2020

Cryo-EM Studies of Bacterial RNA Polymerase

James Chen

CRYO-EM STUDIES OF BACTERIAL RNA POLYMERASE

A Thesis Presented to the Faculty of
The Rockefeller University
in Partial Fulfillment of the Requirements for
the degree of Doctor of Philosophy

by
James Chen
June 2020

CRYO-EM STUDIES OF BACTERIAL RNA POLYMERASE

James Chen, Ph.D.
The Rockefeller University 2020

In bacteria, a single RNA polymerase (RNAP) performs all transcription. The overall structure of bacterial RNAP resembles a crab claw with pincers comprising the β' and β subunits and a large cleft where the active site sits. Structural information about this essential enzyme has mainly been provided by X-ray crystal structures of stable transcription complexes. RNAP crystal structures are difficult to obtain and the captured states may not always represent physiological states of the enzyme due to crystal packing artifacts. Recent advances in electron detectors and software enable near-atomic resolution structures of large biological complexes to be determined by single particle cryo-electron microscopy (cryo-EM). Unlike X-ray crystallography, cryo-EM samples can be directly visualized without crystallization.

In this thesis, I optimized and utilized cryo-EM methodologies to structurally characterize several bacterial RNAP complexes from *Escherichia coli* (*Eco*) and *Mycobacterium tuberculosis* (*Mtb*): (1) *Eco* σ^{70}-holoenzyme ($E\sigma^{70}$) in complex with the *Eco* non-coding RNA (ncRNA) 6S RNA; (2) *Mtb* RNAP bound to the RNAP inhibitor Fidaxomicin (Fdx); (3) $E\sigma^{70}$ bound to the F element-encoded TraR protein; and (4) $E\sigma^{70}$-dependent promoter DNA melting intermediates stabilized by the TraR transcription factor. Using cryo-EM, I captured RNAP structures that were intractable to crystallization, visualized multiple RNAP conformational states populated in solution, deconvoluted RNAP molecular motions, and observed transient complexes. The work in this thesis showcases the power of cryo-EM to examine macromolecular machines in action.

(1) Bacterial 6S RNAs globally regulate transcription by RNAP, directly competing with promoter DNA binding. During transitions between exponential and stationary growth phases, *Eco* 6S RNA plays a key role in the transcriptional reprogramming by interacting specifically with the housekeeping $E\sigma^{70}$. During my initial cryo-EM experiments with the *Eco* 6S RNA-$E\sigma^{70}$ complex, I encountered a severe particle orientation bias in my samples. I discovered that the zwitterionic detergent 3-([3-Cholamidopropyl]dimethylammonio)-2-hydroxy-1-propanesulfonate (CHAPSO) was uniquely effective at solving this issue for bacterial RNAPs. Using this detergent, I determined the cryo-EM structure of the *Eco* 6S RNA/$E\sigma^{70}$ complex in combination with footprinting and crosslinking approaches in order to elucidate the structural mechanism of 6S RNA mediated inhibition of $E\sigma^{70}$. The structure reveals that 6S RNA is composed of duplex RNA segments that have A-form C3'-endo sugar puckers but with widened major groove widths, giving the RNA an overall architecture that mimics B-form promoter DNA. The results showed how 6S RNA specifically targets $E\sigma^{70}$ and how an ncRNA can mimic B-form DNA to directly regulate transcription by the DNA-dependent RNAP.

(2) Fdx is a RNAP-binding drug that is highly effective against *Mtb* RNAP *in vitro*. In collaboration with postdoctoral fellow Hande Boyaci, we solved cryo-EM structures of *Mtb* σ^{A}-holoenzyme (σ^{A}-holo) bound to the actinobacteria general transcription factor RbpA and an upstream promoter DNA fork (us-fork) in the presence and absence of Fdx to identify the structural determinants of Fdx binding to RNAP. The results show that Fdx acts like a doorstop fo jam the RNAP clamp module in an open conformation. The structures define the Fdx binding pocket, which includes contacts with RbpA, explaining the drug's strong effect on *Mtb*.

(3) Under starvation conditions, *Eco* changes the expression of almost one-quarter of its genes within 5 minutes. This global response depends on two factors, the alarmone ppGpp, and

the transcription factor DksA. Unlike most transcription factors, ppGpp and DksA bind directly to RNAP but not to DNA. Another transcription factor, TraR, regulates RNAP the same way as ppGpp and DksA. TraR, and ppGpp/DksA, can activate or inhibit transcription initiation depending on the promoter sequence. TraR inhibits $E\sigma^{70}$-dependent transcription from ribosomal RNA and ribosomal protein promoters and activates amino acid biosynthesis and transport promoters *in vivo* and *in vitro*. To understand how these factors modulate transcription initiation, I solved cryo-EM structures of $E\sigma^{70}$, with and without TraR, and an open promoter complex (RPo) using a ribosomal promoter that is regulated by TraR, *rpsT* P2. My cryo-EM structural analyses show that TraR regulates $E\sigma^{70}$ transcription initiation by binding and inducing major conformational changes to mobile RNAP domains that affect RNAP-promoter DNA interactions.

(4) Transcription initiation is a multi-step process that leads to the formation of RPo. Using TraR to modulate $E\sigma^{70}$ transcription initiation on the ribosomal protein promoter *rpsT* P2, I observed promoter DNA melting intermediates by cryo-EM. These structures and supporting biochemical data delineate steps of the RPo formation pathway and provide insight into transient promoter-RNAP interactions that occur along this pathway. These structures span the initial recognition of the duplex promoter in a closed complex (RPc) to the final RPo.

Findings from this thesis provide structural and biochemical insight into bacterial RNAP regulation and function. The work in this thesis shows how single-particle cryo-EM can be applied to study transcription factors and small molecules that modulate RNAP as well as transient transcriptional states. This thesis research outlines a framework for future studies of bacterial transcription complexes using cryo-EM.

ACKNOWLEDGMENTS

I would like to thank the many individuals who provided mentorship, inspiration, and support throughout my scientific journey at The Rockefeller University. Without their contributions, the work presented in this thesis would not have been possible.

I want to express my deepest gratitude to my Ph.D. advisors, Seth A. Darst and Elizabeth A. Campbell. Their combined mentorship was indispensable during my years as a graduate student at The Rockefeller University. They always supported me in all my endeavors by providing me with sound advice and thorough preparation. Not only did they train me to be a careful and critical structural biologist, they also taught me life lessons outside of the lab that have also been instrumental in my development as a person.

Seth's inquisitive nature pushed me to become a more vigorous scientist. His attention to detail when editing pieces of my writing or when going over my oral presentations has taught me to communicate my ideas in a clearer, more precise manner. Seth's knack for structural biology has always amazed me. In our time together, he patiently showed me how to build RNAP structures in Coot one amino acid at a time and how to look at these structures in novel meaningful ways. He has imparted upon me his creativity as well as his meticulous perseverance. In addition to his teachings, I truly value his trust in me to pioneer cryo-EM in his lab. This opportunity not only expanded my skill set but also allowed me to meet and collaborate with many great scientists. Beyond being a dedicated scientific mentor, Seth always has a ready supply of humor and cat stories.

Liz's ability to think outside of the box has inspired me to constantly look for new, creative solutions to scientific problems. Her valiant effort to push the Darst-Campbell Lab to study pathogenic bacteria and other aspects of transcription regulation is truly motivating. Liz is also a highly methodical scientist. She always manages to find time to help me troubleshoot experiments that went awry; patiently staring at gels or helping me look over protocols that she had developed years ago, all the while she supervises and helps everyone else in the lab. Liz is not only a astute collaborator in lab, but also a warm and compassionate person. She truly connects with everyone in lab during coffee breaks or walks around campus, times that I will always cherish. Her exercise classes also helped me to stay mentally and physically fit and to make health-conscious decisions.

The work presented in this thesis relied on the experimental contributions from fellow colleagues and collaborators. I would like to acknowledge: Thomas Walz's Lab at The Rockefeller University for assisting in initial cryo-EM experiments on 6S RNA; Karan Wassarman from the University of Wisconsin-Madison for providing biochemical data on 6S RNA; Alex J. Noble from the New York Structural Biology Center for performing cryo-ET experiments; postdoctoral fellow in the Darst-Campbell Lab, Hande Boyaci, for biochemical and structural data on Fidaxomicin; Richard L. Gourse's Lab at the University of Wisconsin-Madison for providing biochemical data and helpful discussions for the TraR and promoter DNA melting intermediate projects; and Brian T. Chait's Lab at The Rockefeller University for performing native mass spectrometry experiments. I am also grateful for the individuals who manage and run the cryo-EM faculties, in particular: Mark Ebrahim and Johanna Sotiris at The Rockefeller University's Cryo-Electron Microscopy Resource Center and Edward Eng at the New York Structural Biology Center.

I am thankful to my thesis committee members: Luciano Marraffini, Shixin Liu, and Sebastian Klinge, who have provided informative insight and on-going feedback throughout the

progression of my thesis research. Their scientific passion and words of encouragement inspire me to continue pursuing knowledge and to perform scholarly research. I also would like to thank Dong Wang from the University of California San Diego for participating in the evaluation of this thesis as the external examiner.

I treasure all the working relationships and friendships that I have built in the Darst-Campbell Lab as well as with the Rockefeller community during my time as a graduate student. I would like to give special recognition to: Brian Bae for mentoring me when I first joined the Darst-Campbell Lab as a SURF student; Mengyu Wu and Jin Young Kang for helping me pioneer cryo-EM in the lab; Courtney Chiu for contributions to the TraR project; Eliza Llewellyn for being a pillar of support and providing wisdom and life advice; Ruth Saecker for intellectual conversations and inspiriting words; and Brandon Malone for witty remarks and scientific discussions.

I would like to recognize The Dean's Office for providing a supportive environment throughout my time at Rockefeller and assisting in composition of this thesis. I would especially like to thank Dean Sidney Strickland and Senior Assistant Dean Emily Harms for giving me the opportunity to study at The Rockefeller University, first as a SURF student and later as a graduate student.

Lastly, I thank my family for their unwavering love and support throughout my graduate studies at The Rockefeller University.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

6S-Eσ^{70}	*Eco* 6S RNA bound to *Eco* RNAP σ^{70}-holoenzyme
BH	Bridge helix
BIV	Bovine immunodeficiency virus
BL	Basic linker
BME	2-Mercaptoethanol
BPA	p-Benzoyl-L-phenylalanine
bps	Base pairs
CB	Central bubble
CD	Core domain
Cdf	*Clostridium difficile*
CHAPSO	Cholamidopropyl]dimethylammonio)-2-hydroxy-1-propanesulfonate
CMC	Critical micelle concentration
CPK	Corey-Pauling-Koltun
CRE	Core-recognition element
cryo-EM	Cryo-electron microscopy
cryo-ET	Cryo-electron tomography
CS	Closing stem
CS2	CryoSPARC2
CV	Column volumes
DB	Downstream bulge
DD	Downstream duplex
DD*	Shuffled downstream duplex
DDM	n-Dodecyl-B-D-maltoside
DSR	Discriminator region
DTT	Dithiothreitol
E	*Escherichia coli* RNA polymerase core
Eco	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid
EMSA	Electrophoretic mobility shift assay
E_{od}	Orientation efficiency
EXT	Extended -10
EΔ1.1σ^{70}	*E. coli* σ^{70}-holoenzyme lacking $\sigma^{70}{}_{1.1}$
Eσ^{N}	*E. coli* σ^{N}-holoenzyme
Eσ^{S}	*E. coli* σ^{S}-holoenzyme
Fdx	Fidaxomicin
FeBABE	Iron bromoacetamidobenzyl-EDTA
f_{empty}	Fraction of unsampled Fourier space
FL	Fork loop
FRET	Förster resonance energy transfer
GL	Gate loop
GTF	General transcription factor
his pEC	*his* paused elongation complex
IDR	Intrinsically disordered region
KAN	Kanamycin

kDa	Kilodalton
LSI	Lineage-specific insert
MBR	Multi-body refinement
MDR	Multi-drug resistant
Mgw	Major groove widths
mgw	Minor groove widths
Msm	*Mycobacterium smegmatis*
Mtb	*Mycobacterium tuberculosis*
NCR	Non-conserved region
ncRNA	Non-coding RNA
nMS	Native mass spectrometry
NP40S	Nonidet P40 Substitute
NTP	Nucleotide
nt-strand	Nontemplate-strand
NTT	N-terminal tail
PAGE	Polyacrylamide gel electrophoresis
PCA	Principal component analysis
PCR	Polymerase chain reaction
pEC	Paused elongation complex
pol II	RNAP II
pRNA	Product RNA
PSF	Point spread function
$RbpA^{BL}$	RbpA basic linker
$RbpA^{CD}$	RbpA core domain
$RbpA^{SID}$	RbpA sigma interacting domain
RID	RNAP interacting domain
Rif	Rifampicin
RMSD	Root-mean-square deviation
RNAP	RNA polymerase
rNTP	Ribonucleotides
rot	Rotation angle
RPc	Closed promoter complex
RPitc	Initial transcribing complex
RPo	Open promoter complex
rpsT P2*	*rpsT* P2 mutant
SBHM	Sandwich-barrel hybrid motif
SD	Standard deviation
SEM	Standard error of the mean
SID	Sigma interacting domain
SNR	Signal-to-noise
Sw1	Switch 1
Sw2	Switch 2
Sw3	Switch 3
Sw4	Switch 4
Taq	*Thermus aquaticus*
TB	Tuberculosis

TEC	Transcription elongation complex
TEM	Transmission electron microscope
TIC	Transcription initiation complex
TIC	Transcription initiation complexes
tilt	Tilt angle
TL	Trigger loop
$TraR_C$	C-terminal helix of TraR
$TraR_G$	Globular domain of TraR
$TraR_N$	N-terminal helix of TraR
TSS	Transcription start site
t-strand	Template strand
TTC	Transcription termination complex
Tth	*Thermus thermophilus*
UB1	Upstream bulges 1
UB2	Upstream bulges 2
UD1	Upstream duplexes 1
UD2	Upstream duplexes 2
UD3	Upstream duplexes 3
UP	Upstream
UTL	Upstream terminal loop
WT	Wild-type
ZBH	β' Zinc Binding Domain
αCTD	C-terminal domain of the α subunit
αNTD	N-terminal domain of the α subunit
$\Delta 1.1\sigma^{70}$	*E. coli* σ^{70} with region 1.1 deleted
ΔSi3	Si3 deletion
ΔαCTD-E	*E. coli* RNAP with αCTD deleted
σ_2	σ region 2
σ_3	σ region 3
σ_4	σ region 4
$\sigma^{70}{}_{1.1}$	σ^{70} region 1.1
$\sigma^{70}{}_2$	σ^{70} region 2
$\sigma^{70}{}_3$	σ^{70} region 3
$\sigma^{70}{}_4$	σ^{70} region 4
σ^A-holo	σ^A-holoenzyme
$\sigma^A{}_N$	N-terminus of σ^A

CHAPTER 1. Introduction

1.1 Overview of bacterial transcription

Bacteria are very abundant and widely dispersed organisms (Bunge et al., 2014). They are able to adapt and thrive in many conditions ranging from extreme temperatures to toxins (Roszak and Colwell, 1987). Control of gene expression is a major method that bacteria use to respond to their environment (López-Maury et al., 2008). The major control point for gene expression is transcription.

Transcription is the process by which genetic information is transcribed into RNA. In bacteria, a single DNA-dependent RNA polymerase (RNAP) performs all transcription. Bacterial RNAPs consist of an evolutionarily conserved catalytic core with subunit composition $\alpha_2\beta\beta'\omega$ (Burgess, 1969; Burgess and Travers, 1970). Although the catalytic core is capable of transcription elongation (RNA synthesis), bacterial RNAPs require a promoter-specificity factor, σ, for promoter DNA recognition and transcription initiation (Burgess et al., 1969; Murakami and Darst, 2003). When core RNAP is bound to a σ factor, it is termed RNAP holoenzyme. In *Escherichia coli* (*Eco*), RNAP (E) binds to the essential primary σ factor, σ^{70}, to form the holoenzyme ($E\sigma^{70}$) that initiates at promoters for housekeeping genes (Ishihama, 2000; Lonetto et al., 1992).

Once bound to σ, RNAP can readily start transcription. Transcription consists of at least three stages: initiation, elongation, and termination **(Figure 1.1)**. Upon locating the promoter DNA, RNAP forms a closed promoter complex (RPc) with the promoter (Kovacic, 1987). RNAP initiates the melting of about 13 base pairs (bps) of DNA from -11 to +2 (+1 is the start site for transcription), to form the open promoter complex (RPo). In RPo, the DNA template strand (t-strand) is loaded into the RNAP active site, positioning the transcription start site (TSS) for catalysis (Bae et al., 2015a; Zuo and Steitz, 2015). In the presence of ribonucleotides (rNTP), the RPo begins to synthesize RNA to form an initial transcribing complex (RPitc). Nucleotide substrates diffuse into the secondary channel of RNAP and generate transcripts that vary in length (Krummel and Chamberlin, 1989; Munson and Reznikoff, 1981; Vo et al., 2003). Production of longer transcripts destabilizes σ with the core RNAP resulting in promoter escape (Henderson et al., 2017; Murakami and Darst, 2003). σ is released stochastically and the RNAP transitions into a transcription elongation complex (TEC), continuing the synthesis of the RNA transcript (Shimamoto et al., 1986). Transcription continues until the RNAP encounters a termination signal characterized by an RNA hairpin or a termination factor (Ding Jun Jin et al., 1992; Gusarov and Nudler, 1999). The transcription termination complex (TTC) releases the RNA transcript and DNA. Released RNAP and σ reenter the cycle for another round of transcription. The primary focus of my thesis will be centered around transcription initiation and its regulation.

1.2 Promoter recognition by RNAP holoenzyme bound to primary σ factors

In bacteria, σ factors guide the core RNAP to DNA promoters by providing specific contacts with the DNA (Campbell et al., 2002; Murakami et al., 2002a). All bacteria possess a primary housekeeping σ factor (σ^{70} in *Eco*, σ^{A} in *Thermus* and *Mycobacterium*) that controls the transcription of essential genes. The vast majority of initiation events in bacteria involve RNAP holoenzymes bound to the primary σ (Feklístov et al., 2014; Gruber and Gross, 2003). In addition to the primary σ, alternative σ factors regulate transcription in response to various stresses and environmental conditions (Gruber and Gross, 2003; Lonetto et al., 1992)

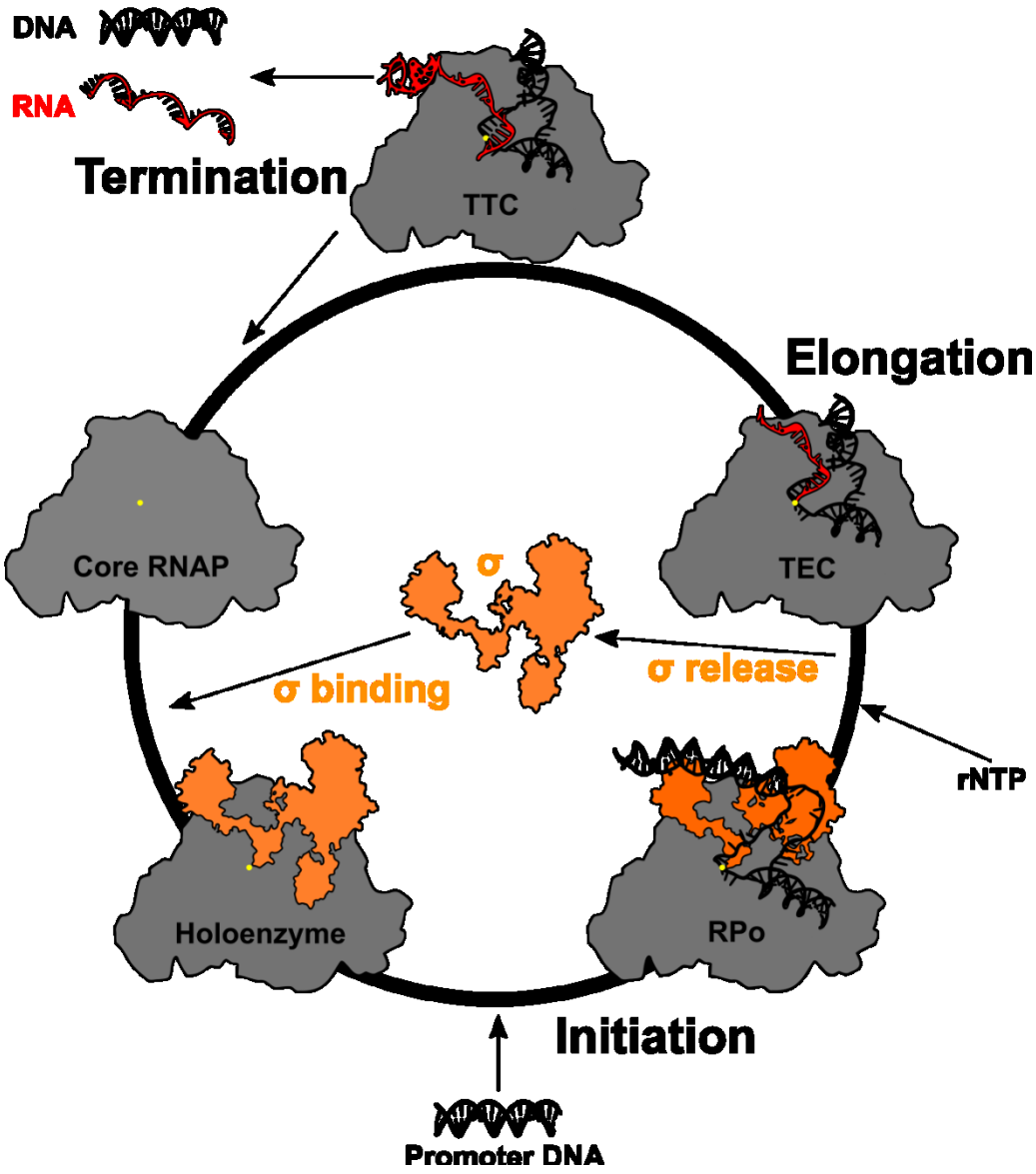


Figure 1.1. Schematic of the bacterial transcription cycle
The core RNAP (colored in grey with active site Mg^{2+} depicted as a yellow dot) binds a promoter specificity factor, σ (colored in orange), generating the holoenzyme. Holoenzyme can specifically recognize and bind promoter DNA (black) to begin transcription initiation. During initiation, the duplex promoter DNA is melted and loaded into the RNAP active site cleft to form RPo. In the presence of rNTP, the initiation complex synthesizes many RNA transcripts (red) until it produces a transcript long enough to cause σ to disassociate from the core RNAP in a process called promoter escape. The RNAP becomes a TEC that continues to transcribe the gene. Elongation continues until the TEC encounters a termination signal that triggers the formation of a TTC that causes the RNAP to release the RNA transcript and the DNA. Core RNAP and σ are recycled and the cycle starts over.

Bacterial promoters that are recognized by the primary σ generally comprise two sequence motifs located upstream of the +1 TSS: the -35 DNA element (consensus sequence: TTGACA) and the -10 DNA element (consensus sequence: TATAAT) (Hawley and Mcclure, 1983; Simpson, 1979). These DNA elements are separated by a spacer region ranging from 16-19 bps in length but with no general consensus sequence (Stefano and Gralla, 1982). A subset of these promoters also carry additional DNA elements such as the upstream (UP) element (Ross et al., 1993), extended -10 element (Barne et al., 1997) and the discriminator (Haugen et al., 2006).

The domain architecture of primary σ factors consists of three highly conserved domains [σ region 2 (σ_2), σ region 3 (σ_3), σ region 4 (σ_4)] and a poorly conserved N-terminal region ($\sigma_{1.1}$) that are connected by flexible linkers (Campbell et al., 2002). Free σ cannot bind DNA because $\sigma_{1.1}$ acts as an autoinhibitory domain that masks the DNA binding domains of σ (Dombroski et al., 1993; Schwartz et al., 2008). Binding of σ to RNAP to form holoenzyme extends σ and organizes its domain architecture for promoter DNA binding and melting (Callaci et al., 1998, 1999; Murakami et al., 2002b; Vassylyev et al., 2002) **(Figure 1.2)**.

During promoter binding, the UP element (located 40-60 bases upstream of the TSS) is recognized by the C-terminal domains of the RNAP α subunits (αCTD) (Ross et al., 1993). The -35 element is bound to σ_4 and the -10 element is bound to σ_2 (Campbell et al., 2002; Murakami et al., 2002a). If present, the extended -10 element (TG at -14 and -13) is recognized by σ_3 (Barne et al., 1997). During RPo formation, the discriminator region (between -6 and -3) makes contacts with $\sigma_{1.2}$ (Haugen et al., 2006) and the core-recognition element (CRE, located at +2) binds in a pocket of the β subunit (Zhang et al., 2012). These interactions are schematically shown in **Figure 1.2.** The combination of these DNA elements and their contacts with the holoenzyme can modulate the strength of a promoter by contributing to the initial RNAP binding and subsequent steps leading to RPo formation (Haugen et al., 2008; Ruff et al., 2015a).

1.3 Structures studies of prokaryotic RNAP

X-ray crystal structures of RNAP from *Eco* (Bae et al., 2013; Murakami, 2013; Zuo and Steitz, 2015), Thermus [*Thermus aquaticus* (*Taq*) and *Thermus thermophilus* (*Tth*)] (Bae et al., 2015b, 2015a; Murakami et al., 2002a, 2002b; Vassylyev et al., 2002; Weixlbaumer et al., 2013; Zhang et al., 1999, 2012) and Mycobacteria [*Mycobacterium smegmatis* (*Msm*) and *Mycobacterium tuberculosis* (*Mtb*)] (Hubin et al., 2017a, 2017b; Lin et al., 2017) have been fundamental in understanding the bacterial transcription cycle in great detail. These structural studies were performed on the most stable, discrete states of the bacterial transcription cycle: the RNAP core, holoenzyme, RPo, RPitc, and TEC. Structures of these states provide us insight into the RNAP and nucleic acid structures at each these stages of the transcription cycle.

The overall RNAP core structure is reminiscent of a crab claw, with one pincer comprising primarily the β subunit, and the other the β' subunit (Zhang et al., 1999). The active site, comprised of the catalytic Mg^{2+}, sits in a large cleft between the β and β' pincers. The bridge helix (BH) bridges the two pincers across the cleft creating two channels: the main channel where nucleic acid substrate resides, and secondary channel where rNTPs enter the RNAP active site **(Figure 1.3A)**. Opposite of the cleft is the RNA exit channel where the nascently transcribed RNA is expelled during elongation **(Figure 1.3A)**. The catalytic core RNAP is conserved in sequence, structure and basic function in all cellular organisms (Lane and Darst, 2010a, 2010b; Minakhin et al., 2001; Polyakov et al., 1995). The multi-subunit RNAPs from the three domains

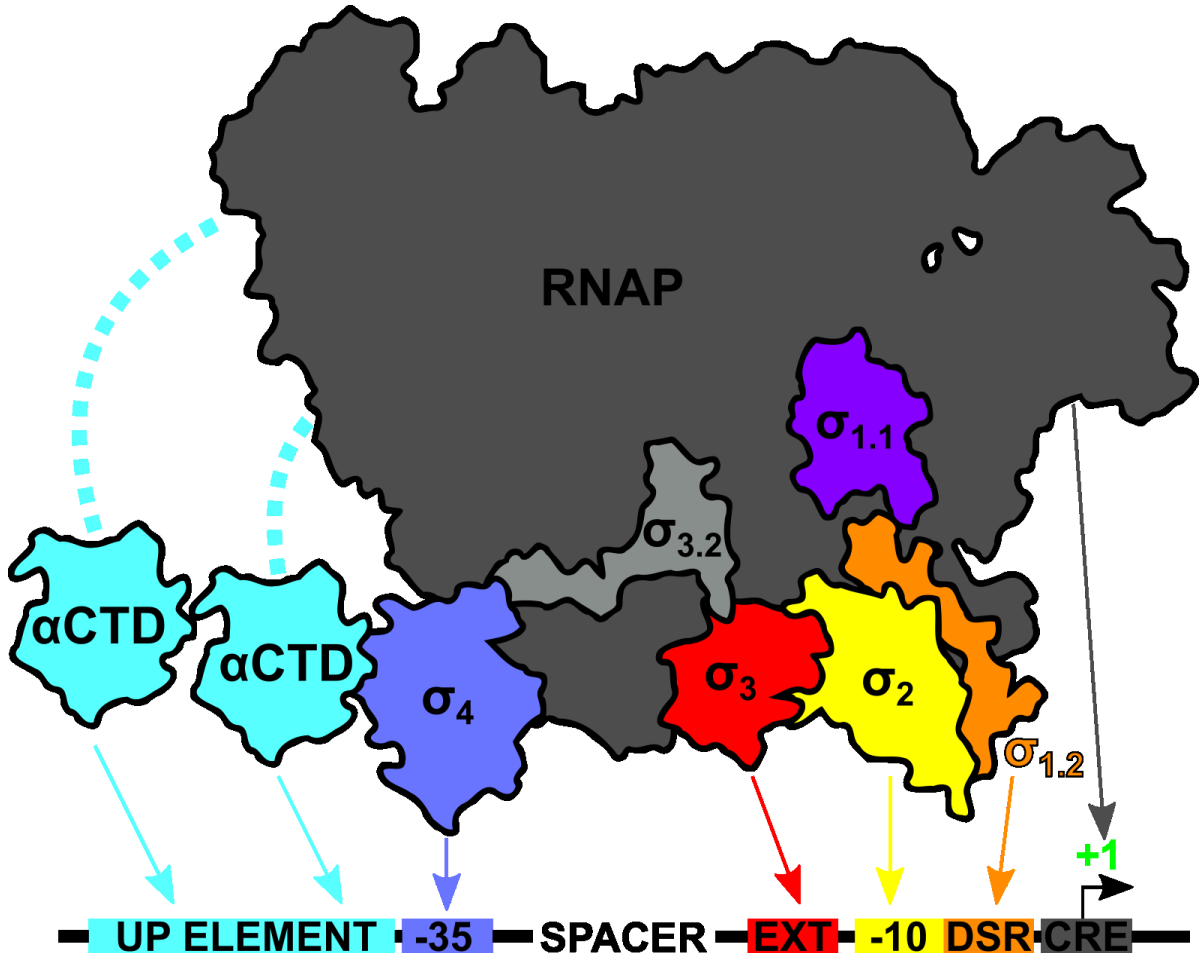


Figure 1.2. Promoter recognition by RNAP holoenzymes bound to the primary σ
RNAP bound to the primary σ generate holoenzymes that specifically recognize sequence elements in housekeeping promoters. RNAP is shown in dark grey and the domain organization of primary σ's is depicted. The flexible linkers that attached αCTDs to α subunits of RNAP are shown as teal dotted lines. The regions of holoenzyme and the promoter elements that these regions interact with are highlighted and color-coded: αCTDs bind the UP element [teal], σ_4 binds -35 element [slate], $\sigma_{3.2}$ [grey] is buried in the RNAP and does not recognize any promoter elements, σ_3 interacts with the extended -10 [EXT, red], σ_2 interacts with the -10 [yellow], $\sigma_{1.2}$ interacts with the discriminator region [DSR, orange], and the β subunit of RNAP recognizes the core recognition element [CRE, dark grey]. The TSS is shown as a green '+1' and the direction of transcription is depicted by the black arrow below the TSS.

of life share a common overall structure with the active site cleft comprised of conserved structural elements including the BH, the trigger loop (TL), the fork loop (FL) and the switches (Murakami, 2015).

Structures of holoenzyme reveal that σ makes extensive interactions with the β and β' subunit **(Figure 1.3B)**. The DNA binding domains of σ are spatially arranged in the holoenzyme for engagement with promoter DNA elements (Murakami et al., 2002a). Structures of RPo and RPitc demonstrate how the melted promoter DNA engages with the holoenzyme **(Figure 1.3C)**. In RPo, the -35 element interacts with σ_4 and the -10 element is engaged with σ_2, marking the upstream edge of the transcription bubble. The t-strand of the transcription bubble is loaded into the active site cleft, while the nontemplate-strand (nt-strand) of the transcription bubble sits between the RNAP clamp module and the βlobe. The downstream edge of the bubble and the remaining duplex DNA exit the main channel (Bae et al., 2015a; Zuo and Steitz, 2015).

These RPo structures also show key bases that interact specifically with RNAP. The universally conserved nt-strand adenine at position -11 [A_{-11}(nt)] flips into a pocket of σ_2 and is then stabilized by the repositioning of two universally conserved tryptophan residues in σ_2 (W-dyad) from an 'edge-on' to a 'chair'-like conformation (Bae et al., 2015a; Feklistov and Darst, 2011). The nearly absolutely conserved T_{-7}(nt) is captured in another pocket of σ_2 (Bae et al., 2015a; Feklistov and Darst, 2011). Additional bases that are captured by RNAP including G_{-5}(nt) within the discriminator region with $\sigma_{1.2}$ (Feklistov et al., 2006; Haugen et al., 2006; Zhang et al., 2012) and the G_{+2}(nt) CRE in a pocket in the β subunit (Zhang et al., 2012).

The downstream portion of the melted transcription bubble is stabilized by extensive interactions with mobile modules of the RNAP β and β' subunits. These modules include the clamp domain and βlobe (Bae et al., 2015a; Zuo and Steitz, 2015) and are common to all bacterial RNAPs. The β and β' subunits have been conserved throughout evolution, containing 16 and 11 shared sequence regions, respectively (Lane and Darst, 2010b). However, these shared regions are often separated by spacers of non-conserved lineage-specific inserts (LSI), that range from 50 to 500 amino acids in size. These inserts are typically independently-folded, located on the surface of the RNAP and are highly mobile (Lane and Darst, 2010b).

1.3.1 Mobile domains in bacterial RNAPs

RNAP is composed of a static core module and mobile domains that play critical roles in the function of RNAP (Murakami et al., 2002a). Since open complex formation in bacteria is driven by binding free energy (no ATP required), these flexible modules contribute to transcription initiation by allowing RNAP conformations that facilitate promoter melting (Saecker et al., 2011).

Early crystal structures defined a mobile structural element of the RNAP termed the clamp (Gnatt et al., 2001). The RNAP switch regions connect the clamp domain to the rest of the RNAP like hinges (Cramer et al., 2001; Gnatt et al., 2001). Single-molecule FRET was implemented to observe the RNAP clamp and its opening/closing dynamics in solution (Chakraborty et al., 2012; Duchi et al., 2018). These dynamics play an important role in promoter melting for all cellular RNAPs (Boyaci et al., 2019a; Chakraborty et al., 2012; Feklistov et al., 2017; He et al., 2013; Schulz et al., 2016). Conformational changes of the mobile βlobe also play a role in allowing DNA access to the RNAP cleft (Boyaci et al., 2019a; Chen et al., 2010c). These mobile domains are highlighted in **Figure 1.4B** for *Eco* RNAP and *Mtb* RNAP.

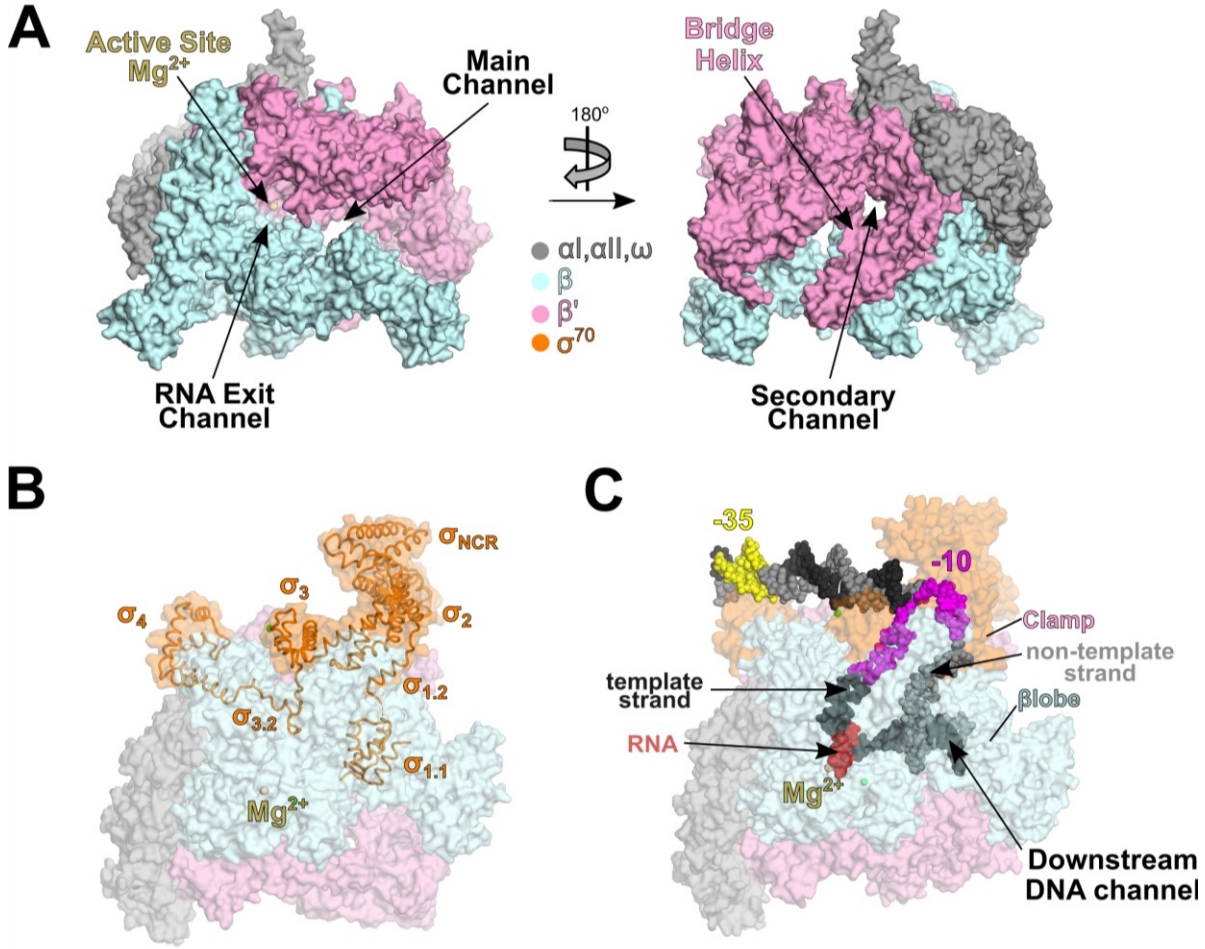


Figure 1.3. Structures of prokaryotic RNAP
Structural models of RNAP generated from crystal structures of Eσ^{70} (PDB ID 4LK1; Bae et al., 2013) and Eσ^{70} bound to promoter DNA as an initiation complex are shown (PDB ID 4YLN; Zuo and Steitz, 2015). **(A)** Overall structural of Eco core RNAP (using 4LK1, with σ^{70} removed for clarity) showing views of the active site Mg^{2+} (sphere colored wheat), main channel, RNA exit channel, and secondary channel. RNAP core is rendered as a molecular surface and colored according to the color legend (αI,αII, ω: grey; β: light cyan; β': pink; σ: orange. **(B)** σ^{70} domain organization is shown for Eσ^{70} (4LK1). Colored based on the legend, Eσ^{70} is rendered as a molecular surface, the active site Mg^{2+} is shown as a sphere, and σ^{70} overlaid and shown as cartoon tubes. **(C)** Structure of transcription initiation complex (4YLN) showing loading of the promoter DNA into RNAP for transcription. Eσ^{70} is shown as a molecular surface with promoter DNA shown as spheres. The active site Mg^{2+} is shown as a wheat sphere, nt-strand DNA is light gray, t-strand DNA is dark gray, -10 is yellow, -35 is magenta and the transcribed RNA is red. The RNAP clamp (pink) and the βlobe (light cyan) are indicated.

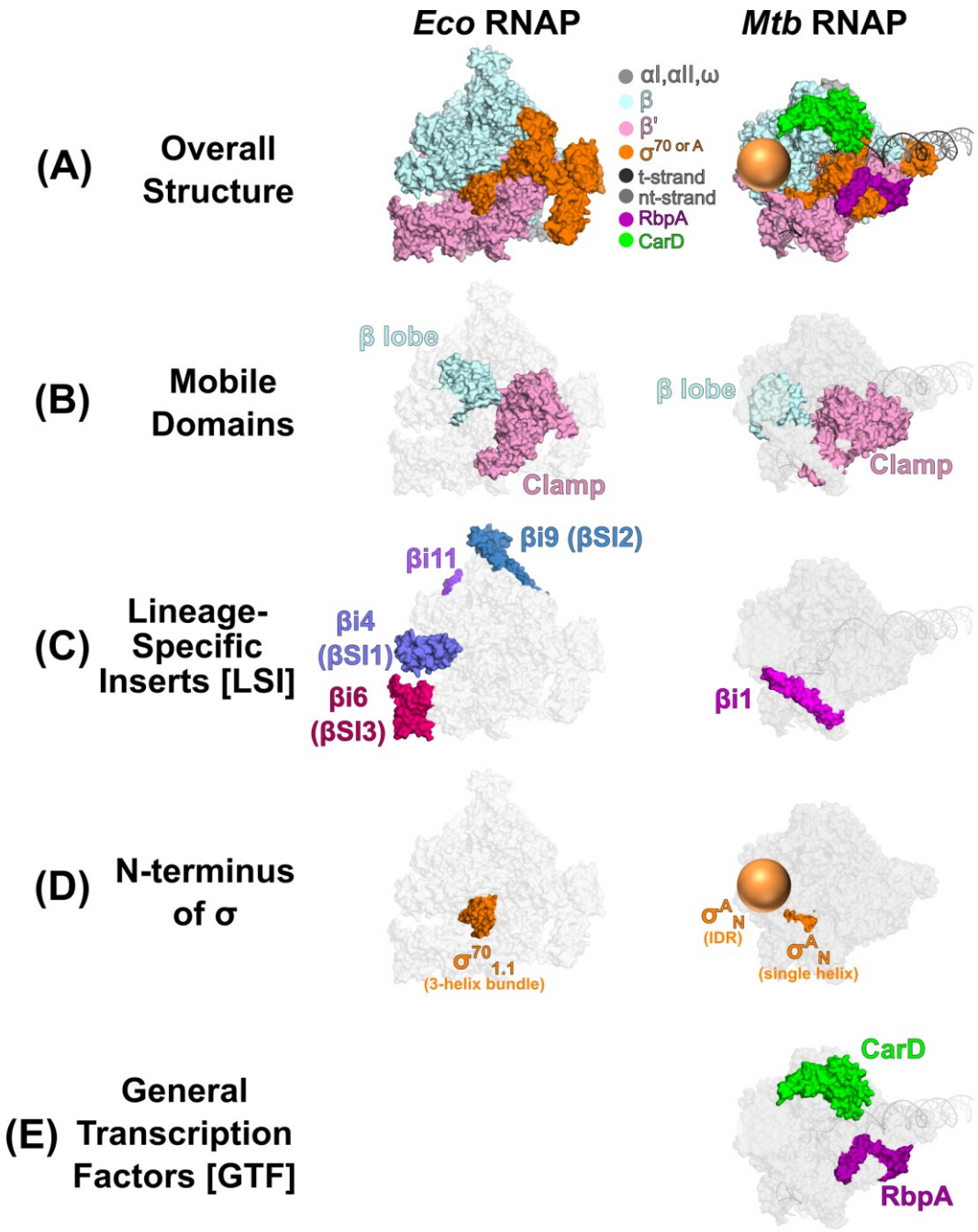


Figure 1.4. Comparison of *Eco* RNAP and *Mtb* RNAP
A crystal structure of Eσ^{70} (PDB ID 4LK1; Bae et al., 2013) was compared to a cryo-electron microscopy (cryo-EM) structure of a complete transcription initiation complex (TIC) with *Mtb* RNAP (PDB ID 6EDT; Boyaci et al., 2019a) in order to highlight differences between the two bacterial RNAPs. **(A)** The overall structure of *Eco* RNAP and *Mtb* RNAP (protein shown as a molecular surface, DNA is shown in cartoon format) colored according to the color legend. **(B)** Mobile domains are highlighted in color (βlobe is light cyan, the clamp is pink) while the rest of RNAP is grey. **(C)** LSIs are highlighted in color while the rest of the RNAP is grey. **(D)** N-terminus of σ is highlighted in orange while the rest of the RNAP is grey. The IDR of *Mtb* σ^{A}_{N} is shown as a orange sphere. **(E)** GTF are highlighted in color while the rest of the RNAP is grey.

1.3.2 Comparison of *Eco* RNAP and *Mtb* RNAP

In my thesis, I will examine the structures of *Eco* and *Mtb* RNAPs. *Eco* belongs to the proteobacteria phylum, which is distinct and evolutionarily distant from that of actinobacteria, the phylum to which *Mtb* belongs. Despite being divergent species, RNAPs from these bacteria overall are structurally similar **(Figure 1.4A-B)** with the exception of the LSIs **(Figure 1.4C)**, differences in their housekeeping σ **(Figure 1.4D)**, and the presence of general transcription factors (GTF) **(Figure 1.4E)**.

Eco has four LSIs that are not present in mycobacteria: βi4, βi9, βi11 and β'i6 (Lane and Darst, 2010a; Opalka et al., 2010). These inserts in *Eco* RNAP are highlighted in **Figure 1.4C (left).** The βi4 and β'i6 sequence insertions are important for initiation by contributing to RPo stability as well as regulation (Artsimovitch et al., 2003). βi4, also known as βSI1 (Artsimovitch et al., 2003), is an insertion within the βlobe located at the downstream entrance of the main channel. The deletion of βSI1 supports basic *in vitro* transcription function (Severinov et al., 1994). However, βSI1 is needed for normal cell growth and growth under nutrient starvation (Artsimovitch et al., 2003; Parshin et al., 2015).

β'i6, also known as β'SI3 (Artsimovitch et al., 2003), is an insertion within the TL, a key structural element in the RNAP nucleotide addition cycle that is conserved in all multi-subunit RNAPs (Lane and Darst, 2010a). β'SI3 plays a central role in *Eco* RNAP function and the deletion of β'SI3 is not viable (Artsimovitch et al., 2003; Zakharova et al., 1998). This insert is known to be highly mobile, moving to accommodate folding and unfolding of the TL at each RNAP nucleotide addition cycle (Kang et al., 2018a; Malinen et al., 2012; Zuo and Steitz, 2015).

Unlike *Eco*, *Mtb* RNAP contains one LSI, located on the β' subunit (β'i1). Initial structures of *Msm* and *Mtb* RNAP reveal the β'i1 insert emerges from the tip of the RNAP clamp module and extends across the entrance RNAP main channel (Hubin et al., 2017a, 2017b; Lin et al., 2017). This insert in *Mtb* RNAP is highlighted in **Figure 1.4C (right).** The insert is coupled to clamp conformations but its role remains unknown (Boyaci et al., 2018, 2019a).

The N-terminal region of the primary σ in *Eco* and *Mtb* bears no sequence or structural similarity. *Eco* $\sigma^{70}_{1.1}$ is composed of a compact three-helix domain attached to the rest of σ^{70} by a flexible linker (Bae et al., 2013; Schwartz et al., 2008). The position of *Eco* $\sigma^{70}_{1.1}$ is shown in **Figure 1.4D (left)**. Single-molecule and structural studies show that *Eco* $\sigma^{70}_{1.1}$ occupies the RNAP active-site channel but is ejected upon entry of the downstream duplex DNA in RPo (Bae et al., 2013; Mekler et al., 2002). $\sigma^{70}_{1.1}$ has promoter-dependent effects on steps in initiation (Ruff et al., 2015a; Vuthoori et al., 2001; Wilson and Dombroski, 1997). In contrast, the N-terminus of *Mtb* σ^{A} (σ^{A}_{N}) is comprised of a single α-helix followed by an intrinsically disordered region (IDR). It is located proximal to the entrance of the active site and the β'i1 (Hubin et al., 2017a, 2017b). The position of *Mtb* σ^{A}_{N} is shown in **Figure 1.4D (right)**.

Mycobacterium RNAP forms unstable open complexes (Davis et al., 2015). Two transcription factors, CarD and RbpA **(Figure 1.4E)**, found in *Mycobacterium* but not in *Eco*, are essential for transcription initiation, leading to the formation and stabilization of RPo (Newell et al., 2006; Paget et al., 2001; Stallings et al., 2009). Biochemical and structural studies have extensively characterized CarD and RbpA. CarD binds the protrusion domain of RNAP via an RNAP interacting domain (RID). CarD stabilizes RPo by wedging a conserved tryptophan residue (W85 in the case of *Mtb* CarD) into the upstream edge of the transcription bubble (Bae et al., 2015b; Srivastava et al., 2013). Crystal structures of RbpA reveal the factor's RNAP binding determinants: the C-terminal sigma interacting domain (SID) binds σ^{A}_{2} (Hubin et al., 2015) and its non-conserved region (NCR) and the core domain (CD) binds the β' Zn binding domain

(ZBD) (Hubin et al., 2017b). This positions the basic linker (BL) of RbpA near the DNA backbone just upstream of the TSS in order to anchor the promoter DNA to RNAP (Hubin et al., 2017b).

1.4 Mechanism of transcription initiation

Structures of bacterial RNAPs have failed to provide snapshots of the intermediate states of RNAP-promoter complexes during transcription initiation. A step-wise model for transcription initiation in bacteria based on structures of the final state (RPo) is outlined in **Figure 1.5** (adopted from Murakami and Darst, 2003). Transcription initiation begins with the recognition of promoter DNA by holoenzyme to form the closed complex, RPc (Kovacic, 1987). In RPc, the duplex DNA makes contact with σ_4 at the -35 element and with σ_2 at the -10 element (Campbell et al., 2002; Murakami et al., 2002a). After binding the DNA, the RNAP-promoter complex isomerizes and undergoes a conformational intermediate to begin melting and loading the DNA into the active site cleft (Buc and McClure, 1985; Saecker et al., 2002). In this intermediate, base-specific interactions between the RNAP and -10 element play a critical role in promoter unwinding (Cook and DeHaseth, 2007; Feklistov and Darst, 2011). As the transcription bubble is formed, RNAP makes key interactions with A_{-11}(nt) (Bae et al., 2015a; Feklistov and Darst, 2011), T_{-7}(nt) (Bae et al., 2015a; Feklistov and Darst, 2011), and the discriminator region (Bae et al., 2015a; Feklistov et al., 2006; Haugen et al., 2006; Zhang et al., 2012). The DNA is fully melted from -11 to +2 to form the transcriptionally competent RPo. In the presence of NTPs, RNAP begins synthesizing RNA (Ritc), releases σ, escapes the promoter and continues to transcribe the DNA as a TEC (Henderson et al., 2017; Krummel and Chamberlin, 1989; Munson and Reznikoff, 1981; Murakami and Darst, 2003; Vo et al., 2003). Kinetic and footprinting studies have been used to further dissect the steps of transcription initiation and to characterize the RNAP-promoter intermediates that populate them.

During initiation, the formation of RPo is a multi-step process, with the RNAP-promoter complex passing through multiple intermediates before RPo is formed (Hubin et al., 2017b; Ruff et al., 2015a; Saecker et al., 2002). Bacterial holoenzymes use binding free energy to isomerize into RPo (Ruff et al., 2015a; Saecker et al., 2011). Therefore, the sequence of the promoter plays a critical role in the rates of each step leading to RPo, resulting in differential kinetics between promoters that lead to varying transcriptional output (Haugen et al., 2008; Ruff et al., 2015a).

Initial kinetic studies with $E\sigma^{70}$ on the lacUV5 and T7A1 promoters demonstrated that RPo formation can be described by a minimal three-step mechanism: an initial "closed complex" intermediate (RPc); a second intermediate melting state (RP2); and then the final transcriptionally-competent open promoter complex (RPo) (Buc and McClure, 1985; Ruff et al., 2015a; Saecker et al., 2011). More recently, detailed kinetic studies with *Mycobacterium* σ^A-holoenzyme also revealed that RPo formation is best explained by a three-step mechanism for the *rrnA* P3 and *vapB* promoters (Hubin et al., 2017b).

Although a minimal three-step mechanism provides a working model for how DNA is melted by bacterial RNAP, additional intermediates likely occupy this pathway. For example, the intermediate states I1, I2, and I3 observed with the λPR promoter are distinct from the intermediates previously mentioned (Roe et al., 1984; Saecker et al., 2002). Additionally, other kinetic analyses used urea, salt, and other perturbants to identify other intermediates of RPo formation (Gries et al., 2010; Kontur et al., 2008, 2010). Due to their transient nature, intermediate states are difficult to capture and characterize structurally.

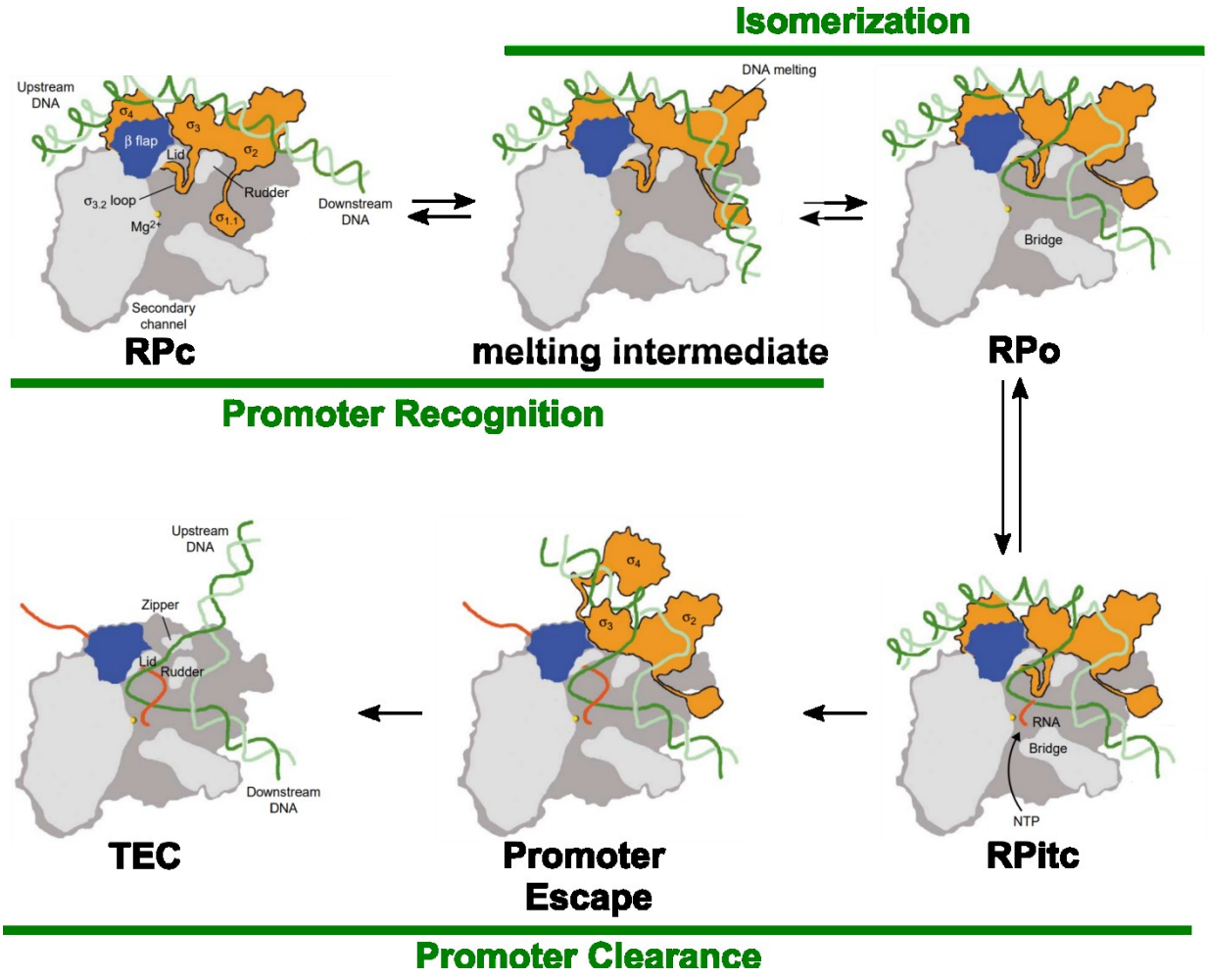


Figure 1.5. Steps for bacterial transcription initiation
2D cartoon depicting the RNAP-promoter structural changes during transcription initiation. Colored as follows: RNAP, light grey; β flap; blue; σ, orange; t-strand, dark green; nt-strand, light green; and RNA, red. Promoter recognition begins with promoter binding and the formation of RPc. An intermediate forms where the DNA is bent and melting of the -10 begins. The RNAP-promoter complex isomerizes and DNA is further melted and loaded into the RNAP active site to form the RPo. In the presence of NTPs, RNAP begins transcribing RNA (Ritc) until it generates a transcript long enough to remove $\sigma_{3.2}$ from the RNA exit channel. Once removed and σ is stochastically lost, RNAP escapes the promoter and continues to transcribe the DNA as a TEC. Adopted from Murakami and Darst, 2003.

1.5 Protein crystallography is a powerful but limited technique for RNAP structural biology

X-ray crystal structures of bacterial RNAPs have been foundational in our understanding of the bacterial transcription cycle. This technique can be used to determine high-resolution structures of biological complexes from a broad range of molecular weights (Shi, 2014). Protein crystallography serves as a powerful tool for RNAP structural biology but has several limitations: 1) crystallization of biological samples is difficult and 2) crystal packing artifacts affect structures making them affecting the interpretation (Luo et al., 2015).

Crystallization is the main bottleneck for samples studied by X-ray crystallography. To favor crystallization, samples have to be very homogenous in terms of biochemical purity and conformational heterogeneity. Impurities from a sample can inhibit successful nucleation and growth of crystals as well as crystal quality by disfavoring crystal packing interactions (Kors et al., 2009; Moreno et al., 2005). Thus, large multi-subunit samples that lack structural symmetry are more difficult to capture by crystallography.

Samples with conformational flexibility or transitional states are also difficult to study by X-ray crystallography because they are less likely to form an array of repeating units required for crystal packing. To reduce conformational heterogeneity in a sample, the flexible parts of a complex are often deleted. In crystal structures of *Eco* RNAP, the αCTDs and $\sigma^{70}_{1.1}$ are thought to be highly flexible and protein constructs harboring deletions of these domains are often used to encourage crystal formation (Bae et al., 2013; Zuo and Steitz, 2015).

X-ray crystallography has only been successfully implemented in the study of stable, long-lived states of bacterial RNAPs. To enrich these discrete states, nucleic scaffolds such as DNA forks or non-complementary transcription bubbles have been used to capture structures of RPo and elongation complexes (Bae et al., 2015a; Hubin et al., 2017b; Weixlbaumer et al., 2013; Zhang et al., 2012; Zuo and Steitz, 2015). These manipulations to the complex limit the interpretability of the crystal structures as they do not represent the de novo formed complex.

Once crystallized, molecules of a given complex arrange into an array of repeating asymmetric units that pack together to form the crystal. This packing with neighboring molecules affects the structure of the complex by influencing native conformations and restricting motions of mobile domains. These crystal packing artifacts affect the interpretability of the structures. Examples of crystal packing can be seen in a number of bacterial transcription complexes **(Figure 1.6)**. Packing of symmetry-related molecules with the highly mobile clamp domain has been observed in crystal structures of *Thermus* RNAP (Vassylyev et al., 2002; Weixlbaumer et al., 2013) **(Figure 1.6A-B)**. More recently, crystal structures of $E\sigma^{70}$ have been solved with crystal packing around the clamp and βlobe (Murakami, 2013) **(Figure 1.6C)**.

Despite these caveats, X-ray crystallography provides a wealth of information about bacterial RNAP structure and function. These structures have been instrumental in our understanding of RNAP interactions with nucleic acids and transcription factors, making it possible to address many mechanistic questions.

