2020

Structural and Functional Studies of the Unconventional Proteobacterial Transcription Activator CRL in Complex with the Transcription Machinery

Alexis Jaramillo Cartagena

STRUCTURAL AND FUNCTIONAL STUDIES OF THE UNCONVENTIONAL PROTEOBACTERIAL TRANSCRIPTION ACTIVATOR CRL IN COMPLEX WITH THE TRANSCRIPTION MACHINERY

A Thesis Presented to the Faculty of

The Rockefeller University

in Partial Fulfillment of the Requirements for

the degree of Doctor of Philosophy

by

Alexis Jaramillo Cartagena

June 2020

STRUCTURAL AND FUNCTIONAL STUDIES OF THE UNCONVENTIONAL PROTEOBACTERIAL TRANSCRIPTION ACTIVATOR CRL IN COMPLEX WITH THE TRANSCRIPTION MACHINERY

Alexis Jaramillo Cartagena, Ph.D.

The Rockefeller University 2020

The σ^S subunit of RNA polymerase (RNAP) is the master regulator of stress responses in many Gram-negative bacteria. This alternative σ factor assembles with the core RNA polymerase to initiate the transcription of genes needed to survive different environmental changes. Crl is a small protein that activates the transcription of σ^S-dependent genes. In contrast to most transcription activators, Crl does not bind DNA to help recruit RNA polymerase and instead interacts directly with σ^S.

At the outset of my research, little was known about how the binding of Crl to σ^S leads to transcription activation. It was not clear if in addition to σ^S, Crl also made specific interactions with core RNAP. Using structural biology, molecular biology, biochemical and biophysical techniques, I gained novel insight into the unusual mechanism of Crl. This research validated and expanded on previous studies delineating the Crl/σ^S interaction and showed how a previously uncharacterized interaction between Crl and the β' subunit of RNAP is critical for full transcription activation by Crl. This work advances our understanding of an unconventional mode of transcription activation in bacteria that might be more widespread than currently known.

Chapter 1 provides background on bacterial transcription, σ factors, aspects of

regulation, and closes with an introduction to Crl. Chapter 2 describes an approach that can be used to gain insight into the regulons of transcription factors like Crl. Most of the research in this thesis is presented in Chapter 3, which uses biochemical and biophysical approaches to elucidate how Crl activates transcription. Appendix A presents an attempt to study the surface of Crl that interacts with β'. Appendix B, briefly shows an attempt to investigate an additional mechanism by which Crl can activate transcription.

Amparo, Orlando, y Robison
Estoy muy agredecido por tenerlos en vida. Los quiero mucho.

Acknowledgements:

In the early Fall of 2016, Nikhil and I had just finished processing our first set of cryo-EM data thanks to the great help of James. We had been working really hard that previous summer, and we finally got the first reconstruction of our complex at a resolution of ~9 Å. Even at this low-resolution, we unambiguously observed density near σ^{S}_{2} that no one had seen before. We showed this reconstruction to Liz and Seth and they were filled with enthusiasm seeing that our activator, Crl, was present in a transcription initiation complex in a way that was consistent with previous evidence. That reconstruction launched a series of mechanistic of studies that have been one of my most rewarding pursuits. This event highlights a lot of elements for which I am profoundly grateful.

First, I am really grateful for the opportunity to have mentored a diligent and dedicated individual like Nikhil. It filled me with pride as I saw his intellectual growth during his stay in our lab. In addition to Nikhil, the work I have accomplished in the Darst/Campbell Lab has been possible thanks to the great help and support of a number of great colleagues: James, Jin-Young, Ruth, Mira, Hande, Lizzy, Fred, Madge, Andrey, Jesse, Nathaniel, Courtney, Deena, Mengyu, Brandon, Ruby, and Josh. I am also extremely fortunate to have been guided by Liz and Seth, who not only are terrific scientists and devoted professors, but they are compassionate and caring individuals who truly want to nurture the scientific and personal the lives of everyone they mentor. Witnessing their brilliant minds at-work has been an inspirational. Similarly, receiving the guidance my committee members Chris, Luciano, and Tarun has been a great honor and I will always be grateful for their contributions to my scientific growth. Additionally,

I am grateful to Derek Tan, Scott Blanchard, Margie, Kathleen, Francine, and again to Tarun for all the effort they put to make the Chemical Biology program a terrific scientific training environment. My project was a collaboration with Rick Gourse, Wilma Ross, and Amy Banta. Throughout the past four years we have had numerous interactions and it has been a pleasure to collaborate with them. More recently, I had the fortune to collaborate with Jamie Link and his lab, which yielded and we had very fruitful scientific results. I would also like to thank Dom, Brian and everyone in the Chait Lab as well and Mark, Johanna, and Tom Walz for their kindness and support in specialized techniques that made my research possible. I am very grateful to all members of the Dean's Office: Cris, Stephanie, Marta, Emily, Sid, and Kristen who have provided unlimited help during my journey. Lastly, I am thankful for the opportunity to have spent the last few years at this superb institution surrounded by phenomenal people, resources and assistance that make Rockefeller an amazing place to conduct science.

Table of Contents

List of Figures

List of Tables

List of abbreviations

BTH	bacterial two-hybrid assay
CT	Clamp toe
CTD	C-terminal domain
Cryo-EM	Cryo-electron microscopy
E	Core RNA polymerase
Eσ	RNA polymerase holoenzyme
Eco	*Escherichia coli*
K_d	Dissociation constant
kDa	Kilodalton
ITC	Isothermal titration calorimetry
MST	Micro-scale thermophoresis
NMR	Nuclear magnetic resonance
NTP	Nucleotide triphosphate
NTD	N-terminal domain
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
Pm	*Proteus mirabilis*
RNAP	RNA polymerase
RNA-Seq	RNA sequencing
RPc	RNAP holoenzyme closed promoter complex
RPo	RNAP holoenzyme open promoter complex
Sty	*Salmonella enterica* serovar Typhimurium
Vh	*Vibrio harveyi*

Chapter 1: **Introduction**

Bacterial RNA polymerase

Transcription is the biological process where genetic information encoded as DNA is used as a template to synthesize RNA. RNA polymerase (RNAP, E) is the central enzyme of transcription (Archambault & Friesen, 1993). The structure of E resembles a crab claw with pincers comprising the large β and β' subunits (Zhang et al., 1999)(**Fig.1.1**). The large channel between the pincers contains the catalytic site for RNA synthesis and accommodates nucleic acids during transcription (Korzheva et al., 2000; Vassylyev et al., 2002). The α subunit dimer serves as a scaffold for the assembly of E (Zhang & Darst, 1998), plays roles in DNA binding, and interacts with transcription factors (Dove, Darst, & Hochschild, 2003; Ishihama, 1981; Niu et al., 1996; Ross et al., 1993). The small ω subunit appears to function as a dedicated chaperone for the β' subunit (Minakhin et al., 2001) that promotes assembly of E and is also a target of transcription regulation (Mathew & Chatterji, 2006; Ross et. al, 2013).

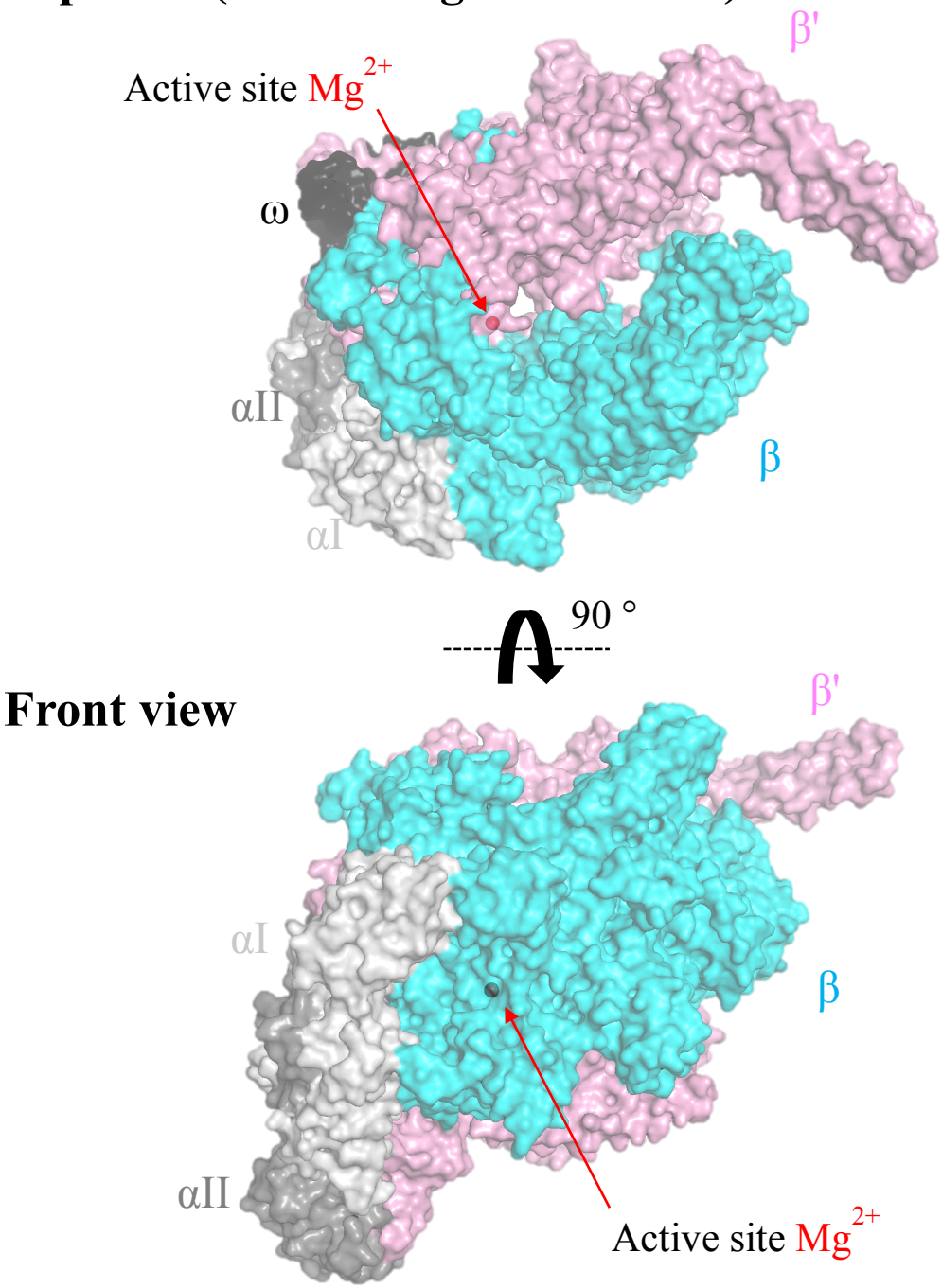


Fig. 1.1. Structure of Bacterial core RNAP. Structural models of core RNA polymerase enzyme (figure made using PDB ID 6EDT, Boyaci et al., 2019). Protein subunits are represented as individually colored surfaces: β is shown in cyan, β' is pink, α1 is light grey, α1 is dark grey, ω is black.

σ-factors and the σ cycle

To bind a specific DNA sequence in duplex DNA (the promoter), E assembles with a promoter recognition subunit called the sigma (σ) factor to form the RNA polymerase holoenzyme (Eσ) (Burgess et al., 1969; Murakami, Masuda, & Darst, 2002; Travers & Burgess, 1969) (**Fig. 1.2**). Bacteria have at minimum one primary σ factor, which directs transcription of housekeeping functions during exponential growth (Gruber & Gross, 2003). Bacteria also contain a repertoire of alternative σ factors which are expressed and activated in specific cellular settings (Österberg et al., 2011). Some bacteria, like *Sorangium cellulosum*, contain 109 σ factors (Han et al., 2013). *Escherichia coli* (*Eco*) contains seven σ factors (σ^{70}, σ^{S}, σ^{F}, σ^{H}, σ^{FecI}, σ^{E}, and σ^{54}) that compete for binding E, each one having different expression levels and binding affinity to E (Burgess, 2001; Maeda, Fujita, & Ishihama, 2000) (**Fig. 1.2**) (**Table 1.1**).

Table 2.1 – Cellular levels of σ and their binding affinities to E in *E. coli*

σ factor	***K*$_d$ (nM)**	**σ molecules/cell in exponential growth**	**Conc. in exponential growth (μM)***	**Conc. in stationary growth (μM)***
σ^{70} (σ^{D})	**0.26**	**700**	**1.16**	**1.16**
σ^{N} (σ^{54})	0.30	110	0.18	0.18
σ^{F} (σ^{28})	0.74	370	0.61	0.61
σ^{H} (σ^{32})	1.24	<10	<0.17	n/a
σ^{FecI} (σ^{18})	1.73	<1	<0.02	n/a
σ^{E} (σ^{24})	2.43	<10	<0.17	n/a
σ^{S} (σ^{38})	**4.26**	**<1**	**<0.02**	**0.37**

From Maeda et. al. 2000. *The molar concentrations were calculated using the reported σ numbers and assuming an *E. coli* cell volume to be 1 fL.

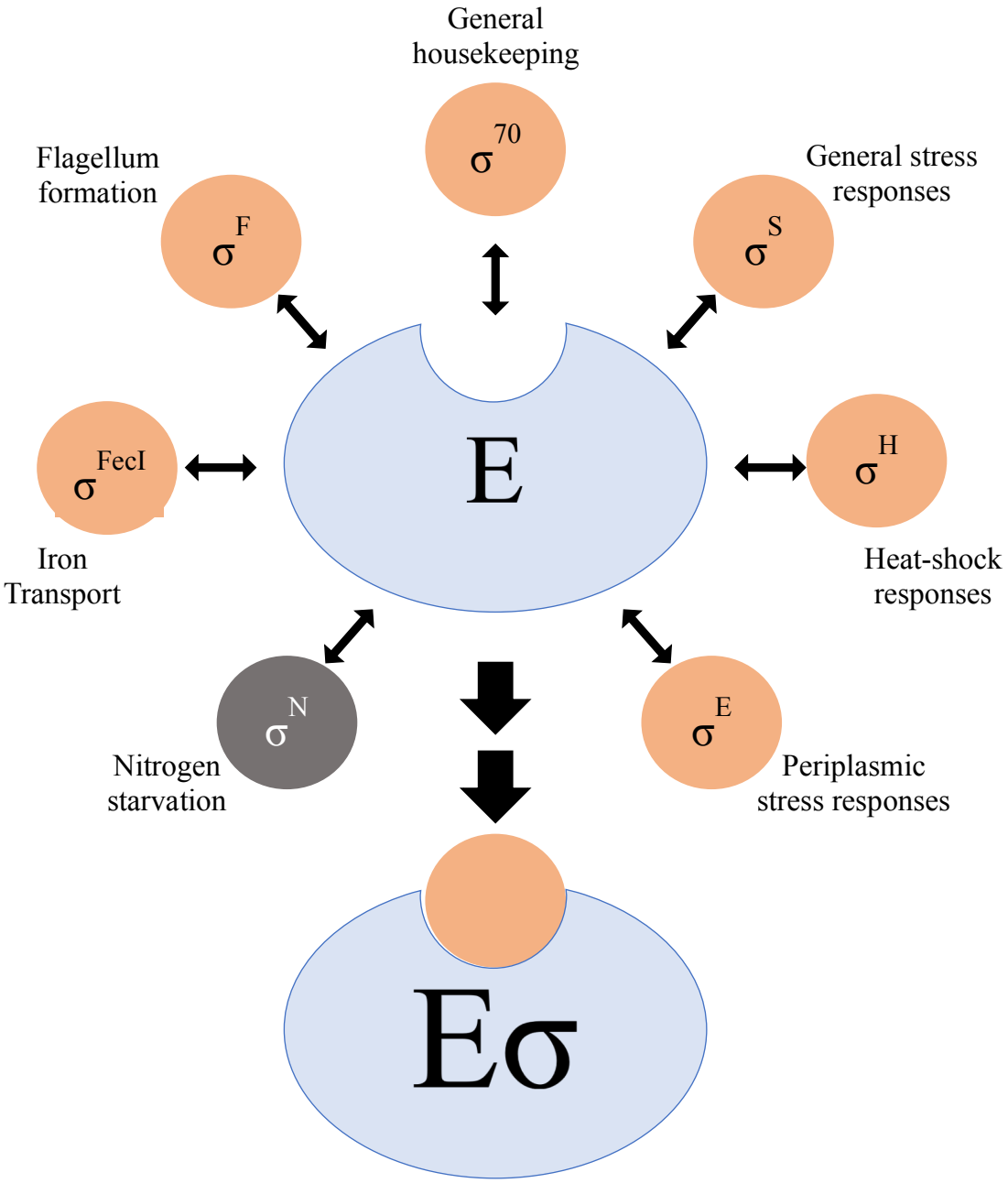


Fig. 1.2. σ factors in *Eco* and holoenzyme. Eco has seven σ factors. They are divided into two families: the σ^{70} family (tan) and the σ^{54} family (grey). The order of binding affinity strength for each σ to E is denoted. E and σ assemble to form an RNA polymerase holoenzyme Eσ. Maeda, Fujita, & Ishihama, 2000.

σ factors are divided into two evolutionarily and structurally unrelated families: the large and diverse σ^{70} family (Lonetto et al., 1992) and the smaller σ^{54} (also called σ^{N}) family (Buck et al., 2000). Eσ containing σ^{70}-family members is able to both recognize specific promoters and to initiate transcription (Lonetto et al., 1992). In contrast, $E\sigma^{54}$ can recognize specific promoter sequences but requires additional enhancers and energy in the form of ATP hydrolysis to initiate transcription (Buck et al., 2000). The vast majority of bacterial transcription is directed by σ^{70}-family members, and they will be the focus of the rest of this chapter and the work presented in this thesis.

The association and dissociation of specific σ factors with E plays important roles in transcription regulation in a process called the σ cycle (Mooney, Darst, & Landick, 2005) (**Fig. 1.3**). This cycle is key to the regulation of gene expression because σ-factor switching can induce changes in the expression of thousands of genes (Gruber & Gross, 2003). In the context of Eσ, σ^{70} factors make several interactions with the promoter principally at elements centered 10 and 35 bases upstream of the transcription start site (Gross et al., 1998) known as the -10 and -35 promoter elements (**Fig. 1.4**). The initial recognition and binding of Eσ to the promoter forms a closed complex (RPc) where the DNA remains double stranded (Saecker, Record, & DeHaseth, 2011). After RPc formation, σ promotes the separation of the DNA duplex strands, a process called DNA melting. In a series of isomerization steps, a set of Watson–Crick base pairing interactions are broken, which results in the formation of the transcription bubble, which extends from the -11 to the +2 base (numbering with respect to the transcription start site at +1) of the template DNA (Siebenlist, 1979; Siebenlist, Simpson, & Gilbert, 1980). This forms an open promoter complex (RPo), which exposes DNA bases to serve as the

template for RNA synthesis (Bae et al., 2015) (**Fig. 1.5**). At this stage, ribonucleoside triphosphate substrates (NTPs) enter the active site and base pair with the template DNA strand to start the synthesis of RNA transcripts.

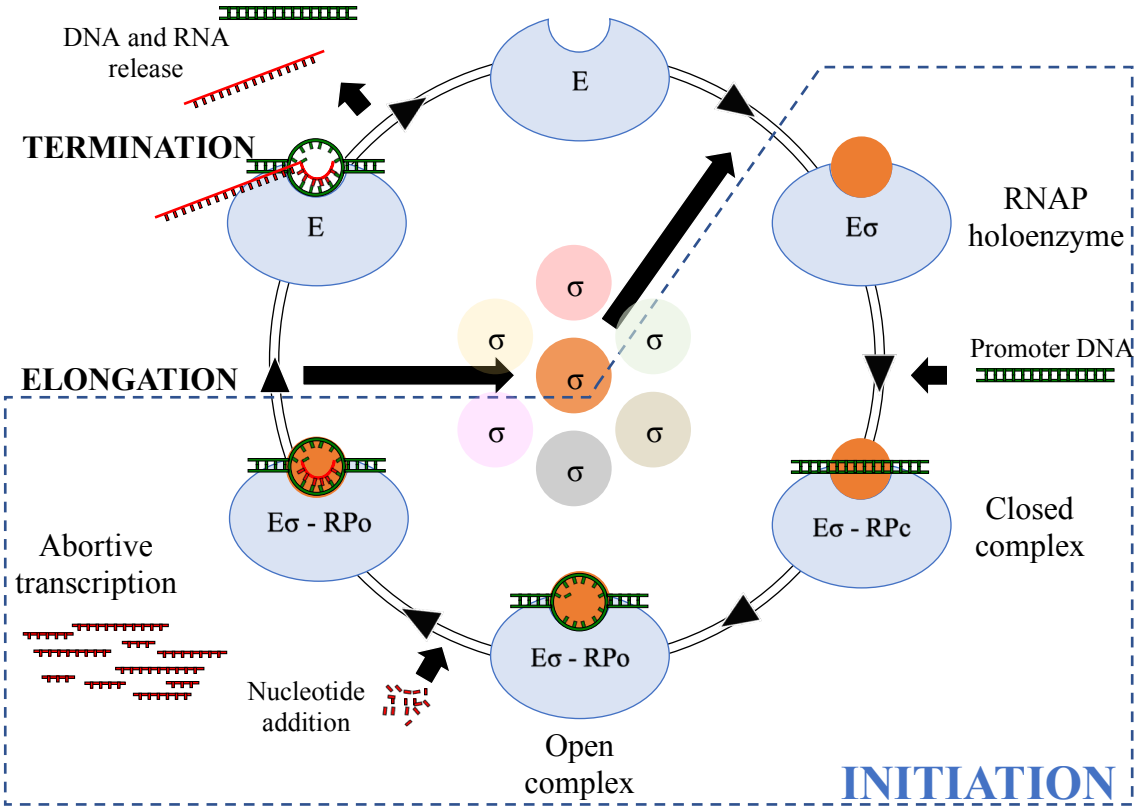


Fig. 1.3. σ cycle in bacteria. Transcription initiation (blue-dashed area) starts with the assembly of E (blue oval) and a σ (colored circle) to form a holoenzyme (Eσ) which recognizes and binds specific promoter elements to form a closed-promoter complex (RPc). σ facilitates melting of promoter DNA and the complex transitions between intermediate states and eventually forms a transcription bubble, yielding a catalytically active open-promoter complex (RPo). In the presence of NTP substrates, the complex engages in abortive transcription where it produces short RNA transcripts until it the interactions between E and σ are disrupted. As the complex enters into productive elongation σ is released and can associate with another E to initiate another cycle. Elongation continues until E encounters a termination sequence or is assisted by termination factors. (Mooney, Darst, & Landick, 2005).

Fig. 1.4. (Previous page) σ^{70}/promoter interactions and σ^{70} structural architecture. The σ factor and the C-terminal domains (CTD) of the α subunits within E are configured to bind promoter DNA through specific interactions.

A. Schematic of promoter elements (rectangles) showing their interactions with domains of σ and α (circles):

- Upstream (US) element interacts with the α CTDs (light red and light orange)
- -35 element interacts σ_4 (pink)
- Extended -10 element (EXT -10) interacts with σ_3 (light green)
- -10 element interacts with σ_2 (light blue)
- Discriminator (DISC) interacts with region $\sigma_{1.2}$ (dark blue), which is located within σ_2

The promoter schematic also shows the transcription start site (TSS, position +1, green) and the downstream (DS) DNA. Ruff et al, 2015.

B. σ^{70} is composed of several structural domains that interact with promoter DNA and core RNAP. The individual domains of σ^{70} (PDB 4LK1) are colored individually and shown as ribbons with superimposed with transparent surfaces. Promoter DNA is included to show the contact interactions between the -10 element (yellow) with σ^{70}_2 and -35 element (yellow) with σ^{70}_4.

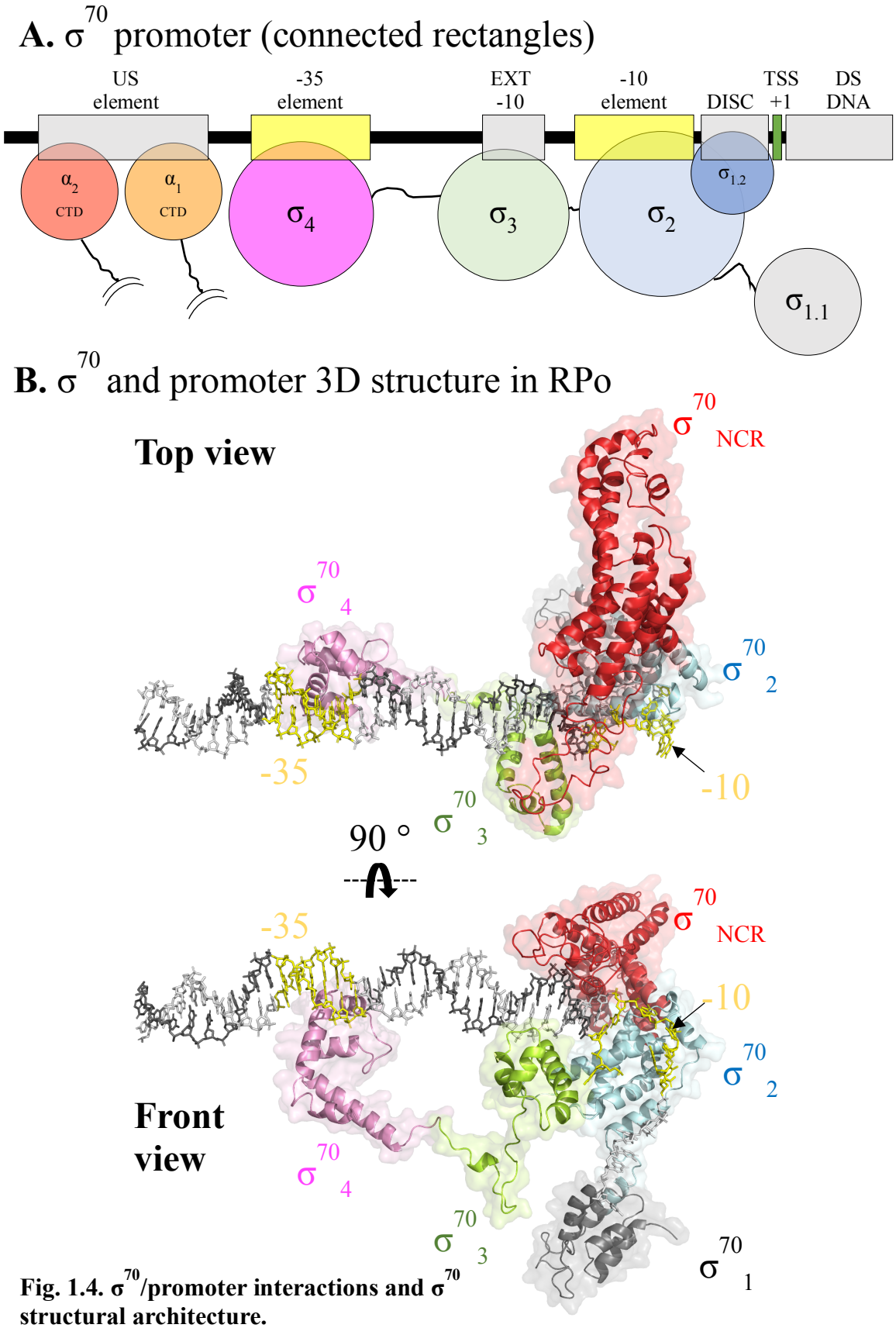


Fig. 1.4. σ^{70}/promoter interactions and σ^{70} structural architecture.

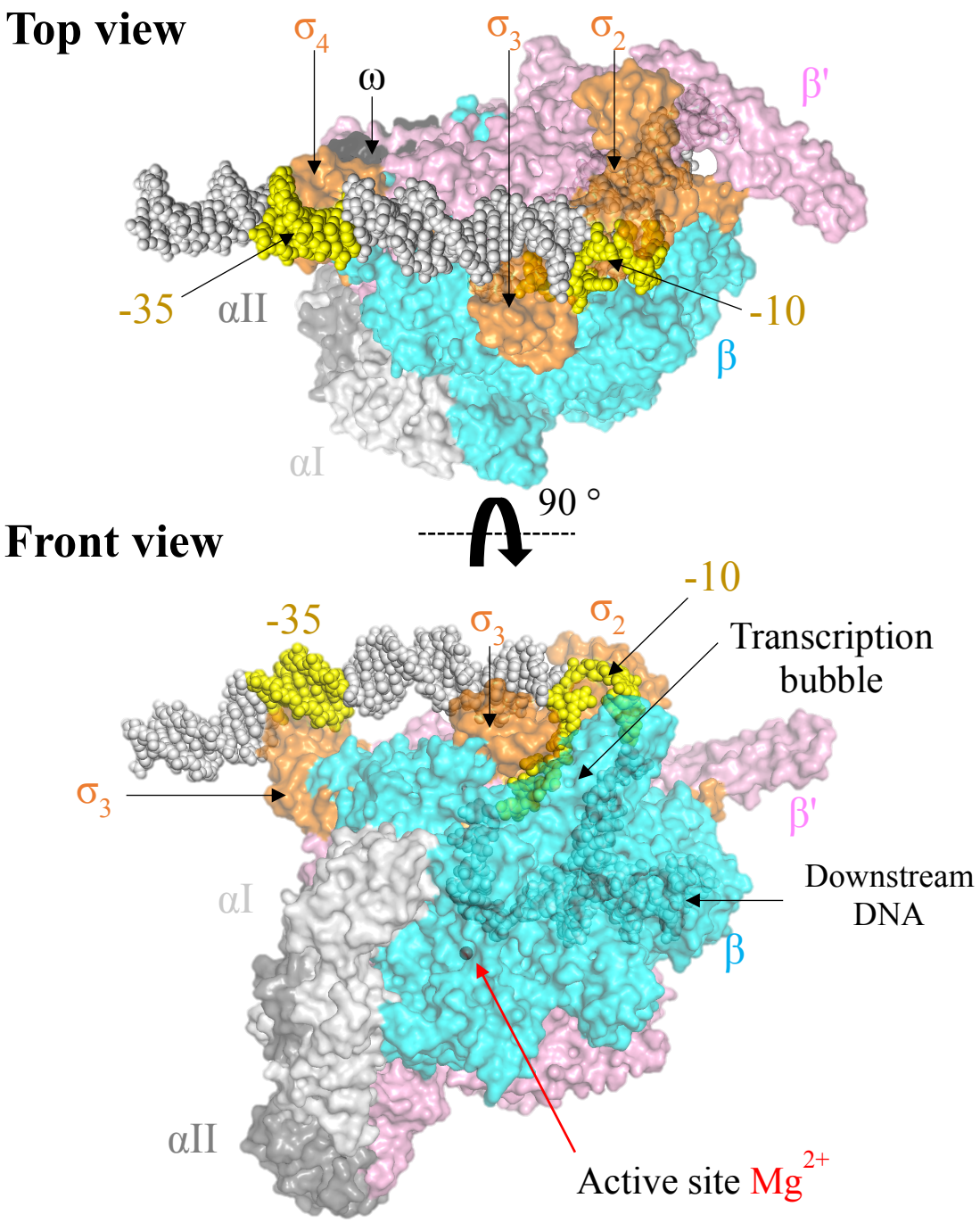


Fig. 1.5. Structure of Bacterial RNAP holoenzyme open promoter complex. Structural models of RNA polymerase holoenzyme open promoter complex (figure made using 6EDT (Boyaci et al., 2019). Protein subunits shown as in Fig. 1.1. Additionally, σ is orange and its domains are denoted. DNA is shown as Corey-Pauling-Koltun (CPK) spheres. The -10 and -35 elements are shown in yellow. Active site Mg^{+2} is indicated in the front view.

At the start of RNA synthesis, Eσ remains strongly bound to the promoter through interactions between σ and the DNA and it cannot move forward. The enzyme engages in a process called abortive transcription where short RNA transcripts are synthesized and released (Carpousis & Gralla, 1980; Goldman, Ebright, & Nickels, 2009; McClure, Cech, & Johnston, 1978). Eventually, the contacts between σ and the DNA are broken in a process called promoter escape and the complex moves forward (Henderson et al., 2017). At this stage the complex enters the elongation phase where the nascent RNA continues to grow and σ is stochastically released from complex (Mooney et al., 2005; Raffaelle et al., 2005). Once dissociated, σ can bind another E to start another cycle of transcription.

During elongation, the RNAP-DNA complex can experience a variety of transcription pauses. Promoter proximal pauses can occur where σ rebinds promoter-like sequences (Ring, Yarnell, & Roberts, 1996). It can also experience transient elemental pauses, where the complex adopts a catalytically inactive conformation and NTP substrate loading is inhibited (Hein et al., 2014; Kang et al., 2018; Saba et al., 2019). Longer-lived pauses can arise from the elemental pause (Kang et al., 2019). Transcription elongation continues until a termination signal is encountered and the RNA transcript is released. Transcription termination can occur by one of two mechanisms: intrinsic termination and Rho-dependent termination (Ray-Soni, Bellecourt, & Landick, 2016). In intrinsic termination, synthesized RNA forms a stem-loop followed by a series of uracil bases which form weak base pairing with adenine bases that lead to the destabilization of the RNAP-DNA complex and the release of RNA. In Rho-dependent termination, the helicase Rho binds to a terminator pause site in the RNA while the complex is engaged in a Rho-sensitive pause site (Farnham & Platt, 1981). Rho uses its ATPase activity to

translocate along the RNA until it reaches the 3' end and unwinds the RNA from the template DNA resulting in RNA release (Bidnenko et al., 2016). After transcription termination and the release of DNA, E is able to associate with another σ factor and initiate another cycle of transcription.

Classification of σ^{70} family and their domains

In the 1980's, it was discovered that, in addition to the primary σ like σ^{70} in *Eco*, other σ factors were present in bacteria and played roles in specific cellular functions such as heat-shock responses and sporulation (Grossman, Erickson, & Gross, 1984; Haldenwang & Losick, 1980; Hunt & Magasanik, 1985). As more of these alternative σ were discovered, their key role in re-programming transcription was established (**Fig. 1.3**). Most alternative σ belong to the σ^{70} family, which is divided into four phylogenic Groups (I-IV) based on protein sequence, domain architecture, and cellular function (Paget, 2015) (**Fig. 1.6**). Members of the σ^{70}-family vary widely in size from ~20 kDa to ~70 kDa, and are composed of two, three, or four structured domains separated by flexible linkers (Campbell et al., 2002) (**Fig. 1.6**). Group I (primary) σ control the majority of transcription for genes involved cellular homeostasis during exponential growth including genes involved in metabolism and ribosome synthesis (Gruber & Gross, 2003). Groups II-IV consist of alternative σ, which conduct transcription for specialized cellular functions (Österberg et al., 2011).

Group I σ factors, like σ^{70} in *Eco*, are essential and contain four helical structured domains ($\sigma_{1.1}$, σ_2, σ_3, σ_4) (Campbell et al., 2002; M. S. Paget, 2015) (**Fig. 1.6**). Although promoter recognition is primarily through σ, the σ^{70}-family members can only bind DNA

once assembled with E. The N-terminal domain $\sigma_{1.1}$, only found in Group I, auto-inhibits promoter recognition in the absence of E (Camarero et al., 2002; Schwartz et al., 2008). Upon assembly with E, $\sigma_{1.1}$ occupies the active site cleft of E where it is thought to prevent non-specific nucleic acids from accessing the E active site (Bae et al., 2013; Mekler et al., 2002). The $\sigma_{1.1}$ is displaced from E active site by promoter DNA during transcription initiation (Mekler et al., 2002).

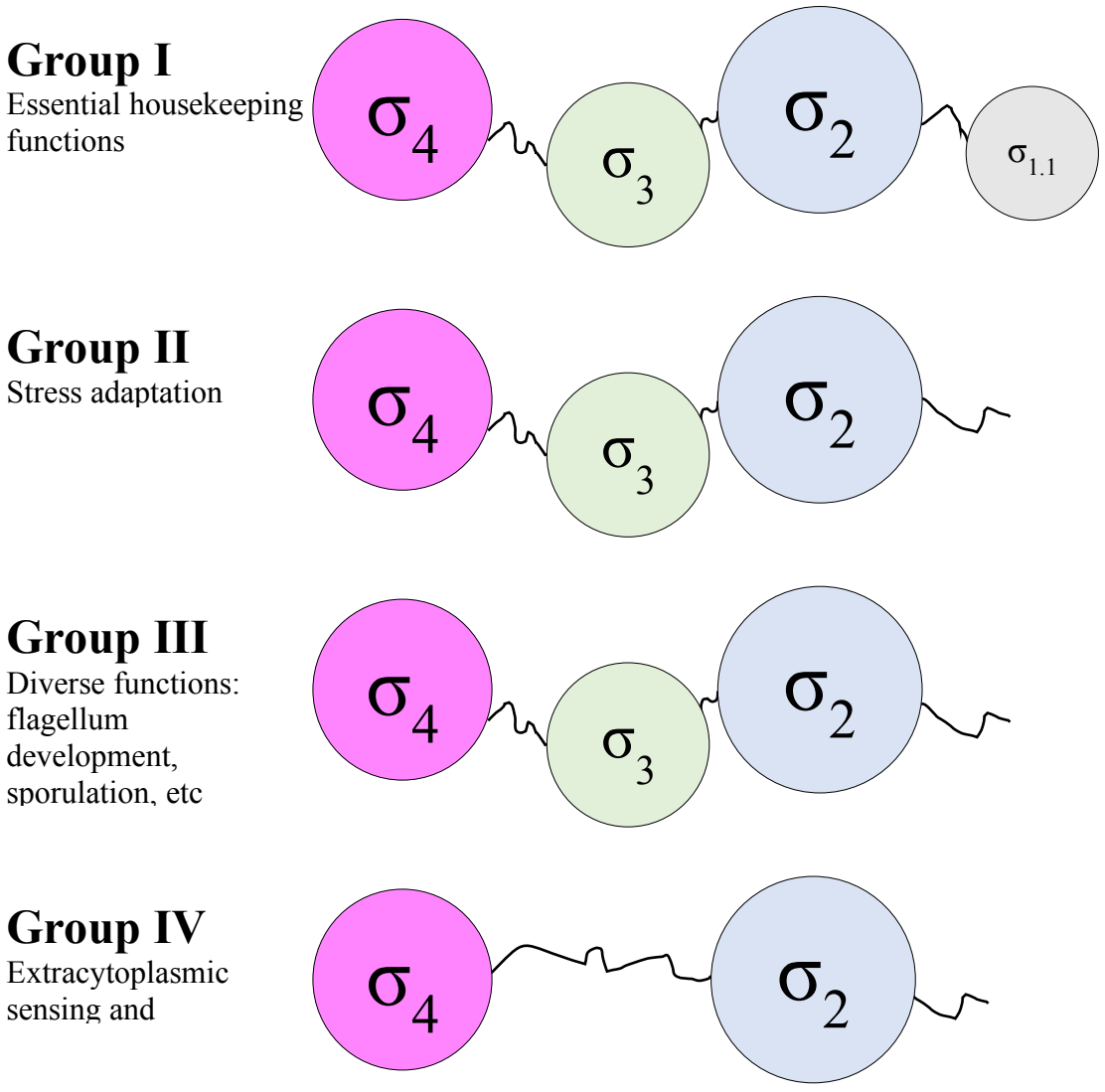


Fig. 1.6. Group classification of σ^{70} family. σ^{70} factors are composed of structured domains (circles) separated by flexible linkers. Each group has several general functions, some of which are listed below the group class.

