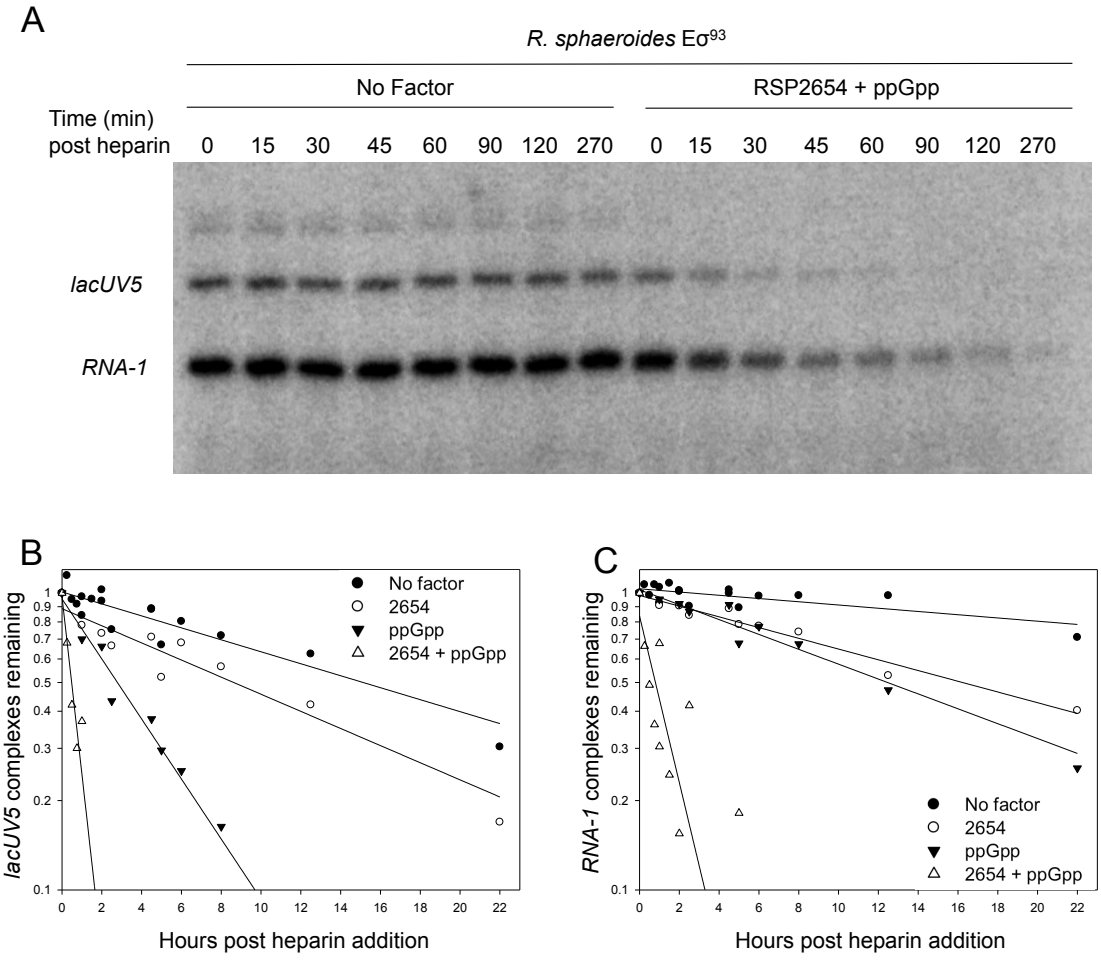

A
32P–DksA + E. coli RNAP
Fe2+/DTT
•OH
32P–N
C
Full Length 32P–DksA
32P–N
~73
Cleaved 32P–DksA
B
DksA
C
N
V73
C
32P-DksA
32P-2654
E. coli RNAP
-
+
-
+
Full Length
Cleaved

Figure 4.6. RSP2654 and ppGpp directly reduce the lifetime of *R. sphaeroides* RNAP-promoter complexes. (A) The ½ life of *R. sphaeroides* RNAP Eσ^{93} *lacUV5* and *RNA-1* promoter complexes is reduced in presence of His_6-HMK-RSP2654 (4 μM) and ppGpp (333 μM). Semilog plots of *R. sphaeroides* RNAP Eσ^{93} *lacUV5* (B) or *RNA-1* (C) promoter complex ½ lives in the presence of no factor, His_6-HMK-RSP2654 (4 μM), ppGpp (333 μM), or His_6-HMK-RSP2654 (4 μM) and ppGpp (333 μM). The decay curves show the fraction remaining complexes at various times following heparin addition. Approximate ½ lives of *lacUV5* (D) or *RNA-1* (E) promoter complexes as quantified from (B) and (C).