1.6 Single particle cryo-EM has many advantages for the study of bacterial RNAP

Cryo-EM is a technique to study the 3D structure of biological samples embedded in a layer of vitreous ice. Recent improvements in electron detectors and computing enable near-atomic resolution structures of large biological complexes to be solved by single particle cryo-EM (Bai et al., 2015a; Danev et al., 2019). Unlike X-ray crystallography, cryo-EM samples are

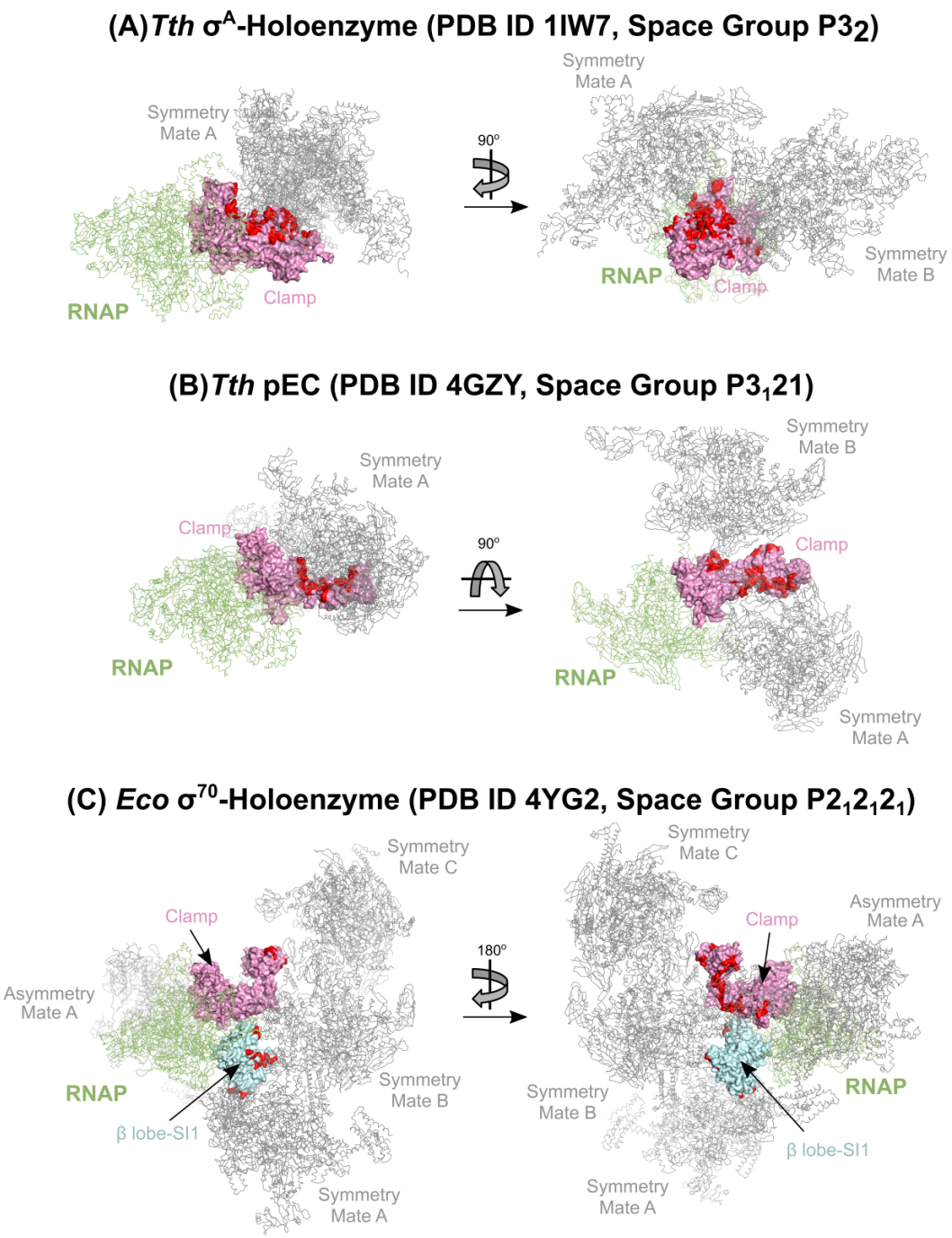


Figure 1.6. Crystal packing in bacterial RNAP structures determined by X-ray crystallography

Crystal packing artifacts affect the structure of RNAP. RNAP is shown as ribbons. The RNAP of interest is color in smudge green and neighboring RNAP molecules (asymmetric and symmetric) are shown in grey. Mobile RNAP domains impacted by packing are shown as molecular surfaces (clamp: pink, β lobe: light cyan) with clashes (distance cutoff of 4 Å) are highlighted in red. **(A)** *Tth* σ^A-holoenzyme (PDB ID 1IW7; space group $P3_2$; Vassylyev et al., 2002. **(B)** *Tth* paused elongation complex (pEC; PDB IF 4GZY; space group $P3_121$; Weixlbaumer et al., 2013. **(C)** $E\sigma^{70}$ (PDB ID 4YG2; space group $P2_12_12_1$; Murakami, 2013.

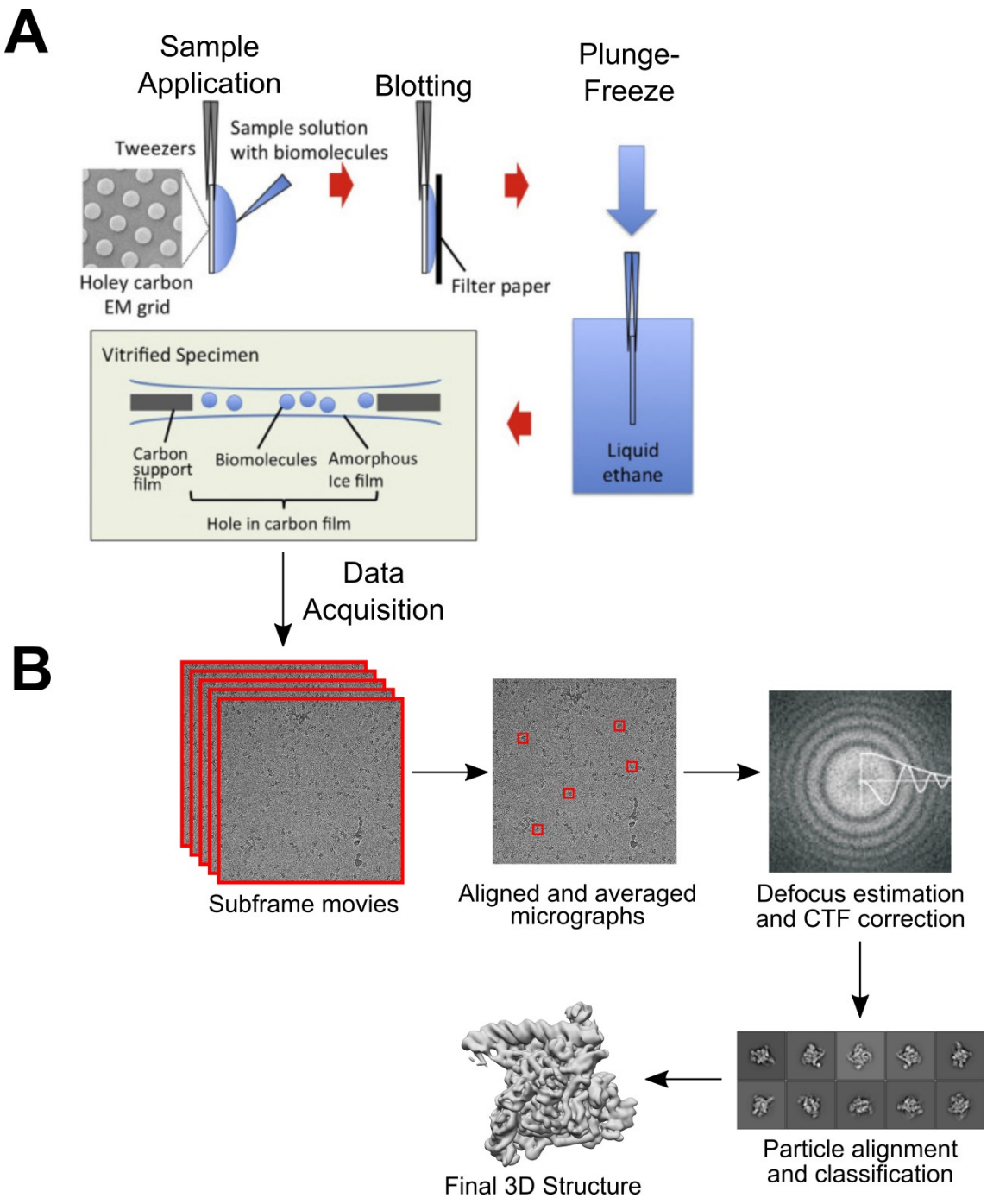


Figure 1.7. Single particle cryo-EM workflow
(A) A solution containing a biological sample is applied on a holey carbon EM grid. Excess solution is blotted off and rapidly plunge-frozen into liquid ethane to generate a vitrified specimen over the holes of the grid. Adopted from Murata and Wolf, 2018. **(B)** Subframe movies of the sample are acquired on a transmission electron microscope (TEM) at cryogenic temperatures. These movies are aligned and averaged and particles are picked. The defocus is estimated and the CTF is corrected for each image. The particles are extracted from each image for subsequent alignment and classification, leading to the final 3D reconstruction.

rapidly plunge-frozen **(Figure 1.7A)** from solution to cryogenic temperatures without the need for crystallization (Dobro et al., 2010). Macromolecular complexes can be studied in native solvent conditions and are free in solution, allowing them to adopt conformations that are not impacted by neighboring molecules. Cryo-EM allows direct visualization of these complexes and provides insight into their native states (Dobro et al., 2010) **(Figure 1.7B)**.

In addition, cryo-EM does not require very pure, homogenous samples. Image classification tools in cryo-EM processing packages can be used to sort heterogeneity that may be present in a particular sample (Grigorieff, 2007; Punjani et al., 2017; Scheres, 2012; Tang et al., 2007). Image classification is very robust and can sort heterogeneity across many levels ranging from particles that are missing components to different conformations of the particles. This allows one to study the structure of large complexes with multiple components as well as complexes that adopt multiple states or conformations.

Cryo-EM can be applied to study the solution structure of large complexes with multiple components, which is challenging for crystallization. Because crystallization is not limiting, it is easier to add many factors to a complex and study their structure. Large complexes are well suited for single particle cryo-EM since they have a better signal-to-noise ratio (SNR) (Danev et al., 2019). Sub-100 kilodaltons (kDa) complexes, which have poorer SNR, are more difficult to resolve to high-resolutions using cryo-EM (Henderson, 1995).

Cryo-EM processing software can distinguish multiple conformations in one sample using image classification (Grigorieff, 2007; Punjani et al., 2017; Scheres, 2012; Tang et al., 2007). Various implementation of classification algorithms such as masked 3D classification with partial signal subtraction (Bai et al., 2015b) and multi-body refinement (MBR) with principal component analysis (PCA) (Nakane et al., 2018) provide robust ways to sort structural heterogeneity in a sample. Therefore, cryo-EM allows insight into the molecular conformations and dynamics of a complex to an extent not possible in static crystal structures.

Cryo-EM is an excellent method to study structures and conformational movements of bacterial RNAP transcription complexes. The catalytic core is ~400 kDa, providing good SNR for 2D and 3D alignments. Many transcription factors and nucleic acids bind directly to RNAP to regulate its function (Browning and Busby, 2004). RNAP is also conformationally dynamic with many mobile domains (Chakraborty et al., 2012; Murakami et al., 2002a; Opalka et al., 2010). In addition, RNAP undergoes many transient states as it progresses through transitions of the transcription cycle. Single particle cryo-EM provides a robust approach to study all these aspects of the bacterial transcription cycle.

In this thesis, I optimized and utilized cryo-EM methodologies to structurally characterize bacterial RNAP complexes in *Eco* and *Mtb*. These include: (1) $E\sigma^{70}$ in complex with the *Eco* non-coding RNA (ncRNA) 6S RNA; (2) *Mtb* RNAP bound to the small molecule, fidaxomicin (Fdx), that inhibits transcription; (3) $E\sigma^{70}$ bound to the F element-encoded TraR protein; and (4) $E\sigma^{70}$-dependent promoter DNA melting intermediates stabilized by the TraR transcription factor. Using cryo-EM, I captured RNAP structures that were intractable to crystallization, visualized multiple RNAP conformational states populated in solution, deconvoluted RNAP molecular motions, and observed transient complexes. The work in this thesis establishes cryo-EM as a powerful tool to study bacterial transcription complexes.

1.7 Acknowledgments

I would like to thank the following people for their contributions to this chapter: E. A. Campbell and S. A. Darst (The Rockefeller University) provided insight, edits, and revisions.

CHAPTER 2. Bacterial RNAP sample preparation for single-particle cryo-EM

2.1 Introduction

Recent advances in electron detectors and software allow near atomic-resolution structures of biological macromolecular complexes to be solved by single-particle cryo-electron microscopy (cryo-EM) (Cheng, 2018). However, a major issue for many biological cryo-EM specimens is particle orientation bias (Glaeser, 2015; Glaeser and Han, 2017). Specimens that suffer this bias have an anisotropic distribution of angular projections leading to under-sampling of Fourier components and an overall loss of structural information in the final 3D reconstruction. This under-sampling leads to a "smearing" artifact (Barth et al., 1989; Naydenova and Russo, 2017; Tan et al., 2017), which affects the reliability and interpretability of the cryo-EM maps. The primary cause of preferred particle orientation is presumed to be due to the adsorption of the particles at the air/water interface during cryo-EM grid preparation (Taylor and Glaeser, 2008).

In collaboration with the Walz Lab (The Rockefeller University), we examined a complex of *Eco* 6S RNA bound to *Eco* RNA polymerase (RNAP) σ^{70}-holoenzyme (referred here as 6S-Eσ^{70}) by single-particle cryo-EM. Detailed analysis of the 6S-Eσ^{70} binary complex will be presented in **Chapter 3**. Initial single-particle cryo-EM experiments showed a severe orientation bias for this complex. To solve the orientation bias problem, we explored a range of solution conditions and detergent additives and discovered that the detergent 3-([3-Cholamidopropyl]dimethylammonio)-2-hydroxy-1-propanesulfonate (CHAPSO) was uniquely effective.

We hypothesized that the preferred orientation was due to adsorption of the particles to the air-water interface (Taylor and Glaeser, 2008) and that CHAPSO mitigated this problem by preventing adsorption at the interface. In collaboration with A. J. Noble (New York Structural Biology Center), we used fiducial-less cryo-electron tomography (cryo-ET) on the single-particle specimens to visualize particle distributions within the vitreous ice (Noble and Stagg, 2015; Noble et al., 2018a). The results showed that the orientation bias arose from interactions of the particles with the surfaces of the ice layer. In the presence of a sufficient concentration of CHAPSO, the particles were excluded from the ice surfaces and distributed within the ice layer with nearly random orientations. Portions of this chapter have been previously published in the Journal of Structural Biology: X (Chen et al., 2019a).

2.2 6S-Eσ^{70} particles exhibit severe orientation bias which is eliminated with CHAPSO

Initial cryo-EM analysis of the 6S-Eσ^{70} complex was performed using a potassium L-glutamate buffer (KGlu, **Table 2.1**). Cryo-EM micrographs [collected with Z. Li (Harvard University) from the Walz Lab (The Rockefeller University)] showed uniform, monodisperse particles that yielded detailed 2D classes (**Figure 2.2A**, **Table 2.3**). However, 3D alignment of these particles resulted in "smeared" maps (**Figure 2.2B**, **Table 2.3**), suggesting particle orientation bias. The angular distribution plot of the particles revealed a distribution corresponding to essentially one orientation (**Figure 2.4A**). A rough characterization of the particle distribution as a Gaussian yielded a peak at about rotation angle (rot) -19°, tilt angle (tilt) -13°, and a standard deviation (SD) (±) of 15°.

To overcome this particle orientation bias, we prepared samples in a number of alternative conditions and screened the samples for complex stability using an electrophoretic mobility shift assay and negative stain EM (**Table 2.5**). Triton X-100 and Tween 20 (at critical micelle concentration, CMC) affected the stability of the 6S-Eσ^{70} complex and were not explored

Table 2.1 Solution conditions for 6S-Eσ^{70} cryo-EM analyses

Abbreviation	Solution components	[RNAP] (μM)
KGlu	10 mM Tris-HCl, pH 8.0, 150 mM K-glutamate, 1 mM $MgCl_2$, 2 mM DTT	0.9
KCl	10 mM Tris-HCl, pH 8.0, 150 mM KCl, 1 mM $MgCl_2$, 2 mM DTT	0.9
KGlu+NP40S	10 mM Tris-HCl, pH 8.0, 150 mM K-glutamate, 1 mM $MgCl_2$, 2 mM DTT, 0.06 mM Nonidet P40 substitute (NP40S)	6
KCl+CHAPSO	10 mM Tris-HCl, pH 8.0, 150 mM KCl, 1 mM $MgCl_2$, 2 mM DTT, 8 mM CHAPSO	9

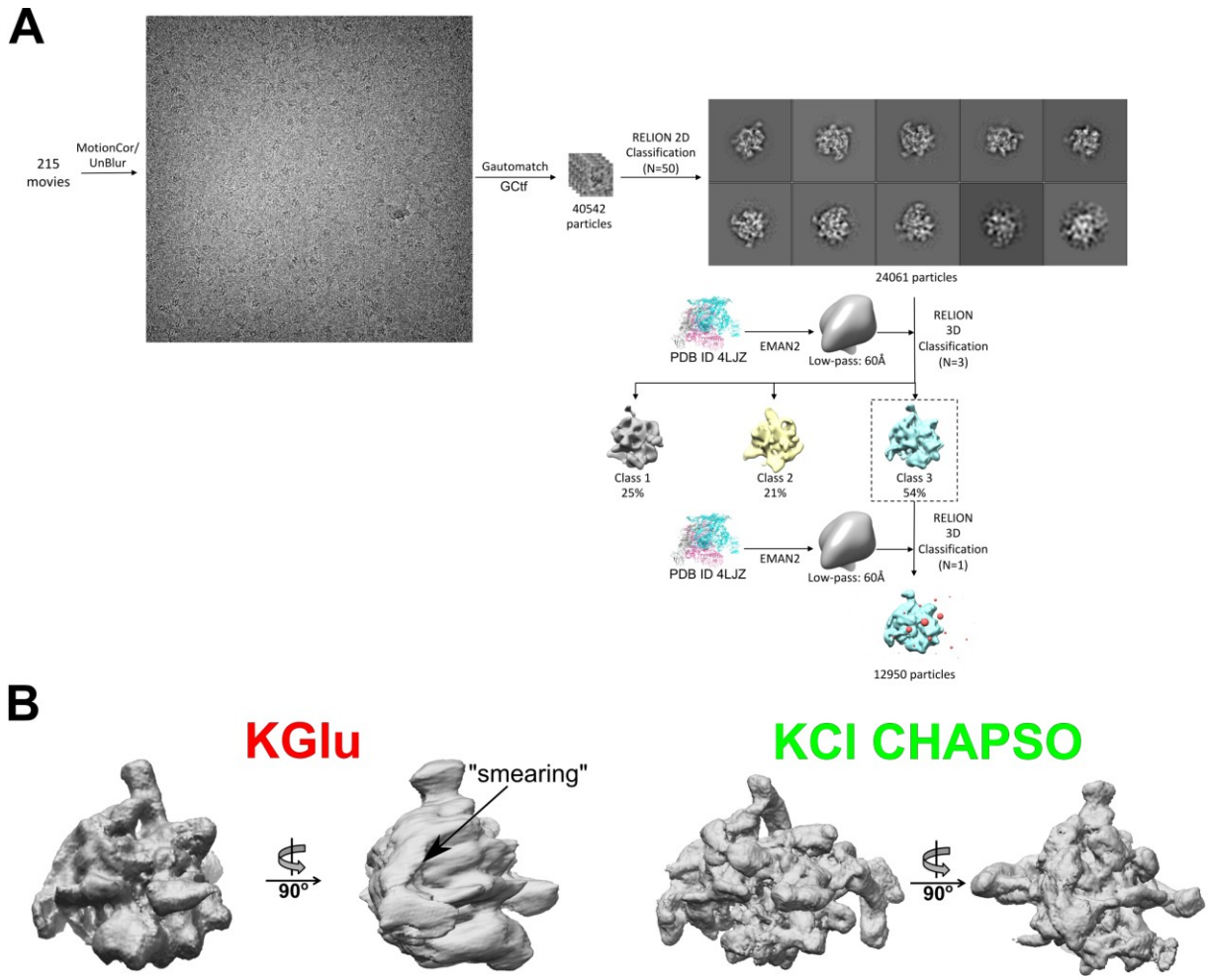


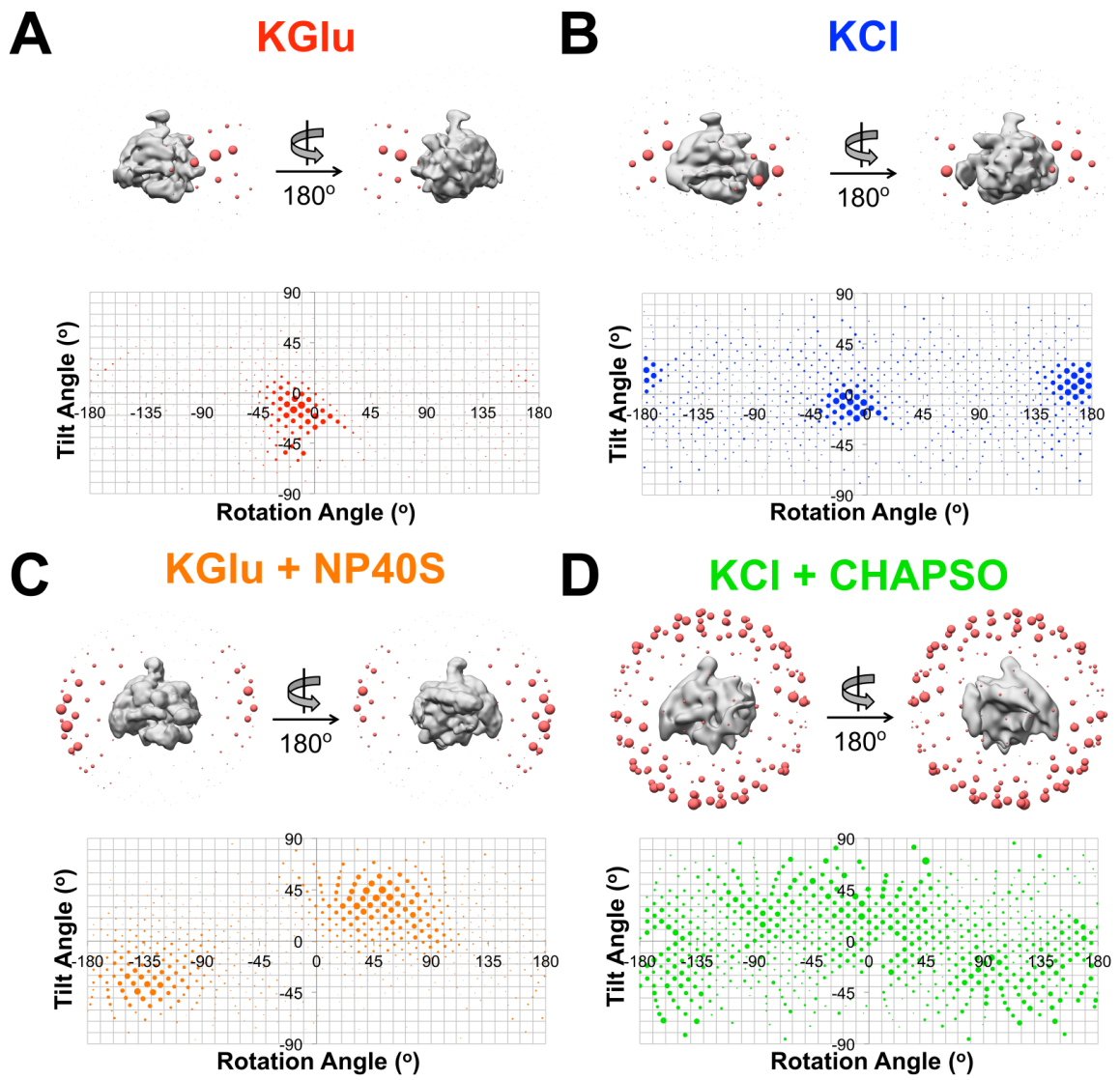
Figure 2.2. Cryo-EM data collection and processing of 6S-Eσ^{70}
(A) Cryo-EM processing pipeline for 6S-Eσ^{70} in KGlu. This processing scheme was applied to the other single-particle cryo-EM datasets (KCl, KGlu+NP40S, KCl+CHAPSO, TEC, TEC+4mM CHAPSO, and TEC+8mM CHAPSO). The final set of particles were aligned in RELION (Scheres, 2012) 3D classification (N = 1) using a 3D template (PDB ID 4LJZ; Bae et al., 2013) to determine angular orientations. **(B)** Demonstration of the "smearing" effect in maps generated from particles with preferred orientation. (*left*) RELION (Scheres, 2012) 3D reconstruction of particles from KGlu showing "smearing" in the plane opposite to that of the preferred orientation. (*right*) RELION (Scheres, 2012) 3D reconstruction of particles from KCl + CHAPSO shows an isotropically uniform map.

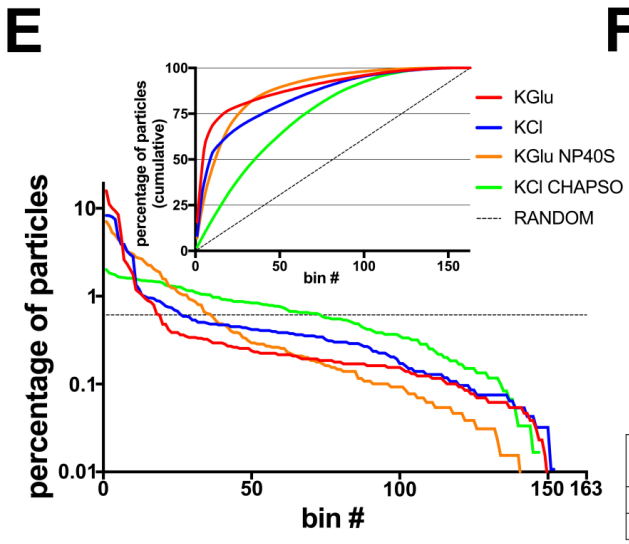
Table 2.3. Data collection and reconstruction information for single-particle cryo-EM datasets (6S-Eσ^{70})

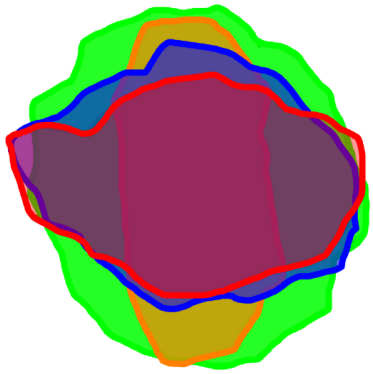
	KGlu	KCl	KGlu+NP40S	KCl+CHAPSO
Sample				
Buffer	10mM Tris-HCl pH 8.0, 150mM KGlu, 1mM $MgCl_2$, 2mM DTT	10mM Tris-HCl pH8.0, 150mM KCl, 1mM $MgCl_2$, 2mM DTT	10mM Tris-HCl pH8.0, 150mM KGlu, 1mM $MgCl_2$, 2mM DTT, 0.06mM NP40S	10mM Tris-HCl pH8.0, 150mM KCl, 1mM $MgCl_2$, 2mM DTT, 8mM CHAPSO
Grid type	Quantifoil	Quantifoil	Quantifoil	Quantifoil
Data Collection				
Microscope	Tecnai G2 Polara	Tecnai G2 Polara	Tecnai G2 Polara	Tecnai G2 Polara
Voltage (keV)	300	300	300	300
Nominal Magnification	31000x	31000x	31000x	31000x
Number of frames/movie (no.)	30	30	30	30
Exposure rate (e-/pixel/sec)	8	8	8	8
Total exposure time (sec)	6	6	6	6
Electron dose (e-/Å^2)	31.7	31.7	31.7	31.7
Detector	K2 Summit	K2 Summit	K2 Summit	K2 Summit
Recording mode	Counting	Counting	Counting	Counting
Pixel size (Å)	1.23	1.23	1.23	1.23
Defocus Range (μm)	-1.3 to -1.7	-1.4 to -2.2	-1.5 to -2.3	-1.4 to -2.3
Reconstruction				
Number of micrographs (no.)	215	62	147	290
Number of particles picked (no.)	40542	26114	33987	39002
Symmetry Imposed	C1	C1	C1	C1
Number of particles after 2D (N=50) (no.)	24061	20487	29153	12540
Number of particles after 3D (N=3) (no.)	12950	9332	16695	5978

Figure 2.4. Single-particle cryo-EM analysis of 6S-Eσ^{70} particle orientation distributions in KGlu, KCl, KGlu + NP40S, and KCl + CHAPSO

(A-D) (*top*) 3D distribution plot of particle orientations. Particles were aligned with a 3D template of *Eco* RNAP (PDB ID 4LJZ; Bae et al., 2013). The resulting 3D reconstruction is shown as a solid grey volume and the angular particle distribution is shown as red spheres. Each sphere represents a particular Euler angle and the sphere volume represents the absolute number of particles at that particular angle. (*bottom*) 2D distribution plot of particle orientations. Particles are plotted on a tilt angle vs rotation angle graph. The size of points represents the percentage of particles at that particular orientation**. (A)** KGlu (red). **(B)** KCl (blue). **(C)** KGlu + NP40S (orange). **(D)** KCl + CHAPSO (green). **(E)** Particles for each dataset were grouped into Euler angle bins (20° rotation angle × 20° tilt angle bins) and then the bins were ranked according to the number of particles populating that bin (bin #1 has the most particles, so on). Plotted on a semi-log scale is the percent of total particles in each dataset by bin #. The horizontal dashed line represents a totally random particle orientation distribution (equal number of particles in each bin). (*inset*) Plotted is the cumulative percent particles by bin #. The random distribution is denoted by the dashed line. (**F**) Cross-sections through the middle of the expected PSFs (calculated using cryoEF Naydenova and Russo, 2017) are superimposed, illustrating the anisotropy for the KGlu (red), KCl (blue), and KGlu + NP40S (orange) samples, while the KCl + CHAPSO sample yields an isotropic PSF (green). Parameters further characterizing the orientation distributions (the orientation efficiency, E_{od}; and the fraction of unsampled Fourier space, f_{empty}; Naydenova and Russo, 2017) are also tabulated. Data courtesy of Z. Li (Harvard University) from the Walz Lab (The Rockefeller University).







	KGlu	KCl	KGlu+ NP40S	KCl+ CHAPSO
E_{od}	0.46	0.54	0.55	0.78
f_{empty}	0.37	0.16	0.07	0.067

Table 2.5. Screening of cryo-EM solution conditions for 6S-Eσ^{70}

Solution Conditions	EMSA[a]	Negative Stain EM	Cryo-EM
10mM Tris-HCl pH 8.0, 150 mM KGlu, 1 mM $MgCl_2$, 2 mM DTT	+	+	+
10 mM Tris-HCl pH 8.0, 150 mM KGlu, 1 mM $MgCl_2$, 2 mM DTT, 0.15 mM DDM[b]	+	-	-
10 mM Tris-HCl pH 8.0, 150 mM KGlu, 1 mM $MgCl_2$, 2 mM DTT, 0.18 mM Triton™ X-100	-	-	-
10 mM Tris-HCl pH 8.0, 150 mM KGlu, 1 mM $MgCl_2$, 2 mM DTT, 0.04mM TWEEN® 20	-	-	-
10 mM Tris-HCl pH 8.0, 150 mM KGlu, 1 mM $MgCl_2$, 2 mM DTT, 0.06 mM NP40S	+	+	+
10 mM Tris-HCl pH8.0, 150 mM KGlu, 1 mM $MgCl_2$, 2 mM DTT, 8 mM CHAPSO	+	-	-
10 mM Tris-HCl pH8.0, 150 mM KCl, 1 mM $MgCl_2$, 2 mM DTT	+	+	+
10 mM Tris-HCl pH8.0, 150 mM KCl, 1mM $MgCl_2$, 2 mM DTT, 8 mM CHAPSO	+	+	+

[a] electrophoretic mobility shift assay
[b] n-Dodecyl-B-D-Maltoside
[†] + indicates experiments that show that the 6S-Eσ^{70} complex is stable in a given buffer condition, whereas – indicates experiments that show destabilization of the 6S-Eσ^{70} complex in a given buffer condition. Green highlights conditions used for single-particle cryo-EM.

further**.** The 6S-Eσ^{70} complex appeared to be stable with DDM, but negative stain EM images showed protein aggregation. From these results, we proceeded with the cryo-EM analysis of the four sample conditions (listed in **Table 2.1**).

Cryo-EM analysis of the 6S-Eσ^{70} particles in the KCl condition also showed severe orientation bias (**Figure 2.4B**, **Table 2.3**), with one peak at an orientation and spread very similar to the KGlu condition (rot -12°, tilt -8°, ±14°), but with an additional peak at approximately (rot 168°, tilt 8°), corresponding to a mirror image projection of the first orientation. The mirror image projection can be treated as a single orientation since it does not contribute any new information to the 3D reconstruction.

Particles in KGlu + NP40S (**Figures 2.4C**, **Table 2.3**) also yielded two mirror image peaks [(rot -130°, tilt -47°) and (rot 50°, tilt 47°)] with standard deviation ±20°. The particles in this condition had slightly less bias compared to particles in KGlu and KCl. In contrast to KGlu, KCl, and KGlu + NP40S, particles prepared in KCl + CHAPSO did not exhibit peaks of preferred orientation; instead, the particle orientations were spread over a large fraction of Euler angles (**Figure 2.4D**, **Table 2.3**), resulting in isotropically uniform 3D reconstructions (**Figure 2.2B**).

For each individual sample, the particles were grouped into bins according to their Euler angles and ranked according to the percentage of particles in each bin (**Figure 2.4E**). Visualizing the orientation distributions this way highlights the orientation bias of the KGlu, KCl, and KGlu + NP40S samples. The KCl + CHAPSO sample, while not completely randomized, approaches the random orientation distribution (dashed horizontal line, **Figure 2.4E**) more closely. A plot of the cumulative % of particles across the bins reveals that 50% of the particles are binned into only 5, 9 and 12 bins for the KGlu, KCl, and KGlu + NP40S conditions, respectively, while 50% of the KCl + CHAPSO particles are spread out over 36 bins (inset of **Figure 2.4E**).

The effects of particle orientation on resulting 3D reconstructions were also assessed using cryoEF (Naydenova and Russo, 2017). A cross-section through the middle of the expected point spread functions (PSFs) calculated from the particle orientation distributions reveals the severe anisotropy of the KGlu, KCl, and KGlu + NP40S PSFs while the KCl + CHAPSO PSF appeared as roughly a spherical ball (**Figure 2.4F**). The orientation efficiency (E_{od}) and the fraction of unsampled Fourier space (f_{empty}) scores were dramatically improved with CHAPSO (Naydenova and Russo, 2017).

2.3 Cryo-ET shows that orientation bias corresponds to adsorption to an ice surface

A. J. Noble (New York Structural Biology Center) applied fiducial-less cryo-ET on the single-particle cryo-EM specimens to visualize the locations of the particles in the vitreous ice layer (Noble and Stagg, 2015; Noble et al., 2018a). Tilt series of cryo-grids of the 6S-Eσ^{70} complex were collected for each solution condition (**Table 2.6**, **Figure 2.7**). The 6S-Eσ^{70} particles prepared in KGlu and KCl completely localized to one of the ice surfaces (**Figures 2.7A-B**). The 6S-Eσ^{70} particles prepared in KGlu + NP40S were restricted to two thin layers corresponding to both ice surfaces (**Figure 2.7C**). By contrast, particles prepared in KCl + CHAPSO were excluded from the air/water interfaces and were evenly distributed throughout the middle of the ice layer (**Figure 2.7D**).

2.4 Effect of CHAPSO on particle orientations is concentration dependent

Our results indicate that adsorption of the 6S-Eσ^{70} particles at air/water interfaces gives rise to the severe orientation bias seen in the KGlu, KCl, and KGlu + NP40S samples (**Figures 2.4A-C**,

Table 2.6. Data collection and reconstruction information for cryo-ET datasets (6S-Eσ^{70})

	KGlu	KCl	KGlu+NP40S	KCl+CHAPSO
Data collection				
Buffer	10mM Tris-HCl pH 8.0, 150mM KGlu, 1mM $MgCl_2$, 2mM DTT	10mM Tris-HCl pH8.0, 150mM KCl, 1mM $MgCl_2$, 2mM DTT	10mM Tris-HCl pH8.0, 150mM KGlu, 1mM $MgCl_2$, 2mM DTT, 0.06mM NP40S	10mM Tris-HCl pH8.0, 150mM KCl, 1mM $MgCl_2$, 2mM DTT, 8mM CHAPSO
Grid type	Quantifoil	Quantifoil	Quantifoil	Quantifoil
Microscope	Titan Krios	Titan Krios	Titan Krios	Titan Krios
Camera	K2 Summit	K2 Summit	K2 Summit	K2 Summit
Magnification	18,000x	18,000x	18,000x	18,000x
Recording mode	Counting	Counting	Counting	Counting
Voltage (keV)	300	300	300	300
Exposure navigation	Stage Position	Stage Position	Stage Position	Stage Position
Total dose (e-/Å^2)	118	118	118	118
Exposure rate (e-/pixel/sec)	8	8	8	8
Pixel size (Å)	1.33	1.33	1.33	1.33
Number of frames/movie (no.)	8-11	8-11	8-11	8-11
Nominal defocus (μm)	5	5	5	5
Tilt range and increment (˚)	-45:45:3	-45:45:3	-45:45:3	-45:45:3
Tilt-series collected	3	3	4	3
Particles in each tomogram shown in **Figure 2.7**	266	307	894	107

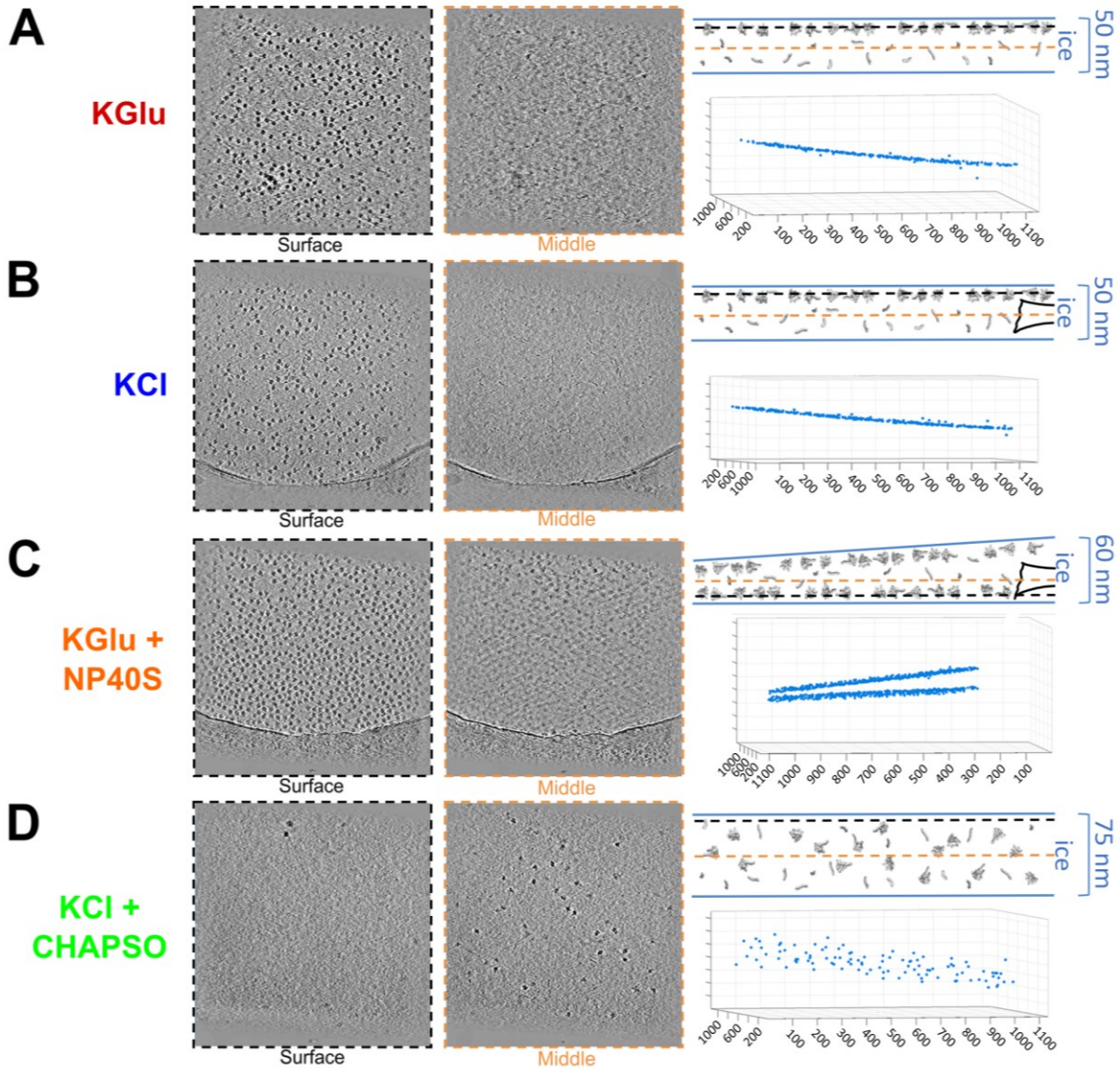


Figure 2.7. Cryo-ET reveals the mechanism for preferred orientation
(A-D) (*left panel*) Surface tomographic cross-section of the vitreous ice layer. (*middle panel*) Middle tomographic cross-section of the vitreous ice layer. (*top-right panel*) Schematic diagram of particle distribution in vitreous ice. The top and bottom surfaces of the ice are shown with a solid blue line. The thickness of ice is indicated on the bracket right of the cartoon. 6S-Eσ^{70} and free 6S RNA particles are shown as grey volumes in the cartoon. (*bottom-right panel*) A spatial plot of particles in the vitreous ice layer oriented orthogonal to the ice surface. Each 6S-Eσ^{70} particle is represented as a blue point and graphed based on 3D position in the ice layer. **(A)** KGlu. **(B)** KCl. **(C)** KGlu + NP40S. **(D)** KCl + CHAPSO. Data and figure courtesy of A. J. Noble (New York Structural Biology Center).

2.7A-C). The addition of CHAPSO at the CMC (8 mM) completely eliminates surface interactions causing the particles to evenly distribute throughout the ice and adopt random particle orientations (**Figures 2.4D**, **2.7D**). To investigate if CHAPSO at CMC was required for the full effect, we compared particle orientations for datasets of *Eco* transcription elongation complex (TEC) [courtesy of J. Y. Kang from the Darst-Campbell Lab (The Rockefeller University); Kang et al., 2017)] collected in 0 mM (0xCMC), 4 mM (0.5xCMC), and 8 mM (1xCMC) CHAPSO (**Table 2.8)**. In the absence of CHAPSO, the particles again exhibited severe orientation bias similar to KCl (**Figure 2.9A**). The particle orientations were significantly spread by the presence of 4 mM CHAPSO (**Figure 2.9A**), but there is a clear difference in the particle spread between 4 mM and 8 mM CHAPSO (compare **Figures 2.9B and 2.9C**) and 8 mM CHAPSO induces a distribution of particle orientations much closer to a random distribution (**Figures 2.9D-E**).

2.5 Cryo-EM maps reveal CHAPSO interacts with specific sites on the *Eco* RNAP surface

Examination of the nominal 3.5 Å resolution cryo-EM map of an *Eco* TEC bound to RfaH (Kang et al., 2018b), the highest resolution cryo-EM map available for an *Eco* RNAP transcription complex at the time this work was performed, revealed three CHAPSO molecules bound to the RNAP surface (**Figure 2.10**). Retrospective analysis of previously published cryo-EM structures of *Eco* RNAP transcription complexes where 8 mM CHAPSO was used to prevent orientation bias revealed CHAPSO molecules consistently bound at the same sites (Cartagena et al., 2019; Chen et al., 2017, 2019b; Kang et al., 2017, 2018a, 2018b). In each case, the cholic acid-derived amphifacial moiety of CHAPSO was bound to the RNAP while the attached flexible chain and the hydrophilic head group were disordered and lacked cryo-EM density.

2.6 Discussion

Eco RNAP transcription complexes adsorb strongly to the air/water interfaces during cryo-EM grid preparation (**Figures 2.7A-C**). Adsorption to this interface causes the particles to become oriented, confounding single-particle reconstruction approaches (**Figures 2.2B**, **2.4**, **2.9**). The addition of CHAPSO at the CMC (8 mM) completely prevents adsorption of the complexes to this interface (**Figure 2.7**). In CHAPSO, particles distribute throughout the ice layer in randomized orientations (**Figures 2.4D-F**, **2.7D**, **2.9C-E**), allowing for the determination of isotropically uniform maps (**Figures 2.2B**, **2.4D**, **2.9C**).

CHAPSO at 0.5 CMC (4 mM) significantly broadens the particle orientation distribution compared to no CHAPSO (**Figures 2.9A**, **2.9B**, **2.9D**). However, the calculated PSF and orientation parameters (**Figure 2.9E;** Naydenova and Russo, 2017) suggest CHAPSO at its CMC (8 mM; **Figure 2.9C**) is required to have the full effect on the relative randomization of the particle orientations. Due to the high CHAPSO concentration, we observed CHAPSO molecules consistently bound to three sites on the surface of *Eco* RNAP (**Figure 2.10**). Our cryo-ET results showed that in the presence of 8 mM CHAPSO, the RNAP complexes are completely excluded from the air/water interfaces and localized in the middle of the ice layer (**Figure 2.7D**) where rotational diffusion allows for nearly randomized particle orientations. Elimination of RNAP adsorption from the air/water interface could be due to: (1) CHAPSO coats the interface and alters its surface properties, and/or (2) CHAPSO molecules binding to the surface of RNAP (**Figure 2.10**), altering the RNAP surface properties to prevent interfacial adsorption. Our result

Table 2.8. Data collection and reconstruction information for single-particle cryo-EM datasets (TEC)

	TEC	TEC + 4mM CHAPSO	TEC + 8mM CHAPSO
Sample			
Buffer	20mM Tris pH8.0, 150mM KCl, 5 mM $MgCl_2$, 5mM DTT	20mM Tris pH8.0, 150mM KCl, 5 mM $MgCl_2$, 5mM DTT, 4mM CHAPSO	20mM Tris pH8.0, 150mM KCl, 5 mM $MgCl_2$, 5mM DTT, 8mM CHAPSO
Grid type	C-Flat	C-Flat	C-Flat
Data Collection			
Microscope	Titan Krios	Talos Arctica	Ttitan Krios
Voltage (keV)	300	200	300
Nominal Magnification	22500x	28000x	22500x
Number of frames/movie (no.)	50	50	50
Exposure rate (e-/pixel/sec)	10	10	10
Total exposure time (sec)	15	15	15
Electron dose (e-/Å^2)	88.6	66.7	88.6
Detector	K2 Summit	K2 Summit	K2 Summit
Recording mode	Super-resolution	Super-resolution	Super-resolution
Pixel size (Å)	1.3	1.5	1.3
Defocus Range (μm)	-1.5 to -3.5	-0.8 to -2.4	-2.5 to -5.0
Reconstruction			
Number of micrographs (no.)	500	558	491
Number of particles picked (no.)	68994	134546	65659
Symmetry Imposed	C1	C1	C1
Number of particles after 2D (N=50) (no.)	23960	41399	28130
Number of particles after 3D (N=3) (no.)	11410	29513	12066

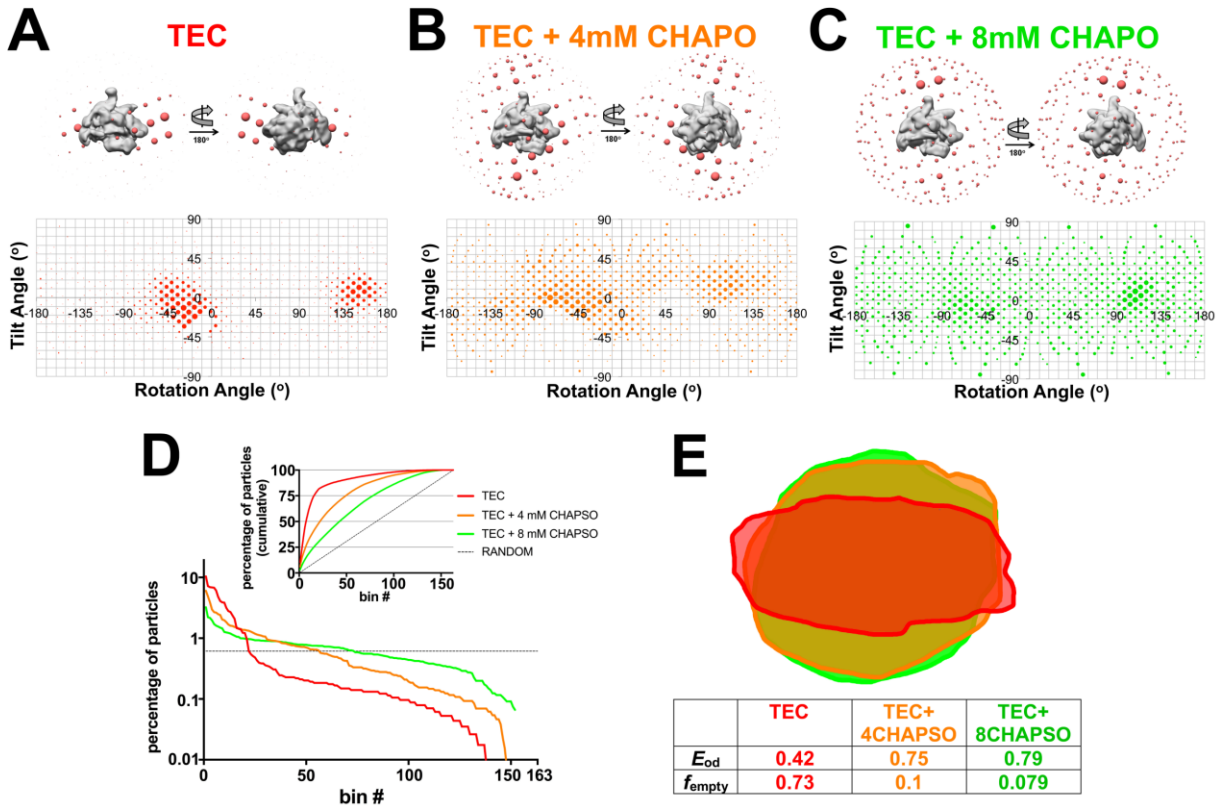


Figure 2.9. Effect of CHAPSO on particle orientations is concentration dependent
(A-C) See **Figure 2.6** for a description of angular distribution plots. **(A)** TEC particles (no CHAPSO). **(B)** TEC + 4 mM CHAPSO (0.5xCMC). **(C)** TEC + 8 mM CHAPSO (1xCMC). **(D)** Particles for each dataset were grouped into Euler angle bins. Plotted on a semi-log scale is the percent of total particles in each dataset by bin #. The horizontal dashed line represents a totally random particle orientation distribution. (*inset*) Plotted is the cumulative percent particles by bin #. The random distribution is denoted by the dashed line. **(E)** Cross-sections through the middle of the expected PSFs (calculated using cryoEF; Naydenova and Russo, 2017)) are superimposed. Parameters further characterizing the orientation distributions (E_{od} and f_{empty}; Naydenova and Russo, 2017) are also tabulated. Data courtesy of J. Y. Kang from the Darst-Campbell Lab (The Rockefeller University).

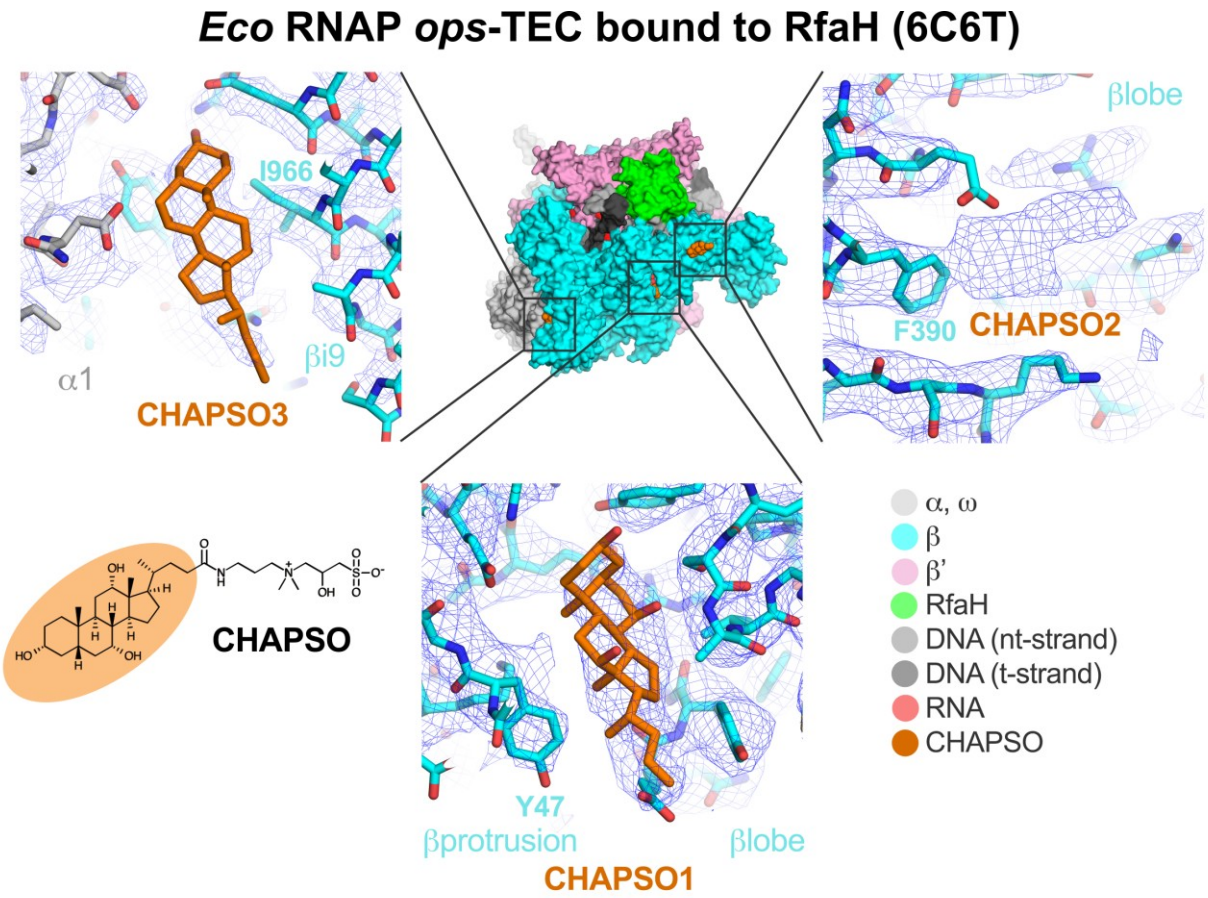


Figure 2.10. CHAPSO molecules interact with RNAP particles
(*top middle*) Overall view of the *Eco* RNAP ops-ternary elongation complex bound to RfaH (PDB ID 6C6T; Kang et al., 2018a). The structure is shown as molecular surfaces color-coded as shown in the color key at the lower right. Shown in orange are three CHAPSO molecules bound to the RNAP surface. (*lower left*) Molecular structure of CHAPSO. The portion highlighted in orange is resolved in the cryo-EM maps. (*bottom middle*) Magnified view showing CHAPSO1 along with the cryo-EM map. (*top right*) Magnified view showing CHAPSO2 along with the cryo-EM map. The cryo-EM density for CHAPSO2 was not of sufficient quality to determine the CHAPSO orientation. (*top left*) Magnified view showing CHAPSO3 along with the nominal 3.5 Å resolution cryo-EM map (blue mesh).

that 8 mM CHAPSO is required for the full effect (**Figure 2.9**) strongly supports the hypothesis (1) as the primary factor, since CHAPSO molecules are observed in RNAP at 4 mM CHAPSO (Chen et al., 2019a).

A number of cryo-EM structures of bacterial RNAP transcription complexes have benefitted from CHAPSO (Boyaci et al., 2018, 2019a; Cartagena et al., 2019; Chen et al., 2017, 2019b; Kang et al., 2017, 2018a, 2018b). The observation of specifically-bound CHAPSO molecules on the *Eco* RNAP surface (**Figure 2.10**) raises the potential for the introduction of structural artifacts. We proposed that the binding energy for these sites is very low and is insufficient to alter the conformational/functional properties of the RNAP. For instance, a conformational change in the *his* pause EC (*his* pEC) altered the relationship between βi9 and the rest of the RNAP, disrupting CHAPSO site 3 (**Figure 2.10**), and no complexes were observed with CHAPSO bound at this site. In biochemical assays, 8 mM CHAPSO had a minimal (less than 2-fold) effect on the pause lifetime (Kang et al., 2018a).

Orientation bias has long been an issue confounding single-particle cryo-EM structure determination (Glaeser, 2015; Glaeser and Han, 2017). Maps generated from oriented samples are limited in their interpretability due to the "smearing" artifact in the final 3D construction (**Figure 2.2B**). It has often been presumed that particle orientation bias arises from adsorption to the air/water interface. A common solution to mitigate particle orientation bias has been to add surfactants during grid preparation in order to reduce particle-interface adsorption (Taylor and Glaeser, 2008). Our cryoET analysis establishes that this is indeed the case for samples comprising *Eco* RNAP transcription complexes and the surfactant CHAPSO (**Figure 2.7**).

Recent studies have established that severe particle orientation bias can be overcome by cryo-EM data collection from tilted grids (Tan et al., 2017) or by rapid grid freezing after sample application (Noble et al., 2018b). Since rapid grid freezing requires special instrumentation, grid tilting is a more practical option to address orientation bias for most cryo-EM samples. However, tilting presents many technical obstacles for optimal high-resolution data collection. More importantly, it does not address the issue of particle adsorption, which has recently been suggested to cause denaturation for most particles (D'Imprima et al., 2019). Denatured particles can affect image classification leading to artificial structures that are not biologically relevant. The addition of CHAPSO during cryo-EM grid preparation of samples comprising *Eco* RNAP transcription complexes is relatively functionally inert, completely eliminates interaction and orientation of the particles at air/water interfaces (**Figure 2.7**), and significantly broadens the particle orientation distributions to allow determination of high-resolution cryo-EM maps with isotropic PSFs (**Figures 2.4F**, **2.9E**). In the subsequent chapters, I use CHAPSO to facilitate high-resolution structure determination of various *Eco* and *Mtb* RNAP transcription complexes by cryo-EM.

2.7 Acknowledgments

The work in this chapter is published in the Journal of Structural Biology: X. My role in this work involved: cryo-EM sample preparation; data collection and processing; structural analysis; and conception of ideas and figures. Citation:

Chen, J., Noble, A.J., Kang, J.Y., and Darst, S.A. (2019). Eliminating effects of particle adsorption to the air/water interface in single-particle cryo-electron microscopy: Bacterial RNA polymerase and CHAPSO. J. Struct. Biol. X *1*, 100005.

I would like to thank the following people for their contributions to this chapter: S. A. Darst (The Rockefeller University) supervised this project, assisted with data analysis, and edited this chapter; M. Ebrahim and J. Sotiris (The Rockefeller University) helped with single-particle cryo-EM data collection; J. Y. Kang (The Rockefeller University) prepared and collected single-particle cryo-EM data on TEC; Z. Li (Harvard University) and T. Walz (The Rockefeller University) helped with EM sample preparation and single-particle cryo-EM data collection of 6S-Eσ^{70}; A. J. Noble (New York Structural Biology Center) collected and processed cryo-ET data on 6S-Eσ^{70} samples.

CHAPTER 3. Structure determination of *Eco* 6S RNA in complex with *Eco* RNAP

3.1 Introduction

Bacterial 6S RNAs globally regulate transcription in a wide range of bacterial species (Barrick et al., 2005; Willkomm et al., 2005). 6S RNAs directly bind to RNA polymerase (RNAP) holoenzyme containing the primary or group I σ factor [$E\sigma^{70}$ in *Escherichia coli* (*Eco*)] but not holoenzymes containing alternative σ factors, thus sequestering the ability of $E\sigma^{70}$ to bind promoter DNA (Cavanagh and Wassarman, 2014; Wassarman and Storz, 2000)

Eco 6S RNA, a 184 nucleotide non-coding RNA (ncRNA), plays a key role in the transcriptional reprogramming upon transitions between exponential and stationary growth phases by specifically targeting and inhibiting the housekeeping holoenzyme, $E\sigma^{70}$. Previous, iron bromoacetamidobenzyl-EDTA (FeBABE) cleavage experiments have mapped specific interactions of the RNA to σ^{70} regions 2 ($\sigma^{70}{}_2$), 3 ($\sigma^{70}{}_3$), and 4 ($\sigma^{70}{}_4$) (Steuten et al., 2013). This preferential binding for $E\sigma^{70}$ allows for the increased activity of the stationary phase σ factor, σ^S (Steuten et al., 2014).

Although not conserved in sequence, 6S RNAs are widely conserved among bacteria through a common secondary structure that includes a prominent large central bubble that is reminiscent of the promoter DNA bubble in the transcription open promoter complex (Barrick et al., 2005; Trotochaud and Wassarman, 2005). Indeed, 6S RNA can serve as a transcription template during outgrowth from stationary phase, likely due to high nucleotide (NTP) concentrations (Wassarman and Saecker, 2006).

Physiological studies demonstrate that 6S RNA is present during both exponential and stationary phases of growth. However, 6S RNA accumulates in high amounts during stationary phase and serves as a substrate for synthesis of various lengths of product RNAs (pRNAs). The formation of these short pRNAs has been shown to disrupt binding of 6S RNA to RNAP, relieving inhibition (Beckmann et al., 2011; Cavanagh et al., 2012). Synthesis of a product RNA (pRNA) induces a structural rearrangement of the RNA that provokes release of $E\sigma^{70}$ from 6S RNA (Beckmann et al., 2012; Panchapakesan and Unrau, 2012), providing a mechanism for rapidly recovering $E\sigma^{70}$ activity in response to nutrient availability (Cavanagh et al., 2012). These two 6S RNA isoforms are shown in **Figure 3.1**. Through this mechanism, 6S RNA plays an important role for long term survival during stationary phase as well as escape from stationary phase back into exponential growth (Hsu et al., 1985; Lee et al., 1985; Trotochaud and Wassarman, 2004).

The structural basis for how 6S RNA interacts with the DNA-dependent RNAP to outcompete promoter DNA has not been elucidated. Previous attempts to crystallize 6S RNA with RNAP holoenzyme did not yield a high resolution structure of an intact complex. Since crystallizing the complex was difficult, I developed sample preparation methodology in **Chapter 2** (Chen et al., 2019a), in order to study bacterial RNAP complexes by single particle cryo-EM. Initial cryo-EM experiments were done in collaboration with the Walz Lab (The Rockefeller University). In this chapter, I report the 3.8 Å resolution structure of a complex between *Eco* 6S RNA and $E\sigma^{70}$ determined by cryo-EM. In collaboration with Karen Wassarman (University of Wisconsin–Madison) and other members of the Darst-Campbell Lab (The Rockefeller University), biochemical approaches were used to valid this structure. Portions of this chapter have been previously published in Molecular Cell (Chen et al., 2017).