All members of the σ^{70} family contain σ_2, which is the most conserved σ domain and makes extensive contacts with the β' clamp helices (β'CH) in E (Murakami et al., 2002). It also recognizes the -10 element (also known as the Pribnow box), which has a consensus sequence $T_{-12}A_{-11}T_{-10}A_{-9}A_{-8}T_{-7}$ for σ^{70} (**Fig. 1.4A**) (Daniels, Zuber, & Losick, 1990; Kenney et al., 1989; Siegele et al., 1989; Waldburger et al., 1990; Zuber et al., 1989). It has been shown that two nucleotides (A_{-11} and T_{-7}) in the non-template strand are extruded from double-stranded promoter DNA, flipped out of the base stack, and bound to protein pockets within σ_2 (Feklistov & Darst, 2011). This explains how σ stabilizes the initial melting of the promoter DNA to form the transcription bubble (deHaseth & Helmann, 1995; Jones & Moran, 1992) (**Fig. 1.4B and Fig. 1.5**). The σ_2 also facilitates a bend in the promoter DNA, which direct it towards the main cleft of E. The $\sigma_{1.2}$ is a region within the structured σ_2 that interacts with the promoter region downstream of the -10 element called the discriminator element, which has a consensus sequence $G_{-6}G_{-5}G_{-4}$ for σ^{70} (Morichaud, Chaloin, & Brodolin, 2016; Pribnow, 1975; Zenkin et al., 2007) (**Fig. 1.4A**). This discriminator element is absent in ribosomal RNA promoters and leads unstable RPo (Haugen et al., 2006).

Some σ factors like σ^{70} contain a large non-conserved region (NCR) between $\sigma_{1.2}$ and σ_2 (**Fig. 1.4B**). This is one of the main differences between σ^{70} whose NCR is 252 amino acids long and its closest relative, the Group II σ^S, whose NCR is only seven amino acids long (Banta et al., 2013). For σ^{70}, this NCR has been shown to play a role in aiding promoter escape by counteracting the interaction between σ^{70}_2 and the β'-CHs (Leibman & Hochschild, 2007; Young et al., 2001). A few transcription factors have been shown to interact with the NCR of their host primary σ factor: RpbA in

Mycobacterium tuberculosis (Hubin et al., 2015), GrgA in *Chlamydia trachomatis* (Bao, Nickels, & Fan, 2012), GcrA in *Caulobacter crescentus* (Haakonsen, Yuan, & Laub, 2015; Wu et al., 2018), and Crl in many γ-proteobacteria (Banta et al., 2014).

The σ_3 is present in Group I, II, and III σ factors, but it is altogether missing in Group IV σ like the extracytoplasmic function (ECF) σ factors such as σ^E (Paget & Helmann, 2003). This compact domain is composed of a three-helix bundle (**Fig. 1.4B**). The σ_3 interacts with the extended -10 promoter elements to convey additional specificity (Bae et al., 2015; Barne et al, 1997; Koo et al,, 2009). The $\sigma^{70}{}_3$ has a consensus sequence $T_{-15}G_{-14}$.

Similar to σ_2, σ_4 is highly conserved and present in all members of the σ^{70} factors (Haugen, Ross, & Gourse, 2008; Lonetto et al., 1992). The σ_4 interacts with the β flap of E (Kuznedelov et al., 2002; Murakami, Masuda, & Darst, 2002). $\sigma^{70}{}_4$ recognizes the promoter DNA duplex -35 element with a consensus sequence $T_{-35}T_{-34}G_{-33}A_{-32}C_{-31}A_{-30}$ (Campbell et al., 2002; Gardella, Moyle, & Susskind, 1989; Hawley & McClure, 1983; Kenney, & Moran, 1991; Murakami, Masuda, & Darst, 2002; Siegele et al., 1989). The σ_4 contacts template and non-template sites in the -35 element that lie in the major groove of the DNA (Campbell et al., 2002). The σ_4 is also the target of regulation for many transcription activators (Gross et al., 1998; Ishihama, 1993), such as the bacteriophage activator λcI (Dove et al., 2003; Jain et al., 2004).

Group II σ are the closest relatives to Group I and are typically involved in stress adaptation (Paget & Helmann, 2003). Structurally, they lack $\sigma_{1.1}$. The most well characterized member of this group is σ^S, the master regulator of transcription during stationary phase in many gram-negative bacteria, which has a regulon of around ~500

genes (Battesti, Majdalani, & Gottesman, 2011; Bouché et al., 1998; Polen et al., 2005). While σ^S is not required for growth, mutants of σ^S renders cells more vulnerable to environmental stresses (Notley-McRobb, King, & Ferenci, 2002). The similarities between σ^{70} and σ^S lead to an overlap in their promoter recognition, and in fact σ^S has the same consensus for -35 and -10 elements as σ^{70} (Gaal et al., 2001). However, its tolerance for deviations in the extended -10 element, and its spacer length confers selectivity for promoters different than σ^{70}-dependent promoters (Typas, Becker, & Hengge, 2007; Typas & Hengge, 2006).

Group III σ are very diverse and more distantly related to Group I compared to Group II (Paget & Helmann, 2003). They recognize different promoter sequences from those of Group I and II (Koo et al., 2009). This group contains σ_3, which also interacts with extended -10 element-like sequences (Koo et al., 2009). This group controls a variety of cellular functions including flagellum development, heat-shock responses, and sporulation (Wösten, 1998). In many gram-positive bacteria, a member of this group called σ^B controls transcription in response to general stresses (Hecker, Pané-Farré, & Uwe, 2007). Furthermore, sporulation in organisms like *Bacillus subtilis* requires concerted efforts of several Group III σ (Hilbert & Piggot, 2004).

Group IV σ, also known as extracytoplasmic σ, is the largest and most phylogenetically diverse group within the σ^{70} family (Helmann, 2002; Lonetto et al., 1994). They play important roles in signal transduction in roles like envelope stress responses and iron import (Staroń et al., 2009). Structurally, Group IV lack both $\sigma_{1.1}$ and σ_3 (Campagne et al., 2014; Lin et al., 2019). They also exhibit different promoter preferences for those of other groups (Staroń et al., 2009).

σ^S discovery and characterization

The alternative σ^S is the master regulator of general stress responses and reprograms transcription in several Gram-negative bacteria (Hengge-Aronis, 2002b). The gene that encodes σ^S (*rpoS*) was discovered independently by several groups investigating cellular functions and phenotypes.

In the 1970's and 1980's, an isolated *Eco* mutant was shown to be particularly sensitive to near ultra-violet radiation compared to wild type *Eco* (Tuveson, 1981; Tuveson & March, 1980). The source of this sensitivity was traced to an allele named *nur* (after its near ultraviolet radiation sensitivity phenotype). Another group investigated a gene they called *katF*, whose protein product was important for the expression of catalase hydroxyperoxidase II (HPII, which is encoded by the gene *katE*) (Loewen & Triggs, 1984). Other researchers determined the product of *katF* was also necessary for the expression of exonuclease III (Sak, Eisenstark, & Touati, 1989), which is important for DNA repair. In addition, σ^S was also identified as a regulator of acid phosphatase and named *appR* (Touati, Dassa, & Boquet, 1986). In 1991, a quest to find carbon-starvation inducible genes identified *rpoS* as a regulator of a variety of proteins present in starvation conditions (Lange & Hengge-Aronis, 1991). Lastly, another group investigated the gene *aidB*, which is involved in DNA repair, and discovered its expression is dependent on a factor they named *abrD* (for *aibB* regulator D) (Volkert, et al., 1994). Eventually, it was recognized that each of these studies had independently identified the same gene, *rpoS*, (Lange & Hengge-Aronis, 1991), which controlled a diverse set of cellular functions. In 1988, the gene *katF* was cloned (Mulvey & Loewen, 1989). This sequence revealed that

this gene shared a high degree of sequence homology with *rpoD*, the gene that encodes σ^{70}, the housekeeping σ factor subunit of RNAP. It was then postulated that the product of *katF* was an alternative σ factor subunit of RNAP, which controlled the transcription of many genes. Due to its importance in stationary phase and stress response transcription programs, the name *rpoS* was suggested for the gene, and RpoS or σ^S for the protein product (Lange & Hengge-Aronis, 1991). After corrections of the initial sequence, σ^S was successfully purified, yielding a protein of approximately 38 kDa (Ivanova et al., 1992; Tanaka et al., 1993). This led to the alternative nomenclature of σ^{38} in line with the common naming convention of the σ^{70} housekeeping σ factor. Tanaka and colleagues also conducted the first *in vitro* reconstruction of σ^S holoenzyme ($E\sigma^S$), which demonstrated how some promoters are equally transcribed by $E\sigma^{70}$ or $E\sigma^S$, whereas other promoters showed preference for one or the other (Tanaka et al., 1993). This suggested an overlap in the recognition sequence preference of σ^S and σ^{70}, which was subsequently confirmed (Gaal et al., 2001; MacIąg et al., 2011; Typas & Hengge, 2006). Further work investigated how σ^S regulates genes that are directly implicated in responses to environmental stresses such as changes in temperature, pH, UV light, osmolarity, nutrient availability, and host invasion (Hengge-Aronis, 2002a).

Multifactorial regulation of σ^S expression

The rapid and efficient expression of genes under σ^S is critical for the survival of bacteria in response to stresses. However, once conditions become favorable for growth the global transcription program by σ^S must be shut down to prevent fitness disadvantages from other organisms that can more effectively use environmental

resources. The reason for this is because σ^S has adverse effects in the expression of genes required for growth housekeeping functions. Additionally, the σ^S transcription program induces wide-range protection against the direct source of stress but also additional stresses. For example, exposure to osmotic stress will induce the expression of proteins that confer additional resistance to increases in temperature and presence of reactive oxygen species (Hengge-Aronis et al., 1993). Thus, the expression of σ^S confers high metabolic costs as it activates genes that are not required for the direct stress that bacteria experience. This might be the evolutionary result of bacteria developing preemptive mechanisms to persist in challenging environments where multiple stresses are encountered such as nutrient scarcity following drops in temperature. For these reasons, the expression of σ^S is highly regulated at transcriptional, translational, and post-translational levels (Landini, et al., 2014) (**Fig. 1.7**).

At the transcriptional level, factors like ArcA directly bind DNA sites near the *rpoS* promoter to repress transcription of *rpoS* (Mika & Hengge, 2005). Phosphorylated ArcA performs this repression during high energy cellular states. The global regulator cAMP-activated global transcriptional regulator (CRP), induced in low glucose, has a binding site in the promoter of *rpoS*, and in some bacteria it appears to activate transcription of *rpoS* while in other bacteria it represses it (Mika & Hengge, 2005).

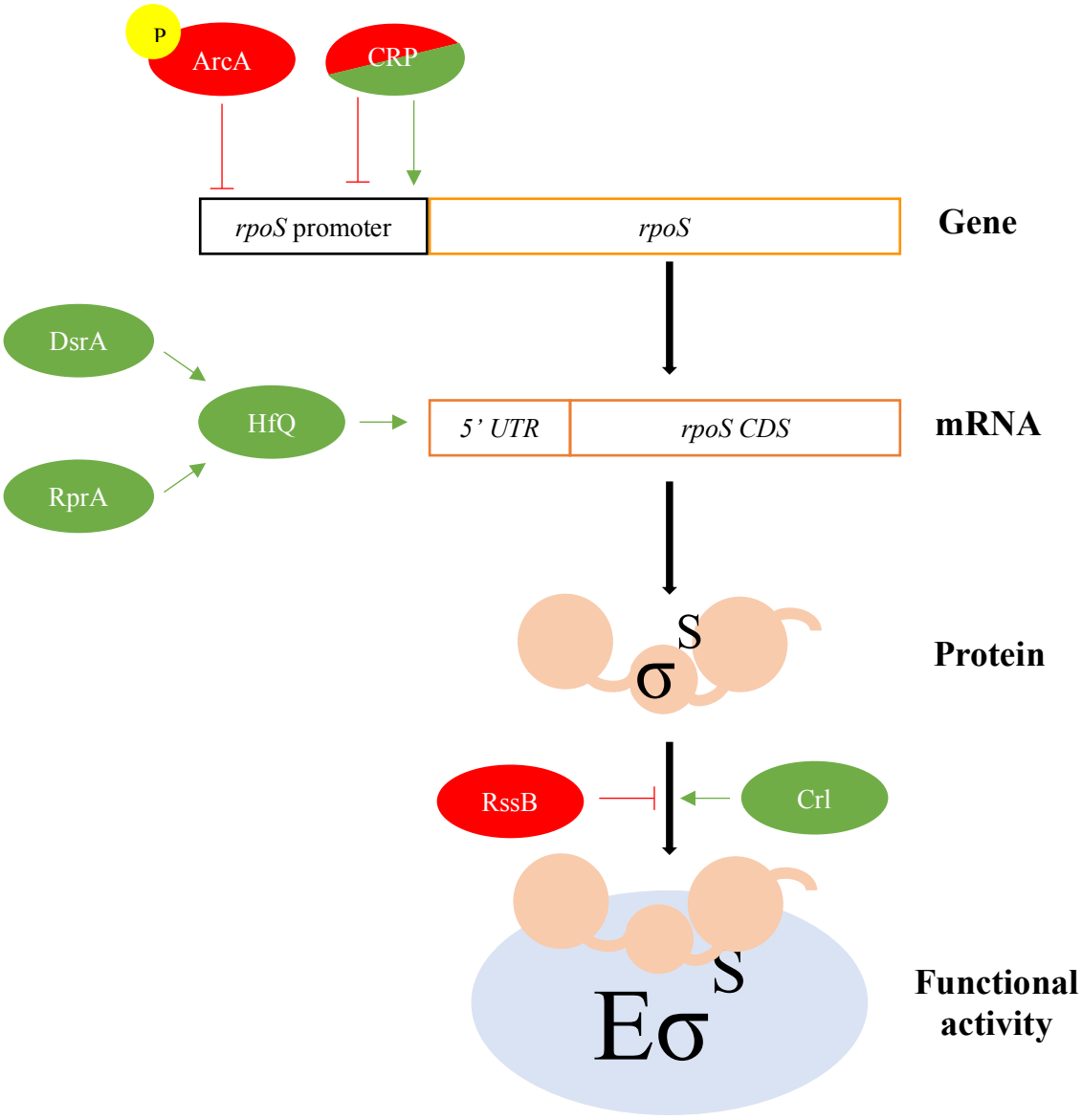


Fig. 1.7. Examples of *rpoS*/σ^S regulation. σ^S (encoded by the gene *rpoS*) is highly regulated at different expression stages. Positive regulation is shown by green arrows, negative regulation by red blunt arrows.

At the translational level, the expression of σ^S is inefficient. The translation initiation site for σ^S is located around 500 bp downstream from its transcription start site. This long 5'-untranslated region is predicted to form structures that prevent the ribosome from binding its initiation site. The short regulatory RNAs, DsrA and RprA, stabilize the 5'-untranslated region of the *rpoS* mRNA through the action of nucleic acid chaperone Hfq (Lease & Belfort, 2000; Majdalani et al., 1998; Wassarman et al., 2001). This results in re-structuring the *rpoS* mRNA and allows the ribosome to initiate translation.

Post-translationally, σ^S is targeted by RssB, which delivers it to the protease ClpXP to be degraded (Klauck, Lingnau, & Hengge-Aronis, 2001). The binding of RssB to σ^S sequesters and prevents it from binding E; in this sense RssB acts as an anti-sigma factor. The inhibition of σ^S proteolysis has been shown to be critical for the accumulation of σ^S, which is required to enter into stationary phase (Lange & Hengge-Aronis, 1994). Several small proteins like IraD become expressed in the presence of stresses like DNA damage and they promote σ^S activity by binding to RssB and preventing it from targeting σ^S (Bougdour, et al., 2008).

Another strategy by which σ factors are post-transcriptionally regulated is by inhibiting their binding to E with anti-σ factors (Campbell, Westblade, & Darst, 2008) (**Fig. 1.8**). The sequestration of certain σ^{70} family members by their cognate anti-σ favor the Eσ for σ like σ^S, which does not have a cognate anti-σ. Structures of anti-σ factors bound to their cognate σ factors revealed that they directly occlude the RNAP interacting surface and/or stabilize inter-domain interactions within the σ factor (Campagne et al., 2012; Campbell & Darst, 2000; Campbell et al., 2003; Campbell et al., 2008; Herrou et al., 2012; Maillard et al., 2014; Shukla et. al, 2014). These anti-σ factors typically favor

closed/occluded conformations of σ factors, so they cannot bind E. Upon relief of inhibition via a variety of mechanisms that dissociate the anti-σ (such as oxidation, proteolysis, phosphorylation), the σ can adopt extended conformations that expose its RNAP binding surfaces, which is needed for Eσ assembly (Campbell et al., 2008). Lastly, the activity of σ^S is also activated by Crl, a small protein that is the subject of this thesis and will be discussed further in the upcoming sections.

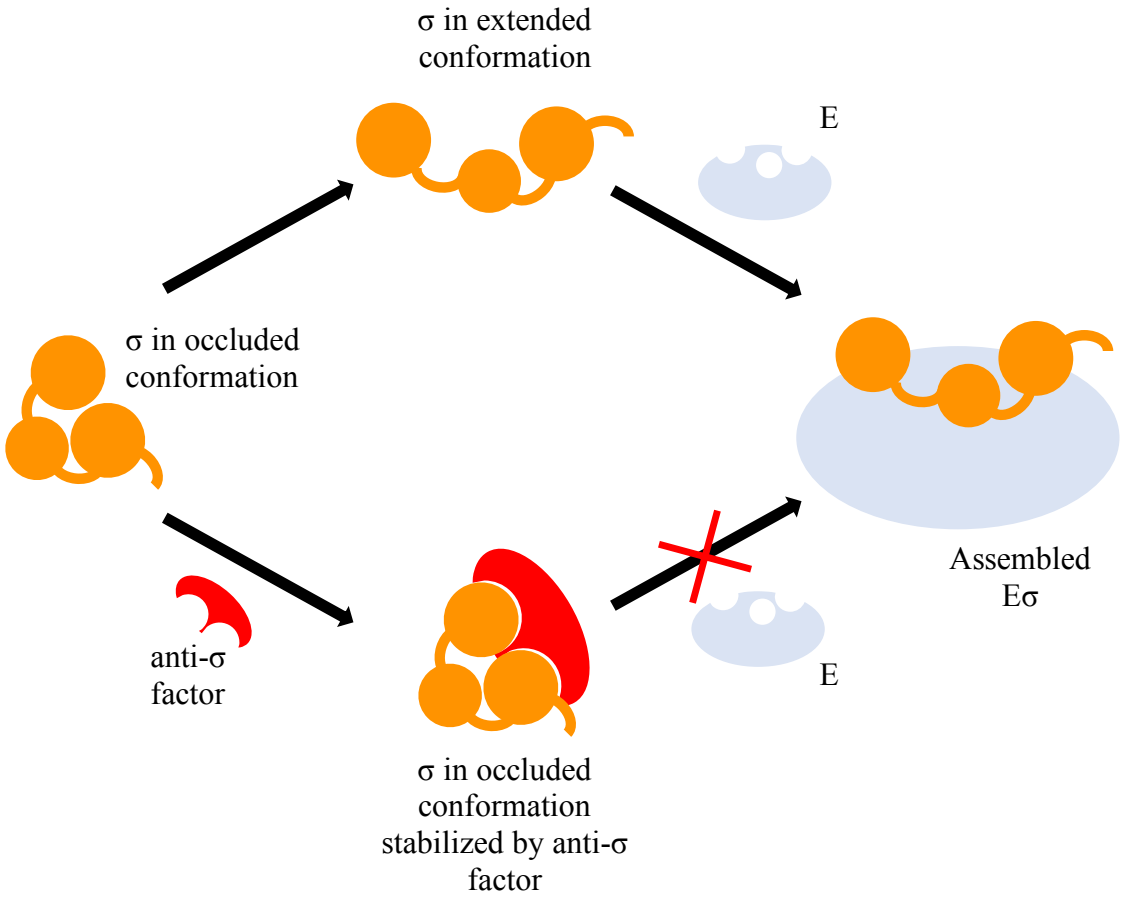


Fig. 1.8. Anti-σ factors as transcription repressors. Cartoon schematically illustrating how anti-σ factors prevent assembly of Eσ by stabilizing occluded conformations of σ

Discovery of Crl

Flagella and pili are two types of functional extracellular structures that allow gram negative bacteria to interact with their environment. In the 1980's researchers reported a third type of fiber-like structure at the surface of gram-negative bacteria called curli (Olsén, Jonsson, & Normark, 1989). These fibers are important for biofilm formation, host cell adhesion, and colonization of inert surfaces (Barnhart & Chapman, 2006; Cao et al., 2014). Olsén and colleagues also discovered a gene they initially believed was the structural protein of curli fibers and called it *crl*, which encodes the small (16 kDa, ~130 amino acid) protein Crl. Subsquently, it was discovered that Crl is not the structural subunit of curli, but rather it is a factor that activates the trancription of many genes under the control of σ^S, including the genes responsible for the synthesis and secretion of curli (Arnqvist et al., 1992).

Regulation of Crl expression

Several studies have reported that Crl is constitutively expressed under a housekeeping transcription program and reaches maximum levels when the bacteria enter late stationary phase, followed by a small decrease (Bougdour, Lelong, & Geiselmann, 2004; Robbe-Saule et al., 2006; Robbe-Saule et al., 2007). In contrast, the levels of σ^S are negligible until bacteria enter stationary phase and then accumulate as cells continue through stationary phase (Robbe-Saule et al., 2006; Tanaka et al., 1993). The transcription of *crl* is under the control of a σ^{70} promoter, which explains why it is constitutively expressed (Arnqvist et al., 1992). Interestingly, the *crl* gene contains an additional σ^{54} promoter sequence, but if *crl* is transcribed by $E\sigma^{54}$, then its mRNA lacks a ribosomal binding site and cannot be translated (Zafar et al., 2014). This intriguing σ-dependent regulation

ensures that Crl is not synthetized as cells execute a σ^{54} transcription program, which activates genes during nitrogen-limiting conditions and other adaptive responses (Wigneshweraraj et al., 2008).

Reports have indicated Crl is naturally expressed at higher levels when bacteria cultures are grown at 30 °C compared to 37 °C (Bougdour et al., 2004). In *Eco*, indole has been proposed to induce the expression of Crl (Lelong et al., 2007). It has also been proposed that transcription of *crl* is repressed by the global transcription regulator Fur (ferric uptake regulator), which binds the *crl* promoter (Lelong et. al, 2007). Finally, it has been proposed that Crl contains a C-terminal signal that targets it to degradation by the ClpXP machinery (Flynn et al., 2003).

Structural conservation of Crl

σ^S homologs are widely present in γ-proteobacteria including *Eco* and *Salmonella enterica* serovar Typhimurium (*Sty*), where the σ^S homologs are 99% identical in protein sequence (Monteil et al., 2010). Even in more distantly related bacterial families such as *Salmonella* and *Vibrio*, the σ^S homologs share 72-82% sequence identity (Monteil et al., 2010). In contrast, Crl is less widespread and less conserved (Banta et al., 2014): *Sty* and *Eco* Crl homologs are 83% identical while *Sty* and *Vibrio* Crl are only 38-50% identical (Monteil et al., 2010). This lack of conservation opened inquiries into whether Crl homologs from different bacteria perform similar functional roles and if their three-dimensional structures were similar. To date, the structures of Crl from *Proteus mirabilis* (*Pm*) and *Sty* have been determined by X-ray crystallography (Banta et al., 2014; Cavaliere et al., 2014) and NMR (Cavaliere et al., 2015) (**Fig. 1.9**). These two Crl

homologs are 48% identical in sequence, but their protein fold is conserved. The regions predicted to interact with σ^S are more significantly conserved than the overall protein. To support similar function among Crl homologs, it has been shown by isothermal titration calorimetry (ITC) that the binding affinity between *Pm* Crl and *Sty* σ^S is very similar to the affinity between *Sty* Crl and *Sty* σ^S (Cavaliere et al., 2014). Most strikingly, *Pm* Crl (48% sequence identity with *Eco* Crl) and *Vh* Crl (38% sequence identity with *Eco* Crl) homologs fully complemented an *Eco* Crl deletion *in vivo* (Banta et al., 2014). In sum, the conservation in structure, interactions and the fact that Crl from different species can rescue Δ*crl* phenotypes, supports the conclusion that the biological function of Crl homologs from different species is conserved.

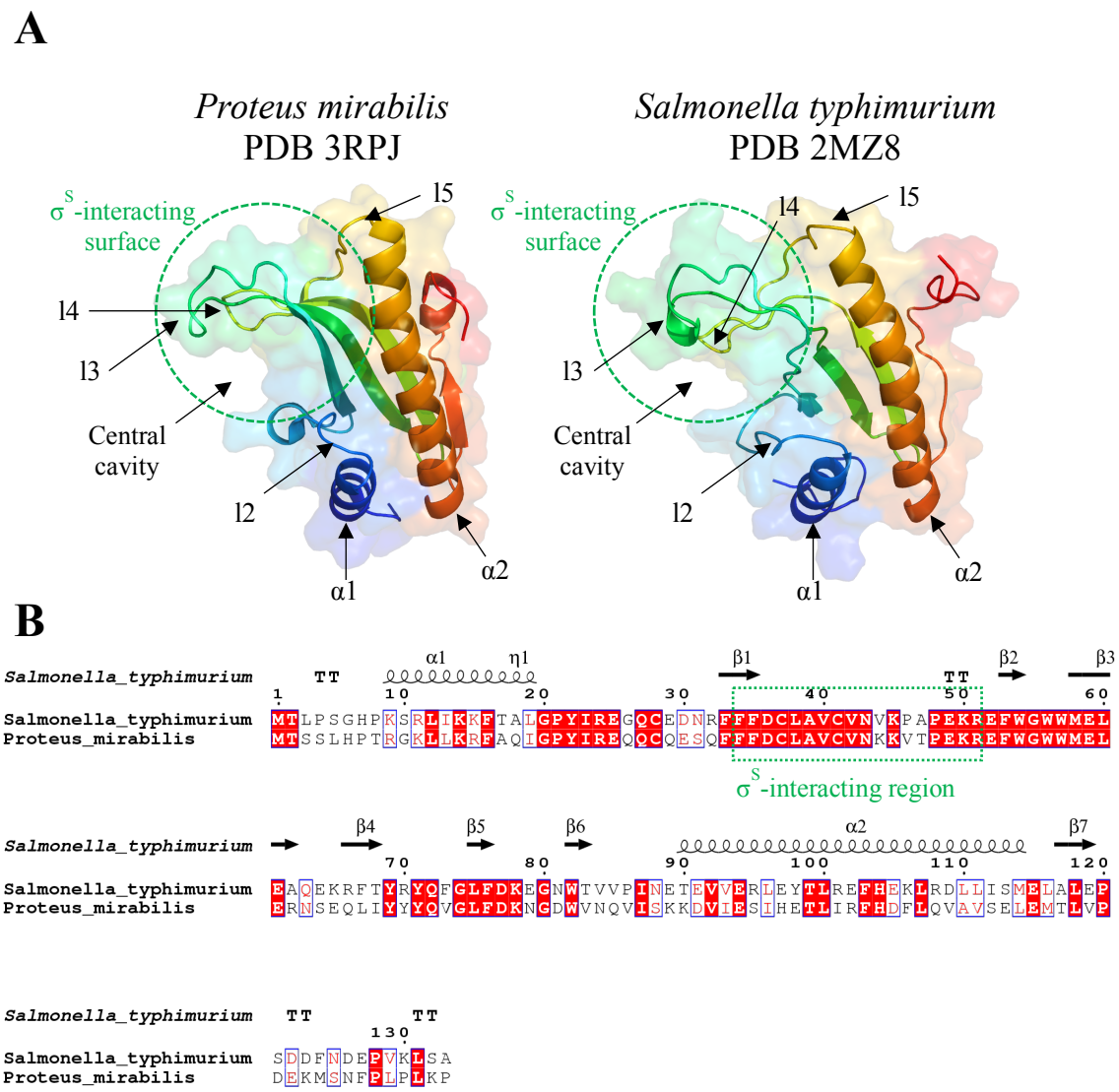


Fig. 1.9. Crl structural conservation. Previously, two x-ray crystal structures of *Proteus mirabilis* have been determined (PDBs: 3RPJ and 4Q11). Additionally, the NMR structure of *Salmonella enterica* has been determined (PDB 2MZ8). **A.** The structures of these two species are shown as cartoons with transparent surfaces and N->C colored, where the N-terminus is show in blue and the C-terminus is shown in red. The region that is predicted to interact with σ^S is outlined in green. **B.** Alignment of the amino acid sequences of these *Salmonella enterica* and *Proteus mirabilis.*

The structure of *Pm* Crl (3RPJ) revealed a globular fold composed of 5 β-strands and 2 α helices (Banta et al., 2014). A β-sheet composed of 4 antiparallel strands forms a conserved central cavity, which is flanked by the intervening loops (l2 in one face, and l3/l4 in the other face). The N-terminus of the protein forms a helix with a number of conserved positively charged amino acids (α1). This helix is followed by a loop that leads to the first β strand that forms the central cavity of the protein. This strand leads to a loop in one surface of the protein that provides important contacts for its interactions with σ^S (Banta et al., 2013) (l3). Loops l3 and l4 along with the β-sheet in central cavity are very important for binding σ^S. A loop follows the last residues of the central cavity (l5), which lead to a long α-helix (α2), which form a surface opposite to the σ^S-surface. The C-terminus of the protein contains a small structured helix, and the last amino acids at the C-terminus appear to be flexible and disordered. Of note, Crl crystallized as a dimer, but NMR and other biophysical studies have revealed that Crl is only present as a monomer in solution (Banta et al., 2014; Cavaliere, Norel, & Sizun, 2015).

Interaction between Crl and σ^S

Crl plays a unique role in the regulation of σ^S-dependent genes by directly enhancing the transcription activity of $E\sigma^S$ (Gaal et al., 2006; Robbe-Saule al., 2007). It has been shown that Crl binds directly and specifically to $\sigma^S{}_2$ and strengthens the binding affinity of σ^S to E (Banta et al., 2013; Monteil et al., 2010). Binding of Crl could also prevent RssB sequestration of σ^S and subsequent degradation. This activation by Crl present a striking contrast with conventional activators, which tend to be DNA binding proteins that act by recruiting Eσ to initiate transcription at specific sites (Dove et al.,

2003). At the onset of this thesis' research there was no available three-dimensional structures of a complex composed of Crl and σ^S. Nonetheless, direct binding of σ^S and Crl has been detected using gel filtration, ITC, surface plasmon resonance (SPR), and bacterial two-hybrid (BTH) assays (Banta et al., 2013, 2014; Bougdour et al., 2004; Cavaliere et al., 2014; Cavaliere et al., 2015; England et al., 2008; Monteil et al., 2010). Genetic, biochemical, and phylogenic conservation analysis along with three-dimensional structures of Crl from *Pm* and *Sty* (PDB: 4Q11 3RPJ, and 2MZ8) provided insight into determinants of Crl for σ^S binding (Banta et al., 2014; Cavaliere et al., 2014, 2015).

Pull down experiments showed that σ^S from cellular extracts co-eluted with affinity-tagged Crl (Bougdour et al., 2004). Biophysical measurements revealed that Crl binds σ^S with dissociation constant (K_d) of 1~3 μM (Cavaliere et al., 2014; England et al., 2008). Follow up studies showed that $\sigma^S{}_2$ is absolutely required for binding of Crl (Banta et al., 2013). Site-directed mutagenesis identified specific sites on $\sigma^S{}_2$ that are essential for its interaction with Crl: D87, D135, P136, and E137 (Banta et al., 2013) (**Fig. 1.10**). Mutating these sites perturbed the σ^S interaction with Crl, but not with the β′-CHs where $\sigma^S{}_2$ binds E. Homology modeling of σ^S based on σ^{70} (PDB 1SIG), showed that these 4 sites lie in the same solvent-accessible surface of σ^S (Banta et al., 2013). This model for σ^S was validated once the structure of a Eσ^S was solved (PDB 5IPL) (Liu, Zuo, & Steitz, 2016). The D87 site in σ^S was of particular interest because it corresponds to the location of the non-conserved region of σ^{70}. This provided insight into the σ selectivity of Crl. Additionally, D87 is part of alpha helix (helix α2), which is conserved in σ^S across multiple species and is thought to be important for binding Crl (Cavaliere et al., 2015). This work also showed that some bacterial species like *Pseudomonas aeruginosa* (*Pae*),

which contain σ^S but lack Crl, have substitutions in this helix and do not bind *Eco* Crl. If mutations in σ^S from *Pae* are introduced to match the sequences found in helix α2 from bacteria that contain Crl, then the σ^S from *Pae* was able to bind *Sty* Crl (Cavaliere et al., 2015).

BTH assays led to the identification of specific Crl residues important for its interaction with σ^S (Y22, R51, F53, W56, and W82) (Banta et al., 2014) (**Fig. 1.11**). Mutations at Crl site D36 have been shown to affect the morphology of *Sty* colonies, resembling those of *Sty* mutants where Crl has been deleted (Cavaliere et al., 2014; Monteil et al., 2010). This site lies in the central cavity of Crl and is one of the most conserved residues among Crl homologs. In addition, an R51A substitution in *Sty* Crl was shown to have no detectable binding to σ^S using ITC, despite evidence that it shows a nearly identical structural profile compared to WT Crl in differential scanning calorimetry (DSC) and circular dichroism (CD) (Cavaliere et al., 2014).

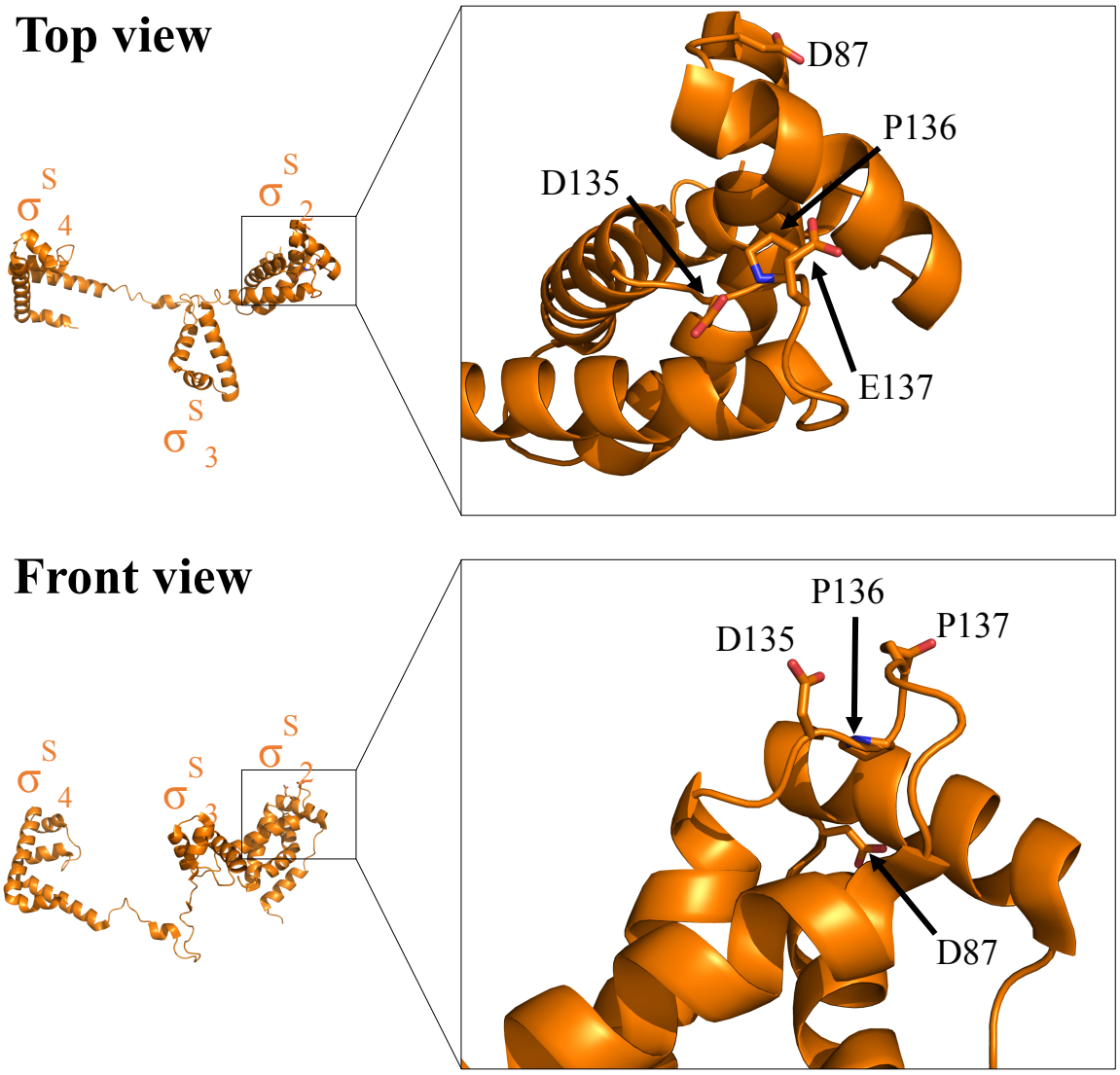


Fig. 1.10. Sites in σ^S known to interact with Crl. Made from PDB 5IPL, Liu et al., 2016.

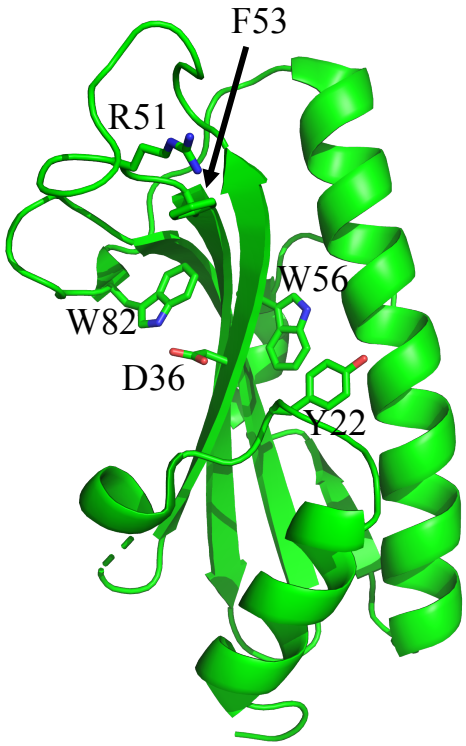


Fig. 1.11. Sites in Crl known to interact σ^S. Made from PDB 3RPJ, Banta et al., 2014.