Figure 4.6



D

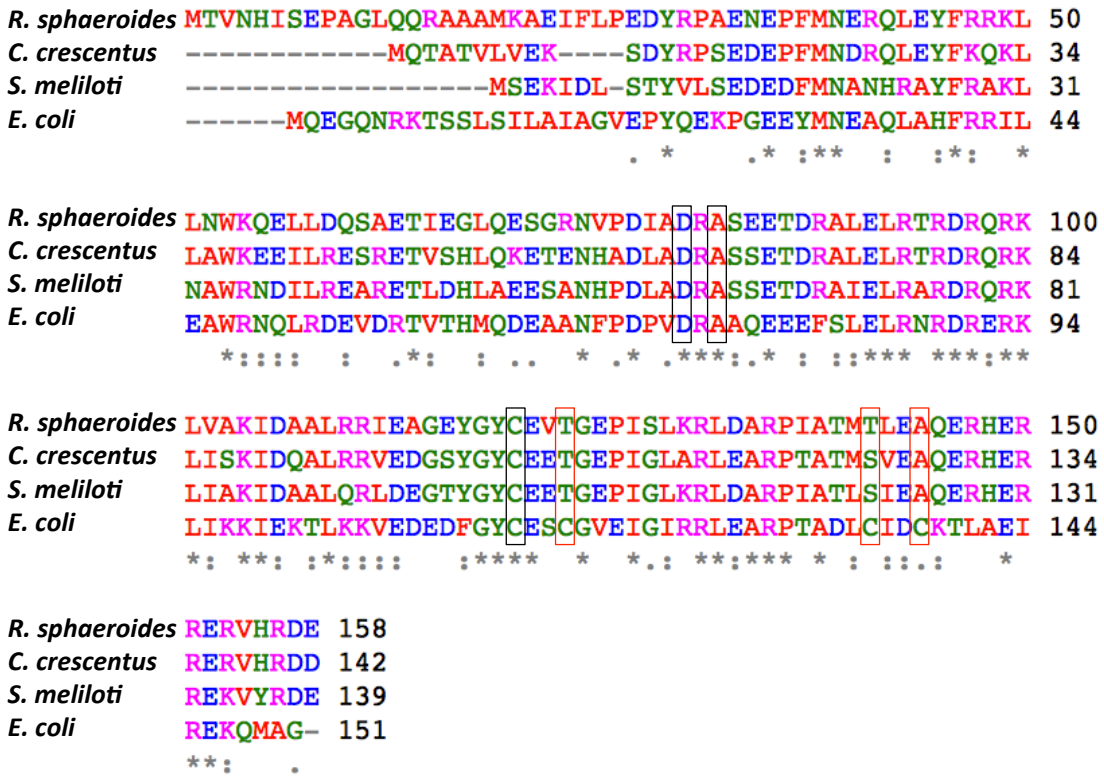
	lacUV5 ½ life (hours)	Fold reduction
No factor	16	-
RSP2654	9	<2
ppGpp	4	4
RSP2654 + ppGpp	0.33	48

E

	RNA-1 ½ life (hours)	Fold reduction
No factor	50	-
RSP2654	17	3
ppGpp	13	4
RSP2654 + ppGpp	1	50

Figure 4.7. Alignment of DksA protein sequences from *E. coli, R. sphaeroides, S. meliloti* and *C. crescentus* using the clustal W2 program (ebi.ac.uk). Brackets specifically align residues corresponding to *E. coli* residues D74, A76, C114, C117, C135 and C138.

Figure 4.7


R. sphaeroides MTVNHISEPAGLQQRAAAMKAEIFLPEDYRPAENEPFMNERQLEYFRRKL 50
C. crescentus ----------MQTATVLVEK----SDYRPSEDEPFMNDRQLEYFKQKL 34
S. meliloti ----------MSEKIDL-STYVLSEDEDFMNANHRAYFRAKL 31
E. coli ------MQEGQNRKTSSLSILAIAGVEPYQEKPGEEYMNEAQLAHFRRIL 44
R. sphaeroides LNWKQELLDQSAETIEGLQESGRNVPDIADRASEETDRALELRTRDRQRK 100
C. crescentus LAWKEEILRESRETVSHLQKETENHADLADRASSETDRALELRTRDRQRK 84
S. meliloti NAWRNDILREARETLDHLAEESANHPDLADRASSETDRAIELRARDRQRK 81
E. coli EAWRNQLRDEVDRTVTHMQDEAANFPDPVDRAAQEEEFSLELRNRDRERK 94
R. sphaeroides LVAKIDAALRRIEAGEYGYCEVTGEPISLKRLDARPIATMTLEAQERHER 150
C. crescentus LISKIDQALRRVEDGSYGYCEETGEPIGLARLEARPTATMSVEAQERHER 134
S. meliloti LIAKIDAALQRLDEGTYGYCEETGEPIGLKRLDARPIATLSIEAQERHER 131
E. coli LIKKIEKTLKKVEDEDFGYCESCGVEIGIRRLEARPTADLCIDCKTLAEI 144
R. sphaeroides RERVHRDE 158
C. crescentus RERVHRDD 142
S. meliloti REKVYRDE 139
E. coli REKQMAG- 151