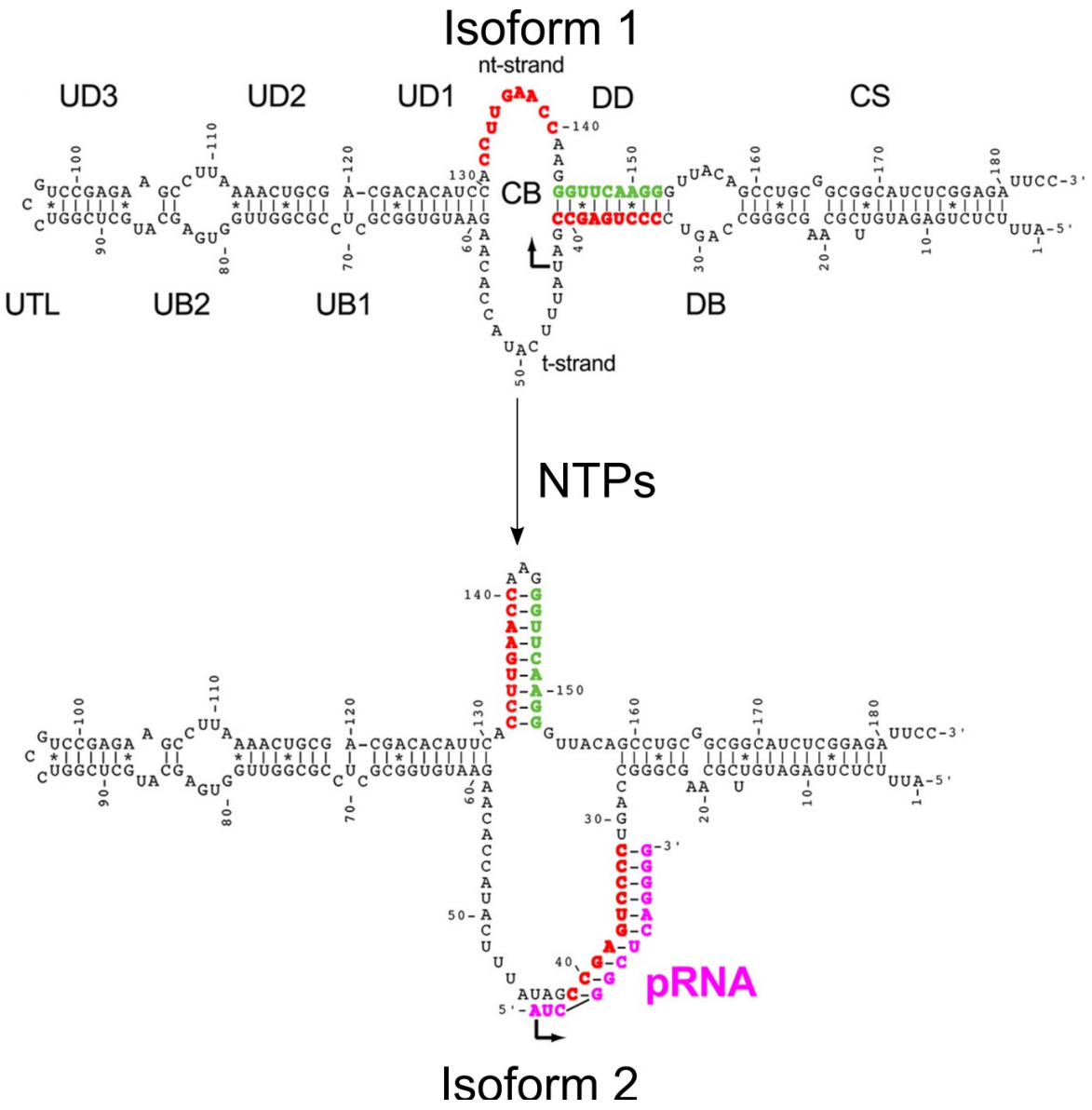


Figure 3.1. 6S RNA secondary structure and transitions
The pRNA-induced structural transition of 6S RNA (Beckmann et al., 2012; Panchapakesan and Unrau, 2012). (*top*) Secondary structure of wild-type (WT) 6S RNA isoform 1 (structural definitions: CS, closing stem; DB, downstream bulge; DD, downstream duplex; CB, central bubble; UD1-3, upstream duplexes 1-3; UB1-2, upstream bulges 1 and 2; UTL, upstream terminal loop) (*bottom*) Secondary structure of WT 6S RNA isoform 2. Isoform 2 forms a hairpin (green-red duplex) and is stabilized by the RNAP mediated synthesis of pRNA (magenta).

3.2 6S RNA footprinting and minimal binding region

To probe 6S RNA/Eσ^{70} interactions and minimal binding requirements, A. Feklistov from the Darst-Campbell Lab (The Rockefeller University) performed nucleic acid footprinting (Chen et al., 2017; summarized in **Figure 3.2A**). Ribonucleases that cleave single stranded RNA at the 3' of C and U (RNase A) and double stranded RNA (RNase V1) were used to probe 6S RNA interactions with Eσ^{70}. RNase footprints of full-length Eσ^{70} and a truncated Eσ^{70} [comprised of *Eco* RNAP with the C-terminal domains of the α subunits (ΔαCTD) deleted (ΔαCTD-E; Twist et al., 2011) and σ^{70} with region 1.1 ($\sigma^{70}_{1.1}$) deleted (Δ1.1σ^{70}; Bae et al., 2013)] were identical, indicating that the αCTDs and $\sigma^{70}_{1.1}$ did not play a role in 6S RNA interactions. The RNA upstream of the central bubble (CB; **Figure 3.1**) showed an alternating pattern of hypersensitivity and protection from RNases, suggesting interaction with the surface of the RNAP (**Figures 3.2A-B**). Additionally, 6S RNA CB and DD (see **Figure 3.1** for definitions of the 6S RNA structural elements) were protected from RNase treatment suggesting enclosure of these regions in the RNAP cleft. This was further illustrated by Fe^{2+}-directed cleavage demonstrating loading of uracil at position 44 (U44, see **Figure 3.1)** near the RNAP active site (Wassarman and Saecker, 2006; **Figures 3.2A-B**). RNase protection did not extend downstream beyond the DB region, indicating that the CS is outside of RNAP cleft and does not play a role in binding, as shown previously (Shephard et al., 2010).

6S RNA is predicted to exist in two nearly isoenergetic isoforms (**Figure 3.1**) even in the absence of pRNA synthesis. Isoform 1 is most compatible for binding with Eσ^{70}. To enrich for isoform 1, I shuffled the base sequence of the DD region of 6S RNA while maintaining complementary base pairing. This shuffling locked 6S RNA into a homogenous population of isoform 1 by preventing formation of isoform 2. This 6S RNA variant, 6S RNA*, (**Figure 3.2A**) was generated by combining a truncation of the CS and DD sequence shuffling (DD*). E. A. Campbell, S. Feng, and K. Leon from the Darst-Campbell Lab (The Rockefeller University) tested 6S RNA* binding and transcription. This 6S RNA construct bound Eσ^{70} specifically and initiated transcription as efficiently as wild-type (WT) 6S RNA from the correct transcription start site (TSS) (**Figure 3.2C**).

3.3 Cryo-EM analysis of *Eco* 6S RNA* bound to Eσ^{70}

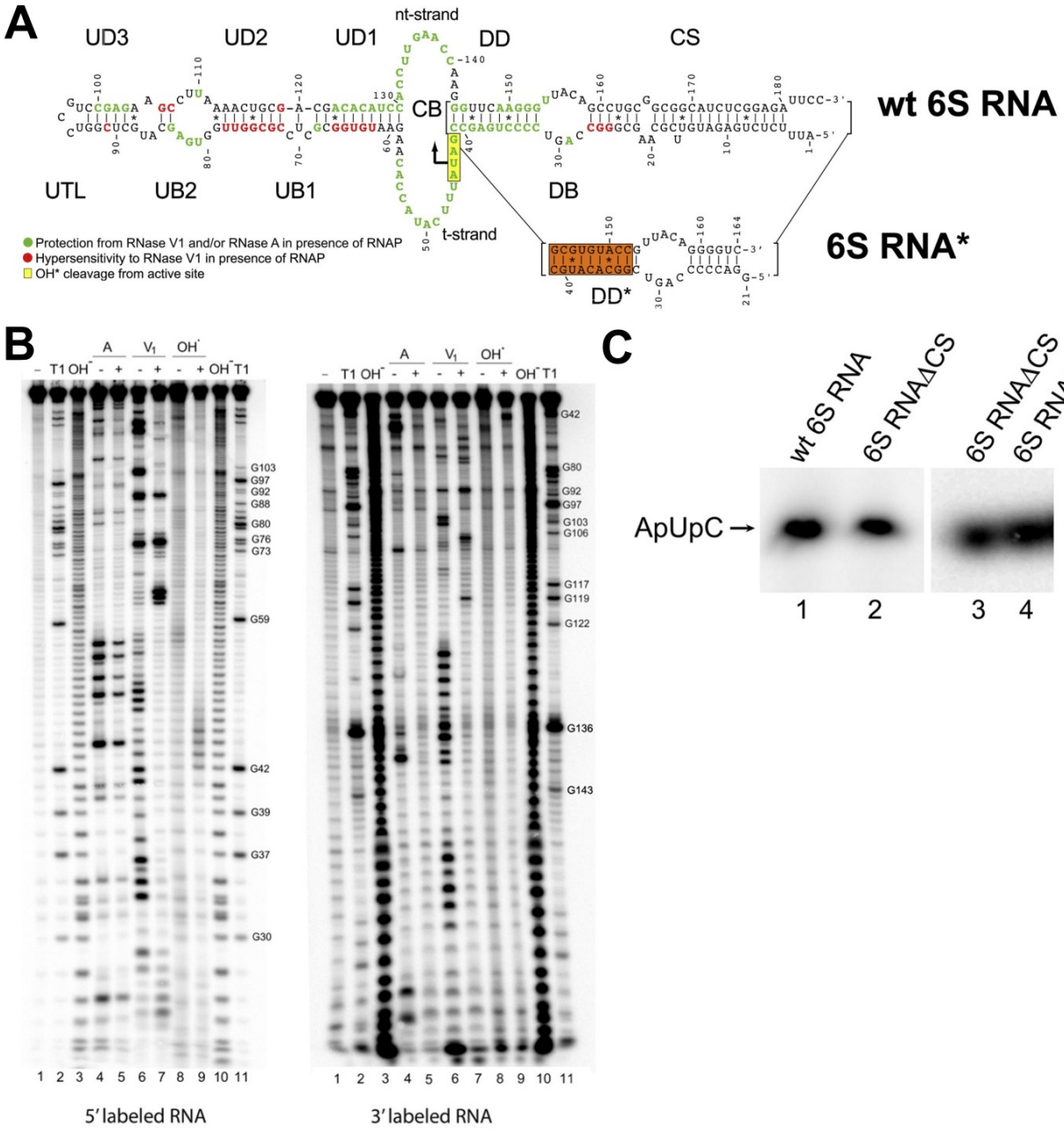
I generated and purified 6S RNA* using *in vitro* transcription and native purification approaches. RNase-free ΔαCTD-E and Δ1.1σ^{70} were assembled to make ΔαCTD-E-Δ1.1σ^{70} holoenzyme (termed Eσ^{70} for the rest of this chapter). 6S RNA* was added to generate the *Eco* 6S RNA-Eσ^{70} complex (termed 6S-Eσ^{70} in **Chapter 2** and in this chaper) for structure determination.

I used single-particle cryo-EM to determine the structure of the 6S-Eσ^{70} binary complex. After 3D classification and refinement (**Figure 3.3A**), the final class contained 362,926 particles with an overall resolution of 3.8 Å using the Gold-standard FSC cutoff of 0.143 in RELION (Scheres, 2012) (**Figure 3.3B**). Local resolution calculations for the final structure indicated that the central core of the structure was determined to 3.3–3.8 Å resolution (Cardone et al., 2013) (**Figure 3.3C**).

3.4 Modeling and validation of 6S-Eσ^{70} cryo-EM structure

I built and refined a model of the 6S-Eσ^{70} complex into the cryo-EM map using PHENIX (Adams et al., 2010) and a combination of RNA modeling tools such as Rcrane (Keating and Pyle, 2010) (**Table 3.4**). The 6S-Eσ^{70} cryo-EM map and model are shown in **Figure 3.5A.** The CB template-strand (t-strand) from nucleotides 48-57 and the downstream portion of the DD* of

Figure 3.2. 6S RNA construct for single particle cryo-EM
(A) (*top*) Secondary structure of WT Eco 6S RNA as observed in the cryo-EM structure. Structural elements of the 6S RNA are labeled based on **Figure 3.1**. The sequence is color coded according to the RNase footprinting and localized hydroxyl-radical cleavage results comparing 6S RNA with and without $E\sigma^{70}$. RNA positions protected from RNase V1 and/or RNase A cleavage in the presence of $E\sigma^{70}$ are colored green. RNA positions showing hypersensitivity to RNase V1 in the presence of $E\sigma^{70}$ are colored red. RNA positions efficiently cleaved by hydroxyl radicals generated from Fe^{2+} in place of the RNAP active site Mg^{2+} are boxed in yellow (Wassarman and Saecker, 2006). The position of pRNA synthesis initiation (U44) is indicated by a bent arrow. (*bottom right*) Secondary structure of substituted downstream stem of 6S RNA* used in cryo-EM structure determination, generated by truncating the CS and shuffling the sequence of the DD to give DD* (orange box). **(B)** RNAse footprinting gels: T1 shows cleavage by RNase T1 specific for G-residues, their positions are indicated on the right side of gels, A and V1 - cleavage by RNAse A and V1, respectively. OH* represents hydroxyl radical cleavage, OH^- - alkaline hydrolysis under single hit conditions, resulting in an RNA 'ladder' with single nucleotide resolution. (*left*) 5'-labeled RNA. '-' and '+' indicates absence or presence of $E\sigma^{70}$, respectively. (*right*) 3'-labeled RNA. Footprinting data courtesy of A. Feklistov from the Darst-Campbell Lab (The Rockefeller University). **(C)** Relative abortive transcription activity of WT 6S RNA (lane 1) and 6S RNA with the CS truncated (6S RNAΔCS, lane 2), and 6S RNAΔCS (lane 3) and 6S RNA* (lane 4). *In vitro* transcription data courtesy of E. A. Campbell, S. Feng, and K. Leon from the Darst-Campbell Lab (The Rockefeller University).


A
nt-strand
UD3
UD2
UD1
DD
CS
CB
wt 6S RNA
UTL
UB2
UB1
DB
t-strand
6S RNA*
DD*
Protection from RNase V1 and/or RNase A in presence of RNAP
Hypersensitivity to RNase V1 in presence of RNAP
OH* cleavage from active site
B
A
V1
OH-
T1
G103
G97
G92
G88
G80
G76
G73
G59
G42
G39
G37
G30
G42
G80
G92
G97
G103
G106
G117
G119
G122
G136
G143
1 2 3 4 5 6 7 8 9 10 11
5′ labeled RNA
3′ labeled RNA
C
wt 6S RNA
6S RNAΔCS
6S RNAΔCS
6S RNA*
ApUpC →
1 2 3 4

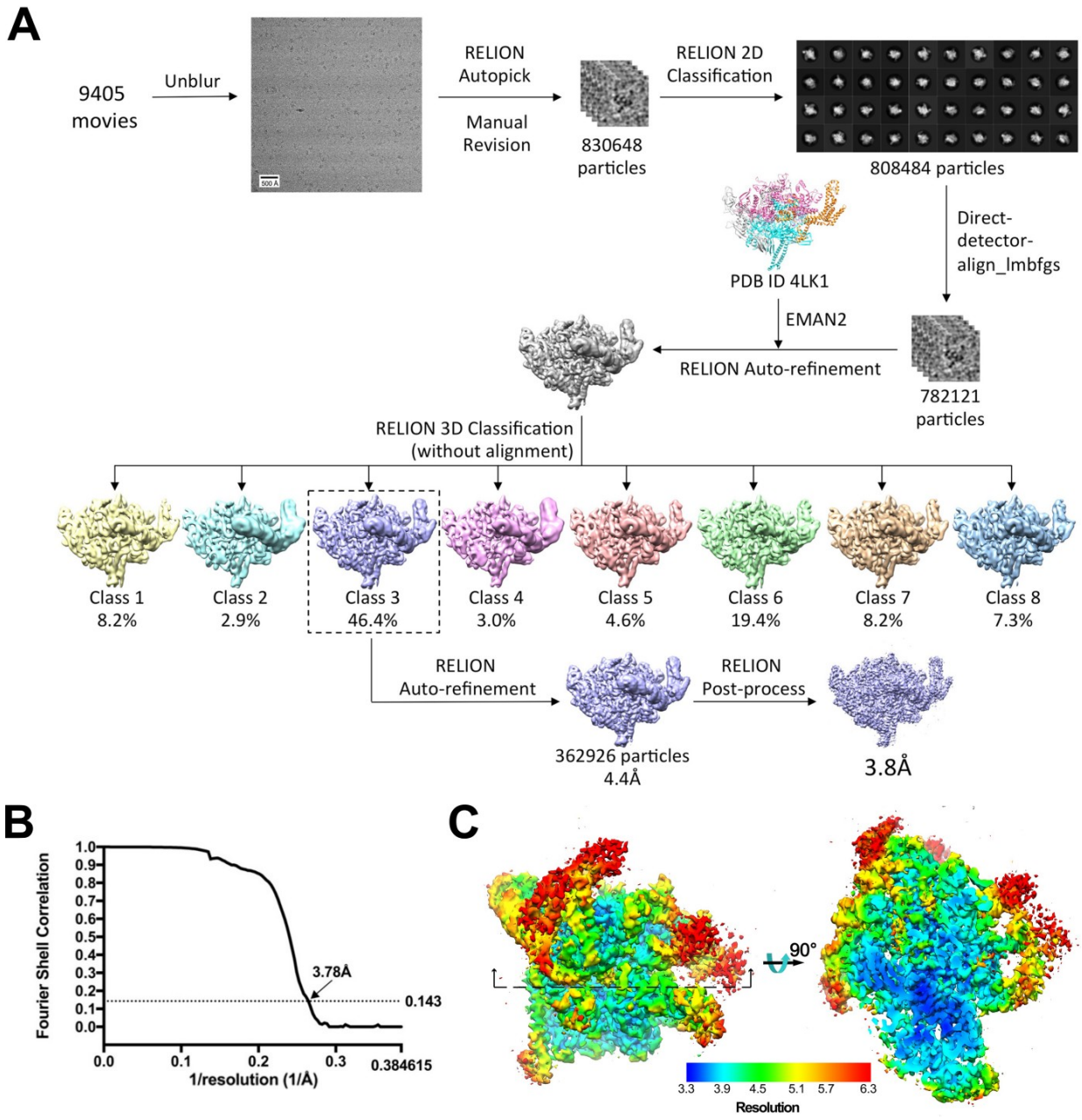


Figure 3.3. Cryo-EM analysis of 6S-Eσ^{70} complex

(A) Flowchart showing the image processing pipeline for the cryo-EM data of 6S-Eσ^{70}. Final structure contains 362,926 particles with a 3.8 Å resolution after auto-refinement and post-processing in RELION. **(B)** Gold-standard FSC of the 6S-Eσ^{70} cryo-EM map. The gold-standard FSC was calculated by comparing the two independently determined half-maps from RELION (Scheres, 2012). The dotted line represents the 0.143 FSC cutoff, which indicates a nominal resolution of 3.78 Å. **(C)** The 3.8-Å resolution cryo-EM density map of the 6S-Eσ^{70} complex colored by local resolution in Å (Cardone et al., 2013). The view on the right shows a cross section at the position indicated in the left panel.

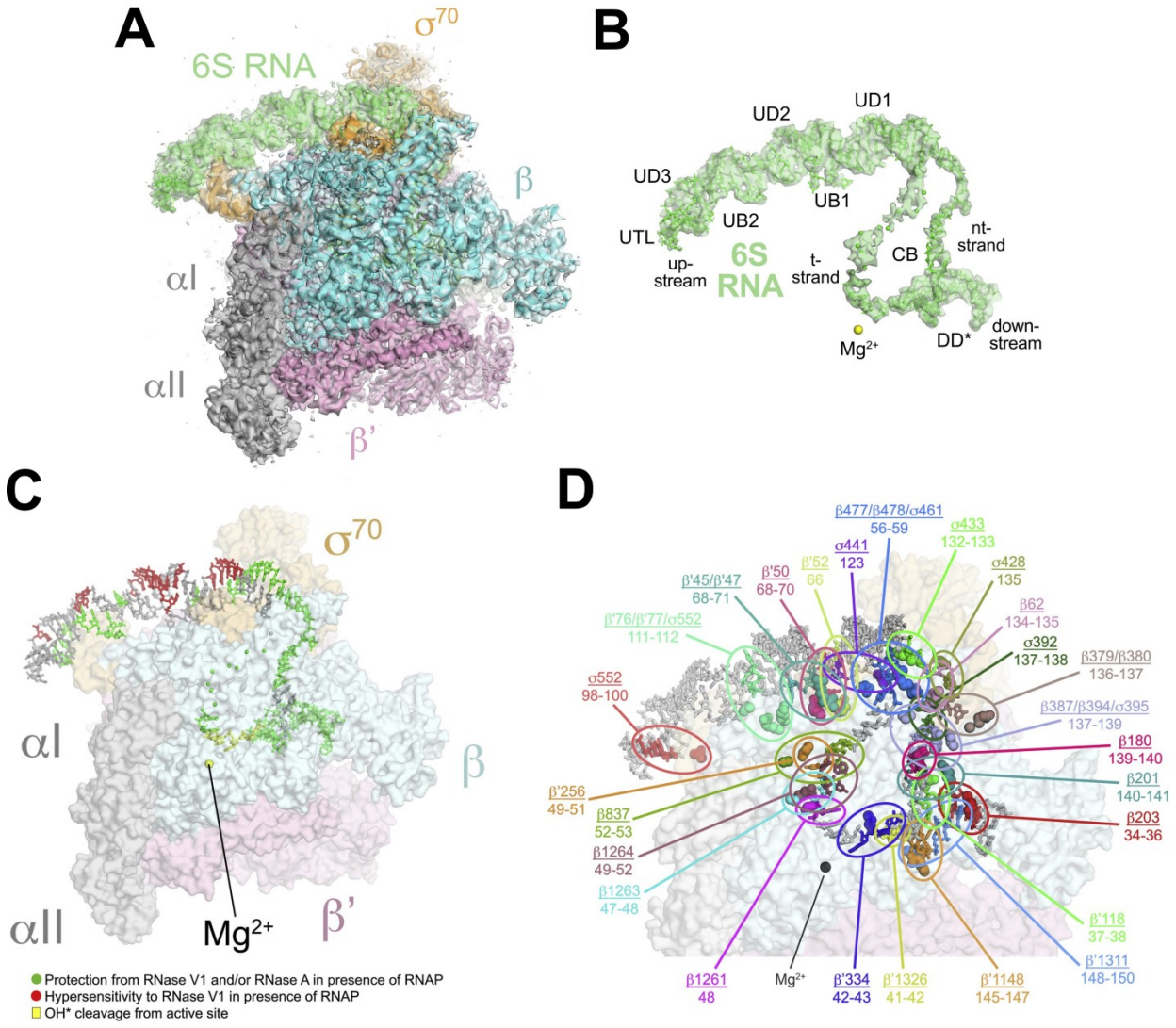
Table 3.4 Model statistics from Molprobity (Chen et al., 2010b)

		6S RNA/Eσ^{70}
Overall	Resolution[a]	3.78 Å
	Molprobity score	1.83
	Clashscore (all atoms)	6.56
	RMS deviations bonds (Å)	0.007
	RMS deviations angles (°)	1.027
Protein (3,132 residues)	Rotamer outliers	0.54%
	Ramachandran favored	92.57%
	Ramachandran outliers	0.08%
RNA[b] (112 nucleotides)	Probably wrong sugar puckers	0
	Bad backbone conformations	3 (2.68%)

[a] Overall resolution calculated using gold-standard FSC 0.143 cutoff (Rosenthal and Henderson, 2003).
[b] Validation of RNA model (Richardson et al., 2008).

Figure 3.5. Cryo-EM and biochemical validation of the 6S-Eσ^{70} structure
(A) The 3.8 Å resolution cryo-EM density map of 6S-Eσ^{70} is rendered as a transparent surface colored as shown. Superimposed is the final refined model; the RNAP is shown as a backbone ribbon; the 6S RNA is shown in stick format. **(B)** The 3.8 Å resolution cryo-EM density map with the superimposed model of only the 6S RNA. Shown for reference is the RNAP active site Mg^{2+} ion (yellow sphere). **(C)** Correspondence of the 6S-Eσ^{70} structure with the RNase footprinting and localized hydroxyl-radical cleavage results (**Figures 3.2A-B**). The 6S-Eσ^{70} structure is shown (RNAP as a transparent molecular surface, 6S RNA in stick format). The RNA is color coded according to **Figure 3.2A** (green, protected from RNase cleavage by Eσ^{70}; red, hypersensitive to RNase V1 cleavage in the presence of Eσ^{70}; yellow, efficiently cleaved by hydroxyl radicals generated at the RNAP active site). Footprinting data courtesy of A. Feklistov from the Darst-Campbell Lab (The Rockefeller University). **(D)** The results of RNAP-BPA/6S RNA crosslink mapping are superimposed on the 6S-Eσ^{70} structure. RNAP residues substituted individually with BPA and crosslinked to the 6S RNA are shown in Corey-Pauling-Koltun (CPK) format. The mapped 6S RNA nucleotides are indicated and color matched with the crosslinked BPA-RNAP residue. The matched RNAP residue and corresponding crosslinked RNA nucleotides are circled and labeled (RNAP residue/6S RNA nucleotides) in the same color. BPA crosslinking data courtesy of K.M Wassarman (University of Wisconsin–Madison).

Cryo-EM Map



RNA Footprinting

RNAP-RNA BPA Crosslinking

the RNA were not model due to a lack of density. Portions of the RNA upstream of UB2 were low resolution and were modeled using MC-fold (Parisien and Major, 2008) (**Figures 3.3C and 3.5B**).

The 6S-Eσ^{70} model was consistent with the RNA footprinting data (**Figures 3.2A-B**). This data was mapped on to the 3D model of 6S RNA binary complex in **Figure 3.5C**, showing nucleotides that were protected or sensitive to RNase treatment and hydroxyl radical cleavage. The model was further validated using nucleic acid-protein crosslinking, performed by our collaborator, K.M Wassarman (University of Wisconsin–Madison). A comprehensive map of RNAP-6S RNA crosslinks was generated by incorporating p-benzoyl-L-phenylalanine (BPA) at 35 distinct sites in the RNAP β, β', and σ^{70} subunits (Chen et al., 2017; **Figure 3.5D**).

3.5 Duplex segments of the 6S RNA show an unusual nucleic acid geometry

An alignment of the open DNA promoter from the *Thermus aquaticus* (*Taq*) RNAP open complex structure (PDB ID 4XLN; Bae et al., 2015a) with the structure of 6S RNA shows that the RNA interacts with RNAP in similar fashion to that of open DNA promoters (**Figure 3.6A**). As expected, the CB mimics a promoter DNA transcription bubble, including the 6S RNA t-strand near the RNAP active site. The U44 nucleotide is positioned near the active site Mg^{2+} for pRNA synthesis similar to the TSS of a DNA transcription bubble (**Figure 3.6B**) . The overall architecture and path of the 6S RNA phosphate backbone matches that of B-form open promoter DNA (Bae et al., 2015a) over essentially the entire length of the RNA (**Figure 3.6A**). Typically, duplex RNA is A-form, characterized by a deep, narrow major groove and a very shallow, wide minor groove. B-form RNA is sterically unfavorable and duplex regions of the 6S RNA do not exhibit B-form helical geometry (**Tables 3.7-3.9**).

Insight into the origin of the overall "B-form" appearance of the 6S RNA comes from a 3DNA analysis of the major and minor groove widths (Mgw and mgw, respectively) of duplex regions in 6S RNA (Lu and Olson, 2008). In duplex stretches long enough to measure the groove widths (UD3, UD2, UD1, and DD∗), the mgws conform to A-form, while the Mgws conform to B-form (**Table 3.7**, **Figure 3.10A**). The sugar pucker conformations (**Table 3.8**), intrastrand phosphate-phosphate distances (**Table 3.9**) and twist (**Table 3.7**) are characteristic of A-form nucleic acid duplexes, but the rise and pitch (**Table 3.7**) are characteristic of B-form duplexes. Thus, overall the duplex regions of the 6S RNA are neither A-form nor B-form but instead have characteristics of both. The segments of duplex 6S RNA (UD3, UD2, UD1, and DD∗) with widened Mgws with B-form characteristics correspond to regions that have extensive RNAP interactions with the 6S RNA phosphate backbone (**Figure 3.10B**).

3.6 Interactions between 6S RNA and Eσ^{70}

The duplex segments of 6S RNA are interrupted by regions of base-pair mismatches, non-canonical base pairs, and flipped-out bases (UB2, UB1, and CB) that are conserved in an alignment of 101 enterobacteriaceae 6S RNA sequences (**Figure 3.11**). The majority of RNAP interactions occurs with the phosphate backbone of 6S RNA. Base-specific 6S RNA/RNAP interactions occur nearly exclusively with flipped-out bases [C69 of UB1; G82 of UB2; and U135/G136/G143 of the CB nontemplate-strand (nt-strand)] (**Figure 3.12A**).

Although not conserved, C69 is flipped out of UB1 and interacts with residues of the RNAP β′ subunit (**Figure 3.11**). Conserved G82 of UB2 (**Figure 3.11**) interacts with βR903 on the flap-tip-helix (**Figure 3.12B**). The residue corresponding to *Eco* RNAP βR903 is 100% identical among enterobacteriaceae RNAPs and 80% identical in nearly 1,000 diverse bacterial

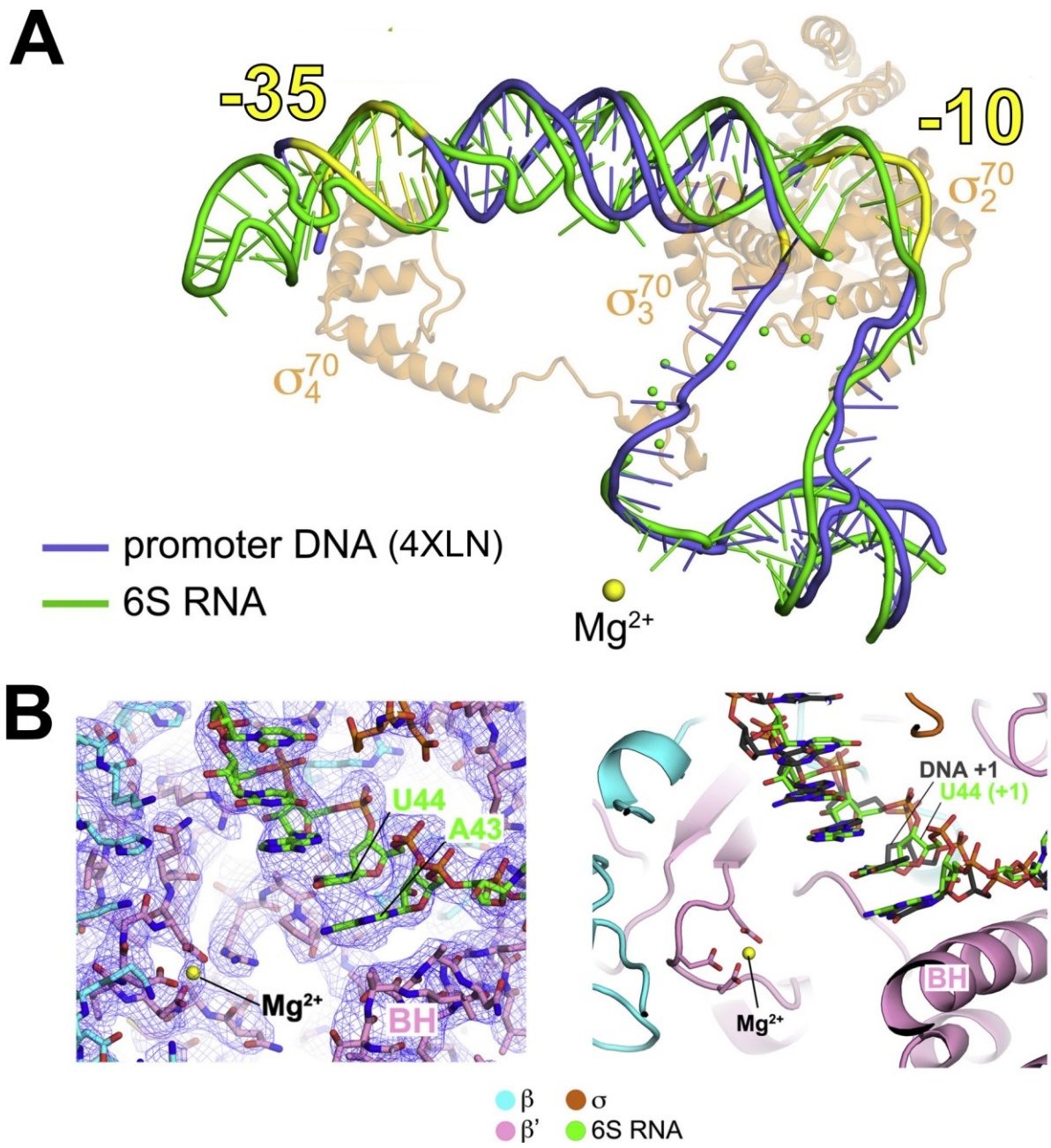


Figure 3.6. Comparison of 6S RNA structure with open promoter DNA
(A) Superimposition of 6S RNA (green) and promoter DNA from an open promoter complex (blue; PDB ID 4XLN; Bae et al., 2015a). The -35 and -10 elements of the promoter DNA are colored yellow. For reference, the Eco σ^{70} from the 6S-Eσ^{70} structure is shown (transparent light orange Cα ribbon), and the RNAP active site Mg^{2+} ion is shown as a yellow sphere. **(B)** (*left*) View of the 6S-Eσ^{70} cryo-EM density map (blue mesh) showing the region around the RNAP active site (Mg^{2+} ion shown as a yellow sphere). The TSS on 6S RNA is U44. (*right*) Shown is the 6S-Eσ^{70} structure near the RNAP active site (Mg^{2+} ion, yellow sphere). Superimposed on the 6S RNA is the promoter DNA t-strand from the promoter open complex structure (Bae et al., 2015a). The TSS on 6S RNA (U44) and promoter DNA (DNA +1) are noted.

Table 3.7. Minor and major groove widths and axial base pair rise for 6S RNA[a]

6S RNA element[b]	Strand I	step	Strand II	Minor groove P-P (Å)[c]	Major groove P-P (Å)[c]	Rise (Å)	Twist (°)	Pitch (Å)
	32-33	CG/CG	152-153			3.6	28	
	33-34	GG/CC	151-152			2.9	32	
	34-35	GC/AC	150-151	11.5	10.4	2.8	33	
	35-36	CA/UA	149-150	11.2	11.9	4.4	29	
DD*	36-37	AC/GU	148-149	10.7	13.2	3.2	31	41
	37-38	CA/UG	147-148	10.5	15.2	3.4	24	
	38-39	AU/GU	146-147	10.6	13.8	3.1	39	
	39-40	UG/CG	145-146			3.7	18	
	40-41	GC/GC	144-145			3.3	31	
	59-60	GA/UC	129-130			2.9	32	
	60-61	AA/UU	128-129	10.9	14.9	3.2	30	
	61-62	AU/AU	127-128	10.7	12.9	3.2	28	
	62-63	UG/CA	126-127	10.7	12.7	3.4	27	
UD1	63-64	GU/AC	125-126	10.8	12	3.2	32	35
	64-65	UG/CA	124-125	10.9	11.1	3.2	31	
	65-66	GG/AC	123-124	10.6	10.6	3.1	30	
	66-67	GC/GA	122-123	9.3	9.9	3.2	31	
	67-68	CG/CG	121-122	9.9	8.8	2.9	33	
	72-73	CG/CG	118-119	10.8	9.2	4.4	30	
	73-74	GC/GC	117-118	10.7	11.2	3.1	35	
	74-75	CG/UG	116-117	11.0	12.1	3.3	26	
UD2	75-76	GG/CU	115-116	11.5	11	2.9	33	33
	76-77	GU/AC	114-115	11.8	10.3	3.1	32	
	77-78	UU/AA	113-114			3.1	29	
	78-79	UG/AA	112-113			1.4	49	
	88-89	GC/GA	103-104	13.3	15	3.2	39	
	89-90	CU/AG	102-103	12.1	10.8	3.0	25	
UD3	90-91	UC/GA	101-102	12.1	11.5	2.9	30	31
	91-92	CG/CG	100-101	11.9	10.9	2.6	34	
	92-93	GG/CC	99-100			3.0	37	
	93-94	GU/UC	98-99			2.9	39	
				Canonical A-form nucleic acid[d]	Canonical B-form nucleic acid[d]	6S RNA (average)		
			Rise (Å)	2.6	3.4	3.2 ± 0.5		
			Twist (°)	33	36	32 ± 6		
			Pitch (Å)	28	34	35 ± 4		

[a] Parameters calculated using 3DNA (Lu and Olson, 2008).
[b] See **Figure 3.1**.
[c] Direct P-P distances (with van der Waals radii of phosphate groups, 5.8 Å, subtracted).
[d] See reference (Saenger, 1984).

Table 3.8. Sugar puckers for 6S RNA[a]

	Strand I			Strand II	
6S RNA element	nucleotide	Sugar pucker		nucleotide	Sugar pucker
DD*	32C	C3'-endo		153G	C3'-endo
	33G	C3'-endo		152C	C3'-endo
	34G	C3'-endo		151C	C3'-endo
	35C	C3'-endo		150A	C3'-endo
	36A	C3'-endo		149U	C3'-endo
	37C	C3'-endo		148G	C3'-endo
	38A	C3'-endo		147U	C3'-endo
	39U	C3'-endo		146G	C3'-endo
	40G	C3'-endo		145C	C3'-endo
	41C	C3'-endo		144G	C3'-endo
UD1	59G	C3'-endo		130C	C3'-endo
	60A	C3'-endo		129U	C3'-endo
	61A	C3'-endo		128U	C3'-endo
	62U	C3'-endo		127A	C3'-endo
	63G	C3'-endo		126C	C3'-endo
	64U	C3'-endo		125A	C3'-endo
	65G	C3'-endo		124C	C3'-endo
	66G	C3'-endo		123A	C3'-endo
	67C	C3'-endo		122G	C3'-endo
	68G	C3'-endo		121C	C3'-endo
UD2	72C	C3'-endo		118G	C3'-endo
	73G	C1'-endo		117C	C3'-endo
	74C	C3'-endo		116G	C3'-endo
	75G	C3'-endo		115U	C3'-endo
	76G	C3'-endo		114C	C3'-endo
	77U	C3'-endo		113A	C3'-endo
	78U	C3'-endo		112A	C3'-endo
	79G	C3'-endo		111A	C2'-endo
UD3	88G	C3'-endo		104A	C3'-endo
	89C	C3'-endo		103G	C3'-endo
	90U	C3'-endo		102A	C3'-endo
	91C	C3'-endo		101G	C3'-endo
	92G	C3'-endo		100C	C2'-exo
	93G	C3'-endo		99C	C3'-endo
	94U	C3'-endo		98U	C3'-endo

[a] Calculated using 3DNA (Lu and Olson, 2008).

Table 3.9. Intrastrand P-P virtual bond distances[a]

6S RNA element[b]	Strand I	step	Intrastrand P—P distance (Å)		Strand II	step	Instrastrand P—P distance (Å)
DD*	32-33	CG	6.0		152-153	CG	5.8
	33-34	GG	6.5		151-152	CC	5.8
	34-35	GC	5.7		150-151	AC	5.9
	35-36	CA	5.6		149-150	UA	5.9
	36-37	AC	5.7		148-149	GU	6.4
	37-38	CA	6.0		147-148	UG	5.6
	38-39	AU	5.9		146-147	GU	6.6
	39-40	UG	5.8		145-146	CG	5.8
	40-41	GC	5.6		144-145	GC	5.4
UD1	59-60	GA	6.3		129-130	UC	5.6
	60-61	AA	7.0		128-129	UU	6.0
	61-62	AU	5.5		127-128	AU	5.8
	62-63	UG	5.9		126-127	CA	5.8
	63-64	GU	5.6		125-126	AC	5.7
	64-65	UG	5.9		124-125	CA	5.8
	65-66	GG	5.8		123-124	AC	6.3
	66-67	GC	5.3		122-123	GA	5.7
	67-68	CG	6.5		121-122	CG	6.2
UD2	72-73	CG	5.7		118-119	CG	5.7
	73-74	GC	6.8		117-118	GC	6.6
	74-75	CG	5.8		116-117	UG	6.1
	75-76	GG	5.7		115-116	CU	5.6
	76-77	GU	6.0		114-115	AC	5.8
	77-78	UU	5.6		113-114	AA	4.9
	78-79	UG	6.2		112-113	AA	8.5
UD3	88-89	GC	5.1		103-104	GA	6.4
	89-90	CU	5.2		102-103	AG	6.2
	90-91	UC	5.3		101-102	GA	5.7
	91-92	CG	5.8		100-101	CG	6.4
	92-93	GG	5.8		99-100	CC	7.0
	93-94	GU	5.3		98-99	UC	6.8

	Canonical A-form nucleic acid[d]	Canonical B-form nucleic acid[d]	6S RNA (average)
Intrastrand P—P distance (Å)[c]	5.9	7.0	6.0 ± 0.8

[a] Parameters calculated using 3DNA (Lu and Olson, 2008).
[b] See **Figure 3.1** for 6S RNA structural definitions.

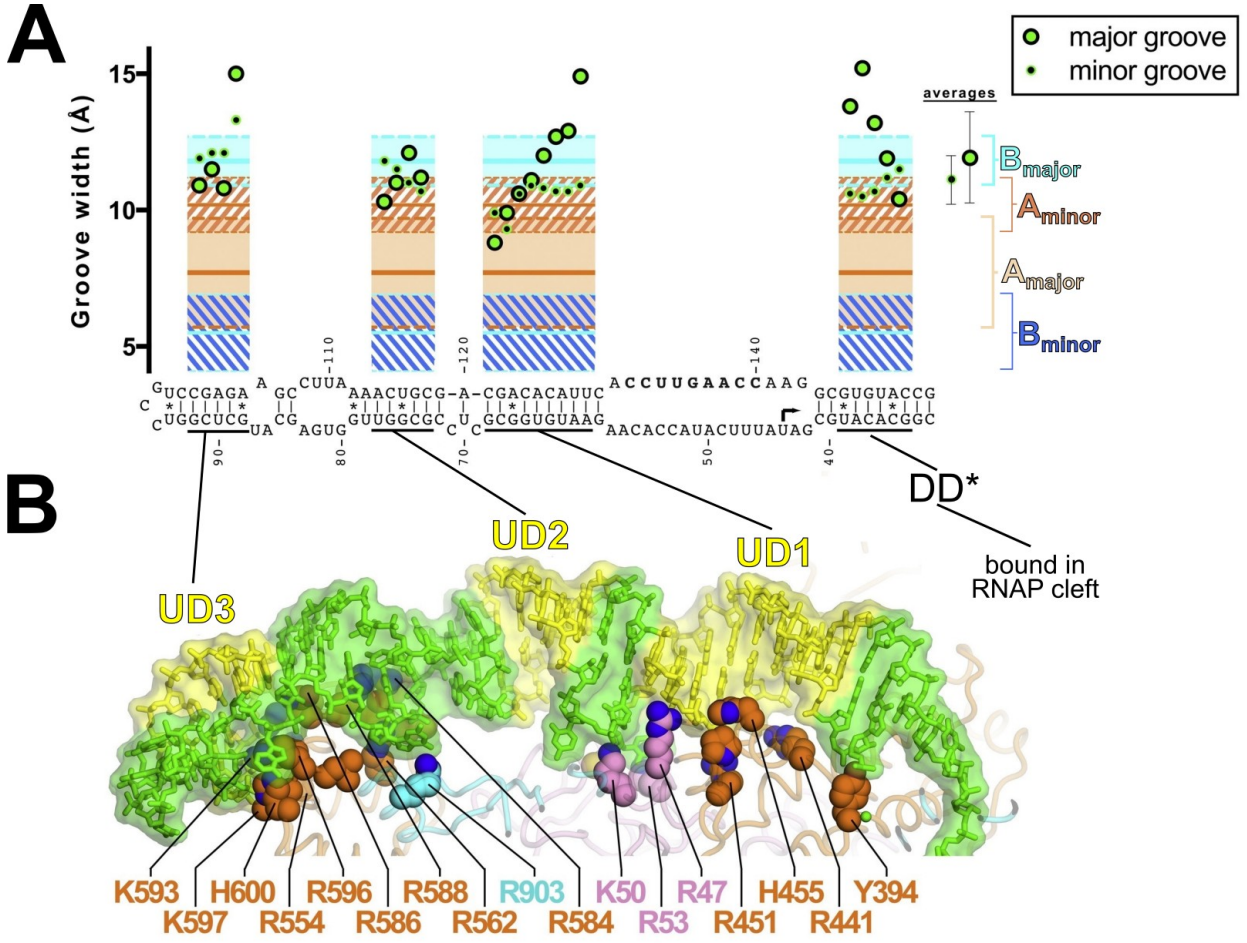


Figure 3.10. 6S RNA mimics B-Form promoter DNA

(A) Major groove (large green dot with black border) and minor groove (small black dot with green border) widths (Mgw and mgw, respectively) for duplex regions of the 6S RNA, calculated using 3DNA (Lu and Olson, 2008) but with the van der Waals radii of the phosphate groups (5.8 Å) subtracted, compared with average values (± SD) for B-form (blue regions) and A-form (orange regions) nucleic acids (PDB IDs: A-form, 1D4R, 1LNT, 1QC0, 1RXB; B-form, 1D23, 1D49, 1D56, 1D57, 158D). The average 6S RNA mgw (11.1 ± 0.9 Å, shown on the right under "averages") is close to the average mgw for A-form nucleic acids (10.2 ± 1.0 Å), while the average Mgw of the 6S RNA (11.9 ± 1.7 Å) is closer to the average Mgw for B-form nucleic acids (11.8 ± 0.9 Å). **(B)** Shown is the upstream portion of the 6S RNA (transparent molecular surface with stick model inside). Yellow indicates duplex regions with widened major grooves; other regions are colored green. Shown in CPK format are amino acid side chains of RNAP residues that interact with the RNA (< 4.5 Å; β residues, blue; β′, pink; σ^{70}, orange). Selected residues that make polar interactions with the RNA backbone are labeled. Structural analysis and figure courtesy of S. A. Darst (The Rockefeller University).

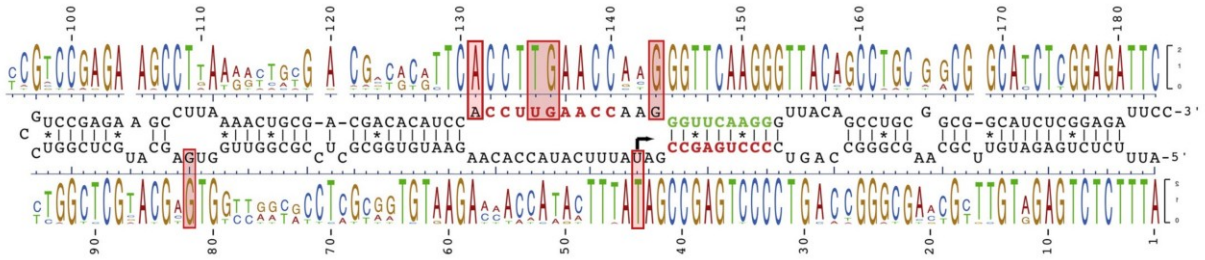


Figure 3.11. Conserved bases in 6S RNA
Shown in the middle is the secondary structure of the WT *Eco* 6S RNA. On top and bottom is aligned a sequence logo (Schneider and Stephens, 1990) derived from a sequence alignment of 101 enterobacteriaceae 6S RNA sequences. Nucleotides in red boxes are examined in light of the structure of the 6S-Eσ^{70} complex.

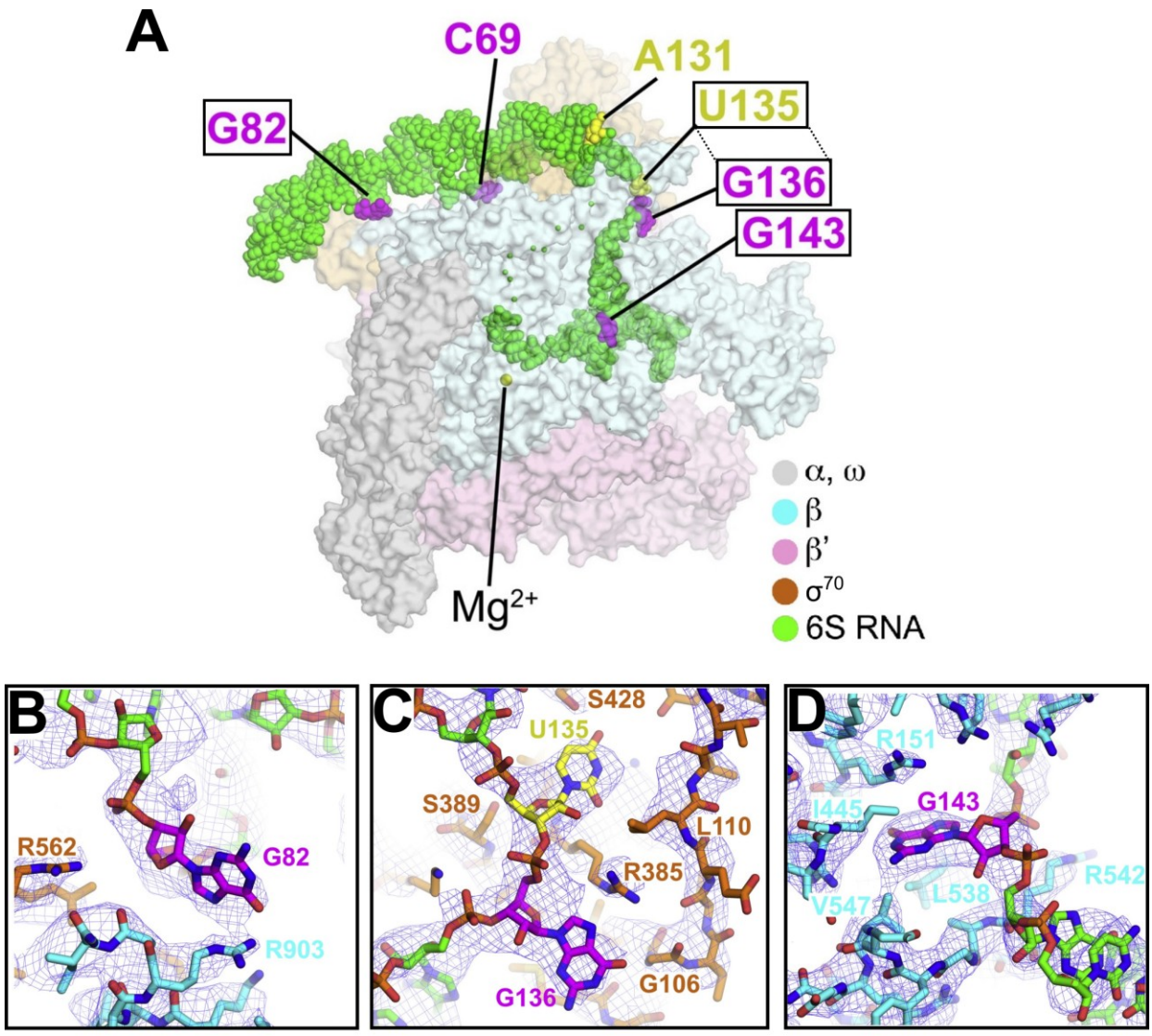


Figure 3.12. Conserved interactions and cryo-EM densities
(A) Overall structure of 6S-Eσ^{70} complex, color coded as shown in the legend. The RNAP is shown as a transparent molecular surface; the RNA is shown as green CPK atoms with flipped-out bases that interact with RNAP colored magenta; and A131/U135, positions equivalent to promoter DNA A_{-11}(nt)/T_{-7}(nt), are colored yellow. The boxed bases are shown below. **(B)** Cryo-EM density (blue mesh) showing the conserved interaction between RNA nucleotide G82 (magenta) and RNAP β R903. **(C)** Cryo-EM density (blue mesh) showing RNA nucleotides U135 (yellow) and G136 (magenta) interacting with σ^{70} in the same manner as promoter DNA T_{-7}(nt) of the -10 element and G_{-6}(nt) of the discriminator (Bae et al., 2015a; Zhang et al., 2012). **(D)** Cryo-EM density (blue mesh) showing RNA nucleotide G143 (magenta) interacting with a G-specific pocket of the RNAP β subunit in the same manner as promoter DNA G_{+2}(nt) CRE (Zhang et al., 2012).

RNAP β subunit sequences (Lane and Darst, 2010a). K.M Wassarman (University of Wisconsin–Madison) generated 6S RNA mutants and derivatives and tested binding (Chen et al., 2017). G82 mutations substantially disrupt 6S RNA binding to $E\sigma^{70}$ in electrophoretic mobility shift assays (EMSA) (**Figure 3.13A**).

The conserved U135, G136, and G143 of the CB nt-strand (**Figure 3.11**) interact with $E\sigma^{70}$ in a fashion that is essentially identical to the corresponding nucleotides of open promoter DNA (**Figures 3.12C-D**). U135 binds the pocket of σ^{70}_2 that normally interacts with the nt-strand thymine at position -7 [T_{-7}(nt)] (**Figure 3.12C**), one of two nearly absolutely conserved bases of the promoter -10 element (Feklistov and Darst, 2011; Shultzaberger et al., 2007). Immediately downstream of U135, G136 interacts with σ^{70}, mimicking the G_{-6}(nt) in the discriminator (Bae et al., 2015a; Zhang et al., 2012) (**Figure 3.12C**). Finally, G143 binds in a G-specific pocket of the RNAP β subunit, normally reserved for the G_{+2}(nt) core recognition element (CRE) (Bae et al., 2015a; Zhang et al., 2012) (**Figure 3.12D**). However, base substitutions at these three positions of 6S RNA have little effect on 6S RNA binding (**Figure 3.13B**). U135/G136 are involved in the formation of 6S RNA isoform 2 (**Figure 3.1**). Their binding to RNAP may stabilize isoform 1 and disfavor spontaneous formation of isoform 2, thus preventing premature release of 6S RNA from $E\sigma^{70}$.

Although the interactions of U135, G136, and G143 with RNAP mimic interactions of promoter DNA T_{-7}(nt), G_{-6}(nt), and G_{+2}(nt), the most important and conserved promoter DNA/RNAP interaction with the A_{-11}(nt) base of the -10 element is not recapitulated in 6S RNA. In open promoter complexes, A_{-11}(nt) is flipped into a pocket of σ^{70} and stabilized by the 'chair' conformation of the absolutely conserved σ^{70} W-dyad (W433/W434) (Bae et al., 2015a; Feklistov and Darst, 2011). Located four bases upstream of the T_{-7}(nt) mimic U135, the conserved A131 (**Figure 3.11**) corresponds to the A_{-11}(nt) of DNA promoters. However, in the cryo-EM structure of 6S RNA, A131 is not flipped out into the σ^{70} pocket and the W-dyad is in its 'edge-on' conformation and has not switched to the 'chair' conformation as seen in promoter DNA complexes (**Figure 3.14A**). Consistent with this observation, A131 does not crosslink with BPA-RNAP substitutions that crosslink promoter DNA near this position (σ^{70} W434) (Winkelman et al., 2015) (**Figure 3.5D**).

The flipping and capture of the A_{-11}(nt) by σ^{70} is thought to be the critical event that initiates melting of the -10 element and formation transcription bubble formation (Feklistov and Darst, 2011). Since the 6S RNA CB is "pre-melted," this interaction is dispensable. A131 substitutions in 6S RNA do not alter RNAP affinity (**Figure 3.13B**). This is further supported by *in vitro* transcription data courtesy of E. A. Campbell, S. Feng, and K. Leon from the Darst-Campbell Lab (The Rockefeller University). An RNAP holoenzyme containing a quadruple-substituted σ^{70} ($\sigma^{70}_{FYWW,}$ Tomsic et al., 2001) that is severely defective in initiating and maintaining the transcription bubble is completely inactive on duplex T7A1, a strong DNA promoter, but is fully active on a pre-melted T7A1 DNA promoter template and on 6S RNA (Chen et al., 2017; **Figures 3.14B-C**).

3.7 Structural basis of 6S RNA specificity for $E\sigma^{70}$

σ^{70} and σ^S are the most closely related σ factors in *Eco*, with overlapping promoter specificity (Gaal et al., 2001). Despite this, 6S RNA shows unique specificity for $E\sigma^{70}$ over *Eco* σ^S-holoenzyme ($E\sigma^S$) (Trotochaud and Wassarman, 2005). A major difference in the two σ factors is the non-conserved region (NCR) between conserved regions 1.2 and 2.1. The NCR of σ^{70} contains 248 residues between regions 1.2 and 2.1 while σ^S only has 3 residues

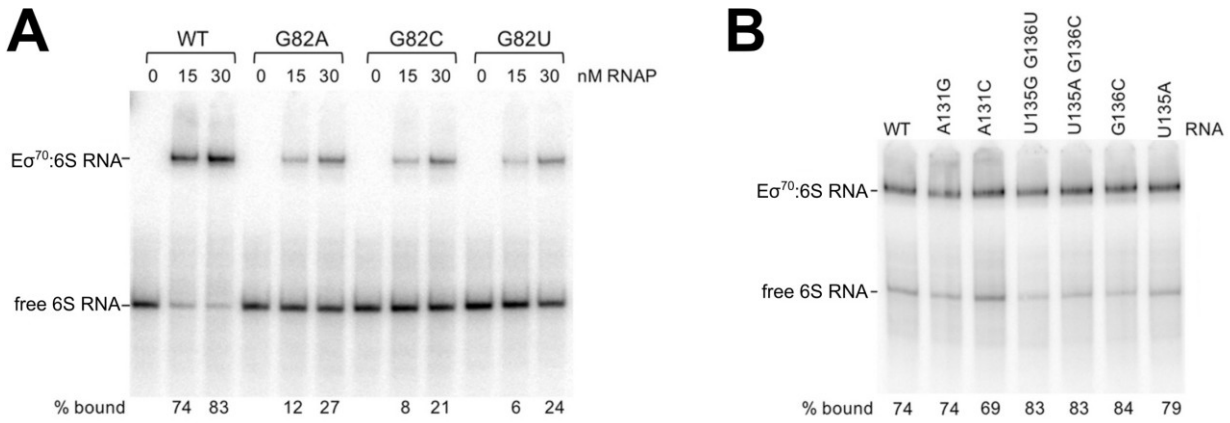


Figure 3.13. Examination of 6S RNA base substitutions on binding
EMSA was used to examine 6S RNA binding to $E\sigma^{70}$. **(A)** G82 substitutions. **(B)** 6S RNA substitutions at A131, U135, or G136 showing absence of altered binding for $E\sigma^{70}$. 6S RNA binding data and figure courtesy of K.M Wassarman (University of Wisconsin–Madison).