Regulation of σ^S-dependent genes by Crl

Genome-wide profiling has shown that σ^S directs the transcription of ~500 genes, more than 10% of all genes in *E. coli* (Weber et al., 2005). A combination of methods have determined that Crl regulates the expression of about 90 σ^S-dependent genes (Santos-Zavaleta et al., 2019). These methods include *in vivo* proteomic studies with a Crl knockout (Lelong, et al., 2007), GFP-transcriptional fusions (Dudin, Lacour, & Geiselmann, 2013), *in vivo* transcription genome-wide microarray analysis with a Crl knockout (Typas et al., 2007), inferences from *in vivo* phenotypes with mutant Crl (Arnqvist et al., 1992; Monteil et al., 2010; Robbe-Saule et al., 2006, 2007), and direct in vitro transcription assays that show activation by Crl (Gaal et al., 2006) and more. Interestingly, most genes in the Crl regulon are positively regulated by Crl, but a small group are negatively regulated by Crl. The complex regulation of σ^S (**Fig. 1.7**) makes it difficult to determine which genes are directly being activated by Crl, which ones are the are the result of indirect effects. Chapter 3 presents a promising method to ascertain which genes are directly activated by Crl.

Possible mechanisms of transcription activation by Crl

Before the work presented in this thesis, it remained unclear how the binding of Crl to σ^S enhanced the transcription of σ^S-dependent genes. *Eco* mutants, where *crl* was knocked out, showed higher levels of σ^S compared to WT, which suggested that Crl does not enhance transcription by enhancing the cellular levels of σ^S to aid its competition against other σ factors in binding E (Monteil et al., 2010; Pratt & Silhavy, 1998). Instead it was proposed and shown that Crl acts by enhancing formation of $E\sigma^S$ (Gaal et al., 2006).

First, genetic experiments with *crl* knockouts revealed that overexpreesion of σ^S complements the deletion of *crl* to restore the morphology of WT *Sty* (Robbe-Saule et al., 2006). This overexpression of σ^S would increase the cellular concentration of σ^S and ultimately enhance the levels of assembled $E\sigma^S$ thus leading to higher transcription of σ^S-dependent genes. Additionally, *in vitro* transcription experiments showed that Crl exerts its largest activation effects at low concentrations of σ^S (Gaal et al., 2006). Therefore, it was concluded that Crl promotes $E\sigma^S$ assembly. This idea was supported by SPR binding measurements that showed the affinity of σ^S to E was strengthened 7-fold in the presence of Crl (England et al., 2008). Additionally, Crl was shown to directly enhance the association of labeled E and σ^S using luminescence resonance energy transfer (LRET) (Banta et al., 2014). The questions that then remained were: What is the molecular mechanism by which Crl enhances holoenzyme formation? Does it only interact with σ^S? Does it make specific interactions with core RNAP? This thesis described my approach and experiments to answer those questions with the aim to further advance our understanding of Crl. **Fig. 1.12** shows potential ways by which Crl could activate transcription.

Fig. 1.12. (Next page) Potential mechanisms of activation by Crl. In solution, σ^S could be making intermolecular σ^S-σ^S interactions, which disfavor their association with E. Crl could help break those interaction to promoter $E\sigma^S$ assembly. The black dashes represent these σ^S-σ^S intermolecular interactions, and the green dashes represent σ^S-Crl interactions. Additionally, Crl could help favor the equilibrium of σ^S conformation towards extended conformations where the E binding surfaces are exposed to readily form $E\sigma^S$. Beyond interacting with σ^S, it is not clear if Crl remains bound to $E\sigma^S$ and if this could play a role in its mechanism.

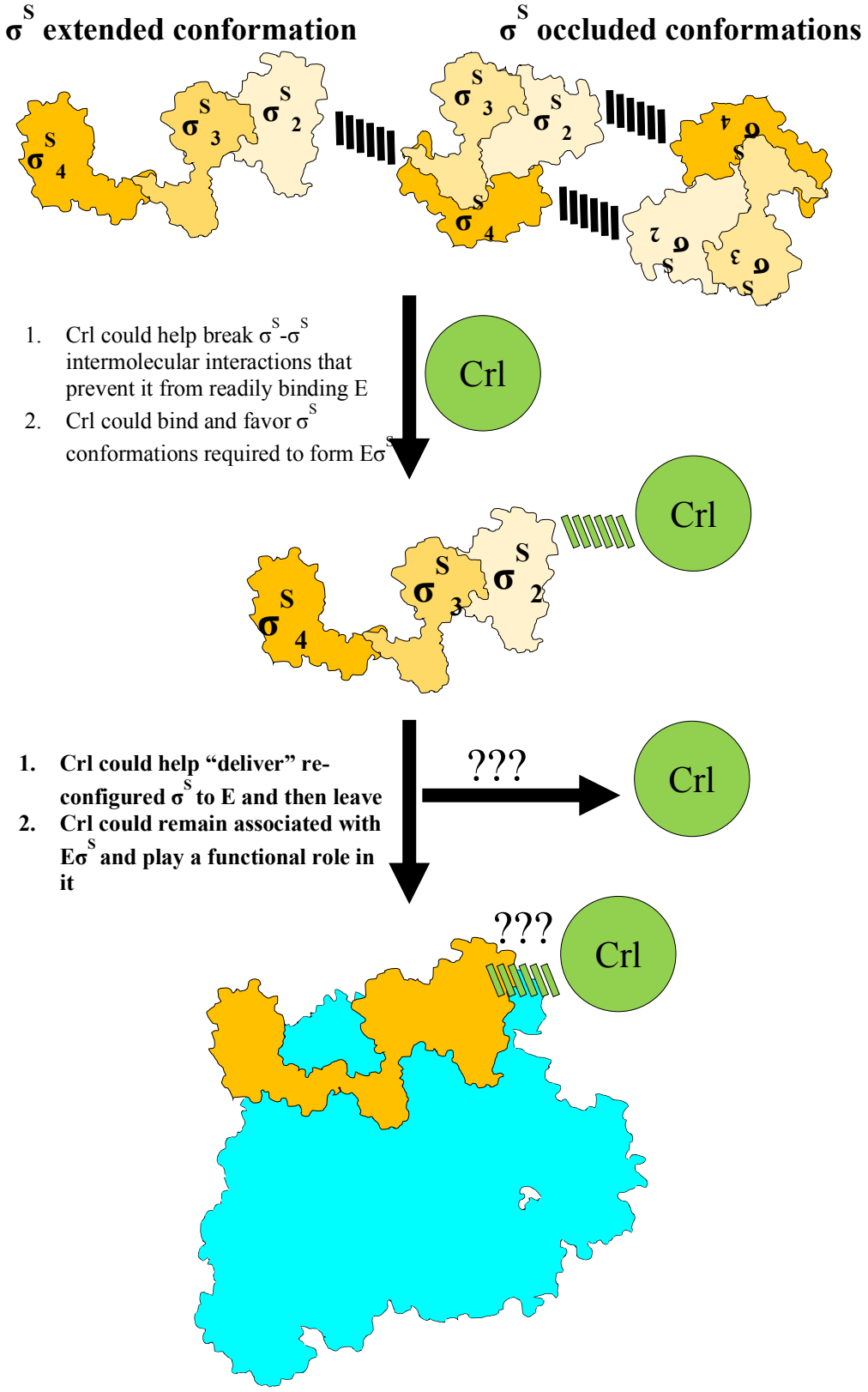


Fig. 1.12. Potential mechanisms of activation by Crl

Chapter 2: **Whole genome *in vitro* transcription investigation of the σ^S and Crl regulons in *Salmonella***

Introduction

The σ^S-regulon in *Salmonella enterica* serovar Typhimurium (*Sty*) is not as well understood as its counterpart in *E. coli* (Chen et al., 1995; Fang et al., 1992; Lelong, Aguiluz, et al., 2007; MacIąg et al., 2011; MacLellan, Eiamphungporn, & Helmann, 2009; Robbe-Saule et al., 2006). There is a significant overlap in the promoters recognized by σ^{70} and σ^S, which complicates the understanding of the sequence determinants of σ^S-specificity. (Peano et al., 2015; Typas et al., 2007; Zafar et al., 2014). Additionally, it is possible that Crl has preference for certain σ^S-dependent promoters but this effect has not been explored at a global genome level.

Global identification of genes directly regulated by transcription factors *in vivo* is difficult due to the complexity of cellular transcriptional networks (Cao et al., 2002). For instance, Crl activates the expression of RssB, which can target σ^S for degradation, which impacts all σ^S-dependent genes. (Santos-Zavaleta et al., 2019; Wurm et al., 2017). In proteomics studies, the direct impact of Crl in gene expression could not be properly assessed as it indirectly leads to downregulation in addition to activation. Reconstituted *in vitro* transcription systems provide the advantage of a reductionist approach by controlling all the components to directly measure the effect of desired transcription factors. Additionally, *in vitro* systems permit the capture of RNA transcripts that might

not be observed from total cellular RNA extractions due to processing and degradation. Furthermore, minor environmental differences could lead to significant global differences obtained from total cellular RNA extractions of biological samples. In order to overcome these problems, techniques like run-off transcription-microarray analysis (ROMA) have been developed (Cao et al., 2002; MacIąg et al., 2011; MacLellan et al., 2009; Zheng et al., 2004). In this method, purified enzymes are used to conduct in vitro transcription with sheared genomic DNA producing RNA transcripts that are converted to cDNA and hybridized to gene microarrays. This technique has two shortcomings: 1) It requires the construction of a microarray based on *a priori* knowledge about the transcriptome of the organism; 2) This technique involves shearing the genomic DNA which can damage the integrity of gene promoters and regulatory sequences.

To directly probe the σ^S-regulon in *Sty* and gain insight into the role of Crl at the whole-genome level I explored the use of an *in vitro* transcription method that uses purified bacterial genomic DNA as the transcription template and RNA sequencing (RNA-Seq) as the read out. Nonetheless, this method also suffers disadvantages: for example nucleotide pools in bacteria are in constant flux, which play direct roles in transcription and affects some promoters more strongly than others (Buckstein, He, & Rubin, 2008). This *in vitro* method cannot answer what happens to the transcriptome as nucleotide pools change with the complexity that it takes place in vivo. Another weakness is that the method will be completely dependent on the activity and purity of the enzymes used; this means that different batches of proteins can lead to different results.

This method capitalizes on the higher integrity of non-sheared DNA to provide more relevant transcription data. It also takes advantage of the high-throughput nature of next generation sequencing that allows multiple conditions to be tested at once. Conditions of interest include DNA modifications like methylation, the presence of different transcription factors, and small modulatory molecules like antibiotics that target RNA polymerase. Most importantly, it gives a readout of which genes are directly regulated by the transcription factor, here Crl, of interest.

Results

The presence of RNase inhibitors is required to obtain RNA transcripts

Our first attempts to conduct whole genome in vitro transcription reactions did not show significant production of transcripts. Although RNase-free reagents were used throughout, we postulated that the lack of products might have been due to small amounts of RNase contaminants that co-purified with our proteins and rapidly degraded the synthesized transcripts. The presence of RNase contaminants was confirmed using an RNase detection kit. The addition of RNase inhibitors led to the detection of significant transcription products in this reaction (**Fig. 2.1**). The reaction with RNase inhibitors appears as a smear in agarose, which is the expected result as transcription reactions will produce a heterogenous mixture of RNA transcript of different sizes.

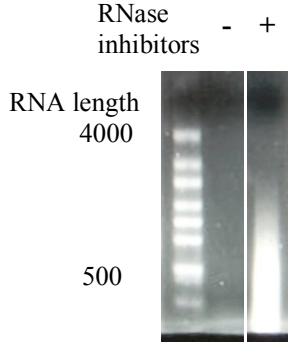


Fig. 2.1. Assessment of whole genome *in vitro* transcription reactions by agarose gel electrophoresis. In the absence of RNase inhibitors (-), no transcription products are seen. A smear is observed in the presence of RNase inhibitors (+), which is expected as transcripts of many lengths would have been synthesized. The products on the reactions on the left were run on a 1% agarose gel in 1x TAE buffer run at 100 V for 1 hour.

RNA transcripts vary depending on the growth stage of the harvested DNA

Studying the effect of DNA modifications on the *in vitro* transcription profiles by holoenzymes with different σ factors is a driving motivation behind the development of this approach (Casadesús & Low, 2006; Kahramanoglou et al., 2012; Low, Weyand, & Mahan, 2001). In these preliminary experiments, we wanted to compare whole genome transcription profiles using DNA harvested during exponential growth compared to DNA purified during late stationary phase where the role of DNA cytosine methylation has been demonstrated (Kahramanoglou et al, 2012). For both exponential and stationary phase DNA, 1 μg of genomic template DNA was added in the reactions. The reactions using genomic DNA purified from exponential growth yielded very similar amounts of RNA transcripts ~600 ng (**Fig. 2.2**). By contrast, reactions using DNA purified from late stationary phase showed significant variability, which would make RNA sequencing results difficult to interpret (**Fig. 2.2**) and were not analyzed further. It is unclear what led to such drastic differences in transcription yields. This variability has important implications for future investigations such as selecting a source of DNA (potentially commercially), to avoid adding variables that could lead to irreproducible results in downstream analysis.

Reaction #	σ	CrI	DNA	RNases inhibitors	RNA yield (ng)
1	S	-	Early	+	670
2	S	+	Early	+	642
3	70	-	Early	+	605
4	S	-	Late	+	2425
5	S	+	Late	+	1315
6	70	-	Late	+	748

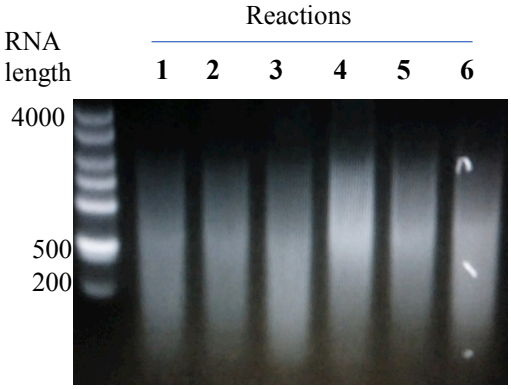


Fig. 3.2. Variability of RNA yield from reactions with DNA harvested at different states. ***Left***, table showing the 6 reaction conditions that were tested and the yield of RNA they produced after purification. ***Right***, the products on the reactions on the left were run on a 1% agarose gel in 1x TAE buffer run at 100 V for 45 min.

σ^S regulon is validated in whole genome *in vitro* transcription

Given the irregularity of results using the DNA purified from stationary phase cells, I only looked at the RNA-Seq from the conditions conducted using DNA extracted from exponential phase. To validate the results of this approach, we mapped the reads and looked at known σ^S-dependent promoters. The *dps* gene has a σ^S-dependent promoter and it is in fact the promoter I used in the structural and biochemical to investigate the mechanism of Crl (Grainger et al. 2008; Lacour & Landini, 2004). This promoter can be recognized by both σ^S and σ^{70}, but shows a marked preference for σ^S (Grainger et al., 2008). The *dps* promoter has a -10 element of sequence $T_{-12}A_{-11}T_{-10}A_{-9}C_{-8}T_{-7}$, which is a strong -10 promoter element for σ^S-dependent promoters (Hiratsu, Shinagawa, & Makino, 1995). Additionally, this promoter has an extended $G_{-14}C_{-13}$ motif, which has been shown to be important for σ^S-selectivity (Lacour, Kolb, & Landini, 2003).

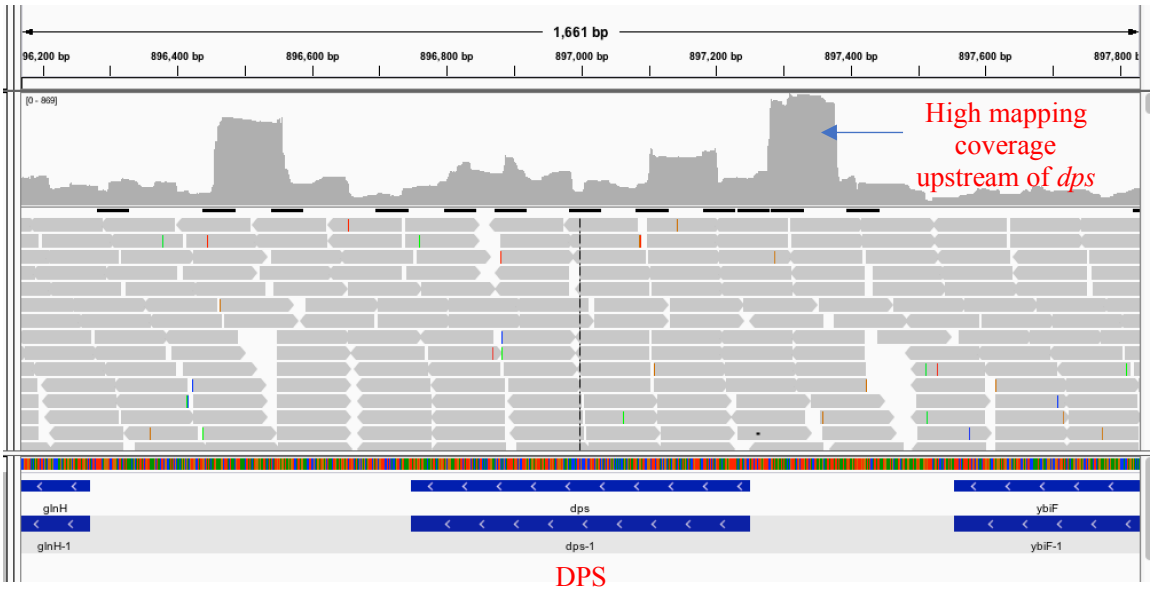


Fig. 2.3 Mapping of RNA-Seq reads for reaction with $E\sigma^S$ + Crl around the *dps* locus. The Integrative Genomics Viewer (IGV) was used to map the RNA-Seq reads into the Sty SL1344 genome. The locus around DPS is shown.

Visualization of the reads from the reaction conducted by $E\sigma^S$ and $E\sigma^S$ + Crl show high coverage of the *dps* gene, in particular upstream of its coding sequence (**Fig. 2.3**).

Differential expression analysis suggests roles for Crl beyond aiding holoenzyme formation

To validate this whole genome *in vitro* transcription approach, we conducted differential expression analysis to compare the hits that showed the most significant differences across the three conditions tested: $E\sigma^S$, $E\sigma^S$ + Crl, and $E\sigma^{70}$ (**Fig. 2.4**). Many of the genes that were preferentially expressed by $E\sigma^{70}$ were housekeeping genes that encode ribosome assembly proteins like *rpmH* and *rpsB* (Aseev, Levandovskaya, Tchufistova, Scaptsova, & Boni, 2008; Old, Margarita, & Girons, 1992). Reactions conducted with $E\sigma^S$ transcribed many more genes involved in response to stresses like *dps* and *trxA* (Ceci et al., 2004; Song et al., 2016) than $E\sigma^{70}$ did; the transcription of these and many other genes was enhanced in the presence of Crl. Interestingly, there are a group of genes that were preferentially transcribed by $E\sigma^S$ over $E\sigma^{70}$, but were directly repressed by Crl. If validated, direct repression by Crl would be a striking result that would add an intriguing component to its role as an unconventional transcription activator, possibly alluding to Crl having a role in promoter escape. Chapter 3 addresses the possibility of this role more thoroughly but I note that other unconventional transcription activators such as RbpA and CarD in *Mtb* have been suggested to repress some genes by inhibiting promoter escape (Jensen and Galburt 2019- NAR).

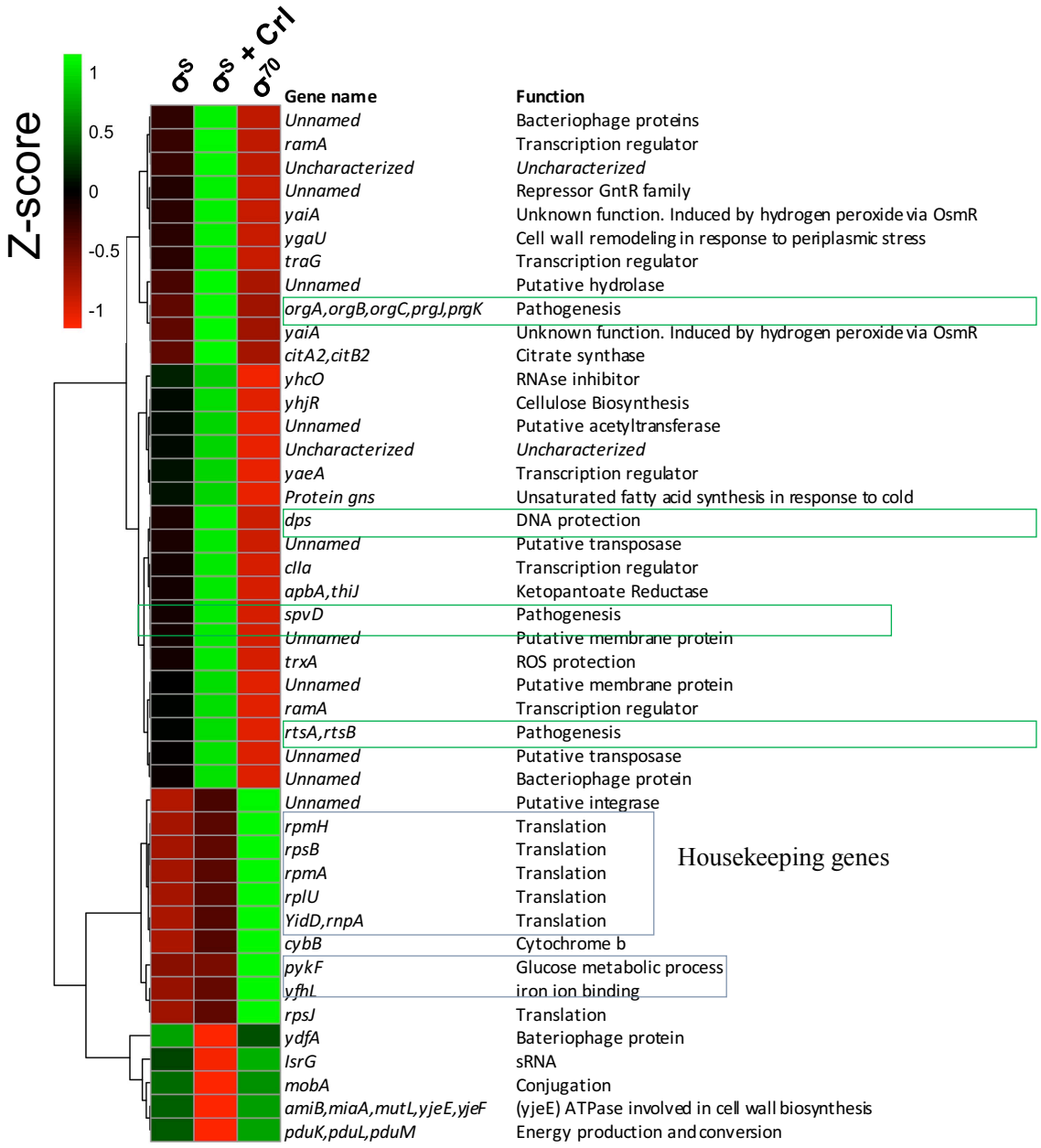


Fig. 2.4 Heatmap of genes most highly differentially transcribed by $E\sigma^S$ (left), $E\sigma^S$ + Crl (middle), and $E\sigma^{70}$ (right). Values are presented as the Z-score of each gene for the three conditions. Where known, the gene name and its biological function were noted. Green boxes denote interesting genes preferentially transcribed by $E\sigma^S$ + Crl including DPS and genes involved in pathogenesis.

Conclusion

These preliminary results suggest that *in vitro* transcription of whole genomes coupled with RNA-Seq (*in vitro* RNA-Seq) holds great promise to identify direct regulons of transcription factors. Many of the hits from the specific reactions with $E\sigma^S$, $E\sigma^S$ + Crl, and $E\sigma^{70}$ were indeed genes that fall under their specific regulons. The biggest advantage is that components of the reaction are controlled and experiments can be design to determine the global effect of one factor over the whole genome. The power of this approach can be further enhanced by coupling it with novel methods to identify transcription start sites, which can be used to extract information about promoter DNA motifs that might be targeted by specific factors (Ju, Li, & Liu, 2019; Ramachandran et al., 2012).

Materials and methods

Purification of *Salmonella enterica* serovar Typhimurium genomic DNA

Salmonella enterica serovar Typhimurium strain SL1344 was cultured in 10 mL LB and upon reaching an OD600nm of 0.25 (early stationary phase) and 2.5 (late stationary phase) the genomic DNA was purified using GenElute Bacterial Genomic DNA Kit (Sigma-Aldrich, Cat. No. NA2110). DNA concentrations were measured using a NanoDrop 2000 spectrophotometer.

Whole genome in vitro transcription reactions

200 µL transcription reactions were conducted using 100 nM core RNAP, 500 nM σS or σ70, 0 or 2 µM Crl, 1 µg Salmonella typhimurium genomic DNA, 500 µM rNTP

mix (NEB, Cat. N0466L), 1 unit/µL SUPERase RNase inhibitor (Invitrogen, Cat. AM2694). Reactions were run for 2 hours at 37 °C and transcription was stopped by incubating samples at 65 °C for 5 minutes and template DNA was digested by incubating the reactions with 10 units of DNase I (NEB, Cat. M0303S) for 20 minutes. Synthesized RNA was purified from the reactions using RNeasy Mini kit (Qiagen, 74104). RNA concentrations and quality were measured using a BioAnalyzer (Agilent) instrument.

RNA-Seq library preparation and sequencing

Purified RNA was sheered by sonication to generate ~200 bp fragment. A TruSeq (Illumina) RNA sample preparation approach was followed. The first cDNA strand was synthesized using reverse transcriptase, random DNA primers and the sheered RNA. The second DNA strand was synthesized using DNA Polymerase I in the presence of RNase H to degrade the original RNA. The DNA went through an end repair process and 3'end adenylation. DNA sequencing adapters were ligated which contain unique barcodes from individual samples. Samples were sequenced to 100 bases (single strand) using a HiSeq 2500 sequencing system (Illumina).

RNA sequencing read mapping and expression quantification

A differential RNA-Seq analysis was conducted using a pipeline based on the Tuxedo suite tools (Ghosh & Chan, 2016). Note, these preliminary experiments were only conducted with one set of replicates. Sequence reads were de-barcoded, trimmed to 80 bases, and mapped to the Salmonella enterica serovar Typhimurium strain SL1344

reference sequence genome (ASM21085v2)[1] using Bowtie2 (v 2.2.8) and TopHat (v 2.1.1) with default settings. Cufflinks (v 2.1.1) was used to calculate reads per kilobase of transcript per million fragments mapped (FPKM) abundance of the transcripts, compare the transcript assemblies to annotations, and find differential expressed transcripts. Data for the 50 highest differentially regulated genes was extracted and visualized using CummeRbund.

[1] http://bacteria.ensembl.org/Salmonella_enterica_subsp_enterica_serovar_typhimurium_str_sl1344/Info/Index

Chapter 3: **Structural Basis of Transcription Activation by Crl**

In bacteria, the primary σ factor associates with the core RNA polymerase (RNAP, E) to initiate the transcription of housekeeping genes. Bacteria also harbor alternative σ factors, which coordinate expression of additional regulons in response to environmental changes. Many alternative σ factors are negatively regulated by anti-σ factors. In *Escherichia coli*, *Salmonella enterica*, and other γ-proteobacteria, the transcription factor Crl positively regulates the alternative σ^S regulon by promoting the association of σ^S with RNAP. The molecular mechanism for activation by Crl is not completely understood. Here, we used single-particle cryo-electron microscopy to determine a structure of Crl-σ^S-RNAP in an open promoter complex with a σ^S-activated promoter. In addition to supporting previously predicted interactions between Crl and domain 2 of σ^S ($\sigma^S{}_2$), the structure, along with p-benzoylphenylalanine crosslinking, reveals that Crl interacts with a structural element of the RNAP β' subunit we term the β'-clamp-toe (β'CT). Deletion of the β'CT decreases activation by Crl without affecting basal transcription, highlighting the functional importance of the Crl-β'CT interaction. We conclude that Crl activates σ^S-dependent transcription in part by stabilizing σ^S-RNAP interactions through a tethering mechanism of $\sigma^S{}_2$ to RNAP via the β'CT.

Introduction

Bacteria contain a repertoire of σ factors, which compete against each other to bind a limited amount of available E (Feklístov et al., 2014; Maeda et al., 2000). *Escherichia coli* (*Eco*) has seven σ factors; σ^{70} is the primary (housekeeping) σ, while σ^S (encoded by *rpoS*) is the master regulator of transcription programs in the stationary phase of growth as well as in response to various stresses including antibiotics, UV light, low temperature, osmolarity changes, acidity changes, and nutrient depletion (Battesti et al., 2011). In certain conditions, the rapid and efficient expression of genes under σ^S control is critical for the survival of bacteria. However, once conditions become favorable for growth the σ^S transcription program must be shut down for optimal fitness. For these reasons, the expression of σ^S is highly regulated at transcriptional, translational, and post-translational levels (Landini et al., 2014).

Transcription from σ^S-dependent promoters can be limited by the $E\sigma^S$ concentration. To form $E\sigma^S$, σ^S must compete against other σ factors to assemble with E, for which σ^S has the weakest binding affinity (Maeda et al., 2000). Maeda and colleagues measured the dissociation constant K_d of σ^S to E to be 4.26 nM while σ^{70} to E is 0.26 nM (Maeda et al., 2000). In addition, the maximum cellular concentration of σ^S is only 30% that of σ^{70}. (Jishage & Ishihama, 1995). Thus, σ^S is at a significant disadvantage competing against σ^{70} to bind E. Crl is a ~16 kDa protein, widely distributed in γ-proteobacteria, that specifically activates $E\sigma^S$ transcription (Arnqvist et al., 1992; Pratt & Silhavy, 1998). Crl does not bind DNA like conventional transcription factors (Lee, Minchin, & Busby, 2012) but rather acts by directly binding $\sigma^S{}_2$ (Banta et al., 2013) and

stimulating expression of stress response genes, genes required for formation of amyloid curli fibers involved in adhesion and biofilm formation (Arnqvist et al., 1992), and many other genes in the σ^S regulon (Pratt & Silhavy, 1998).

Crl accumulates during bacterial exponential growth and reaches peak levels as bacteria enter stationary phase, with levels dropping as cells progress into late stationary phase (Bougdour et al., 2004; Robbe-Saule et al., 2006). By contrast, σ^S is not detectable until bacteria begin to enter stationary phase, and the level of σ^S continues to increase until late stationary phase (Jishage & Ishihama, 1995). This interplay in the levels of Crl and σ^S suggests a critical role for Crl when the levels of σ^S are very low. This is consistent with *in vitro* experiments demonstrating that transcription activation by Crl is most pronounced when σ^S concentrations are lowest (Gaal et al., 2006; Robbe-Saule et al., 2007; Typas et al., 2007). These findings have led to proposals that Crl functions by facilitating the assembly of E and σ^S into Eσ^S (Banta et al., 2013; Gaal et al., 2006).

Previous studies have determined structures of Eσ^S (Liu et al., 2016) and Crl (Banta et al., 2014; Cavaliere et al., 2014) separately. However, understanding the molecular mechanism of Crl has been hindered by the lack of structures of complexes of Crl with σ^S or with Eσ^S. Therefore, for my thesis I employed single-particle cryo-electron microscopy (cryo-EM) to determine the structure of Crl bound to an Eσ^S open promoter complex (RPo) containing a σ^S-regulon promoter. Initial analysis of the structure revealed that Crl simultaneously interacts with σ^S and E in the complex, stabilizing Eσ^S by tethering σ^S with RNAP. I therefore then tested the importance of this tethering interaction by performing function studies and confirmed this mechanism was critical for Crl activating transcription. This finding leads us to propose that Crl, and other

unconventional transcription activators that use a similar mechanism, be designated as σ-activators.

Results

Overexpression and purification of soluble σ^S

Several σ factors have a high propensity to aggregate when overexpressed. Previous purification protocols typically involve solubilizing σ^S found in inclusion bodies, denaturing, and re-folding (Burgess, 1996; Tanaka et al., 1993). Unfolding and refolding σ^S could lead to conformations that are not physiologically relevant, which could lead to misinterpretations in mechanistic studies of σ^S and Crl. At the outset of this research, attempts to overexpress and purify large amounts of *Eco* σ^S in our laboratory yielded low amounts as σ^S. Thus, we focused on the *Salmonella enterica* serovar Typhimurium (*Sty*) σ^S homolog, which is 99.1% identical to *Eco* σ^S and could be purified without the need for refolding (Cavaliere et al., 2014).

Initially, we explored if co-expressing σ^S with Crl could improve its solubility. Although this approach improved the solubility of σ^S we discovered that σ^S was proteolyzed by the PreScission Protease used to remove the His-tag fused to σ^S (**Fig. 3.1**).

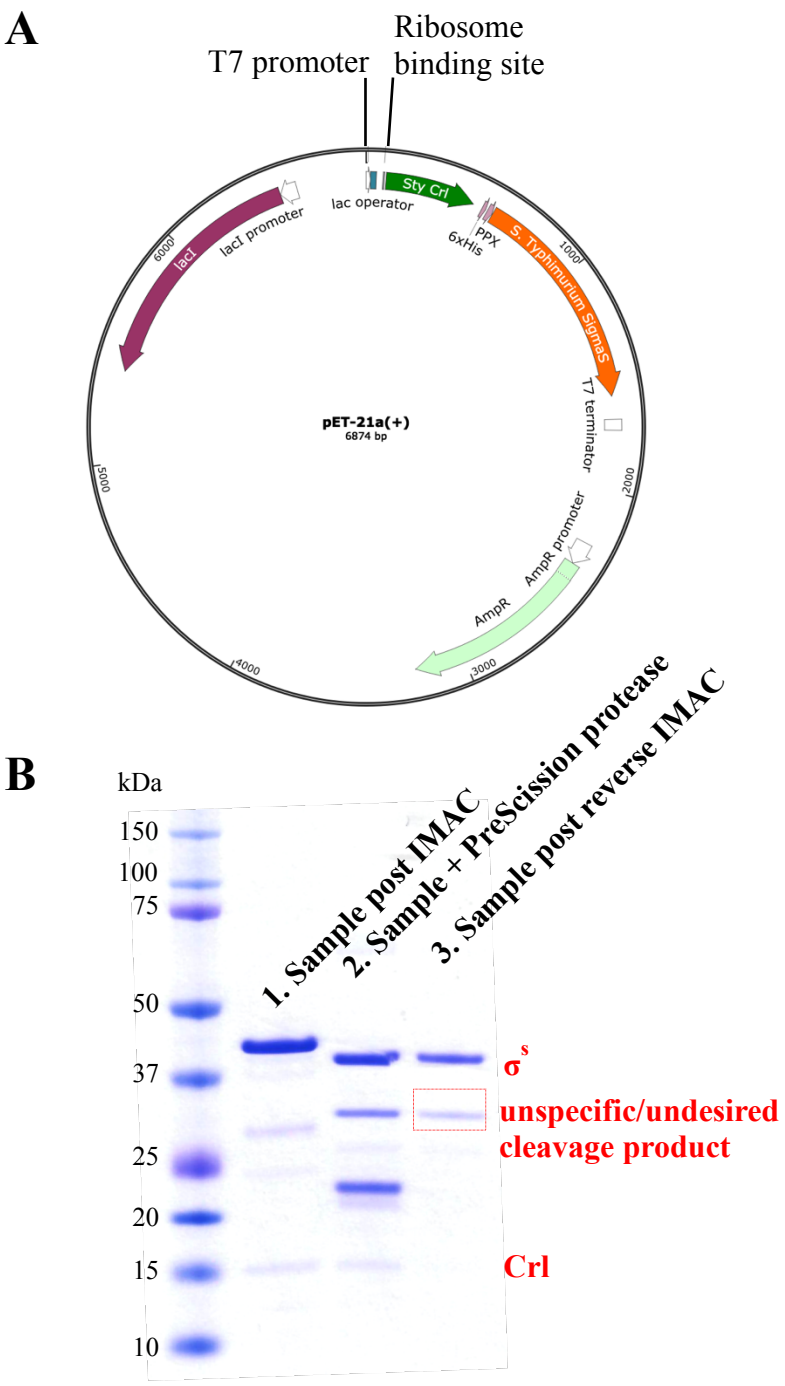


Fig. 3.1. *Sty* Crl/σ^S co-expression plasmid. A. A co-expression pET-21a construct was used to co-express untagged *Sty* Crl with a His-ppx-Sty-σ^s. **B.** SDS-PAGE of sample post IMAC, after adding PreScission protease (ppx) and incubating for 2 hours, and after running it again on IMAC to separate from ppx and any σ^s that remains tagged. Unspecific cleavage product is highlights in red.

This prompted us to developed an alternate expression and purification approach where we fused σ^S to His-SUMO, which would eliminate the issue of unspecific cleavage during His-tag removal (Mossessova & Lima, 2000; Reverter & Lima, 2004). This approach has been reported to enhance the solubility of expressed fusion proteins (Hubin et al., 2017; Malakhov et al., 2004) (US patent 6,872,551) (**Fig. 3.2**). This expression and purification methodology yielded large amounts of highly purified σ^S in the absence of Crl so we were no longer required to co-express the two proteins to obtain soluble σ^S **(Fig. 3.3).**

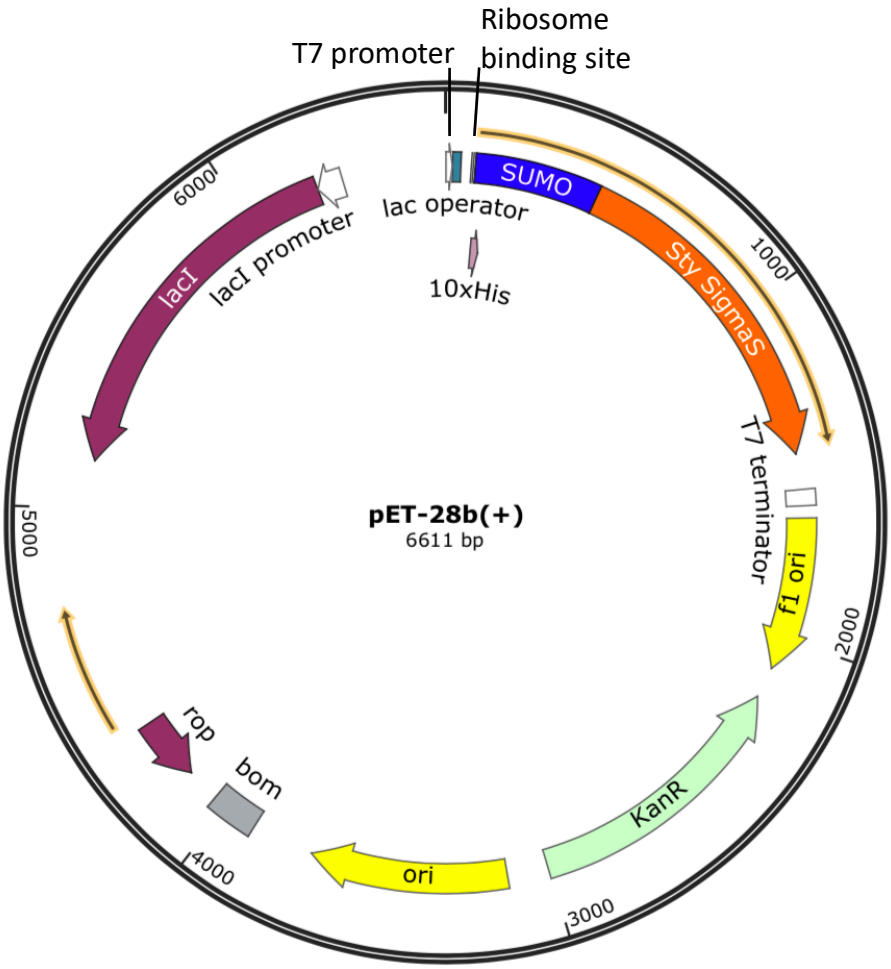


Fig. 3.2. *Sty* SUMO-σ^S expression plasmid. pET expressions vectors were created using Gibson assembly to generate a plasmid expressing polyhistidine-SUMO tagged *Sty* σ^S.