Chapter 5

Conclusions and Future Directions

Conclusions

Data presented in this thesis has significantly advanced our understanding of the molecular mechanism of DksA action. In chapter 2, I describe an assay to quantitatively measure the affinity of DksA for RNAP (Lennon et al. 2009). This advance was important because it allowed us to more precisely dissect the effects of DksA and RNAP substitutions. Prior to this assay, we could only test how substitutions influenced the function of DksA. If we found that a DksA or RNAP substitution disrupted the effect of DksA, we could only speculate as to whether this was a result of decreased binding affinity or a loss of function that resulted from a step after binding. With this assay, we can explicitly test whether a substitution changed the affinity of DksA for RNAP. As such, data gathered using this assay has been useful to the conclusions of numerous studies (Blankschien et al. 2009, Lee et al. 2012, Lennon et al. 2012). Another finding described in chapter 2 is that DksA has an affinity for free RNAP (core and holo, ~100 nM) that is approximately 10-fold higher than RNAP in a competitor resistant open-complex (RNAP bound to the full consensus promoter, ~1000 nM). This result led us to the idea that DksA is selectively targeted to certain RNAP-promoter complexes over others. In line with this finding, it was recently found that DksA has decreased affinity for certain transcription elongation complexes when compared to free RNAP (Furman et al. 2012).

Chapter 3 describes the main focus during my graduate work (Lennon et al. 2012). Our lab has been interested in understanding the detailed mechanism of DksA-mediated transcriptional regulation. Because there was no structure of

the DksA-RNAP complex (and there is still no structure), we decided to build an evidence-based model of the DksA-RNAP complex using an alternative approach to traditional structural methods. Through site-specific incorporation of the cross-linkable amino acid benzophenylalanine (BPA) into DksA, we were able to establish physical constraints between DksA and RNAP that made it possible to employ a computer modeling program called HADDOCK to generate a model of the DksA-Core RNAP complex. This model fit nicely with numerous pieces of independent biochemical data and allowed us to hypothesize about the mechanism of DksA action. We subsequently established that there is a direct contact between the coiled-coil tip of DksA and the trigger loop (TL) domain of RNAP, and that this contact is required for RNAP to respond to DksA, an idea first postulated in an earlier publication from our lab (Rutherford et al. 2009). Finally, based upon the results above and the behavior of an interesting RNAP TL mutant (β' L930P/T931P), we propose that DksA mediates it's effects by targeting/inducing a partially folded form of the TL, allowing for allosteric changes in RNAP outlined previously (Rutherford et al. 2009).

One especially useful contribution to our lab from the work described in chapter 3 is the successful incorporation of BPA site-specifically into DksA and RNAP, which was not a trivial technique to establish. Our lab is currently using this technique to investigate a number of fundamental, long-standing questions in bacterial molecular biology, and I am very happy to have been able to establish this method in the lab.

In chapter 4, I describe the genetic and biochemical characterization of a *dksA*-like gene from *Rhodobacter sphaeroides* (*2654*). Plasmid supplied 2654 can complement Δ*dksA* cells for growth on media lacking amino acids and inhibits *rrnB* P1-*lacZ* activity *in* vivo. Using *E. coli* RNAP, 2654 directly inhibits *rrnB* P1 *in vitro*, and residues corresponding to the coiled-coil tip of DksA in 2654 are required for this function. 2654 can even activate transcription from the *hisG* promoter with ppGpp. Most importantly, both 2654 and ppGpp directly reduce the ½ life of *R. sphaeroides* RNAP-*lacUV5* or *RNA1* promoter complexes, a characteristic shared by DksA and ppGpp with *E. coli* RNAP. This effect is amplified when both factors are present, suggesting that *R. sphaeroides* and other alphaproteobacteria employ the DksA/ppGpp system to directly target RNAP to reduce promoter complex lifetime.

Future Directions

A major prediction of our model explaining DksA-mediated direct negative regulation (Chapter 3) is that the ultimate role of DksA is to influence downstream RNAP-DNA contacts. While we know that DksA changes the footprint of RP_I (Rutherford et al. 2009), incorporating BPA into RNAP and crosslinking to DNA has the potential to monitor this process at amino-acid/base resolution. Jared Winkelman of our lab has been working on a related project and either has or can synthesize the products needed to address. Along with Wilma Ross, they are currently testing whether DksA/ppGpp change promoter DNA-clamp contacts and/or RNAP core-shelf module contacts consistent with an open clamp and/or ratcheted RNAP conformation. This could provide precise details of how DksA/ppGpp are influencing contacts DNA-RNAP contacts during initiation, and more broadly, detect other predicted and unknown molecular motions during initiation.