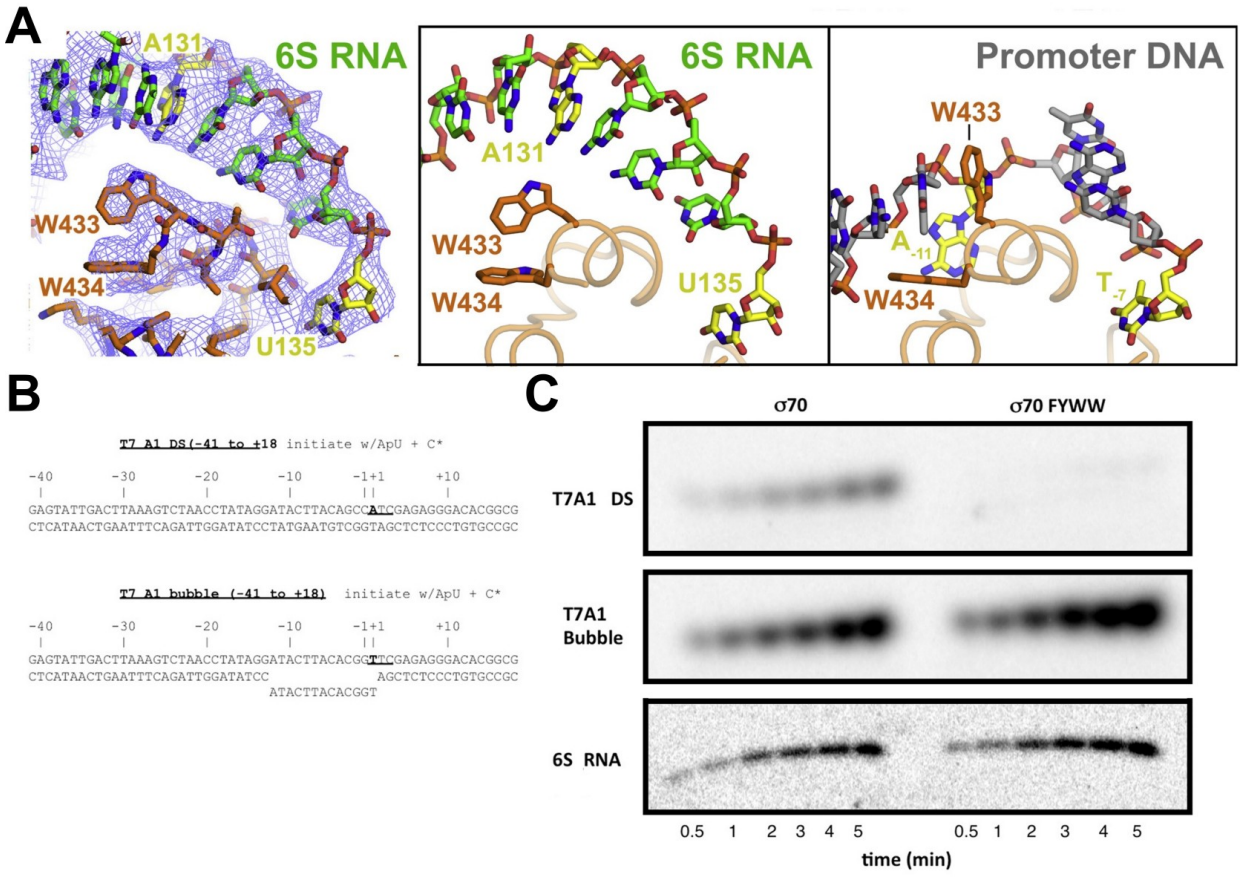


Figure 3.14. Role of W-dyad in 6S RNA binding

(A) (*left*) View of cryo-EM density (blue mesh) showing RNA nucleotide A131 (yellow) in a duplex RNA base stack passing over the σ^{70} W-dyad. The tryptophan residues maintain their 'edge-on' conformation characteristic of σ^{70} alone (Malhotra et al., 1996) when not interacting with the double-strand/single-strand fork at the upstream edge of the transcription bubble (Bae et al., 2015a). (*right*) Comparison of 6S RNA and promoter DNA interacting with the absolutely conserved σ^{70} W-dyad. Shown is only the nt-strand of the RNA or DNA. Corresponding nucleotides A131/U135 of 6S RNA or conserved A_{-11}(nt)/T_{-7}(nt) of the -10 element (Shultzaberger et al., 2007) are colored yellow. **(B)** Sequences of the duplex T7A1 promoter (*top*) and the T7A1 bubble promoter (*bottom*). **(C)** Abortive transcription activity (ApUpC synthesis) of $E\sigma^{70}$ (*left*) and melting deficient $E\sigma^{70}$ FYWW (*right*; Tomsic et al., 2001) on duplex T7A1 (T7A1 DS, *top row*), T7A1 bubble (*middle row*), and 6S RNA (*bottom row*). *In vitro* transcription data courtesy of E. A. Campbell, S. Feng, and K. Leon from the Darst-Campbell Lab (The Rockefeller University).

(Lonetto et al., 1992). The cryo-EM structure shows that 6S RNA does not interact with this non-conserved region of σ^{70}. The base-specific interactions between 6S RNA and σ^{70} all involve highly conserved regions that are common to both σ factors (**Figure 3.15A**). I mutated and tested for residues of σ^{70} that interact with 6S RNA but were divergent from the corresponding residue in σ^{S} for binding via transcription assays on a σ^{70} regulated promoter (T7A1), a σ^{S} regulated promoter (*dps*) and 6S RNA. Swap mutants with the strongest effects on 6S RNA binding were located on σ^{70}_4 domain (**Figure 3.15B**).

The σ^{70}_4 domain, responsible for recognition of the promoter DNA -35 element, is required for the 6S RNA/Eσ^{70} interaction and has been implicated in mediating 6S RNA specificity between σ^{70} and σ^{S} (Cavanagh et al., 2008; Klocko and Wassarman, 2009). Notable base-specific interactions were not observed between σ^{70}_4 and 6S RNA, but the σ^{70} binding surfaces for promoter DNA and 6S RNA differ significantly (Klocko and Wassarman, 2009) (**Figures 3.16A-B**). The -35 element sits at the N-terminal end of the σ^{70}_4 helix-turn-helix motif forming an interface composed of base-specific and backbone interactions that buries 634 Å^2 of molecular surface area (Hubin et al., 2017b). By contrast, the 6S RNA wraps around σ^{70}_4, burying 791 Å^2 of protein surface (**Figure 3.16A**), consistent with BPA-RNAP crosslinks to 6S RNA (σ^{70} 552; **Figure 3.5D**) that are not observed with promoter DNA (Winkelman et al., 2015). Multiple polar interactions are established between the RNA phosphate backbone and positively charged protein side chains, mostly with residues that are conserved between σ^{70} and σ^{S} (**Figure 3.15A**).

Two σ^{70}_4 residues that interact with the 6S RNA backbone, K593 and H600, correspond to negatively charged Glu residues in σ^{S} (E308 and E315; **Figure 3.15A**). Substitution σ^{70} K593E is strongly defective for 6S RNA binding (Klocko and Wassarman, 2009). Substitutions of these positions in σ^{S} with the corresponding residues of σ^{70} (σ^{S} E308K, σ^{S} E315H, σ^{S} E308K/E315H) increases the transcription activity of the resulting Eσ^{S} on 6S RNA ~2-fold and 1.5-fold, ~3-fold, respectively (**Figures 3.15B**, **3.16C**). The σ^{S} single or double substitutions do not have dramatic effects on activity on the σ^{70}-specific T7A1 promoter, nor on the σ^{S}-specific DPS promoter (**Figures 3.15B**, **3.16C**). Thus, the identities of two residues in σ^{70}_4 (out of an interface that buries ~4,500 Å^2 surface area and involves at least 68 residues) account for a significant fraction of the 6S RNA preference for σ^{70} over σ^{S}.

3.8 Discussion

In summary, I used single particle cryo-EM to solve the structure of a complex of *Eco* 6S RNA bound to Eσ^{70} that was previously intractable for study by X-ray crystallography. The structure shows duplex segments of 6S RNA have A-form C3'-endo sugar puckers but widened major groove widths, giving the RNA an overall architecture that mimics B-form promoter DNA. The structure also shows 6S RNA specificity for Eσ^{70} over Eσ^{S} is largely determined by two amino acids. The results help explain the specificity of *Eco* 6S RNA for Eσ^{70}, and show how an ncRNA can mimic B-form DNA to directly regulate transcription by the DNA-dependent RNAP.

The deep and narrow major groove of A-form RNA duplexes is inaccessible for sequence-specific readout by protein side chains. The bovine immunodeficiency virus (BIV) expresses a peptide that forms a β-hairpin that inserts into the major groove of the TAR RNA, widening the RNA major groove to allow sequence-specific RNA recognition (Puglisi et al., 1995; Ye et al., 1995). Bacterial RNAP uses a completely different mechanism to widen the 6S RNA major groove; RNAP interactions with the RNA backbone pry open the major groove, not

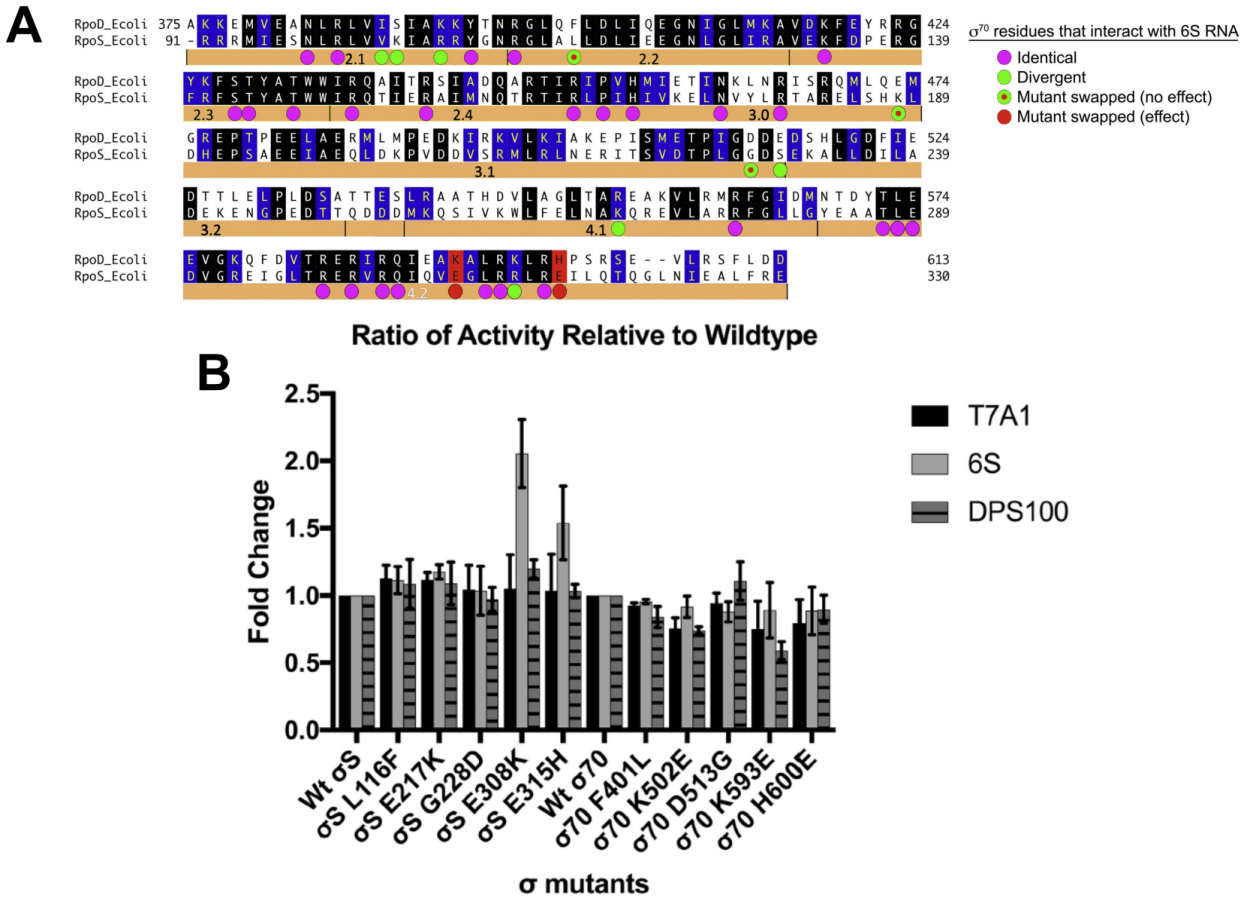


Figure 3.15. Interactions between σ^{70} with 6S RNA

(A) Sequence alignment of Eco σ^{70} (RpoD) and σ^{S} (RpoS). Identical residues are shaded black, and homologous residues are shaded blue. *Eco* σ^{70} residues that interact with 6S RNA are denoted by a colored dot underneath: magenta dots indicate positions that are identical between σ^{70}/σ^{S} and green dots indicate positions that diverge in sequence. Green dots with red dots inside are positions that were swapped between σ^{70} and σ^{S} but had no apparent effect on 6S RNA preference. Two loci, denoted in red (σ^{70} K593/H600; σ^{S} E308/E315), altered 6S RNA preference when swapped (Klocko and Wassarman, 2009). **(B)** Histogram showing the normalized abortive transcription initiation activity for the indicated σ's (*bottom*) on the indicated templates [σ^{70}-specific T7A1 promoter; σ^{70}-specific 6S RNA; σ^{S}-specific *dps* promoter (Grainger et al., 2008)]. The Eσ^{S} mutant activities are shown normalized with respect to WT Eσ^{S}, while the Eσ^{70} mutant are shown normalized to WT Eσ^{70}. The histograms represent the average of at least three measurements, the error bars denote standard error of the mean (SEM).

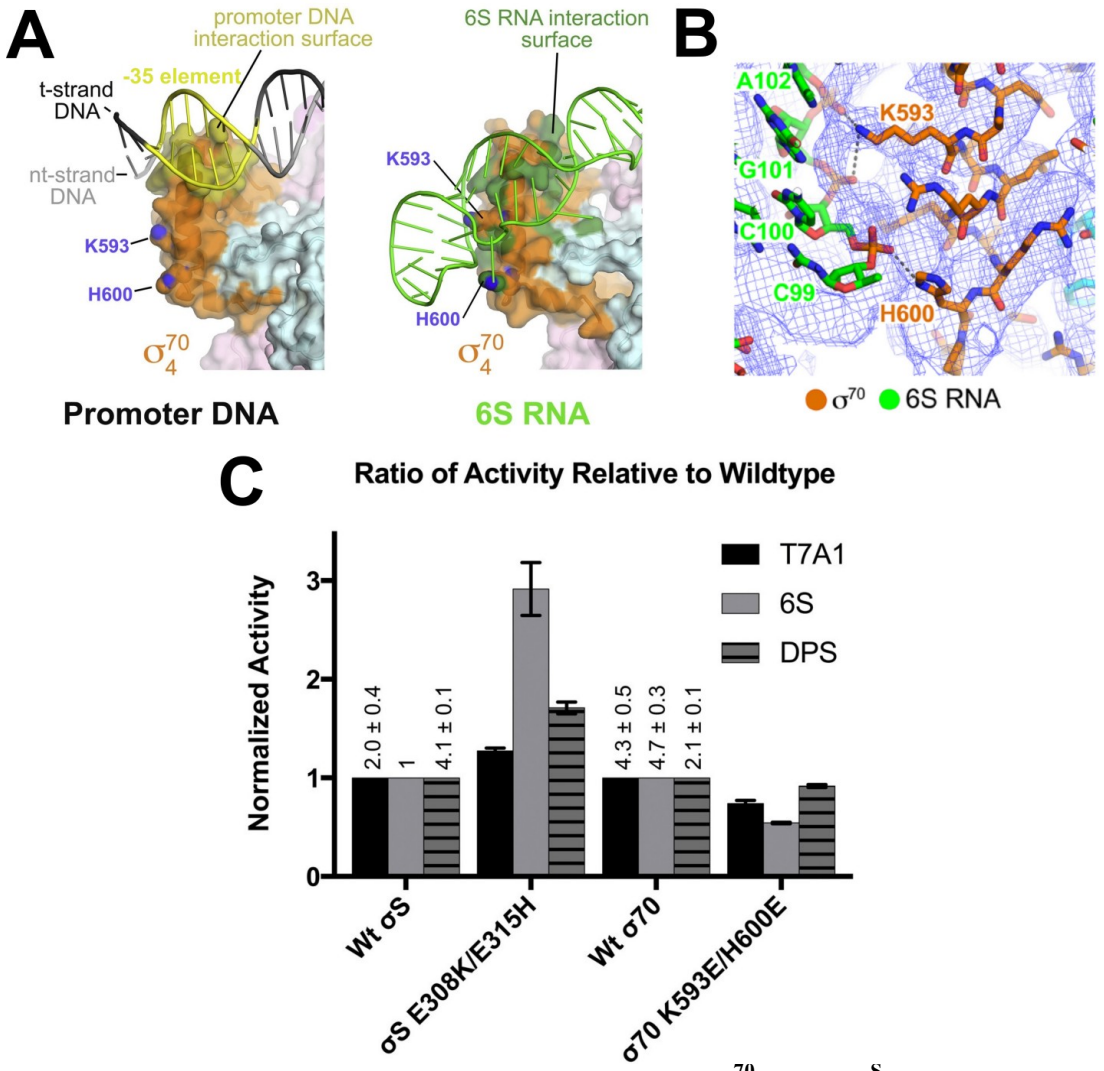


Figure 3.16. Structural Basis for 6S RNA Specificity for Eσ^{70} over Eσ^{S}

(A) Interaction surfaces for promoter DNA (*left*; PDB ID 5WT1; Hubin et al., 2017b) and 6S RNA (*right*) on the σ_4 domain. The RNAP holoenzyme is shown as a molecular surface (β, light cyan; β′, light pink; σ^{70}, orange). The surface of σ_4 that contacts the promoter DNA is colored dark yellow (*left*). Marked with their blue nitrogen atoms are K593 and H600 (using *Eco* σ^{70} numbering), which are solvent-exposed. On the right, the surface of $\sigma^{70}{}_4$ that contacts 6S RNA is colored dark green, illustrating that K593/H600 are engaged with the 6S RNA. **(B)** View of the 6S-Eσ^{70} cryo-EM density map (blue mesh) showing the interaction of σ^{70} K593 and H600 with the 6S RNA phosphate backbone. **(C)** Histogram showing the normalized abortive transcription initiation activity for the indicated σ's (bottom) on the indicated templates (σ^{70}-specific T7A1 promoter; σ^{70}-specific 6S RNA; σ^{S}-specific *dps* promoter). The Eσ^{S} mutant activities are shown normalized with respect to WT Eσ^{S}, while the Eσ^{70} mutants are shown normalized to WT Eσ^{70}. The relative values of the WT Eσ^{70} and Eσ^{S} transcription activities on the different templates (with respect to 6S RNA with σ^{S}) are listed. The histograms represent the average of at least three independent measurements, the error bars denote SEMs. Mutation of σ^{S} positions 308 and 315 to the corresponding residues in σ^{70} (σ^{S} E308K/E315H) increased the relative activity on 6S RNA with less effect on the relative activity on the other templates.

for the purposes of sequence-specific RNA recognition but in the service of B-form mimicry (**Figure 3.10**).

Overall, 6S RNA represents a unique regulator of bacterial RNAP. However, direct regulation of RNAP activity by ncRNAs is not limited to the bacterial 6S RNA. Mouse B2 and human Alu ncRNAs both repress mRNA transcription by binding to RNAP II (pol II) and blocking contacts with promoter DNA (Yakovchuk et al., 2009). Remarkably, pol II can utilize the B2 RNA as both a transcription template and a substrate, using an RNA-dependent RNAP activity to extend the 3' end of the B2 RNA in an internally templated reaction (Wagner et al., 2013). Extension of the B2 RNA in this way destabilizes the RNA and relieves the pol II inhibition. The functional parallels between the bacterial and eukaryotic ncRNA RNAP inhibitors suggest that the structural principles for 6S RNA binding and inhibition of *Eco* RNAP are widely applicable.

3.9 Acknowledgements

The work in this chapter is published in Molecular Cell. My role in this work involved: protein and RNA purification, cryo-EM sample preparation; data collection and processing; structural model building and analysis; biochemistry with mutant σ^{70} and σ^{S}; and conception of ideas and figures. Citation:

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

I would like to thank the following people for their contributions to this chapter: D. Acehan and K. Uryu (The Rockefeller University) assisted with EM sample preparation; E. A. Campbell and S. A. Darst (The Rockefeller University) provided mentorship and analyzed the structures and biochemical data; S. A. Darst (The Rockefeller University) edited this chapter; M. Ebrahim and J. Sotiris (The Rockefeller University) helped with cryo-EM data collection; A. Feklistov (The Rockefeller University) performed 6S RNA footprinting experiments; S. Feng and K. Leon (The Rockefeller University) helped with biochemical experiments; Z. Li. (Harvard University) and T. Walz (The Rockefeller University) for assistance with initial cryo-EM data collection and processing; W. T. McAllister (SUNY Health Science Center) provided mutant T7 RNAP for *in vitro* RNA synthesis; A. M. Pyle (Yale University) helped with RNA structural analysis; R. M. Saecker (The Rockefeller University) provided insightful discussions; K.M Wassarman (University of Wisconsin–Madison) performed RNA/protein crosslinking experiments, RNA binding assays, and RNA mutational analysis; J. T. Winkelman (University of Wisconsin–Madison) purified BPA-RNAP variants.

CHAPTER 4. Structural and functional analysis of Fidaxomicin bound to *Mtb* RNAP

4.1 Introduction

Mycobacterium tuberculosis (*Mtb*) is the causative agent of tuberculosis (TB), an infectious disease that causes active TB and kills almost two million people every year (World Health Organization, 2018). Bacterial RNA polymerase (RNAP) is a proven target for antibiotics (Boyaci et al., 2019b). For instance, Rifampicin (Rif), one of the first-line antibiotics to treat TB, targets the bacterial RNAP and inhibits RNA synthesis (Chakraborty and Rhee, 2015). The emergence of multi-drug resistant (MDR) TB has increased the demand for novel antibiotics to treat this disease (Zumla et al., 2015).

Fidaxomicin (Fdx) is an antibiotic used to treat infections caused by the Gram-positive bacterium, *Clostridium difficile* (*Cdf*) (Venugopal and Johnson, 2012). The chemical structure of Fdx is depicted in **Figure 4.1** (Serra et al., 2017). Like Rif, Fdx functions by inhibiting RNAP activity (Talpaert et al., 1975). However, Fdx does not display cross resistance with Rif and is effective against Rif-resistant bacteria (Gualtieri et al., 2006, 2009; Kurabachew et al., 2008; O'Neill et al., 2000), since it targets the switch region of RNAP, a site distinct from the Rif binding pocket (Srivastava et al., 2011). The switch region controls RNAP clamp motions that play an essential role in promoter melting (Boyaci et al., 2019a; Chakraborty et al., 2012; Feklistov et al., 2017).

In culture, Fdx is an effective drug against *Mtb* RNAP (Kurabachew et al., 2008; Srivastava et al., 2011). However, the clinical use of Fdx is restricted to treating *Cdf* infections due to poor absorption in the intestinal tract (Venugopal and Johnson, 2012). Addressing this limitation requires understanding the basis for Fdx binding and mechanism. Since Fdx binds the switch region, which is coupled to the mobile clamp domain, it was of interest to capture RNAP structures without the constraints of crystal packing. Cryo-electron microscopy (cryo-EM) serves as a powerful tool to observe multiple conformations of single molecules. Therefore, I assisted H. Boyaci, a post-doctoral fellow in the Darst-Campbell Lab (The Rockefeller University), in the determination of cryo-EM structures of *Mtb* transcription initiation complexes (TIC) with and without Fdx.

The cryo-EM structures reveal the molecular interactions between *Mtb* RNAP and Fdx as well as the basis of inhibition. The results show that RbpA, an Actinobacteria-specific general transcription factor (GTF) in *Mtb*, interacts with Fdx explaining its potency on *Mtb* RNAP in culture, relative to that of *Escherichia coli* (*Eco*). The structures also show that Fdx traps the clamp in an open conformation preventing the clamp from securing the DNA in the active site. In addition, the different structures [without DNA (±Fdx) and with DNA in the active site] also reveal the range of conformational states that the RNAP clamp domain adopts. Portions of this chapter have been previously published in eLife (Boyaci et al., 2018).

4.2 Cryo-EM analysis of *Mtb* RbpA/σ^A-holoenzyme

Using *in vitro* transcription assays, E. A. Campbell and H. Boyaci (The Rockefeller University) determined the IC_{50} of Fdx for *Mtb* RNAP (0.21 μM; Boyaci et al., 2018). H. Boyaci and I optimized freezing buffer conditions for *Mtb* RNAP using negative-stain EM and then prepared and froze cryo-EM grids of the *Mtb* RbpA and σ^A-holoenzyme (σ^A-holo) complex in the presence of 8mM cholamidopropyl]dimethylammonio)-2-hydroxy-1-propanesulfonate (CHAPSO) (**Chapter 2**; Chen et al., 2019a). Complexes of *Mtb* RbpA/σ^A-holo with and without Fdx were visualized using single-particle cryo-EM (**Figure 4.2**). Although the particles of *Mtb*

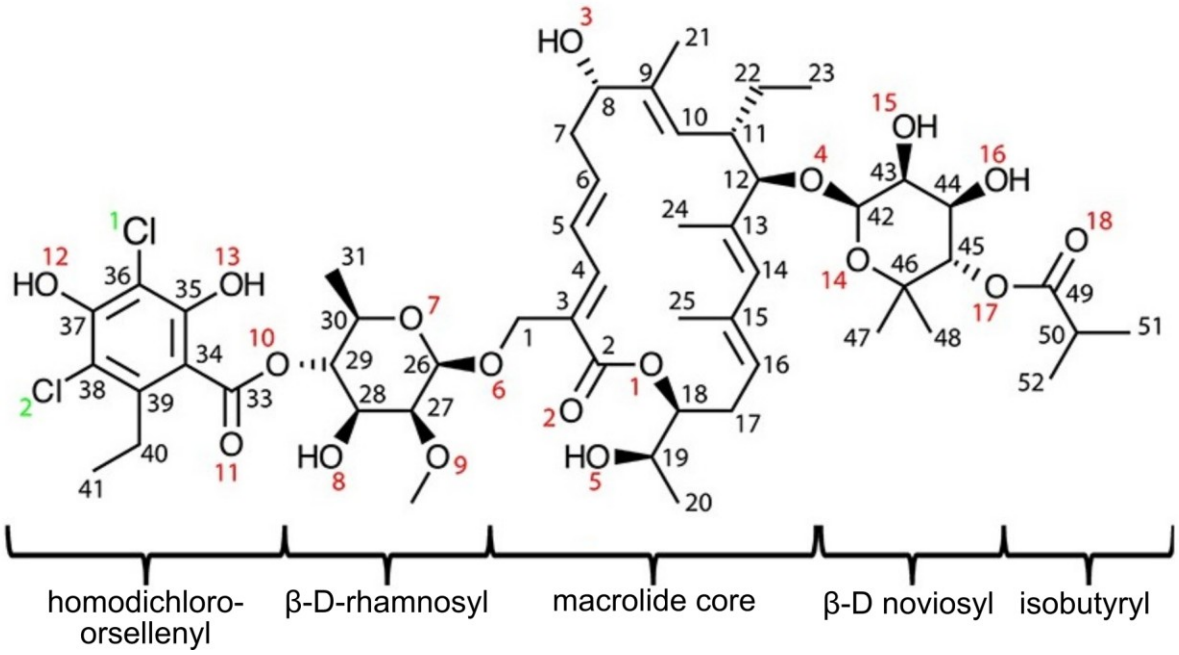


Figure 4.1. Chemical structure of Fidaxomicin (Serra et al., 2017)

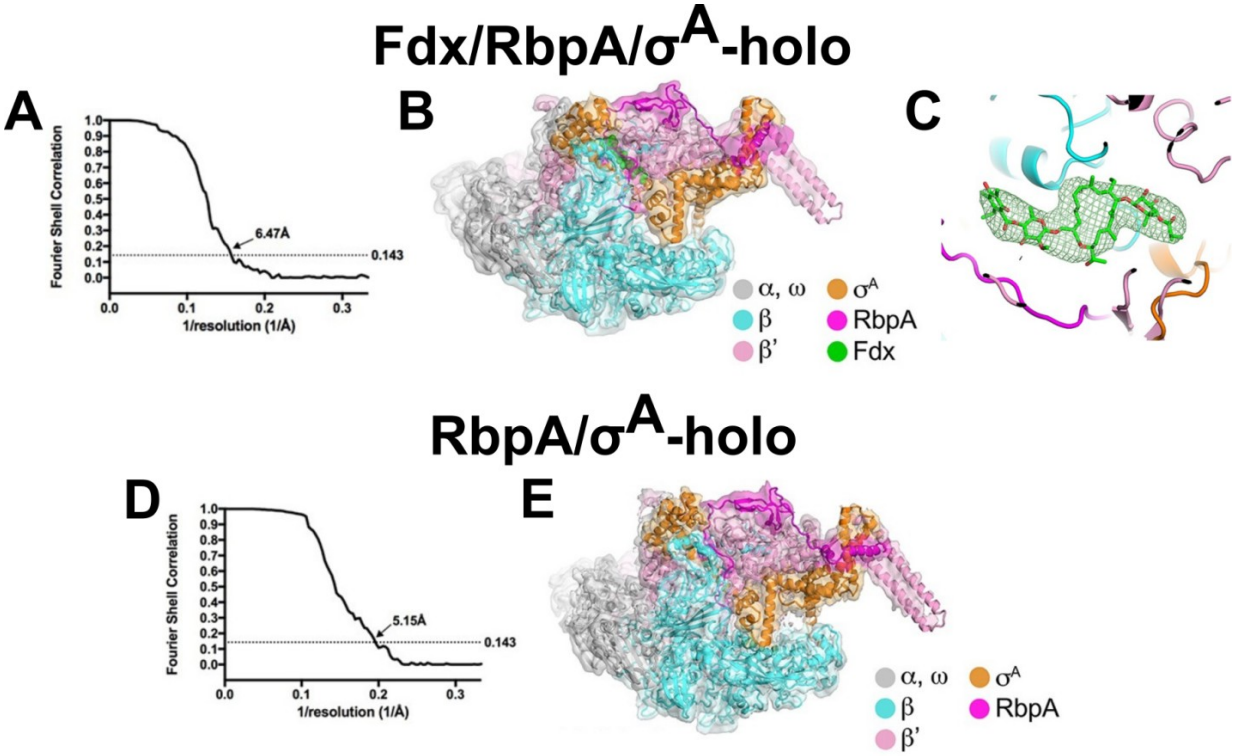


Figure 4.2. Cryo-EM structures of the *Mtb* RbpA/σ^A-holo with and without Fdx
(*top*) Cryo-EM of *Mtb* Fdx/RbpA/σ^A-holo. **(A)** Gold-standard FSC of the *Mtb* Fdx/RbpA/σ^A-holo complex. The gold-standard FSC was calculated by comparing the two independently determined half-maps from cryoSPARC (Punjani et al., 2017). The dotted line represents the 0.143 FSC cutoff which indicates a nominal resolution of 6.47 Å. **(B)** Cryo-EM density maps and superimposed refined models for *Mtb* Fdx/RbpA/σ^A-holo. **(C)** The difference map showing Fdx is bound in 6.47 Å resolution structure. (*bottom*) Cryo-EM of *Mtb* RbpA/σ^A-holo. **(D)** Gold-standard FSC of the *Mtb* RbpA/σ^A-holo complex. The gold-standard FSC was calculated by comparing the two independently determined half-maps from cryoSPARC (Punjani et al., 2017). The dotted line represents the 0.143 FSC cutoff which indicates a nominal resolution of 5.15 Å. **(E)** Cryo-EM density maps and superimposed refined models for and *Mtb* RbpA/σ^A-holo.

RbpA/σ^A-holo were prone to oligomerization, single particles were isolated using image classification and then used to determine 3D reconstructions of *Mtb* Fdx/RbpA/σ^A-holo and *Mtb* RbpA/σ^A-holo. The map of Fdx/RbpA/σ^A-holo (6.5 Å nominal resolution from 21,000 particles, **Figure 4.2A**) had sufficient resolution to generate a model using rigid-body refinement in PHENIX (Adams et al., 2010) (**Figure 4.2B**) and to visualize cryo-EM density for Fdx (**Figure 4.2C**). The cryo-EM map without Fdx (5.2 Å nominal resolution from 88,000 particles; **Figure 4.2D**) was used to generate a model from PHENIX rigid-body refinement (Adams et al., 2010) (**Figure 4.2E**)**.** Overall, the cryo-EM density maps were of sufficient detail to visualize the bound antibiotic in the Fdx complex and to determine the domain organization of the RNAP (**Figures 4.2B and 4.2E**)**.**

4.3 Cryo-EM structures of *Mtb* RbpA/σ^A-holo/(us-fork)$_2$ [RbpA/RPo mimic] and *Mtb* Fdx/RbpA/σ^A-holo/us-fork

To reduce oligomerization, H. Boyaci (The Rockefeller University) added a promoter DNA fragment containing the upstream-fork junction (us-fork; **Figure 4.3A**) to the *Mtb* Fdx/RbpA/σ^A-holo samples. I assisted her in data collection and processing (**Figure 4.3B**). After particle classification of nearly 600,000 cryo-EM images of individual particles, two major classes were observed: (1) RbpA/σ^A-holo/(us-fork)$_2$ and (2) Fdx/RbpA/σ^A-holo/us-fork (**Figure 4.3B**). The first class, RbpA/σ^A-holo/(us-fork)$_2$, comprised of *Mtb* RbpA/σ^A-holo bound to two us-fork promoter fragments with a nominal resolution of 3.3 Å (**Figure 4.4A**). Local resolution calculations of this map indicated that the central core of the structure was determined between 2.8-3.2 Å in resolution (**Figure 4.4B**)**.** There was no observable density for Fdx but the N-terminal tail (NTT) of RbpA was visible (outlined in red, **Figure 4.4B**). The second class, Fdx/RbpA/σ^A-holo/us-fork, comprised of *Mtb* Fdx/RbpA/σ^A-holo bound to the us-fork as expected with a nominal resolution of 3.4 Å (**Figure 4.4C**). Local resolution calculations of this map indicated that the central core of the structure was resolved to 2.9-3.4 Å (**Figure 4.4D**). Density for the bound Fdx was easily visible in this class (outlined in red, **Figure 4.4D**). Atomic models were built for both classes (**Table 4.5**).

The *Mtb* RbpA/σ^A-holo/(us-fork)$_2$ structure consists of one us-fork promoter fragment bound upstream from the RNAP active site cleft and a second us-fork promoter fragment bound to the RNAP downstream DNA channel (**Figure 4.6A**). The binding of two DNA forks was unexpected since the fork (**Figure 4.3A**) was designed to bind outside of the RNAP cleft and not to affect the clamp conformation. The 5-nucleotide 3'-overhang of the second fork DNA (**Figure 4.3A**) is engaged with the RNAP active site [as the template-strand (t-strand)] like previously characterized 3'-tailed templates (Gnatt et al., 2001; Kadesch and Chamberlin, 1982). The overall conformation of this protein complex and its engagement with the upstream and downstream DNA fragments were very similar to the crystal structure of a full *Mycobacterium smegmatis* (*Msm*) open promoter complex (RPo) (Hubin et al., 2017a). However, unlike previous crystal structures of *Msm* RNAP, the cryo-EM map reveals density for the RbpA NTT (RbpA^{NTT}) (**Figures 4.4B**, **4.6A**). This complex will be denoted as an *Mtb* RbpA/RPo mimic (**Figure 4.6A**).
The cryo-EM reconstruction of *Mtb* Fdx/RbpA/σ^A-holo/us-fork comprised of one us-fork promoter fragment and bound to Fdx. The us-fork promoter fragment was bound outside the RNAP active site cleft, as expected, with the -35 and -10 promoter elements engaged with the $\sigma^A{}_4$ and $\sigma^A{}_2$ domains, respectively (**Figure 4.6B**). Like RbpA/σ^A-holo/(us-fork)$_2$, this structure also contains density for RbpA^{NTT} in the RNAP active site cleft.

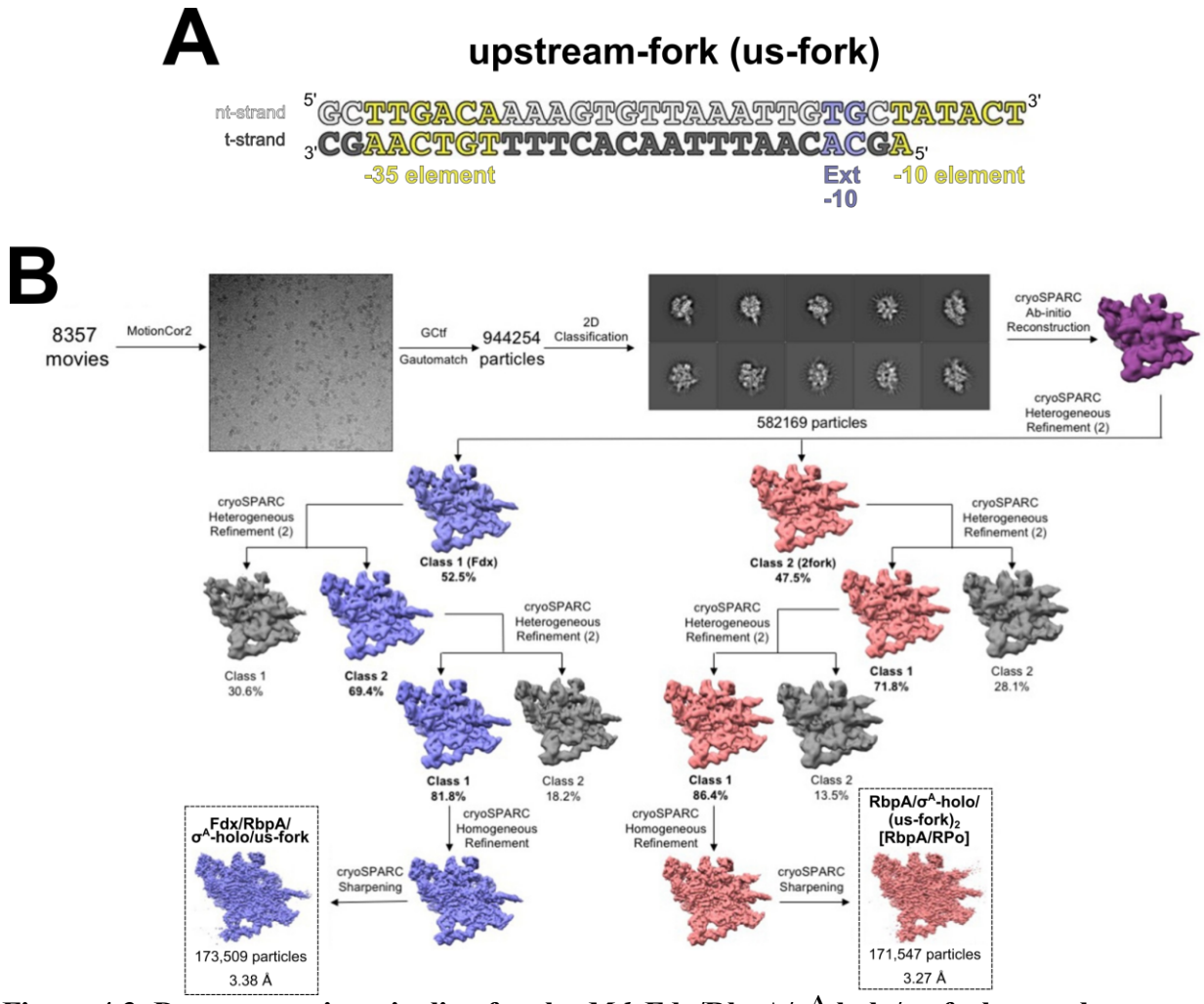


Figure 4.3. Data processing pipeline for the *Mtb* Fdx/RbpA/σ^A-holo/us-fork complexes
(A) Synthetic fragment of an upstream promoter fork used for cryo-EM experiments. The DNA sequence is derived from the full con promoter (Gaal et al., 2001). The nontemplate-strand (nt-strand) DNA (top strand) is colored light gray and the t-strand DNA (bottom strand) is colored dark grey. The -35 and -10 elements are highlighted in yellow. The extended -10 (Keilty and Rosenberg, 1987) is colored slate. **(B)** Flowchart showing the image processing pipeline for the cryo-EM data of Mtb Fdx/RbpA/σ^A-holo/us-fork complexes. Extracted particles were sorted in a binomial-like fashion using 3D heterogeneous refinement in cryoSPARC (Punjani et al., 2017). Multiple rounds of 3D classification were performed; colored classes represent the selected classes while grey classes were thrown out. The final sets of particles for each class were refined using cryoSPARC homogenous refinement and then sharpened for model building. Two major classes were observed: Fdx/RbpA/σ^A-holo/us-fork containing 173,509 particles at a nominal resolution of 3.38 Å (blue) and RbpA/σ^A-holo/(us-fork)$_2$ [RbpA/RPo] containing 171,547 particles at a nominal resolution of 3.27 Å (red).

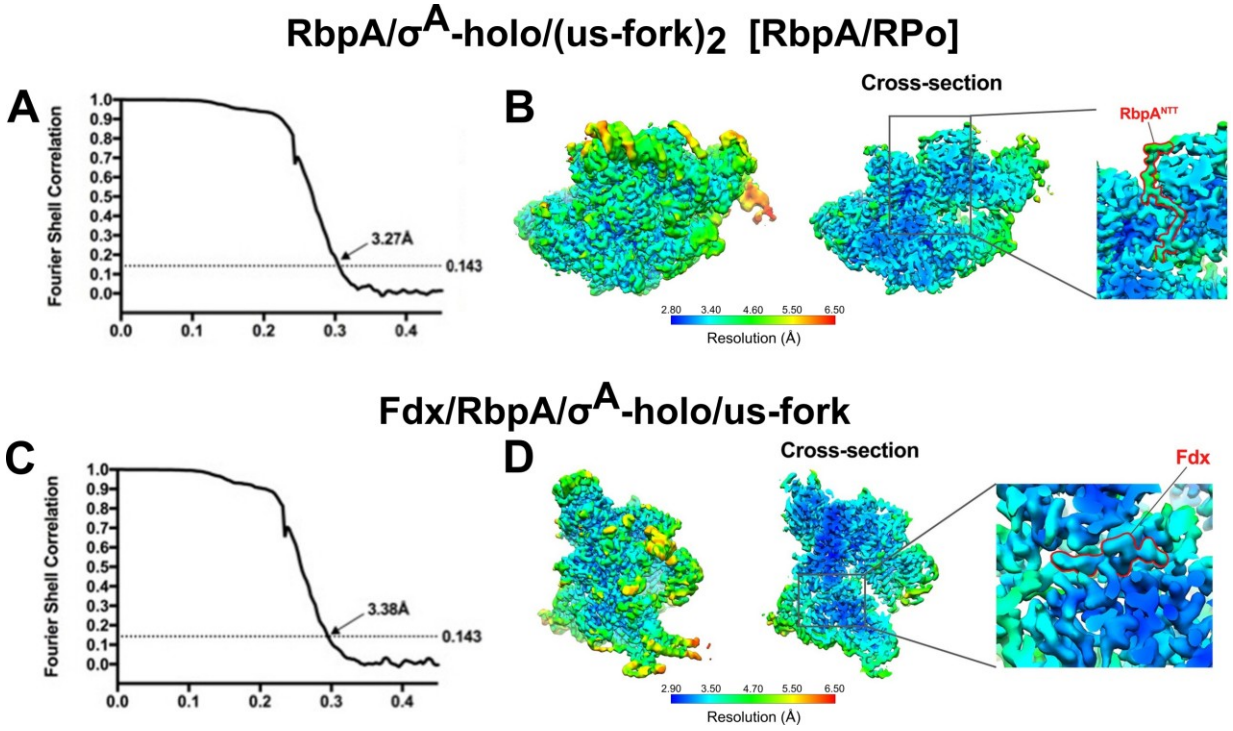


Figure 4.4. FSC and local resolution calculations for RbpA/σ^A-holo/(us-fork)$_2$ [RbpA/RPo] and Fdx/RbpA/σ^A-holo/us-fork

(*top*) RbpA/σ^A-holo/(us-fork)$_2$ [RbpA/RPo]. **(A)** Gold-standard FSC, calculated by comparing the two independently determined half-maps from cryoSPARC (Punjani et al., 2017). The dotted line represents the 0.143 FSC cutoff which indicates a nominal resolution of 3.27 Å. **(B)** Views of the *Mtb* RbpA/RPo mimic cryo-EM map colored by local resolution (Cardone et al., 2013). The left view shows the entire map, while the middle view shows a cross-section of the map sliced at the level of the RbpA^{NTT}. The boxed region is magnified on the right. Density for the RbpA^{NTT} is outlined in red. (*bottom*) Fdx/RbpA/σ^A-holo/us-fork. **(C)** Gold-standard FSC, calculated by comparing the two independently determined half-maps from cryoSPARC (Punjani et al., 2017). The dotted line represents the 0.143 FSC cutoff which indicates a nominal resolution of 3.38 Å. **(D)** Views of the cryo-EM map colored by local resolution (Cardone et al., 2013). The left view shows the entire map, while the middle view shows a cross-section of the map sliced at the level of the Fdx binding pocket. The boxed region is magnified on the right. Density for the Fdx molecule is outlined in red.

Table 4.5. Model statistics calculated in Molprobity (Chen et al., 2010b)

	Fdx/RbpA/σ^A-holo/us-fork	RbpA/σ^A-holo/(us-fork)$_2$
Resolution[a]	3.38 Å	3.27 Å
Molprobity score	2.29	2.42
Clashscore (all atoms)	10.79	5.88
Rotamer outliers	7.13%	13.48%
RMS deviations bonds (Å)	0.002	0.007
RMS deviations angles (°)	0.491	0.837
Ramachandran favored	89.71%	92.14%
Ramachandran outliers	0.30%	0.03%

[a] Gold standard FSC 0.143 cutoff criteria (Rosenthal and Henderson, 2003)

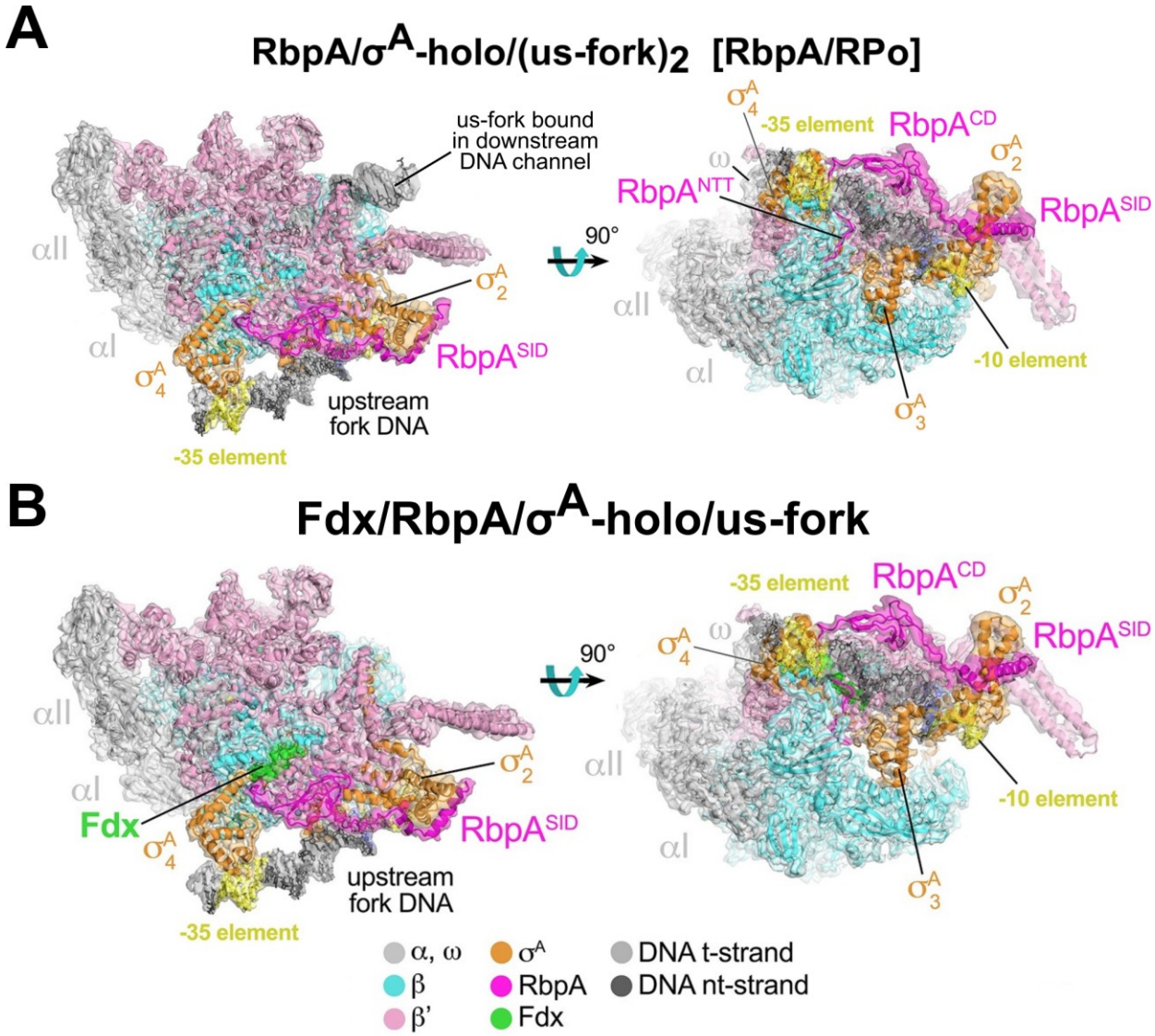


Figure 4.6. Cryo-EM structures of *Mtb* RbpA/σ^A-holo/(us-fork)$_2$ [RbpA/RPo] and *Mtb* Fdx/RbpA/σ^A-holo/us-fork

(*top*) RbpA/σ^A-holo/(us-fork)$_2$ [RbpA/RPo]. **(A)** The 3.3 Å resolution cryo-EM density map of the *Mtb* RbpA/σ^A-holo/(us-fork)$_2$ complex (RbpA/RPo mimic) is rendered as a transparent surface colored as labeled. Superimposed is the final refined model; proteins are shown as a backbone ribbon, nucleic acids are shown in stick format. Two views of the structure are shown. (*bottom*) Fdx/RbpA/σ^A-holo/us-fork. **(B)** The 3.4 Å resolution cryo-EM density map of the *Mtb* Fdx/RbpA/σ^A-holo/us-fork complex is rendered as a transparent surface colored as labeled. Superimposed is the final refined model; proteins are shown as a backbone ribbon, Fdx and the nucleic acids are shown in stick format. Two views of the structure are shown.

4.4 The N-terminal tail of RbpA is located the RNAP active site cleft

In addition to the NTT, RbpA also contains a core domain ($RbpA^{CD}$), a basic linker ($RbpA^{BL}$), and a sigma interacting domain ($RbpA^{SID}$) (Bortoluzzi et al., 2013; Hubin et al., 2017b; Tabib-Salazar et al., 2013). Previous crystal structures of *Msm* TICs containing RbpA showed that the $RbpA^{SID}$ interacts with the σ^{A}_{2} domain, the $RbpA^{BL}$ contacts with the phosphate backbone of the promoter DNA just upstream of the -10 element, and the $RbpA^{CD}$ interacts with the RNAP β' Zinc-Binding-Domain (ZBD) (Hubin et al., 2017a, 2017b). Density for the $RbpA^{NTT}$ (residues 1-25) was never observed in the crystal structures and presumed disordered. The cryo-EM structure of RbpA/σ^{A}-holo/(us-fork)$_2$ reveal density for the $RbpA^{NTT}$, which threads through a narrow channel, created by the ZBD and σ^{A}_{4} domains, into the RNAP active site (**Figures 4.7A-B**). Conserved residues of the $RbpA^{NTT}$ make interactions with the σ-finger ($\sigma^{A}_{3.2}$-linker), ZBD and β'lid (Boyaci et al., 2018).

The N-terminal RbpA residues of the tail are too far from the active site Mg^{2+} to have a direct role in RNAP catalytic activity or nucleotide (NTP) substrate binding. The σ-finger plays an indirect role in transcription initiation by positioning the t-strand DNA for *de novo* phosphodiester bond formation (Kulbachinskiy and Mustaev, 2006; Zhang et al., 2012). In *Eco,* the σ-finger also plays a major role in abortive initiation and promoter escape by physically blocking the path of the elongating RNA transcript before σ release and promoter escape (Cashel et al., 2003; Murakami et al., 2002b). The intimate association of the $RbpA^{NTT}$ with the σ-finger (**Figure 4.7B**) suggests that it may participate in promoter escape. A recent study, examining *Mtb* RNAP transcription initiation and promoter escape by fluorescence kinetic assays, validates this proposed idea (Jensen et al., 2019). This is consistent with findings that the $RbpA^{NTT}$ does not strongly affect RPo formation but plays a significant role in *Msm* gene expression *in vivo* (Hubin et al., 2017b).

4.5 Fdx interacts with the RNAP β and β' subunits, σ^{A}, and RbpA

The structure of the *Mtb* Fdx/RbpA/σ^{A}-holo/us-fork class (**Figure 4.6B**) reveals unambiguous density for Fdx. Fdx binding interacts with four protein components of the complex and buries a large accessible surface area of 4,800 Å^2 (β, 2,100 Å^2; β', 2,000 Å^2; σ^{A}, 300 Å^2; RbpA, 330 Å^2). The structure clearly identified Fdx-interacting residues (**Figures 4.8A-B**). Notably, the Fdx/RNAP interaction is stabilized by two cation-π interactions: (1) β' R84 and the aromatic ring of the Fdx homodichloro-orsellinic acid moiety and (2) β' R89 and the conjugated double-bond system centered between C4 and C5 of the macrolide core (**Figures 4.1**, **4.8B**). Resolution in this region was sufficient to model six water molecules, four of which mediate Fdx/RNAP interactions (β' R89, β' D404, β' Q415, and RbpA E17). (**Figure 4.8B**). Fdx interacts with residues from eight distinct structural elements of the *Mtb* RNAP: β Switch 3 (Sw3), β Switch 4 (Sw4), β residues belonging to the clamp domain, β'ZBD, β'lid, β' Switch 2 (Sw2), the σ-finger, and $RbpA^{NTT}$ (**Figure 4.8B**).

Amino-acid substitutions conferring Fdx resistance (Fdx^{R}) have been identified in RNAP β or β' subunits from *Bacillus subtilis* (Gualtieri et al., 2006), *Cdf* (Kuehne et al., 2018), *Enterococcus faecalis* (Gualtieri et al., 2009), and *Mtb* (Kurabachew et al., 2008). These mutations correspond to β Q1054 (Sw3), β V1100, β V1123 (clamp), β' R89 (ZBD), β' P326 (lid), and β' R412 (Sw2) *Mtb* RNAP. These residues make direct interactions with Fdx in the cryo-EM and support our findings (**Figure 4.8B).** All five chemical moieties of Fdx (**Figure 4.1**) interact with at least one RNAP residue that confers Fdx^{R} when mutated (**Figure 4.8B**),

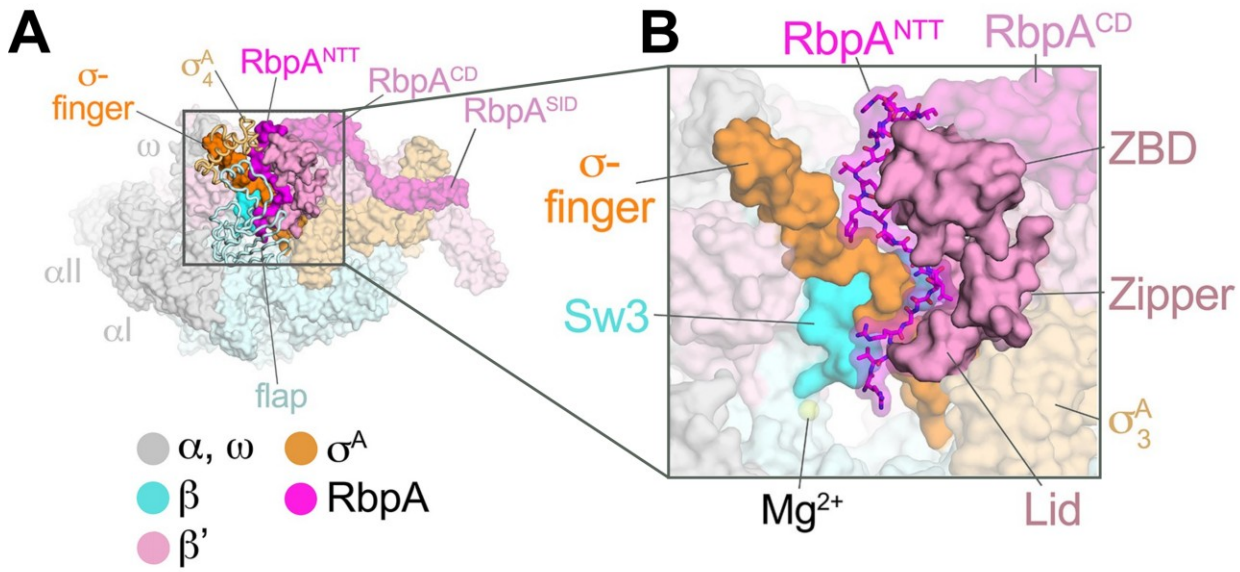


Figure 4.7. The RbpA^{NTT} interacts with conserved structural elements in the RNAP active site cleft

(A) The overall RbpA/RPo structure is shown as a color-coded molecular surface (based on the key). For clarity, the DNA is not shown and the β flap and σ^A_4 domain are shown as backbone worms to reveal the RbpA^{NTT} (magenta) underneath. The boxed region is magnified in panel **(B)**. **(B)** Magnified view of the boxed region from panel **(A)**. The RbpA^{NTT} is shown in stick format with a transparent molecular surface. Conserved RNAP structural elements that interact with the RbpA^{NTT} are highlighted at non-transparent molecular surfaces: βSw3, β'ZBD, β'Zipper, β'Lid, and σ-finger). Figure courtesy of S. A. Darst (The Rockefeller University).

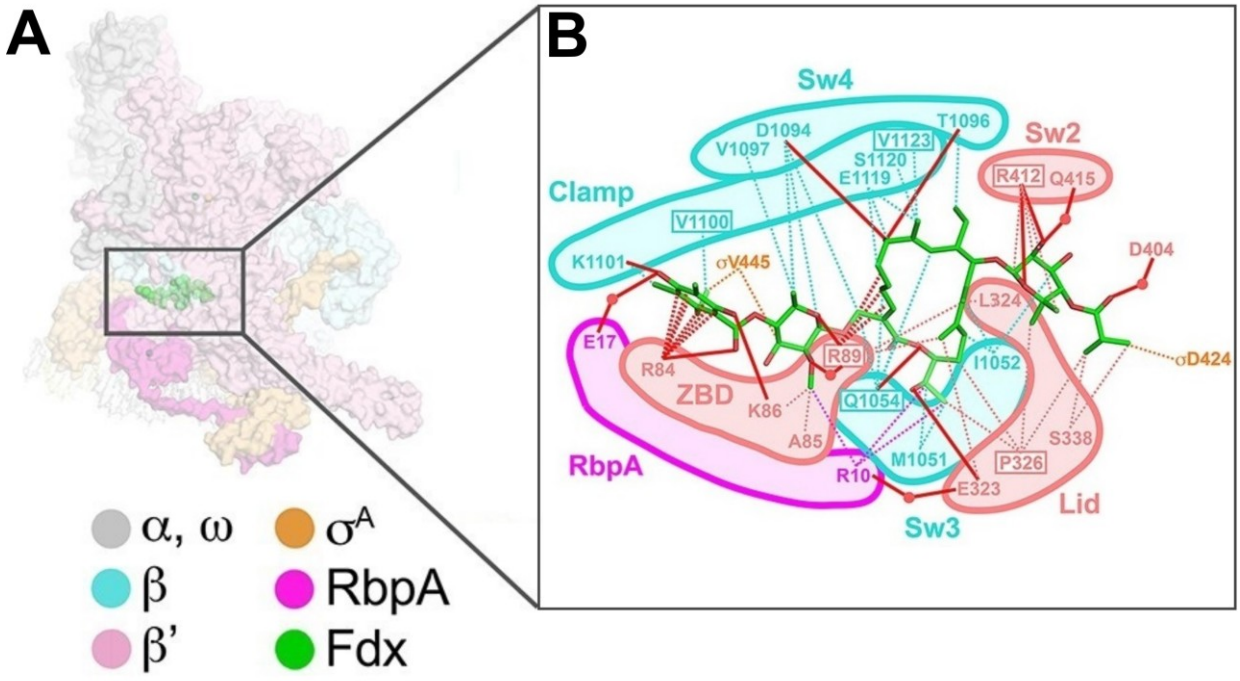


Figure 4.8. Structural basis for Fdx inhibition of *Mtb* transcription

(A) Overview of the Fdx/RbpA/σ^A-holo/us-fork structure, shown as a molecular surface and colored as indicated by the key. The DNA fork is omitted for clarity. The boxed region is magnified on the right. **(B)** Schematic summary of the Fdx contacts with σ^A-holo and RbpA. Fdx is shown in stick format with green carbon atoms. Thin dashed lines represent van der Waals contacts (≤ 4.5 Å), thick red lines represent hydrogen bonds (< 4 Å). The thin dashed red lines denote cation-π interactions. Residues interacting with Fdx are clustered together and color based on the RNAP structural element that they belong to. Figure courtesy of E. A. Campbell and S. A. Darst (The Rockefeller University).

suggesting that each moiety may be important for Fdx action. In addition to the β and β' subunits, Fdx interacts with residues of the σ-finger (D424 and V445) and the $RbpA^{NTT}$ (**Figure 4.8B**).

4.6 $RbpA^{NTT}$ is critical for Fdx potency against mycobacterial RNAP *in vitro* and *in vivo*

$RbpA^{NTT}$ in the RNAP interacts with Fdx and we hypothesized that that interaction contributes to the high Fdx sensitivity of *Mtb* RNAP compared to other bacterial RNAPs. H. Boyaci (The Rockefeller University) tested this hypothesis using *in vitro* transcription assays comparing the IC_{50} of Fdx on *Mtb* RNAP with RbpA, without RbpA, and with a ΔNTT RbpA mutant ($RbpA^{\Delta NTT}$). Truncation of the NTT caused a 33-fold increase in resistance to Fdx confirming the significance of the interaction between $RbpA^{NTT}$ and Fdx (Boyaci et al., 2018).

RbpA is essential in *Mtb* and *Msm* but strains carrying $RbpA^{\Delta NTT}$ are slow-growing but viable in *Msm* (Hubin et al., 2017b). Therefore, we tested the role of $RbpA^{NTT}$ in Fdx growth inhibition in *Msm* cells. E. A. Campbell (The Rockefeller University) performed zone of inhibition assays on two isogenic *Msm* strains: one harboring wild-type (WT) RbpA ($RbpA^{WT}$) and the other $RbpA^{\Delta NTT}$ (Hubin et al., 2017b). The *Msm* $RbpA^{\Delta NTT}$ strain grew considerably slower on plates, taking approximately twice the time as WT *Msm* to reach confluency. Despite the growth defect, the $RbpA^{\Delta NTT}$ strain was significantly less sensitive to Fdx. At 500 μM Fdx, the inhibition zone for $RbpA^{\Delta NTT}$ was significantly smaller than for $RbpA^{WT}$. By contrast, a control disk containing 860 μM streptomycin, a protein synthesis inhibitor, produced equal inhibition zones for the $RbpA^{WT}$ and $RbpA^{\Delta NTT}$ strains (Boyaci et al., 2018). Thus, the essential role of RbpA in *Mtb* transcription is key to the relatively high sensitivity of *Mtb* cells to Fdx.