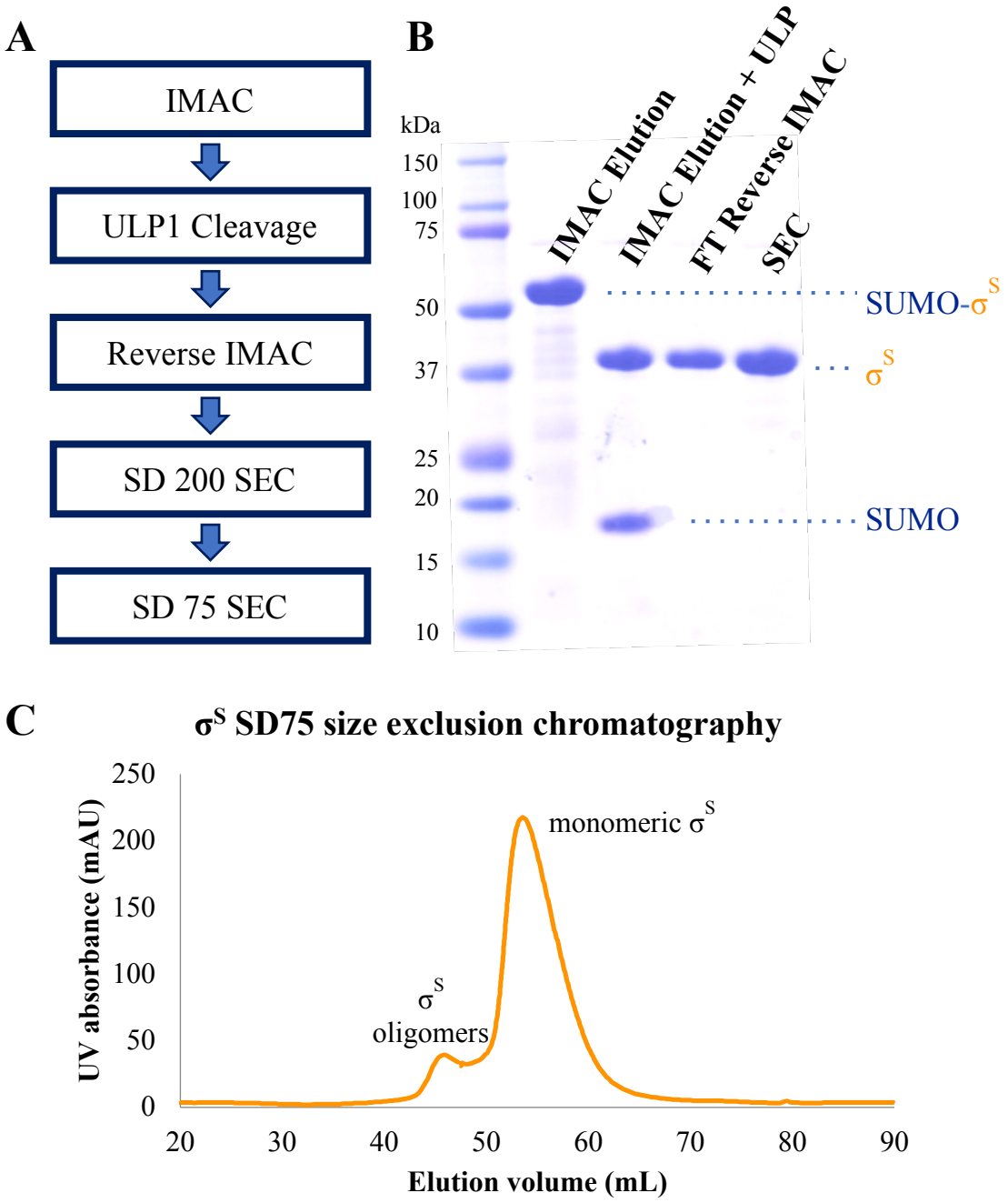


Fig. 3.3. *Sty* σ^S purification. A. Purification scheme: σ^S was recombinantly expressed in *E. coli*. The cells were harvested and lysed. The lysate was centrifuged and supernatant loaded into nickel-charged chelating column. The protein was eluted using 250 mM imidazole. The sample was incubated with ubiquitin-like protease I (ULP) to remove the SUMO tag. The sample was further purified using size exclusion chromatography. **B.** Coomassie blue stained SDS-PAGE of σ^S purification. **C.** Size exclusion chromatography of purified σ^S.

Sty Crl-σ^S activates *Eco* core RNAP

For our structural and functional analyses, we studied a complex between *Sty* Crl, *Sty* σ^S, and *Eco* E lacking the C-terminal domains of the α subunits. This chimeric approach allowed us to capitalize on the increased solubility of *Sty* σ^S compared to that of *Eco* σ^S. Note, Crl-σ^S in *Eco* and *Sty* have 95% sequence identity over 463 residues. The entire 443 kDa Crl-Eσ^S complex has 98.3% sequence identity over 4,576 residues between the two organisms (**Table 3.1**). *Proteus mirabilis* Crl (48% sequence identity with *Eco* Crl) and *Vibrio harveyi* Crl (only 39% sequence identity with *Eco* Crl) homologs fully complemented an *Eco* Crl deletion *in vivo* (Banta et al., 2014).

Purified *Sty* Crl activated *Eco* E/*Sty*-σ^S transcription ~5-fold (compared to no Crl) using an *in vitro* abortive initiation assay on a linear fragment of an *Eco* σ^S-regulon promoter, *dps* (Grainger et al., 2008) (**Fig. 3.4**), indicating that our Crl-Eσ^S complex is structurally and functionally relevant. This σ^S-dependent promoter encodes the DNA protection during starvation protein, which binds to the DNA and protects it from oxidative damage in stationary phase.

Table 3.1. Sequence identity between *Eco* and *Sty* components of RNAP

Subunit	Alignment length (residues)	% sequence Identity
α	329	100.0
β	1342	98.7
β'	1407	98.6
ω	91	100.0
σ^{70}	615	97.6
σ^S	330	99.1
Crl	133	84.2
overall	4,576	98.3

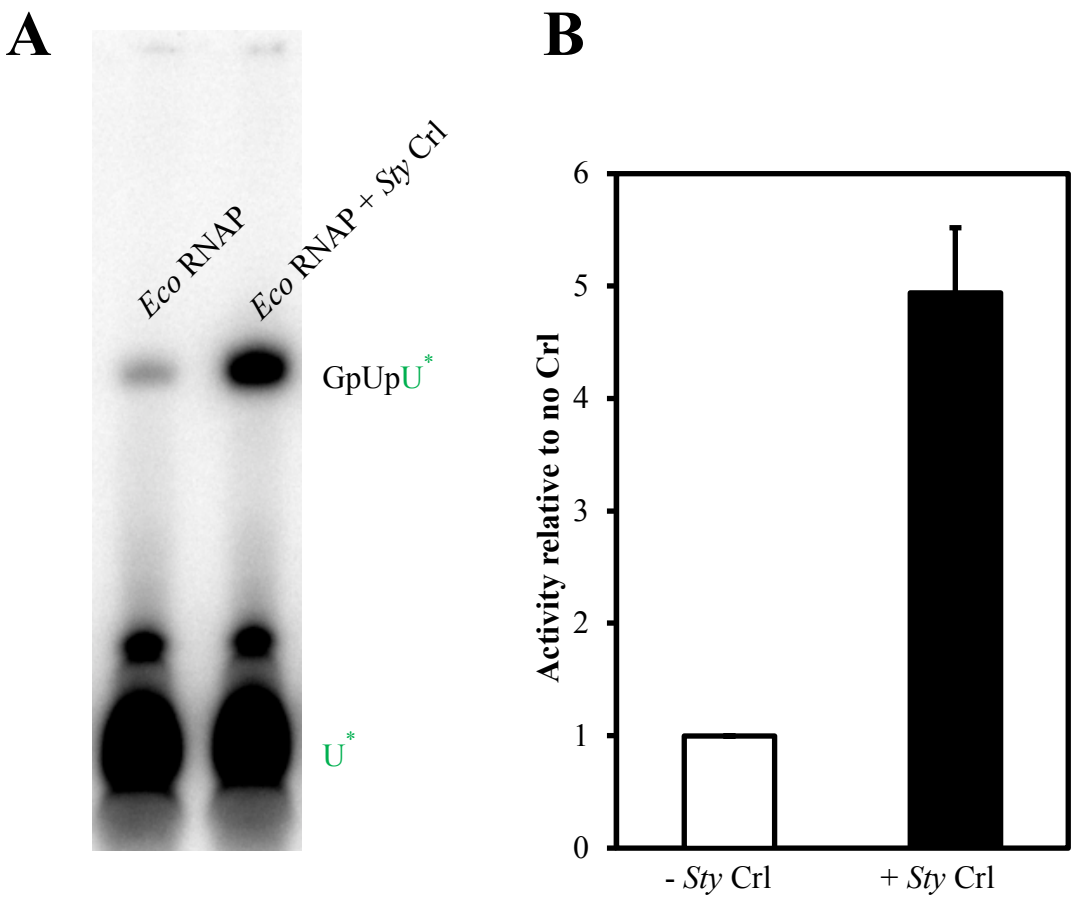


Fig. 3.4. *Sty* Crl-σ^S activates *Eco* core RNAP. A. Abortive initiation transcription assays measured GpUpU synthesis using a *dps* promoter template. [^{32}P]-labeled abortive transcript production was monitored by polyacrylamide gel electrophoresis and autoradiography. The intensity of the bands is representative of the amounts of products synthesized and the enzymatic activity of the sample. *Sty* Crl increases the enzymatic activity of *Eco* Eσ^S as evidence by the difference in intensity in reactions with and without Crl. **B.** The average effect of Crl on transcription across three experiments. Error bars denote the standard error. Under the condition tested, Crl enhances transcription of *Eco* Eσ^S five-fold.

Crl-Eσ^S-*dps*-RPo cryo-EM complex preparation

To prepare a complex for structural determination, we incubated Crl-Eσ^S with a *dps* promoter construct (-46 to +20) containing a non-complementary 'seed' bubble from -7 to -4 (**Fig. 3.5A**) to pre-nucleate the transcription bubble and favor the formation of a homogenous open promoter complex (RPo) as described previously (Liu et al., 2016). An RNA oligonucleotide (CUCG) complimentary to the template strand from -3 to +1 was also added, but it was not observed in the final reconstruction. The entire 477 kDa complex (Crl-Eσ^S-*dps*-RPo) was purified by size-exclusion chromatography (**Fig. 3.5B and C**) and cryo-EM grids were prepared as described in Materials and Methods. The structure of the complex was determined by single-particle cryo-EM (**Fig. 3.6**).

Crl-Eσ^S-*dps*-RPo cryo-EM structure determination and resolution optimization

Analysis of the cryo-EM data yielded a single structural class (**Fig. 3.6**) at a nominal resolution of 3.3 Å (**Fig. 3.7A**). Note, these datasets were initially processed with Relion 2.0, which only yielded a reconstruction of 4.0 Å resolution. Relion 3.0 implements two new features that improved the resolution from 4.0 Å to 3.3 Å: 1) a per-particle defocus and beam tilt estimation and 2) a Bayesian approach to correct beam-induced motion (Zivanov, Nakane, & Scheres, 2019). An improved CTF estimation can improve the Bayesian polishing process, and conversely, polished particles with higher resolution information will allow better CTF estimations. Thus, an iterative process of CTF refinement and Bayesian polishing for multiple rounds can greatly improve the resolution of the reconstruction. In our case, the pre-polishing resolution was 4.01 Å. After the first round of per-particle CTF-estimation and Bayesian polishing the resolution

improved to 3.66 Å, after the second round it was 3.50 Å, after the third 3.32 Å, after the fourth 3.30 Å, and it remained at 3.30 Å after the fifth round (**Fig. 3.6**). The sample exhibited a good distribution of particle orientations (**Fig. 3.7B**). The local resolution ranged from 2.8 Å in the well-ordered core of the complex to 6.5 Å at the flexible periphery (**Fig. 3.7C**). A structural model was built and refined into the cryo-EM map (**Table 3.2, Fig. 3.8**). Initial examination of the cryo-EM structure revealed three key features (**Fig. 3.8**): First, expected interactions occurred between σ^{S}_{4} and the -35 promoter element, which were not observed in a previously determined crystal structure of a σ^{S} transcription initiation complex due to crystal packing restraints (PDB 5IPL) (Liu et al., 2016); second, Crl bound σ^{S}_{2} in a manner predicted from the results of previous studies (Banta et al., 2014; Cavaliere et al., 2015) and is located at the periphery of the complex near the upstream edge of the transcription bubble; third, Crl also interacted with the RNAP β' subunit.

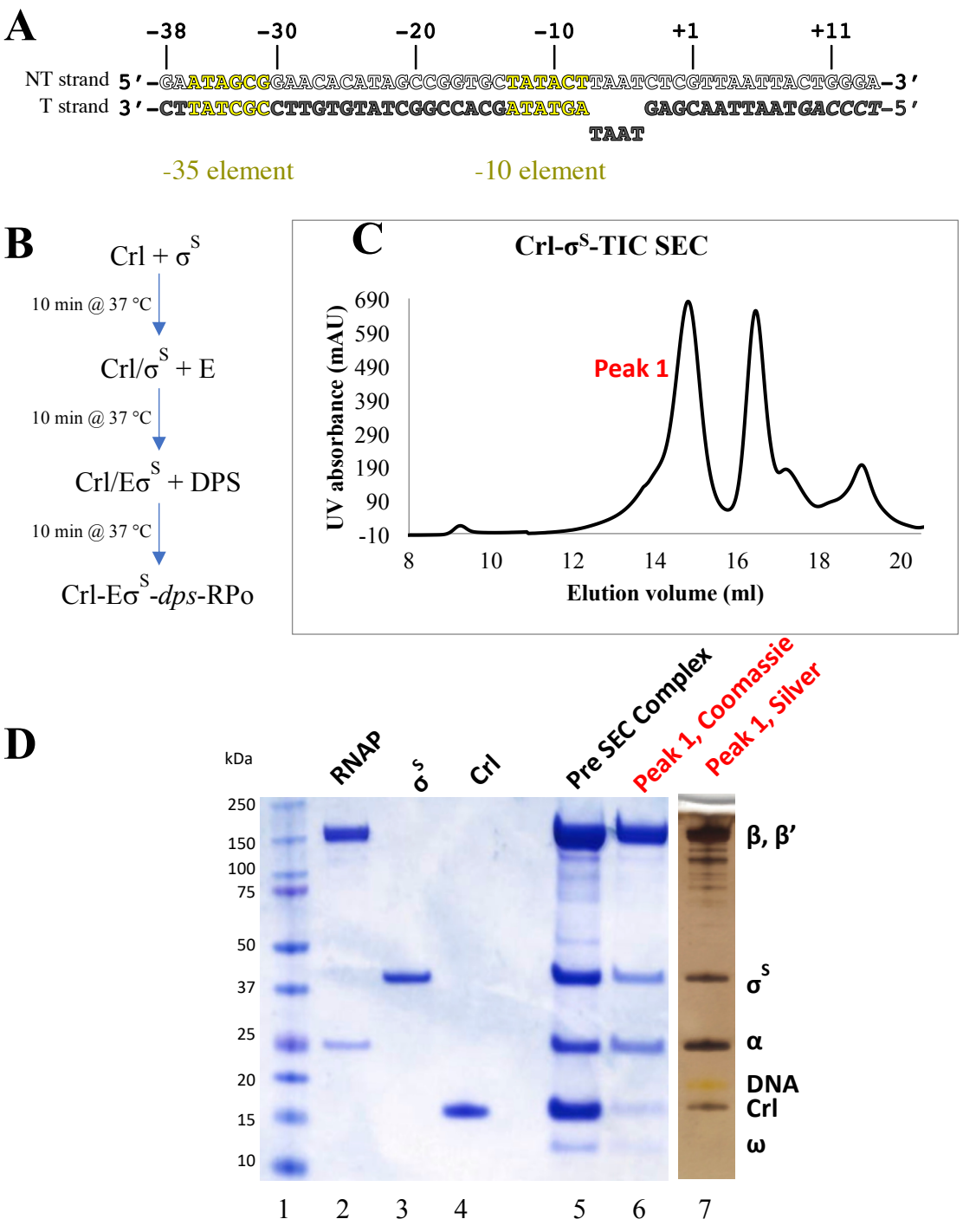
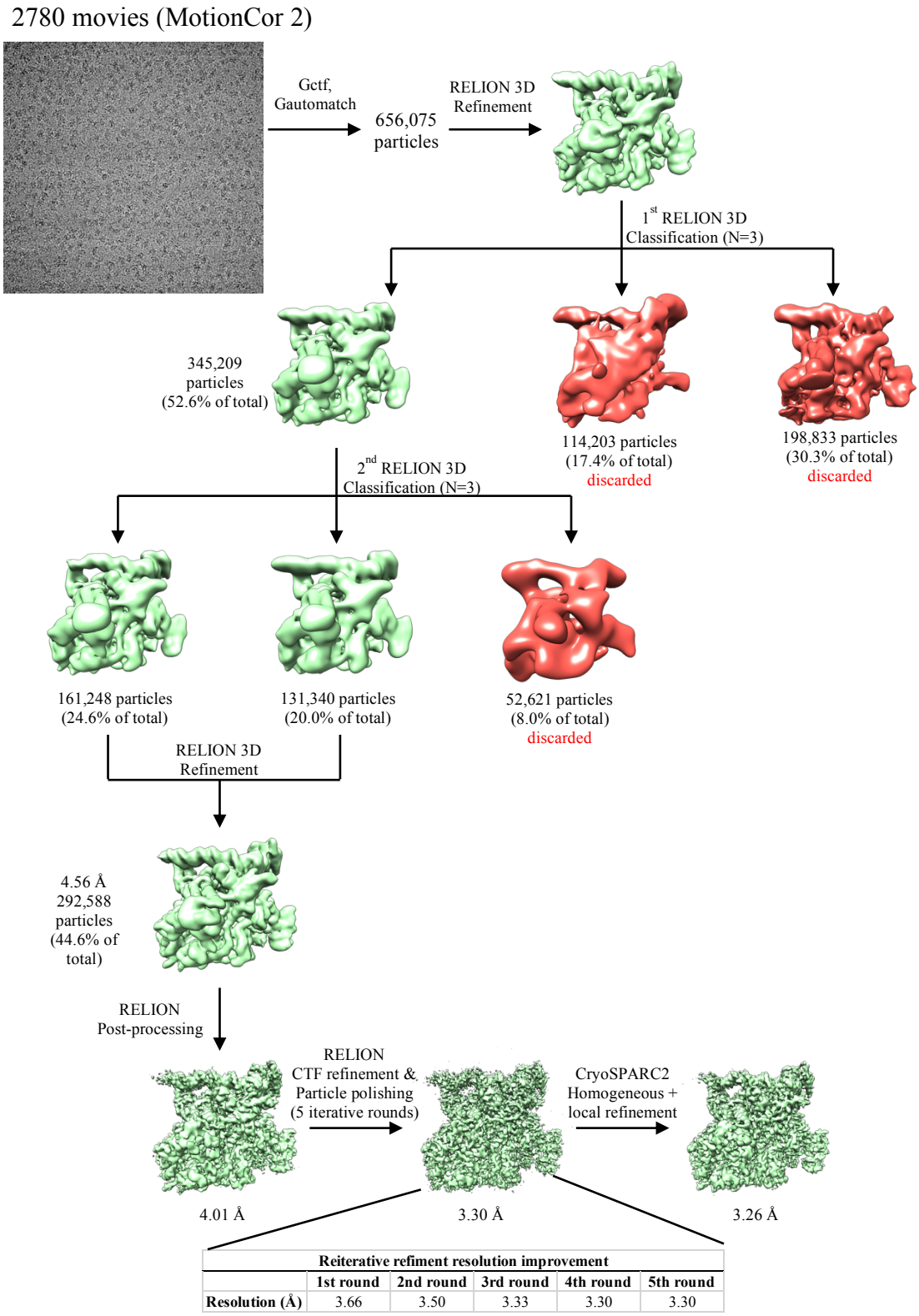


Fig. 3.5. Assembly of Crl-Eσ^S-*dps*-RPo. A. The *Eco dps* promoter construct used for cryo-EM. The template (T) and non-template (NT) strands are shown. **B.** Order of component addition to form Crl-Eσ^S-*dps*-RPo. **C.** Size exclusion chromatography (SEC) profile of Crl-Eσ^S-*dps*-RPo. Peak 1 is analyzed by SDS-PAGE. **D.** SDS-PAGE analysis of Crl-Eσ^S-*dps*-RPo. Lane 1, molecular weight markers; lane 2, purified *Eco* core ΔαCTD-RNAP; lane 3, purified *Sty* σ^S; lane 4, purified *Sty* Crl; lane 5, SEC load; lane 6, Peak 1 from SEC elution [see (**C**)], Coomassie blue stained; lane 7, Peak 1 from SEC elution, silver stained.



Reiterative refiment resolution improvement					
	1st round	2nd round	3rd round	4th round	5th round
Resolution (Å)	3.66	3.50	3.33	3.30	3.30

Fig. 3.6. Crl-Eσ^{S}-*dps*-RPo cryo-EM reconstruction pipeline. The Cryo-EM data was processed as described in Materials and Method to yield a final density map with a nominal resolution of 3.3 Å.

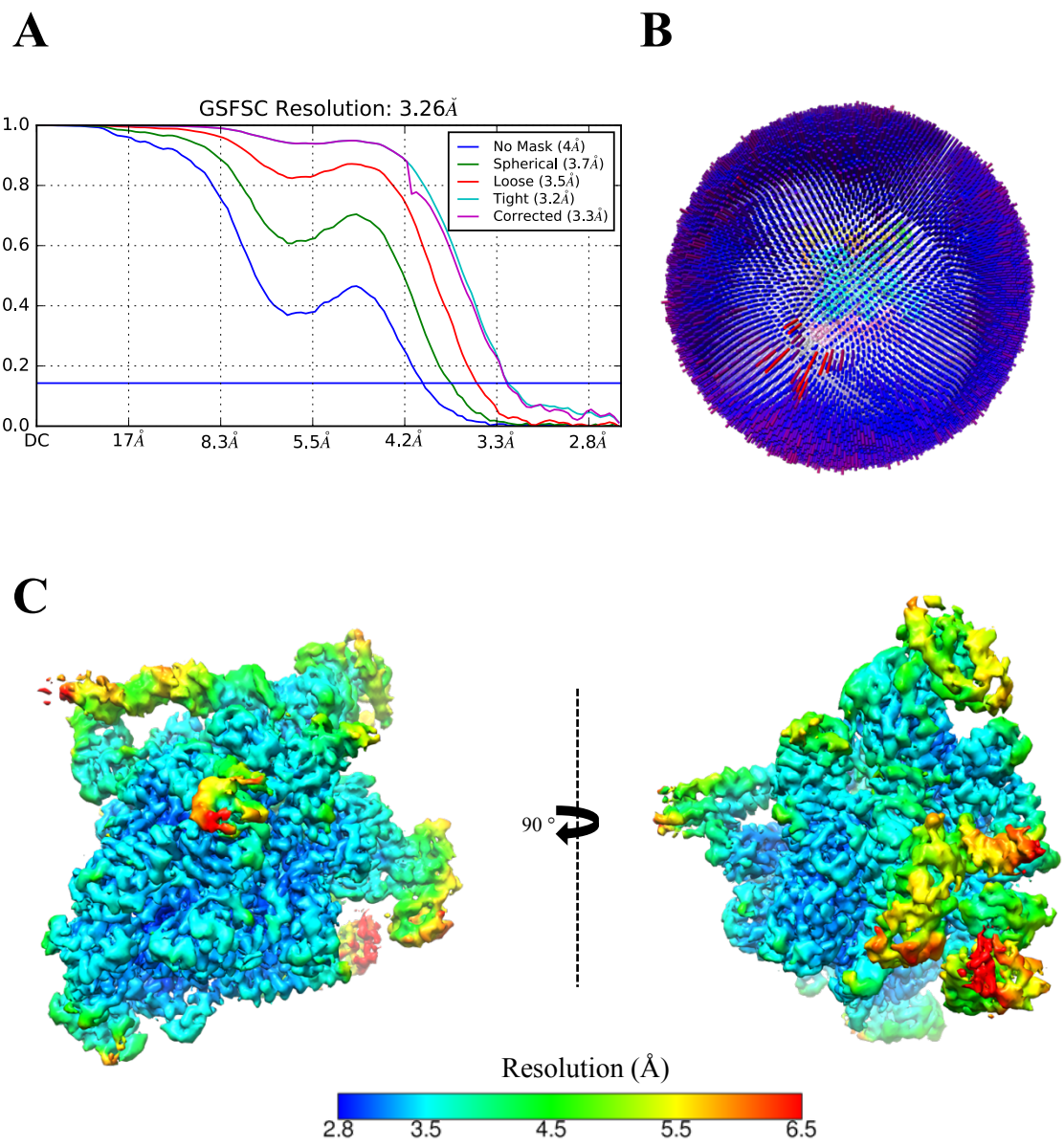


Fig. 3.7. Cryo-EM density map of Crl-Eσ^S-*dps*-RPo resolution information. A. Gold-standard Fourier shell correlation calculated by cryoSPARC. The 0.143 FSC cutoff (blue horizontal line) corresponds to a nominal resolution of 3.26 Å. **B.** Angular distribution histogram Crl-σ^S-TIC particle projections from Relion3. Bar height represents number of particles of each viewing angle. **C.** Reconstruction density map colored by local resolution. From Jaramillo Cartagena, 2019.

Table 3.2. Cryo-EM data collection, refinement and validation statistics

	Crl-Eσ^S RPo (EMDB-20090) (PDB 6OMF)
Data collection and processing	
Magnification	22,500
Voltage (kV)	300
Electron exposure (e–/Å^2)	71
Defocus range (μm)	0.8 – 2.4
Pixel size (Å)	1.3
Symmetry imposed	C1
Initial particle images (no.)	656,075
Final particle images (no.)	292,588
Map resolution (Å) - FSC threshold 0.143	3.3
Map resolution range (Å)	2.8 - 6.5
Refinement	
Initial model used (PDB code)	RPo, 5IPL/Crl, 3RPJ
Model resolution (Å)	
FSC threshold 0.5	5IPL 3.6, Crl 1.9
Model resolution range (Å)	3.1 – 7
Map sharpening *B* factor (Å^2)	141.1
Model composition	
Non-hydrogen atoms	30,250
Protein residues	3,616
Nucleic acid residues	92
Ligands	3 (1 Mg^{2+}, 2 Zn^{2+})
B factors (Å^2)	
Protein	144.3
Nucleic acid	219.7
Ligands	143.1
R.m.s. deviations	
Bond lengths (Å)	0.011
Bond angles (°)	0.978
Validation	
MolProbity score	2.78
Clashscore	9.32
Poor rotamers (%)	12.31
Ramachandran plot[a]	
Favored (%)	82.5
Allowed (%)	17.6
Disallowed (%)	0

[a] Ramachandran plot parameters from PROCHECK (Laskowski et al., 1993)

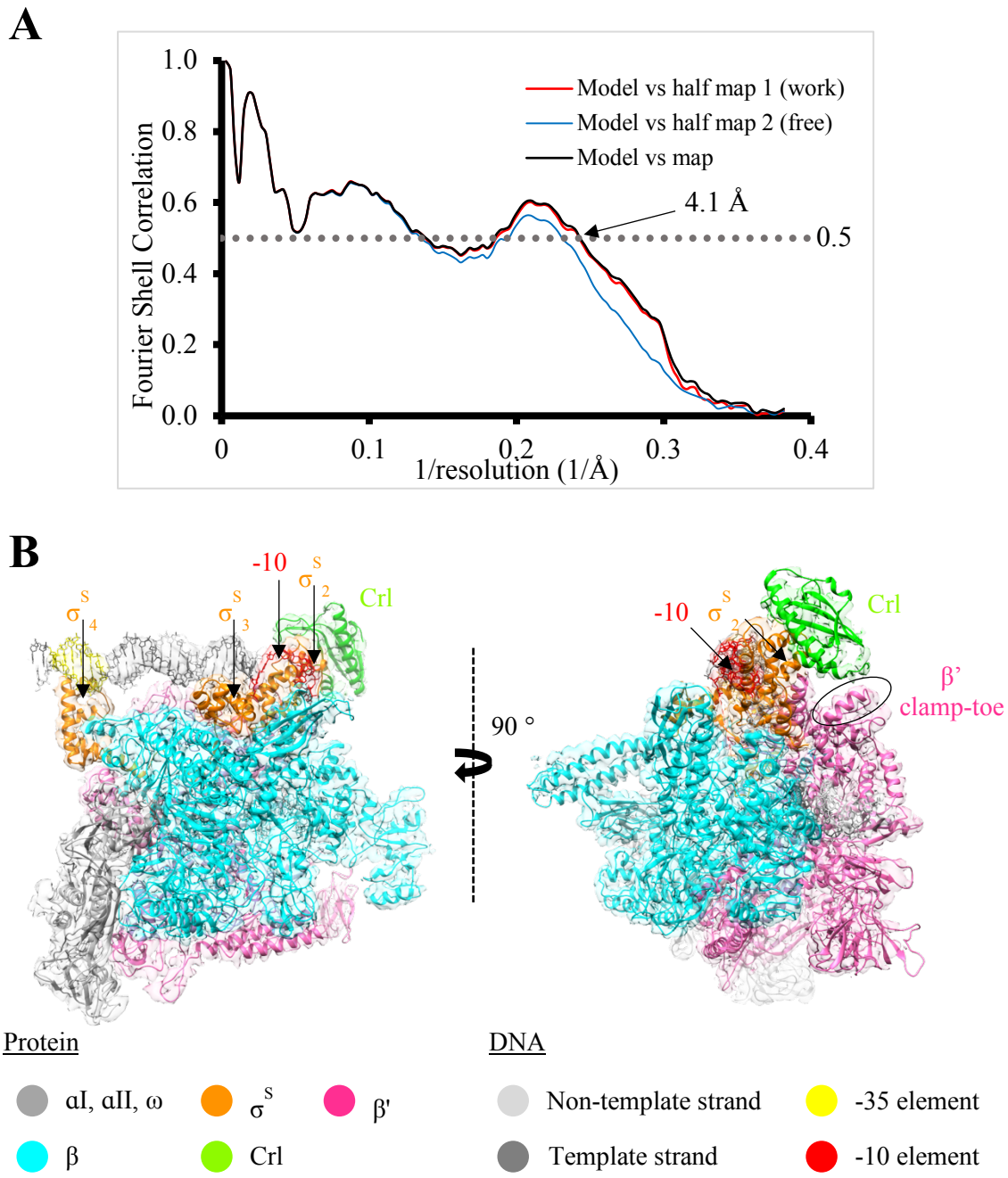


Fig. 3.8. Cryo-EM structure of Crl-Eσ^S-*dps*-RPo. A. FSC calculated between the refined structure and the half map used for refinement (work, red), the other half map (free, blue) and the full map (black). **B.** The 3.3 Å nominal resolution cryo-EM density map of Crl-Eσ^S-*dps*-RPo is rendered as a transparent surface and colored as labeled. The map is low-pass filtered according to the local resolution (Cardone G, Heyman JB, Steven AC, 2013) . Superimposed is the final refined model. The proteins are shown as backbone ribbons, and the nucleic acids are shown in stick format.

Structural basis for selective activation of σ^S by Crl

The σ^S is the closest relative of σ^{70} in terms of sequence, domain architecture, and promoter recognition properties (Paget & Helmann, 2003). Our structure and the Eσ^S-RPo crystal structure (Liu et al., 2016) revealed the expected structural similarity between domains 2, 3, and 4 of σ^S and σ^{70} (**Fig. 3.9**). Despite these similarities, Crl specifically activates σ^S. The main difference between σ^S and σ^{70} is the non-conserved region (NCR) of σ^{70}, a 250 amino acid insertion located between regions 1.2 and 2.1 that is absent in σ^S (**Fig. 3.9**).

As previously reported, Crl is small arc-shaped protein with a shallow concave surface composed of four antiparallel β-strands and flanked by intervening loops (Banta et al., 2014; Cavaliere et al., 2014). This cavity makes extensive electrostatic, polar, and hydrophobic interactions with helix α2 (A73 to R85) of σ^S, which resides within conserved region $\sigma^S_{1.2}$ (Lonetto et al., 1992) (**Fig. 3.9**). In σ^{70}, the corresponding helix extends to become part of the NCR and would sterically clash with Crl binding (Banta et al., 2013), explaining why Crl selectively binds and activates σ^S but not σ^{70} (**Fig. 3.10**).

Fig. 3.9. (Next page) *Sty* σ^{70} and σ^{S} sequence alignment, secondary structure, and domain architecture. Identical residues between the two sequences are boxed red, homologous residues colored red and boxed in white. Above and below the sequences is schematically illustrated the σ^{70} (above) *Sty* σ^{S} (below) secondary structures (α-helices indicated by coils). Above the secondary structure is illustrated the domain architecture. $\sigma_{1.1}$ and σ_{NCR} are unique to σ^{70}. The σ^{S} residues that interact directly with Crl in the cryo-EM structure are marked with green stars and shaded green.

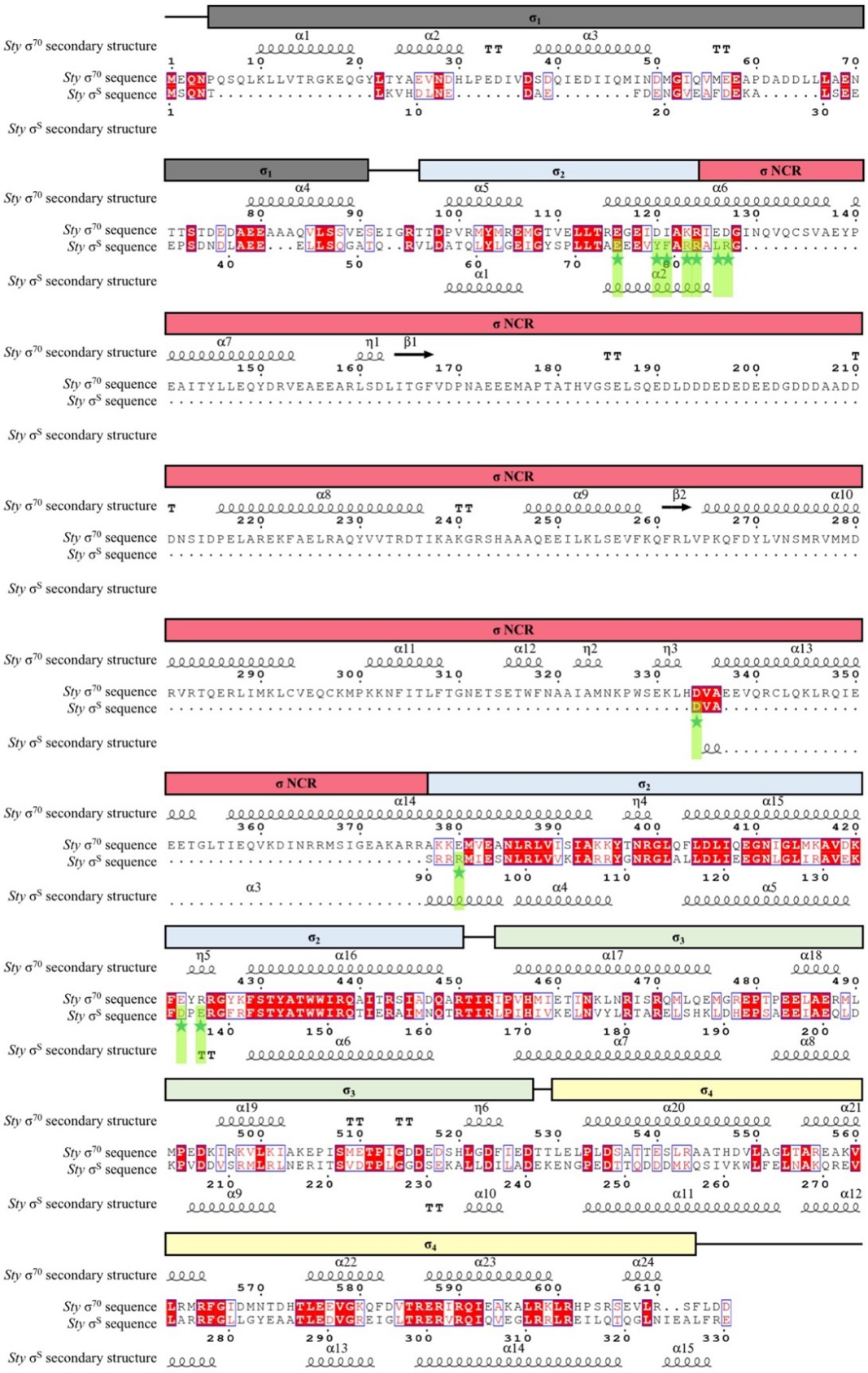


Fig. 3.9. *Sty* σ^{70} and σ^{S} sequence alignment, secondary structure, and domain architecture

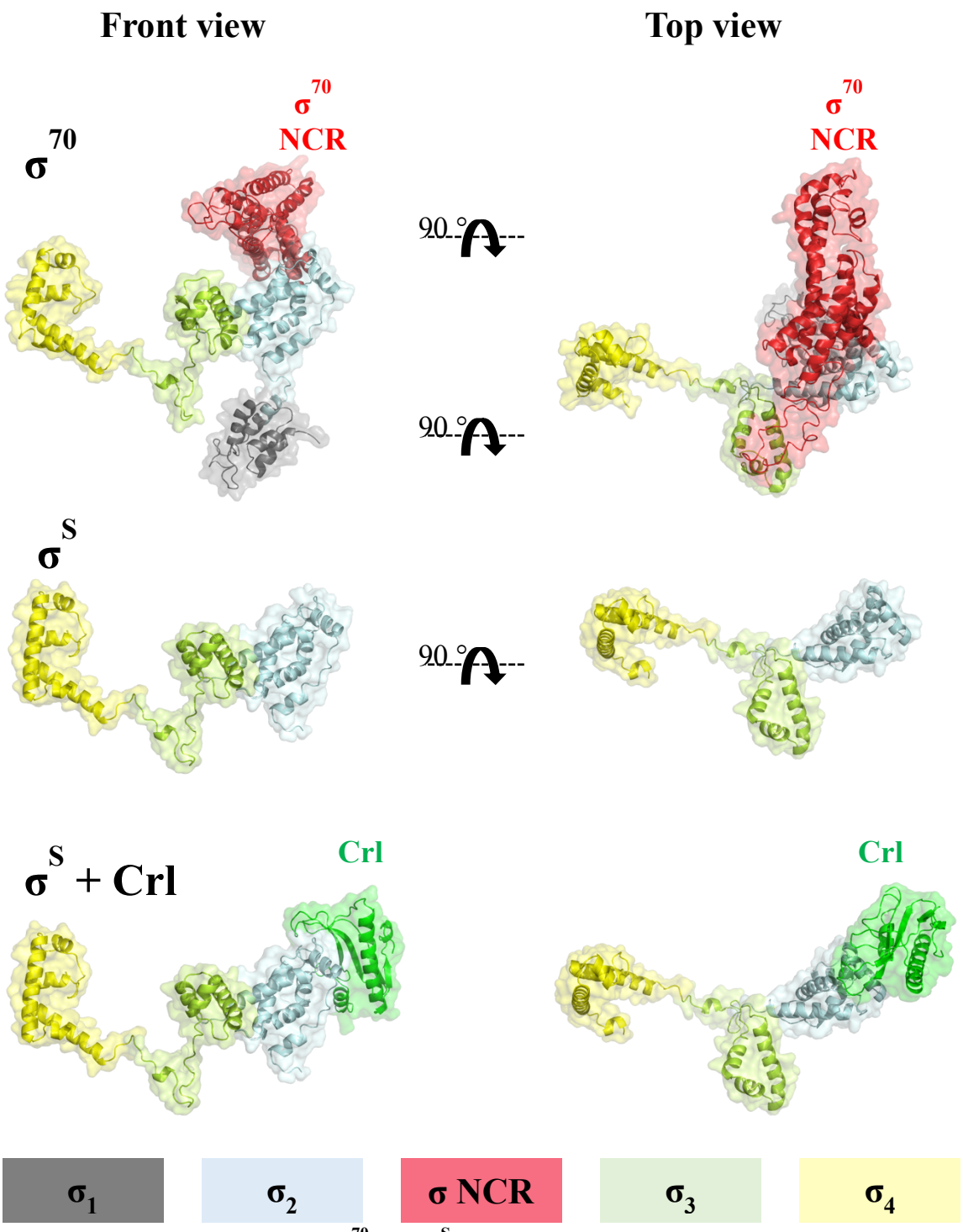


Fig. 3.10. Comparison of *Sty* σ^{70} and σ^{S} structural domain architecture. The *Sty* σ^{70} model was generated by homology modeling using the *Sty* σ^{70} sequence and the structure of *Eco* σ^{70} (from PDB 4LK1) (Bae et al. 2013) using SWISS-MODEL Workspace (Waterhouse A., et al. 2018). *Sty* and *Eco* σ^{70} have 97.6% sequence identity (Table 1).