Another remaining question related to DksA function is the molecular mechanism of positive control. In the case of promoters that are sensitive to transcription inhibition by DksA/ppGpp, each factor can cause inhibition to some degree in the absence of the other (Paul et al. 2004). In contrast to this mechanism, positive control by DksA/ppGpp requires that both factors be present for effect (Paul et al. 2005). It remains to be determined if DksA may interact differently with RNAP when binding to different promoters, if DksA itself may form a different conformation to control activity, why positive control requires ppGpp and negative control does not, etc. Further, it remains to be determined how

much of the genome is directly repressed or activated by DksA and ppGpp. While quantitative conclusions will require genome-wide approaches and *in vitro* analysis of each predicted promoter target individually, my work should provide greater detail about how DksA and ppGpp function. Potentially, this could also allow for the prediction of promoters directly inhibited or activated by DksA and ppGpp.

Another interesting area of research will be determining roles of DksA following transcription initiation in *E. coli*. One very exciting finding was that DksA plays a significant role in DNA replication under amino acid starvation (Tehranchi et al. 2010). In Δ*dksA* cells treated with serine hydroxamate to induce amino acid starvation, DNA replication elongation is significantly inhibited compared to WT (Tehranchi et al. 2010). The authors argue that DksA serves to prevent conflicts between the transcription and replication machinery (Tehranchi et al. 2010). While I accept the findings of this work and overall framework for how this is occurring, the molecular details of this process remain to be determined. This discovery is quite significant because it not only shows that DksA is used in replication, but that unexpected functions of DksA almost certainly remain to be elucidated. Recently, it has also been reported that DksA has effects on transcription elongation and termination (Furman et al. 2012, Furman et al. 2013). Unfortunately, these two reports are somewhat contradictory. In the case of Furman et al. 2012, the authors argue that DksA has broad effects on transcription elongation and termination complexes (using very high concentrations of a tighter binding DksA variant, N88I). However in Furman et al.

2013, the authors argue that DksA is restricted from interacting with elongation complexes by the lineage specific i6 domain of RNAP and that all bacterial genomes sequenced to date that contain a DksA homologue also contain i6. Therefore, more investigation is required to determine the precise effect of DksA on post-initiation transcription complexes.

Along these lines, we know that the affinity of DksA for RNAP core and holoenzyme is much higher than for RP_O (Lennon et al. 2009) and certain elongation complexes (Furman et al. 2013). However, we don't know the affinity of DksA for RP_C and RP_I. This is technically challenging because determining affinity requires that the RNAP population be homogenous, which is relatively simple to achieve for core, holo and RP_O, but not for RP_C and RP_I. Further, how other secondary channel binding proteins influence this process, and vice-versa, is unclear.

Not surprisingly, DksA may target different genes in bacteria other than *E. coli*. Our work with *R. sphaeroides* RSP2654 (a DksA homologue) in chapter 4 shows that DksA and ppGpp directly destabilize RNAP-promoter complexes in α-proteobacteria. We suggest that the overall molecular mechanism of DksA/ppGpp is conserved in alphaproteobacteria, but that the system has been rewired to regulate different genes. This model is consistent with previous studies showing that *R. sphaeroides* synthesizes ppGpp, but does not accumulate ppGpp upon leucine starvation (Acosta and Lueking, 1987), and that *Caulobacter crescentus* and *Sinorhizobium meloiti* (α-proteobacteria) synthesize ppGpp in response to glucose or ammonium starvation, but not after amino acid starvation

alone, and can halt rRNA synthesis in the absence of ppGpp accumulation (Boutte and Crosson, 2011, Krol and Becker 2011). The promoters regulated by DksA and ppGpp have yet to be determined in *R. sphaeroides*; planned genome-wide studies should shed light on this issue.

Our lab's work suggests that the concentration of DksA always remains constant in *E. coli* because DksA feedback inhibits its own promoter (Paul et al. 2004, Rutherford et al. 2007, Chandrangsu et al. 2011). While there no evidence indicating that DksA levels change, additional growth conditions should be tested.

One longstanding observation is that Δ*dksA* and $ppGpp^0$ *E. coli* cells cannot grow in the absence of amino acids. However, it is unclear why this occurs and if this phenotype is related to the activation of amino acid biosynthesis promoters. Elucidating the mechanism of this phenotype may provide insights into new targets of DksA/ppGpp repression/activation, or perhaps even new roles of DksA/ppGpp in the cell when dealing with less than optimal nutritional conditions.

References

Acosta R, Lueking DR. 1987. Stringency in the absence of ppGpp accumulation in *Rhodobacter sphaeroides*. *J Bacteriol* **169**: 908-912.

Blankschien M, Lee J, Grace E, Lennon C, Halliday J, Ross W, Gourse R, Herman C. 2009. Super DksAs: Substitutions In DksA enhancing its effects on transcription initiation. *EMBO J* **28:** 1720–1731.