4.7 Fdx acts like a doorstop to stabilize the open-clamp conformation

The RNAP switch regions connect the mobile clamp domain to the rest of the RNAP (Gnatt et al., 2001; Lane and Darst, 2010a). A number of bacterial RNAP inhibitors (ex. myxopyronin, corallopyronin, and ripostatin) directly bind the Sw1 and Sw2 regions, stabilizing the clamp domain into a closed conformation (Belogurov et al., 2009; Boyaci et al., 2019a; Mukhopadhyay et al., 2008). Although the Fdx binding determinants do not overlap the sites for these other inhibitors, Fdx-RNAP interactions around the β'Sw2, βSw3, and βSw4 regions (**Figure 4.8B**) suggest that Fdx plays a role in clamp conformations.

To understand the role of Fdx in clamp movement, we compared the RNAP conformational states from the four solution cryo-EM structures that were solved in this chapter. These structures can be directly compared since they are from the same bacterial RNAP and are not influenced by crystal packing forces. The four structures were superimposed by the structural core module (**Table 4.9**), comprising the ω subunit and highly conserved β and β′ regions in or near the active center that have not been observed to undergo significant conformational changes in dozens of RNAP structures. Using the RbpA/RPo structure (**Figure 4.6A**) as a reference, the structures superimposed with root-mean-square deviations (RMSDs) < 0.4 Å over at least 898 aligned alpha carbon (Cα) atoms of the structural core module, but RMSDs > 9 Å for 461 Cα-atoms of the clamp modules (**Table 4.9**), indicating large shifts of the clamp module with respect to the rest of the RNAP in the different complexes.

In the aligned structures, the clamp conformations can be characterized as rigid body rotations about a common rotation axis (**Figure 4.10**). The closed clamp of the RbpA/RPo structure was used as the reference and was assigned a clamp rotation angle of 0° (blue, **Figure 4.10**). The RbpA/σ^A-holo clamp is rotated open by about 12° (green, **Figure 4.10).** Because the complex is not bound to any ligands that alter the clamp conformation (such as Fdx or DNA), we

Table 4.9. Superimposition of cryo-EM structures

Structure	Total # of Cα's aligned (structural core)[a]	RMSD (Å) (structural core alignment)	Clamp[b] RMSD (Å) (461 Cα's)
RbpA/σ^A-holo/(us-fork)$_2$ (reference)	1,034	0	0
RbpA/σ^A-holo	900	0.386	9.01
Fdx/RbpA/σ^A-holo/us-fork	898	0.384	12.4
Fdx/RbpA/σ^A-holo	900	0.386	13.5

[a]Structural core module: β residues 30-53, 177-182, 370-380, 445-639, 705-747, 879-1116; β' 414-443, 496-863, 1247-1282; ω.
[b]Clamp module: β residues 1117-1140; β' residues 1-140, 231-413, 1219-1245.

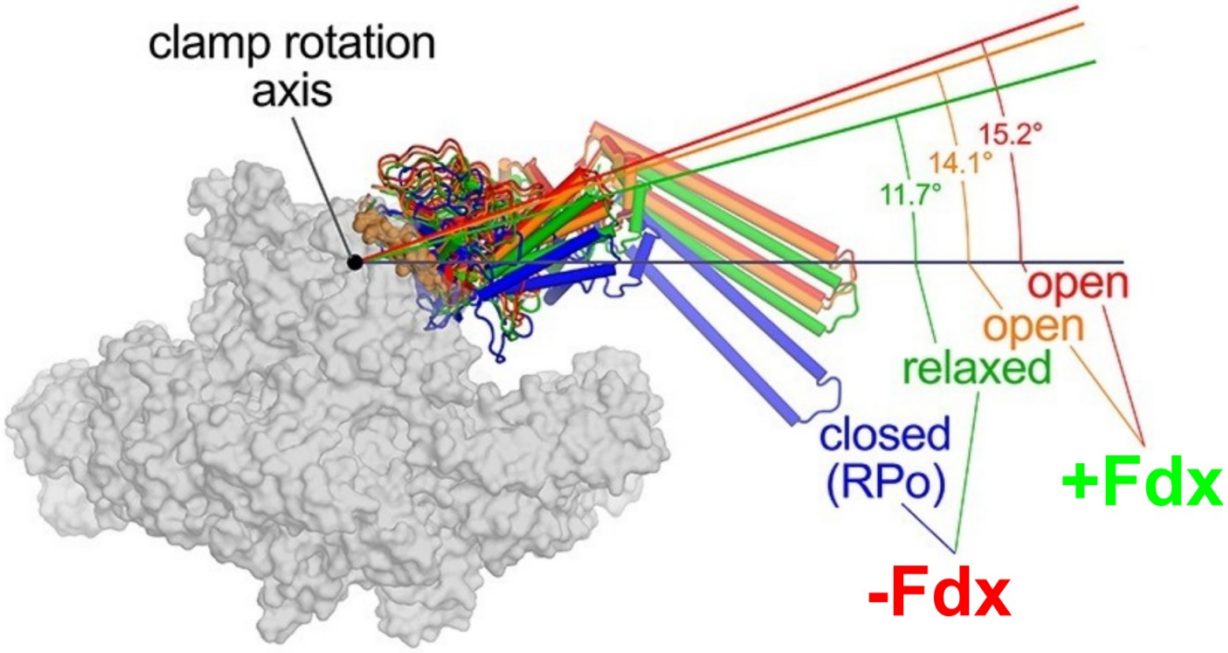


Figure 4.10. Clamp positions of cryo-EM structures of *Mtb* RbpA/σ^A-holo
RNAP clamp conformational changes for four cryo-EM structures determined in this chapter. The RbpA/RPo mimic **(Figure 4.6A)** structure was used as a reference for the superimposition of the other structures. Structures were aligned by α-carbon atoms of the structural core module (shown as a gray molecular surface), revealing large shifts in the clamp modules **(Table 4.9)**. The clamp modules are shown as backbone cartoons with cylindrical helices and color-coded (closed clamp of RbpA/RPo mimic, blue; relaxed clamp of RbpA/σ^A-holo, green; open clamp of Fdx/RbpA/σ^A-holo/us-fork, orange; open clamp of Fdx/RbpA/σ^A-holo, red). The clamp conformational changes can be characterized as rigid body rotations about a rotation axis perpendicular to the page (denoted by a black point). The angles of clamp opening, relative to the closed RbpA/RPo mimic clamp (blue; 0° opening) for the different structures are shown.

termed this the 'relaxed' conformation. The two Fdx-bound complexes, with or without us-fork DNA, show further opening of the clamp (14° and 15°, respectively; orange and red in **Figure 4.10**).

In the high-resolution Fdx/TIC structure (**Figure 4.10**), Fdx binds in a narrow gap between the open clamp module and the rest of the RNAP, via the switch regions (**Figure 4.11A**). The examination of the high-resolution RPo (closed clamp with DNA in the active site) structure reveals that clamp closure required to secure promoter DNA in the channel clashes with the Fdx binding pocket (**Figure 4.11B**). Fdx binding is only compatible with the open clamp conformation of RNAP. Thus, Fdx acts like a doorstop, wedging open the clamp.

4.8 Discussion

Clamp dynamics play multiple important roles in the transcription cycle. Changes in the switch regions and their correlation to the clamp conformations were first noted by comparing crystal structures of free RNAPs (Cramer et al., 2001; Zhang et al., 1999) with the crystal structure of an elongation complex (Gnatt et al., 2001). In the elongation complex, the downstream duplex DNA and RNA/DNA hybrid occupy the RNAP active-site cleft making interactions that close the clamp around the nucleic acids. This closing of the clamp on the nucleic acid substrate was proposed as an explanation for the high processivity of the transcription elongation complex. Numerous subsequent crystal structures have supported the idea that stable, transcription-competent complexes of RNAP with nucleic acids, either RPo (Bae et al., 2015a; Hubin et al., 2017a; Zuo and Steitz, 2015) or elongation complexes (Gnatt et al., 2001; Kettenberger et al., 2004; Vassylyev et al., 2007), correlate with the closed-clamp conformation. However, the effects of crystal packing forces on the observed clamp conformation could not be excluded. Observations of clamp positions by solution Förster resonance energy transfer (FRET) (Chakraborty et al., 2012), and more recently in cryo-EM structures (Bernecky et al., 2016; Boyaci et al., 2019a; Chen et al., 2017, 2019b; Hoffmann et al., 2015; Kang et al., 2017; Neyer et al., 2016) have confirmed the relationship between stable nucleic-acid RNAP complexes and the closed clamp conformation. Clamp motions have also been shown to play a critical role in promoter melting to form RPo during transcription initiation (Feklistov et al., 2017).

The cryo-EM structures show that Fdx traps the open clamp RNAP conformation in the absence of crystal packing (**Figure 4.11A**). Thus, Fdx inhibits transcription by hindering stable RNAP-nucleic acid binding or by preventing clamp motions required for RPo formation, or both (**Figure 4.11B**). These results are broadly consistent with mechanistic analyses demonstrating that Fdx blocks early steps of duplex promoter melting but providing RNAP with a pre-melted template overcomes this block (Morichaud et al., 2016; Tupin et al., 2010). It was proposed that Fdx likely prevented the clamp from closing, which is consistent with the structural findings of this work.

A recent study used cryo-EM to solve a structure of *Mtb* RNAP bound Fdx in combination with solution FRET (Lin et al., 2018). Although this study shows the same open clamp RNAP conformation in the presence of Fdx, this study has a number of caveats. The authors do not determine the normal state of the *Mtb* RNAP clamp (in the absence of ligands), rather they infer that Fdx opens the clamp by comparing FRET studies of *Eco* RNAP. Crucially, their structure and experiments do not include RbpA and therefore do not address the role of RbpA in Fdx activity.

The work presented in this chapter establishes Fdx molecular interactions with *Mtb* RNAP (**Figure 4.8**) and provides a mechanism of action for Fdx (**Figure 4.11**). The essential

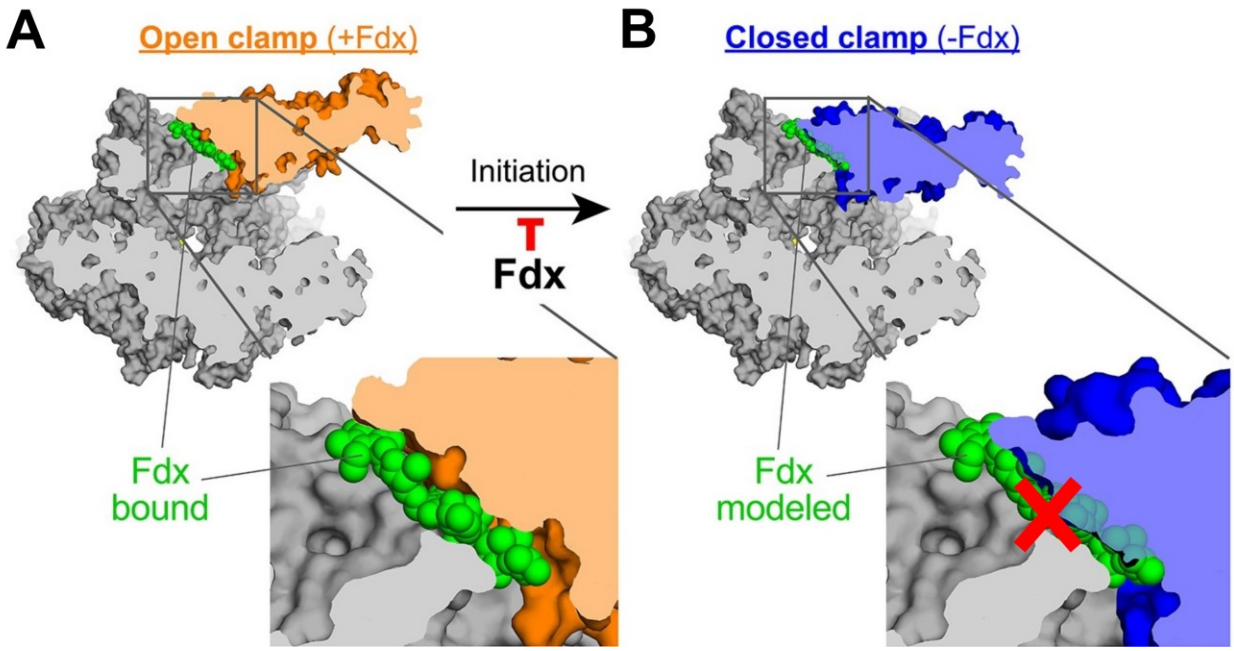


Figure 4.11. Mechanism of Fdx inhibition of bacterial RNAP
(A) The core module of RNAP from the 3.4 Å resolution Fdx/RbpA/σ^A-holo/us-fork structure is shown as a gray molecular surface but with the open clamp colored orange. The structure is sliced at the level of the Fdx binding pocket (the bound Fdx is shown in green spheres). The boxed region is magnified below, showing the tight fit of the Fdx molecule in a narrow gap between the clamp and the rest of the RNAP. **(B)** The core module of RNAP from the 3.3 Å resolution RbpA/RPo structure is shown as a gray molecular surface but with the closed clamp colored blue. The structure is sliced at the level of the empty Fdx binding pocket. Fdx, modeled from the structure shown in **(A)**, is docked into the pocket and shown in green. The boxed region is magnified below. Fdx cannot bind to RNAP with a closed clamp because clamp closure pinches off the Fdx binding site. This clash is indicated by a red 'X'.

actinobacterial GTF RbpA is responsible for the high sensitivity of mycobacterial RNAP to Fdx both *in vitro* and *in vivo* (Boyaci et al., 2018). This new knowledge provides a structural platform for the development of antimicrobials that target the Fdx binding determinants and underscores the need to define structure-activity relationships of drug leads using near-native states, in this case using cryo-EM with the RbpA/σ^A-holo complex to guide the development of effective *Mtb* treatments.

4.9 Acknowledgments

The work in this chapter is published in eLife. My role in this work involved: cryo-EM sample preparation; data collection and processing; structural analysis; conception of ideas and figures. Citation:

Boyaci, H., Chen, J., Lilic, M., Palka, M., Mooney, R.A., Landick, R., Darst, S.A., and Campbell, E.A. (2018). Fidaxomicin jams *Mycobacterium tuberculosis* RNA polymerase motions needed for initiation via RbpA contacts. Elife 7, 1–18.

I would like to thank the following people for their contributions to this chapter: H. Boyaci, E. A. Campbell, and M. Lilic (The Rockefeller University) performed biochemical and zone of inhibition assays; E. A. Campbell (The Rockefeller University) edited this chapter; E. A. Campbell and S. A. Darst (The Rockefeller University) provided mentorship and analyzed structures and data; M. Ebrahim. and J. Sotiris (The Rockefeller University) helped with cryo-EM data collection; E. Eng, M. Kopylov, and L.Yen (New York Structural Biology Center) helped with cryo-EM data collection; R. A. Mooney, R. Landick, and M. Palka (University of Wisconsin–Madison) provided *Mtb* RNAP constructs and insightful discussions; K. Uryu (The Rockefeller University) assisted with EM sample preparation.

CHAPTER 5. Structural examination of *Eco* TraR in transcription initiation

5.1 Introduction

Transcription initiation is a major control point for gene expression. A number of transcription factors regulate this process by binding to the promoter DNA and/or to RNA polymerase (RNAP) directly (Browning and Busby, 2016; Haugen et al., 2008). Bacterial RNAP-binding factors interact with several structural elements of the RNAP to regulate its functions (Haugen et al., 2008). An example of a RNAP-binding factor is DksA, an essential protein for the bacterial stringent response to cellular stress. During the stringent response, bacteria reprogram their metabolism by changing the expression of hundreds of genes that affect processes such as ribosome biogenesis and amino acid synthesis. Effects of DksA on transcription are potentiated by the alarmone guanosine pentaphosphate (ppGpp), a molecule synthesized by RelA during amino acid starvation (Brown et al., 2016; Cashel and Gallant, 1969; Ryals et al., 1982). Together, ppGpp and DksA alter the expression of many genes (Paul et al., 2004a, 2005; Sanchez-Vazquez et al., 2019), for example, inhibiting promoters responsible for ribosome biogenesis and activating promoters responsible for amino acid synthesis.

DksA regulates transcription by allosterically binding to RNAP in the secondary channel (Lennon et al., 2012; Molodtsov et al., 2018; Perederina et al., 2004) (**Figure 5.1A**). ppGpp binds directly to RNAP at two binding sites: site 1, located at the interface of the β′ and ω subunits (Ross et al., 2013; Zuo et al., 2013), and site 2, located at the interface of β′ and DksA (Molodtsov et al., 2018; Ross et al., 2016) (**Figure 5.1A**). The ppGpp bound at site 1 inhibits transcription ~2-fold, however, the effects of ppGpp bound at both sites together with DksA are as much as 20-fold (Paul et al., 2004b; Ross et al., 2016). By contrast, DksA and ppGpp bound at site 2 are necessary and sufficient for activation (Ross et al., 2016). Since ppGpp concentrations vary, this binding synergy allows DksA and ppGpp to regulate transcription in accordance to the stress despite unchanging DksA concentrations (Rutherford et al., 2009).

TraR is a distant homolog of DksA that is encoded by the conjugative F plasmid and is expressed from the pY promoter as part of the major *tra* operon transcript (Frost et al., 1994; Maneewannakul and Ippen-Ihler, 1993). Although only half the length of DksA, TraR regulates transcription from *Escherichia coli* (*Eco*) RNAP σ^{70}-holoenzyme (termed Eσ^{70}) by binding to the RNAP secondary channel and mimicking the combined effects of DksA and ppGpp (Blankschien et al., 2009; Gopalkrishnan et al., 2017) (**Figure 5.1B**). TraR inhibits Eσ^{70}-dependent transcription from ribosomal RNA promoters (e.g. *rrnB* P1) and ribosomal protein promoters (e.g. *rpsT* P2, expressing S20), and activates amino acid biosynthesis and transport promoters (e.g. p*thrABC*, p*hisG*, p*argI*, p*livJ*) *in vivo* and *in vitro* (Blankschien et al., 2009; Gopalkrishnan et al., 2017). The affinity of TraR for RNAP is only slightly higher than that of DksA, yet its effect on promoters negatively regulated by ppGpp/DksA *in vitro* are as large or larger than those of ppGpp/DksA (Gopalkrishnan et al., 2017). The effect of TraR on promoters positively regulated by ppGpp/DksA are also independent of ppGpp (Gopalkrishnan et al., 2017).

Models for DksA/ppGpp and TraR binding to RNAP have been proposed based on biochemical and genetic approaches (Gopalkrishnan et al., 2017; Parshin et al., 2015; Ross et al., 2013, 2016). Crystal structures of ppGpp/DksA-RNAP (**Figure 5.1A**) and TraR-RNAP (**Figure 5.1B**) were obtained by soaking the factors into crystals of Eσ^{70} (Molodtsov et al., 2018; Murakami, 2013). These structures confirmed the general features of the models and provided

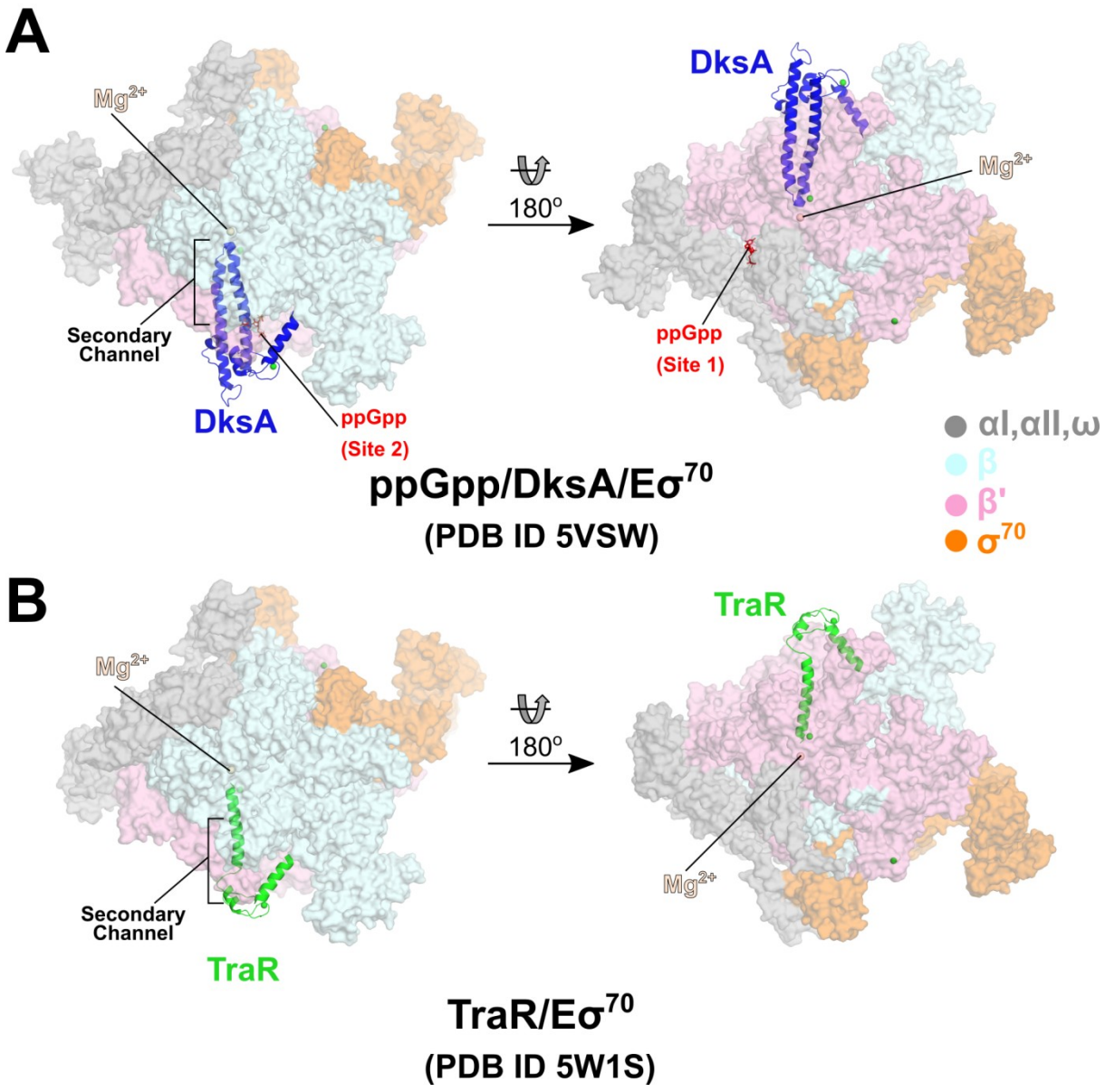


Figure 5.1. Crystal structures of ppGpp/DksA-Eσ^{70} and TraR-Eσ^{70}
(A) Two views of the crystal structure of ppGpp/DksA-Eσ^{70} (PDB ID 5VSW; space group $P2_1P2_1P2_1$; Molodtsov et al., 2018). Subunits of RNAP are shown as molecular surfaces and color coded according to key. The active site Mg^{2+} is colored wheat and is shown as a sphere. DksA is colored blue and is shown in cartoon format. ppGpp is colored red and shown in stick format. **(B)** Two views of the crystal structure of TraR-Eσ^{70} (PDB ID 5W1S; space group $P2_1P2_1P2_1$; Molodtsov et al., 2018). RNAP and active site Mg^{2+} are shown and colored similar to **(A)**. TraR is colored green and is shown in cartoon format. In both structures, the secondary channel is labeled.

additional details about their interactions with RNAP. However, due to crystal packing of mobile regions of the complex (**Chapter 1, Figure 1.6C**), the crystal structures did not reveal the mechanism of inhibition or activation (Molodtsov et al., 2018). Thus, the structural basis for the effects of ppGpp/DksA or TraR on transcription have remained elusive.

In **Chapter 4**, H. Boyaci (The Rockefeller University) and I utilized cryo-electron microscopy (cryo-EM) to study the different conformational states of the mobile clamp domain in *Mycobacterium tuberculosis* (*Mtb*) RNAP (Boyaci et al., 2018). In this chapter, I wanted to reexamine the TraR/RNAP structure in the absence of crystal packing by cryo-EM. Cryo-EM allows the visualization of a complex in native conditions. Multiple discrete conformational states populated in solution can be sorted using image classification. Furthermore, new software tools allow for the analysis of continuous conformational heterogeneity in the cryo-EM data (Nakane et al., 2018).

Therefore, I used cryo-EM to examine structures of $E\sigma^{70}$ alone, $E\sigma^{70}$ bound to TraR (TraR-$E\sigma^{70}$), and $E\sigma^{70}$ bound to a promoter inhibited by TraR [*rpsT* P2; Gopalkrishnan et al., 2017] in order to understand TraR regulation and principles of the regulation of transcription initiation in general. These structures show TraR binding in the secondary channel of RNAP, consistent with the previous crystal structures of the complex (Molodtsov et al., 2018). However, the cryo-EM structures also reveal that TraR binding induces major changes in RNAP conformations and dynamics that were not observed in the previous crystallographic analyses.

TraR affects a trigger-loop (TL) lineage-specific insert (LSI), the βlobe and an associated LSI, the bridge helix (BH), and the clamp. Changes in these mobile regions of RNAP affect promoter DNA interactions, critical for transcription activation or inhibition. A. Y. Chen and S. Gopalkrishnan from the Gourse Lab (University of Wisconsin-Madison) used mutational approaches to show that these structural changes, as well as effects on $\sigma^{70}_{1.1}$, are critical for transcription activation or inhibition by TraR. Based on the combined structural and functional analyses, we proposed a model in which TraR accelerates multiple steps along the open promoter complex (RPo) formation pathway and at the same time modulates the relative stability of intermediates in the pathway. Whether a promoter is activated or inhibited by TraR is determined by the intrinsic kinetic properties of the promoter (Galburt, 2018; Haugen et al., 2008; Paul et al., 2005). Portions of this chapter have been previously published in eLife (Chen et al., 2019b).

5.2 Cryo-EM structures of TraR-$E\sigma^{70}$, $E\sigma^{70}$, and *rpsT* P2 RPo

I used single-particle cryo-EM to examine the structure of the *Eco* TraR-$E\sigma^{70}$ complex in the absence of crystal packing interactions that could restrict conformational states of the complex. Additionally, I determined cryo-EM structures of $E\sigma^{70}$ alone and the $E\sigma^{70}$-*rpsT* P2 promoter RPo for comparison. By comparing the TraR-$E\sigma^{70}$, $E\sigma^{70}$, and *rpsT* P2-RPo cryo-EM structures, key structural differences in RNAP conformations revealed how TraR activates and inhibits transcription and generated testable mechanistic hypotheses for TraR function.

5.2.1 Cryo-EM structures of TraR-$E\sigma^{70}$

TraR function was examined by A. Y. Chen and S. Gopalkrishnan from the Gourse Lab (University of Wisconsin) in conditions that were previously used to solve cryo-EM structures of bacterial RNAPs (**Chapter 2**; Chen et al., 2019a). In the cryo-EM conditions, TraR function was indistinguishable from function under standard *in vitro* transcription assay conditions (**Figures 5.2A-C**; Chen et al., 2019b).

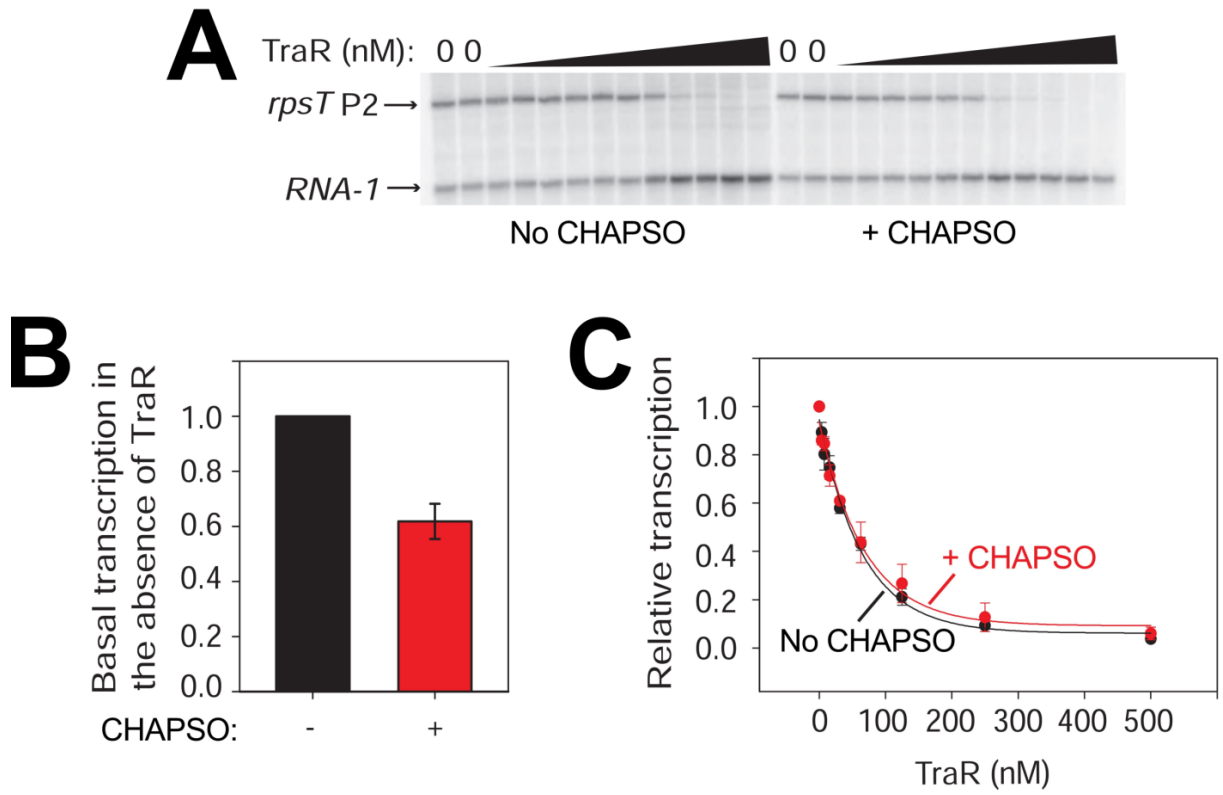


Figure 5.2. Cryo-EM solution conditions do not affect TraR function
(A) Multi-round *in vitro* transcription of *rpsT* P2 by $E\sigma^{70}$ (20 nM) at a range of TraR concentrations (wedge indicates 4 nM - 4 μM) in the absence or presence of 8 mM cholamidopropyl]dimethylammonio)-2-hydroxy-1-propanesulfonate (CHAPSO) as indicated. Plasmid templates also contained the *RNA-1* promoter. Transcription from the *RNA-1* promoter is slightly affected by TraR. **(B)** Transcription in the absence of TraR is plotted, relative to the same reactions without CHAPSO. CHAPSO reduced transcription slightly. Averages with range from two independent experiments are plotted. **(C)** Quantification of transcripts from experiments like those in (**A**) plotted relative to values in the absence of TraR. The IC_{50} for inhibition by TraR was ~50 nM for both ± CHAPSO data sets, demonstrating CHAPSO has little to no effect on the concentration of TraR required for half-maximal inhibition. Averages with range from two independent experiments are plotted. Transcription assay data and figure courtesy of A. Y. Chen and S. Gopalkrishnan from the Gourse Lab (University of Wisconsin–Madison).

Analysis of the cryo-EM data for the TraR-Eσ^{70} complex gave rise to three distinct conformational classes (**Figure 5.3**). The structures of each class were nearly identical except for the disposition of β'Si3 [also called β'i6; (Lane and Darst, 2010a)], a 188-residue LSI in the TL of *Eco* RNAP (Chlenov et al., 2005) (**Figure 5.3**). The first class [TraR-Eσ^{70}(I)] contained 153,295 particles (~41% of the particles) and resolved to a nominal resolution of 3.7 Å (**Figures 5.3**, **5.4A-B**). The second class [TraR-Eσ^{70}(II)] contained 123,607 particles (~33% of the particles) and resolved to a nominal resolution of 3.8 Å (**Figures 5.3**, **5.4C-D**). The third class [TraR-Eσ^{70}(III)] contained 95,767 particles (~26% of the particles) and resolved to a nominal resolution of 3.9 Å but lacked density for Si3 (**Figure 5.3**). Atomic models were built for TraR-Eσ^{70}(I) and TraR-Eσ^{70}(II), both of which had discrete cryo-EM density for Si3. (**Table 5.5**, **Figure 5.6A-B**). With Si3 (β' residues 948–1126) excluded, the structures superimpose with a root-mean-square deviation (RMSD) of 0.495 Å over 3,654 alpha carbon (Cα) atoms.

The binding mode of TraR in the cryo-EM structures (**Figures 5.6A-D**) is consistent with the effects of TraR or RNAP substitutions on TraR function (Gopalkrishnan et al., 2017). TraR can be divided into three structural elements: an N-terminal helix ($TraR_N$, residues 2-27), a globular domain ($TraR_G$, residues 28-57) and a C-terminal helix ($TraR_C$, residues 58-73) (**Figure 5.6D**). Consistent with the X-ray structure (Molodtsov et al., 2018), $TraR_N$ extends from the RNAP active site out through the RNAP secondary channel to the β'rim-helices, located at the entrance to the RNAP secondary channel. The N-terminal tip of $TraR_N$ (TraR residue S2) is only 4.3 Å from the active site Mg^{2+}. $TraR_N$ makes extensive interaction with structural elements surrounding the active site, including the -NADFDGD- motif that chelates the active site Mg^{2+} (Zhang et al., 1999), the F-loop (Miropolskaya et al., 2009), and the bridge-helix (**Figure 5.6D**).

$TraR_G$ interacts primarily with the β'rim-helices, previously observed in the crystal structures (**Figure 5.6D**). However, the interactions of $TraR_C$ with RNAP in the cryo-EM structures differ substantially with the X-ray crystal structures (Molodtsov et al., 2018) due to conformational changes in RNAP induced by TraR binding in the native cryo-EM conditions (**Figure 5.6D**). Alignment of the cryo-EM and X-ray structures report an RMSD of 4.26 Å over 3,471 α-carbons, indicating significant conformational differences between the structures.

5.2.2 Cryo-EM structures for structural comparisons: Eσ^{70} and *rpsT* P2 RPo

Cryo-EM data for Eσ^{70} alone were collected for comparison with TraR-Eσ^{70}. Previously published cryo-EM structures of Eσ^{70} were crosslinked to reduce conformational heterogeneity so not suited for comparative analysis (Narayanan et al., 2018). After image classification (**Figure 5.7**), the final reconstruction of Eσ^{70} contained 358,725 particles and resolved to a nominal resolution of 4.1 Å (**Figure 5.8A**) with local resolution estimates between 2.9-3.4 Å for the central part of RNAP (**Figure 5.8B**). The overall structure of Eσ^{70} is consistent with a previously determined X-ray crystal structure of Eσ^{70} (Bae et al., 2013) (**Table 5.9**).

Cryo-EM data were collected on a complex of Eσ^{70} with an *rpsT* P2 promoter fragment (in the absence of TraR) in an RPo (*rpsT* P2-RPo). The *rpsT* P2-RPo was generated *de novo* by incubating Eσ^{70} with a linear duplex fragment of *rpsT* P2 (**Figure 5.10**). Analysis of the *rpsT* P2-RPo cryo-EM data (**Figure 5.11**) gave rise to two conformational classes that differed only in the disposition of the upstream promoter DNA and C-terminal domains of the α subunits (αCTDs) (**Figure 5.11**). I focused on the highest resolution class, resolved to a nominal resolution of 3.4 Å (**Figures 5.12A-B**). An atomic model was built and refined for this map (**Table 5.13**, **Figure 5.14**).

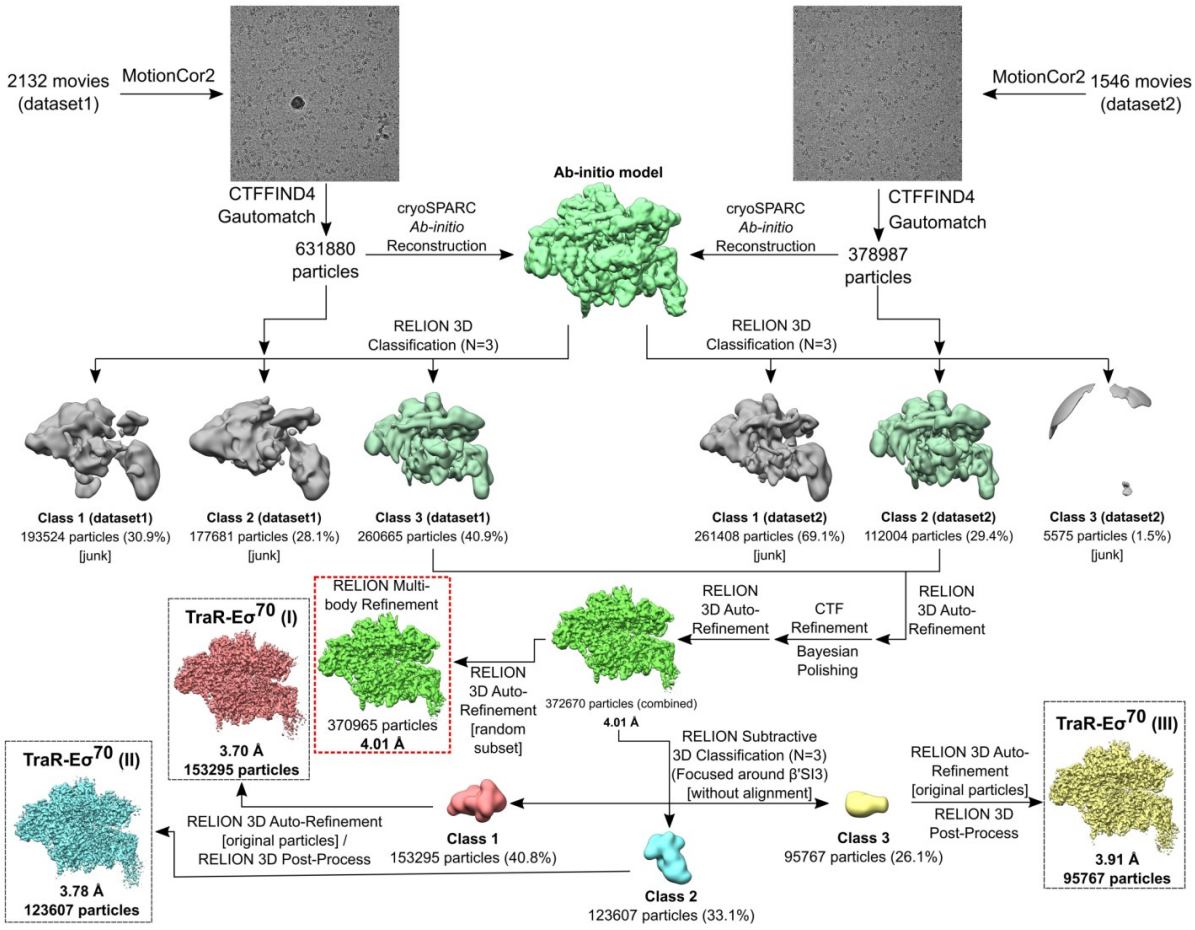


Figure 5.3. TraR-Eσ^{70} cryo-EM processing pipeline

Flowchart showing the image processing pipeline for the cryo-EM data of TraR-Eσ^{70}. Multi-body refinement (Nakane et al., 2018) was performed on particles highlighted in the red dashed box. After image classification, three structures were observed that are highlighted in black dashed boxes: TraR-Eσ^{70}(I), TraR-Eσ^{70}(II), and TraR-Eσ^{70}(III). After auto-refinement and post-processing in RELION (Zivanov et al., 2018), TraR-Eσ^{70}(I) (red) contains 153,295 particles at 3.70 Å resolution, TraR-Eσ^{70}(II) (cyan) contains 123,607 particles at 3.78 Å resolution, TraR-Eσ^{70}(III) (yellow) contains 95,767 particles at 3.91 Å resolution. Atomic models were generated for TraR-Eσ^{70}(I) and TraR-Eσ^{70}(II) for analysis.

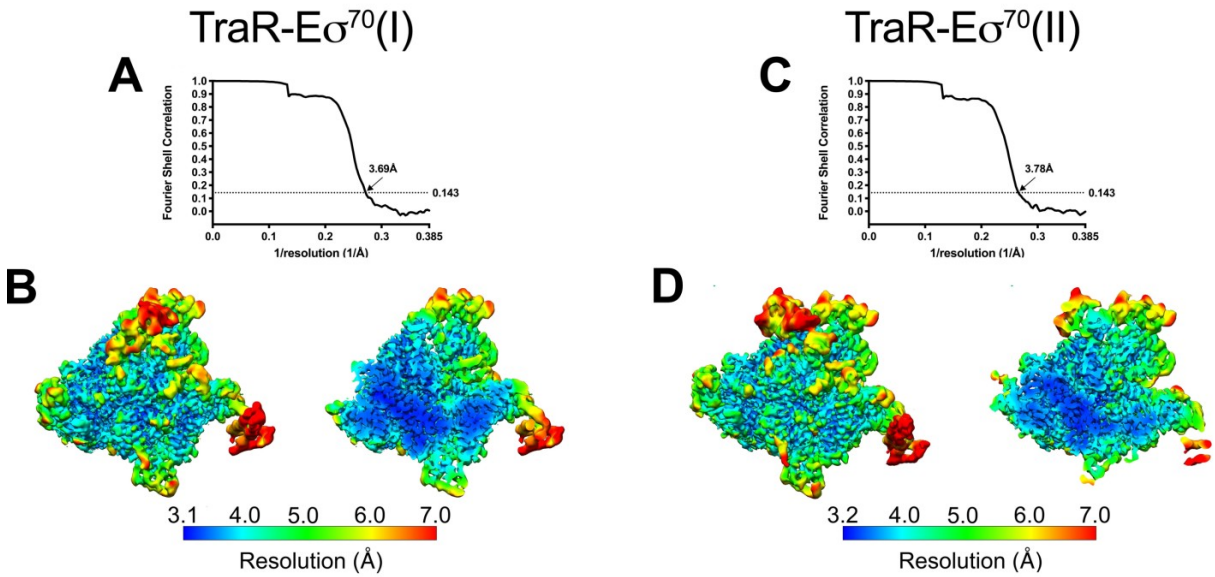


Figure 5.4. Overall FSC and local resolution calculations for TraR-Eσ^{70} cryo-EM maps
(*left)* TraR-Eσ^{70}(I) **(A)** Gold-standard FSC of TraR-Eσ^{70}(I). The gold-standard FSC was calculated by comparing the two independently determined half-maps from RELION (Zivanov et al., 2018). The dotted line represents the 0.143 FSC cutoff, which indicates a nominal resolution of 3.7 Å. **(B)** The 3.7 Å resolution cryo-EM density map of TraR-Eσ^{70}(I) is colored by local resolution (Cardone et al., 2013). The right view is a cross-section of the left view. (*right)* TraR-Eσ^{70}(II). **(C)** Gold-standard FSC of TraR-Eσ^{70}(II). The gold-standard FSC was calculated by comparing the two independently determined half-maps from RELION (Zivanov et al., 2018). The dotted line represents the 0.143 FSC cutoff, which indicates a nominal resolution of 3.8 Å. **(D)** The 3.8 Å resolution cryo-EM density map of TraR-Eσ^{70}(II) is colored by local resolution (Cardone et al., 2013). The right view is a cross-section of the left view.

Table 5.5. Cryo-EM data collection and refinement parameters for TraR-Eσ^{70}

Sample	TraR-Eσ^{70}(I)	TraR-Eσ^{70}(II)	TraR-Eσ^{70}(III)
EMDB	EMD-0348	EMD-0349	EMD-20231
PDB	6N57	6N58	-
Data collection and processing			
Microscope	FEI Titan Krios		
Voltage (kV)	300		
Detector	K2 summit		
Electron exposure (e^{-}/Å^2)	71		
Defocus range (μm)	0.8-2.4		
Data collection mode	Counting		
Nominal Magnification	22,500x		
Pixel size (Å)	1.3		
Symmetry imposed	C1		
Initial particle images (no.)	101,0867		
Final particle images (no.)	153,295	123,607	95,767
Map resolution (Å) - FSC threshold 0.143	3.7	3.78	3.91
Map resolution range (Å)	3.1-7.0	3.2-7.0	3.5-8.5
Refinement[c]			
Initial model used (PDB code)	4LJZ (EΔ1.1σ^{70})[a] 4LK1 ($\sigma^{70}_{1.1}$)[a] 5W1S (TraR)[b]		
Map sharpening B factor (Å^2)	-95	-94	-79
Model composition			
Non-hydrogen atoms	30,124	30,188	
Protein residues	3,841	3,833	
Nucleic acid residues	0	0	
Ligands	1 Mg^{2+} 3 Zn^{2+} 4 CHAPSO	1 Mg^{2+} 3 Zn^{2+} 4 CHAPSO	
B factors (Å^2)			
Protein	47.80	39.95	
Nucleic acid	-	-	
Ligands	51.05	42.64	
R.m.s. deviations			
Bond lengths (Å)	0.011	0.012	
Bond angles (°)	1.062	1.147	
Validation			
MolProbity score	2.19	2.14	

Clashscore	9.67	8.96	
Poor rotamers (%)	0.25	0.37	
Ramachandran plot			
Favored (%)	83.86	84.92	
Disallowed (%)	0.29	0.37	

[a](Bae et al., 2013)

[b](Molodtsov et al., 2018)

[c]Refinement: PHENIX real_space_refine (Adams et al., 2010). Validation: MOLPROBITY (Chen et al., 2010b).

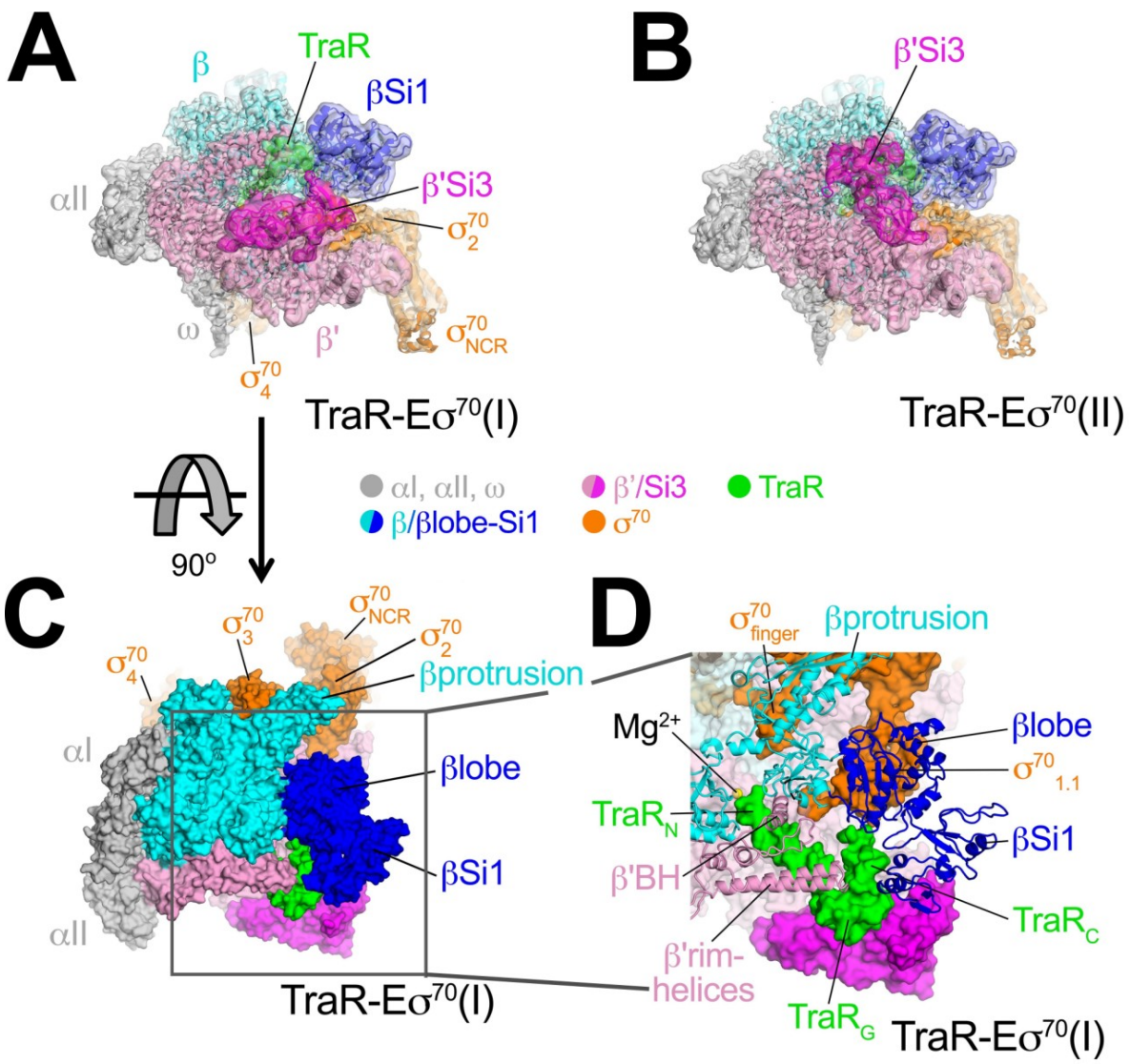


Figure 5.6. Cryo-EM structure of TraR-Eσ^{70}(I) and TraR-Eσ^{70}(II)
(A) Structure of TraR-Eσ^{70}(I). Cryo-EM density map (3.7 Å nominal resolution, low-pass filtered to the local resolution) is shown as a transparent surface and colored according to the key. The final model is superimposed. (**B**) Structure of TraR-Eσ^{70}(II). Cryo-EM density map (3.8 Å nominal resolution, low-pass filtered to the local resolution) is shown as a transparent surface and colored according to the key. The final model is superimposed. (**C**) Top view of TraR-Eσ^{70}(I), orthogonal to **(A)**. The area boxed in grey is magnified in (**D**). (**D**) Magnified top view of TraR-Eσ^{70}(I) showing architecture of TraR: TraR$_N$ (starting near RNAP active site Mg^{2+}, extending out secondary channel), TraR$_G$ (interacting primarily with β'rim-helices), and TraR$_C$ (interacting with βlobe-Si1).

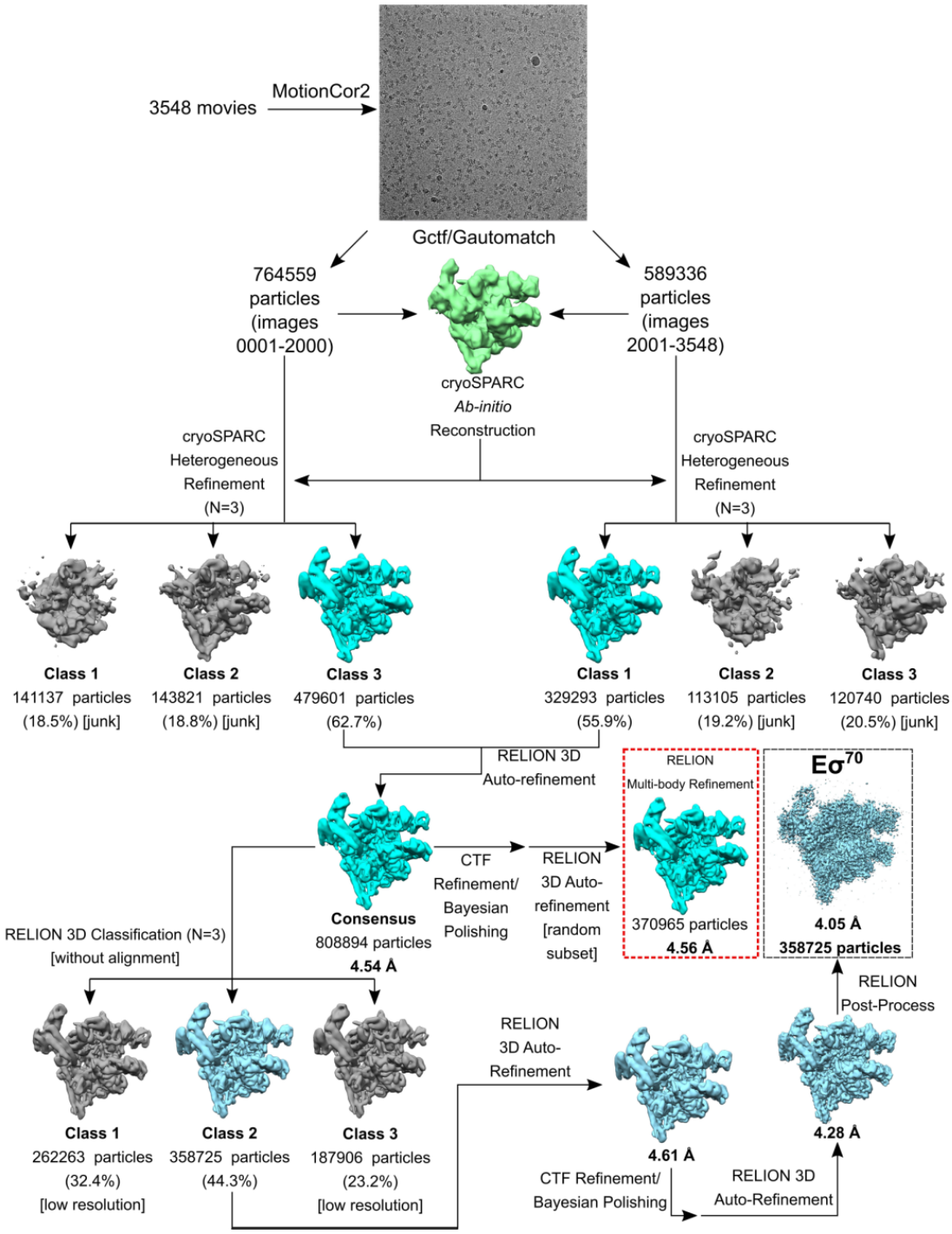


Figure 5.7. Eσ^{70} cryo-EM processing pipeline
Flowchart showing the image processing pipeline for the cryo-EM data of TraR-Eσ^{70}. Multi-body refinement (Nakane et al., 2018) was performed on particles highlighted in the red dashed box. After image classification, the final structure (highlighted in black dashed boxes) contains 358,725 particles with a nominal resolution of 4.05 Å after auto-refinement and post-processing in RELION (Zivanov et al., 2018).

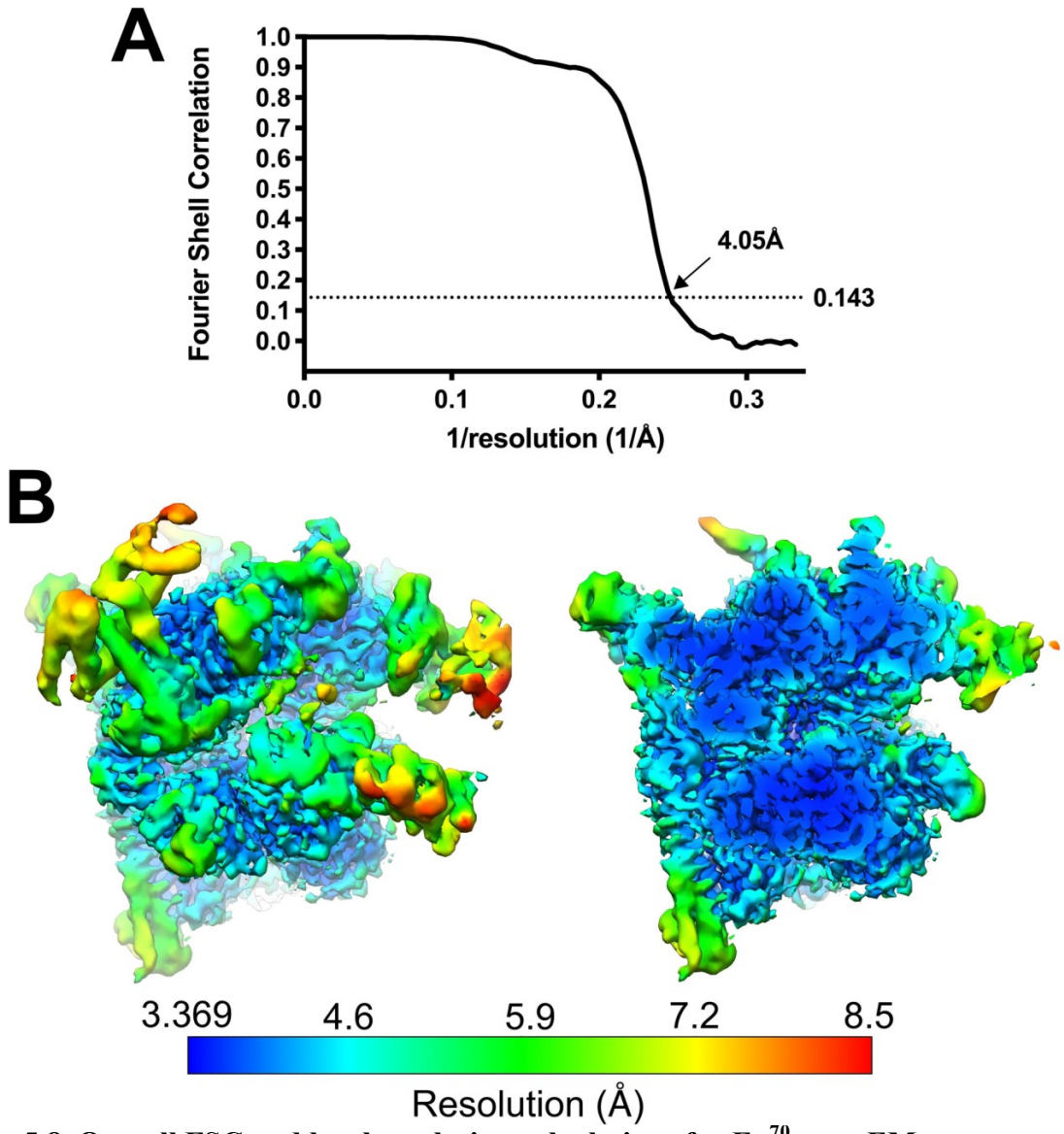


Figure 5.8. Overall FSC and local resolution calculations for Eσ^{70} cryo-EM map
(A) Gold-standard FSC of Eσ^{70}. The gold-standard FSC was calculated by comparing the two independently determined half-maps from RELION (Zivanov et al., 2018). The dotted line represents the 0.143 FSC cutoff, which indicates a nominal resolution of 4.1 Å. **(B)** The 4.1 Å resolution cryo-EM density map of Eσ^{70} is colored by local resolution (Cardone et al., 2013). The right view is a cross-section of the left view.

Table 5.9. Cryo-EM data collection and refinement parameters for $E\sigma^{70}$

Sample	**$E\sigma^{70}$**
EMDB	EMD-20230
PDB	6P1K
Data collection and processing	
Microscope	Talos Arctica
Voltage (kV)	200
Detector	K2 summit
Electron exposure (e–/Å^2)	37.3
Defocus range (μm)	1.0-2.5
Data collection mode	Counting
Nominal Magnification	28,000x
Pixel size (Å)	1.5
Symmetry imposed	C1
Initial particle images (no.)	1387166
Final particle images (no.)	358,725
Map resolution (Å) - FSC threshold 0.143	4.05
Map resolution range (Å)	3.4-8.5
Refinement[c]	
Initial model used (PDB code)	4LJZ ($E\Delta1.1\sigma^{70}$)[a] 4LK1 ($\sigma^{70}_{1.1}$)[a]
Map sharpening B factor (Å^2)	-194
Model composition	
Non-hydrogen atoms	29,160
Protein residues	3,714
Nucleic acid residues	0
Ligands	1 Mg^{2+} 2 Zn^{2+}
B factors (Å^2)	
Protein	37.68
Nucleic acid	-
Ligands	49.76
R.m.s. deviations	
Bond lengths (Å)	0.008
Bond angles (°)	1.131
Validation	
MolProbity score	1.75
Clashscore	3.60
Poor rotamers (%)	0.66
Ramachandran plot	

Favored (%)	88.16
Disallowed (%)	0.16

[a](Bae et al., 2013)
[b](Molodtsov et al., 2018)
[c]Refinement: PHENIX real_space_refine (Adams et al., 2010). Validation: MOLPROBITY (Chen et al., 2010b).

Figure 5.10. *Eco rpsT* P2 promoter fragment

Linear fragment of *Eco rpsT* P2 promoter (expressing S20 ribosomal protein) used for cryo-EM studies. The nontemplate strand (nt-strand) is colored light grey, t-strand is colored dark grey, -35 element is colored yellow, and -10 is colored magenta. Numbering above the sequence indicates position with respect to the +1 transcription state site (TSS) (shown as arrow pointing downstream).