The central role of σ^S helix α2 confirms previous reports that identified key residues in σ^S for its interaction with Crl. *Sty* σ^S R82 is not conserved in *Pseudomonas aeruginosa* (*Pae*) σ^S, and *Pae* σ^S does not interact with *Sty* Crl in *in vivo* bacterial two-hybrid assays unless the corresponding amino acid (Leu) is mutated to an arginine (Cavaliere et al., 2015). Mutations at this site lead to bacterial colony morphology changes, which is consistent with the interaction of Crl and σ^S being important for processes like biofilm formation. In our structure, *Sty* σ^S R82 is positioned towards the central cavity of Crl and forms an extensive network of electrostatic, polar, and hydrophobic interactions with Crl-P21, Y22, I23, D36, and C37 (**Fig. 3.11**). Our structure also validates the importance of other sites in helix α2 like Y78 and F79, which have been substituted with benzoyl-L-phenylalanine (BPA) and shown to crosslink to Crl (Banta et al., 2013).

Adjacent to σ^S helix α2 is a loop within $\sigma^S_{2.3}$ that also makes significant interactions with Crl, in particular σ^S residues D135 and E137 (**Fig. 3.11**). D135 makes favorable electrostatic interactions with Crl R51, which is absolutely conserved among Crl homologs (Monteil et al., 2010). Substitutions Crl R51A or R51K were totally defective in Crl function *in vitro* and *in vivo* (Banta et al., 2014). D135 and E137 of σ^S, along with P136, have been referred to as the DPE motif (**Fig. 3.11**) and is a key difference from σ^{70} (Banta et al., 2013). In bacterial two-hybrid assays, substitutions in the DPE motif significantly decrease the interactions between Crl and σ^S, and a chimeric σ^{70} mutant containing this region of σ^S can interact with Crl (Banta et al., 2013). Thus, substitutions that alter either partner of the Crl/σ^S-DPE motif interaction interface highlight the importance of this interaction for Crl function.

Altogether, the interaction between Crl and σ^S forms an interfacial area of 785 $Å^2$ (Krissinel & Henrick, 2007) and is completely consistent with previous analyses of Crl-σ^S interactions (Banta et al., 2013, 2014; Cavaliere et al., 2018; Monteil et al., 2010).

In summary, our structure: 1) confirms the σ^S residues previously proposed to interact with Crl based on genetic and biochemical data; 2) identifies additional residues in σ^S and Crl involved in the intermolecular interaction; 3) reveals that 6 out of 11 (55%) of the residues of σ^S contacting Crl are not conserved in σ^{70} (**Fig. 3.9 and Fig. 3.11)** shows that the NCR of σ^{70} occupies some of the volume corresponding to the location of Crl in a Crl/σ^S complex (Banta et al., 2013) (**Fig. 3.11**), thereby explaining the molecular basis of Crl discrimination for regulating σ^S and not σ^{70}.

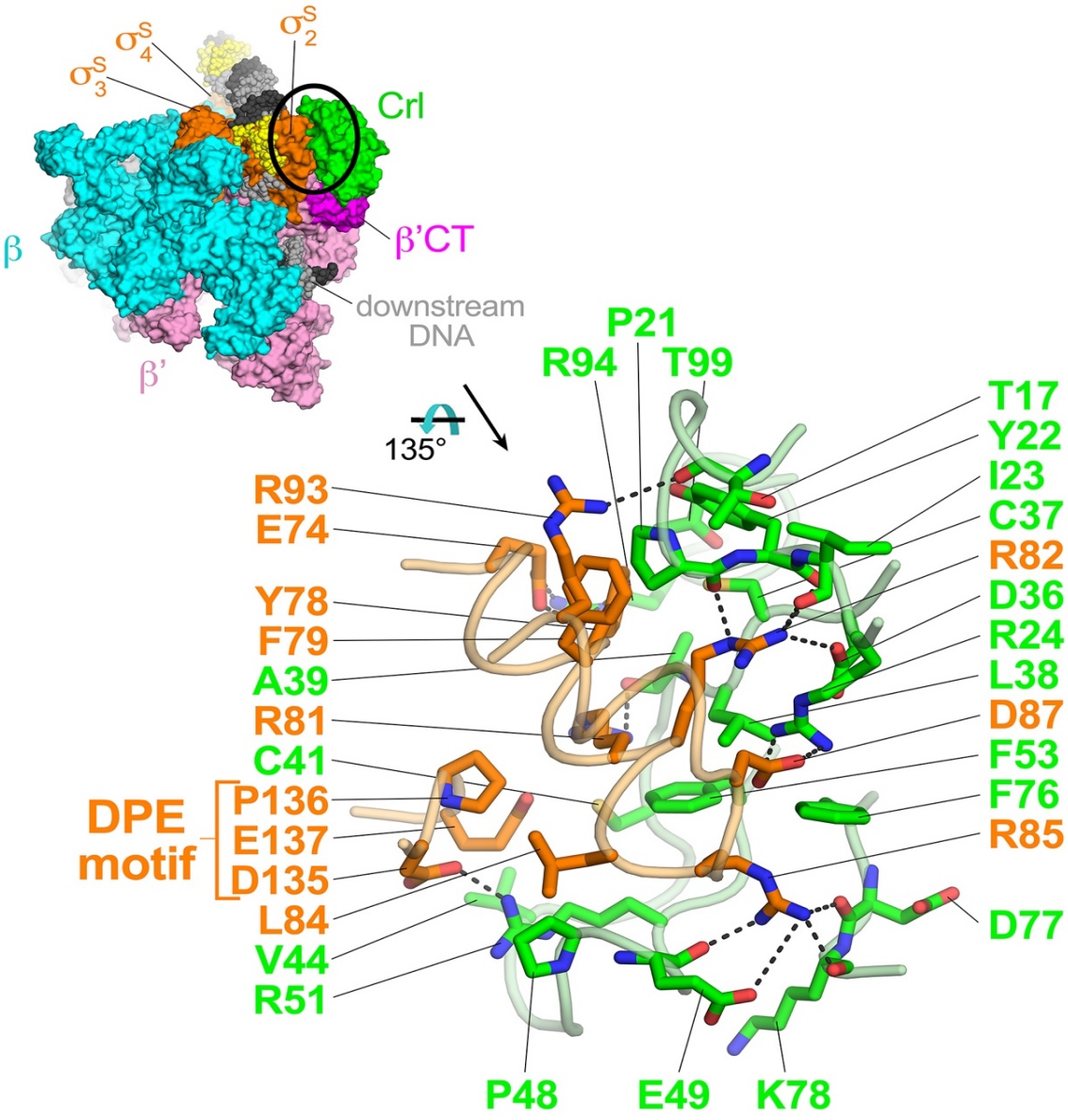


Fig. 3.11. Crl-σ^S_2 interactions. ***Top left:*** The overall structure of Crl-Eσ^S-*dps*-RPo. Proteins are shown as molecular surfaces with subunits colored as labeled. The DNA is shown as Corey-Pauling-Koltun (CPK) spheres and colored according to Fig. 2.7. The circled region is magnified below. ***Bottom right:*** Crl-σ^S_2 interactions. Crl (green) and σ^S_2 (orange) are shown as backbone worms. Residues that interact are shown in stick format. Polar interactions are denoted by dashed gray lines. From Jaramillo Cartagena, 2019.

Crl tethers σ^S to core RNAP to help activate transcription

In addition to extensive contacts with σ^S_2 (**Fig. 3.11**), Crl in the Crl-Eσ^S-*dps*-RPo structure interacts with a small domain of the *Eco* RNAP β' subunit we call the β' clamp-toe (β'CT; β' residues 144-179; **Fig. 3.8B**). A structure-guided sequence alignment of evolutionarily diverse Crl homologs reveals conservation of basic amino acids in the region of Crl that interacts with the β'CT, corresponding to *Sty* Crl K9, R11, K14, and K15 (**Fig. 3.15**) (Crooks et al., 2004; Robert & Gouet, 2014). The β'CT is not strictly conserved among bacterial RNAP β' subunits as it is the site of lineage-specific-insertions in many bacterial clades, including *Deinococcus-Thermus* and Actinobacteria (Lane & Darst, 2010). However, the sequence of the β'CT is conserved among RNAP β' sequences from γ-proteobacteria, including residues that interact directly with Crl: *Eco* β'L166, as well as two acidic residues, D167 and E170 (**Fig. 3.15**). The α-helix of the β'CT that interacts with Crl harbors five conserved acidic residues corresponding to *Eco* β' E162, E163, D167, E170, and E171 (**Fig. 3.15**). The observation that Crl interacts with E was verified by our collaborators R. Gourse, W. Ross , and A. Banta by photo-crosslinking experiments (**Fig. 3.16**). Crl crosslinked to β' with BPA substitutions at D167 and F172 but not E142 in a σ^S-dependent manner, consistent with our structure (**Fig. 3.16A**). Surface electrostatic analysis supports favorable interactions between Crl and β' (Baker et al., 2001) (**Fig. 3.16A**).

Testing the role of the β'CT in Crl activation

These observations suggest that Crl may assist Eσ^S assembly by interacting with σ^S_2 and RNAP simultaneously. To test this hypothesis, we investigated Crl function with

a mutant RNAP in which the entire β'CT was deleted (Δβ'CT-E) (**Fig. 3.12 – 3.14**) using the quantitative abortive initiation assay with the *dps* promoter. While wt-Eσ^S and Δβ'CT-Eσ^S had essentially the same transcription activity in the absence of Crl, the presence of a saturating concentration of Crl activated wt-Eσ^S ~5-fold compared to only ~2.3-fold with Δβ'CT-Eσ^S (**Fig. 3.17**). We conclude that the simultaneous interaction of Crl with σ^S_2 and the β'CT helps tether σ^S to the E increasing the stability of Eσ^S, accounting for partial, but not full, transcription activation function of Crl.

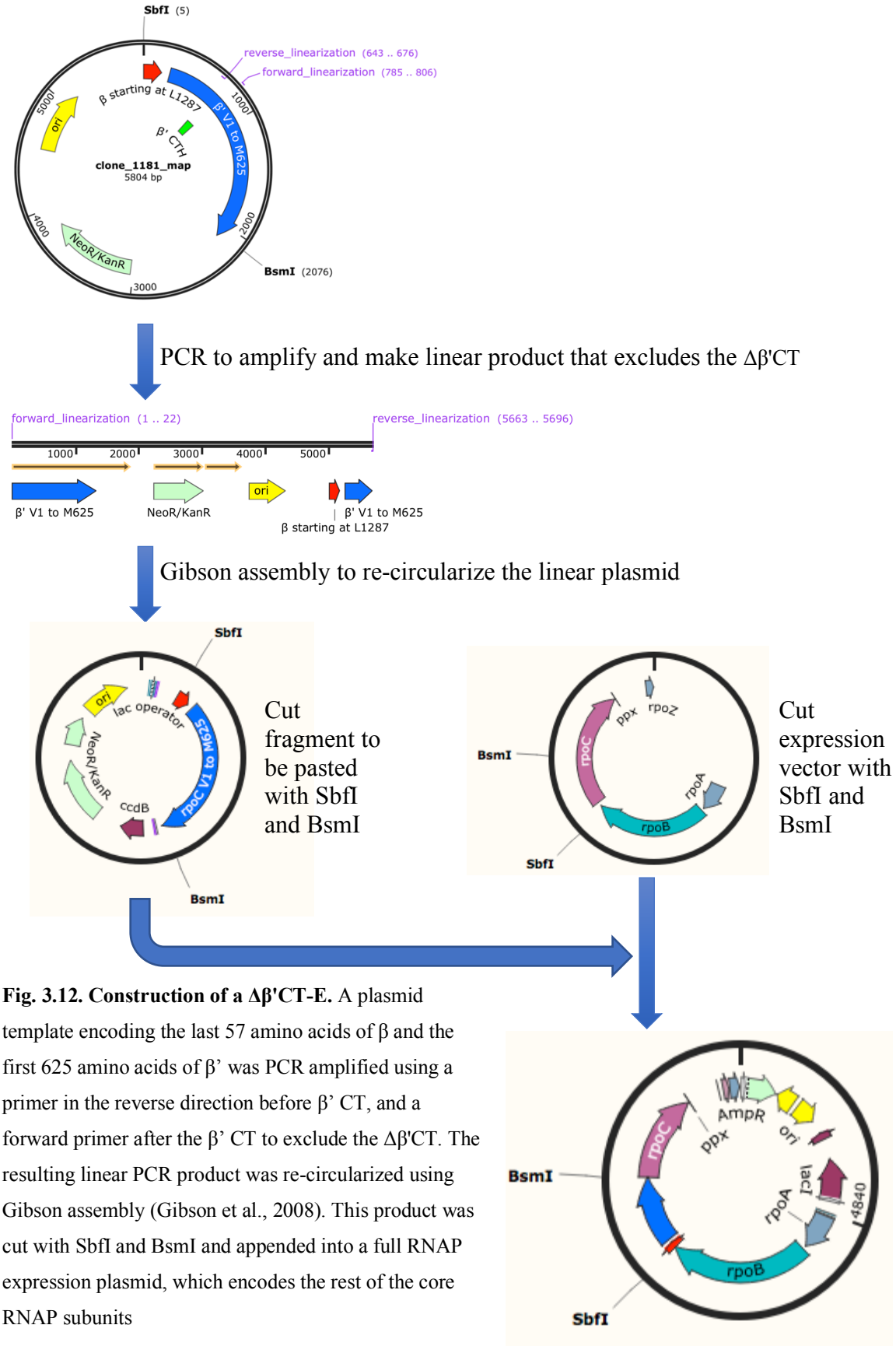


Fig. 3.12. Construction of a Δβ'CT-E. A plasmid template encoding the last 57 amino acids of β and the first 625 amino acids of β' was PCR amplified using a primer in the reverse direction before β' CT, and a forward primer after the β' CT to exclude the Δβ'CT. The resulting linear PCR product was re-circularized using Gibson assembly (Gibson et al., 2008). This product was cut with SbfI and BsmI and appended into a full RNAP expression plasmid, which encodes the rest of the core RNAP subunits

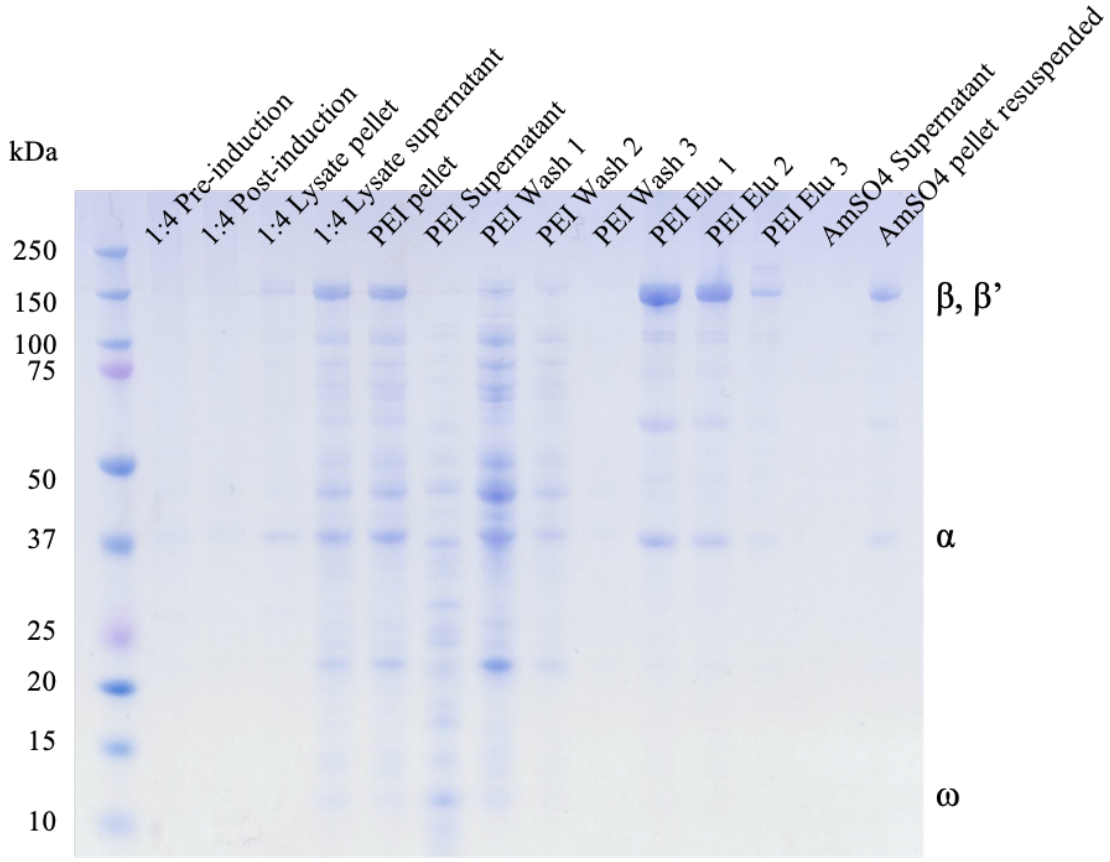


Fig. 3.13.A. Purification of Δβ'CT-E I

PEI: polyethyleneimine

AmSO4: ammonium persulfate

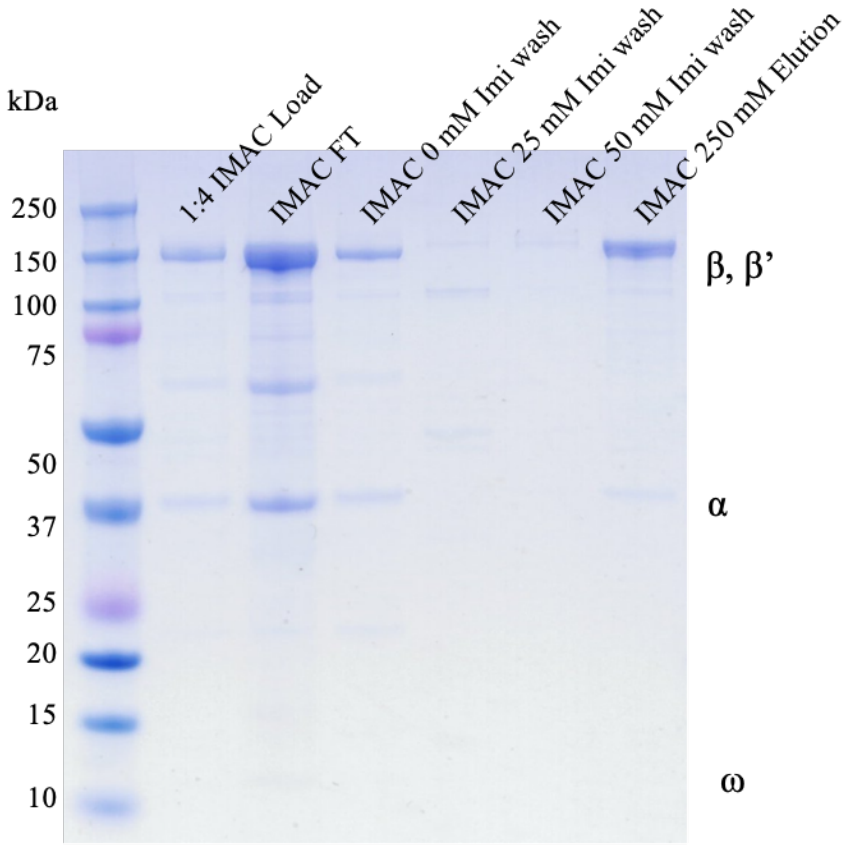


Fig. 3.13.B. Purification of Δβ'CT-E II

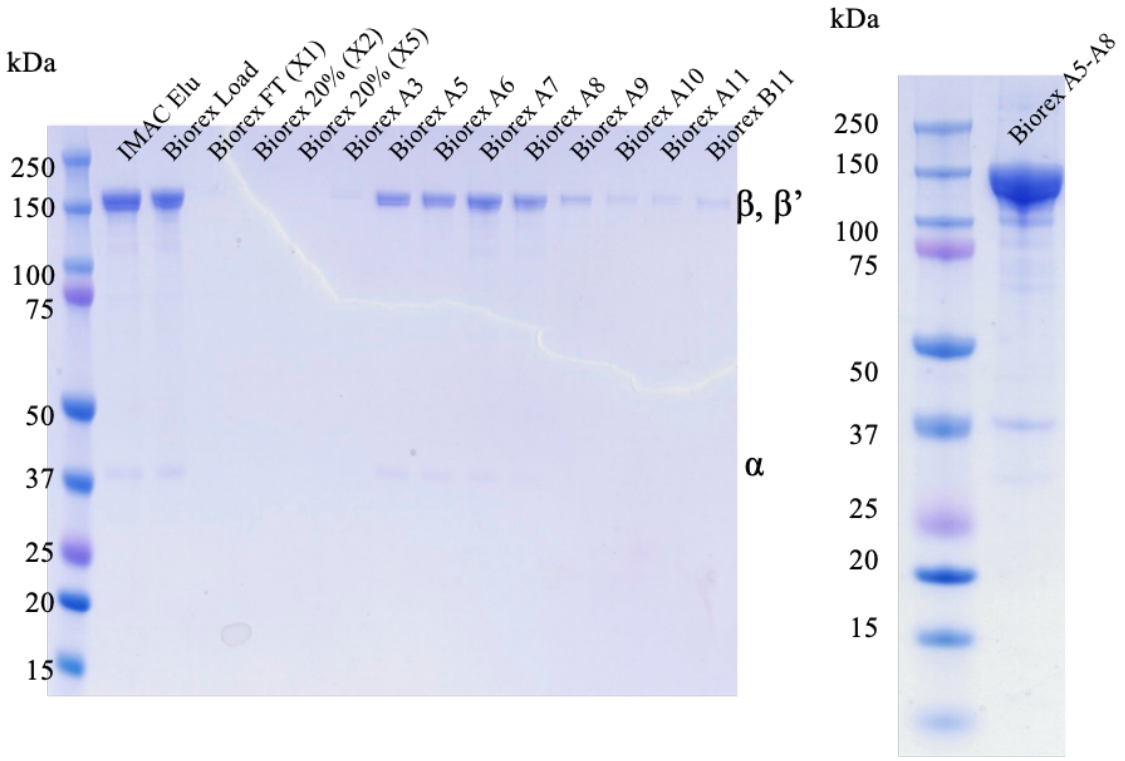


Fig. 3.13.C. Purification of Δβ'CT-E III

Biorex fractions A5-A8 were pooled and concentrated.

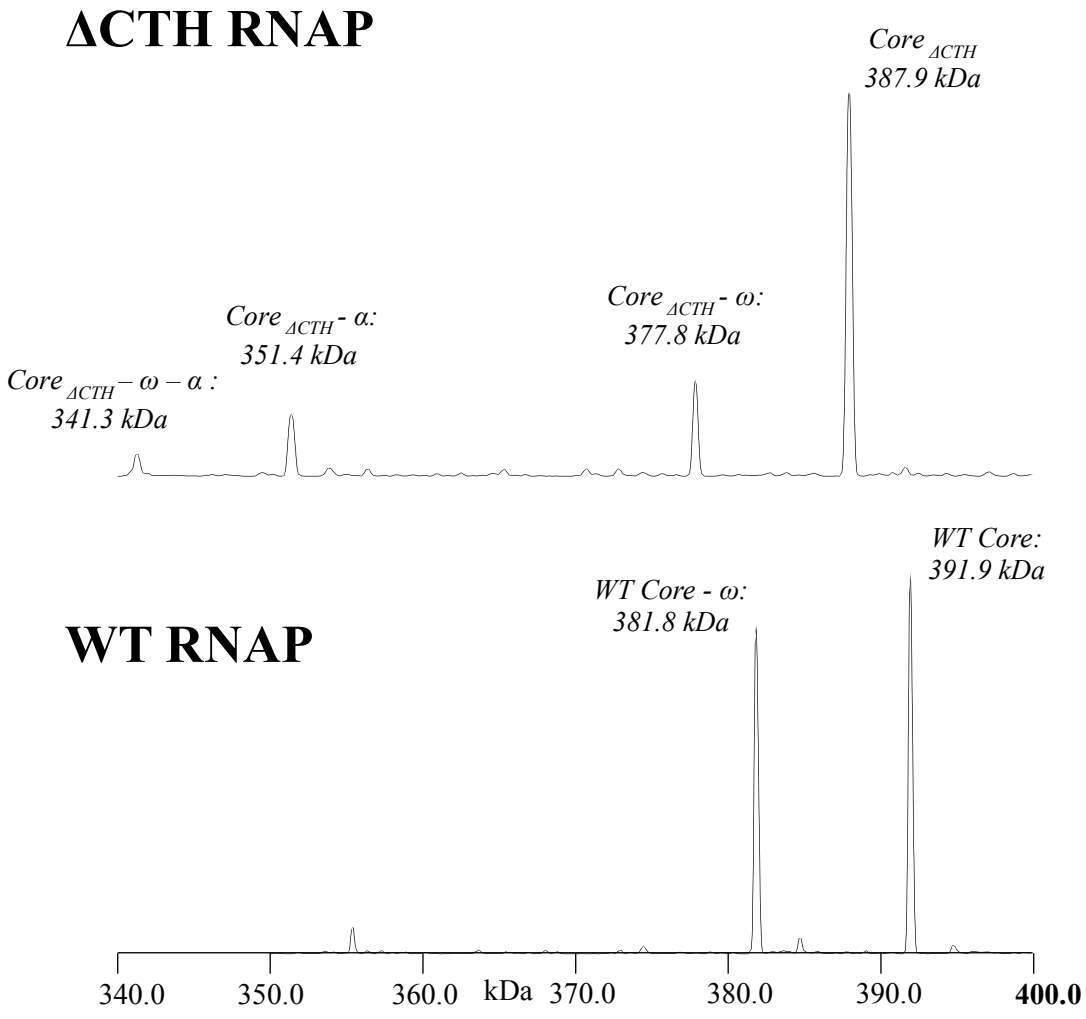


Fig. 3.14. Native mass spectrometry of Δβ'CT-E. Molecular weight of the complex is consistent with the deletion of Δβ'CT.

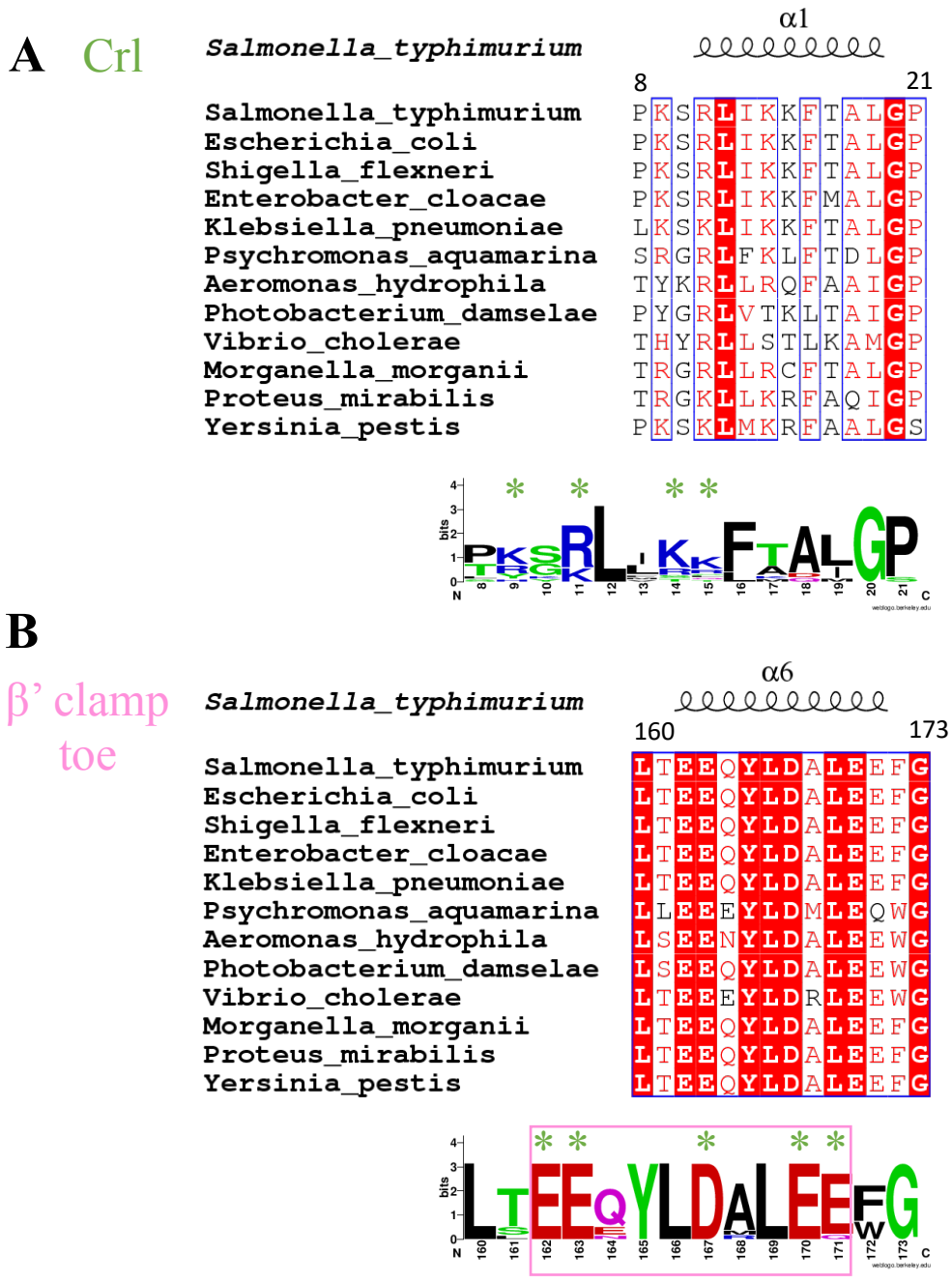


Fig. 3.15 Sequence conservation at the interface of Crl and β'. A. Sequence alignment in the region of Crl that interacts with β'. This alignment is comprised of 12 distinct bacterial species representing 4 orders of γ-proteobacteria. Below the sequences is a logo (Crooks et al. 2004) that shows the consensus of these sequences. Conserved basic amino acids poised to interact with the β' CTH are denoted by green asterisks. **B.** Sequence alignment in the region of β' that interacts with Crl for the same bacteria as in **A**. The magenta box in the logo denotes the central helix within the β' clamp toe.

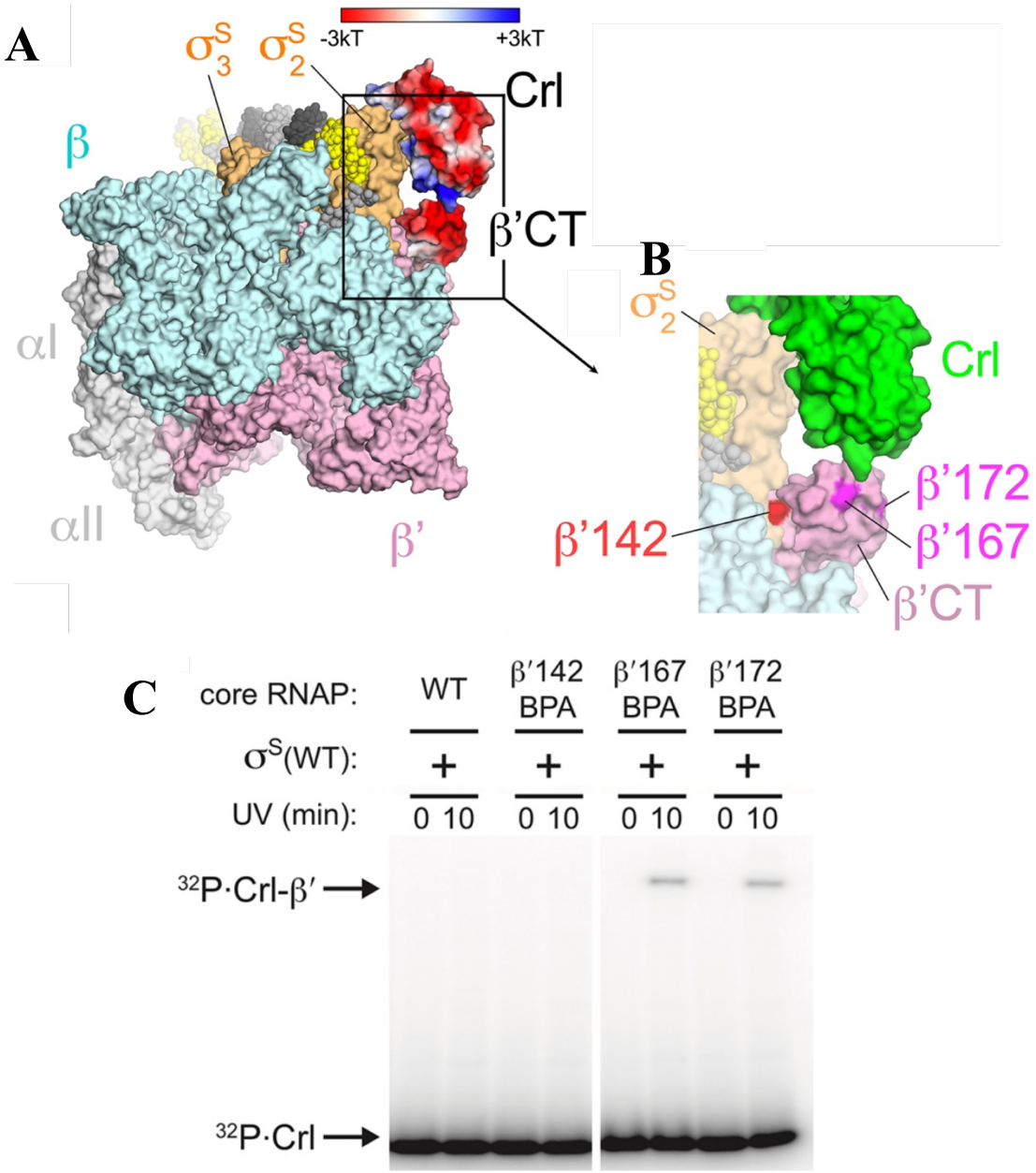


Fig. 3.16. The Crl-β'CT interaction. A. The overall structure of Crl-Eσ^S-*dps*-RPo. Proteins are shown as molecular surfaces with subunits colored as labeled, except Crl and the β'CT are colored according to surface electrostatic potential (red -3kT; blue, +3kT) (Baker NA., et al, 2001) . DNA is shown as CPK spheres and colored according to Fig. 3.5. **B,** Magnified view from **A** focusing on Crl (green) - β'CT (pink) interaction. **C.** SDS-PAGE gel showing effect of UV exposure on RNAP core with β' BPA substitutions incubated with ^{32}P-Crl. Crosslinked complexes and free ^{32}P-Crl are indicated. RNAP β' BPA substitutions at residues 167 and 172 crosslink to Crl [magenta in (**B**)]; BPA substitution at 142 [red in (**B**)] does not crosslink to Crl. From Jaramillo Cartagena, 2019.