Boutte CC, Crosson S. 2011. The complex logic of stringent response regulation in *Caulobacter crescentus*: starvation signaling in an oligotrophic environment. *Mol Microbiol* **80**: 695-714.

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

Furman R, Sevostyanova A, Artsimovitch I. 2012. Transcription initiation factor DksA has diverse effects on RNA chain elongation. *Nucleic Acids Res* **40:** 3392-3402.

Furman R, Tsodikov OV, Wolf YI, Artsimovitch I. 2013. An insertion in the catalytic trigger loop gates the secondary channel of RNA polymerase. *J Mol Biol* **425:** 82-93.

Krol E, Becker A. 2011. ppGpp in *Sinorhizobium meliloti*: biosynthesis in response to sudden nutritional downshifts and modulation of the transcriptome. *Mol Microbiol* **81** 1233-1254.

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

Lennon C, Gaal T, Ross W, Gourse R. 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 CW, Ross W, Martin-Tumasz S, Toulokhonov I, Vrentas CE, Rutherford ST, Lee JH, Butcher SE, Gourse RL. 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.

Paul B, Barker M, Ross W, Schneider D, Webb C, Foster J, Gourse R. 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, Barker M, Gourse R. 2005. DksA potentiates direct activation of amino acid promoters by ppGpp. *Proc Natl Acad Sci USA* **102:** 7823-7828.

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

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

Tehranchi A, Blankschien M, Zhang Y, Halliday J, Srivatsan A, Peng J, Herman C, Wang J. 2010. The transcription factor DksA prevents conflicts between DNA replication and transcription machinery. *Cell* **141:** 595-605.

Appendix A

Blankschien M, Lee J, Grace E, Lennon C, Halliday J, Ross W, Gourse R, Herman C. 2009. Super DksAs: Substitutions in DksA enhancing its effects on transcription initiation. *EMBO J* **28**: 1720-1731.

I performed experiments used in this manuscript helping to explain the mechanism by which “Super DksA” mutants exhibit a gain of function. This paper was a collaboration between our laboratory and that of Dr. Christophe Herman (Baylor College of Medicine). Using a binding assay developed to measure the affinity of DksA for RNAP (Lennon et al. 2009 and Chapter 2), I was able to show that DksAs with gain of function phenotypes isolated by the Herman lab bound to RNAP tighter than WT DksA. Below are excerpts directly from the paper based on my results.

Abstract

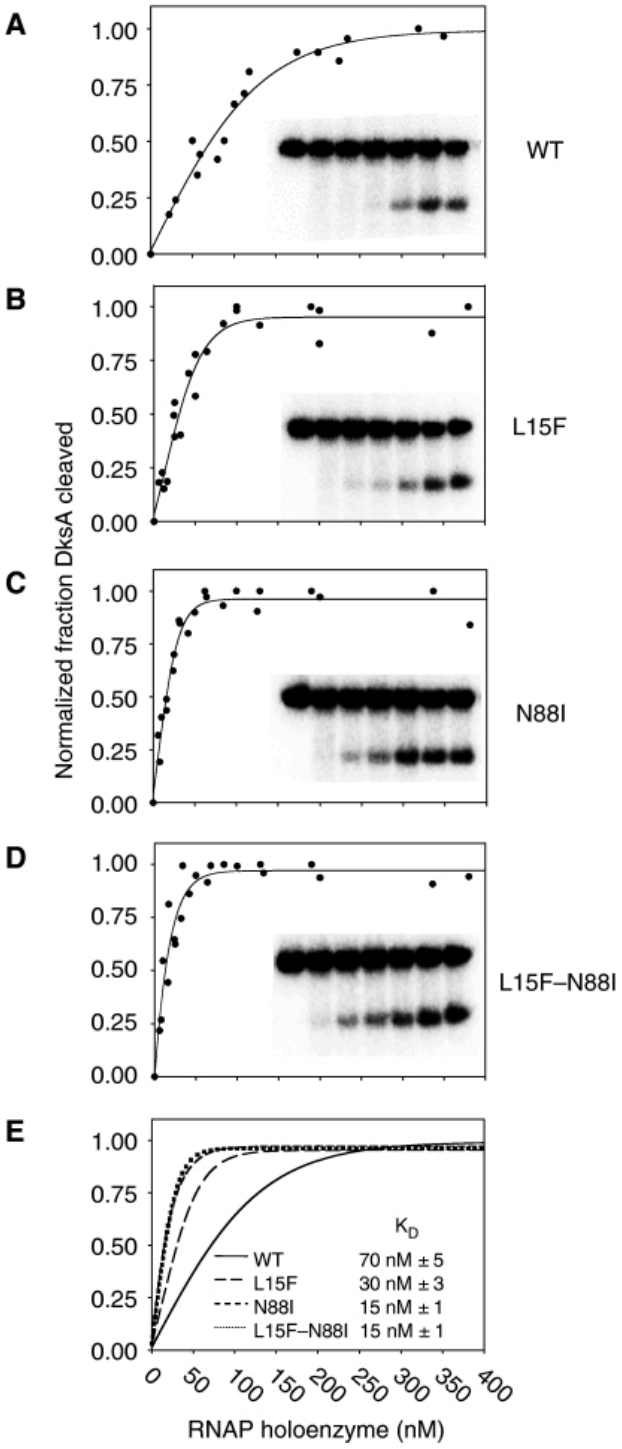
At specific times during bacterial growth, the transcription factor DksA and the unusual nucleotide regulator ppGpp work synergistically to inhibit some Escherichia coli promoters (e.g. rRNA promoters) and to stimulate others (e.g. promoters for amino-acid synthesis and transport). However, the mechanism of DksA action remains uncertain, in part because DksA does not function like conventional transcription factors. To gain insights into DksA function, we identified mutations in dksA that bypassed the requirement for ppGpp by selecting for growth of cells lacking ppGpp on minimal medium without amino acids. We show here that two substitutions in DksA, L15F and N88I, result in higher DksA activity both in vivo and in vitro, primarily by increasing the apparent affinity of DksA for RNA polymerase (RNAP). The mutant DksA proteins suggest potential roles for ppGpp in DksA function, identify potential surfaces on DksA crucial for RNAP binding, and provide tools for future studies to elucidate the mechanism of DksA action.