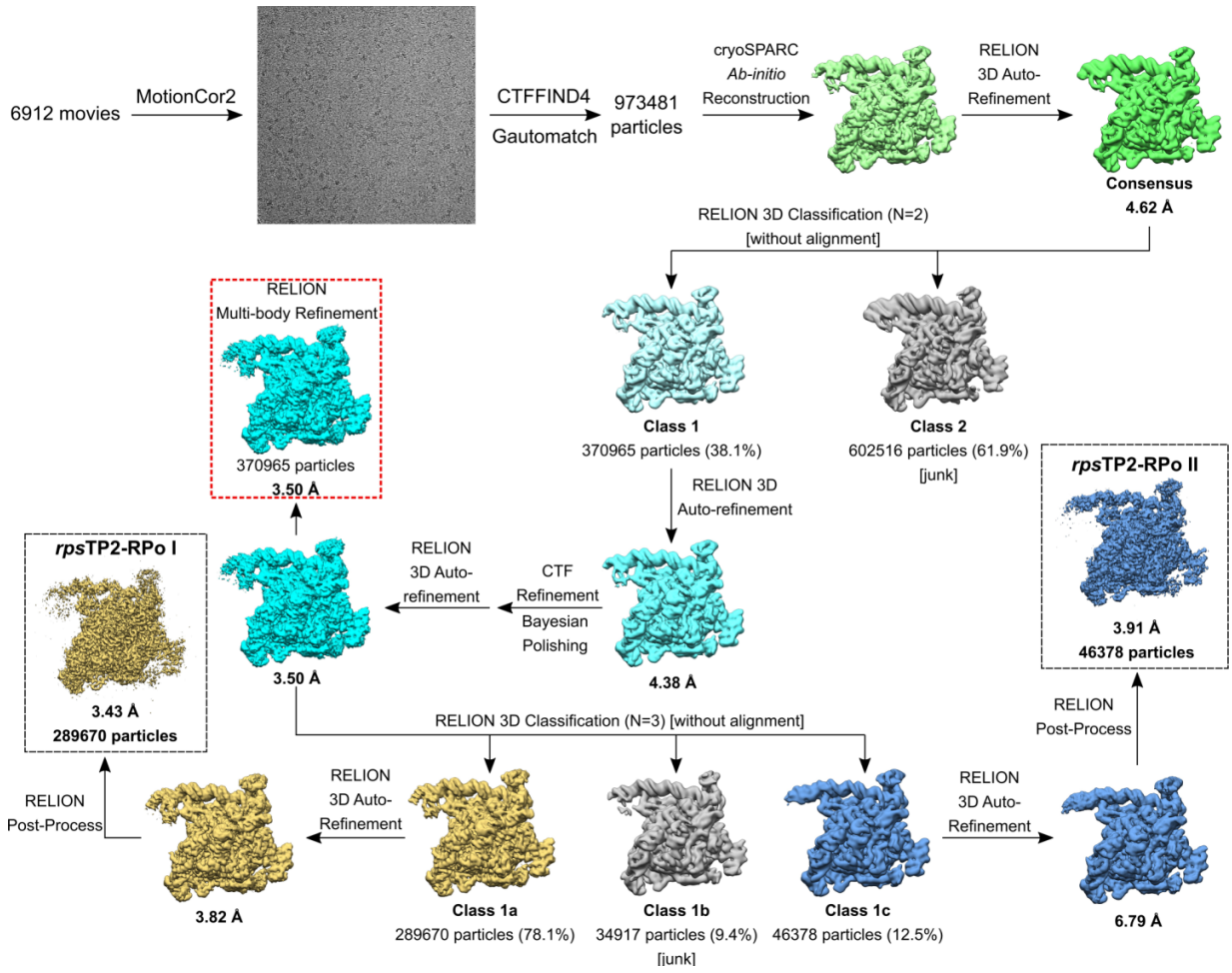


Figure 5.11. *rpsT* P2 RPo cryo-EM processing pipeline
Flowchart showing the image processing pipeline for the cryo-EM data of ***rpsT* P2-RPo**. Multi-body refinement (Nakane et al., 2018) was performed on particles highlighted in the red dashed box. After image classification, two structures were observed that are highlighted in black dashed boxes: *rpsT* P2-RPo (I) and *rpsT* P2-RPo (II). After auto-refinement and post-processing in RELION (Zivanov et al., 2018), *rpsT* P2-RPo (I) (yellow) contains 289,670 particles with a nominal resolution of 3.43 Å and *rpsT* P2-RPo (II) (blue) contains 46,378 particles with a nominal resolution of 3.91 Å. An atomic model was built for *rpsT* P2-RPo (I).

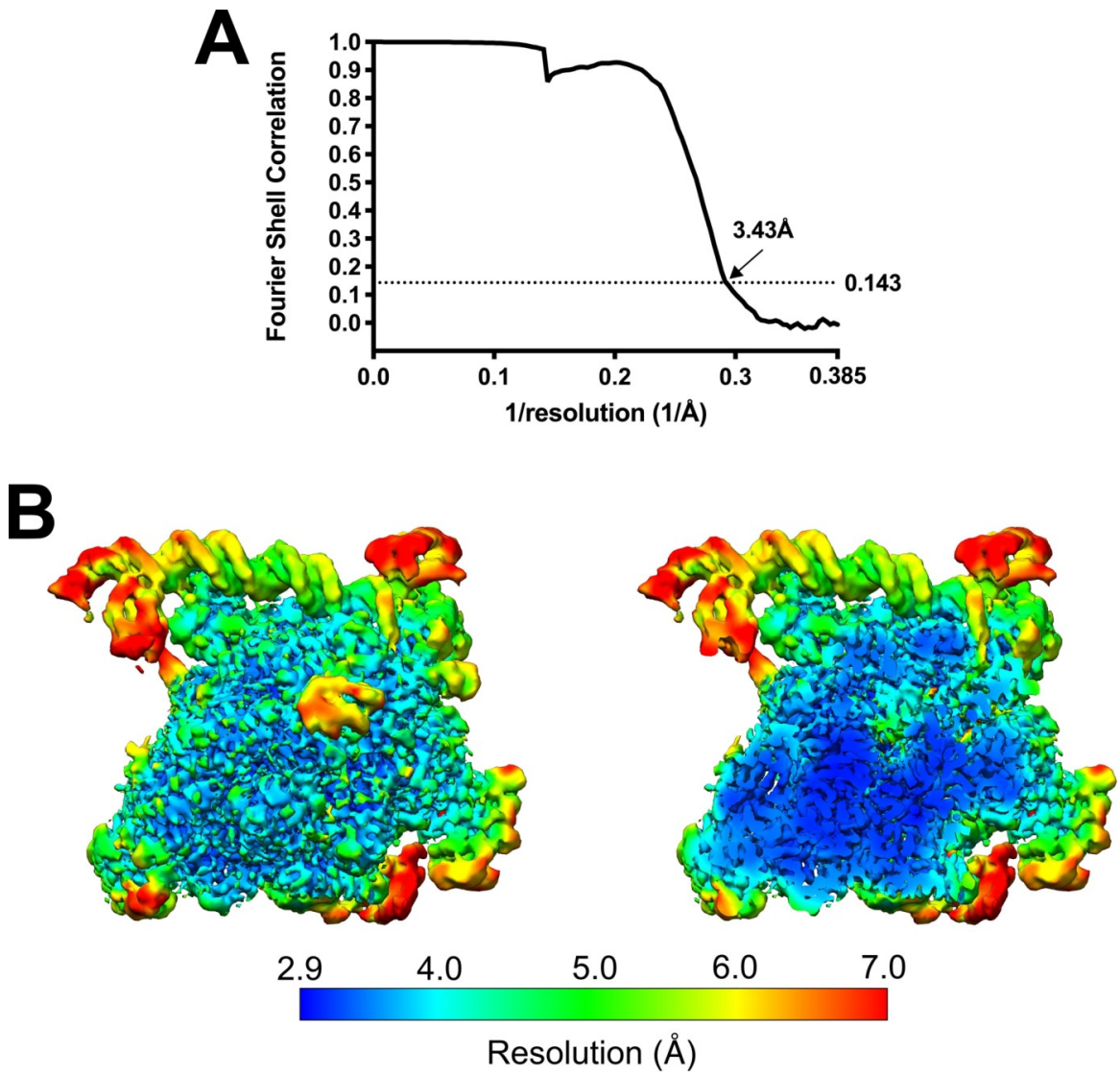


Figure 5.12. Overall FSC and local resolution calculations for *rpsT* P2-RPo (I) cryo-EM map

(A) Gold-standard FSC of *rpsT* P2-RPo (I). The gold-standard FSC was calculated by comparing the two independently determined half-maps from RELION (Zivanov et al., 2018). The dotted line represents the 0.143 FSC cutoff, which indicates a nominal resolution of 3.4 Å. **(B)** The 3.4 Å resolution cryo-EM density map of *rpsT* P2-RPo (I) is colored by local resolution (Cardone et al., 2013). The right view is a cross-section of the left view.

Table 5.13. Cryo-EM data collection and refinement parameters for *rpsT* P2-RPo

Sample	***rpsT* P2-RPo (I)**	***rpsT* P2-RPo (II)**
EMDB	EMD-20203	EMD-20232
PDB	6OUL	-
Data collection and processing		
Microscope	FEI Titan Krios	
Voltage (kV)	300	
Detector	K2 summit	
Electron exposure (e–/Å^2)	47.3	
Defocus range (μm)	0.5-2.5	
Data collection mode	Super-resolution	
Nominal Magnification	22,500x	
Pixel size (Å)	1.3	
Symmetry imposed	C1	
Initial particle images (no.)	973,481	
Final particle images (no.)	289,670	46,378
Map resolution (Å) - FSC threshold 0.143	3.43	3.91
Map resolution range (Å)	3.3-8.0	2.9-8.8
Refinement[c]		
Initial model used (PDB code)	4LJZ ($E\Delta1.1\sigma^{70}$)[a]	
Map sharpening B factor (Å^2)	-95	-66
Model composition		
Non-hydrogen atoms	31,926	
Protein residues	3,751	
Nucleic acid residues	117	
Ligands	1 Mg^{2+} 2 Zn^{2+} 3 CHAPSO	
B factors (Å^2)		
Protein	36.90	
Nucleic acid	106.08	
Ligands	34.18	
R.m.s. deviations		
Bond lengths (Å)	0.014	
Bond angles (°)	1.130	
Validation		
MolProbity score	2.07	
Clashscore	8.43	

Poor rotamers (%)	0.31	
Ramachandran plot		
Favored (%)	87.50	
Disallowed (%)	0.27	

[a](Bae et al., 2013)
[b](Molodtsov et al., 2018)
[c]Refinement: PHENIX real_space_refine (Adams et al., 2010). Validation: MOLPROBITY (Chen et al., 2010b).

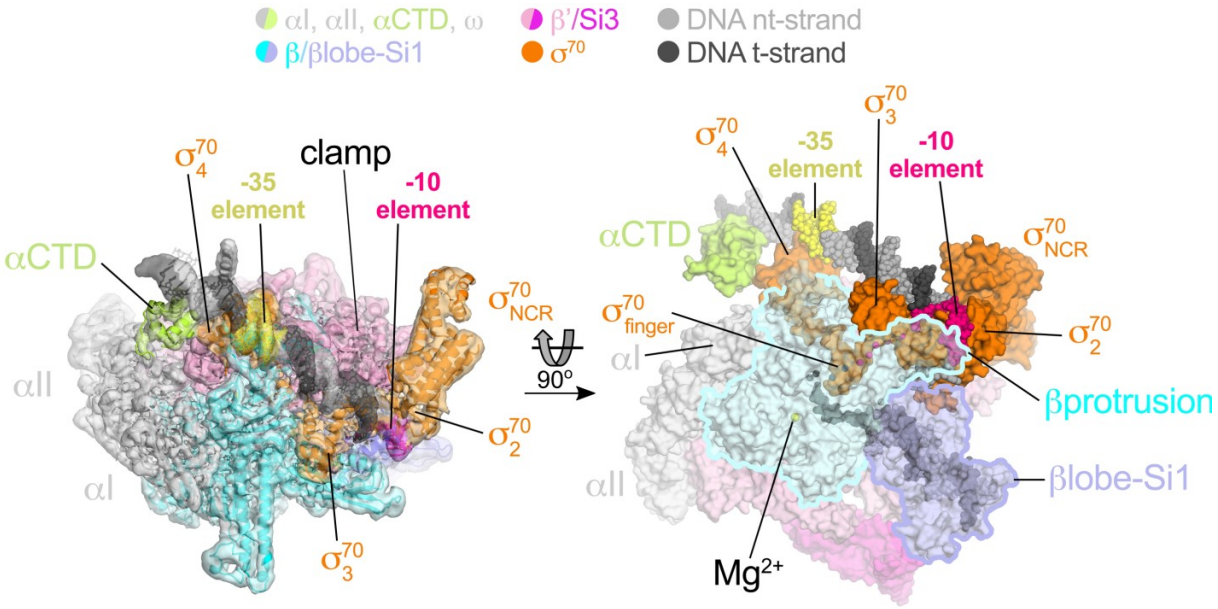


Figure 5.14. Cryo-EM structure of *rpsT* P2-RPo (I)
(*top*) Color-code key. (*left*) *rpsT* P2-RPo (I) cryo-EM density map (3.4 Å nominal resolution, low-pass filtered to the local resolution) is shown as a transparent surface and colored according to the key. The final model is superimposed. The DNA was modeled from -45 to +21. The t-strand DNA from -10 to -2, and the nt-strand DNA from -3 to +2 were not modeled due to poor cryo-EM density for these nucleotides. (*right*) Top view of *rpsT* P2-RPo (I). DNA is shown as spheres. Proteins are shown as molecular surfaces. Much of the β subunit is transparent to reveal the active site Mg^{2+} (yellow sphere), σ^{70}_{finger}, and DNA inside the RNAP active site cleft.

The closed-clamp RNAP in the *rpsT* P2-RPo cryo-EM structure (**Figure 5.14**) interacts with the melted DNA promoter in the same manner observed in RPo structures determined by X-ray crystallography (Bae et al., 2015b, 2015a; Hubin et al., 2017a) or cryo-EM (Boyaci et al., 2019a) and is consistent with the DNase I footprint of the *rpsT* P2-RPo (Gopalkrishnan et al., 2017). In the *rpsT* P2-RPo structure, an αCTD (Ross et al., 1993) is bound to the promoter DNA minor groove (Benoff et al., 2002; Ross et al., 2001) just upstream of the promoter -35 element [-38 to -43, corresponding to the proximal upstream (UP) element subsite; Estrem et al., 1999]. This αCTD interacts with $\sigma^{70}{}_4$ through an interface previously characterized through genetic analyses (Ross et al., 2003) (**Figure 5.14**). The αCTDs are linked to the α-N-terminal domains (αNTDs) by ~15 residue flexible linkers (Blatter et al., 1994; Jeon et al., 1995), however density for these linkers was not observed in the cryo-EM map.

5.3 β'Si3 adopts two conformations in TraR-Eσ^{70}

The three TraR-Eσ^{70} structures differ from each other only in the disposition of β'Si3. Si3 comprises two tandem repeats of the sandwich-barrel hybrid motif (SBHM) fold (Chlenov et al., 2005; Iyer et al., 2003), SBHMa and SBHMb, that are linked to the TL-helices (**Figure 5.15A**). In TraR-Eσ^{70}(I) and TraR-Eσ^{70}(II), Si3 is in two distinct positions with respect to the RNAP (**Figures 5.6A-B**, **5.15A**), while in TraR-Eσ^{70}(III) Si3 is disordered (**Figure 5.3**). Si3 in the TraR-Eσ^{70}(I) structure [Si3(I)] interacts primarily with the β'shelf (via SBHMa) and the β'jaw (via SBHMb) in a manner seen in many previous X-ray (Bae et al., 2013) and cryo-EM (Chen et al., 2017; Kang et al., 2017; Liu et al., 2017) structures of *Eco* RNAP. Si3 in the TraR-Eσ^{70}(II) structure [Si3(II)] is rotated 121° with respect to [Si3(I)] such that SBHMa interacts with the β'jaw and SBHMb interacts with $TraR_G$ (**Figures 5.15A-C**), a disposition of Si3 that has not been observed previously.

A. Y. Chen and S. Gopalkrishnan from the Gourse Lab (University of Wisconsin) performed *in vitro* transcription assays to test the relevance of the [Si3(II)] conformation for TraR function. We examined TraR function with wild-type (WT) and β'Si3 deleted (Δβ'Si3) RNAPs at promoters where TraR inhibits or activates transcription (Chen et al., 2019b). Deletion of Si3 had only minor effects on TraR-mediated inhibition of *rrnB* P1 and *rpsT* P2 (**Figures 5.16A-B**) but transcription by ΔSi3-RNAP was activated only ~50% compared with WT-RNAP on three different TraR-activated promoters (p*thrABC*, **Figure 5.16C**; p*argI*, **Figure 5.16D**; p*hisG*, **Figure 5.16E**).

Three $TraR_G$ residues (TraR residues E46, R49, and K50) make up the main Si3-$TraR_G$ interface (**Figures 5.15B-C**). Alanine mutations of these residues (TraR variants: E46A, R49A, or K50A) gave rise to similar results as deleting Si3. Inhibition of *rrnB* P1 was similar to WT-TraR for TraR-K50A, and mildly impaired for TraR-E46A or R49A (Chen et al., 2019b) (**Figure 5.17A**, see legend for IC_{50} values). However, these same variants exhibited at least ~2 fold reduced activation at the *thrABC* promoter, similar to deletion of Si3 (**Figure 5.17B).**

The combination of the TraR-Si3 interface mutants and ΔSi3-RNAP was epistatic; the same ~2 fold reduction in activation was observed as with the Si3-TraR interface mutants or the ΔSi3-RNAP individually (Chen et al., 2019b) (**Figure 5.17C**). These results indicate that the Si3(SBHMb)-$TraR_G$ interaction enabled by the Si3(II) conformation accounts for part of the TraR-mediated effect on activation.

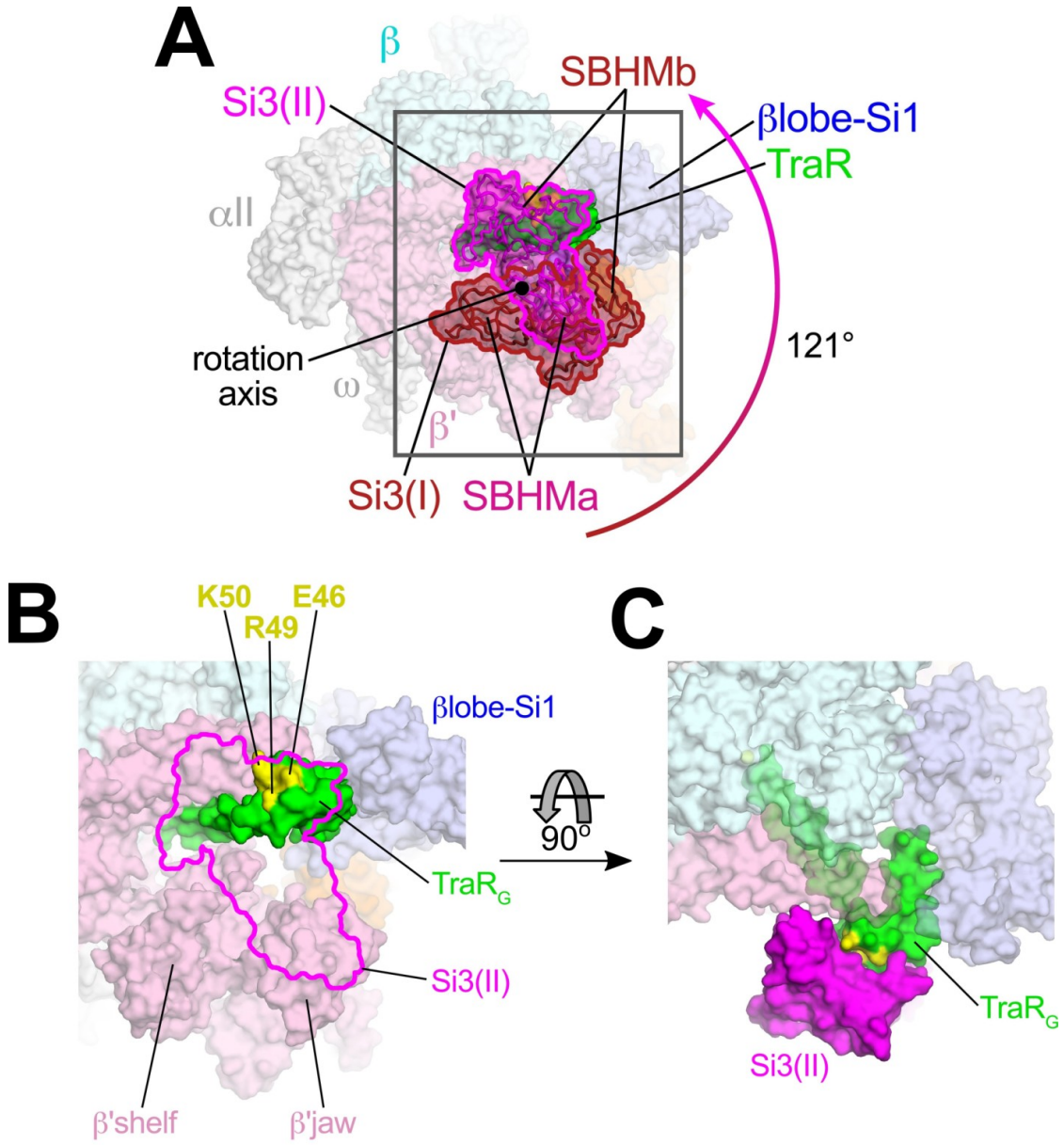
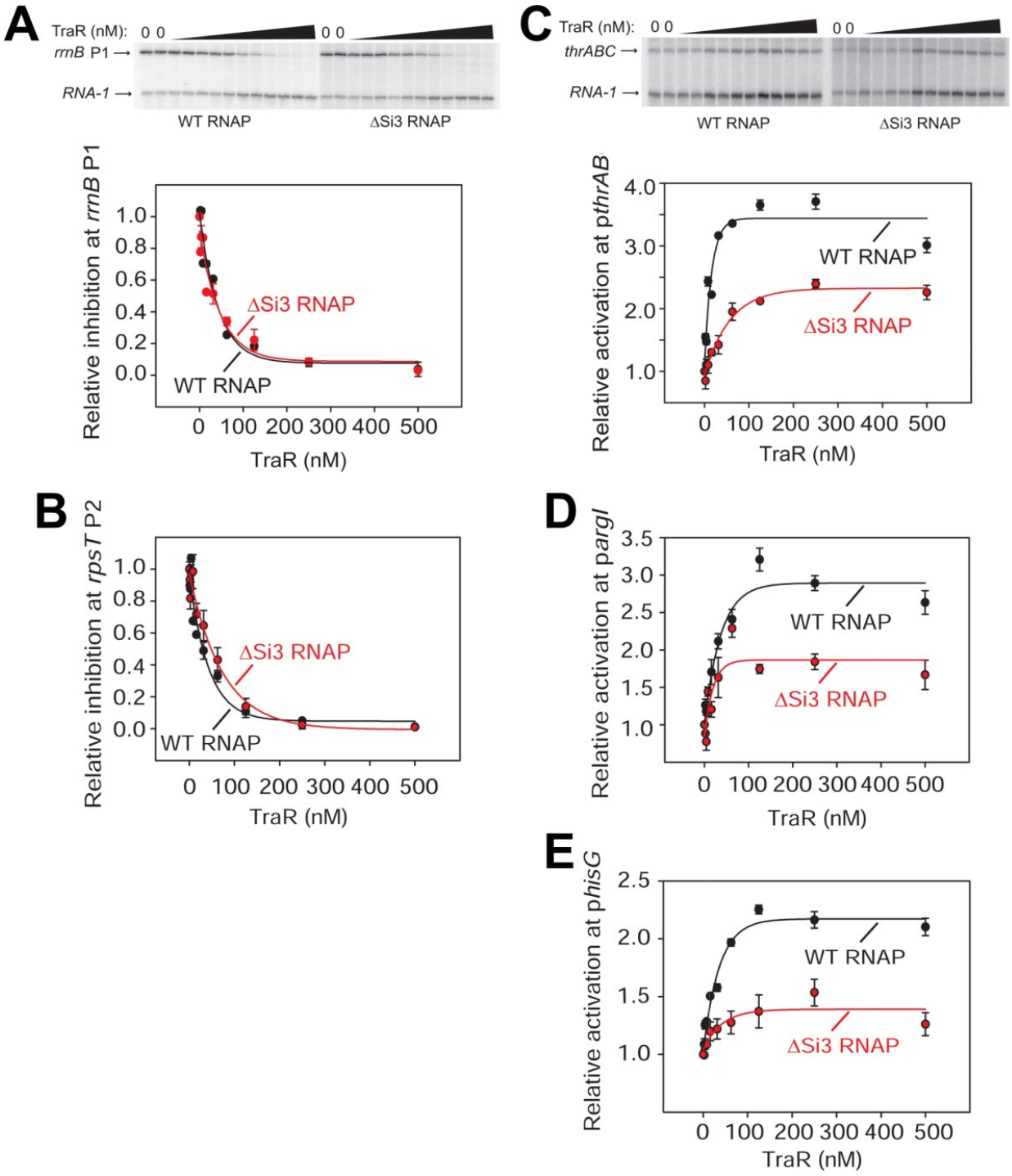


Figure 5.15. Conformational flexibility of β'Si3 in TraR-Eσ^{70}
(A) Overall view of TraR-Eσ^{70} structure with alternative positions of Si3 seen in TraR-Eσ^{70}(I) and TraR-Eσ^{70}(II). RNAP is shown as transparent molecular surface. Si3(I) is shown as cartoon tubes and outlined in brown. A ~ 121° rotation about the rotation axis shown gives rise to the position of Si3(II) shown as cartoon tubes and outlined in magenta. Si3 comprises two SBHM domains (Chlenov et al., 2005; Iyer et al., 2003), denoted SBHMa and SBHMb. **(B)** Magnified view of TraR-Eσ^{70}(II) [boxed area in **(A)**]. The position of Si3(II) is outlined in magenta but the rest of Si3 is removed, revealing TraR (green molecular surface) behind. Three residues central to the TraR-Si3(II) interface (TraR-E46, R49, and K50) are colored yellow. **(C)** Orthogonal view as **(B)**, showing the extensive TraR-Si3(II) interface.

Figure 5.16. Deletion of β'Si3 affects activation but not inhibition by TraR
(A) (*top*) Multi round *in vitro* transcription of *rrnB* P1 over a range of TraR concentrations (wedge indicates 2 nM - 2 μM) in the presence of WT-RNAP or ΔSi3-RNAP as indicated. Plasmid templates also contained the RNA-1 promoter. (*bottom*) Quantification of transcripts plotted relative to values in the absence of TraR. The IC_{50} for inhibition of *rrnB* P1 by TraR was ~40 nM for both data sets. **(B)** Plot of multi round *in vitro* transcription of *rpsT* P2 over a range of TraR concentrations in the presence of WT-RNAP or ΔSi3-RNAP as indicated. The IC_{50} for inhibition by TraR was ~60 nM with WT-RNAP and ~90 nM with ΔSi3-RNAP. **(C)** (*top*) Multi round *in vitro* transcription of *thrABC* over a range of TraR concentrations (wedge indicates 2 nM - 2 μM) in the presence of 20 nM WT-RNAP or ΔSi3-RNAP as indicated. Plasmid templates also contained the RNA-1 promoter. (*bottom*) Quantification of transcripts from *in vitro* transcription assay (*top*) plotted relative to values in the absence of TraR. **(D)** p*argI*. **(E)** p*hisG*. Quantifications show averages with range from two independent experiments. Transcription assay data and figure courtesy of A. Y. Chen and S. Gopalkrishnan from the Gourse Lab (University of Wisconsin–Madison).


A
TraR (nM): 0 0
0 0
rrnB P1
RNA-1
WT RNAP
ΔSi3 RNAP
Relative inhibition at rrnB P1
1.0
0.8
0.6
0.4
0.2
0.0
ΔSi3 RNAP
WT RNAP
0 100 200 300 400 500
TraR (nM)
B
Relative inhibition at rpsT P2
1.0
0.8
0.6
0.4
0.2
0.0
ΔSi3 RNAP
WT RNAP
0 100 200 300 400 500
TraR (nM)
C
TraR (nM): 0 0
0 0
thrABC
RNA-1
WT RNAP
ΔSi3 RNAP
Relative activation at pthrAB
4.0
3.0
2.0
1.0
WT RNAP
ΔSi3 RNAP
0 100 200 300 400 500
TraR (nM)
D
Relative activation at pargI
3.5
3.0
2.5
2.0
1.5
1.0
WT RNAP
ΔSi3 RNAP
0 100 200 300 400 500
TraR (nM)
E
Relative activation at phisG
2.5
2.0
1.5
1.0
WT RNAP
ΔSi3 RNAP
0 100 200 300 400 500
TraR (nM)

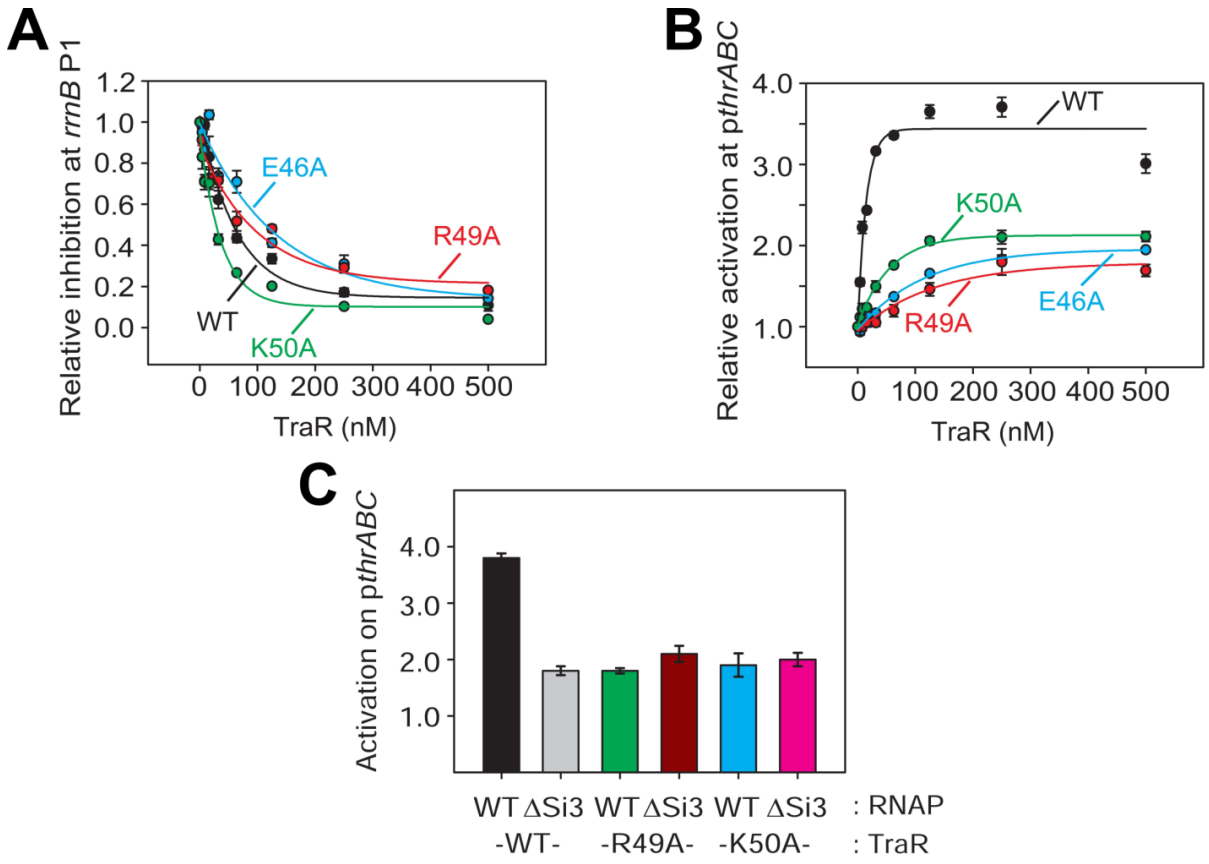


Figure 5.17. Si3 interaction with $TraR_G$ affects activation but not inhibition
(A) and **(B)** Multi round *in vitro* transcription of *rrnB* P1 **(A)** or p*thrABC* **(B)** was performed with 20 nM WT-Eσ^{70} at a range of concentrations of WT or variant TraR (2 nM -2 µM). TraR variants include single alanine mutations at residue positions E46, R49, K50 (highlighted in **Figures 5.15B-C**). Transcripts were quantified and plotted relative to values in the absence of any factor (n = 2). For *rrnB* P1 **(A)**, IC_{50} for inhibition by WT-TraR was ~50 nM, by E46A TraR was ~115 nM, R49A TraR was ~85 nM and by K50A TraR was ~30 nM. **(C)** Transcription experiments were carried out with 20 nM WT- or ΔSi3-RNAP with 250 nM WT- or variant TraR as indicated. Values are relative to basal transcription by WT-RNAP without factor (normalized to1.0). Averages with range from two independent experiments are shown. Transcription assay data and figure courtesy of A. Y. Chen and S. Gopalkrishnan from the Gourse Lab (University of Wisconsin–Madison).

5.4 Binding of TraR induces a ~18° rotation of βlobe-Si1

In RPo, the downstream duplex DNA exits the RNAP through the main channel created between the β'shelf and the clamp on one side, and the βlobe-Si1 domains on the other (**Figure 5.14**). In $E\sigma^{70}$ without nucleic acids, this channel is occupied by the $\sigma^{70}_{1.1}$ domain, which is ejected upon entry of the downstream duplex DNA during RPo formation (Bae et al., 2013; Mekler et al., 2002) (**Figures 5.6D**, **5.14**). TraR binding induces a ~ 18° rotation of the RNAP βlobe domain along with the attached LSI, βSi1 [also called βi4; (Lane and Darst, 2010a)]. This rotation towards TraR, allows the βlobe-Si1 to establish a 615 $Å^2$ molecular interface with $TraR_G$ and $TraR_C$ (**Figure 5.18A**).

Most of the TraR/βlobe-Si1 interface (77%) is between TraR and Si1. Deleting Si1 from RNAP strongly effects inhibition [*rpsT* P2; **Figure 5.18B,** (Chen et al., 2019b)] and nearly abolishes activation function [p*argI,* **Figure 5.18C**; *thrABC*, Gopalkrishnan et al., 2017], even at saturating concentrations of TraR (Gopalkrishnan et al., 2017). These results suggest that the βlobe-Si1 rotation induced by TraR is critical for inhibition and activation of transcription by TraR.

In addition, the rotation of the βlobe-Si1 widens the gap between the βprotrusion and the βlobe (**Figure 5.18A**) and changes the shape of the downstream RNAP channel where $\sigma^{70}_{1.1}$ of $E\sigma^{70}$ resides (**Figure 5.6D**). This rotation alters RNAP contacts with $\sigma^{70}_{1.1}$ and may facilitate $\sigma^{70}_{1.1}$ ejection during RPo formation, contributing to activation of promoters that are limited at this step. To test this, A. Y. Chen and S. Gopalkrishnan from the Gourse Lab (University of Wisconsin) examined TraR function on an inhibited (*rrnB* P1) and an activated (*thrABC*) promoter with holoenzyme lacking $\sigma^{70}_{1.1}$ ($E\Delta1.1\sigma^{70}$) (Chen et al., 2019b).

In the absence of TraR, $E\sigma^{70}$ exhibited weak basal transcription from the *thrABC* promoter. In the presence of TraR, transcription from this promoter was stimulated about ~4 fold (**Figure 5.19A**). Strikingly, $E\Delta1.1\sigma^{70}$ exhibited a large increase in basal transcription activity from this promoter (~32 fold) compared to WT-$E\sigma^{70}$ (**Figure 5.19A**). Only a small additional increase in transcription was observed upon the addition of TraR (**Figure 5.19A**). These results suggest that $\sigma^{70}_{1.1}$ is an obstacle to promoter DNA entering the RNAP channel and that TraR partially overcomes this barrier. In contrast to deletion of region $\sigma^{70}_{1.1}$, which almost entirely bypassed the requirement for TraR, rotation of the βlobe-Si1 did not weaken $\sigma^{70}_{1.1}$-RNAP contacts sufficiently to release $\sigma^{70}_{1.1}$ completely (**Figure 5.6D**). Rather, the βlobe-Si1 rotation may facilitate the competition between promoter DNA and $\sigma^{70}_{1.1}$ for the occupancy of the downstream DNA channel during RPo formation.

In contrast to the dramatic effect of $E\Delta1.1\sigma^{70}$ on activated promoters, deletion of $\sigma^{70}_{1.1}$ had only a small effect on basal transcription from the TraR-inhibited *rrnB* P1 promoter (**Figure 5.19B**). Inhibition of *rrnB* P1 by TraR with $E\Delta1.1\sigma^{70}$ was only slightly defective, with an IC_{50} of ~90 nM versus an IC_{50} of 50 nM with WT-$E\sigma^{70}$ (**Figure 5.19C**). Thus, unlike the effects of βlobe-Si1 domains on activation by TraR, the effects of the βlobe-Si1 rotation on TraR-mediated inhibition are not strongly coupled to $\sigma^{70}_{1.1}$. We propose that TraR-mediated stimulation of $\sigma^{70}_{1.1}$ release still occurs at inhibited promoters like *rrnB* P1 and *rpsT* P2, but this has little effect on transcription because these promoters are limited by their unstable RPo (Barker et al., 2001).

5.5 TraR induces a β'shelf rotation and kinks the bridge helix

TraR binding induces a ~ 4.5° rotation of the β'shelf module (**Figures 5.20A-B**). The shelf module is connected to the BH, a long α-helix that traverses the RNAP active site cleft from one

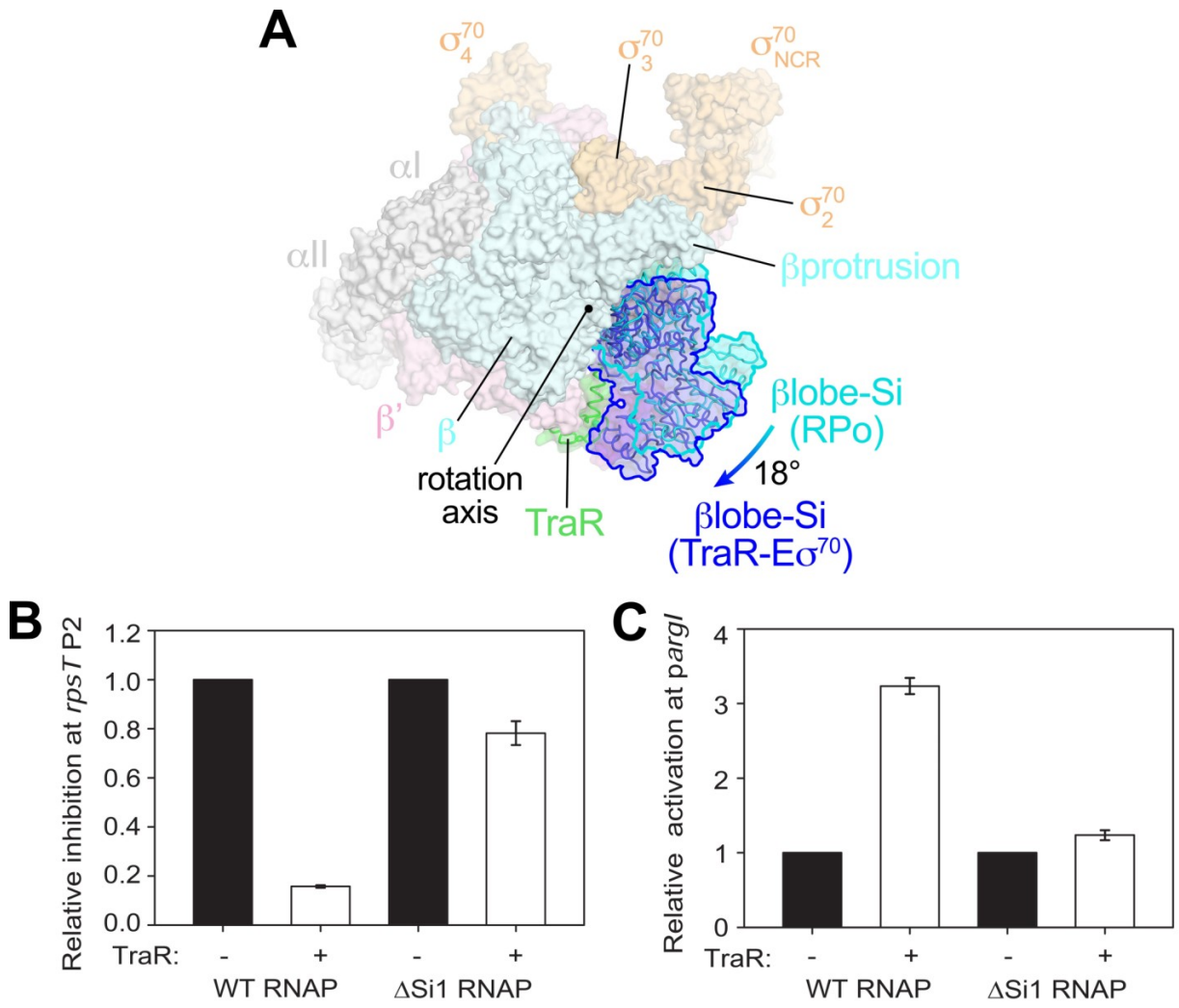


Figure 5.18. TraR induces a rotation in the βlobe-Si1 domain
(A) Overall top view of the TraR-Eσ^{70} structure with the βlobe-Si1 shown in cartoon tubes and outlined in dark blue. The corresponding position of the βlobe-Si1 in the *rpsT* P2-RPo structure (**Figure 5.14**) is superimposed and outlined in light blue. The βlobe-Si1 of the *rpsT* P2-RPo structure (light blue) undergoes an ~18° rotation about the rotation axis shown to the βlobe-Si1 position in the TraR-Eσ^{70} structure (dark blue), generating an extensive TraR-βlobe-Si1 interface. **(B)** Transcription of inhibited promoter *rpsT* P2 by 20 nM WT-RNAP or ΔSi1-RNAP with (+) or without (-) 250 nM TraR as indicated. **(C)** Transcription of activated promoter p*argI* by 20 nM WT-RNAP or ΔSi1-RNAP with (+) or without (-) 250 nM TraR as indicated. Error bars denote standard deviation of three independent measurements. Transcription assay data courtesy of A. Y. Chen and S. Gopalkrishnan from the Gourse Lab (University of Wisconsin–Madison).

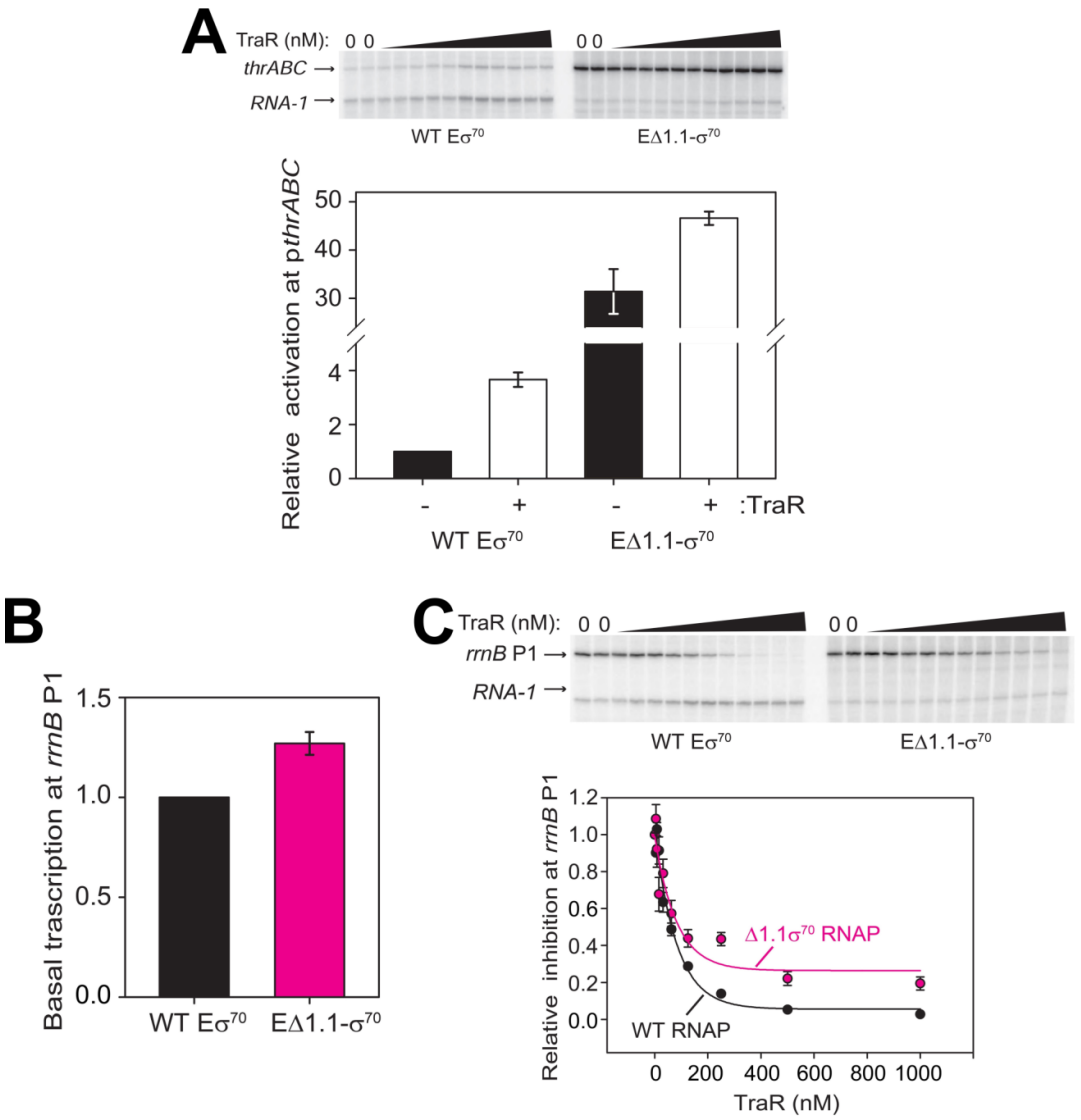


Figure 5.19. EΔ1.1σ^{70} has small defects for inhibition of *rrnB* P1 but large effects for activation of p*thr*ABC by TraR

(A) (*top*) Multi-round *in vitro* transcription at *pthrABC* was carried out at a range of TraR concentrations (wedge indicates 4 nM - 4 μM) in the presence of 20 nM WT-Eσ^{70} or EΔ1.1σ^{70} as indicated. Plasmid template also contained the RNA-1 promoter. (*bottom*) Transcripts were quantified and plotted relative to values in the absence of TraR with WT-Eσ^{70} or EΔ1.1σ^{70} with (+) or without (-) 250 nM TraR as indicated. Averages with range from two independent experiments are shown. **(B)** Basal level of transcription from *rrnB* P1 is only slightly affected by Δ1.1σ^{70}. Error bars denote standard deviation of three independent measurements. **(C)** (*top*) Multi round *in vitro* transcription at *rrnB* P1 was carried out at a range of TraR concentrations (wedge indicates 4 nM - 4 μM) in the presence of 20 nM WT-Eσ^{70} or EΔ1.1σ^{70} as indicated. Plasmid templates also contained the *RNA-1* promoter. (*bottom*) Transcripts were quantified and plotted relative to values in the absence of TraR. For *rrnB* P1, the IC_{50} for inhibition by TraR was ~50 nM with WT-RNAP and ~90 nM with Δ1.1σ^{70}-RNAP. Transcription assay data and figure courtesy of A. Y. Chen and S. Gopalkrishnan from the Gourse Lab (University of Wisconsin–Madison).

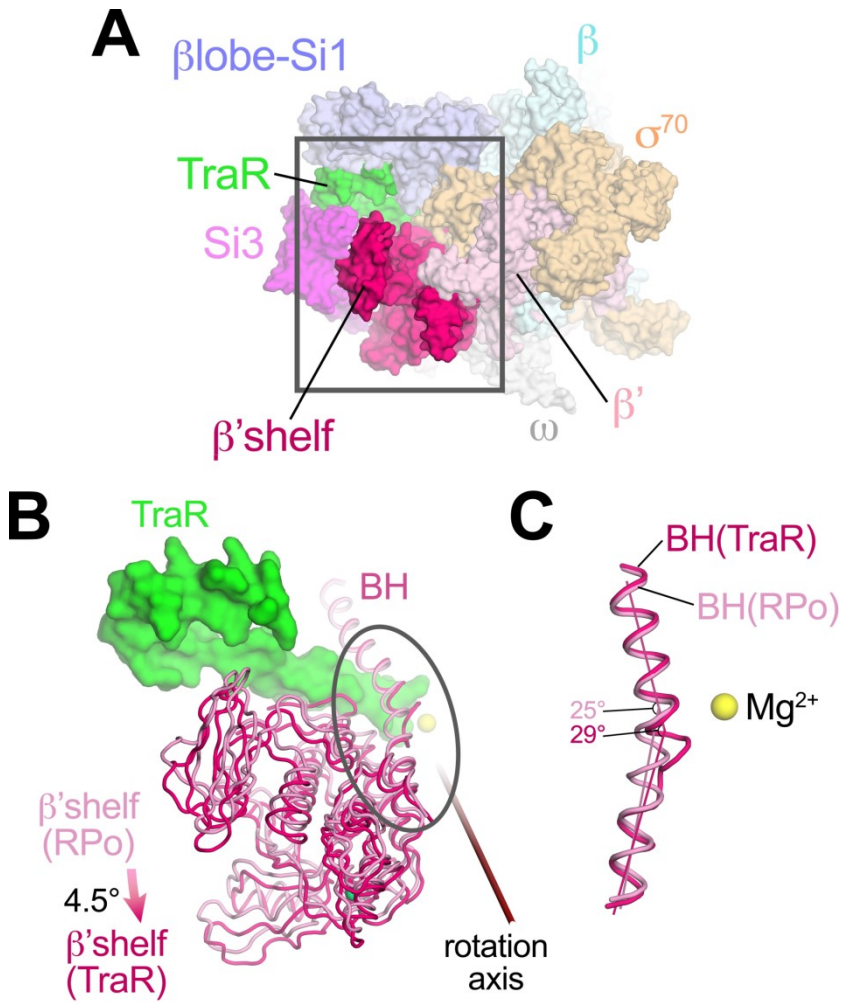


Figure 5.20. TraR rotates the β'shelf and kinks the BH
(A) Overall view if the TraR-Eσ^{70}(I) structure, shown as a molecular surface. The β'shelf domain is highlighted in hot pink. The β'shelf (which here includes the β'jaw) comprises *Eco* β' residues 787-931, 1135-1150 and 1216-1317. **(B)** Magnified view of boxed area in **(A)** showing the comparison of the *rpsT* P2-RPo BH-β'shelf (pink) and the TraR-Eσ^{70} BH-β'shelf (hot pink). Binding of TraR induces an ~4.5° rotation (about the rotation axis shown) of the RPo-β'shelf to the position of the TraR-Eσ^{70} β'shelf and a kink in the BH (circled region). **(C)** Magnified view of circled region in **(B)** with focus on the BH kink, which is centered near β' L788. The kink in the RPo BH is about 25°, while the kink in the TraR-Eσ^{70} BH is about 29°.

pincer to the other and lies directly across from the active site Mg^{2+}. The TraR-mediated rotation of the β'shelf introduces a kink in the BH (**Figures 5.20B-C**). The BH plays critical roles in the RNAP nucleotide addition cycle (Lane and Darst, 2010b), including interacting with the template strand (t-strand) DNA at the active site (**Figure 5.21A**). TraR causes the BH (**Figure 5.20C**) to bend towards the t-strand DNA, similar to BH kinks observed previously (Tagami et al., 2010, 2011; Weixlbaumer et al., 2013; Zhang et al., 1999), resulting in a steric clash with the normal position of the t-strand nucleotide at +2 (**Figure 5.21B**). Thus, the TraR-induced BH kink would sterically prevent the proper positioning of the t-strand DNA in RPo, likely contributing to inhibition of transcription.

5.6 TraR binding dampens clamp motions in $E\sigma^{70}$

TraR induces discrete conformational changes in RNAP structural elements: β'Si3 (**Figures 5.6A-B**, **5.15A**), βlobe-Si1 (**Figure 5.18A**), β'shelf, and BH (**Figure 5.20**). However, only modest changes in clamp positions were observed (**Table 5.22**). Dynamic clamp behavior has been observed by single-molecule Förster resonance energy transfer (FRET) in solution (Duchi et al., 2018). It was likely that the clamp module of $E\sigma^{70}$ adopts a continuous distribution of positions that could not be easily classified into discrete states. This conformational heterogeneity of $E\sigma^{70}$ limits the resolution of the single particle reconstruction (**Figure 5.8**). Therefore, I examined the heterogeneity of RNAP clamp positions between the $E\sigma^{70}$, TraR-$E\sigma^{70}$, and *rpsT* P2-RPo datasets using multibody refinement (MBR) as implemented in RELION 3.0 (Nakane et al., 2018).

In order to compare clamp motions using RELION MBR, the three datasets were processed equivalently (**Figures 5.3**, **5.7**, **5.11**)**.** After an initial round of 3D classification and particle polishing, a random subset of particles (370,965) was selected from each dataset (TraR-$E\sigma^{70}$, $E\sigma^{70}$, and *rpsT* P2-RPo) so that each map for MBR was generated from the same number of particles. The final maps used for MBR had nominal resolutions of 4.0 Å (TraR-$E\sigma^{70}$; red dashed box in **Figure 5.3**), 4.6 Å ($E\sigma^{70}$; red dashed box in **Figure 5.7**), and 3.5 Å (*rpsT* P2-RPo; red dashed box in **Figure 5.11**).

For $E\sigma^{70}$, three major components (Eigenvectors) of clamp changes were revealed (**Figures 5.23A-D**). For each Eigenvector, the histogram of Eigenvalues closely approximated a Gaussian distribution (**Figures 5.23B-D**). To quantitate the range of clamp conformations represented by the Eigenvalues, I divided the particles into three bins according to their Eigenvalues such that each bin contained an equal number of particles (red, gray, and blue in (**Figures 5.23B-D**). Three-dimensional alignments and reconstructions were then calculated for each bin.

For component 1, the red and blue particles gave rise to reconstructions that differed in clamp positions by a rotation angle of 2.7° in a direction we termed opening/closing (**Figure 5.23E**). The negative Eigenvalue particles yielded a closed clamp structure (red), while the positive Eigenvalue particles (blue) gave an open clamp structure. In the middle, the particles having intermediate Eigenvalues (gray) gave a clamp position half-way in between the red and the blue, as expected. Component 2 gave rise to clamp positions that differed by a 4.6° rotation about a rotation axis roughly perpendicular to the open/close rotation axis, a conformational change we termed twisting (**Figure 5.23F**). Finally, component 3 gave rise to clamp positions that differed by a 2.0° rotation about a third rotation axis parallel with the long axis of the clamp, a conformational change we termed rolling (**Figure 5.23G**).

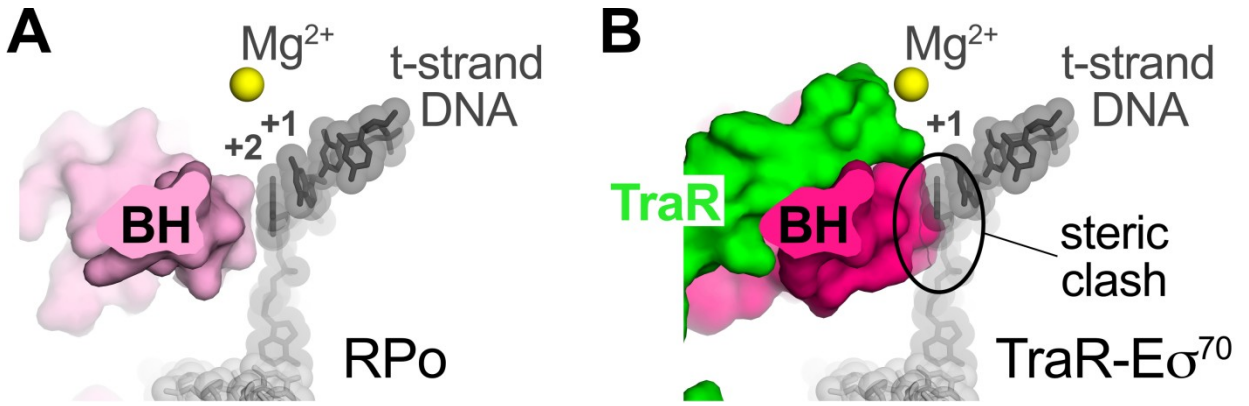


Figure 5.21. Kink in BH clashes with t-strand loading
(**A**) View down the axis of the *rpsT* P2-RPo BH. The t-strand DNA (shown as sticks superimposed on transparent sphere, colored dark grey), positioned at the RNAP active site (marked by the Mg^{2+} ion, yellow sphere), closely approaches the BH. (**B**) View down the axis of the TraR-Eσ^{70} BH. The BH kink induced by TraR binding sterically clashes with the position of the t-strand DNA [superimposed from the RPo structure, **(A)**].

Table 5.22. Comparison of RNAP conformational changes

structure	reference structure	PyMOL align[a]	PyMOL rms_cur[b]						align	rms_cur
		structural core[c]	structural core	overall (-structural core)	clamp[d]	β'shelf-jaw[e]	βlobe-Si1[f]		BH[g]	
Eσ^{70}	RPo	0.647 Å (1273 Cαs)	0.839 Å (1356 Cαs)	2.480 (1804)	3.396 3.3° open	1.735 2.7° twisted open	2.912 3.4° DNA->		0.484 (33)	0.55 (35)
TraR-Eσ^{70}(I)	RPo	0.454 (1249)	0.673 (1380)	4.931 (1817)	2.880 3.6° open	2.675 4.1° open	10.213 18° ->TraR		0.81 (30)	1.478 (35)
	H		0.873 (1355)	4.567 (1806)	1.413 1.8° roll-CCW	1.755 2.7° open	9.993 17° ->TraR		0.764 (30)	1.485 (35)
PDB 5w1s-A (TraR-Eσ^{70} complex A)[h]	TraR1	0.632 (1251)	0.871 (1377)	6.323 (1646)	3.837 5.7° open	2.523 3.2° closed	13.968 24° TraR->			
	RPo		0.604 (1256)	3.354 (1638)	5.906 7.8° open	2.211 1.8° twisted open	3.912 6.1° TraR->		0.585 (35)	0.585 (35)
	H		0.680 (1241)	2.699 (1629)	3.226 4.4° open	2.030 0.86°	5.007 8.1° TraR->			
	4yg2a	0.344 (1306)	0.541 (1482)	1.089 (1648)	1.587 1.7° open	1.174 1.2°	1.351 2.4° TraR->			
PDB 5w1s-B (TraR-Eσ^{70} complex B)[h]	TraR1	0.731 (1279)	0.963 (1369)	5.780 (1637)	2.583 4.8° open	2.948 3.7° closed	12.498 22° TraR->			
	RPo		0.938 (1370)	2.865 (1635)	4.376 5.7° open	2.826 3.9° twisted	2.955 5.7° DNA->		0.509 (33)	0.576 (35)
	H		0.989 (1355)	2.7 (1627)	2.256 3.4° roll-CCW	2.365 1.8° closed	4.998 8.8° TraR->			
	4yg2b	0.358 (1283)	0.47 (1372)	0.73 (1638)	0.864 0.89°	0.887 1.2°	0.622 1.2°			
PDB 4yg2-A (Eσ^{70} complex A)[i]	RPo								0.519 (35)	0.519 (35)
PDB 4yg2-B (Eσ^{70} complex B)[i]	RPo								0.415 (31)	0.568 (35)

[a]The structure in the first column (structure) was aligned to the structure in the second column (reference structure) by Cα atoms. The entries list the resulting RMSD (Å, top row) and the number of Cα atoms aligned (in parentheses). For the clamp, βlobe-Si1, and β'shelf-jaw entries, the rotational angle and axis between the elements was. The rotation angle is listed in the second row. The direction of movement (from the target structure to the reference structure) is denoted below.

[b]The structures/structural elements denoted were compared (but not aligned) by Cα atoms by calculating the RMSD for all of the specified atoms.

[c]*Eco* RNAP structural core: αI, αII, β (1-27, 147-152, 445-455, 520-713, 786-828, 1060-1240), β' (343-368, 421-786), ω.

[d]*Eco* RNAP clamp: β (1319-1342), β' (1-342, 1318-1344), σ^{70} (92-137, 353-449).

[e]*Eco* RNAP β'shelf-jaw: β' (787-931, 1135-1315).