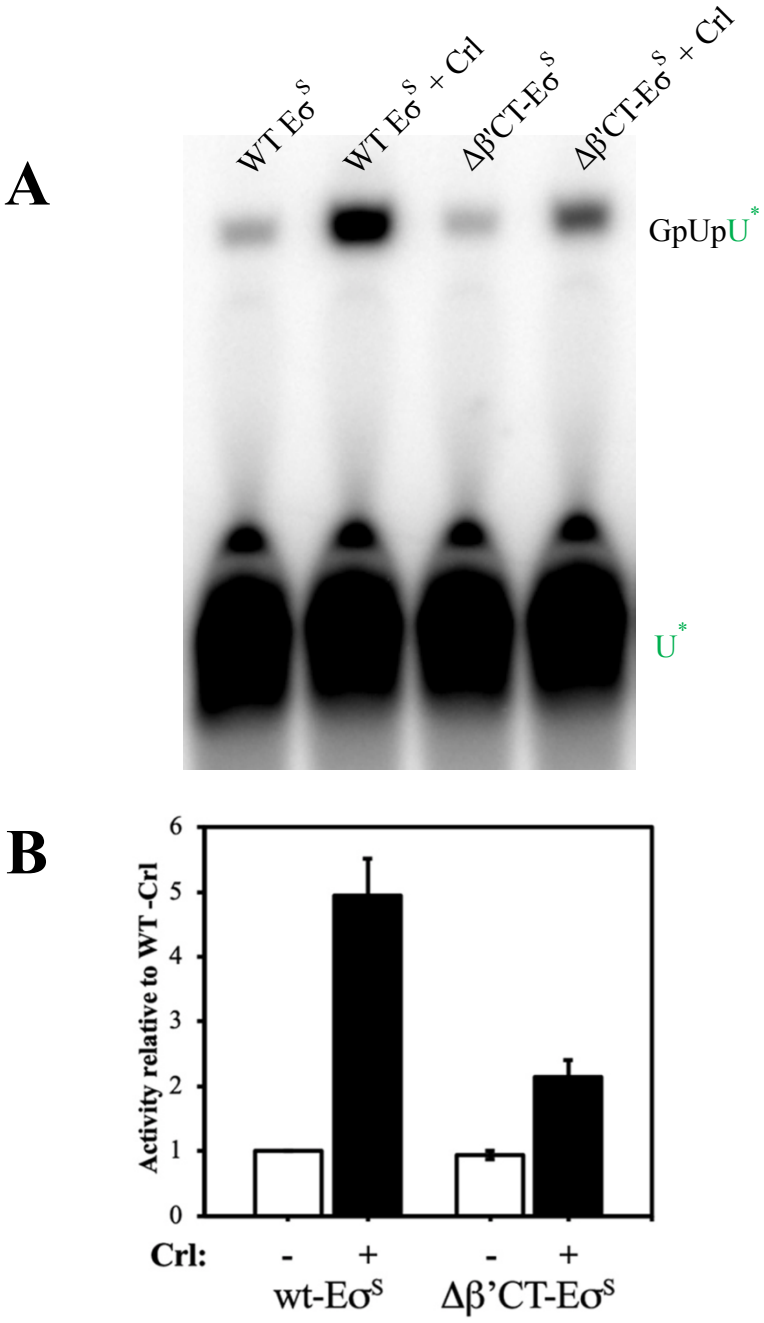


Fig. 3.17. The β'CT is required for full Crl activation. A. abortive transcription assays were conducted as in Fig. 2.3. Gel shows the bands for the GpUpU product for WT $E\sigma^S$ +/- Crl and $\Delta\beta$'CT-$E\sigma^S$ +/- Crl. **B.** Plotted is RNA trinucleotide synthesis without Crl (-, white bars) or with Crl (+, black bars) for wt-$E\sigma^S$ or $\Delta\beta$'CT-$E\sigma^S$. The values are normalized with respect to wt-$E\sigma^S$ (-Crl). The error bars denote standard deviation of three measurements.

Discussion

Our results provide insights into the mechanisms by which Crl promotes $E\sigma^S$ assembly. The Crl-$E\sigma^S$-*dps*-RPo cryo-EM structure is consistent with, and expands upon, previous information about the interactions of Crl with $\sigma^S{}_2$ (Banta et al., 2013, 2014; Cavaliere et al., 2014; Monteil et al., 2010; Robbe-Saule et al., 2006) and includes a previously unidentified interaction with the RNAP β'CT. To our knowledge, this region of E has not been previously described as a binding determinant for any transcription factors and may represent a target for other uncharacterized transcription factors. Deletion of the β'CT had little effect on basal transcription (-Crl) but Crl was unable to fully activate the Δβ'CT-RNAP, suggesting that one mechanism by which Crl activates $E\sigma^S$ transcription is to stabilize the $E\sigma^S$ complex by tethering σ^S and RNAP. This mechanism is consistent with previous studies showing that Crl activation function is most pronounced when σ^S concentrations are low .

The tethering mechanism only accounts for partial Crl activity (**Fig. 3.17**). Our studies do not exclude a post-$E\sigma^S$ assembly role for Crl in transcription activation, such as facilitating promoter melting (RPo formation) or promoter escape. Crl was shown to promote full transcription bubble formation at the *Sty katN* promoter, particularly at 20°C where $E\sigma^S$ without Crl could only form partially melted intermediates (Bougdour et al., 2004). In our structure, Crl does not interact with the promoter DNA. However, the σ^S DPE motif (D135/P136/E137), critical for the Crl-σ^S interaction (**Fig. 3.18**), is part of the conserved region 2.3 of σ^S, comprising a short loop that forms a part of the binding pocket for the non-template strand -11A, the most conserved position of bacterial promoters (Feklistov & Darst, 2011; Shultzaberger et al., 2007). In fact, the -11A base

forms a hydrogen-bond with the α-carbon backbone NH of σ^S D135, the side chain of which interacts with Crl (**Fig. 3.18**). Thus, Crl might stabilize a conformation of the σ^S DPE motif that facilitates transcription bubble formation.

Our results suggest Crl exerts direct transcription activation through contacts with the β'CT. The β'CT was previously shown to interact with the σ^{70}_{NCR}, which antagonized the σ^{70}-β'clamp-helices interaction to enhance promoter escape and reduce early elongation pausing (Leibman & Hochschild, 2007). The roles of Crl in promoter escape and early elongation pausing have not been examined, but the similarity in spatial arrangement between Crl and the σ^{70}_{NCR} with respect to the β'CT point towards possible analogous roles. Furthermore, the conservation of the β'CT in Crl-containing γ-proteobacteria (**Fig. 3.15**) highlights its mechanistic importance and suggests it might be the target of regulation by other transcription factors or play other important roles.

Previous studies determined crystal structures of Eσ^S bound to a promoter fragment (highest resolution of 3.6 Å; 5IPL) containing σ^{70} consensus -10 and -35 elements (Liu et al., 2016). These structures showed the engagement of Eσ^S with the promoter -10 element and downstream part of the promoter, but the upstream promoter -35 element was not bound to σ^S_4 due to exclusion by crystal packing. The protein components of our structure align well with this previous structure (0.96 rmsd over 2,600 Cα's), and σ^S_2 (σ^S residues 53-167) aligns with an rmsd of 0.51 Å over 104 Cα's, indicating that Crl binding does not induce significant conformational changes in σ^S_2 once Eσ^S is formed. Similarly, Crl in the Crl-Eσ^S-*dps*-RPo cryo-EM structure aligns with an rmsd of 0.8 Å over 113 Cα's with the Crl crystal structure [3RPJ; (Banta et al., 2014)].

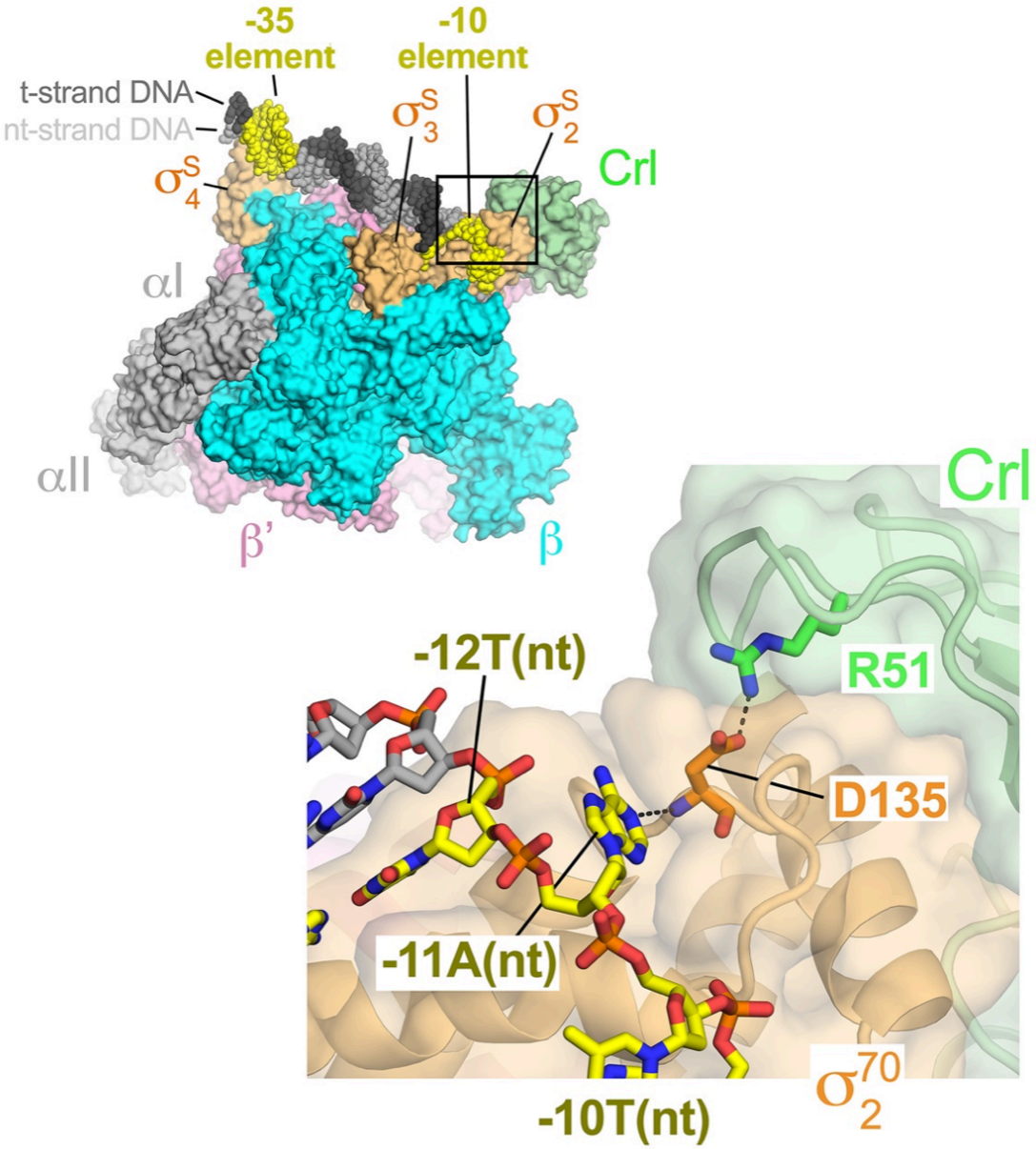


Fig. 3.18. Crl interacts with σ^S residues involved in non-template strand -11A capture. ***Top left.*** The overall structure of Crl-Eσ^S-*dps*-RPo. Proteins are shown as molecular surfaces with subunits colored as labeled. The DNA is shown as CPK spheres and colored according to Fig. 2.5. The boxed region is magnified below. ***Bottom right.*** Crl (light green) and σ^S_2 (light orange) are shown as backbone worms. The promoter DNA is shown in stick format. The side chains of Crl-R51 (green) and σ^S-D135 (orange) are also shown in stick format. The side chain and backbone amide of σ^S-D135 make simultaneous polar interactions with Crl-R51 and -11A(nt), respectively (denoted by dashed gray lines). From Jaramillo Cartagena, 2019.

Our cryo-EM structure contains a *dps* promoter DNA fragment. The *dps* promoter is part of the σ^S regulon (Lacour & Landini, 2004; Lelong et al., 2007). The *dps* promoter is transcribed by both $E\sigma^{70}$ and $E\sigma^S$ *in vitro* but shows a marked preference for $E\sigma^S$ (Grainger et al., 2008). However, our RPo structure does not reveal striking differences in promoter DNA interactions comparing σ^{70} and σ^S, which suggests that differences in promoter preference may not be simply due to σ-promoter interactions in the final RPo, but to differences in the kinetics of RPo formation. Since Crl is induced at low temperatures, it would be interesting to determine if activation by Crl has more mechanistic functional importance for transcription initiation at low temperature where promoter melting by RNAP is inhibited (Bougdour et al., 2004; England et al., 2008).

Transcription activators Crl, GcrA, GrgA, and RbpA interact with σ_2 and/or the NCR of σ factors and represent an emerging paradigm in bacterial transcription regulation (Bao et al., 2012; Hubin et al., 2017; Tabib-Salazar et al., 2013; Wu et al., 2018). Crl plays a unique role as it specifically activates transcription of σ^S-dependent genes, which bacteria express in order to respond to changes in their environment. We propose that Crl be termed a σ-activator, which may represent a class of factors with similar function. We note that mycobacterial RbpA has been shown to tether σ^A to E via the β' Zinc binding domain (Hubin et al., 2017). It remains to be seen whether GcrA or GrgA interact with RNAP to tether the σ factors they regulate.

Materials and Methods

Protein expression and purification

Sty Crl

His_6-ppx-Crl was recombinantly expressed from a pET-21a plasmid transformed into *Eco* BL21 LOBSTR strain, which removes common contaminants in His-tag purifications (Andersen, Leksa, & Schwartz, 2013). LB media (2-liters) was inoculated with cell colonies grown overnight on LB agar plates with 100 μg/ml. Cultures were incubated with shaking at 200 rpm at 37 ºC until OD_{600nm} = 0.4 when the temperature was reduced to 30 ºC and at OD_{600nm} = 0.6 protein expression was induced by adding isopropyl β-D-1-thiogalactopyranoside (IPTG, final concentration 0.5 mM) with further incubation at 30°C for 2 hours.

Cells were harvested and resuspended in 60 mL of 200 mM NaCl, 50 mM Tris-HCl, pH 8.0, 10 mM imidazole, 5% glycerol (v/v). Cells were lysed using a continuous flow French press (Avestin) at 15,000 psi. Cell lysate was centrifuged at 15000 rpm (~12000 g) for 40 minutes. Cell lysis supernatant was loaded into a 5 mL Ni^{2+}-charged HiTrap IMAC column (GE Healthcare). The column was washed with 200 mL of 30 mM imidazole, then proteins were eluted with 300 mM imidazole. PreScission protease (GE Healthcare) was added and the sample was dialyzed (8 kDa molecular weight cutoff dialysis tubing) against 1 L of 200 mM NaCl, 50 mM Tris-HCl, pH 8.0, 5% glycerol (v/v). After overnight incubation, a subtractive IMAC was performed to separate untagged Crl from the cleaved His_6-tag and uncleaved His_6-Crl. The untagged Crl was further purified by size exclusion chromatography (SEC) on a Superdex 200 HiLoad column (GE Healthcare). Fractions containing Crl were pooled, concentrated to 160 μM

by centrifugal filtration (Amicon Ultra), supplemented with glycerol to 20% (v/v), then divided into 40 μL aliquots and stored at -80 ºC.

***Sty* σ^S**

Gibson assembly (Gibson et al., 2008) was used to generate His_{10}-SUMO-σ^S in pET-21a. The Crl purification protocol was used for this construct, except the His_{10}-SUMO-tag was removed using ULP-1. Protein was divided into 40 μL of 200 μM aliquots in the presence of 20% glycerol and stored at -80 ºC.

***Eco* Δα-C-terminal domain RNAP**

In vivo assembled *Eco* RNAP lacking the α C-terminal domains was expressed and purified as previously described (Twist et al., 2011).

Construction of RNAP β' ΔCT mutant

We generated an *Eco* RNAP expression construct with the β'CT deleted and replaced with a Gly-Ser linker **[**Δβ'(Tyr144-Lys179)::(Gly144-Ser145)]. The desired deletion was constructed in three steps. First, a plasmid template that encoded the last 57 amino acids in the sequence of the β and the first 625 amino acids of β' was PCR amplified using a primer in the reverse direction before β' CT, and a forward primer after the β' CT. Second, the resulting linear PCR product was re-circularized using Gibson assembly (Gibson et al., 2008). Third, this product was cut with SbfI and BsmI and appended into a full RNAP expression plasmid, which encodes the rest of the E subunits. This expression plasmid is different than the one used for the cryo-EM studies and leads to the expression of RNAP harboring full-length α-subunit. The plasmid encodes full-

length WT-α, β, β'(Δβ'CT)-ppx-His_{10}, and ω. DNA sequencing and native mass spectrometry confirmed the correct construct was designed and purified. For transcription assays this construct was purified alongside its WT-RNAP parent with full-length β'.

Transcription assays

Abortive initiation assays were adapted from previous protocols (Davis et al., 2015). The reactions were performed at 37°C in a buffer containing 50 mM Tris-HCl, pH 8, 100 mM K-glutamate, 10 mM $MgCl_2$, 10 mM Dithiothreitol (DTT), and 100 μg/mL BSA. Reactions contained final concentrations of 50 nM core RNAP, 100 nM σ^S, 3.2 μM Crl, and 10 nM *dps* promoter DNA (-39 to +42; Fig. S1A).

First, σ^S was incubated with Crl or buffer for 10 minutes at 37°C. Next, σ^S or Crl-σ^S were incubated with WT-RNAP or Δβ'CT-RNAP for 10 minutes at 37°C to form holoenzyme. Lastly, promoter DNA was added to the holoenzymes and incubated for 10 minutes at 37°C to allow RPo formation. Abortive transcription was initiated by adding a mix of GpU dinucleotide primer (250 μM, TriLink), [α-^{32}P]UTP (1.25 μCi), and unlabeled GTP (50 μM). After 20 minutes, reactions were quenched with 2X stop buffer (8 M urea, 89 mM Tris-HCl, 89 mM boric acid, 2 mM EDTA, 0.05% bromophenol blue, 0.05% xylene cyanol) and separated on a 23% (w/v) urea-polyacrylamide gel. Abortive products were visualized by phosphorimaging and quantified using ImageJ (Schneider, Rasband, & Eliceiri, 2012).

Preparation of Crl- σ^S-dps-RPo for Cryo-EM

A 400 uL sample was prepared containing final concentrations of 40 μM RNAP, 80 μM σ^S, 160 μM Crl, and 80 μM promoter DNA (Fig. 1B). First, σ^S was incubated with

Crl for 10 minutes at 37°C. Next, RNAP was added and incubated for 10 minutes at 37°C to form Crl-Eσ^S. Lastly, promoter DNA was added and incubated for 10 minutes at 37°C to allow RPo formation. The sample was injected into a 24 mL Superose 6 Increase column (GE Healthcare Life Sciences, Pittsburgh, PA) pre-equilibrated with 20 mM Tris-HCl, pH 8.0, 150 mM K-Glutamate, 10 mM $MgCl_2$, 5 mM DTT. Fractions containing the complex were pooled and concentrated to 5 mg/mL protein by centrifugal filtration (EMD-Millipore, Darmstadt, Germany). An RNA oligonucleotide (CUCG) was added to a final concentration of 20 μM. CHAPSO (3-([3-cholamidopropyl]dimethylammonio)-2-hydroxy-1-propanesulfonate) was then added to a final concentration of 8 mM (Chen et al., 2018) and the sample was kept at room temperature before grid preparation.

Cryo-EM grid preparation

C-flat holey carbon grids (CF-1.2/1.3-4Au) were glow-discharged for 30 s before the application of 3.5 μl of the sample. After 4 seconds the grids were plunge-frozen in liquid ethane using an FEI Vitrobot Mark IV (FEI, Hillsboro) with 100% chamber humidity at 25°C. Grids were stored in liquid nitrogen.

Cryo-EM data acquisition and processing

Grids were imaged using a 300 keV Titan Krios (FEI) microscope equipped with a K2 Summit direct electron detector. Images were collected using Serial EM (7) in super-resolution counting mode with a super-resolution pixel size of 0.65 Å (22,500 X) and a defocus range of 0.8 to 2.4 μm with 0.2 μm steps. Images were collected at a dose rate of 8 electron/pixel/s. 50 frames were collected over 15 seconds (0.3 s/frame),

yielding a dose of 71 electron/Å^2. Dose-fractionated frames were 2 X 2 binned (giving a pixel size of 1.3 Å).

Data processing was initially carried out in Relion3 using MortionCor2 to bin and align the 50 frames (Zheng et al., 2017; Zivanov et al., 2019) to obtain single dose filtered micrographs. Particles in the micrographs were automatically picked using Gautomatch (K. Zhang, MRC Laboratory of Molecular Biology, Cambridge, UK, http://www.mrc-lmb.cam.ac.uk/kzhang/Gautomatch), which resulted in ~658,000 particles. We built an initial model using a previously published structural model of a σ^S-TIC [PDB 5IPL; (Liu et al., 2016)] with a docked model of Crl from an x-ray crystal structure [PDB 3RPJ; (Banta et al., 2014)]. The particles were classified into three classes using this model (**Fig. 3.6**). One class contained ~53% of the particles and was refined to 4.6 Å, the particles from the other two classes refined to much lower resolution and/or did not contain density for Crl and were discarded. The promising class was furthered classified into 3 additional classes. Two of the classes showed good reconstructions and they were combined to obtain ~292,000 high quality, while particles from the third class (8% of the particles) gave a very low-resolution reconstruction and were discarded. A series of five iterative 3D auto-refinements of the high-quality particles were conducted using particle CTF refinement and Bayesian particle polishing in Relion3. The final polished particles were imported into CryoSPARC2 (Punjani et al., 2017) to perform a homogenous refinement, followed by a local refinement, which yielded the final density map with an overall resolution of 3.26 Å. Local resolution calculations were also conducted using cryoSPARC.

Model building and refinement

An initial structural model was assembled from the core RNAP component of PDB 5VT0 (Chen et al., 2017), *Eco* σ^S from PDB 5IPL (Liu et al., 2016), *Proteus mirabilis* Crl [PDB 3RPJ; (Banta et al., 2014)], and DNA from PDB 4XLN (Bae et al., 2015). The σ^S, Crl, and DNA components were mutated to correspond to *Sty* σ^S, *Sty* Crl, and the *dps* promoter construct used for cryo-EM (**Fig. 3.6**). The model was fit into the cryo-EM density map using UCSF Chimera (Pettersen et al., 2004). Appropriate domains of the complex were rigid-body refined, then subsequently refined with secondary structure and nucleic acid restraints using PHENIX real space refinement (Adams et al., 2010), along with rounds of manual adjustment using COOT (Emsley et al., 2004).

Benzoyl-L-phenylalanine crosslinking

Benzoyl-L-phenylalanine (BPA)-mediated crosslinking was carried out using modifications of a previously described procedure (Banta et al., 2014). *Eco* σ^S and *Eco* N-HMK-Crl were purified and HMK-Crl was ^{32}P-labeled as described (Banta et al., 2014). WT or BPA-substituted *Eco* core RNAPs were overexpressed in BL21DE3 from multi-subunit plasmid pIA299 derivatives and purified as described (Winkelman et al., 2015) (pRLG7814, WT RNAP; pRLG11773, β'142BPA RNAP; RLG11777, β'167BPA RNAP, pRLG13513, β'172BPA RNAP). σ^S and Crl were preincubated for 10 min in buffer containing 40 mM Tris-HCl, pH 8, 100 mM KCl, 10 mM $MgCl_2$, 5 mM DTT and 100 μg/ml BSA. RNAP core enzyme was then added and incubated for an additional 10 min to assemble Eσ^S. Samples were exposed to UV (360 nm) for 0-10 min (as indicated), mixed with 4 μl 4X LDS (Invitrogen) + 25 mM DTT, heated at 85°C for 3 min, loaded

onto a 4-12% Bis-Tris gel (Invitrogen) and run with 1X MES buffer at 120V for 1.5-2 hrs. Gels were transferred to Whatman paper and dried for 1 h at 80°C under vacuum. Crosslinked proteins were visualized using a Typhoon 9400 and ImageQuant software (GE Healthcare).

Conclusions and future directions

The work conducted in this research presents a structural and functional characterization of Crl, an activator of σ^S, an alternative σ factor and the master regulator of stress responses in many proteobacteria (Banta et al., 2013, 2014). The cryo-EM structure of a transcription initiation complex containing σ^S and Crl revealed an important interaction between Crl and the β' subunit of RNA polymerase. Functional transcription studies showed this interaction allows Crl to enhance the transcription activity of $E\sigma^S$.

There are still several structural and biochemical questions that remain to be answered. First, the structure of free σ^S and its conformational equilibrium in the absence of core RNA polymerase (E) are unknown. Furthermore, how Crl alters the structure of free σ^S and its conformational equilibrium is also unknown. It would be interesting to determine if free σ^S adopts a conformation where domains 2 and 4 are close together, and if the distance between these domains is extended in the presence of Crl. If Crl stabilizes extended conformations of σ^S, then it is worth exploring if a Crl/σ^S complex can bind DNA in the absence of E. Other work using bacterial two hybrid assays has determined that σ^S makes σ^S-σ^S interactions *in vivo* (Banta, 2013). It is interesting to determine *in vivo* if another component in the mechanism of transcription activation by Crl is its role in breaking σ^S-σ^S interactions so σ^S can assemble with E. Potassium permanganate DNA-footprinting experiments conducted at -20 °C showed that $E\sigma^S$ forms a partial transcription bubble and that addition of Crl allows the full bubble to be formed (England et al., 2008). This precedent is interesting because Crl is induced at low temperatures (Bougdour et al., 2004) and it would be significant to discover the molecular basis of this

promoter melting enhancement by Crl and to characterize how Crl affects open-promoter complex formation for different promoters.

To date, transcription initiation studies have used a small number of promoters compared to the vast multitude of promoters in the bacterial world. The *in vitro* RNA-Seq method presented in chapter 2 presents a unique opportunity to analyze thousands of promoters in a single experiment where every component is controlled. Thus, this method can be used to identify promoters with interesting characteristics for mechanistic studies, such as potential promoters that could be repressed by Crl; this would present an interesting quest in light of the results of this research where Crl is poised to interact as an activator and not a repressor. For example, motif analysis of "repressed" promoters could define promoter sequences where Crl could help prevent promoter escape. Additionally, the hits from this method could report on promoters with unusual stability and activity that would make them subjects of interesting kinetic and structural studies.

Figure 3.17.B showed that event after the clamp toe of β, the complex was still activated by Crl two-fold. This remaining activation prompts future studies to understand how Crl is still activating transcription. It is conceivable that just the interaction between Crl and $\sigma^S{}_2$ is could have a role in promoter melting. If this is the case, then subjecting Crl to the type of kinetic studies that have been conducted with RpbA and CarD is a worthwhile endeavor (Hubin et al., 2017). Additionally, the role of Group II σ, which lack $\sigma_{1.1}$, in binding promoter DNA on its own has not been extensively explored. It could be possible that σ^S could have a different conformational equilibrium than σ^{70} (Group I, contains $\sigma_{1.1}$), and that difference could lead to σ^S potentially interacting with DNA in the absence of E (Schwartz et al., 2008). The temperature-dependence activation

by Crl has not been extensively studied. Cold temperatures, is one of the stresses that activate Crl and σ^S-dependent genes (Bougdour et al., 2004; Dudin et al., 2013). It is possible that at low temperatures, Crl could have a stronger effect in steps like promoter melting. Systematic functional transcription experiments +/- Crl at different temperatures could provide important insight into how different activators.

Investigating the evolution/emergence of alternative σ factors and activators like Crl could also reveal new insight into transcription regulation and answer important questions: did alternative σ (Groups II, III, and IV) emerged as gene duplication of primary σ (Group I) and later lose structural domains as they evolved to specialize and control small regulons? Or, did the domain architecture of ancestral primary σ factors show low complexity similar to the current Group IV and later evolved more complex architectures? An interesting speculation is that σ^S and Crl could once have been part one single σ factor that closely resembled σ^{70}. In this scenario, Crl would have been the non-conserved region of σ^{70}, and after the gene "divorce" event both Crl and σ^S would have continue to evolve. The essentiality of σ^S in bacteria compared to Crl, explains why σ^S is more conserved and widespread than Crl. It is also conceivable that the bacteria that kept factors like Crl are bacteria that live in complex dynamic environments and need several different strategies to regulate transcription. If this is true, then bacterial with complex life-cycles like living inside and outside hosts as well as bacteria that employ aerial dispersal could tend to have transcription "fine-tuners" like Crl, compare to bacteria whose niche is more stable. For bacteria in very dynamic environments, factors like Crl present an appealing strategy to activate the most important genes as bacteria respond to specific environmental changes. Having discrete transcription "fine-tuners" like Crl could

also present a strategy for bacteria to modularly "shut" these activators as transcription responses to specific stresses is no longer needed. Similarly, "fine-tuners" could also reduce the transcription of certain genes whose expression might initially not be needed, but whose expression might subsequently be important. It is fascinating to explore if Crl could both be an activator and repressor that "fine tunes" transcription.

The work conduced in this thesis also ignites further investigations into the role of the β' clamp toe (β' CT). The extent to conservation in the β' CT helix in proteobacteria is striking as this region is almost 100% conserved in identity. The central helix in the β' CT, which binds Crl, is also structurally present in all the bacterial polymerases whose structures have been solved. The β' CT has also been implicated in promoter escape, which raises several questions: 1) does this region have additional roles? 2) could other bacterial families have evolved other transcription regulators that interact with this region? 3) what happens to cellular physiology and transcription responses if this region is deleted? 4) can this region be used as a target of novel antimicrobial therapeutics?

This research has highlighted the mechanism of Crl, an unconventional bacterial transcription activator, and motivates the discovery and characterization of additional unconventional bacterial transcription activators, which might be a widespread mode of transcription regulation in bacteria.

Appendix A: **Additional Functional studies with Crl mutants designed to disrupt the β'-CT binding surface**

Our Crl-Eσ^S-*dps*-RPo cryo-EM structure revealed an electrostatic interaction between Crl and the clamp toe (CT) of β' (**Fig. 3.16**). Deleting the clamp CT of β' reduced the extent to which Crl activates Eσ^S. In an attempt to complement those results, we also purified Crl with single point mutations at some the sites that interact that β'.

Results

Purification of Crl mutants

Guided by our structure and sequence conservation of Crl homologs, we designed four Crl constructs with mutations at positively conserved amino acids poised to interact with β' (K9E, R11E, K14E, and K15E) (**Fig. A.1**). At first, the parallel purification of these mutants was challenging and often aggregated, which might be indicative instability and the importance of these sites for the folding of Crl. However, performing the protein purification in one day instead of the typical two days yielded high levels of soluble and highly pure proteins (**Fig. A.1C**). Crl mutants were purified following the same protocol as for WT Crl.

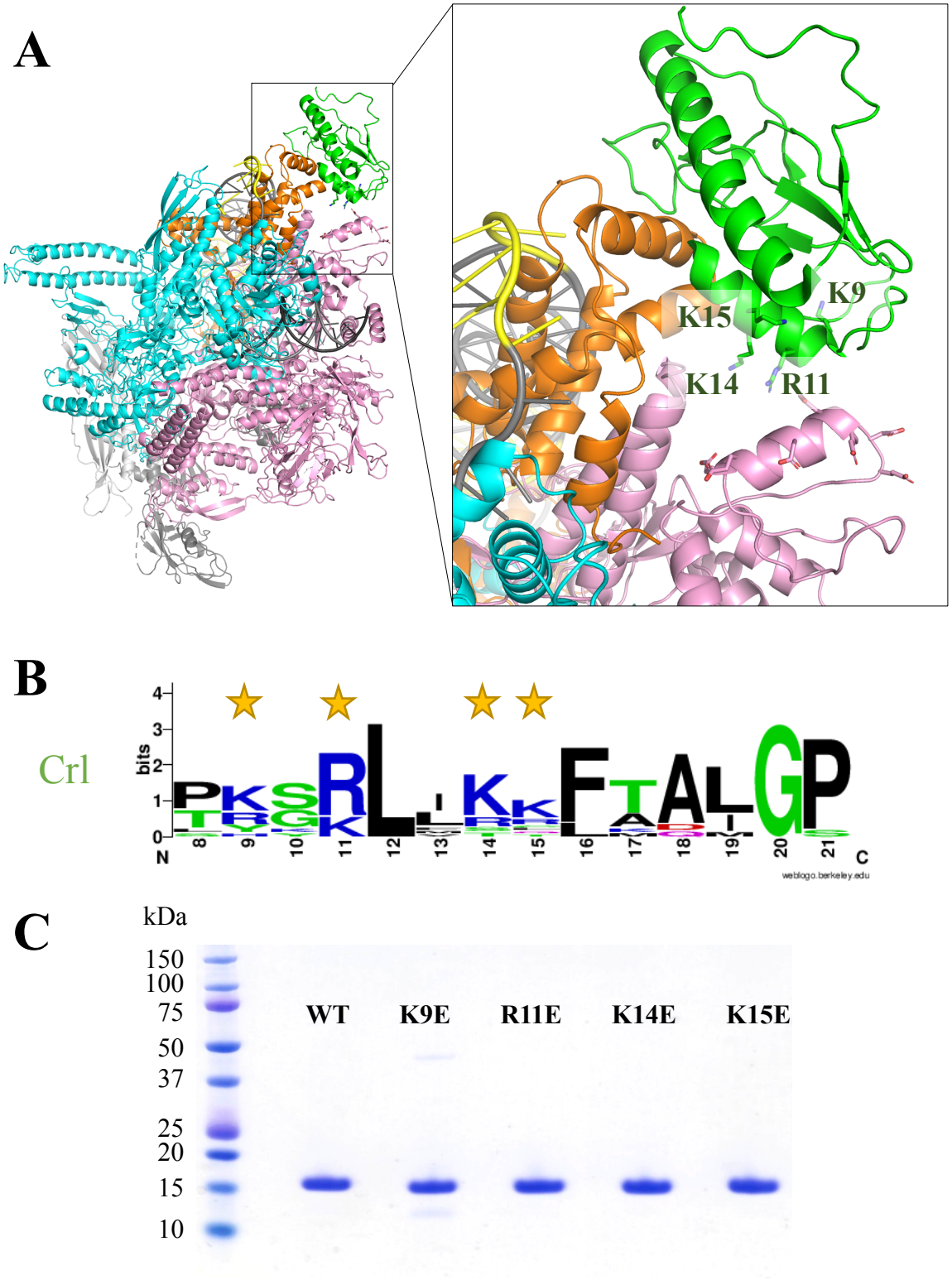


Fig. A.1. Basic amino acids in Crl poised to interact with the β'CT. A. Structure highlighting basic amino in the Crl surface that interacts with the β'CT. **B.** Sequence logo showing the consensus sequence for Crl homologs from 12 distinct bacteria. **C.** SDS PAGE showing purified Crl mutants.

Crl mutants do not activate Eσ^S as strongly as WT Crl

Under the conditions tested, WT Crl activated transcription ~2.5 fold compared to reactions without Crl. The three Crl mutants tested (R11E, K14E, K15E) activated transcription ~1.5 fold (**Fig. A.2**). This result aligns with the idea that these mutants cannot readily interact with the β', but can still activate transcription to a small extent as Crl has at least one additional mechanism that leads to activation of σ^S-dependent genes;

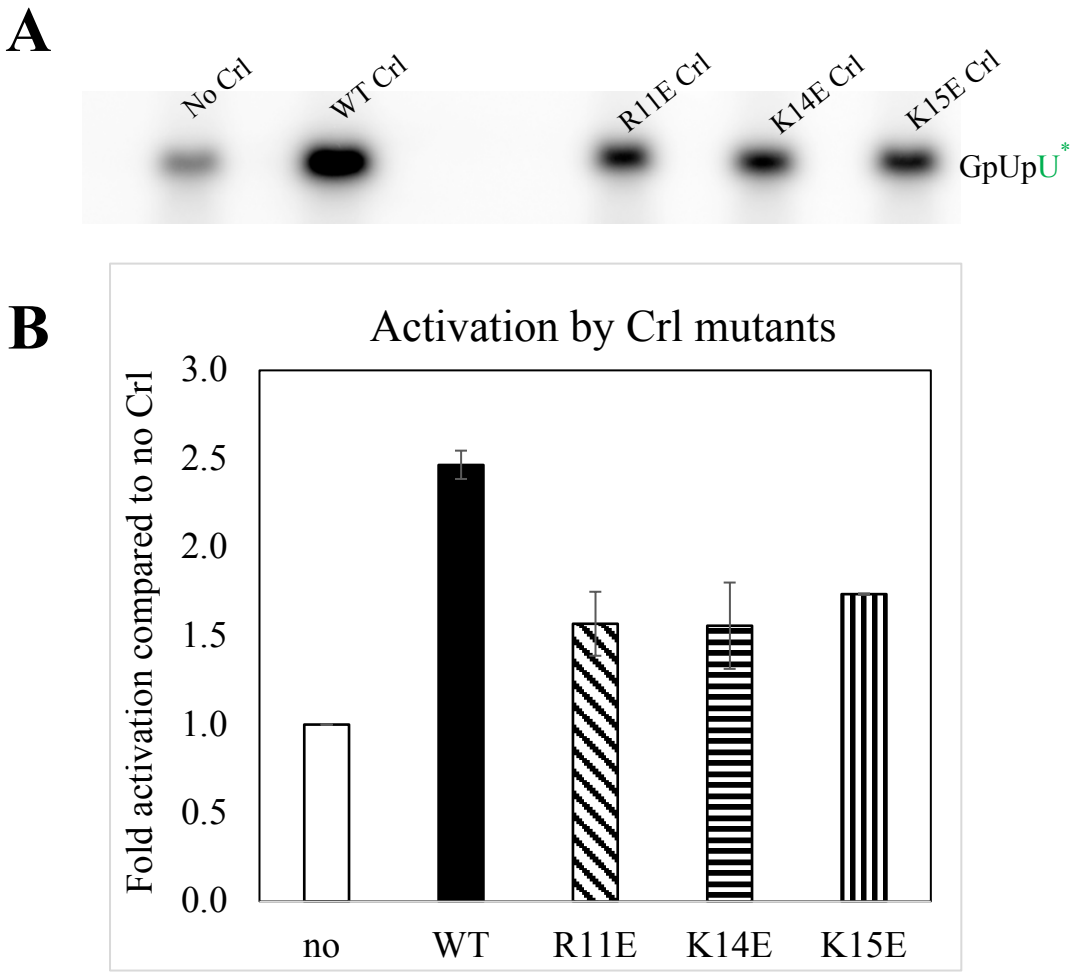


Fig. A.2 Transcription activation by WT and mutant Crl β'. A. A portion of an acrylamide gel showing the abortive transcripts from reactions with different Crl mutant. **B.** Quantitation of triplicate experiments

this could be favoring the equilibrium of σ^S towards conformation states that can bind RNAP more stably.

Crl mutants unexpectedly bind σ^S with weaker affinity

To determine if the decrease in activation was solely the result of disrupting the interaction between Crl and β' (and not between Crl and σ^S), we used microscale thermophoresis (Seidel et al., 2013) to measure the binding affinity of the Crl mutants to σ^S in comparison to wild type Crl. To employ this technique, we introduced a cysteine into $\sigma^S{}_4$ (R311C), which is far away from surface interface between Crl and σ^S (**Fig. A.3A**). We labeled this cysteine with an excitable fluorescent probe and used it in microscale thermophoresis binding experiments. To our surprise, all the Crl mutants bound σ^S with weaker affinity compared to WT Crl (**Table A.1**). This complicated interpretation of the results of the transcription experiments as we could not distinguish if the decrease in activation was due to disrupting the Crl/β' interface, or whether they disrupted the Crl/σ^S interaction. **Fig. A.3A** illustrates that these mutations are not near the Crl/σ^S interface. One possible explanation is that these mutants disturbed the native folding of Crl in comparison to WT. We note that we do not know the conformation $\sigma^S{}_4$ relative to $\sigma^S{}_2$ when Crl is bound and **Fig. A.3A** is a model of Crl with σ^S in the conformation bound to E.