DksA variants bind more tightly to RNAP than wild-type DksA

To alter RNAP function, DksA binds in the secondary channel of RNAP and reduces promoter DNA interactions by an allosteric mechanism (Rutherford et al, 2009). In theory, the increased function of the mutant DksA proteins could result from improved binding to RNAP and/or more effective transmission of effects of binding to the sections of RNAP that interact with the promoter. As an estimate of the relative binding affinities of the wild-type and mutant DksA proteins for RNAP holoenzyme, we utilized an assay in which Fe^{2+} was substituted for Mg^{2+} at the RNAP active site, resulting in production of hydroxyl radicals that cleave the coiled-coil tip of DksA (Perederina et al, 2004) (Figure 8A and D). Apparent binding constants for RNAP with the wild-type and mutant DksA proteins were determined by measurement of the fraction of ^{32}P-labelled, N-terminally heart muscle kinase (HMK)-tagged DksA cleaved at different RNAP concentrations, normalized to the fraction cleaved at saturating RNAP concentration (Lennon et al. 2009 and see chapter 2). Apparent binding affinities of the wild-type and mutant DksA proteins for RNAP are compared in Figure 8E. We determined that N88I–DksA and N88I–L15F–DksA bind to RNAP about five-fold more tightly and L15F–DksA binds about two-fold more tightly than wild-type DksA. These results suggest that much (or all) of the increase in DksA function derives from increased affinity of the mutant DksA proteins for RNAP.

Figure AA.1 (Figure 8 from paper). The L15F, N88I, and L15F–N88I substitutions increase the affinity of DksA for RNAP holoenzyme. The binding affinity of ^{32}P-labelled HMK-His6-tagged DksA (wild-type, L15F, N88I, or L15F–N88I) was measured quantitatively using an RNAP-dependent localized iron cleavage assay (CWL and RLG, unpublished). In panels A–D, the normalized fraction of DksA (wild-type or mutant) cleaved is plotted versus the concentration of RNAP holoenzyme in the reaction. Curves were generated from at least six separate experiments. Inset: representative gel images in which the lanes from left to right contain 0, 4, 8, 16, 32, 64, or 128nM RNAP holoenzyme. (E) Binding curves are displayed together for compar- ison. WT (solid line), L15F (long dashes), N88I (short dashes), and L15F–N88I DksA (filled circles). Apparent binding constants are shown in the inset.

Figure AA.1


A
B
C
D
E
1.00
0.75
0.50
0.25
0.00
WT
L15F
N88I
L15F–N88I
Normalized fraction DksA cleaved
K_D
WT 70 nM ± 5
L15F 30 nM ± 3
N88I 15 nM ± 1
L15F–N88I 15 nM ± 1
0 50 100 150 200 250 300 350 400
RNAP holoenzyme (nM)

References

Lennon C, Gaal T, Ross W, Gourse R. 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.

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

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

Appendix B

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

My contribution to this paper was mainly the preparation of numerous DksA and Gre factor proteins containing different substitutions for *in vitro* analysis. This required construction of various plasmids and purification of numerous proteins. Some of the experiments in the paper were initially performed by me, but were ultimately repeated by JHL prior to publication.