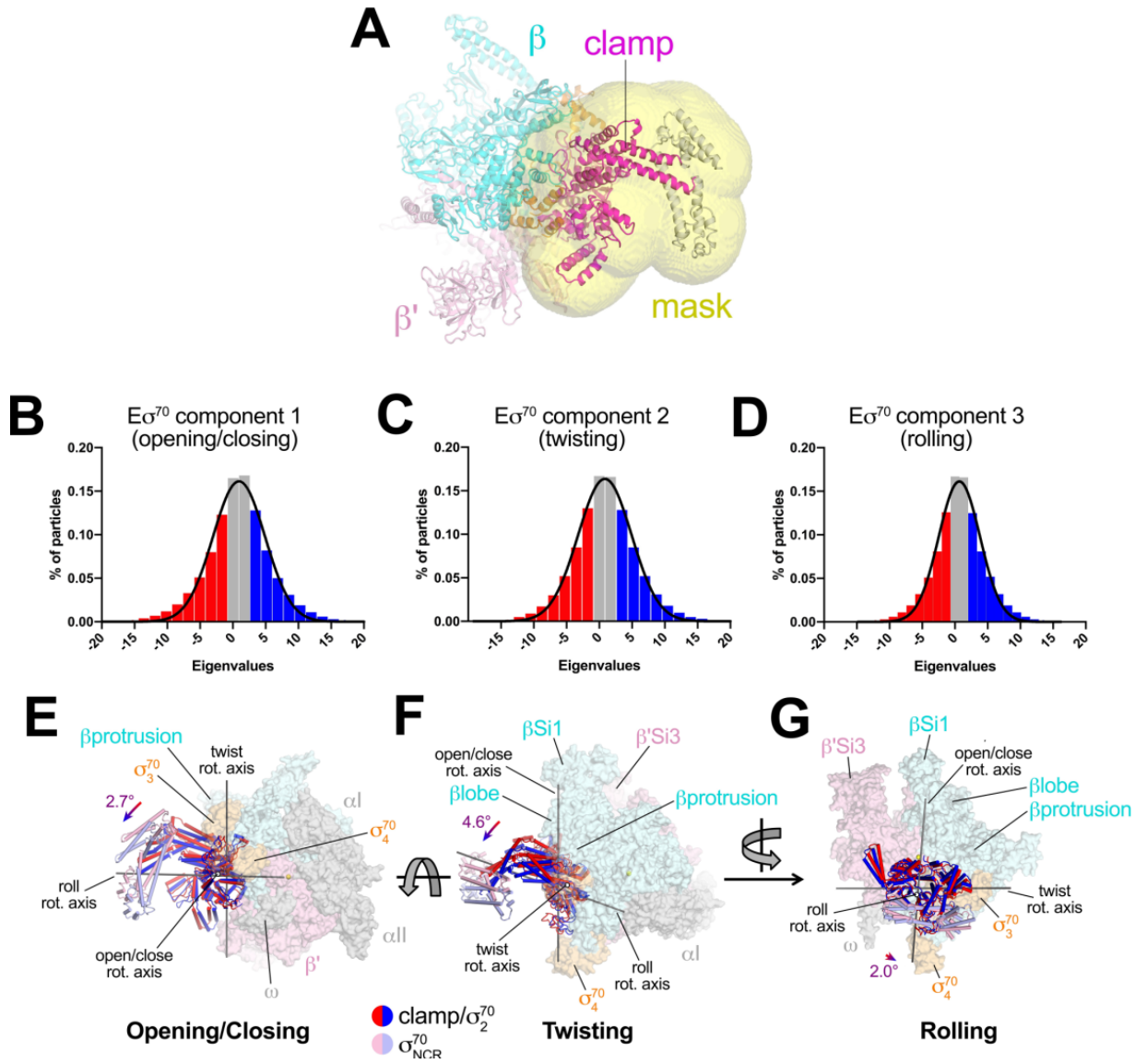
[f]*Eco* RNAP βlobe-Si1: β (153-444).

[g]*Eco* RNAP BH: β' (769-803).

[h](Molodtsov et al., 2018).

[i](Murakami, 2013).

Figure 5.23. Multi-body analysis of $E\sigma^{70}$ clamp conformational changes
(**A**) Model of $E\sigma^{70}$ refined into the consensus cryo-EM map (nominal 4.1 Å resolution). The RNAP clamp is highlighted in magenta. The clamp (which in the context of $E\sigma^{70}$ includes $\sigma^{70}{}_2$) comprises the following *Eco* RNAP residues: β 1319–1342; β' 1–342, 1318–1344; σ^{70} 92–137, 353–449. The mask used to analyze clamp motions by multi-body refinement (Nakane et al., 2018) is shown as a transparent yellow surface. (**B - D**) Histograms of Eigenvalue distributions (% of particles assigned to each Eigenvalue from the dataset) for the first three major principle components (Eigenvectors) generated from the multi-body analysis (Nakane et al., 2018). Each set of particles were divided into three equal-sized bins based on eigenvalue (red; gray; blue). The solid lines denote Gaussian fits to the histograms. (**B**) Component 1 - clamp opening/closing. (**C**) Component 2 - clamp twisting. (**D**) Component 3 - clamp rolling. (**E - G**) Three-dimensional reconstructions were calculated from the red and blue-binned particles for each principal component and models were generated by rigid body refinement in PHENIX (Adams et al., 2010). The models were aligned using the Cα of the RNAP (shown as a faded molecular surface) structural core module, revealing the alternate clamp positions shown (red and blue α-carbon ribbons with cylindrical helices). The $\sigma^{70}{}_{NCR}$, attached to the clamp but not included in the clamp motion analyses, is shown in faded colors. For each component, the clamp rotation and the direction of the rotation axis were determined (rotation axes are shown in gray). (**E**) Component 1 - clamp opening/closing. (**F**) Component 2 - clamp twisting. (**G**) Component 3 - clamp rolling.


A
β
clamp
β'
mask
B
Eσ70 component 1
(opening/closing)
C
Eσ70 component 2
(twisting)
D
Eσ70 component 3
(rolling)
% of particles
Eigenvalues
E
βprotrusion
twist
rot. axis
σ3 70
2.7°
αI
σ4 70
roll
rot. axis
open/close
rot. axis
αII
β'
ω
Opening/Closing
F
βSi1
β'Si3
open/close
rot. axis
βlobe
βprotrusion
4.6°
twist
rot. axis
roll
rot. axis
αI
σ4 70
clamp/σ2 70
σNCR 70
Twisting
G
β'Si3
βSi1
open/close
rot. axis
βlobe
βprotrusion
roll
rot. axis
twist
rot. axis
ω
2.0°
σ3 70
σ4 70
Rolling

Using the parameters of the Gaussian fits to the Eigenvalue histograms (**Figures 5.23B-D**), the full range of clamp rotations was calculated for each component, which I defined as the rotation range that accounted for 98% of the particles (excluding 1% of the particles at each tail; **Figure 5.24**).

These same components (opening/closing, twisting, rolling) explained the major clamp changes for the TraR-Eσ^{70} and *rpsT* P2-RPo particles as well. The same analyses revealed that TraR binding significantly reduced the range of clamp positions for each of the three clamp motions (**Figure 5.24**). As expected, the clamp positions for RPo, with nucleic acids stably bound in the downstream duplex channel, were restricted even further for all three of the major clamp motions (**Figure 5.24**).

5.7 Discussion

My cryo-EM structural analyses show that TraR binding modulates *Eco* RNAP transcription by altering the discrete conformation and dynamics of the RNAP in four major ways: (1) manipulation of the disposition of β'Si3 (**Figures 5.6**, **5.15**); (2) alteration of the shape of the RNAP active site cleft through a large rearrangement of the βlobe-Si1 (**Figure 5.18**); (3) induction of a significant kink in the BH (**Figure 5.20**); and (4) dampening the range of clamp positions (**Figures 5.23**, **5.24**). None of these four major TraR-mediated conformational changes seen in the cryo-EM analysis were observed in the crystal structure of TraR-Eσ^{70} (Molodtsov et al., 2018) (**Table 5.22**). Comparing RNAP conformations, the TraR-Eσ^{70} crystal structure (5W1S) matches the Eσ^{70} crystal structure [PDB ID 4YG2, the same crystal form from which the TraR complex was derived; Murakami, 2013] much more closely than the TraR-Eσ^{70} cryo-EM structure (**Table 5.22**). Thus, crystal packing constraints prevented the conformation of the RNAP from properly responding to TraR binding (**Chapter 1, Figure 1.6C**).

The results highlight important advantages of cryo-EM over crystallography for structural analysis of large, conformationally dynamic molecular machines such as RNAP (Bai et al., 2015a). First, single-particle cryo-EM analysis does not require crystallization and avoids limitations imposed by crystal packing. Second, multiple, discrete conformational states, such as TraR-Eσ^{70}(I), TraR-Eσ^{70}(II), and TraR-Eσ^{70}(III) (**Figures 5.6A-B**, **5.3**), can be revealed from a single sample (Bai et al., 2015b). Third, when a continuous conformational motion does not parse into discrete states such as movement of the RNAP clamp, the range of conformational states can nevertheless be assessed quantitatively (Nakane et al., 2018) (**Figures 5.23**, **5.24**).

The consequences of the TraR-induced conformational changes for promoter function (activation or inhibition) depend on the distinctly different properties of the promoters, which are attributable to differences in DNA sequence (Haugen et al., 2008; Sanchez-Vazquez et al., 2019). Eσ^{70} can melt and form RPo more readily for some promoters versus others. Likewise, the stability of the RPo can vary greatly between different promoters. The large range of promoter properties gives rise to activities of bacterial transcription initiation that vary over ~4 orders of magnitude and provide rich targets for regulation (Galburt, 2018; McClure, 1985).

Mechanistic studies of ppGpp/DksA- and TraR-dependent regulation of initiation revealed general characteristics of promoters that are either activated or inhibited by these factors (Gopalkrishnan et al., 2017; Gourse et al., 2018; Paul et al., 2004a, 2005; Rutherford et al., 2009). In the absence of factors, inhibited promoters generate RPo very rapidly (Rao et al., 1994), but the final transcription-competent RPo is relatively unstable. For example, the half-life of RPo for the inhibited promoter *rrnB* P1 is short, lasting only seconds to a few minutes (Barker et al., 2001). In the absence of either transcription factors or high initiating nucleotide (NTP)

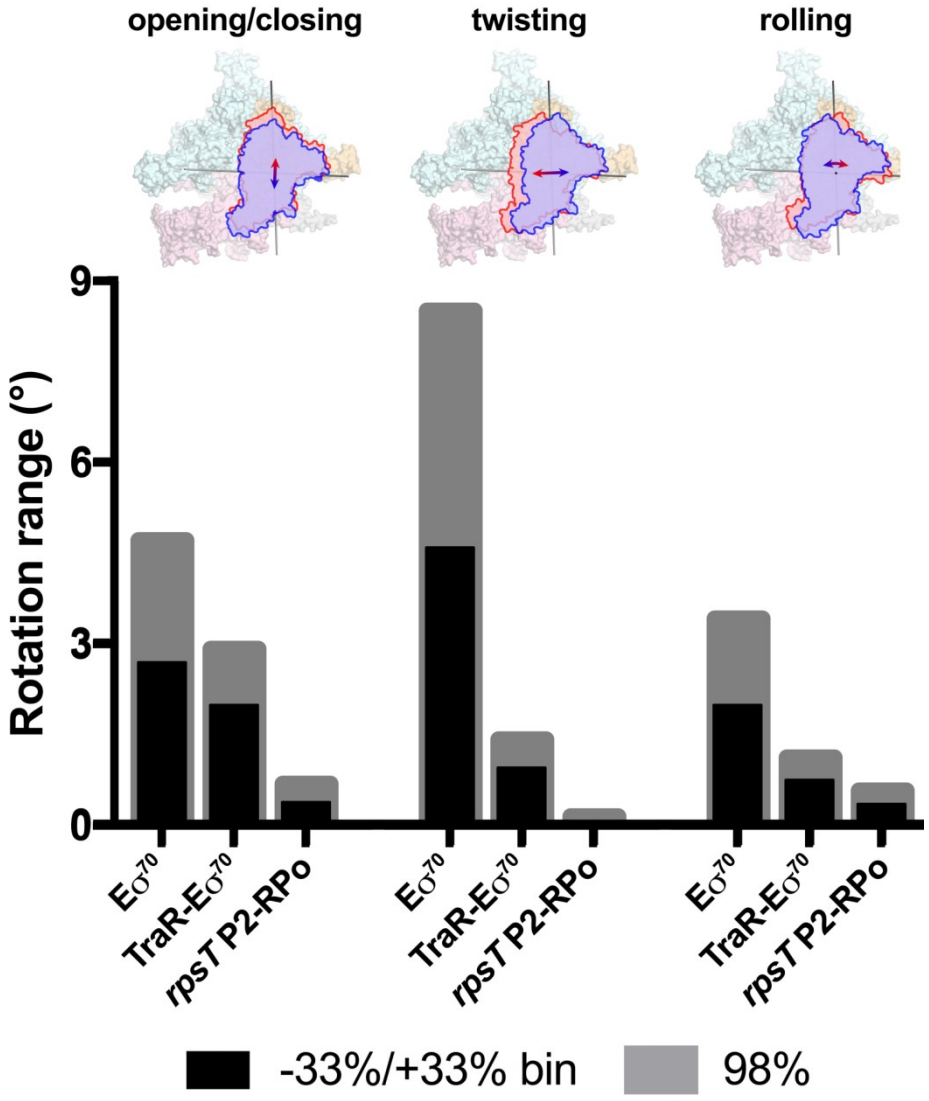


Figure 5.24. Range of clamp conformations for *Eco* RNAP complexes
(*top*) $E\sigma^{70}$ is shown as a molecular surface (α, ω, light gray; β, light cyan; β', light pink; σ^{70}, light orange) except the clamp/$\sigma^{70}{}_2$ module is shown schematically as blue or red outlines (the $\sigma^{70}{}_{NCR}$ is omitted for clarity) to illustrate the direction and approximate range for the three major components of the clamp conformational changes (left, opening/closing; middle, twisting; left, rolling). (*bottom*) Histograms denote the range of clamp conformational changes for $E\sigma^{70}$, TraR-$E\sigma^{70}$, and *rpsT* P2-RPo, as indicated. The black bars denote the rotation range defined by dividing the Eigenvalue histograms into three equal bins and determining the clamp position for the red and blue bins (−33 %/+33% bin; see **Figure 5.23**). The gray bars denote the estimated rotation range to include 98% of the particles calculated from the Gaussian fits to the Eigenvalue histograms (1% of the particles excluded from each tail; see **Figure 5.23**).

concentrations, RPo at *rrnB* P1 exists in equilibrium with earlier intermediates along the pathway to RPo formation (Gopalkrishnan et al., 2017; Rutherford et al., 2009). The very short RPo half-life at inhibited promoters means that initiation of RNA chain synthesis competes with dissociation of RPo. High NTP concentrations can shift the equilibrium in favor of RPo by mass action by populating complexes that follow RPo in the transcription cycle (Barker and Gourse, 2001; Murray et al., 2003).

By contrast, activated promoters generate RPo very slowly (Barker et al., 2001; Paul et al., 2005) but the RPo that is ultimately formed is stable. For example, the activated promoters p*argI*, p*hisG*, and p*thrABC* have RPo half-lives measured in many hours [15 hr,>13 hr, and 6.7 hr, respectively (Barker et al., 2001).

In order for a transcription factor, such as TraR, to achieve differential regulation of promoters while inducing the same conformational effects on RNAP, the factor must affect more than one feature of the multi-step pathway of RPo formation (Galburt, 2018) (**Figure 5.26A**). In our model for TraR function, TraR acts on all promoters similarly. TraR relieves kinetic barriers to accelerate RPo formation but at the same time stabilizes an intermediate prior to RPo formation (Galburt, 2018). Whether TraR activates or inhibits a promoter depends on the basal kinetic landscape for RPo formation at that promoter (**Table 5.25**, **Figure 5.26**). Our structural analysis and biochemical assessment of the conformational changes imparted on Eσ^{70} by TraR binding provide insight into the molecular mechanism of TraR-mediated activation and inhibition and TraR effects on the RPo formation pathway.

5.7.1 Structural mechanism for TraR-mediated activation

Previous data showed that TraR or its homolog DksA (in conjunction with ppGpp) enhance the transcription output from activated promoters relative to that with RNAP alone (Gopalkrishnan et al., 2017; Paul et al., 2005). In addition, kinetic analyses demonstrated that the TraR homolog ppGpp/DksA enhances the rate and amount of RPo formation at an activated promoter (Paul et al., 2005) primarily by increasing the rates of step(s) subsequent to RNAP binding to the promoter.

In the absence of TraR, RNAP with a deletion of $\sigma^{70}_{1.1}$ displayed greatly increased promoter activity on the *thrABC* promoter relative to the activity with WT RNAP (**Figure 5.19A**). This large effect on promoter activity was not observed with *rrnB* P1, an inhibited promoter (**Figure 5.19C**). These results are consistent with previous reports that deletion of $\sigma^{70}_{1.1}$ had varying effects on transcription from different promoters (Hook-Barnard and Hinton, 2009; Ruff et al., 2015a; Vuthoori et al., 2001).

The results of this work indicate that the presence of $\sigma^{70}_{1.1}$ in the RNAP channel poses the most significant barrier to the formation of RPo for activated promoters. However, addition of TraR to RNAP lacking $\sigma^{70}_{1.1}$ resulted in a small TraR-dependent increase in transcription of the *thrABC* promoter, suggesting the presence of additional barriers that TraR can act on (**Figure 5.19A**). We suggest that this small increase could result from the restriction of clamp motion by TraR at a step prior to $\sigma^{70}_{1.1}$ ejection, perhaps transcription bubble nucleation (Feklistov et al., 2017).

TraR binding allosterically alters and weakens $\sigma^{70}_{1.1}$ interactions in the RNAP channel by causing a large rotation of the βlobe-Si1 module (**Figure 5.18A**). In addition, the interaction of $TraR_G$ with the SBHMa motif of Si3 (**Figures 5.15-5.17**) in [Si3(II)] may also contribute to weakening of $\sigma^{70}_{1.1}$ interactions with the main channel. Si3 is inserted within the TL and its interactions with TraR could perturb interactions of the TL/BH/Switch one (Sw1) region with

Table 5.25. Details of flux calculator (Galburt, 2018) calculations

constants		inhibited promoter	activated promoter
(Galburt, 2018) nomenclature	Current nomenclature	basal = no TraR reg = TraR	
k^{basal}_{on} ($nM^{-1}s^{-1}$)	k^{basal}_{2} (s^{-1})	1	0.1
k^{reg}_{on} (s^{-1})	k^{reg}_{2} (s^{-1})	5	0.5
k^{basal}/k^{reg}		5	5
k^{basal}_{off} (s^{-1})	k^{basal}_{-2} (s^{-1})	10	1
k^{reg}_{off} (s^{-1})	k^{reg}_{-2} (s^{-1})	10	1
k^{basal}/k^{reg}		1	1
k^{basal}_{open} (s^{-1})	k^{basal}_{3} (s^{-1})	1	0.1
k^{reg}_{open} (s^{-1})	k^{reg}_{3} (s^{-1})	10	1
k^{basal}/k^{reg}		10	10
k^{basal}_{close} (s^{-1})	k^{basal}_{-3} (s^{-1})	50	0.05
k^{reg}_{close} (s^{-1})	k^{reg}_{-3} (s^{-1})	2500	2.5
k^{basal}/k^{reg}		50	50
k^{basal}_{esc} (s^{-1})	k^{basal}_{4} (s^{-1})	5	5
k^{reg}_{esc} (s^{-1})	k^{reg}_{4} (s^{-1})	5	5
k^{basal}/k^{reg}		1	1
$flux^{basal}$ (min^{-1})		5.3	5.8
$flux^{reg}$ (min^{-1})		1.2	35
flux ratio (basal/reg)		0.23	6.0

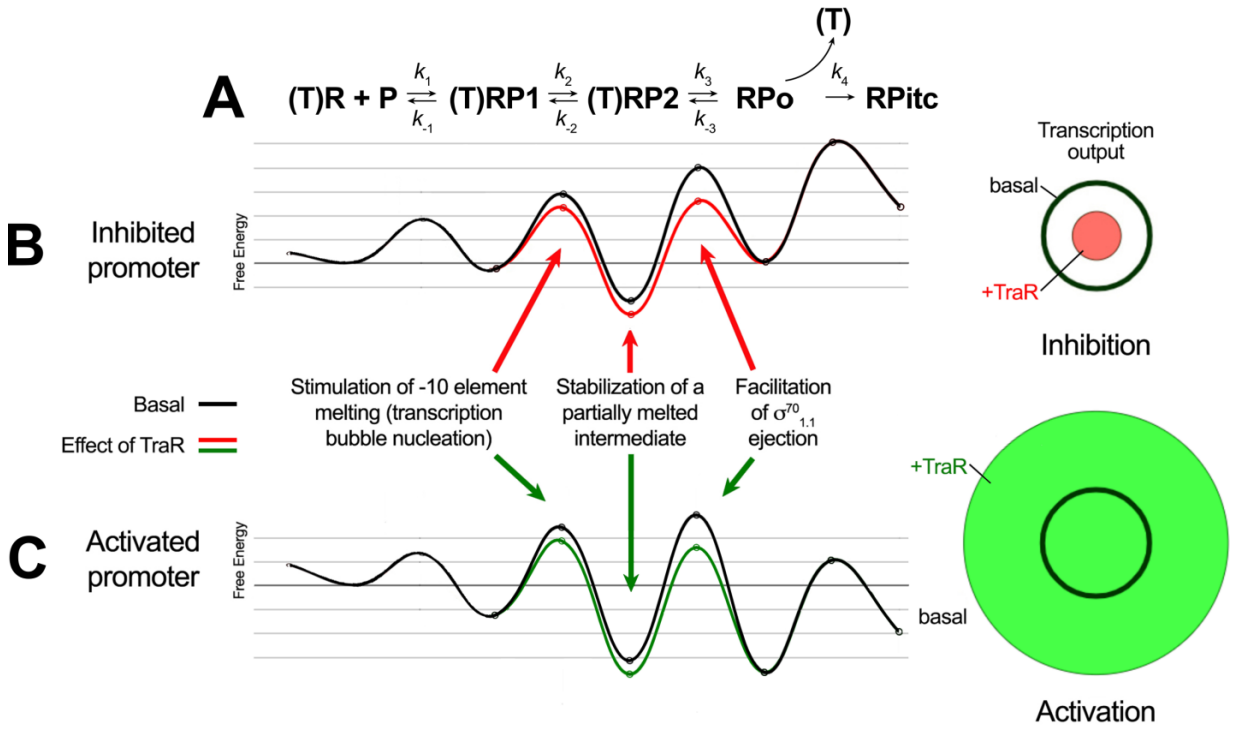


Figure 5.26. Proposed effects of TraR on the free energy diagram for hypothetical inhibited and activated promoters

(A) Proposed three-step linear kinetic scheme for RPo formation (Hubin et al., 2017b) with an added fourth irreversible step (formation of RP_{itc}) once RNA synthesis begins. (T) denotes the presence of TraR, which must dissociate to allow the transition of RPo → RPitc. The basal (WT-RNAP) free energy diagrams for hypothetical inhibited **(B)** and activated **(C)** promoters are shown in black (adapted from Galburt, 2018). The proposed influence of TraR on the energy diagram (lowering the kinetic barrier for the transition RP1 ⇄ RP2; lowering the free energy of RP2 relative to RPo; lowering the kinetic barrier for the transition RP2 ⇄ RPo) is shown (inhibited promoter, red curve; activated promoter, green curve) along with proposed links with the structural effects of TraR binding to RNAP described here. The steady-state transcription output [calculated with the transcription flux calculator (Galburt, 2018) is represented by the circles on the right. The precise values for the inputs and outputs for the flux calculator are tabulated in **Table 5.25**. (*right*) The area inside the black circle represents the basal transcription output. The red or green circles (inhibited or activated promoters, respectively) represent the effect of TraR on the transcription output. Kinetic analysis and figure courtesy of S. A. Darst (The Rockefeller University).

$\sigma^{70}_{1.1}$. Deletion of β'Si3 has been shown to reduce the lifetime of open complexes by 3-10 fold (Ruff et al., 2015a), suggesting the possibility that the reduced activation observed with the Si3 deletion RNAP could also result from effects on RPo stability.

The TraR-induced conformational changes that alter and weaken $\sigma^{70}_{1.1}$ interactions in the main channel are proposed to facilitate displacement of $\sigma^{70}_{1.1}$ by promoter DNA (**Figures 5.6D**, **5.18A**). However, effects of TraR on clamp conformation may also contribute to the increase in RPo by increasing bubble nucleation, as proposed by (Feklistov et al., 2017). Together, these effects could be sufficient to account for the TraR-dependent increase in the amount of RPo formed at an activated promoter.

The structural data indicate that TraR binding to the RNAP secondary channel would block NTP access to the active site and kink the BH to sterically clash with proper t-strand positioning for catalysis (**Figure 5.21B**; also see Molodtsov et al., 2018). Precise positioning of the t-strand DNA at the active site is critical for efficient catalysis of phosphodiester bond formation by RNAP in the S_N2 mechanism (Yee et al., 1979). Therefore, TraR must dissociate from the RNAP complex in order to allow transcription from the activated promoter. The proposal that TraR dissociates from the complex at an activated promoter prior to the dissociation of RNAP from the promoter is supported by previous data on the lifetimes of these complexes. The lifetime of RPo at activated promoters (measured in hours; Barker et al., 2001)) is much longer than the lifetime of the transcription factors on RNAP (GreB, DksA and by extension TraR; measured in seconds; Stumper et al., 2019)).

The kinetics of transcription initiation are illustrated by a schematic energy landscape in **Figure 5.26,** courtesy of S. A. Darst (The Rockefeller University), where specific structural effects of TraR are correlated with changes in free energy at positively and negatively regulated promoters. The TraR-dependent changes in RNAP conformation and the resulting effects on $\sigma^{70}_{1.1}$ would lower the kinetic barrier to RPo formation. At an activated promoter, the free energy of the transition state between RP2 and RPo is rate-limiting, and it is reduced by the TraR-induced weakening of $\sigma^{70}_{1.1}$ interactions in the main channel, leading to an increase in transcription (**Figure 5.26C**).

5.7.2 Structural mechanism for TraR-mediated inhibition

TraR (and ppGpp/DksA)-inhibited promoters form RPo rapidly (Rao et al., 1994) but their intrinsically unstable RPo can reverse to earlier melting intermediate states (Gopalkrishnan et al., 2017; Rutherford et al., 2009). TraR effects on βlobe-Si1 and clamp dynamics likely stimulate bubble nucleation and $\sigma^{70}_{1.1}$ ejection at inhibited promoters but these steps are already rapid for these promoters so the overall transcription output is unaffected. Deletion of $\sigma^{70}_{1.1}$ had much smaller effects on basal transcription and inhibition by TraR on an inhibited promoter (**Figures 5.19B-C**).

However, TraR binding induces two distinct conformational changes in RNAP that we propose disfavor RPo formation, accounting for inhibition. Most prominent is the rotation of βlobe-Si1 that alters the shape of the main channel and widens the gap between the βprotrusion. This large conformational change may exposes surfaces on the RNAP that stabilize DNA contacts in an intermediate prior to RPo (such as RP2; **Figure 5.26**). Stabilization of an intermediate prior to formation of a transcription-competent RPo would have a dramatic effect on transcription output at inhibited promoters. Consistent with these hypotheses, ΔSi1-RNAP has reduced TraR-mediated inhibition (**Figure 5.18B**), and footprints with RNAP on negatively

regulated promoters like *rrnB* P1 and *rpsT* P2 have a shortened downstream boundary of DNase I protection (Gopalkrishnan et al., 2017).

In addition to causing rotation of the βlobe-Si1, TraR binding inhibits TL folding and blocks NTPs from entering the secondary channel (Molodtsov et al., 2018). It also induces a kinked BH (**Figure 5.20**) that sterically clashes with the t-strand DNA near the active site (**Figure 5.21**; also see Molodtsov et al., 2018). Therefore, any RNAP complexes containing TraR that proceeded to RPo cannot initiate transcription. Recent single molecule fluorescence studies (Stumper et al., 2019) indicated that secondary channel binding factors (GreB, DksA and by extension TraR) formed an RNAP-factor complex prior to binding to promoter DNA. These factors dissociate together with RNAP from an inhibited promoter complex (*rrnB* P1) because the lifetime of RNAP on *rrnB* P1 is shorter than the lifetime of the factors on RNAP (Stumper et al., 2019). Binding of the factors to preformed RPo was not observed (Stumper et al., 2019), consistent with the ten-fold reduced affinity of DksA for RPo (Lennon et al., 2009).

These data are consistent with our model for inhibition, suggesting that TraR would bind together with RNAP to an inhibited promoter and severely reduce RPo formation by stabilizing an intermediate (**Figure 5.26B**). TraR would remain associated with the complex until RNAP dissociated. TraR's presence in the complex for the entire time of RNAP occupancy of the promoter would leave little opportunity for RPo formation and catalysis.

5.7.3 TraR manipulates *Eco* RNAP lineage-specific insertions to modulate transcription initiation

A key feature of the mechanism of TraR function is modulation of *Eco* RNAP transcription initiation through conformational changes brought about by interactions with two of the *Eco* RNAP LSIs, βSi1 (**Figure 5.18A**) and β'Si3 (**Figures 5.15A-C**). In our cryo-EM and biochemical analysis, TraR interacts with βSi1 as well as with the nearby βlobe to distort the RNAP channel (**Figure 5.18A**), effecting both inhibition (**Figure 5.18B**) and activation (**Figure 5.18C**) by TraR. β'Si3 is often disordered in *Eco* RNAP crystal structures [for example, see Molodtsov et al., 2018]. In our cryo-EM structures, TraR engages with β'Si3, stabilizing a previously undetected conformation of Si3 that plays a critical role in TraR-mediated activation (**Figure 5.15**). Si3 has been implicated previously in RPo formation since the Δβ'Si3-RNAP forms an unstable RPo (Artsimovitch et al., 2003).

5.8 Acknowledgements

The work in this chapter is published in eLife. My role in this work involved: protein purification, cryo-EM sample preparation; data collection and processing; structural model building and analysis; conception of ideas and figures. Citation:

Chen, J., Gopalkrishnan, S., Chiu, C., Chen, A.Y., Campbell, E.A., Gourse, R.L., Ross, W., and Darst, S.A. (2019). *E . coli* F element-encoded TraR protein allosterically regulates transcription initiation by altering RNA polymerase conformation and conformational dynamics. eLife *8*, 1–81.

I would like to thank the following people for their contributions to this chapter: E. A. Campbell and S. A. Darst (The Rockefeller University) provided mentorship and structural data analysis; A. Y. Chen and S. Gopalkrishnan (University of Wisconsin–Madison) performed biochemical and mutational analyses; C. Chiu assisted with cryo-EM sample preparation; S. A. Darst (The

Rockefeller University) edited this chapter; M. Ebrahim and J. Sotiris (The Rockefeller University) helped with cryo-EM data collection; A. Feklistov and R. M. Saecker (The Rockefeller University) provided helpful discussions; R. L. Gourse and W. Ross (University of Wisconsin–Madison) collaborated with our laboratory and supervised various aspects of the project.

CHAPTER 6. Structural studies of stepwise promoter melting by *Eco* RNAP

6.1 Introduction

Transcription initiation requires the formation of the open promoter complex (RPo). To generate RPo, the RNAP holoenzyme binds a duplex promoter DNA, unwinds the DNA to form the transcription bubble, and loads the template-strand (t-strand) into the RNAP active site for catalysis (Abascal-Palacios et al., 2018; Bae et al., 2015a; He et al., 2016; Nagy et al., 2015; Plaschka et al., 2016; Tafur et al., 2016; Vorländer et al., 2018; Zuo and Steitz, 2015). The vast majority of initiation events in bacteria involve the RNAP (E) associated with the primary σ factor, σ^{70} in *Escherichia coli* (*Eco)* (Feklístov et al., 2014; Gruber and Gross, 2003). Eσ^{70} utilizes binding free energy to generate RPo through a multi-step pathway involving multiple RNAP-promoter intermediates of unknown structure (Hubin et al., 2017b; Ruff et al., 2015b; Saecker et al., 2002).

Structures of bacterial RPo have been well characterized (Bae et al., 2015a; Boyaci et al., 2019a; Hubin et al., 2017a; Narayanan et al., 2018; Zuo and Steitz, 2015), but the structural basis for RPo formation is poorly understood. Intermediates along the RPo formation pathway, from the initial recognition of the duplex promoter in the so-called closed complex (RPc) to the final RPo (Ruff et al., 2015b), are difficult to study due to their transient and dynamic nature. Our laboratory recently reported the structure of a late melting intermediate on the RPo formation pathway (deemed RP2). This intermediate was obtained with *Mtb* RNAP, which forms a relatively unstable RPo in equilibrium with RP2 (Boyaci et al., 2019a). Additionally, previous kinetic analyses used salt, urea, and other perturbants to identify intermediates of RPo formation in solution (Gries et al., 2010; Kontur et al., 2008, 2010).

In collaboration with the Gourse Lab (University of Wisconsin–Madison) and the Chait Lab (The Rockefeller Univerity), I investigated RPo formation with Eσ^{70} on the S20 ribosomal protein promoter, *rpsT* P2 (Lemke et al., 2011). To populate RPo intermediates for visualization, I added TraR, an F plasmid-encoded transcription factor that inhibits *rpsT* P2 activity (Blankschien et al., 2009; Gopalkrishnan et al., 2017). TraR binds the RNAP directly without contacting promoter DNA directly, altering the RNAP conformation and increasing the occupancy of intermediates on the RPo pathway (**Chapter 5**; Chen et al., 2019b). Cryo-electron microscopy (cryo-EM) is a useful technique to examine these intermediate states in solution because it can sort conformational heterogeneity within a sample into distinct classes (Bai et al., 2015a).

Using single-particle cryo-EM and image classification approaches, I solved the structures of five intermediates formed by the wild-type (WT) *rpsT* P2 promoter, and two additional intermediates formed by a mutant *rpsT* P2 promoter. The structures span the RPo formation pathway from the initial Eσ^{70} recognition of the duplex promoter to the final RPo. The structures and supporting biochemical data, provided by the Gourse Lab (University of Wisconsin-Madison), define RNAP and promoter DNA conformational changes that delineate steps on the pathway, including previously undetected transient promoter-RNAP interactions that contribute to populating the intermediates but do not occur in RPo. Features of the structures allow their placement in an ordered pathway that provides a structural basis for understanding RPo formation in all organisms, a major point of regulation for gene expression Portions of this chapter are adopted from a manuscript that is currently in press at Molecular Cell (Chen et al., in press).

6.2 TraR stabilizes a partially melted intermediate on the *rpsT* P2 promoter

TraR-Eσ^{70}-*rpsT* P2 complexes were formed by incubating TraR-Eσ^{70} with the linear duplex fragment of WT *rpsT* P2 used in **Chapter 5** (**Figure 5.10**). TraR-Eσ^{70}-*rpsT* P2 complexes were detected by footprinting (**Figure 6.1**) and native mass spectrometry (nMS, **Figure 6.2**).

A. Y. Chen, S. Gopalkrishnan, M. F. Maloney, W. Ross and J. T. Winkelman from the Gourse Lab (University of Wisconsin–Madison) employed footprinting approaches to examine intermediates of TraR-Eσ^{70}-*rpsT* P2. A shift in the occupancy of the *rpsT* P2 promoter from RPo to earlier intermediates was detected by DNase I cleavage patterns with RNAP upon lowering temperature and/or addition of TraR (Gopalkrishnan et al., 2017; **Figure 6.1**). $KMnO_4$ footprinting (Ross and Gourse, 2009) suggested that under some conditions the transcription bubble was partially melted (**Figure 6.1**). The substitution of the nearly absolutely conserved nontemplate-strand (nt-strand) thymine at position -7 [T_{-7}(nt)] (Shultzaberger et al., 2007) to adenine [T_{-7}A(nt)], in combination with lower temperature, shifted the $KMnO_4$ signal upstream (**Figure 6.1**, lanes 6, 11). These perturbations (T_{-7}A(nt) and reduced temperature) in conjunction with TraR shift the RPo population to earlier DNA melting intermediates (**Figure 6.1**, lanes 7, 11).

In collaboration with P. D. B. Olinares from the Chait Lab (The Rockefeller University), the TraR-Eσ^{70}-*rpsT* P2 complexes were analyzed by nMS. Stable TraR-Eσ^{70}-*rpsT* P2 complexes were detected by nMS (**Figure 6.2A**). Interestingly, a higher percentage of Eσ^{70} was incorporated into complexes with *rpsT* P2 DNA with TraR than without (compare **Figures 6.2A** and **Figure 6.2B**). Additionally, a RNAP construct (β'Δ215-220-RNAP; Bartlett et al., 1998) that cannot maintain the transcription bubble on WT *rpsT* P2, can still form a partial bubble but only in the presence of TraR, as detected by KMnO4 footprinting [data courtesy of the Gourse Lab (University of Wisconsin–Madison); Chen et al., in press]. These results suggest that rather than populating an earlier intermediate indirectly by destabilizing RPo, TraR stabilizes the earlier intermediate directly (Chen et al., 2019b; Galburt, 2018).

6.3 Cryo-EM structures of TraR-Eσ^{70}-*rpsT* P2 delineate the promoter melting pathway

TraR inhibited transcription from the WT *rpsT* P2 promoter fragment under conditions used for cryo-EM (**Chapter 5**, **Figure 5.2**). I visualized the TraR-Eσ^{70} complexes with the WT *rpsT* P2 promoter by cryo-EM. Steps of maximum-likelihood classification in RELION (Bai et al., 2015b; Scheres, 2012) revealed five TraR-Eσ^{70}-DNA structures (T-RPc, T-RPi1, T-RPi2, T-preRPo, T-RPo; **Figure 6.3**) at 3.4 - 3.9 Å nominal resolution, with the central core of the structures resolved to 3.0 - 3.4 Å (**Figures 6.4A-E**, **Table 6.5**).

Additionally, an *rpsT* P2 promoter fragment with a two-nucleotide non-complementary bubble from -11 to -10 and with the T_{-7}A(nt) substitution (*rpsT* P2*; **Figure 6.6A**) was engineered to stabilize the intermediate detected by $KMnO_4$ footprinting (**Figure 6.1**, lane 7). Biochemical assays from S. Gopalkrishnan of the Gourse Lab (University of Wisconsin–Madison) show that basal transcription (without TraR) from *rpsT* P2* was very weak compared with WT *rpsT* P2 (**Figure 6.6B**), even though *rpsT* P2* had a pre-melted bubble. This result is consistent with the stabilization of an intermediate on the pathway to RPo formation, thereby depopulating RPo.

TraR-Eσ^{70}-*rpsT* P2* complexes were visualized by cryo-EM. Image classification gave rise to two distinct structures (T-RPi1.5a, T-RPi1.5b; **Figure 6.7**) at 3.5 and 3.0 Å nominal resolution, with the central core of the structures resolved to 3.0 and 2.6 Å, respectively (**Figures 6.8A-B**, **Table 6.9**). For structural analysis, I compared these intermediate structures with the 3.4

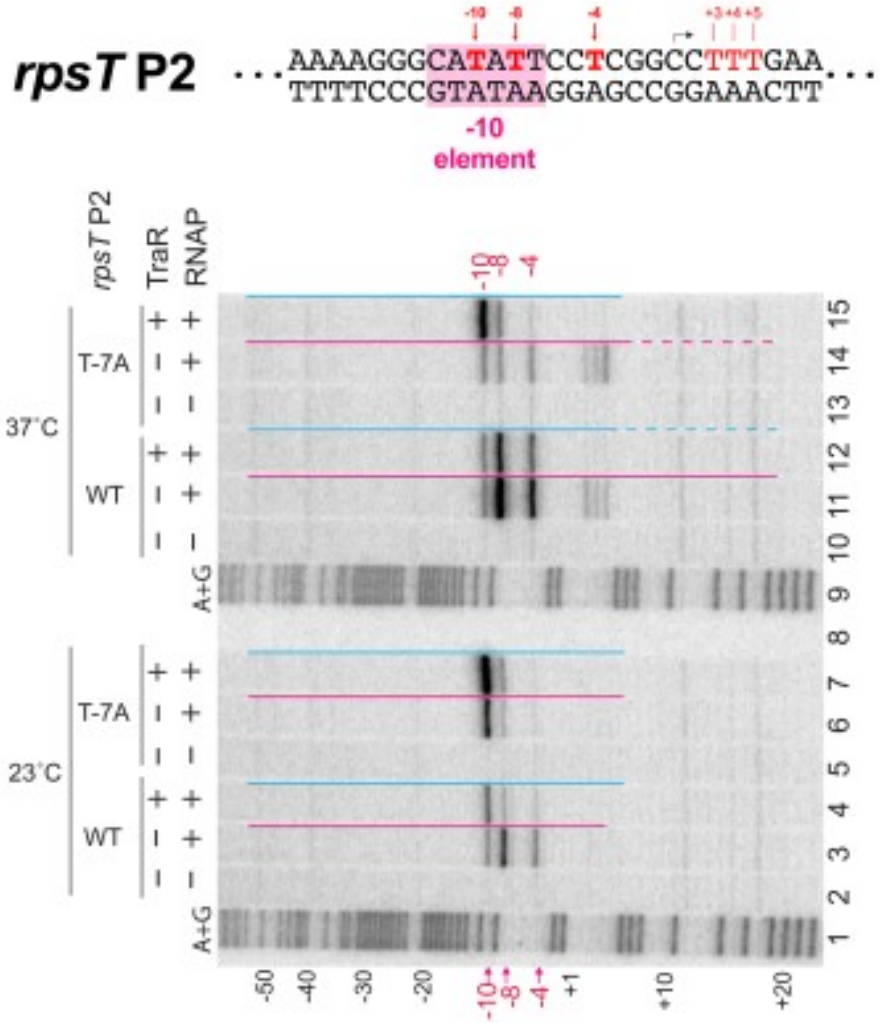


Figure 6.1. *Eco* TraR-Eσ^{70} forms stable, partially melted complexes with *rpsT* P2 in footprinting assays

Detection of unpaired thymines by $KMnO_4$ footprinting of Eσ^{70} complexes formed with the WT *rpsT* P2 or T-7A promoters ± TraR, and DNase I footprint protection ranges, shown by red or blue lines above each lane (dashed lines: partial protection). Strand cleavage of modified thymines at 23˚C (lanes 2-7) or 37˚ (lanes 10-15) was detected by gel electrophoresis of DNA fragments ^{32}P end-labeled in the nt-strand. Lanes 1, 9: A+G sequence ladder. Modified thymines at -10, -8 and -4 are indicated in red above and below gel, and on the section of the WT *rpsT* P2 sequence shown above the gel (-10 element shaded in pink). Black arrow: transcription start site (TSS). Footprinting data and figure courtesy of A. Y. Chen, S. Gopalkrishnan, M. F. Maloney, W. Ross and J. T. Winkelman from the Gourse Lab (University of Wisconsin-Madison).

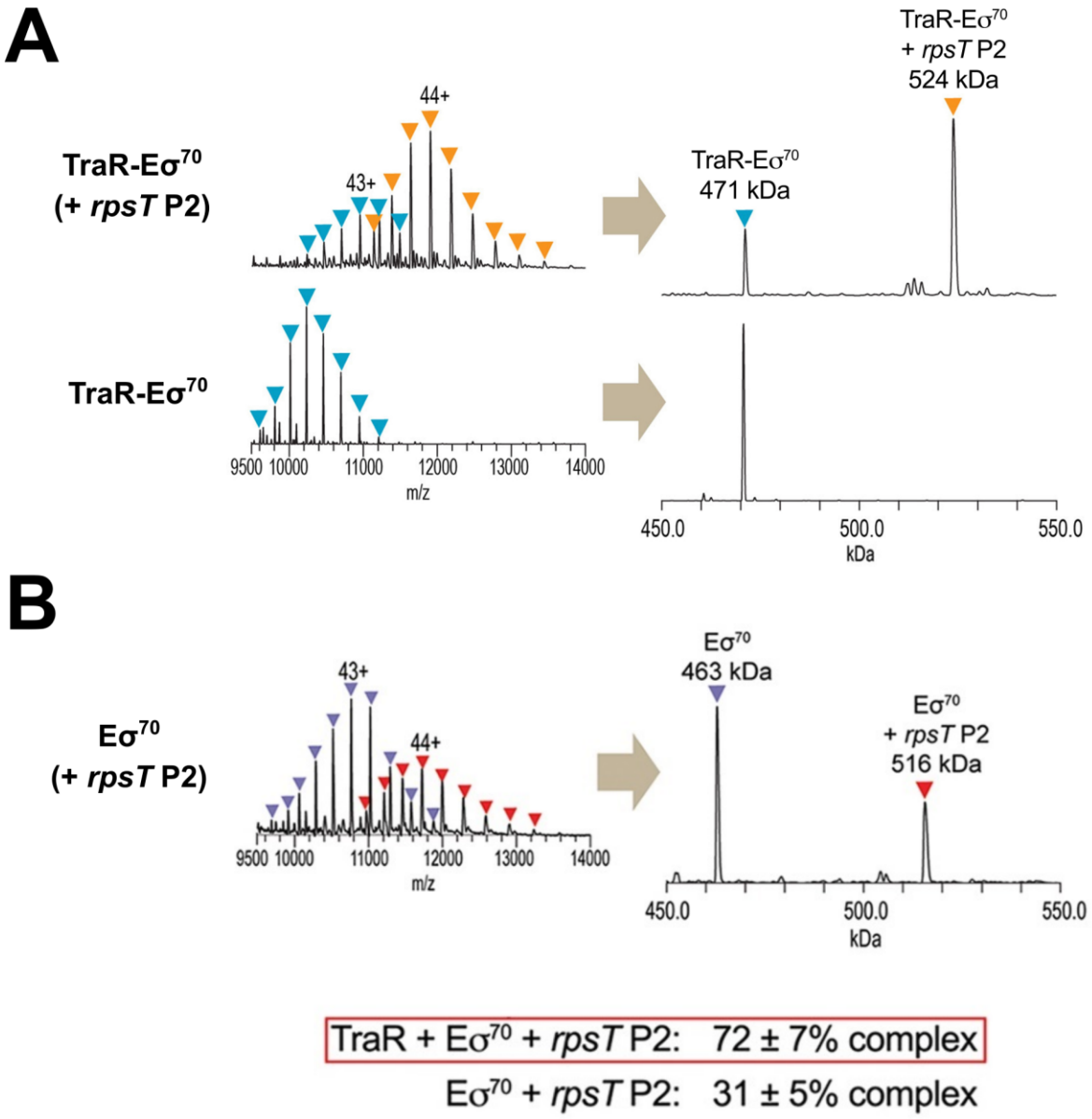


Figure 6.2. Detection of stable complexes of TraR-Eσ^{70} with *rpsT* P2 by nMS
(A) nMS spectra (*left*) and the corresponding deconvolved spectra (*right*) for TraR-Eσ^{70} complexes with the WT *rpsT* P2 promoter fragment (**Figure 5.10**). TraR binds to Eσ^{70} in a 1:1 stoichiometry, forming a 471 kDa complex. Upon incubation of this complex with the promoter DNA (52 kDa), a predominant charge state series for the TraR-Eσ^{70}-promoter assembly (524 kDa) was observed. **(B)** nMS spectrum (*left*) and the corresponding deconvolved spectrum (*right*) for Eσ^{70} incubated with the WT *rpsT* P2 promoter fragment (**Figure 5.10**). (*bottom*) Percentage of RNAP-*rpsT* P2 complex in each sample with and without TraR (calculated according to the peak intensities for each protein complex after deconvolution). Mass spectrometry data and figure courtesy of P. D. B. Olinares from the Chait Lab (The Rockefeller University).

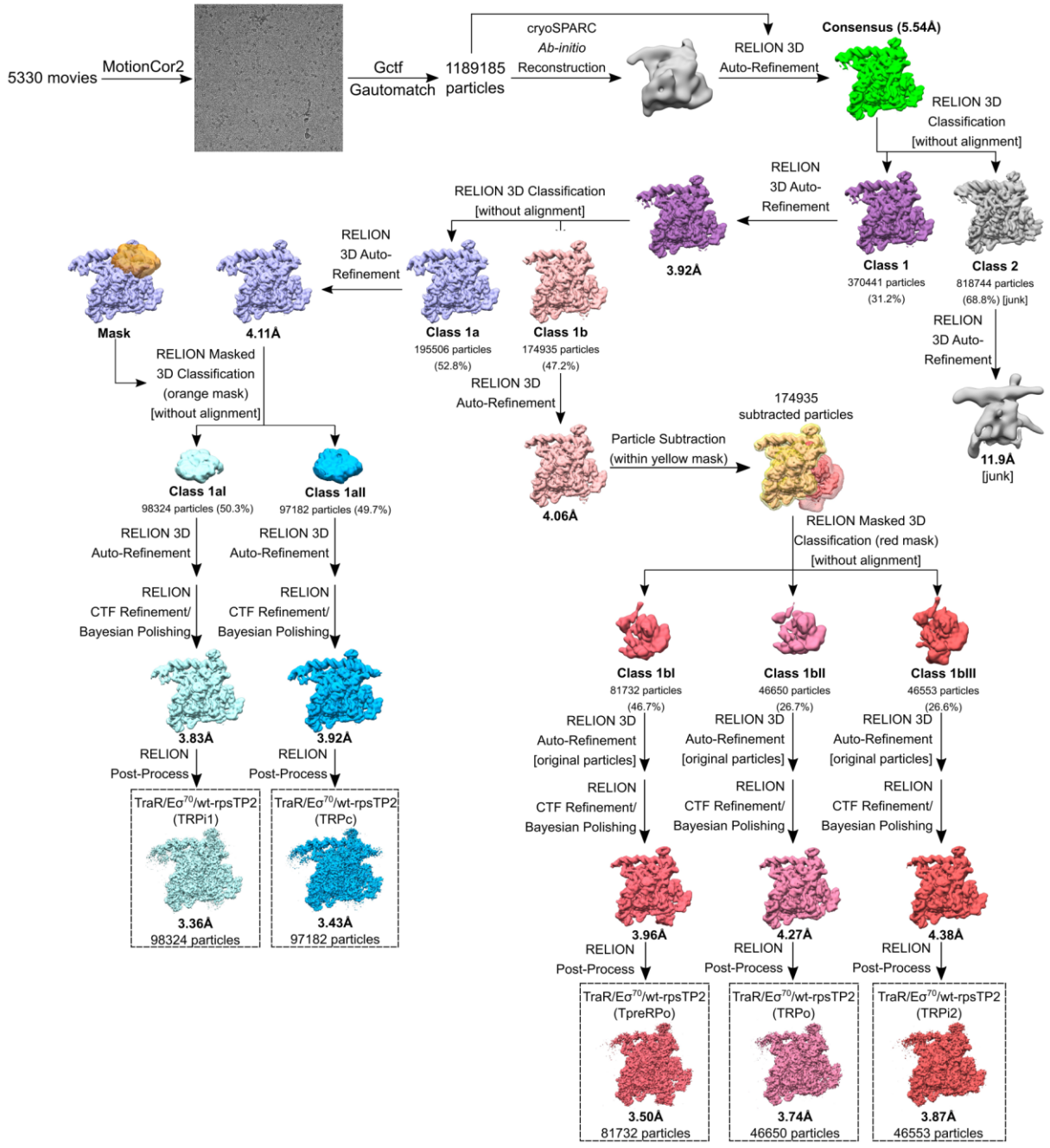


Figure 6.3. Cryo-EM processing pipeline for TraR-Eσ^{70}-WT-*rpsT* P2 complexes
Flowchart showing the image processing pipeline for the cryo-EM data of TraR-Eσ^{70}-WT-*rpsT*. After image classification, five structures were observed that are highlighted in black dashed boxes: TRPc, TRPi1, TRPi2, TpreRPo, and TRPo. After auto-refinement and post-processing in RELION (Zivanov et al., 2018): TRPc (blue) contains 97,182 particles at 3.43 Å resolution, TRPi1 (cyan) contains 98,324 particles at 3.36 Å resolution, TRPi2 (dark red) contains 46,553 particles at 3.87 Å resolution, TpreRPo (light red) contains 81,732 particles at 3.50 Å resolution, and TRPo (pink) contains 46,650 particles at 3.74 Å resolution.

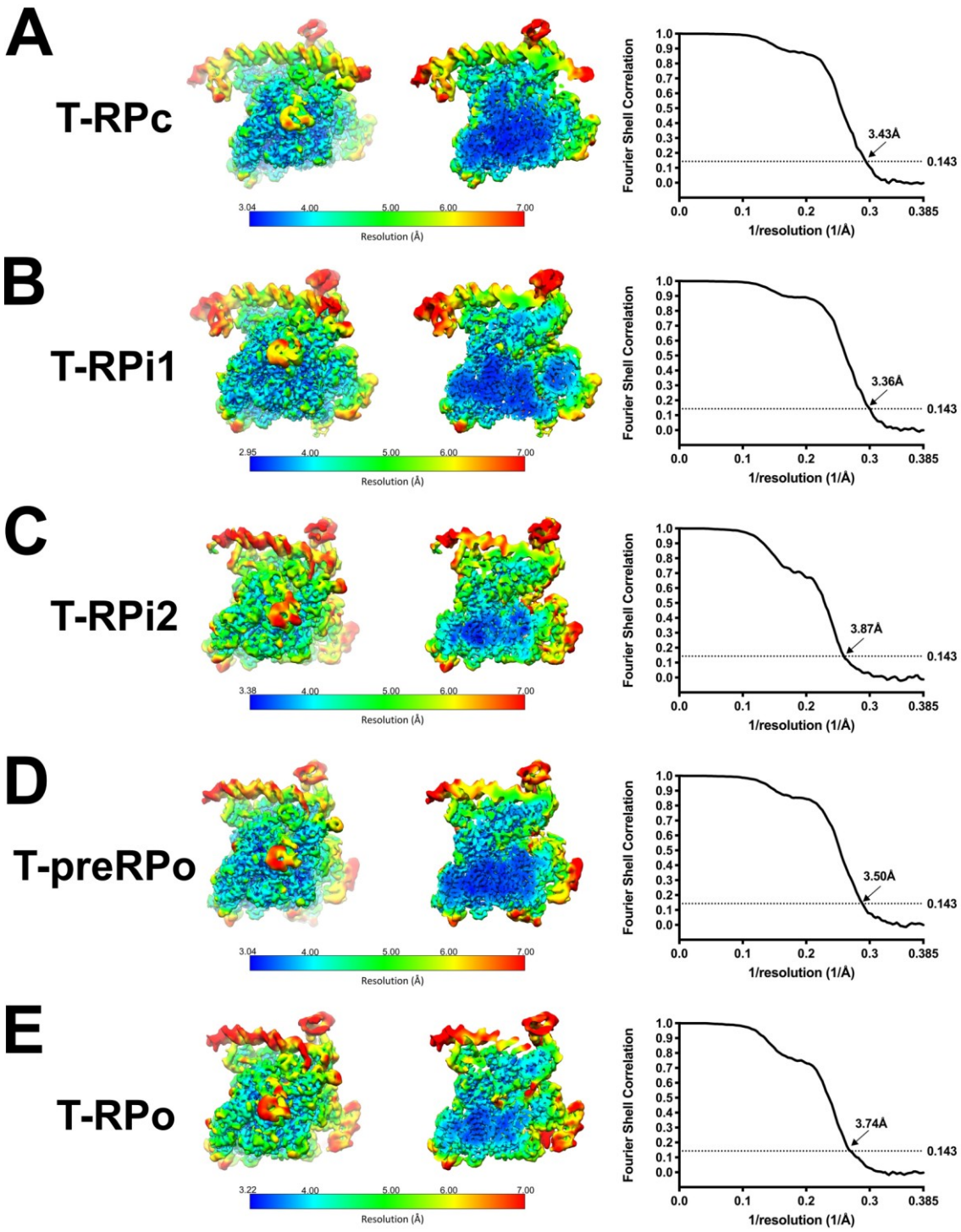


Figure 6.4. Resolution of TraR-Eσ^{70}-WT-*rpsT* P2 classes
(A-E) From left to right; i) class name, ii) two views of the cryo-EM density map colored by local resolution (Cardone et al., 2013). The right view is a cross-section through the left view, iii) gold-standard FSC, calculated by comparing the two independently determined half-maps from RELION (Zivanov et al., 2018). The dotted line represents the 0.143 FSC cutoff. **(A)** T-RPc. **(B)** T-RPi1. **(C)** T-RPi2. **(D)** T-preRPo. **(E)** T-RPo.

Table 6.5. Cryo-EM data collection, refinement and validation statistics for TraR-Eσ^{70}-WT-*rpsT* P2 complexes

	TraR-Eσ^{70}-WT-*rpsT* P2				
Data collection and processing					
Magnification	22,500				
Voltage (kV)	300				
Electron exposure (e–/Å^2)	47.3				
Defocus range (μm)	0.5 – 3.0				
Pixel size (Å)	1.3				
Symmetry imposed	C1				
Initial particle images (no.)	1,189,185				
	T-RPc	T-RPi1	T-RPi2	T-preRPo	T-RPo
Final particle images (no.)	97,182	98,324	46,553	81,732	46,650
Map resolution (Å) FSC threshold 0.143	3.4	3.4	3.9	3.5	3.7
Map resolution range (Å)	3.0 - 7.0	3.0 - 7.0	3.4 - 8.0	3.0 - 7.0	3.2 - 7.0
Refinement					
Initial model used (PDB code) Model resolution (Å) FSC threshold 0.5 Model resolution range (Å)	6N57 4.0 3.1 - 7.0	6N57 4.0 3.1 - 7.0	6N57 4.0 3.1 - 7.0	6N57 4.0 3.1 - 7.0	6N57 4.0 3.1 - 7.0
Map sharpening *B* factor (Å^2)	-58	-59	-80	-76	-71
Model composition Non-hydrogen atoms Protein residues Nucleic acid residues Ligands	 32,787 3,904 100 8 (1 Mg^{2+}, 3 Zn^{2+}, 4 CHAPSO)	 32,196 3,902 71 8 (1 Mg^{2+}, 3 Zn^{2+}, 4 CHAPSO)	 32,197 3,877 79 8 (1 Mg^{2+}, 3 Zn^{2+}, 4 CHAPSO)	 32,456 3,830 105 8 (1 Mg^{2+}, 3 Zn^{2+}, 4 CHAPSO)	 32,703 3,831 124 6 (1 Mg^{2+}, 3 Zn^{2+}, 2 CHAPSO)
B factors (Å^2) Protein Nucleic acid Ligands	 48.23 171.3 46.56	 41.61 247.7 41.34	 30.14 128.9 29.48	 38.56 215.7 31.4	 41.45 139.7 53.66

R.m.s. deviations					
Bond lengths (Å)	0.009	0.008	0.012	0.008	0.007
Bond angles (°)	0.878	0.885	1.221	1.896	0.879
Validation					
MolProbity score	2.91	2.84	2.49	2.8	2.11
Clashscore	10.72	9.49	16.09	9.14	9.92
Poor rotamers (%)	11.43	10.54	1.45	9.85	0.34
Ramachandran plot[a]					
Favored (%)	81.2	81.7	77.7	81.0	80.5
Allowed (%)	18.5	18.3	22.3	19.0	19.5
Disallowed (%)	0	0	0	0	0

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

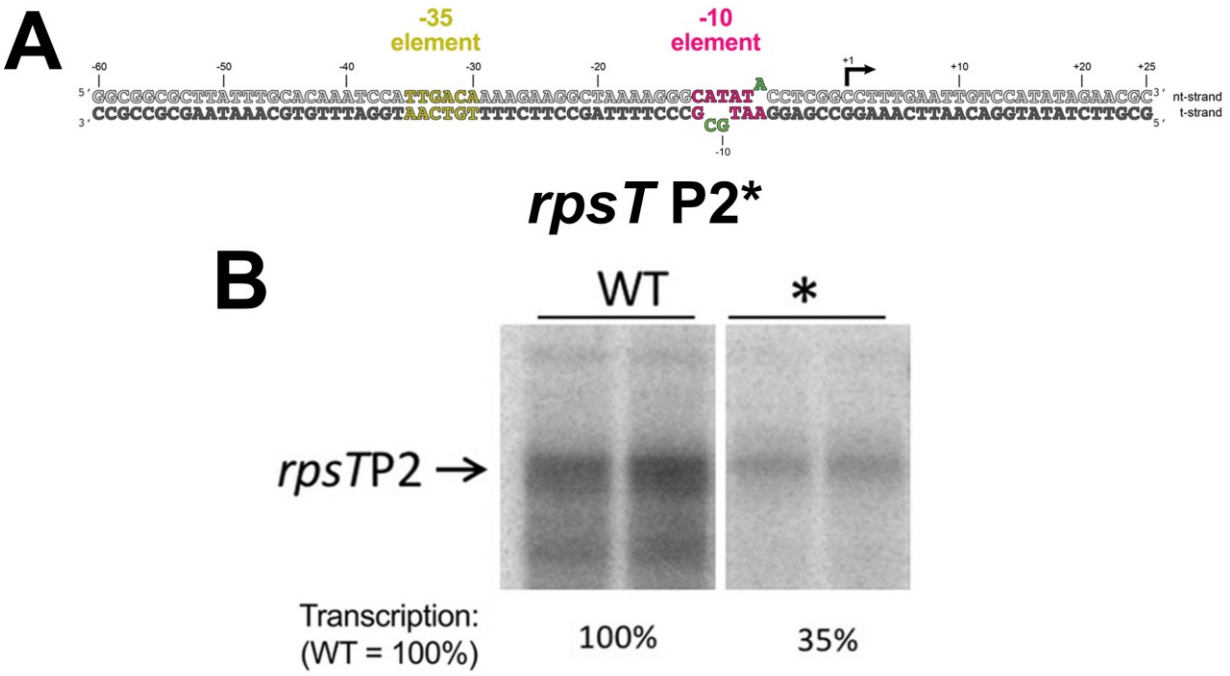


Figure 6.6. Modified *Eco rpsT* P2 promoter fragment (*rpsT* P2*) for cryo-EM
(A) Linear fragment of modified *Eco rpsT* P2 promoter (*rpsT* P2*) used for cryo-EM studies. The nt-strand is colored light grey, t-strand is colored dark grey, -35 element is colored yellow, -10 is colored magenta, and nucleotide substitutions in *rpsT* P2* are colored green. Numbering above the sequence indicates position with respect to the +1 TSS (shown as an arrow pointing downstream). **(B)** Multi-round transcription assays on WT *rpsT* P2 (WT) and *rpsT* P2* (*) fragments. Transcription assay data courtesy of S. Gopalkrishnan from the Gourse Lab (University of Wisconsin–Madison).