Table A.1 Crl mutants binding affinity to σ^S

	Day 1	Day 2	Day 3		
	K_d (μM)	K_d (μM)	K_d (μM)	Avg K_d (μM)	K_d STD (μM)
WT	4.30	2.91	2.12	3.11	1.10
K9E	8.42	5.02	4.34	5.93	2.19
R11E	6.29	6.72		6.51	0.30
K14E	7.44	2.63		5.04	3.40
K15E	9.40	10.60		10.00	0.85

Conclusion

The binding measurement of WT Crl to σ^S with a K_d of 3 μM is consistent with previously reported values (England et al., 2008). However, the binding affinity of all Crl mutants is weaker than WT. Since Crl strengthens the affinity of σ^S to core RNAP, we could not determine if the decreased of activation in the transcription assays was due to the Crl mutations affecting their binding to β' or σ^S.

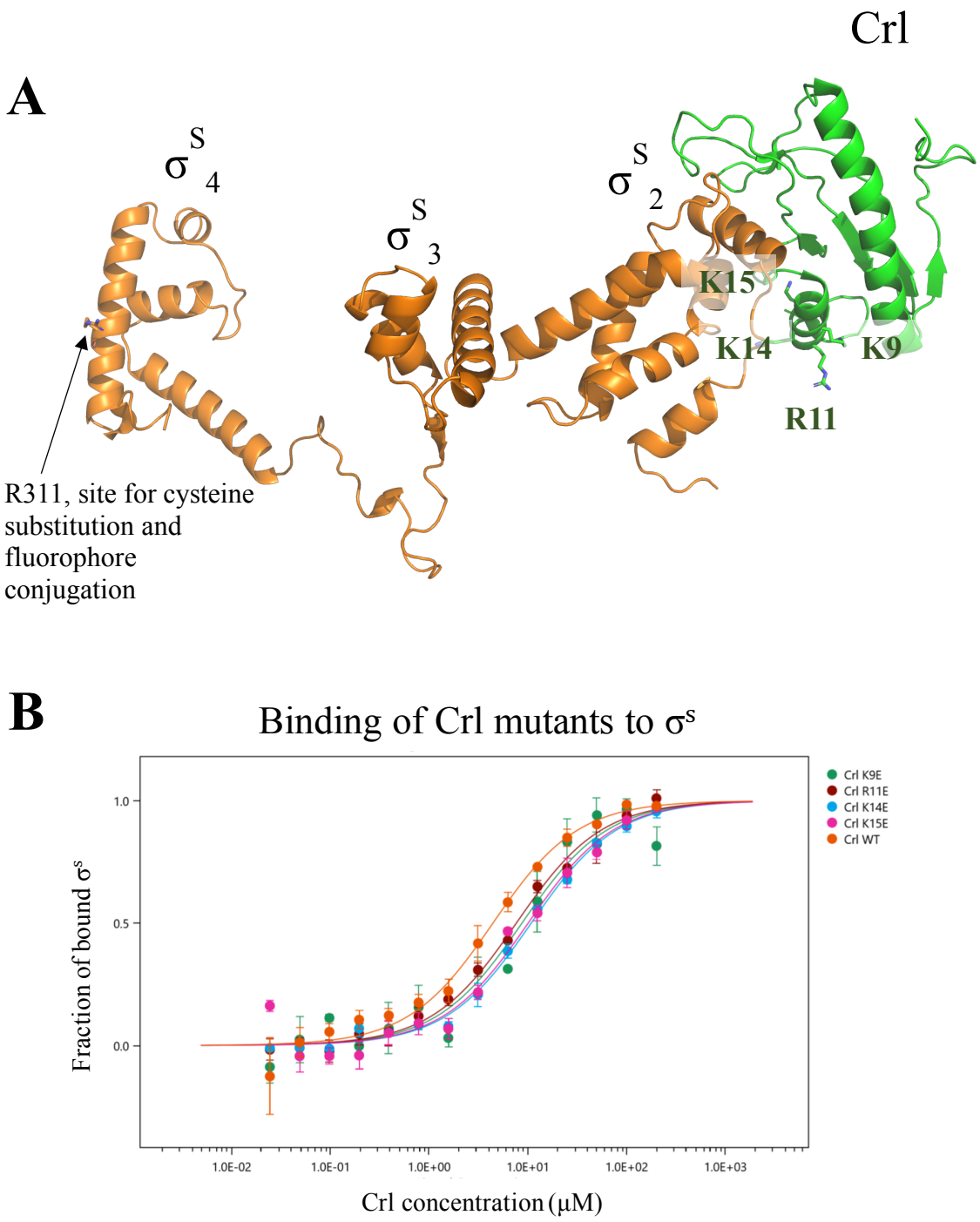


Fig. A.3. Microscale thermophoresis with σ^S and Crl mutants. A. Guided by our structure, we introduced a cysteine site at σ^S R311, which is far away from the Crl/σ^S binding interface, which was subsequently labeled with an Alexa 658 fluorophore. **B.** Binding isotherm for σ^S as a function of Crl (WT or mutants) concentration. Experiments were performed in triplicate for the WT and K9E Crl mutant and duplicate for the R11E, K14E, and K15E mutants.

Materials and methods

In vitro abortive transcription using Crl mutants

An open-promoter holoenzyme complex was formed by pre-incubating RNAP (10 nM), σ^S (20 nM), Crl (2 μM), and DPS promoter DNA (10 nM). Abortive transcription was initiated by adding 250 uM GpU, 50 uM UTP, 1 uCi [γ-32P] UTP. The transcription products were separated by urea PAGE. The band corresponding to the GUU abortive transcription product was quantified by phosphoimagery using the program ImageQuant. The activation of WT and the Crl mutants was compared to that against the products from reactions without Crl.

Binding measurements of Crl mutants to σ^S

Microscale thermophoresis (MST) was used to determine the binding affinity of Crl mutants to σ^S. To conduct these experiments, a R311C mutation was introduced into σ^S to label the protein with an Alexa-647 dye. R311 lies within $\sigma^S{}_4$ and it is significantly distant from $\sigma^S{}_2$, which is the region that interacts with Crl. In MST, the formation of a complex is assessed by monitoring the heat-induced diffusion of the labeled protein as a function of the concentration of its binding partner. In this case, σ^S bound to Crl diffuses faster than unbound σ^S. Experiments were conducted using a fixed concentration of labeled σ^S and a range of Crl concentrations. The average results of triplicate (WT and R9E) or duplicate (R11E, K14E, K15E) were plotted as fraction bound σ^S against the Crl concentration (**Fig. A.3.B**). The measured affinity of WT Crl to σ^S with dissociation constant (K_d) of 3.11 μM in these MST experiments is comparable to the 2.46 μM K_d reported from surface plasmon resonance experiments (England et al., 2008).

Appendix B: Crl interactions with σ^S in the absence of core RNAP that could lead to transcription activation

In surface plasmon resonance experiments, σ^S bound to E with a dissociation constant, K_d, of 68 nM, while the Crl-σ^S complex bound with a K_d of 9 nM (England et al., 2008). The enhanced affinity of Crl-σ^S was largely due to an increased association rate. Crl could be causing this effect by a number of ways, among them: 1) by breaking intermolecular σ^S- σ^S interactions that prevent σ^S in oligomers from binding E as readily as monomers (Banta et al., 2013); 2) Crl could alter the conformational equilibrium of σ^S toward an open conformation, "preparing" σ^S for E binding. To test if these effects could be happening, we used limited proteolysis to investigate the conformation of σ^S with or without Crl to ask whether σ^S would be in an open conformation and therefore more sensitive to proteolysis when bound to Crl.

Results

Crl increases the proteolytic sensitivity of σ^S, but only at high σ^S concentrations

We monitored the disappearance of full-length σ^S as a function of time during treatment with proteinase K, a broad-spectrum serine protease, in the presence and absence of Crl (**Fig. B.1A**). As a negative control, we also monitored proteolysis of σ^{70} ± Crl (**Fig. B.1B**) since Crl does not interact with σ^{70} (Banta et al., 2013). The decay of full-length σ^S (**Fig. B.1A**) or σ^{70} (**Fig. B.1B**) as a function of time was analyzed to calculate the half-life ($t_{1/2}$) of the full-length σ ± Crl. For σ^{70}, the rates of proteolysis were essentially identical ± Crl, with half-lives of 6.4 ± 0.1 min (+Crl) or 6.3 ± 0.1 min (-Crl). By contrast, the half-life of

full-length σ^S was significantly reduced by the presence of Crl; 3.9 ± 0.1 min (+Crl) vs. 6.3 ± 0.2 min (-Crl). Thus, although binding of a partner protein would normally be expected to occlude protease sites and protect the target protein from proteolysis (Campbell & Darst, 2000; Campbell et al., 2003), Crl reduced the lifetime of σ^S. These data support the hypothesis that Crl binding to σ^S disrupts either the putative compact conformation of σ^S or intermolecular σ^S-σ^S interactions, thus preparing σ^S in an open conformation that may facilitate σ^S binding to E. However, to provide a strong signal in SDS-PAGE, these experiments were conducted using 25 μM σ^S, which is significantly higher than the physiologically relevant concentrations of σ^S (0.37 μM, **Table 1.1**, Maeda et al. 2011). Additionally, analytical ultracentrifugation has shown that σ^S can form dimers and higher-order oligomers at such high concentrations (Cavaliere et al., 2018). Thus, we repeated the experiments at lower σ^S concentration (10 μM) (**Fig B.2**). Three different experiments showed that the proteolysis of σ^S is not accelerated at this concentration. Thus, it appears the acceleration observed in the first set of results (**Fig B.1**) might have been due to Crl breaking intermolecular σ^S-σ^S interactions that protected Crl from proteolysis.

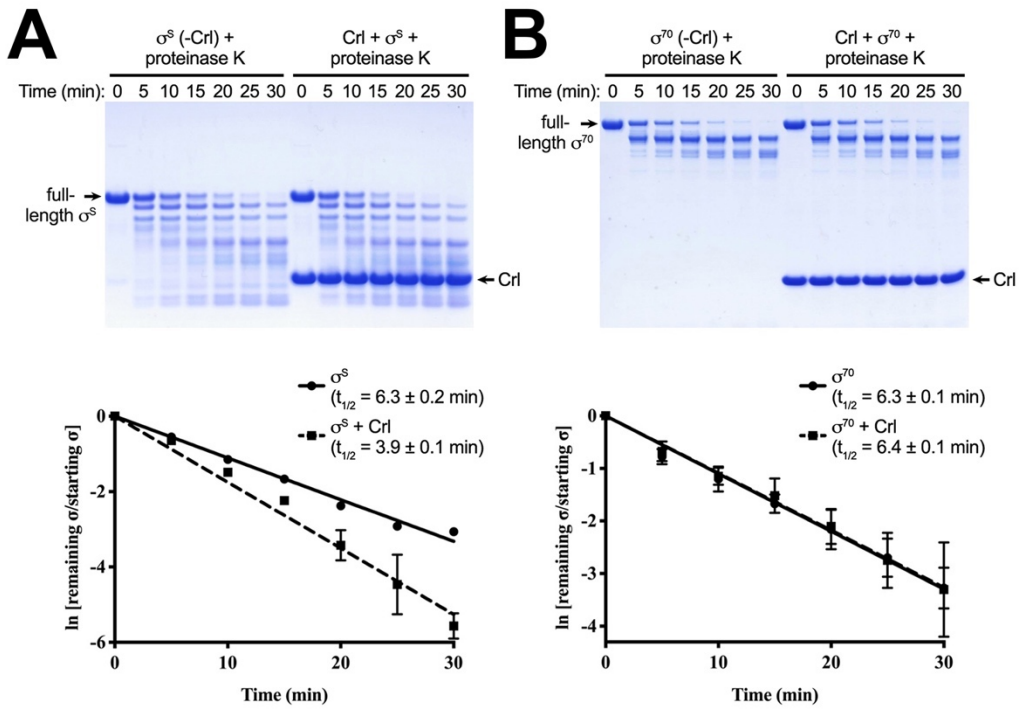


Fig. B.1. Crl sensitizes σ^S but not σ^{70} to proteolysis. ***Top***. Limited proteolysis of (**A**) 25 μM (115 mg) σ^S, or (**B**) 13.6 μM (115 mg) σ^{70} ± 100 μM Crl with 3.3 fM proteinase K for the indicated amounts of time was analyzed by SDS-PAGE and Coomassie blue staining on a 4-12% gradient Bis-Tris acrylamide gel. ***Bottom.*** Bands corresponding to full-length σ^S (**A**) or full-length σ^{70} (**B**) were quantified and plotted on a semi-log scale (normalized by the band of s at time 0). The lines indicate single exponential curve fits from which the half-lives of the full-length s ($t_{1/2}$) were calculated.

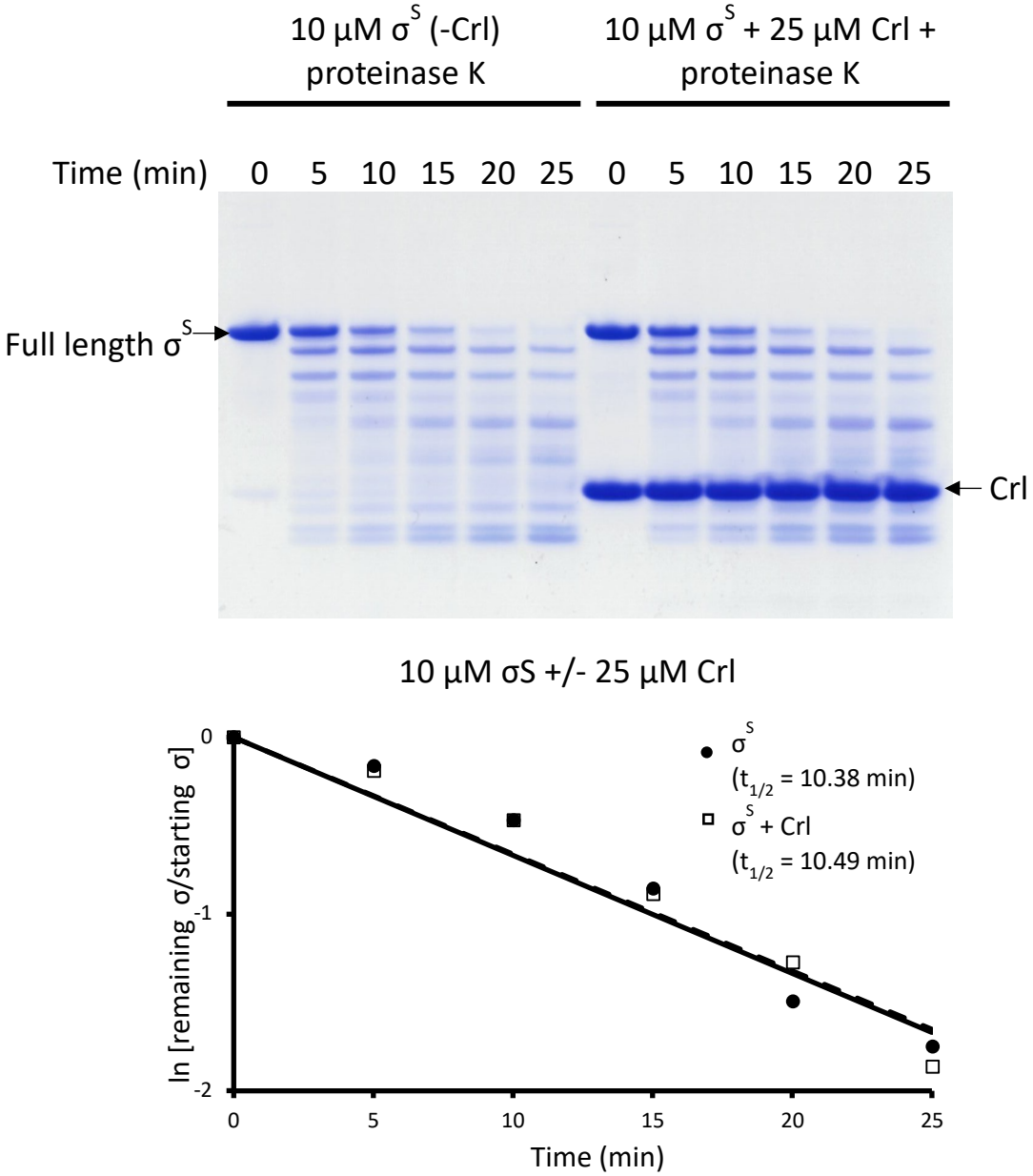


Fig. B.2. Crl does not sensitize σ^S proteolysis at low σ^S concentrations. ***Top***. Limited proteolysis of 10 µM ± 25 mM Crl with 1.33 fM proteinase K for the indicated amounts of time was analyzed by SDS-PAGE and Coomassie blue staining on a 4-12% gradient Bis-Tris acrylamide gel. ***Bottom.*** Bands corresponding to full-length σ^S (were quantified and plotted on a semi-log scale (normalized by the band of σ at time 0). The lines indicate single exponential curve fits from which the half-lives of the full-length s ($t_{1/2}$) were calculated.

Materials and methods

Limited proteolysis time courses

Proteolysis reactions (Fig. 5.1) were conducted in 100 mM K-Glutamate, 50 mM Tris-HCl, pH 8.0, 1 mM EDTA, 5% Glycerol (v/v), and 10 mM β-mercaptoethanol. Reactions of 120 μL contained 25 μM σ^S (total σ^S mass = 115 μg) and either no Crl or 100 μM Crl. Control reactions with σ^{70} contained 13.6 μM σ^{70} (total σ^{70} mass = 115 μg) and either no Crl or 100 μM Crl. A 20 μL aliquots was removed from each sample before adding proteinase K. To each 100 μL reaction volume remaining, 20 μL of a proteinase K solution was added to give a final concentration of 3.33 fM. At 5, 10, 15, 20, 25, and 30 minutes, 12 μL aliquots were removed and immediately incubated at 95°C for 10 minutes. The volume of the samples taken before adding proteinase K were adjusted by adding additional buffer to account to the dilution of the other samples once the proteinase K was added. These samples served as the timepoint for 0 minutes. SDS loading dye was added to each sample, and the samples were analyzed by SDS-PAGE on NuPAGE Novex 4-12% Bis-Tris protein gels (ThermoFisher Scientific). The gels were stained with Coomassie blue and digitized using an HP LaserJet Enterprise 500 MFP M525 scanner. The intensities of the bands for full-length σ^S or σ^{70} were quantified for each timepoint using ImageJ (Schneider et al., 2012). These experiments were conducted in triplicate for each sample.

References

Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., … IUCr. (2010). *PHENIX* : a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D Biological Crystallography*, *66*(2), 213–221. https://doi.org/10.1107/S0907444909052925

Andersen, K. R., Leksa, N. C., & Schwartz, T. U. (2013). Optimized E. coli expression strain LOBSTR eliminates common contaminants from His-tag purification. *Proteins*, *81*(11), 1857–1861. https://doi.org/10.1002/prot.24364

Archambault, J., & Friesen, J. D. (1993). Genetics of eukaryotic RNA polymerases I, II, and III. *Microbiological Reviews*, *57*(3), 703–724. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8246845

Arnqvist, A., Olsén, A., Pfeifer, J., Russell, D. G., & Normark, S. (1992). The Crl protein activates cryptic genes for curli formation and fibronectin binding in Escherichia coli HB101. *Molecular Microbiology*, *6*(17), 2443–2452. https://doi.org/10.1111/j.1365-2958.1992.tb01420.x

Aseev, L. V, Levandovskaya, A. A., Tchufistova, L. S., Scaptsova, N. V, & Boni, I. V. (2008). A new regulatory circuit in ribosomal protein operons: S2-mediated control of the rpsB-tsf expression in vivo. *RNA (New York, N.Y.)*, *14*(9), 1882–1894. https://doi.org/10.1261/rna.1099108

Bae, B., Davis, E., Brown, D., Campbell, E. A., Wigneshweraraj, S., & Darst, S. A. (2013). Phage T7 Gp2 inhibition of Escherichia coli RNA polymerase involves misappropriation of σ70 domain 1.1. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(49), 19772–19777. https://doi.org/10.1073/pnas.1314576110

Bae, B., Feklistov, A., Lass-Napiorkowska, A., Landick, R., Darst, S. A., Adams, P., … Steitz, T. (2015). Structure of a bacterial RNA polymerase holoenzyme open promoter complex. *ELife*, *4*, 213–221. https://doi.org/10.7554/eLife.08504

Baker, N. A., Sept, D., Joseph, S., Holst, M. J., & McCammon, J. A. (2001). Electrostatics of nanosystems: Application to microtubules and the ribosome. *Proceedings of the National Academy of Sciences*, *98*(18), 10037–10041. https://doi.org/10.1073/pnas.181342398

Banta, A. B. (2013). *Molecular interactions between the transcription factor crl and sigma s rna polymerase holoenzyme in escherichia coli*. University of Wisconsin-Madion.

Banta, A. B., Chumanov, R. S., Yuan, A. H., Lin, H., Campbell, E. a, Burgess, R. R., & Gourse, R. L. (2013). Key features of σS required for specific recognition by Crl, a transcription factor promoting assembly of RNA polymerase holoenzyme. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(40), 15955–15960. https://doi.org/10.1073/pnas.1311642110

Banta, A. B., Cuff, M. E., Lin, H., Myers, A. R., Ross, W., Joachimiak, A., & Gourse, R. L. (2014). Structure of the RNA polymerase assembly factor Crl and identification of its interaction surface with sigma S. *Journal of Bacteriology*, *196*(18), 3279–3288. https://doi.org/10.1128/JB.01910-14

Bao, X., Nickels, B. E., & Fan, H. (2012). Chlamydia trachomatis protein GrgA activates transcription by contacting the nonconserved region of 66. *Proceedings of the National Academy of Sciences*, *109*(42), 16870–16875. https://doi.org/10.1073/pnas.1207300109

Barne, K. A., Bown, J. A., Busby, S. J. W., & Minchin, S. D. (1997). Region 2.5 of the Escherichia coli RNA polymerase σ70 subunit is responsible for the recognition of the "extended -10" motif at promoters. *EMBO Journal*, *16*(13), 4034–4040. https://doi.org/10.1093/emboj/16.13.4034

Barnhart, M. M., & Chapman, M. R. (2006). Curli biogenesis and function. *Annual Review of Microbiology*, *60*, 131–147. https://doi.org/10.1146/annurev.micro.60.080805.142106

Battesti, A., Majdalani, N., & Gottesman, S. (2011). The RpoS-Mediated General Stress Response in *Escherichia coli*. *Annual Review of Microbiology*, *65*(1), 189–213. https://doi.org/10.1146/annurev-micro-090110-102946

Bervoets, I., Van Brempt, M., Van Nerom, K., Van Hove, B., Maertens, J., De Mey, M., & Charlier, D. (2018). A sigma factor toolbox for orthogonal gene expression in Escherichia coli. *Nucleic Acids Research*, *46*(4), 2133–2144. https://doi.org/10.1093/nar/gky010

Bidnenko, E., Bidnenko, V., Grylak-Mielnicka, A., & Bardowski, J. (2016). Transcription termination factor Rho: a hub linking diverse physiological processes in bacteria. *Microbiology*, *162*(3), 433–447. https://doi.org/10.1099/mic.0.000244

Bouché, S., Klauck, E., Fischer, D., Lucassen, M., Jung, K., & Hengge-Aronis, R. (1998). Regulation of RssB-dependent proteolysis in Escherichia coli: a role for acetyl phosphate in a response regulator-controlled process. *Molecular Microbiology*, *27*(4), 787–795. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9515704

Bougdour, A., Cunning, C., Baptiste, P. J., Elliott, T., & Gottesman, S. (2008). Multiple pathways for regulation of σS (RpoS) stability in Escherichia coli via the action of multiple anti-adaptors. *Molecular Microbiology*, *68*(2), 298–313. https://doi.org/10.1111/j.1365-2958.2008.06146.x

Bougdour, A., Lelong, C., & Geiselmann, J. (2004). Crl, a Low Temperature-induced Protein in Escherichia coli That Binds Directly to the Stationary Phase σ Subunit of RNA Polymerase. *Journal of Biological Chemistry*, *279*(19), 19540–19550. https://doi.org/10.1074/jbc.M314145200

Buck, M., Gallegos, M., Studholme, D. J., Guo, Y., Gralla, J. D., Gallegos, A., & Gralla, J. a Y. D. (2000). The Bacterial Enhancer-Dependent ς54(ςN) Transcription Factor. *Journal of Bacteriology*, *54*(15), 4129–4136. https://doi.org/10.1128/JB.182.15.4129-4136.2000.Updated

Buckstein, M. H., He, J., & Rubin, H. (2008). Characterization of nucleotide pools as a function of physiological state in Escherichia coli. *Journal of Bacteriology*, *190*(2), 718–726. https://doi.org/10.1128/JB.01020-07

Burgess, R.R. (2001). Sigma Factors. *Encyclopedia of Genetics*, 1831–1834. https://doi.org/10.1006/RWGN.2001.1192

Burgess, Richard R. (1996). Purification of overproduced Escherichia coli RNA polymerase sigma factors by solubilizing inclusion bodies and refolding from Sarkosyl. *Methods in Enzymology*, *273*, 145–149. https://doi.org/10.1016/S0076-6879(96)73014-8

Burgess, Richard R., Travers, A. A., Dunn, J. J., & Bautz, E. K. F. (1969). Factor stimulating transcription by RNA polymerase. *Nature*, *221*(5175), 43–46. https://doi.org/10.1038/221043a0

Camarero, J. A., Shekhtman, A., Campbell, E. A., Chlenov, M., Gruber, T. M., Bryant, D. A., … Muir, T. W. (2002). Autoregulation of a bacterial sigma factor explored by using segmental isotopic labeling and NMR. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(13), 8536–8541. https://doi.org/10.1073/pnas.132033899

Campagne, S., Damberger, F. F., Kaczmarczyk, A., Francez-Charlot, A., Allain, F. H.-T., & Vorholt, J. A. (2012). Structural basis for sigma factor mimicry in the general stress response of Alphaproteobacteria. *Proceedings of the National Academy of Sciences*, *109*(21), E1405–E1414. https://doi.org/10.1073/pnas.1117003109

Campagne, Sébastien, Marsh, M. E., Capitani, G., Vorholt, J. A., & Allain, F. H.-T. (2014). Structural basis for −10 promoter element melting by environmentally induced sigma factors. *Nature Structural & Molecular Biology*, *21*(3), 269–276. https://doi.org/10.1038/nsmb.2777

Campbell, E. A., & Darst, S. A. (2000). The anti-σ factor SpoIIAB forms a 2:1 complex with σF, contacting multiple conserved regions of the σ factor. *Journal of Molecular Biology*, *300*(1), 17–28. https://doi.org/10.1006/JMBI.2000.3838

Campbell, E. A., Muzzin, O., Chlenov, M., Sun, J. L., Olson, C. A., Weinman, O., … Darst, S. A. (2002). Structure of the bacterial RNA polymerase promoter specificity σ subunit. *Molecular Cell*, *9*(3), 527–539. https://doi.org/10.1016/S1097-2765(02)00470-7

Campbell, E. A., Tupy, J. L., Gruber, T. M., Wang, S., Sharp, M. M., Gross, C. A., & Darst, S. A. (2003). Crystal structure of Escherichia coli sigmaE with the cytoplasmic domain of its anti-sigma RseA. *Molecular Cell*, *11*(4), 1067–1078. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12718891

Campbell, E. A., Westblade, L. F., & Darst, S. A. (2008). Regulation of bacterial RNA polymerase σ factor activity: a structural perspective. *Current Opinion in Microbiology*, *11*(2), 121–127. https://doi.org/10.1016/j.mib.2008.02.016

Cao, B., Zhao, Y., Kou, Y., Ni, D., Zhang, X. C., & Huang, Y. (2014). Structure of the nonameric bacterial amyloid secretion channel. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(50), E5439-44.

https://doi.org/10.1073/pnas.1411942111

Cao, M., Kobel, P. A., Morshedi, M. M., Wu, M. F. W., Paddon, C., & Helmann, J. D. (2002). Defining the Bacillus subtilis σW regulon: a comparative analysis of promoter consensus search, run-off transcription/macroarray analysis (ROMA), and transcriptional profiling approaches. *Journal of Molecular Biology*, *316*(3), 443–457. https://doi.org/10.1006/jmbi.2001.5372

Carpousis, A. J., & Gralla, J. D. (1980). Cycling of Ribonucleic Acid Polymerase to Produce Oligonucleotides During Initiation in Vitro at the Lac UV5 Promoter. *Biochemistry*, *19*(14), 3245–3253. https://doi.org/10.1021/bi00555a023

Casadesús, J., & Low, D. (2006). Epigenetic gene regulation in the bacterial world. *Microbiology and Molecular Biology Reviews : MMBR*, *70*(3), 830–856. https://doi.org/10.1128/MMBR.00016-06

Cavaliere, P, Levi-Acobas, F., Mayer, C., Saul, F. A., England, P., Weber, P., … Norel, F. (2014). Structural and functional features of Crl proteins and identification of conserved surface residues required for interaction with the RpoS/sigma(S) subunit of RNA polymerase. *Biochemical Journal*, *463*, 215–224. https://doi.org/Doi 10.1042/Bj20140578

Cavaliere, Paola, Brier, S., Filipenko, P., Sizun, C., Raynal, B., Bonneté, F., … Norel, F. (2018). The stress sigma factor of RNA polymerase RpoS/σ^{S} is a solvent-exposed open molecule in solution. *Biochemical Journal*, *475*(1), 341–354. https://doi.org/10.1042/BCJ20170768

Cavaliere, Paola, Norel, F., & Sizun, C. (2015). (1)H, (13)C and (15)N resonance assignments of σ(S) activating protein Crl from Salmonella enterica serovar Typhimurium. *Biomolecular NMR Assignments*, *9*(2), 397–401. https://doi.org/10.1007/s12104-015-9617-z

Cavaliere, Paola, Sizun, C., Levi-Acobas, F., Nowakowski, M., Monteil, V., Bontems, F., … Norel, F. (2015). Binding interface between the Salmonella σS/RpoS subunit of RNA polymerase and Crl: hints from bacterial species lacking crl. *Scientific Reports*, *5*(July), 13564. https://doi.org/10.1038/srep13564

Ceci, P., Cellai, S., Falvo, E., Rivetti, C., Rossi, G. L., & Chiancone, E. (2004). DNA condensation and self-aggregation of Escherichia coli Dps are coupled phenomena related to the properties of the N-terminus. *Nucleic Acids Research*, *32*(19), 5935–5944. https://doi.org/10.1093/nar/gkh915

Chen, C. Y., Buchmeier, N. A., Libby, S., Fang, F. C., Krause, M., & Guiney, D. G. (1995). Central regulatory role for the RpoS sigma factor in expression of Salmonella dublin plasmid virulence genes. *Journal of Bacteriology*, *177*(18), 5303–5309. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7665519

Chen, J., Noble, A. J., Kang, J. Y., & Darst, S. A. (2018). Eliminating effects of particle adsorption to the air/water interface in single-particle cryo-electron microscopy: Bacterial RNA polymerase and CHAPSO. *BioRxiv*, *1*, 457267. https://doi.org/10.1101/457267

Chen, J., Wassarman, K. M., Feng, S., Leon, K., Feklistov, A., Winkelman, J. T., …

Darst, S. A. (2017). 6S RNA Mimics B-Form DNA to Regulate Escherichia coli RNA Polymerase. *Molecular Cell*, *68*(2), 388-397.e6. https://doi.org/10.1016/j.molcel.2017.09.006

Crooks, G. E., Hon, G., Chandonia, J.-M., & Brenner, S. E. (2004). WebLogo: a sequence logo generator. *Genome Research*, *14*(6), 1188–1190. https://doi.org/10.1101/gr.849004

Daniels, D., Zubert, P., & Losick, R. (1990). Two amino acids in an RNA polymerase, *87*(October), 8075–8079.

Davis, E., Chen, J., Leon, K., Darst, S. A., & Campbell, E. A. (2015). Mycobacterial RNA polymerase forms unstable open promoter complexes that are stabilized by CarD. *Nucleic Acids Research*, *43*(1), 433–445. https://doi.org/10.1093/nar/gku1231

DeHaseth, P. L., & Helmann, J. D. (1995). Open complex formation by Escherichia coli RNA polymerase: the mechanism of polymerase-induced strand separation of double helical DNA. *Molecular Microbiology*, *16*(5), 817–824. https://doi.org/10.1111/j.1365-2958.1995.tb02309.x

Dove, S. L., Darst, S. A., & Hochschild, A. (2003). Region 4 of σ as a target for transcription regulation. *Molecular Microbiology*, *48*(4), 863–874. https://doi.org/10.1046/j.1365-2958.2003.03467.x

Dudin, O., Lacour, S., & Geiselmann, J. (2013). Expression dynamics of RpoS/Crl-dependent genes in Escherichia coli. *Research in Microbiology*, *164*(8), 838–847. https://doi.org/10.1016/j.resmic.2013.07.002

Emsley, P., Cowtan, K., & IUCr. (2004). *Coot* : model-building tools for molecular graphics. *Acta Crystallographica Section D Biological Crystallography*, *60*(12), 2126–2132. https://doi.org/10.1107/S0907444904019158

England, P., Westblade, L. F., Karimova, G., Robbe-Saule, V., Norel, F., & Kolb, A. (2008). Binding of the unorthodox transcription activator, Crl, to the components of the transcription machinery. *Journal of Biological Chemistry*, *283*(48), 33455–33464. https://doi.org/10.1074/jbc.M807380200

Fang, F. C., Libby, S. J., Buchmeier, N. A., Loewen, P. C., Switala, J., Harwood, J., & Guiney, D. G. (1992). The alternative sigma factor katF (rpoS) regulates Salmonella virulence. *Proceedings of the National Academy of Sciences of the United States of America*, *89*(24), 11978–11982. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1465428

Farnham, P. J., & Platt, T. (1981). Rho-independent termination: dyad symmetry in DNA causes RNA polymerase to pause during transcription in vitro. *Nucleic Acids Research*, *9*(3), 563–577. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7012794

Feklistov, A., & Darst, S. A. (2011). Structural Basis for Promoter −10 Element Recognition by the Bacterial RNA Polymerase σ Subunit. *Cell*, *147*(6), 1257–1269. https://doi.org/10.1016/j.cell.2011.10.041

Feklístov, A., Sharon, B. D., Darst, S. A., & Gross, C. A. (2014). Bacterial Sigma

Factors: A Historical, Structural, and Genomic Perspective. *Annual Review of Microbiology*, *68*(1), 357–376. https://doi.org/10.1146/annurev-micro-092412-155737

Flynn, J. M., Neher, S. B., Kim, Y.-I., Sauer, R. T., & Baker, T. A. (2003). Proteomic Discovery of Cellular Substrates of the ClpXP Protease Reveals Five Classes of ClpX-Recognition Signals. *Molecular Cell*, *11*(3), 671–683. https://doi.org/10.1016/S1097-2765(03)00060-1

Gaal, T, Ross, W., Estrem, S. T., Nguyen, L. H., Burgess, R. R., & Gourse, R. L. (2001). Promoter recognition and discrimination by EsigmaS RNA polymerase. *Molecular Microbiology*, *42*(4), 939–954. https://doi.org/2703 [pii]

Gaal, Tamas, Mandel, M. J., Silhavy, T. J., & Gourse, R. L. (2006). Crl facilitates RNA polymerase holoenzyme formation. *Journal of Bacteriology*, *188*(22), 7966–7970. https://doi.org/10.1128/JB.01266-06

Gardella, T., Moyle, H., & Susskind, M. M. (1989). A mutant Escherichia coli σ70 subunit of RNA polymerase with altered promoter specificity. *Journal of Molecular Biology*, *206*(4), 579–590. https://doi.org/10.1016/0022-2836(89)90567-6

Ghosh, S., & Chan, C.-K. K. (2016). Analysis of RNA-Seq Data Using TopHat and Cufflinks. *Methods in Molecular Biology (Clifton, N.J.)*, *1374*, 339–361. https://doi.org/10.1007/978-1-4939-3167-5_18

Gibson, D. G., Benders, G. A., Axelrod, K. C., Zaveri, J., Algire, M. A., Moodie, M., … Hutchison, C. A. (2008). One-step assembly in yeast of 25 overlapping DNA fragments to form a complete synthetic Mycoplasma genitalium genome. *Proceedings of the National Academy of Sciences*, *105*(51), 20404–20409. https://doi.org/10.1073/pnas.0811011106

Goldman, S. R., Ebright, R. H., & Nickels, B. E. (2009). Direct detection of abortive RNA transcripts in vivo. *Science*, *324*(5929), 927–928. https://doi.org/10.1126/science.1169237

Grainger, D. C., Goldberg, M. D., Lee, D. J., & Busby, S. J. W. (2008). Selective repression by Fis and H-NS at the *Escherichia coli dps* promoter. *Molecular Microbiology*, *68*(6), 1366–1377. https://doi.org/10.1111/j.1365-2958.2008.06253.x

Gross, C. A., Chan, C., Dombroski, A., Gruber, T., Sharp, M., Tupy, J., & Young, B. (1998). The functional and regulatory roles of sigma factors in transcription. *Cold Spring Harbor Symposia on Quantitative Biology*, *63*, 141–155. https://doi.org/10.1101/sqb.1998.63.141

Grossman, A. D., Erickson, J. W., & Gross, C. A. (1984). The htpR gene product of E. coli is a sigma factor for heat-shock promoters. *Cell*, *38*(2), 383–390. https://doi.org/10.1016/0092-8674(84)90493-8

Gruber, T. M., & Gross, C. A. (2003). Multiple Sigma Subunits and the Partitioning of Bacterial Transcription Space. *Annual Review of Microbiology*, *57*(1), 441–466. https://doi.org/10.1146/annurev.micro.57.030502.090913

Haakonsen, D. L., Yuan, A. H., & Laub, M. T. (2015). The bacterial cell cycle regulator

GcrA is a σ70 cofactor that drives gene expression from a subset of methylated promoters. *Genes & Development*, *29*(21), 2272–2286. https://doi.org/10.1101/gad.270660.115

Haldenwang, W. G., & Losick, R. (1980). Novel RNA polymerase sigma factor from Bacillus subtilis. *Proceedings of the National Academy of Sciences*, *77*(12), 7000–7004. https://doi.org/10.1073/pnas.77.12.7000

Han, K., Li, Z., Peng, R., Zhu, L., Zhou, T., Wang, L., … Li, Y. (2013). Extraordinary expansion of a Sorangium cellulosum genome from an alkaline milieu. *Scientific Reports*, *3*(1), 2101. https://doi.org/10.1038/srep02101

Haugen, S. P., Berkmen, M. B., Ross, W., Gaal, T., Ward, C., & Gourse, R. L. (2006). rRNA Promoter Regulation by Nonoptimal Binding of σ Region 1.2: An Additional Recognition Element for RNA Polymerase. *Cell*, *125*(6), 1069–1082. https://doi.org/10.1016/j.cell.2006.04.034

Haugen, S. P., Ross, W., & Gourse, R. L. (2008). Advances in bacterial promoter recognition and its control by factors that do not bind DNA. *Nature Reviews Microbiology*, *6*(7), 507–519. https://doi.org/10.1038/nrmicro1912

Hawley, D. K., & McClure, W. R. (1983). Compilation and analysis of *Escherichia coli* promoter DNA sequences. *Nucleic Acids Research*, *11*(8), 2237–2255. https://doi.org/10.1093/nar/11.8.2237

Hecker, M., Pané-Farré, J., & Uwe, V. (2007). SigB-Dependent General Stress Response in *Bacillus subtilis* and Related Gram-Positive Bacteria. *Annual Review of Microbiology*, *61*(1), 215–236. https://doi.org/10.1146/annurev.micro.61.080706.093445

Hein, P. P., Kolb, K. E., Windgassen, T., Bellecourt, M. J., Darst, S. A., Mooney, R. A., & Landick, R. (2014). RNA polymerase pausing and nascent-RNA structure formation are linked through clamp-domain movement. *Nature Structural & Molecular Biology*, *21*(9), 794–802. https://doi.org/10.1038/nsmb.2867

Helmann, J. D. (2002). The extracytoplasmic function (ECF) sigma factors. *Advances in Microbial Physiology*, *46*, 47–110. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12073657

Henderson, K. L., Felth, L. C., Molzahn, C. M., Shkel, I., Wang, S., Chhabra, M., … Record, M. T. (2017). Mechanism of transcription initiation and promoter escape by E. coli RNA polymerase. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(15), E3032–E3040. https://doi.org/10.1073/pnas.1618675114

Hengge-Aronis, R, Lange, R., Henneberg, N., & Fischer, D. (1993). Osmotic regulation of rpoS-dependent genes in Escherichia coli. *Journal of Bacteriology*, *175*(1), 259–265. https://doi.org/10.1128/JB.175.1.259-265.1993

Hengge-Aronis, Regine. (2002a). Recent insights into the general stress response regulatory network in Escherichia coli. *Journal of Molecular Microbiology and Biotechnology*, *4*(3), 341–346. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11931567

Hengge-Aronis, Regine. (2002b). Stationary phase gene regulation: What makes an Escherichia coli promoter ??S-selective? *Current Opinion in Microbiology*, *5*(6), 591–595. https://doi.org/10.1016/S1369-5274(02)00372-7

Herrou, J., Rotskoff, G., Luo, Y., Roux, B., & Crosson, S. (2012). Structural basis of a protein partner switch that regulates the general stress response of -proteobacteria. *Proceedings of the National Academy of Sciences*, *109*(21), E1415–E1423. https://doi.org/10.1073/pnas.1116887109

Hilbert, D. W., & Piggot, P. J. (2004). Compartmentalization of gene expression during Bacillus subtilis spore formation. *Microbiology and Molecular Biology Reviews : MMBR*, *68*(2), 234–262. https://doi.org/10.1128/MMBR.68.2.234-262.2004

Hiratsu, K., Shinagawa, H., & Makino, K. (1995). Mode of promoter recognition by the Escherichia coli RNA polymerase holoenzyme containing the sigmas subunit: identification of the recognition sequence of the fic promoter. *Molecular Microbiology*, *18*(5), 841–850. https://doi.org/10.1111/j.1365-2958.1995.18050841.x

Hubin, E. A., Fay, A., Xu, C., Bean, J. M., Saecker, R. M., Glickman, M. S., … Campbell, E. A. (2017). Structure and function of the mycobacterial transcription initiation complex with the essential regulator RbpA. *ELife*, *6*. https://doi.org/10.7554/eLife.22520

Hubin, E. A., Tabib-Salazar, A., Humphrey, L. J., Flack, J. E., Olinares, P. D. B., Darst, S. A., … Paget, M. S. (2015). Structural, functional, and genetic analyses of the actinobacterial transcription factor RbpA. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(23), 7171–7176. https://doi.org/10.1073/pnas.1504942112

Hunt, T. P., & Magasanik, B. (1985). Transcription of glnA by purified Escherichia coli components: core RNA polymerase and the products of glnF, glnG, and glnL. *Proceedings of the National Academy of Sciences*, *82*(24), 8453–8457. https://doi.org/10.1073/pnas.82.24.8453

Ishihama, A. (1981). Subunit of assembly of Escherichia coli RNA polymerase. *Advances in Biophysics*, *14*, 1–35. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7015808

Ishihama, A. (1993). Protein-protein communication within the transcription apparatus. *Journal of Bacteriology*, *175*(9), 2483–2489. https://doi.org/10.1128/JB.175.9.2483-2489.1993

Ivanova, A., Renshaw, M., Guntaka, R. V, & Eisenstark, A. (1992). DNA base sequence variability in katF (putative sigma factor) gene of Escherichia coli. *Nucleic Acids Research*, *20*(20), 5479–5480. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1437569

Jain, D., Nickels, B. E., Sun, L., Hochschild, A., & Darst, S. A. (2004). Structure of a Ternary Transcription Activation Complex. *Molecular Cell*, *13*(1), 45–53. https://doi.org/10.1016/S1097-2765(03)00483-0

Jishage, M., & Ishihama, A. (1995). *Regulation of RNA Polymerase Sigma Subunit*

Synthesis in Escherichia coli: Intracellular Levels of 70 and 38 Downloaded from. *JOURNAL OF BACTERIOLOGY* (Vol. 177).