Abstract

Escherichia coli DksA works in conjunction with the small-molecule ppGpp to regulate transcription initiation negatively or positively, depending on the identity of the promoter. DksA is in a class of transcription factors that do not bind directly to DNA such as classical repressors or activators but rather bind in the RNA polymerase (RNAP) secondary channel such as the transcription elongation factors GreA and GreB in E. coli and TFIIS in eukaryotes. We found that substitution for either of two residues in its coiled-coil tip, D74 or A76, eliminates DksA function without affecting its apparent affinity for RNAP. The properties of DksA-Gre factor chimeras indicated that the coiled-coil tip is responsible for the DksA-specific effects on open complex formation. A conservative substitution at position 74, D74E, resulted in a loss of DksA function in both negative and positive control, and an E44D substitution at the analogous position in GreA resulted in a gain of function in both negative and positive control. That a single methylene group has such an extraordinary effect on these transcription factors highlights the critical nature of the identity of coiled-coil tip interactions with RNAP for open-complex formation.

Appendix C

What is the physical location of σ^{70} region 1.1 during transcription initiation?

Jared Winkleman, Michael Maloney, Albert Chen and I were involved in this project. Albert Chen will be continuing this work.

Region 1.1 of *E. coli* σ^{70} is proposed to be an auto-inhibitory domain, preventing DNA binding by free σ^{70} in the absence of core RNAP (Dombroski et al. 1993). σ^{70} region 1.1 is thought to prevent DNA binding by associating directly with σ^{70} regions 2 and 4, which bind the -10 and -35 promoter elements, respectively (Dombroski et al. 1993, Schwartz et al. 2008). The position of σ^{70} 1.1 is predicted to undergo major changes during transcription initiation. Briefly, upon holoenzyme formation, σ^{70} 1.1 is expected to occupy the main channel of RNAP and exit to accommodate DNA entry prior to open-complex formation.

The goal of this project is to describe the physical location of σ^{70} region 1.1 during different stages of transcription initiation. This question is important because σ^{70} region 1.1 has been shown to drastically effect transcription initiation, but the mechanism by which this occurs is completely unknown. Deletion of this domain can result in strong inhibition or activation of transcription from some promoters when compared to full-length σ^{70} (Wilson and Dombroski 1997, Vuthoori et al. 2001). Interestingly, our lab has even observed that σ^{70} Δ1.1 holoenzyme forms much more stable complexes with the *rrnB* P1 promoter (unpublished).

As a first approach, we cross-linked BPA containing RNAPs to ^{32}P-labeled σ^{70}. I will summarize the main findings of this work below (these results should be repeated if included in a manuscript). While this approach has yielded some useful results, we have decided to incorporate BPA into σ^{70} region 1.1, rather than into RNAP, to better address this question. The main driving force behind this change is that we are interested in determining where σ^{70} region 1.1 is upon

DNA addition. We have concluded that BPA incorporation directly into σ^{70} region 1.1 gives us a better chance to determine contacts between this domain and RNAP during transcription initiation.

Approximately 40 nM ^{32}P- σ^{70} (label at internal kinase recognition site) was incubated with ~100 nM BPA substituted β' or β RNAP and exposed to 365 nm UV light for 5 minutes. We observed a cross-link between σ^{70} and β' when BPA was present at positions 133, 142, 200, 202, 212, 314, 334, 1148 and 1151, but not when BPA was at positions 74, 264, 312, 1149, 1154, 1160, 1170, 1176, 1178, 1189, 1192, 1194, 1203, 1211 and 1224 (Figure AC.1 and data not shown). We observed a cross-link between σ^{70} and β when BPA was present at positions 62, 63, 180, 197, 201, 202, 203, 268, 324, 352, 356, 367, 368, 371, 374 and 378, but not when BPA was at positions 54, 151, 200, 211, 227, 331 and 332 (Figure AC.2). The fraction of σ^{70} cross-linked varied considerably, with some as low as 1% and others as high as 20%. We cannot conclude that these cross-links are specifically to region 1.1, but it is likely given that most of these positions are in the main DNA binding channel. I also examined many of these cross-links in the presence of an excess of the σ^{70} full-consensus promoter in an attempt to locate region 1.1 in the open complex. Results were initially intriguing and indicated that region 1.1 undergoes a major conformational change as many holoenzyme cross-links were abolished or reduced in the open complex. Unfortunately, the results were not consistent, possibly due to the age of RNAP-BPA preps. These experiments with DNA are potentially useful and interesting, but should be interpreted with caution. Recently (Jan 2013) BPA was

incorporated directly into σ^{70} region 1.1 (Figure AC.3), and should allow us to better address the question of where this domain is during transcription initiation.

References

Dombroski AJ, Walter WA, Gross CA. 1993. Amino-terminal amino acids modulate sigma-factor DNA-binding activity. *Genes Dev* **7**: 2446-2455.

Vuthoori S, Bowers CW, McCracken A, Dombroski AJ, Hinton DM. 2001. Domain 1.1 of the sigma70 subunit of Escherichia coli RNA polymerase modulates the formation of stable polymerase/promoter complexes. *J Mol Biol* **309**: 561-572.

Schwartz EC, Shekhtman A, Dutta K, Pratt MR, Cowburn D, Darst S, Muir TW. 2008. A full-length group 1 bacterial sigma factor adopts a compact structure incompatible with DNA binding. *Chem Biol* **15**: 1091-1103.

Wilson C, Dombroski AJ. 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.

Figure AC.1

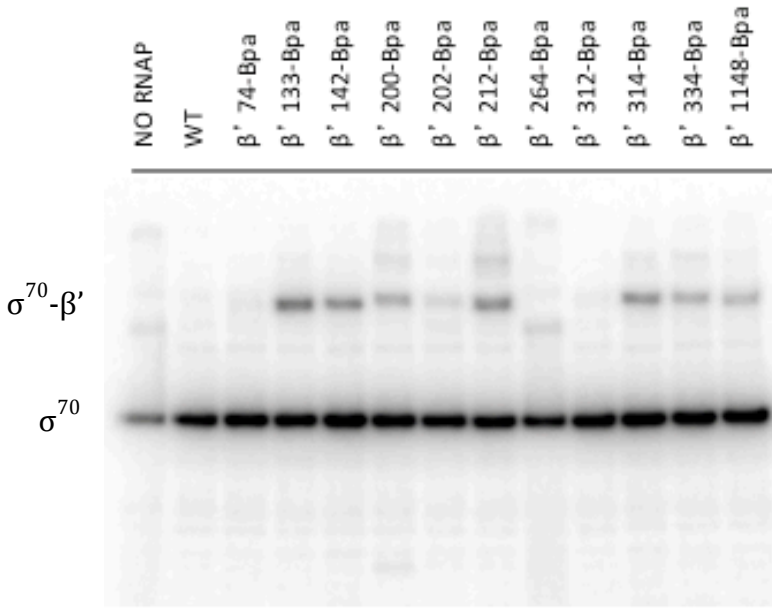

NO RNAP
WT
β' 74-Bpa
β' 133-Bpa
β' 142-Bpa
β' 200-Bpa
β' 202-Bpa
β' 212-Bpa
β' 264-Bpa
β' 312-Bpa
β' 314-Bpa
β' 334-Bpa
β' 1148-Bpa
σ70-β'
σ70

Figure AC.2

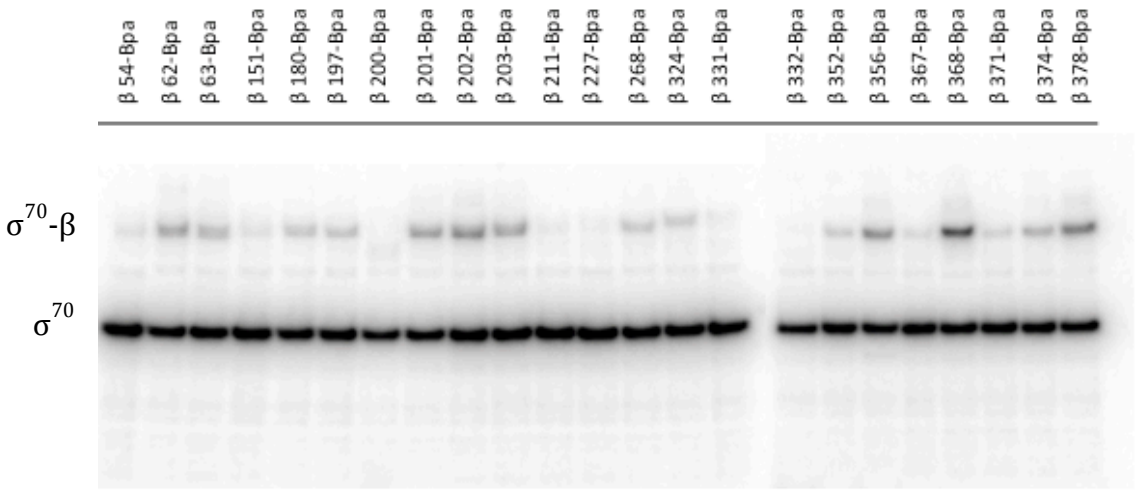

β 54-Bpa
β 62-Bpa
β 63-Bpa
β 151-Bpa
β 180-Bpa
β 197-Bpa
β 200-Bpa
β 201-Bpa
β 202-Bpa
β 203-Bpa
β 211-Bpa
β 227-Bpa
β 268-Bpa
β 324-Bpa
β 331-Bpa
β 332-Bpa
β 352-Bpa
β 356-Bpa
β 367-Bpa
β 368-Bpa
β 371-Bpa
β 374-Bpa
β 378-Bpa
σ70-β
σ70

Figure AC.3