Jones, C. H., & Moran, C. P. (1992). Mutant σ factor blocks transition between promoter binding and initiation of transcription. *Proceedings of the National Academy of Sciences of the United States of America*, *89*(5), 1958–1962. https://doi.org/10.1073/pnas.89.5.1958

Ju, X., Li, D., & Liu, S. (2019). Full-length RNA profiling reveals pervasive bidirectional transcription terminators in bacteria. *Nature Microbiology*, 1–12. https://doi.org/10.1038/s41564-019-0500-z

Kahramanoglou, C., Prieto, A. I., Khedkar, S., Haase, B., Gupta, A., Benes, V., … Seshasayee, A. S. N. (2012). Genomics of DNA cytosine methylation in Escherichia coli reveals its role in stationary phase transcription. *Nature Communications*, *3*(1), 886. https://doi.org/10.1038/ncomms1878

Kang, J. Y., Mishanina, T. V., Bellecourt, M. J., Mooney, R. A., Darst, S. A., & Landick, R. (2018). RNA Polymerase Accommodates a Pause RNA Hairpin by Global Conformational Rearrangements that Prolong Pausing. *Molecular Cell*, *69*(5), 802-815.e1. https://doi.org/10.1016/J.MOLCEL.2018.01.018

Kang, J. Y., Mishanina, T. V., Landick, R., & Darst, S. A. (2019). Mechanisms of Transcriptional Pausing in Bacteria. *Journal of Molecular Biology*. https://doi.org/10.1016/J.JMB.2019.07.017

Kenney, T. J., Moran, C. P., & Jr. (1991). Genetic evidence for interaction of sigma A with two promoters in Bacillus subtilis. *Journal of Bacteriology*, *173*(11), 3282–3290. https://doi.org/10.1128/jb.173.11.3282-3290.1991

Kenney, T. J., York, K., Youngman, P., & Moran, C. P. (1989). Genetic evidence that RNA polymerase associated with σ(A) factor uses a sporulation-specific promoter in Bacillus subtilis. *Proceedings of the National Academy of Sciences of the United States of America*, *86*(23), 9109–9113. https://doi.org/10.1073/pnas.86.23.9109

Klauck, E., Lingnau, M., & Hengge-Aronis, R. (2001). Role of the response regulator RssB in sigma recognition and initiation of sigma proteolysis in Escherichia coli. *Molecular Microbiology*, *40*(6), 1381–1390. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11442836

Koo, B.-M., Rhodius, V. A., Campbell, E. A., & Gross, C. A. (2009). Mutational analysis of *Escherichia coli* σ^{28} and its target promoters reveals recognition of a composite −10 region, comprised of an 'extended −10' motif and a core −10 element. *Molecular Microbiology*, *72*(4), 830–843. https://doi.org/10.1111/j.1365-2958.2009.06691.x

Korzheva, N., Mustaev, A., Kozlov, M., Malhotra, A., Nikiforov, V., Goldfarb, A., & Darst, S. A. (2000). A structural model of transcription elongation. *Science*, *289*(5479), 619–625. https://doi.org/10.1126/science.289.5479.619

Krissinel, E., & Henrick, K. (2007). Inference of Macromolecular Assemblies from Crystalline State. *Journal of Molecular Biology*, *372*(3), 774–797. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/17681537

Kuznedelov, K., Minakhin, L., Niedziela-Majka, A., Dove, S. L., Rogulja, D., Nickels, B. E., … Severinov, K. (2002). A role for interaction of the RNA polymerase flap domain with the σ subunit in promoter recognition. *Science*, *295*(5556), 855–857. https://doi.org/10.1126/science.1066303

Lacour, S., Kolb, A., & Landini, P. (2003). Nucleotides from -16 to -12 determine specific promoter recognition by bacterial sigmaS-RNA polymerase. *The Journal of Biological Chemistry*, *278*(39), 37160–37168. https://doi.org/10.1074/jbc.M305281200

Lacour, S., & Landini, P. (2004). S -Dependent Gene Expression at the Onset of Stationary Phase in Escherichia coli : Function of S -Dependent Genes and Identification of Their Promoter Sequences. *Society*, *186*(21), 7186–7195. https://doi.org/10.1128/JB.186.21.7186

Landini, P., Egli, T., Wolf, J., & Lacour, S. (2014). sigmaS, a major player in the response to environmental stresses in *E scherichia coli* : role, regulation and mechanisms of promoter recognition. *Environmental Microbiology Reports*, *6*(1), 1–13. https://doi.org/10.1111/1758-2229.12112

Lane, W. J., & Darst, S. a. (2010). Molecular evolution of multisubunit RNA polymerases: structural analysis. *Journal of Molecular Biology*, *395*(4), 686–704. https://doi.org/10.1016/j.jmb.2009.10.063

Lange, R., & Hengge-Aronis, R. (1991). Identification of a central regulator of stationary-phase gene expression in Escherichia coli. *Molecular Microbiology*, *5*(1), 49–59. https://doi.org/10.1111/j.1365-2958.1991.tb01825.x

Lange, R., & Hengge-Aronis, R. (1994). The cellular concentration of the sigma S subunit of RNA polymerase in Escherichia coli is controlled at the levels of transcription, translation, and protein stability. *Genes & Development*, *8*(13), 1600–1612. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7525405

Lease, R. A., & Belfort, M. (2000). A trans-acting RNA as a control switch in Escherichia coli: DsrA modulates function by forming alternative structures. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(18), 9919–9924. https://doi.org/10.1073/pnas.170281497

Lee, D. J., Minchin, S. D., & Busby, S. J. W. (2012). Activating Transcription in Bacteria. *Annual Review of Microbiology*, *66*(1), 125–152. https://doi.org/10.1146/annurev-micro-092611-150012

Leibman, M., & Hochschild, A. (2007). A sigma-core interaction of the RNA polymerase holoenzyme that enhances promoter escape. *The EMBO Journal*, *26*(6), 1579–1590. https://doi.org/10.1038/sj.emboj.7601612

Lelong, C., Aguiluz, K., Luche, S., Kuhn, L., Garin, J., Rabilloud, T., & Geiselmann, J. (2007). The Crl-RpoS regulon of Escherichia coli. *Molecular & Cellular Proteomics : MCP*, *6*(4), 648–659. https://doi.org/10.1074/mcp.M600191-MCP200

Lelong, C., Rolland, M., Louwagie, M., Garin, J., & Geiselmann, J. (2007). Mutual regulation of Crl and Fur in Escherichia coli W3110. *Molecular & Cellular Proteomics : MCP*, *6*(4), 660–668. https://doi.org/10.1074/mcp.M600192-MCP200

Lin, W., Mandal, S., Degen, D., Cho, M. S., Feng, Y., Das, K., & Ebright, R. H. (2019). Structural basis of ECF-σ-factor-dependent transcription initiation. *Nature Communications*, *10*(1), 710. https://doi.org/10.1038/s41467-019-08443-3

Liu, B., Zuo, Y., & Steitz, T. A. (2016). Structures of *E . coli* σ S -transcription initiation complexes provide new insights into polymerase mechanism. *Proceedings of the National Academy of Sciences*, 201520555. https://doi.org/10.1073/pnas.1520555113

Loewen, P. C., & Triggs, B. L. (1984). Genetic mapping of katF, a locus that with katE affects the synthesis of a second catalase species in Escherichia coli. *Journal of Bacteriology*, *160*(2), 668–675. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6094482

Lonetto, M. A., Brown, K. L., Rudd, K. E., & Buttner, M. J. (1994). Analysis of the Streptomyces coelicolor sigE gene reveals the existence of a subfamily of eubacterial RNA polymerase sigma factors involved in the regulation of extracytoplasmic functions. *Proceedings of the National Academy of Sciences*, *91*(16), 7573–7577. https://doi.org/10.1073/pnas.91.16.7573

Lonetto, M., Gribskov, M., & Gross, C. A. (1992). The σ70 family: Sequence conservation and evolutionary relationships. *Journal of Bacteriology*, *174*(12), 3843–3849. https://doi.org/10.1128/jb.174.12.3843-3849.1992

Low, D. A., Weyand, N. J., & Mahan, M. J. (2001). Roles of DNA adenine methylation in regulating bacterial gene expression and virulence. *Infection and Immunity*, *69*(12), 7197–7204. https://doi.org/10.1128/IAI.69.12.7197-7204.2001

MacIąg, A., Peano, C., Pietrelli, A., Egli, T., De Bellis, G., Landini, P., … Landini, P. (2011). In vitro transcription profiling of the σ s subunit of bacterial RNA polymerase: Re-definition of the σ s regulon and identification of σ s-specific promoter sequence elements. *Nucleic Acids Research*, *39*(13), 5338–5355. https://doi.org/10.1093/nar/gkr129

MacLellan, S. R., Eiamphungporn, W., & Helmann, J. D. (2009). ROMA: An in vitro approach to defining target genes for transcription regulators. *Methods*, *47*(1), 73–77. https://doi.org/10.1016/j.ymeth.2008.10.009

Maeda, H., Fujita, N., & Ishihama, a. (2000). Competition among seven Escherichia coli sigma subunits: relative binding affinities to the core RNA polymerase. *Nucleic Acids Research*, *28*(18), 3497–3503.

Maillard, A. P., Girard, E., Ziani, W., Petit-Härtlein, I., Kahn, R., & Covès, J. (2014). The Crystal Structure of the Anti-σ Factor CnrY in Complex with the σ Factor CnrH Shows a New Structural Class of Anti-σ Factors Targeting Extracytoplasmic Function σ Factors. *Journal of Molecular Biology*, *426*(12), 2313–2327. https://doi.org/10.1016/j.jmb.2014.04.003

Majdalani, N., Cunning, C., Sledjeski, D., Elliott, T., & Gottesman, S. (1998). DsrA RNA regulates translation of RpoS message by an anti-antisense mechanism, independent of its action as an antisilencer of transcription. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(21), 12462–12467.

https://doi.org/10.1073/pnas.95.21.12462

Malakhov, M. P., Mattern, M. R., Malakhova, O. a., Drinker, M., Weeks, S. D., & Butt, T. R. (2004). SUMO fusions and SUMO-specific protease for efficient expression and purification of proteins. *Journal of Structural and Functional Genomics*, *5*(1–2), 75–86. https://doi.org/10.1023/B:JSFG.0000029237.70316.52

Mathew, R., & Chatterji, D. (2006). The evolving story of the omega subunit of bacterial RNA polymerase. *Trends in Microbiology*, *14*(10), 450–455. https://doi.org/10.1016/j.tim.2006.08.002

McClure, W. R., Cech, C. L., & Johnston, D. E. (1978). A steady state assay for the RNA polymerase initiation reaction. *Journal of Biological Chemistry*, *253*(24), 8941–8948.

Mekler, V., Kortkhonjia, E., Mukhopadhyay, J., Knight, J., Revyakin, A., Kapanidis, A. N., … Ebright, R. H. (2002). Structural organization of bacterial RNA polymerase holoenzyme and the RNA polymerase-promoter open complex. *Cell*, *108*(5), 599–614. https://doi.org/10.1016/S0092-8674(02)00667-0

Mika, F., & Hengge, R. (2005). A two-component phosphotransfer network involving ArcB, ArcA, and RssB coordinates synthesis and proteolysis of S (RpoS) in E. coli. *Genes & Development*, *19*(22), 2770–2781. https://doi.org/10.1101/gad.353705

Minakhin, L., Bhagat, S., Brunning, A., Campbell, E. A., Darst, S. A., Ebright, R. H., & Severinov, K. (2001). Bacterial RNA polymerase subunit ω and eukaryotic polymerase subunit RPB6 are sequence, structural, and functional homologs and promote RNA polymerase assembly. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(3), 892–897. https://doi.org/10.1073/pnas.98.3.892

Monteil, V., Kolb, A., D'Alayer, J., Beguin, P., & Norel, F. (2010a). Identification of conserved amino acid residues of the Salmonella ??S chaperone Crl involved in Crl-??S interactions. *Journal of Bacteriology*, *192*(4), 1075–1087. https://doi.org/10.1128/JB.01197-09

Monteil, V., Kolb, A., D'Alayer, J., Beguin, P., & Norel, F. (2010b). Identification of conserved amino acid residues of the Salmonella sigmaS chaperone Crl involved in Crl-sigmaS interactions. *Journal of Bacteriology*, *192*(4), 1075–1087. https://doi.org/10.1128/JB.01197-09

Monteil, V., Kolb, A., Mayer, C., Hoos, S., England, P., & Norel, F. (2010). Crl binds to domain 2 of σSand confers a competitive advantage on a natural rpoS mutant of Salmonella enterica serovar typhi. *Journal of Bacteriology*, *192*(24), 6401–6410. https://doi.org/10.1128/JB.00801-10

Mooney, R. A., Darst, S. A., & Landick, R. (2005). Sigma and RNA polymerase: An on-again, off-again relationship? *Molecular Cell*, *20*(3), 335–345. https://doi.org/10.1016/j.molcel.2005.10.015

Morichaud, Z., Chaloin, L., & Brodolin, K. (2016). Regions 1.2 and 3.2 of the RNA Polymerase σ Subunit Promote DNA Melting and Attenuate Action of the Antibiotic Lipiarmycin. *Journal of Molecular Biology*, *428*(2), 463–476.

https://doi.org/10.1016/j.jmb.2015.12.017

Mossessova, E., & Lima, C. D. (2000). Ulp1-SUMO crystal structure and genetic analysis reveal conserved interactions and a regulatory element essential for cell growth in yeast. *Molecular Cell*, *5*(5), 865–876. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10882122

Mulvey, M. R., & Loewen, P. C. (1989). Nucleotide sequence of katF of Escherichia coli suggests KatF protein is a novel sigma transcription factor. *Nucleic Acids Research*, *17*(23), 9979–9991. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2690013

Murakami, K. S., Masuda, S., Campbell, E. A., Muzzin, O., & Darst, S. A. (2002). Structural basis of transcription initiation: An RNA polymerase holoenzyme-DNA complex. *Science*, *296*(5571), 1285–1290. https://doi.org/10.1126/science.1069595

Murakami, K. S., Masuda, S., & Darst, S. a. (2002). Structural basis of transcription initiation: RNA polymerase holoenzyme at 4 A resolution. *Science (New York, N.Y.)*, *296*(5571), 1280–1284. https://doi.org/10.1126/science.1069594

Niu, W., Kim, Y., Tau, G., Heyduk, T., & Ebright, R. H. (1996). Transcription activation at class II CAP-dependent promoters: two interactions between CAP and RNA polymerase. *Cell*, *87*(6), 1123–1134. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8978616

Notley-McRobb, L., King, T., & Ferenci, T. (2002). rpoS mutations and loss of general stress resistance in Escherichia coli populations as a consequence of conflict between competing stress responses. *Journal of Bacteriology*, *184*(3), 806–811. https://doi.org/10.1128/JB.184.3.806

Old, lain G., Margarita, D., & Girons, I. Saint. (1992). Nucleotide sequence of the *Borrelia burgdorferi rpmH* gene encoding ribosomal protein L34. *Nucleic Acids Research*, *20*(22), 6097–6097. https://doi.org/10.1093/nar/20.22.6097

Olsén, A., Jonsson, A., & Normark, S. (1989). Fibronectin binding mediated by a novel class of surface organelles on Escherichia coll. *Nature*, *338*(6217), 652–655. https://doi.org/10.1038/338652a0

Österberg, S., Peso-Santos, T. Del, Shingler, V., del Peso-Santos, T., & Shingler, V. (2011). Regulation of alternative sigma factor use. *Annual Review of Microbiology*, *65*(1), 37–55. https://doi.org/10.1146/annurev.micro.112408.134219

Paget, M. S. (2015). Bacterial Sigma Factors and Anti-Sigma Factors: Structure, Function and Distribution. *Biomolecules*, *5*(3), 1245–1265. https://doi.org/10.3390/biom5031245

Paget, M. S. B., & Helmann, J. D. (2003). The sigma70 family of sigma factors. *Genome Biology*, *4*(1), 203. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12540296

Peano, C., Wolf, J., Demol, J., Rossi, E., Petiti, L., De Bellis, G., … Gross, C. A. (2015). Characterization of the Escherichia coli σ(S) core regulon by Chromatin Immunoprecipitation-sequencing (ChIP-seq) analysis. *Scientific Reports*, *5*(April), 10469. https://doi.org/10.1038/srep10469

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E.

C., & Ferrin, T. E. (2004). UCSF Chimera?A visualization system for exploratory research and analysis. *Journal of Computational Chemistry*, *25*(13), 1605–1612. https://doi.org/10.1002/jcc.20084

Pratt, L. A., & Silhavy, T. J. (1998). Crl stimulates RpoS activity during stationary phase. *Molecular Microbiology*, *29*(5), 1225–1236. https://doi.org/10.1046/j.1365-2958.1998.01007.x

Pribnow, D. (1975). Nucleotide sequence of an RNA polymerase binding site at an early T7 promoter. *Proceedings of the National Academy of Sciences of the United States of America*, *72*(3), 784–788. https://doi.org/10.1073/pnas.72.3.784

Punjani, A., Rubinstein, J. L., Fleet, D. J., & Brubaker, M. A. (2017). cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nature Methods*, *14*(3), 290–296. https://doi.org/10.1038/nmeth.4169

Raffaelle, M., Kanin, E. I., Vogt, J., Burgess, R. R., & Ansari, A. Z. (2005). Holoenzyme Switching and Stochastic Release of Sigma Factors from RNA Polymerase In Vivo. *Molecular Cell*, *20*(3), 357–366. https://doi.org/10.1016/j.molcel.2005.10.011

Ramachandran, V. K., Shearer, N., Jacob, J. J., Sharma, C. M., & Thompson, A. (2012). The architecture and ppGpp-dependent expression of the primary transcriptome of Salmonella Typhimurium during invasion gene expression. *BMC Genomics*, *13*(1), 25. https://doi.org/10.1186/1471-2164-13-25

Ray-Soni, A., Bellecourt, M. J., & Landick, R. (2016). Mechanisms of Bacterial Transcription Termination: All Good Things Must End. *Annual Review of Biochemistry*, *85*(1), 319–347. https://doi.org/10.1146/annurev-biochem-060815-014844

Reverter, D., & Lima, C. D. (2004). A Basis for SUMO Protease Specificity Provided by Analysis of Human Senp2 and a Senp2-SUMO Complex. *Structure*, *12*(8), 1519–1531. https://doi.org/10.1016/j.str.2004.05.023

Ring, B. Z., Yarnell, W. S., & Roberts, J. W. (1996). Function of E. coli RNA polymerase sigma factor sigma 70 in promoter-proximal pausing. *Cell*, *86*(3), 485–493. https://doi.org/10.1016/S0092-8674(00)80121-X

Robbe-Saule, V., Jaumouillé, V., Prévost, M.-C. C., Guadagnini, S., Talhouarne, C., Mathout, H., … Norel, F. (2006). Crl activates transcription initiation of RpoS-regulated genes involved in the multicellular behavior of Salmonella enterica serovar typhimurium. *Journal of Bacteriology*, *188*(11), 3983–3994. https://doi.org/10.1128/JB.00033-06

Robbe-Saule, V., Lopes, M. D., Kolb, A., & Norel, F. (2007). Physiological effects of Crl in Salmonella are modulated by sigmaS level and promoter specificity. *Journal of Bacteriology*, *189*(8), 2976–2987. https://doi.org/10.1128/JB.01919-06

Robert, X., & Gouet, P. (2014). Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Research*, *42*(W1), W320–W324. https://doi.org/10.1093/nar/gku316

Ross, W, Gosink, K. K., Salomon, J., Igarashi, K., Zou, C., Ishihama, A., … Gourse, R.

L. (1993). A third recognition element in bacterial promoters: DNA binding by the alpha subunit of RNA polymerase. *Science (New York, N.Y.)*, *262*(5138), 1407–1413. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8248780

Ross, Wilma, Vrentas, C. E., Sanchez-Vazquez, P., Gaal, T., & Gourse, R. L. (2013). The Magic Spot: A ppGpp Binding Site on E. coli RNA Polymerase Responsible for Regulation of Transcription Initiation. *Molecular Cell*, *50*(3), 420–429. https://doi.org/10.1016/j.molcel.2013.03.021

Saba, J., Chua, X. Y., Mishanina, T. V, Nayak, D., Windgassen, T. A., Mooney, R. A., & Landick, R. (2019). The elemental mechanism of transcriptional pausing. *ELife*, *8*. https://doi.org/10.7554/eLife.40981

Saecker, R. M., Record, M. T., & Dehaseth, P. L. (2011). Mechanism of bacterial transcription initiation: RNA polymerase - Promoter binding, isomerization to initiation-competent open complexes, and initiation of RNA synthesis. *Journal of Molecular Biology*, *412*(5), 754–771. https://doi.org/10.1016/j.jmb.2011.01.018

Sak, B. D., Eisenstark, A., & Touati, D. (1989). Exonuclease III and the catalase hydroperoxidase II in Escherichia coli are both regulated by the katF gene product. *Proceedings of the National Academy of Sciences of the United States of America*, *86*(9), 3271–3275. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2541439

Santos-Zavaleta, A., Pérez-Rueda, E., Sánchez-Pérez, M., Velázquez-Ramírez, D. A., & Collado-Vides, J. (2019). Tracing the phylogenetic history of the Crl regulon through the Bacteria and Archaea genomes. *BMC Genomics*, *20*(1), 299. https://doi.org/10.1186/s12864-019-5619-z

Schneider, C. A., Rasband, W. S., & Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nature Methods*, *9*(7), 671–675. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/22930834

Schwartz, E. C., Shekhtman, A., Dutta, K., Pratt, M. R., Cowburn, D., Darst, S., & Muir, T. W. (2008). A Full-Length Group 1 Bacterial Sigma Factor Adopts a Compact Structure Incompatible with DNA Binding. *Chemistry and Biology*, *15*(10), 1091–1103. https://doi.org/10.1016/j.chembiol.2008.09.008

Seidel, S. A. I., Dijkman, P. M., Lea, W. A., van den Bogaart, G., Jerabek-Willemsen, M., Lazic, A., … Duhr, S. (2013). Microscale thermophoresis quantifies biomolecular interactions under previously challenging conditions. *Methods*, *59*(3), 301–315. https://doi.org/10.1016/j.ymeth.2012.12.005

Shukla, J., Gupta, R., Thakur, K. G., Gokhale, R., & Gopal, B. (2014). Structural basis for the redox sensitivity of the *Mycobacterium tuberculosis* SigK–RskA σ–anti-σ complex. *Acta Crystallographica Section D Biological Crystallography*, *70*(4), 1026–1036. https://doi.org/10.1107/S1399004714000121

Shultzaberger, R. K., Chen, Z., Lewis, K. A., & Schneider, T. D. (2007). Anatomy of Escherichia coli σ 70 promoters. *Nucleic Acids Research*, *35*(3), 771–788. https://doi.org/10.1093/nar/gkl956

Siebenlist, U. (1979). RNA polymerase unwinds an 11-base pair segment of a phage T7 promoter [23]. *Nature*, *279*(5714), 651–652. https://doi.org/10.1038/279651a0

Siebenlist, U., Simpson, R. B., & Gilbert, W. (1980). E. coli RNA polymerase interacts homologously with two different promoters. *Cell*, *20*(2), 269–281. https://doi.org/10.1016/0092-8674(80)90613-3

Siegele, D. A., Hu, J. C., Walter, W. A., & Gross, C. A. (1989). Altered promoter recognition by mutant forms of the σ 70 subunit of Escherichia coli RNA polymerase. *Journal of Molecular Biology*, *206*(4), 591–603. https://doi.org/10.1016/0022-2836(89)90568-8

Song, M., Kim, J.-S., Liu, L., Husain, M., & Vázquez-Torres, A. (2016). Antioxidant Defense by Thioredoxin Can Occur Independently of Canonical Thiol-Disulfide Oxidoreductase Enzymatic Activity. *Cell Reports*, *14*(12), 2901–2911. https://doi.org/10.1016/j.celrep.2016.02.066

Staroń, A., Sofia, H. J., Dietrich, S., Ulrich, L. E., Liesegang, H., & Mascher, T. (2009). The third pillar of bacterial signal transduction: classification of the extracytoplasmic function (ECF) σ factor protein family. *Molecular Microbiology*, *74*(3), 557–581. https://doi.org/10.1111/j.1365-2958.2009.06870.x

Tabib-Salazar, A., Liu, B., Doughty, P., Lewis, R. A., Ghosh, S., Parsy, M. L., … Paget, M. S. (2013). The actinobacterial transcription factor RbpA binds to the principal sigma subunit of RNA polymerase. *Nucleic Acids Research*, *41*(11). https://doi.org/10.1093/nar/gkt277

Tanaka, K., Takayanagi, Y., Fujita, N., Ishihama, A., & Takahashi, H. (1993). Heterogeneity of the principal sigma factor in Escherichia coli: the rpoS gene product, sigma 38, is a second principal sigma factor of RNA polymerase in stationary-phase Escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America*, *90*(8), 3511–3515. https://doi.org/10.1073/PNAS.90.8.3511

Touati, E., Dassa, E., & Boquet, P. L. (1986). Pleiotropic mutations in appR reduce pH 2.5 acid phosphatase expression and restore succinate utilisation in CRP-deficient strains of Escherichia coli. *Molecular & General Genetics : MGG*, *202*(2), 257–264. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3517593

Travers, A. A., & Burgess, R. R. (1969). Cyclic re-use of the RNA polymerase sigma factor. *Nature*, *222*(5193), 537–540. https://doi.org/10.1038/222537a0

Tuveson, R. W. (1981). THE INTERACTION OF A GENE (nur) CONTROLLING NEAR-UV SENSITIVITY AND THE polA1 GENE IN STRAINS OF E. COLI K12. *Photochemistry and Photobiology*, *33*(6), 919–923. https://doi.org/10.1111/j.1751-1097.1981.tb05513.x

Tuveson, R. W., & March, M. E. (1980). PHOTODYNAMIC AND SUNLIGHT INACTIVATION OF ESCHERICHIA COLI STRAINS DIFFERING IN NEAR-UV SENSITIVITY AND RECOMBINATION PROFICIENCY. *Photochemistry and Photobiology*, *31*(3), 287–289. https://doi.org/10.1111/j.1751-1097.1980.tb03720.x

Twist, K.-A., Husnain, S. I., Franke, J. D., Jain, D., Campbell, E. a, Nickels, B. E., … Westblade, L. F. (2011). A novel method for the production of in vivo-assembled,

recombinant Escherichia coli RNA polymerase lacking the α C-terminal domain. *Protein Science : A Publication of the Protein Society*, *20*(6), 986–995. https://doi.org/10.1002/pro.622

Typas, A., Barembruch, C., Possling, A., & Hengge, R. (2007). Stationary phase reorganisation of the Escherichia coli transcription machinery by Crl protein, a fine-tuner of sigmas activity and levels. *The EMBO Journal*, *26*(6), 1569–1578. https://doi.org/10.1038/sj.emboj.7601629

Typas, A., Becker, G., & Hengge, R. (2007). The molecular basis of selective promoter activation by the sigmaS subunit of RNA polymerase. *Molecular Microbiology*, *63*(5), 1296–1306. https://doi.org/10.1111/j.1365-2958.2007.05601.x

Typas, A., & Hengge, R. (2006). Role of the spacer between the -35 and -10 regions in sigmas promoter selectivity in Escherichia coli. *Molecular Microbiology*, *59*(3), 1037–1051. https://doi.org/10.1111/j.1365-2958.2005.04998.x

Vassylyev, D. G., Sekine, S., Laptenko, O., Lee, J., Vassylyeva, M. N., Borukhov, S., & Yokoyama, S. (2002). Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 Å resolution. *Nature*, *417*(6890), 712–719. https://doi.org/10.1038/nature752

Volkert, M. R., Hajec, L. I., Matijasevic, Z., Fang, F. C., & Prince, R. (1994). Induction of the Escherichia coli aidB gene under oxygen-limiting conditions requires a functional rpoS (katF) gene. *Journal of Bacteriology*, *176*(24), 7638–7645. https://doi.org/10.1128/JB.176.24.7638-7645.1994

Waldburger, C., Gardella, T., Wong, R., & Susskind, M. M. (1990). Changes in conserved region 2 of Escherichia coli σ70 affecting promoter recognition. *Journal of Molecular Biology*, *215*(2), 267–276. https://doi.org/10.1016/S0022-2836(05)80345-6

Wassarman, K. M., Repoila, F., Rosenow, C., Storz, G., & Gottesman, S. (2001). Identification of novel small RNAs using comparative genomics and microarrays. *Genes & Development*, *15*(13), 1637–1651. https://doi.org/10.1101/gad.901001

Weber, H., Polen, T., Heuveling, J., Wendisch, V. F., & Hengge, R. (2005). Genome-wide analysis of the general stress response network in Escherichia coli: sigmaS-dependent genes, promoters, and sigma factor selectivity. *Journal of Bacteriology*, *187*(5), 1591–1603. https://doi.org/10.1128/JB.187.5.1591-1603.2005

Wigneshweraraj, S., Bose, D., Burrows, P. C., Joly, N., Schumacher, J., Rappas, M., … Buck, M. (2008). Modus operandi of the bacterial RNA polymerase containing the σ^{54} promoter-specificity factor. *Molecular Microbiology*, *68*(3), 538–546. https://doi.org/10.1111/j.1365-2958.2008.06181.x

Winkelman, J. T., Winkelman, B. T., Boyce, J., Maloney, M. F., Chen, A. Y., Ross, W., & Gourse, R. L. (2015). Crosslink Mapping at Amino Acid-Base Resolution Reveals the Path of Scrunched DNA in Initial Transcribing Complexes. *Molecular Cell*, *59*(5), 768–780. https://doi.org/10.1016/j.molcel.2015.06.037

Wösten, M. M. S. . (1998). Eubacterial sigma-factors. *FEMS Microbiology Reviews*, *22*(3), 127–150. https://doi.org/10.1016/S0168-6445(98)00011-4

Wu, X., Haakonsen, D. L., Sanderlin, A. G., Liu, Y. J., Shen, L., Zhuang, N., … Zhang, Y. (2018). Structural insights into the unique mechanism of transcription activation by Caulobacter crescentus GcrA. *Nucleic Acids Research*, *46*(6), 3245–3256. https://doi.org/10.1093/nar/gky161

Wurm, P., Tutz, S., Mutsam, B., Vorkapic, D., Heyne, B., Grabner, C., … Reidl, J. (2017). Stringent factor and proteolysis control of sigma factor RpoS expression in Vibrio cholerae. *International Journal of Medical Microbiology*, *307*(3), 154–165. https://doi.org/10.1016/J.IJMM.2017.01.006

Yim, H. H., Brems, R. L., & Villarejo, M. (1994). Molecular characterization of the promoter of osmY, an rpoS-dependent gene. *Journal of Bacteriology*, *176*(1), 100–107. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8282684

Young, B. A., Anthony, L. C., Gruber, T. M., Arthur, T. M., Heyduk, E., Lu, C. Z., … Gross, C. A. (2001). A coiled-coil from the RNA polymerase beta' subunit allosterically induces selective nontemplate strand binding by sigma(70). *Cell*, *105*(7), 935–944. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11439189

Zafar, M. A., Carabetta, V. J., Mandel, M. J., & Silhavy, T. J. (2014). Transcriptional occlusion caused by overlapping promoters. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(4), 1557–1561. https://doi.org/10.1073/pnas.1323413111

Zenkin, N., Kulbachinskiy, A., Yuzenkova, Y., Mustaev, A., Bass, I., Severinov, K., & Brodolin, K. (2007). Region 1.2 of the RNA polymerase σ subunit controls recognition of the −10 promoter element. *The EMBO Journal*, *26*(4), 955–964. https://doi.org/10.1038/sj.emboj.7601555

Zhang, G., Campbell, E. A., Minakhin, L., Richter, C., Severinov, K., & Darst, S. A. (1999). Crystal structure of thermus aquaticus core RNA polymerase at 3.3 å resolution. *Cell*, *98*(6), 811–824. https://doi.org/10.1016/S0092-8674(00)81515-9

Zhang, G., & Darst, S. A. (1998). Structure of the Escherichia coli RNA polymerase α subunit amino- terminal domain. *Science*, *281*(5374), 262–266. https://doi.org/10.1126/science.281.5374.262

Zheng, D., Constantinidou, C., Hobman, J. L., & Minchin, S. D. (2004). Identification of the CRP regulon using in vitro and in vivo transcriptional profiling. *Nucleic Acids Research*, *32*(19), 5874–5893. https://doi.org/10.1093/nar/gkh908

Zheng, S. Q., Palovcak, E., Armache, J.-P. P., Verba, K. A., Cheng, Y., & Agard, D. A. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy, 14 Nature Methods § (2017). Nature Publishing Group. https://doi.org/10.1038/nmeth.4193

Zivanov, J., Nakane, T., & Scheres, S. H. W. (2019). A Bayesian approach to beam-induced motion correction in cryo-EM single-particle analysis. *IUCrJ*, *6*(1), 5–17. https://doi.org/10.1107/S205225251801463X

Zuber, P., Healy, J., Carter, H. L., Cutting, S., Moran, C. P., & Losick, R. (1989). Mutation changing the specificity of an RNA polymerase sigma factor. *Journal of Molecular Biology*, *206*(4), 605–614. https://doi.org/10.1016/0022-2836(89)90569