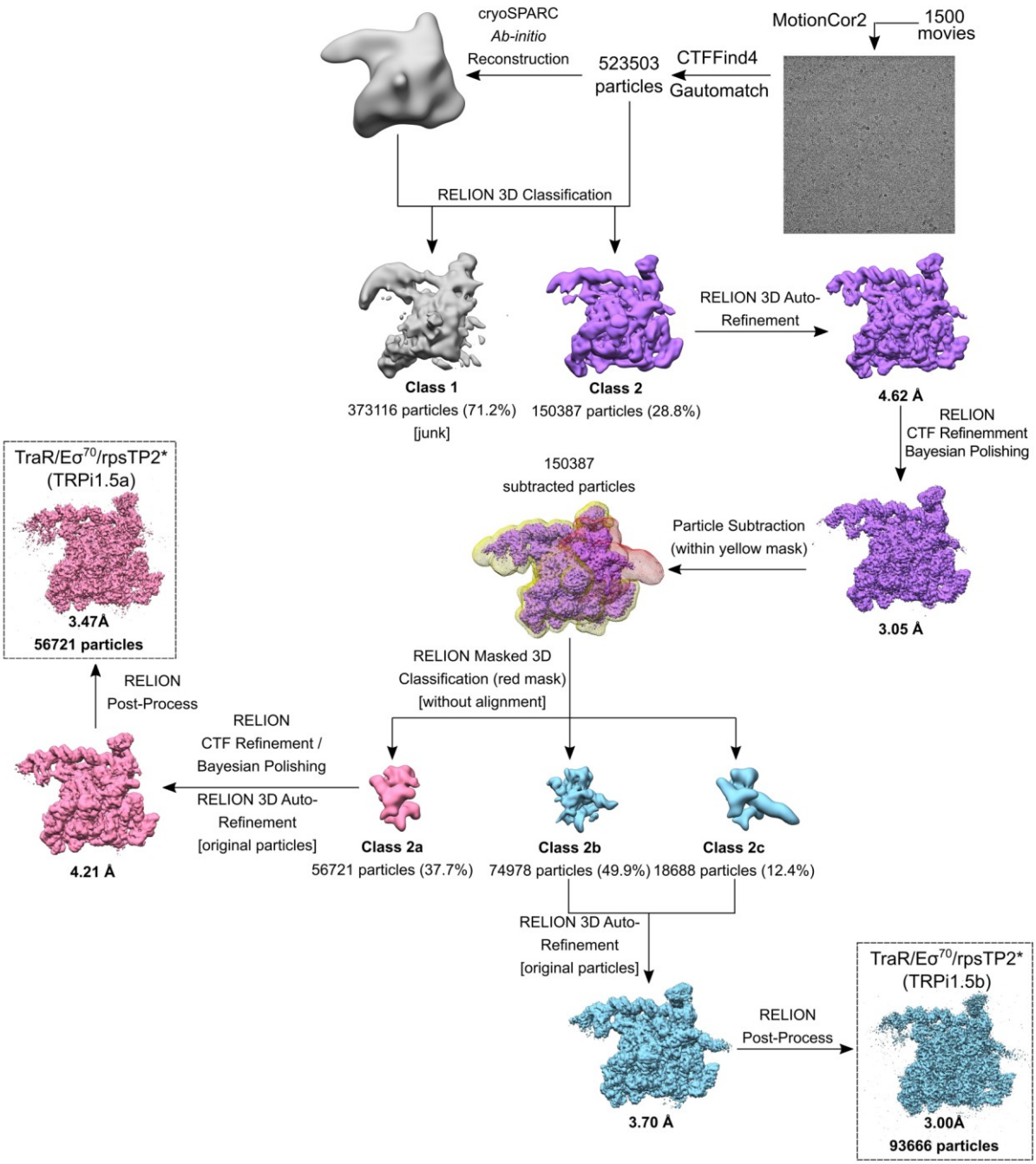


Figure 6.7. Cryo-EM of TraR-Eσ^{70}-*rpsT* P2* complexes
Flowchart showing the image processing pipeline for the cryo-EM data of TraR-Eσ^{70}-*rpsT* P2*. After image classification, two structures were observed that are highlighted in black dashed boxes: TRPi1.5a and TRPi1.5b. After auto-refinement and post-processing in RELION (Zivanov et al., 2018): TRPi15.a (red) contains 56,721 particles at 3.47 Å resolution and TRPi1.5b (blue) contains 93,666 particles at 3.00 Å resolution.

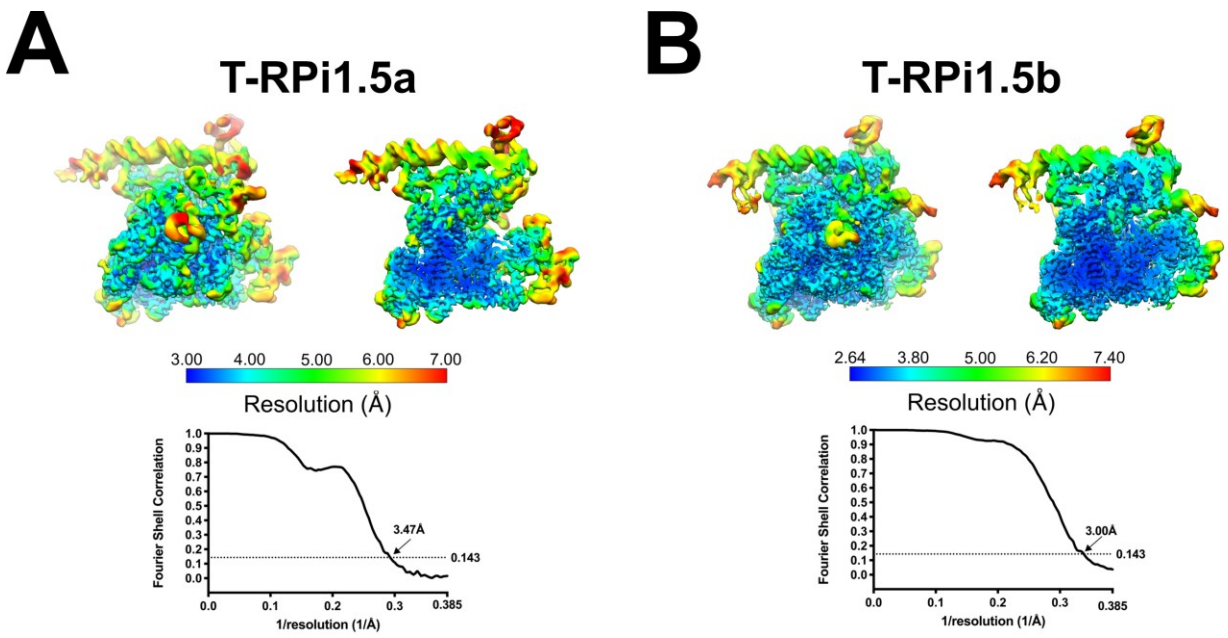


Figure 6.8. Resolution of TraR-Eσ^{70}-*rpsT* P2* classes
(A-B) (*top*) Class name. (*middle*) two views of the cryo-EM density map colored by local resolution (Cardone et al., 2013). The right view is a cross-section through the left view. (*bottom*) Gold-standard FSC, calculated by comparing the two independently determined half-maps from RELION (Zivanov et al., 2018). The dotted line represents the 0.143 FSC cutoff. **(A)** TRPi1.5a. **(B)** TRPi1.5b.

Table 6.9. Cryo-EM data collection, refinement and validation statistics for TraR-Eσ^{70}-*rpsT* P2* complexes

	TraR- Eσ^{70}-*rpsT* P2*	
Data collection and processing		
Magnification	22,500	
Voltage (kV)	300	
Electron exposure (e–/Å^2)	50	
Defocus range (μm)	0.5 – 2.0	
Pixel size (Å)	1.3	
Symmetry imposed	C1	
Initial particle images (no.)	523,503	
	TRPi1.5a	TRPi1.5b
Final particle images (no.)	56,721	93,666
Map resolution (Å) FSC threshold 0.143	3.5	3.0
Map resolution range (Å)	3.0 - 7.0	2.6 - 7.4
Refinement		
Initial model used (PDB code) Model resolution (Å) FSC threshold 0.5 Model resolution range (Å)	6N57 4.0 3.1 - 7.0	6N57 4.0 3.1 - 7.0
Map sharpening *B* factor (Å^2)	-63	-33
Model composition Non-hydrogen atoms Protein residues Nucleic acid residues Ligands	 31,036 3,726 86 4 (1 Mg^{2+}, 3 Zn^{2+})	 31,606 3,726 101 8 (1 Mg^{2+}, 3 Zn^{2+}, 4 CHAPSO)
B factors (Å^2) Protein Nucleic acid Ligands	 47.42 167.52 49.66	 47.51 144.73 56.45
R.m.s. deviations Bond lengths (Å)	 0.008	 0.01

Bond angles (°)	1.069	0.851
Validation		
MolProbity score	2.73	2.85
Clashscore	9.12	9.65
Poor rotamers (%)	10.99	13.10
Ramachandran plot[a]		
Favored (%)	83.8	84.9
Allowed (%)	16.2	15.1
Disallowed (%)	0	0

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

Å structure of an RPo complex between $E\sigma^{70}$ and the WT *rpsT* P2 promoter fragment prepared in the absence of TraR (**Chapter 5**, **Figure 5.14**; Chen et al., 2019b).

The eight complexes observed by cryo-EM with WT *rpsT* P2 and *rpsT* P2* were ordered in the pathway (**Figure 6.10A**) such that the DNA/σ^{70} interface area, the downstream boundary of the DNA/RNAP contacts, and the extent of the transcription bubble monotonically increased, while the root-mean-square deviation (RMSD) of alpha carbons (Cα) positions of each complex compared to RPo decreased monotonically, with progress along the pathway (**Figure 6.10B**). A clear demarcation between early and late complexes could be made based on the presence of $\sigma^{70}{}_{1.1}$ (early complexes) or downstream duplex DNA (late complexes) in the RNAP channel (Bae et al., 2013; Mekler et al., 2002).

In all eight structures, $E\sigma^{70}$ interacts with the upstream promoter DNA (from -43 to -17) in the same manner: i) $\sigma^{70}{}_4$ engages specifically with the major groove of the promoter -35 element from -37 to -30 (Campbell et al., 2002), ii) C-terminal domain of the α subunits (αCTD) binds just upstream of $\sigma^{70}{}_4$, interacting with the DNA minor groove from -43 to -38 (Benoff et al., 2002; Ross et al., 1993, 2001) and interfaces with $\sigma^{70}{}_4$ (Ross et al., 2003), and iii) conserved residues of the β'zipper (β'Y46 and R47) interact with the DNA backbone from -18 to -17 (Bae et al., 2015a; Yuzenkova et al., 2011). By contrast, $E\sigma^{70}$ interacts with the promoter DNA downstream of -17 in diverse configurations that our laboratory propose represent steps on the RPo formation pathway (**Figure 6.10A**).

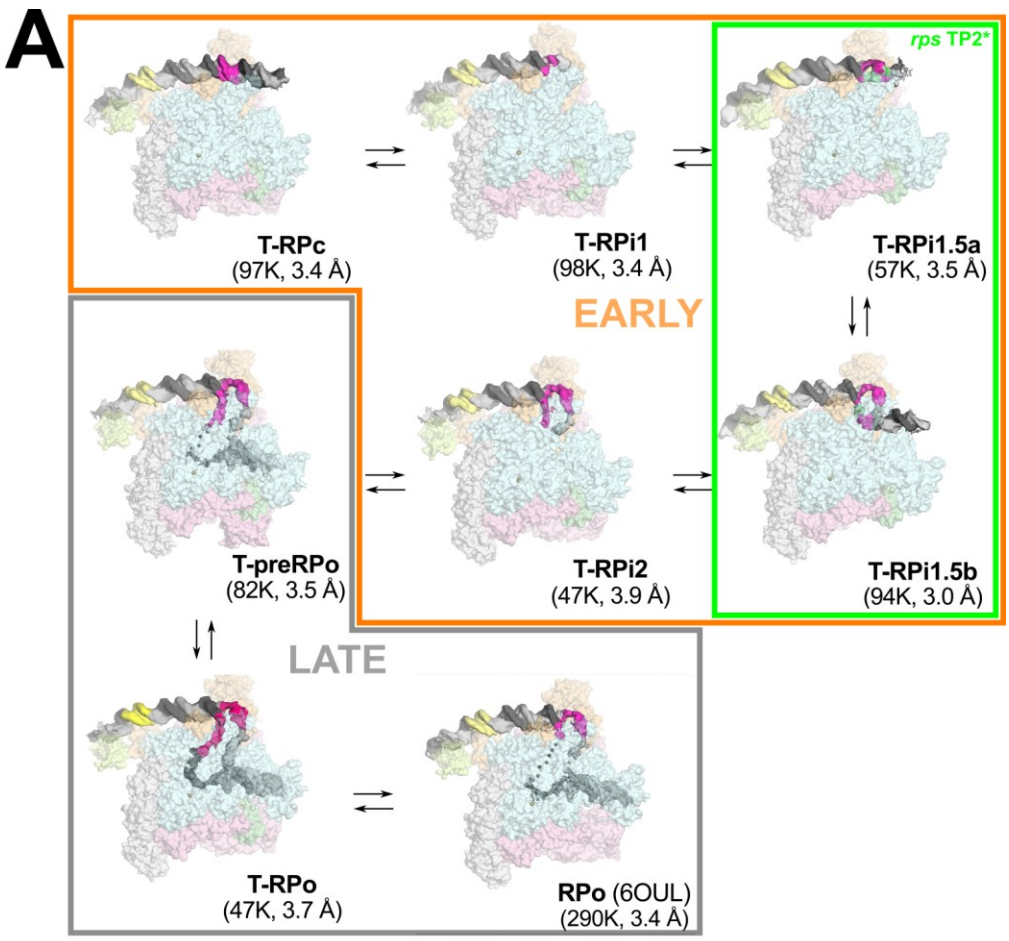
6.4 T-RPc: Structure of a closed complex

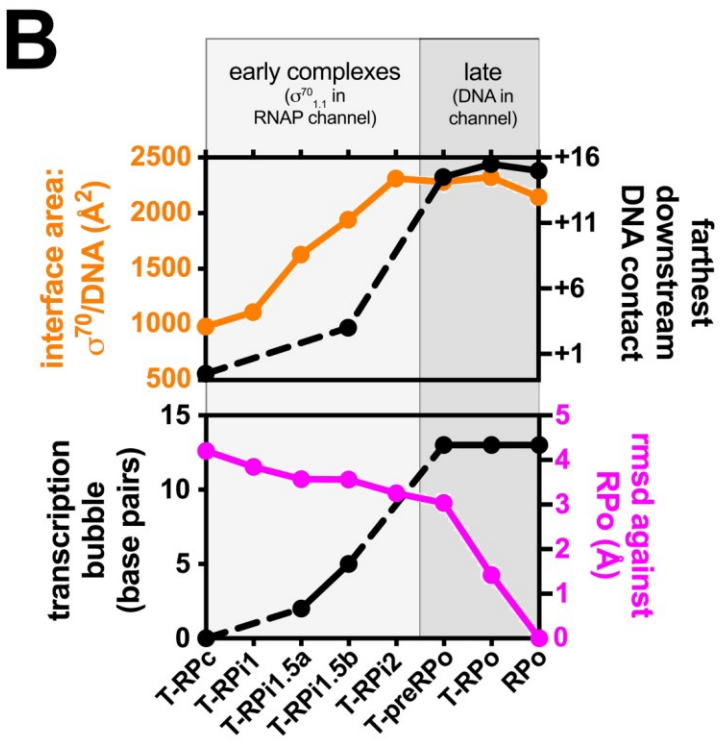
Initial recognition of the duplex promoter sequence prior to melting is thought to give rise to the closed complex, or RPc (Ruff et al., 2015b). DNase I or hydroxyl-radical footprinting of RPc revealed an upstream footprinting pattern similar to RPo but with much less downstream protection (Kovacic, 1987; Schickor et al., 1990), indicating that the duplex DNA downstream of the -10 element was mostly solvent-exposed. The earliest complex in the pathway (T-RPc, **Figures 6.10A**, **6.11**), which contains entirely duplex DNA, forms $E\sigma^{70}$-promoter interactions consistent with these earlier footprinting results.

In T-RPc, the αCTD proximal to $\sigma^{70}{}_4$ and the distal αCTD are bound to the upstream DNA (**Figure 6.11A**). Focused classification of this region upstream of the -35 element revealed two dispositions of the αCTDs, head-to-head (about 53% of the particles) and head-to-tail (about 47%), with altered upstream DNA trajectory (**Figure 6.11A**; compare box iii and iv). The two αCTDs play diverse roles in transcription initiation, binding to DNA upstream (UP) of the promoter -35 element (with a preference for AT-rich UP elements), as well as establishing protein-protein interactions with $\sigma^{70}{}_4$ and/or transcription factors with variably positioned upstream DNA recognition sites (Benoff et al., 2002; Estrem et al., 1998; Lee et al., 2012; Ross and Gourse, 2005; Ross et al., 1993, 2003).

The promoter DNA is completely duplex in T-RPc, thus T-RPc precedes the nucleation of transcription bubble melting. Sequence-specific recognition of the -35 element by $\sigma^{70}{}_4$ fixes the register of the DNA with respect to $E\sigma^{70}$. Base-specific protein-DNA interactions do not occur within the -10 element (**Figures 6.11B**), confirming the conclusion of (Feklistov and Darst, 2011) that recognition of the -10 element sequence is coupled with melting. The duplex -10 element DNA is drawn to a shallow, basic channel on the $E\sigma^{70}$ surface (**Figure 6.11C**) by phosphate backbone interactions with invariant basic residues of $\sigma^{70}{}_2$ (R436, R441, R451) and $\sigma^{70}{}_3$ (K462, R465) (**Figure 6.11B**). Recognition of the -35 and these phosphate backbone interactions positions the critical A_{-11}(nt) in line with the $\sigma^{70}{}_2$ residues that ultimately capture the

Figure 6.10. *Eco* Eσ^{70} promoter melting intermediates on the *rpsT* P2 promoter
(A) Overall structures of promoter melting intermediates obtained by cryo-EM. Proteins are shown as transparent surfaces (αI, αII, ω, light gray; αCTD, pale limon; β, pale cyan; β', light pink; σ^{70}, light orange; TraR, pale green). The Eσ^{70} active site Mg^{2+} is shown as a sand-colored sphere. The promoter DNA is shown as cryo-EM difference density (nt-strand, gray; t-strand, dark gray; -35 element, yellow; -10 element, hot pink). The eight structures were derived from three samples: Sample 1) T-RPc, T-RPi1, T-RPi2, T-preRPo, and T-RPo structures were obtained with TraR and the WT *rpsT* P2 fragment **(Figure 5.10)**; Sample 2) T-RPi1.5a and T-RPi1.5b were obtained with TraR and *rpsT* P2* **(Figure 6.6)** and are boxed in green; Sample 3) RPo was determined previously with WT *rpsT* P2 without TraR (**Figure 5.14**, Chen et al., 2019). In the Early complexes (boxed in orange), $\sigma^{70}_{1.1}$ occupies the Eσ^{70} channel. In the Late complexes (boxed in gray), downstream duplex DNA occupies the channel. **(B)** Structural properties used to order the complexes in the RPo formation pathway. (*top panel*) Plotted in orange (left scale) is the σ^{70}/DNA interface area (Å^2) (Krissinel and Henrick, 2007). Plotted in black (right scale) is the most downstream protein/duplex DNA contact. For T-RPi1, T-RPi1.5a, and T-RPi2, most or all of the downstream duplex DNA was disordered so no point is included. (*bottom panel*) Plotted in black (left scale) is the extent of the transcription bubble. For T-RPi1 and T-RPi2, the downstream fork of the transcription bubble was disordered so no point is included. Plotted in magenta (right scale) is the RMSD of Cα (Å) for each complex superimposed with RPo.


A
rps TP2*
T-RPc
(97K, 3.4 Å)
T-RPi1
(98K, 3.4 Å)
T-RPi1.5a
(57K, 3.5 Å)
EARLY
T-preRPo
(82K, 3.5 Å)
T-RPi2
(47K, 3.9 Å)
T-RPi1.5b
(94K, 3.0 Å)
LATE
T-RPo
(47K, 3.7 Å)
RPo (6OUL)
(290K, 3.4 Å)


B
early complexes
(σ70 1.1 in
RNAP channel)
late
(DNA in
channel)
interface area:
σ70/DNA (Å2)
2500
2000
1500
1000
500
+16
+11
+6
+1
farthest
downstream
DNA contact
transcription
bubble
(base pairs)
15
10
5
0
5
4
3
2
1
0
rmsd against
RPo (Å)
T-RPc
T-RPi1
T-RPi1.5a
T-RPi1.5b
T-RPi2
T-preRPo
T-RPo
RPo

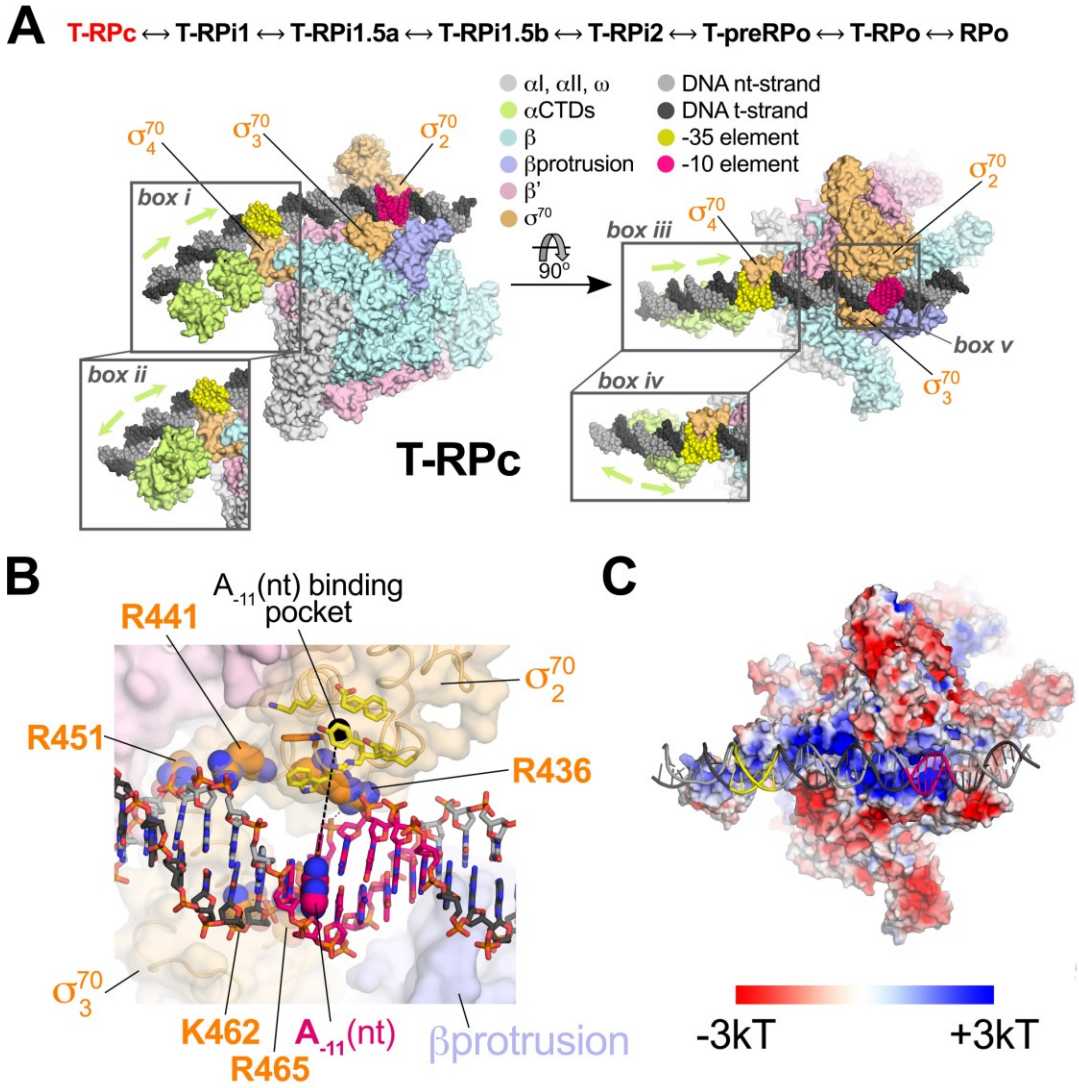


Figure 6.11. T-RPc: Structure of the TraR-Eσ^{70} closed promoter complex
(*top*) The structures determined here are ordered through the RPo formation pathway **(Figure 6.10)**. T-RPc, highlighted in red, is the focus of this figure. **(A-B)** Color-coding is shown in the key. **(A)** Orthogonal views of T-RPc. Proteins are shown as molecular surfaces, DNA is shown as atomic spheres. The proximal (adjacent to σ^{70}_4) and distal (further upstream) αCTDs were visualized in two co-existing dispositions on the DNA upstream of the -35 element, head-to-tail (*box i and iii*) and head-to-head (*box ii* and *iv*). The region around the duplex -10 element (*box v*) is magnified in **(B)**. **(B)** Magnified view of Eσ^{70} interactions with the duplex -10 element showing the absence of sequence-specific interactions (Feklistov and Darst, 2011). The DNA is shown as sticks with the A_{-11}(nt) base highlighted in spheres, and the location of the cognate binding pocket in σ^{70}_2 (yellow side chains) occupied by A_{-11}(nt) in subsequent intermediates indicated by a dashed black line connecting A_{-11}(nt) to the pocket. RNAP is shown as a transparent molecular surface. The side chains shown as spheres (σ^{70}_2 R436, R441, R451; σ^{70}_3 K462, R465), absolutely conserved among primary σ's (Gruber and Bryant, 1997), interact with the duplex DNA phosphate backbone. **(C)** The electrostatic charge distribution (Baker et al., 2001) is shown on the molecular surface of the T-RPc RNAP [same view as the right view of **(A)**]. The DNA is shown in cartoon format.

flipped out base in RPo (yellow residues in **Figure 6.11B**). A_{-11}(nt) is the most highly conserved base in Eσ^{70} promoters (Shultzaberger et al., 2007) and is likely the first base to flip out of the duplex DNA to nucleate transcription bubble formation (Chen and Helmann, 1997; Feklistov and Darst, 2011; Feklistov et al., 2006; Lim et al., 2001).

6.5 T-RPc ↔ T-RPi1 ↔ T-RPi1.5a: Transcription bubble nucleation and the σ^{70} W-dyad

The key event in the nucleation of promoter melting is thought to be flipping of the A_{-11}(nt) base from the duplex DNA into its σ^{70}_2 pocket (Chen and Helmann, 1997; Feklistov and Darst, 2011; Heyduk et al., 2006; Lim et al., 2001) and isomerization of the invariant W-dyad of σ^{70}_2 (W433/W434 in *Eco*) from an 'edge-on' to a 'chair'-like conformation. In the chair conformation, the W433 side chain rotates away from W434, fills the space vacated by the flipped-out A_{-11}(nt), and forms a π-stack with the face of the exposed -12(nt) base (Bae et al., 2015a).

In our pathway (**Figure 6.10A**), T-RPc contains duplex DNA and the W-dyad is in the 'edge-on' conformation (**Figures 6.12A-B**). The next intermediate, T-RPi1, yields a view of transcription bubble nucleation, which begins before DNA enters into the RNAP cleft. In T-RPi1, the A_{-11}(nt) base is flipped and entering its cognate σ^{70}_2 pocket (**Figure 6.12C**). Notably, the W-dyad remains in its 'edge-on' conformation (**Figure 6.12D**). The 'edge-on' orientation of the W433 side chain in T-RPi1 sterically clashes with the -12 base pair, and the cryo-EM density indicates transient melting of the -12 base pair in this intermediate (**Figures 6.12C-D**). In all subsequent structures in the pathway (T-RPi1.5a -> RPo), the flipped-out A_{-11}(nt) base is fully engaged in its pocket, the -12 nucleotides are clearly base-paired, and the -12(nt) base is stacked with W433 in the 'chair'-like conformation (**Figures 6.12E-F**).

6.6 T-RPi1.5a ↔ T-RPi1.5b: Transcription bubble propagation and the protrusion pocket

In T-RPi1.5a (obtained with *rpsT* P2*), the nascent transcription bubble is only two nucleotides (the engineered bubble from -11 to -10), and density for about seven base pairs of downstream duplex DNA is interpretable (**Figure 6.13A**). Further along the pathway, in the transition from T-RPi1.5a to T-RPi1.5b, the transcription bubble extends to 5 nucleotides (-11 to -7), the flipped-out A_{-11}(nt) completely engages its cognate σ^{70}_2 pocket, and nt-strand phosphate-backbone interactions from -10 to -8 with σ^{70}_2 are established as in RPo (**Figure 6.13B**). Presumably because of the $T_{-7}A$(nt) substitution in *rpsT* P2* (**Figure 6.6A**), the mutant A_{-7}(nt) base is not engaged in the σ^{70}_2 pocket normally occupied by the conserved T_{-7}(nt) and the entire -7(nt) nucleotide is disordered (**Figure 6.13B**).

In all of the structures with TraR bound (except for T-RPo), TraR induces an ~18° rotation of the βlobe-Si1 towards TraR, widening the gap between the βlobe and the βprotrusion (Chen et al., 2019b) (**Figures 6.13A-B**). This rotation alters a number of βlobe-$\sigma^{70}_{1.1}$ contacts (Chen et al., 2019b), but interactions between the βlobe-Gate-Loop (GL; β residues 371-376) and $\sigma^{70}_{1.1}$ that creates a barrier for DNA entry into the RNAP cleft are maintained. The GL-barrier and the widened gap between the βlobe and the βprotrusion create a channel that can accommodate the downstream duplex DNA in T-RPi1.5b (**Figure 6.13B**).

Because of the short transcription bubble and the novel disposition of the downstream duplex DNA, the single-stranded t-strand DNA in T-RPi1.5b follows a unique path between σ^{70}_2 and the βprotrusion (**Figure 6.13C**). This path of the DNA allows the -9 t-strand base [T_{-9}(t)] to flip towards the βprotrusion and bind in a distinct pocket in the underside of the βprotrusion that has not been described previously (referred to here as the protrusion pocket; **Figure 6.13C**). The residues that form the protrusion-pocket and interact with T_{-9}(t) (**Figure 6.13C**) are conserved

Figure 6.12. T-RPc ↔ T-RPi1 ↔ T-RPi1.5a transitions
(*top*) The order of structures through the RPo formation pathway **(Figure 6.10)**. The progression from T-RPc ↔ T-RPi1 ↔ T-RPi1.5a, highlighted in red, is the focus of this figure. **(A, C, E)** (*left*) Overall view of T-RPc **(A)**, T-RPi1 **(C)**, and T-RPi1.5a **(E)**. Eσ^{70} is shown as a molecular surface with promoter DNA in cartoon format (color-coded based on key). The σ^{70} W-dyad is colored dark orange. The boxed region is magnified on the right. (*right*) Magnified view of promoter -10 element and W-dyad. Promoter DNA is shown in stick format with the A_{-11}(nt) base highlighted with spheres. σ^{70} is shown as a backbone worm (pale orange) but with side chains of residues that interact with A_{-11}(nt) in RPo shown (orange). The W-dyad is also shown, highlighted with transparent CPK spheres. **(B, D, F)** Cryo-EM maps, filtered according to the local resolution (Cardone et al., 2013), are shown as blue meshes. The final refined models are superimposed, colored according to the key below. The DNA positions are indicated **(A-B)** T-RPc: The -10 element is completely duplex and the W-dyad is in the edge-on conformation. **(C-D)** T-RPi1 and transcription bubble nucleation: A_{-11}(nt) is flipped out of the duplex towards its cognate $\sigma^{70}{}_2$ pocket, nucleating -10 element melting. Steric clash with the 'edge-on' conformation of the W-dyad disrupts the -12 base pair. Downstream DNA lacks cryo-EM density and is presumed to be highly dynamic. **(E-F)** T-RPi1.5a: The flipped out A_{-11}(nt) more fully engages with its cognate $\sigma^{70}{}_2$ pocket. W-dyad is modeled in its 'chair' conformation (Bae et al., 2015a), allowing the -12 base pair to reform. The T-RPi1.5a structure was obtained with the mutant *rpsT* P2* promoter (base substitutions colored green).

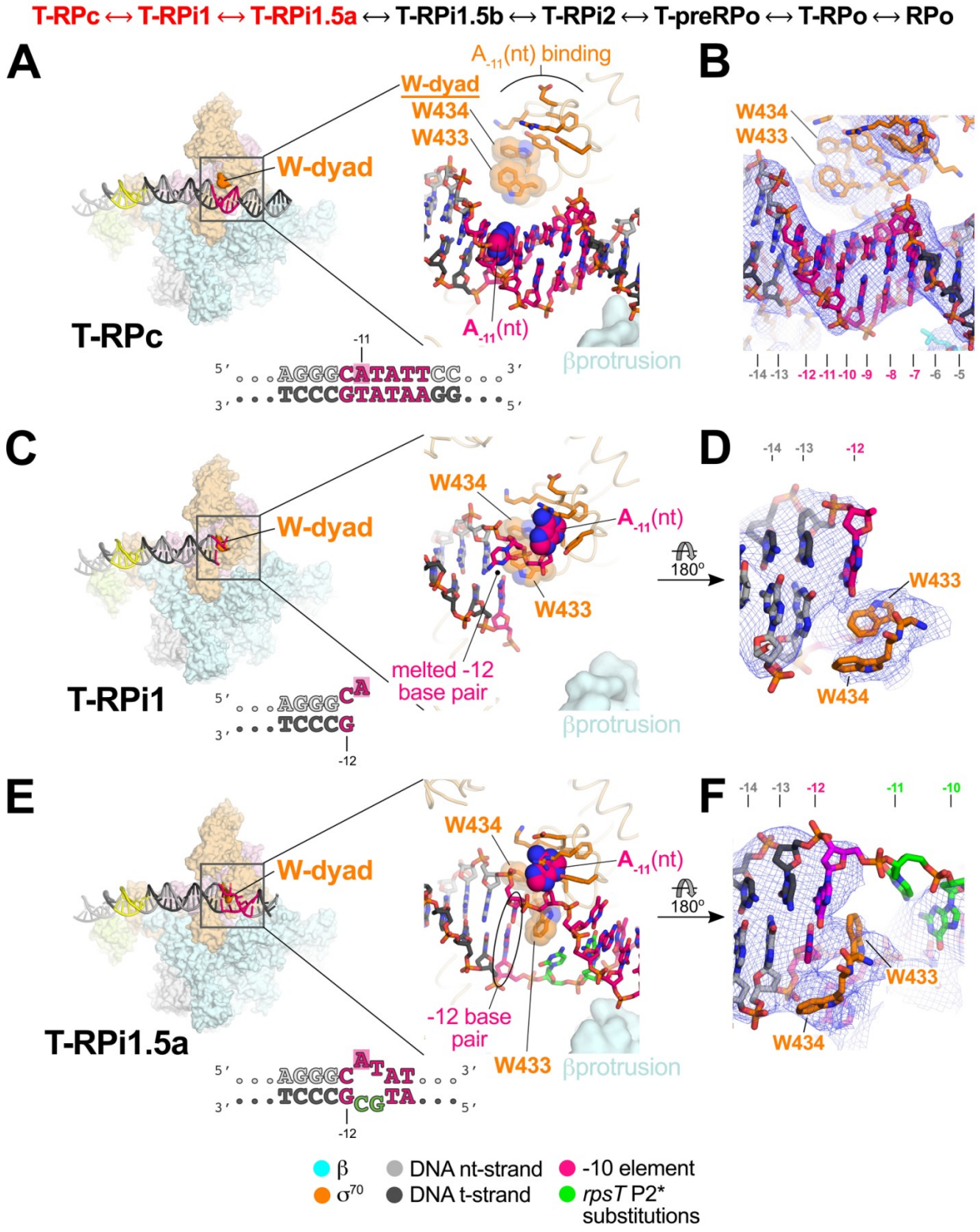
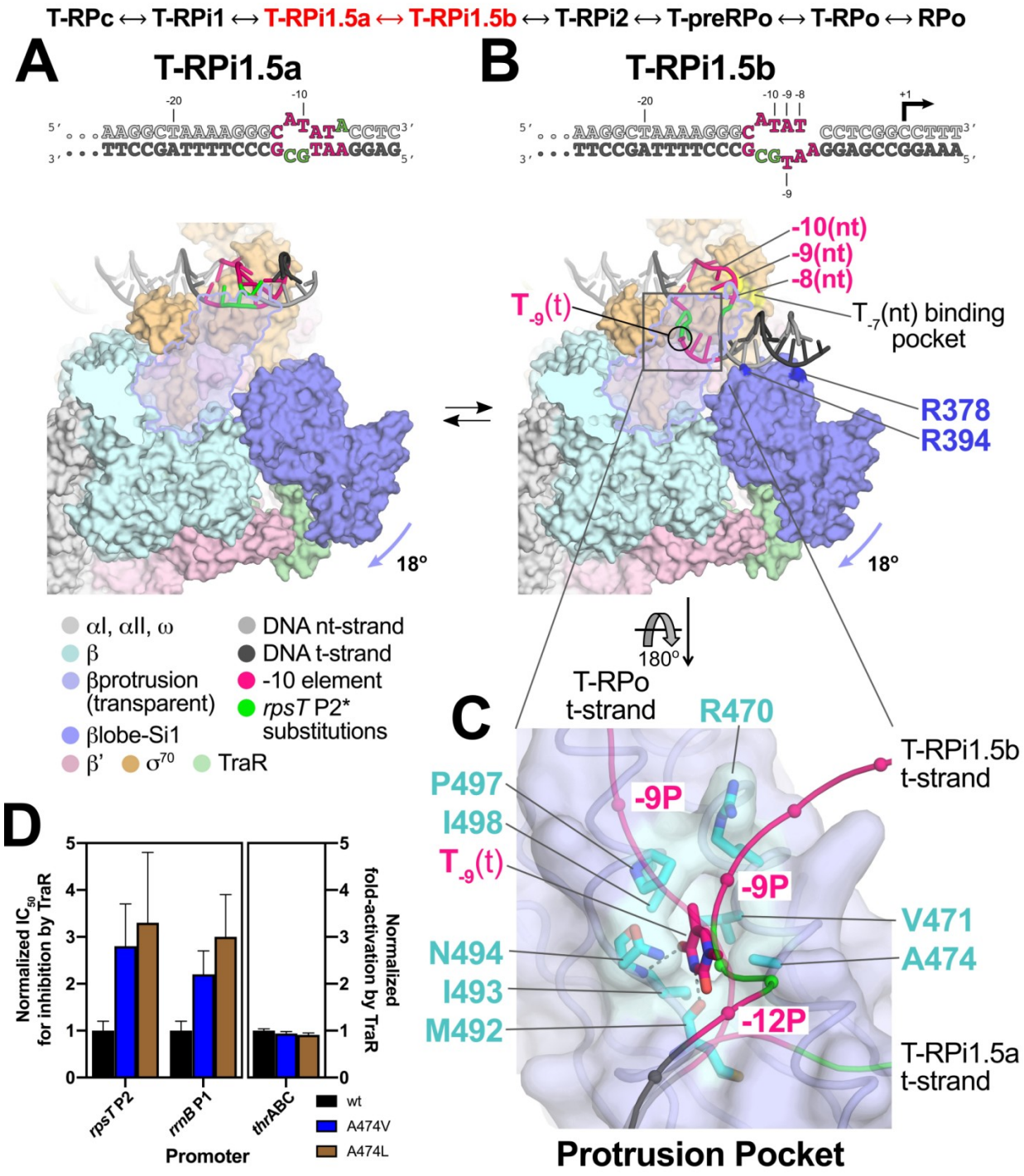

T-RPc ↔ T-RPi1 ↔ T-RPi1.5a ↔ T-RPi1.5b ↔ T-RPi2 ↔ T-preRPo ↔ T-RPo ↔ RPo
A
A-11(nt) binding
W-dyad
W434
W433
W-dyad
A-11(nt)
βprotrusion
T-RPc
-11
5′ ...AGGGCATATTCC... 3′
3′ ...TCCCGTATAAGG... 5′
B
W434
W433
-14 -13 -12 -11 -10 -9 -8 -7 -6 -5
C
W-dyad
W434
A-11(nt)
W433
melted -12 base pair
βprotrusion
T-RPi1
5′ ...AGGGCA 3′
3′ ...TCCCG 3′
-12
D
-14 -13 -12
180°
W433
W434
E
W-dyad
W434
A-11(nt)
-12 base pair
W433
βprotrusion
T-RPi1.5a
5′ ...AGGGCATATAT... 3′
3′ ...TCCCGCGTA... 5′
-12
F
-14 -13 -12 -11 -10
180°
W433
W434
β
σ70
DNA nt-strand
DNA t-strand
-10 element
rpsT P2* substitutions

Figure 6.13. T-RPi1.5a ↔ T-RPi1.5b transitions
(*top*) The order of structures through the RPo formation pathway **(Figure 6.10)**. The progression from T-RPi1.5a ↔ T-RPi1.5b, highlighted in red, is the focus of this figure. **(A-B)** Overall view of T-RPi1.5a **(A)** and T-RPi1.5b **(B)**. Eσ^{70} is shown as a molecular surface with promoter DNA in cartoon format (color-coded as in the key). The βprotrusion (light blue) is transparent with an outline. The rotation of the βlobe-Si1 domains (slate blue) induced by TraR is indicated by the slate blue arrow. **(A)** T-RPi1.5a. **(B)** T-RPi1.5b. The T_{-9}(t) base flips up and is bound in the protrusion-pocket on the underside of the βprotrusion (boxed). βlobe residues R378 and R394 interact with the DNA in T-RPi1.5b but not in RPo, are highlighted (dark blue). **(C)** The protrusion pocket, viewed from the underside of the βprotrusion. The βprotrusion is shown as a backbone worm with a transparent molecular surface. The T-RPi1.5b t-strand DNA is shown as a thin backbone worm with phosphate atom positions denoted by spheres. The T_{-9}(t) base, bound in the protrusion-pocket, is shown as sticks. The t-strand DNA backbone paths for T-RPi1.5a (precedes T-RPi1.5b in the RPo formation pathway) and T-RPo (follows T-RPi1.5b) are shown for comparison. Protrusion-pocket residues that interact with the T_{-9}(t) base are shown as sticks and colored cyan. **(D)** Effect of βA474 substitutions on TraR-mediated inhibition of *rpsT* P2 and *rrnB* P1 promoters (*left*) or activation of *thr*ABC (*right*). For *rpsT* P2 and *rrnB* P1, IC_{50} values for TraR inhibition of WT RNAP (black bar), βA474V-RNAP (blue bar), and βA474L-RNAP (brown bar) are plotted relative to WT RNAP (normalized to 1.0). For *thr*ABC, fold-activation (relative to no TraR) at 500 nM TraR is plotted relative to WT RNAP (normalized to 1.0). Averages with standard deviation from three independent experiments are shown. Transcription assay data courtesy of S. Gopalkrishnan from the Gourse Lab (University of Wisconsin–Madison).


T-RPc ↔ T-RPi1 ↔ T-RPi1.5a ↔ T-RPi1.5b ↔ T-RPi2 ↔ T-preRPo ↔ T-RPo ↔ RPo
A
T-RPi1.5a
B
T-RPi1.5b
-10(nt)
-9(nt)
-8(nt)
T-9(t)
T-7(nt) binding pocket
R378
R394
18°
180°
αI, αII, ω
β
βprotrusion (transparent)
βlobe-Si1
β'
σ70
TraR
DNA nt-strand
DNA t-strand
-10 element
rpsT P2* substitutions
T-RPo t-strand
C
R470
T-RPi1.5b t-strand
P497
I498
T-9(t)
N494
I493
M492
-9P
-12P
V471
A474
T-RPi1.5a t-strand
Protrusion Pocket
D
Normalized IC50 for inhibition by TraR
Normalized fold-activation by TraR
rpsT P2
rrnB P1
thrABC
Promoter
wt
A474V
A474L

among bacterial RNAPs, especially among proteobacteria such as *Eco* (**Figure 6.14A**), pointing to functional importance. The backbone carbonyl of βM492 and the backbone amide of βN494 form hydrogen-bonds with the T_{-9}(t) base, suggesting that the protrusion pocket is thymine specific (**Figure 6.14B**).

To test whether binding of T_{-9}(t) in the protrusion pocket has functional consequences, S. Gopalkrishnan from the Gourse Lab (University of Wisconsin) constructed structure-guided RNAP protrusion pocket mutants, βA474V and βA474L. These substitutions were expected to fill the pocket and exclude the thymine base (**Figure 6.13C**). These mutants affected inhibition of WT *rpsT* P2 and *rrnB* P1 transcription but not activation of p*thrABC* by TraR (**Figure 6.13D**). These results support the functional importance of these pocket residues for T_{-9}(t) interaction and for TraR-mediated stabilization of an intermediate essential for inhibition by TraR (Chen et al., in press).

In T-RPi1.5b, the promoter DNA establishes many $E\sigma^{70}$ contacts that are unique to this intermediate; the contacts are either altered or absent in the subsequent RPo-like complexes (T-RPo or RPo; **Table 6.15**). The substitution of these residues would be expected to affect RPo formation by altering the multi-step energy landscape of RPo formation even though they do not interact with promoter DNA in the final RPo. To test this hypothesis, S. Gopalkrishnan from the Gourse Lab (University of Wisconsin–Madison) constructed Ala and Glu substitutions of βR378 and βR394, two basic residues of the βlobe that interact with the promoter DNA in T-RPi1.5b but not in RPo (**Figure 6.13B**). The mutant RNAPs are not only defective in TraR-mediated inhibition (**Figure 6.16B**) but also in basal transcription (**Figure 6.16A**), consistent with a model where $E\sigma^{70}$-promoter interactions that stabilize the T-RPi1.5b intermediate are important for TraR inhibition but also play a role in basal transcription in the absence of TraR.

6.7 T-RPi1.5b ↔ T-RPi2 ↔ T-preRPo: transcription bubble completion and $\sigma^{70}_{1.1}$ ejection

In T-RPi2, the intermediate following T-RPi1.5b, T_{-7}(nt) is engaged in its σ^{70}_{2} pocket, the single-stranded nt-strand DNA from -11 to -5 interacts with the RNAP in a similar manner as in RPo, but in contrast to RPo, $\sigma^{70}_{1.1}$ remains in the RNAP cleft (**Figures 6.17A-B**). The extent of the T-RPi2 transcription bubble cannot be determined because the downstream DNA in this complex lacks cryo-EM density (**Figure 6.17B**).

In the next intermediate (T-preRPo), the transcription bubble is fully formed (13 nucleotides, from -11 to +2) and the downstream duplex DNA occupies the RNAP channel, displacing $\sigma^{70}_{1.1}$ (**Figure 6.17C**) and initiating the late stages of the pathway (**Figure 6.10**). However, the βlobe-Si1 is still in its rotated conformation and interacting with the TraR that remains bound (**Figure 6.17C**).

Complexes following in the pathway are T-RPo, where TraR remains bound but the rest of the RNAP, including the βlobe-Si1, attains an RPo-like conformation (**Figure 6.17D).** Although the DNA is in an RPo-like state, T-preRPo and T-RPo would not be transcriptionally active because the presence of TraR sterically blocks trigger loop (TL)-folding (Vassylyev et al., 2007; Wang et al., 2006; Windgassen et al., 2014) and entry of the nucleotide (NTP) substrates (Chen et al., 2019b).

6.8 Discussion

Using cryo-EM, seven different intermediate structures were observed that delineate changes in the conformation of both $E\sigma^{70}$ and the *rpsT* P2 promoter on the pathway to forming transcription-capable RPo. To facilitate structure determination, TraR was used to inhibit *rpsT*

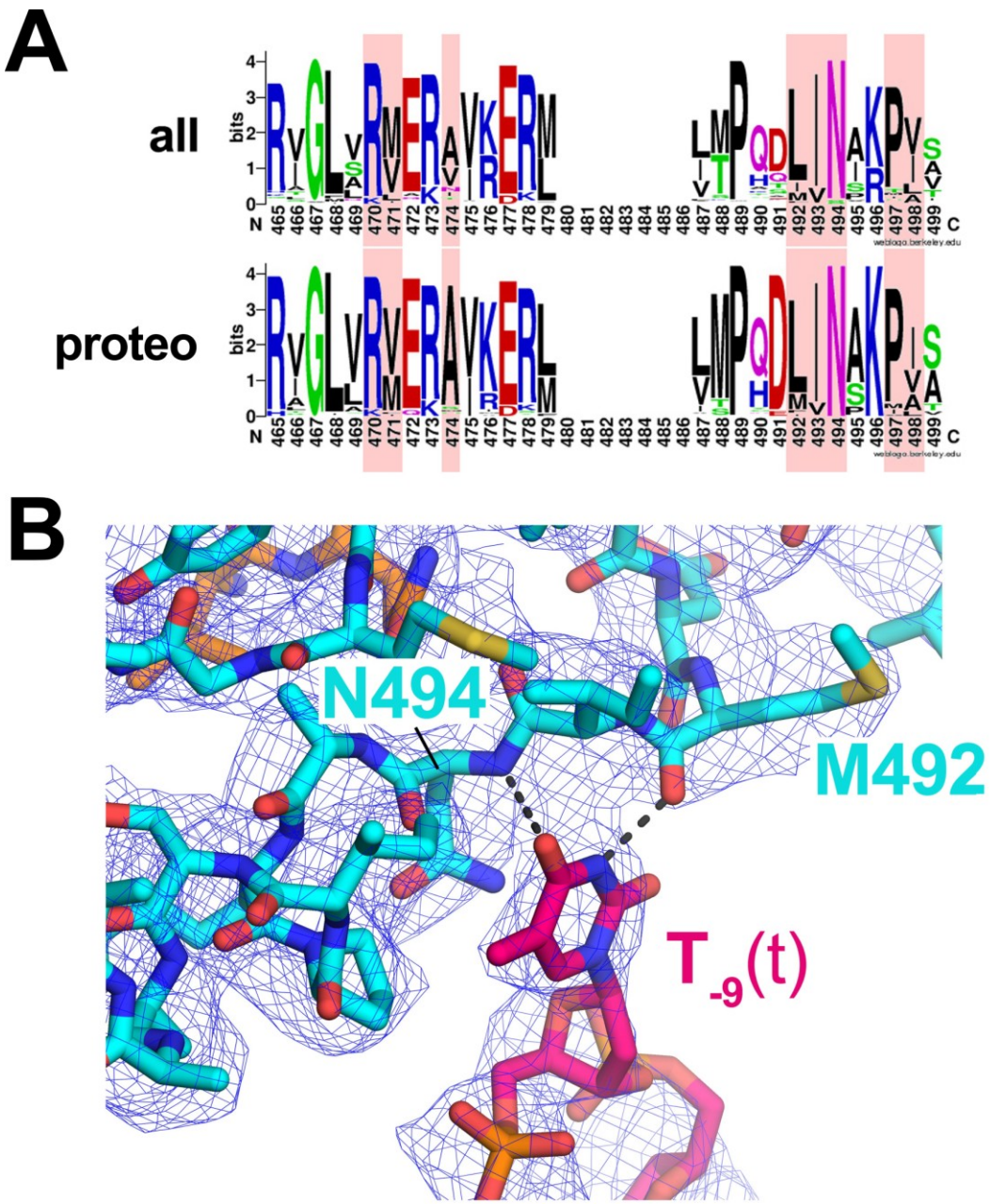


Figure 6.14. Conservation of the protrusion pocket
(A) Sequence logos showing sequence conservation of bacterial RNAP β subunit residues near the protrusion pocket. Residues of the protrusion pocket itself **(Figure 6.13C)** are highlighted with red shading. (*top*) Sequence logo calculated from a sequence alignment of 958 bacterial RNAP RpoB sequences (Lane and Darst, 2010a). (*bottom*) Sequence logo calculated from a sequence alignment of 527 proteobacterial RNAP RpoB sequences (Lane and Darst, 2010a). **(B)** Cryo-EM map of T-RPi1.5b, filtered according to the local resolution (Cardone et al., 2013), is shown as a blue mesh. The final refined model is superimposed. Shown is $T_{-9}(t)$ flipped into the protrusion pocket (residues of the βprotrusion are removed for clarity). Thymine-specific hydrogen-bonds between βM492(O) - $T_{-9}(t)$(N3) and βN494(N) - $T_{-9}(t)$(O4) are indicated by dashed gray lines.

Table 6.15. RNAP/DNA interactions unique to the T-RPi1.5b intermediate

RNAP structural module	residue[a]	conservation[b]	T-RPi1.5b	T-RPo/ RPo
βprotrusion	Y62	nc[c]	nt -5	-
	V471	99% LVMI	t -9	-
	R473	99.7% KR	t -8	nt -4
	A474	nc	t -9	-
	K476	95% KR	nt -4	-
	L481	nc	nt -6	-
	M492	nc	t -9	-
	I493	0.91	t -9	-
	I498	hydrophobic	t -9	-
βlobe	Y367	94% Y	nt -1	-
	R371	98% KR	nt -2/-1	nt -4
	R378	nc	t-4	-
	R394	89% KR	nt -3/-2	-
$\sigma_{1.2}$	R113	nc	nt -9	-
σ_2	K393	100% KR	t -10	-
	T395	nc	t -8	-
	N396	nc	t -8	
σ_3	R468	99.5% R	t -10	-

[a] According to *Eco* σ^{70} numbering.
[b] Calculated from an alignment of 1,002 diverse primary σ sequences (Bae et al., 2015a).
[c] 'nc' = non-conserved

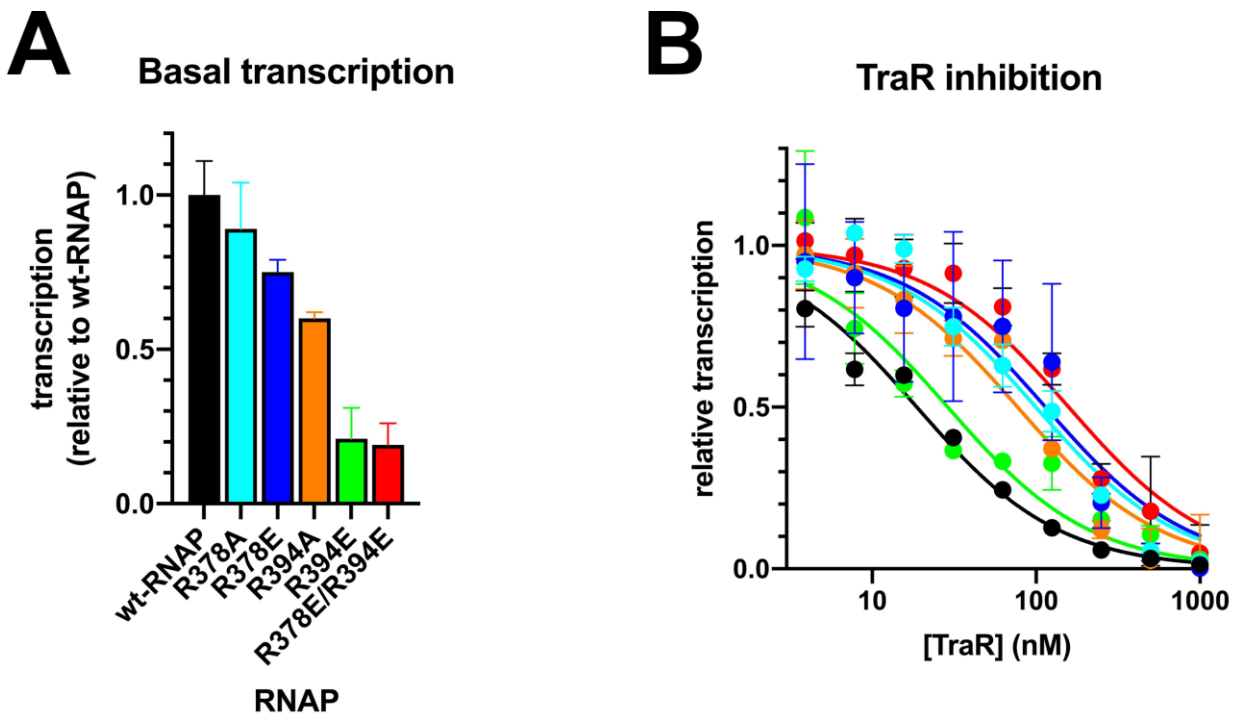


Figure 6.16. βlobe interactions in T-RPi1.5b
(A) Basal transcription (without TraR) of *rpsT* P2 by WT RNAP or RNAPs with βR378 and/or βR394 substitutions as indicated. Averages with standard deviation from three independent experiments are plotted. **(B)** Multi round *in vitro* transcription of *rpsT* P2 was performed with WT RNAP or RNAPs with βR378 and/or βR394 substitutions at a range of concentrations of TraR (4 nM-1 μM). Transcripts were quantified and plotted relative to values in the absence of TraR. The color-coding is the same as in the bar graph at the right. The IC_{50}'s for inhibition by TraR were: WT RNAP, ~19 nM; βR394E-RNAP, ~29 nM; βR394A-RNAP, ~77 nM; βR378A-RNAP, ~100 nM; βR378E-RNAP, ~114 nM; βR378E/R394E-RNAP, ~160 nM. Averages with standard deviation from three independent experiments are plotted. Transcription assay data and figure courtesy of S. Gopalkrishnan from the Gourse Lab (University of Wisconsin–Madison).

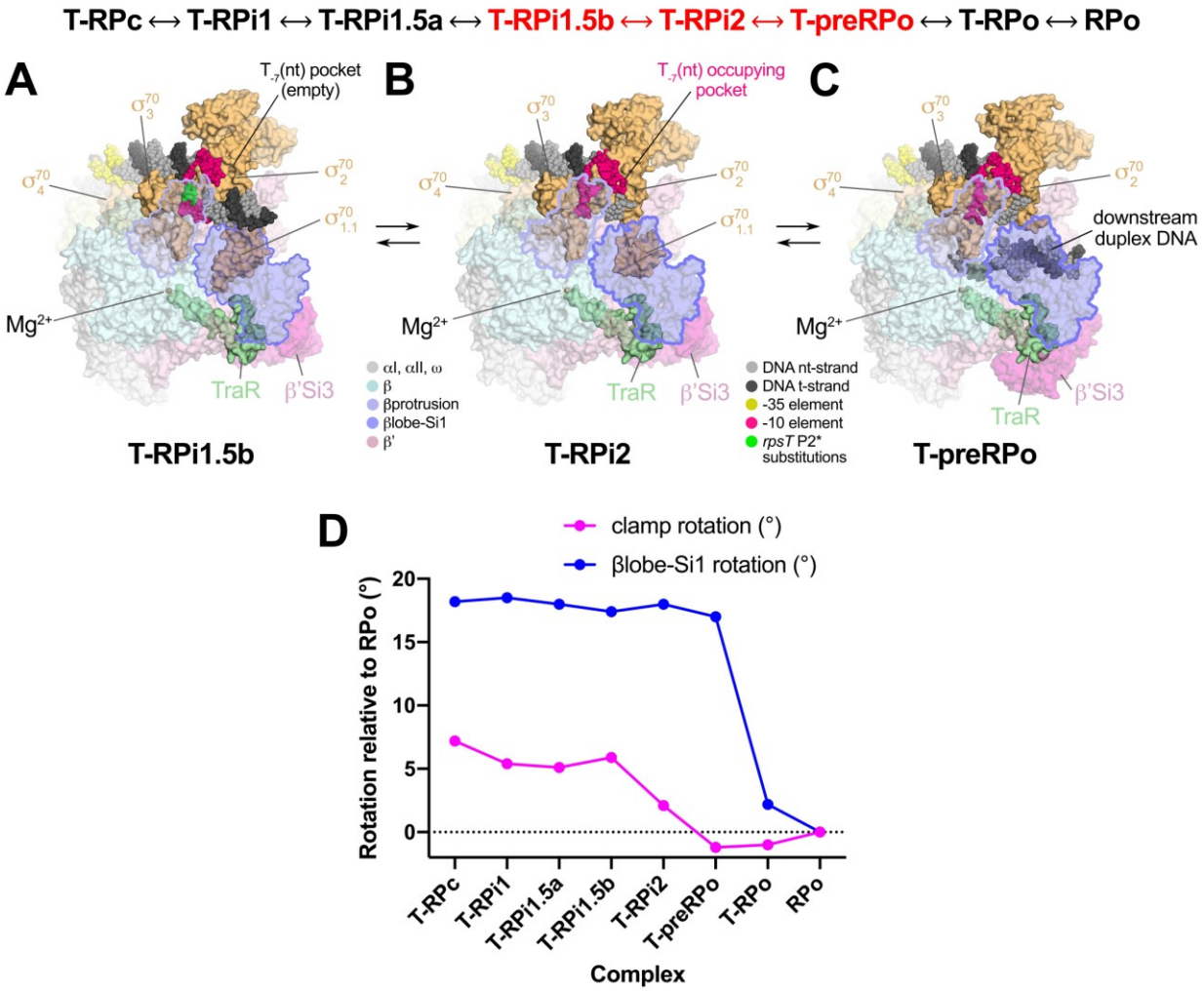


Figure 6.17. T-RPi1.5b ↔ T-RPi2 ↔ T-preRPo transitions
(*top*) The order of structures through the RPo formation pathway **(Figure 6.10)**. The progression from T-RPi1.5b ↔ T-RPi2 ↔ T-preRPo, highlighted in red, is the focus of this figure. **(A-C)** Overall view of T-RPi1.5b **(A)**, T-RPi2 **(B)**, and T-preRPo **(C)**. Eσ^{70} is shown as molecular surfaces, with core RNAP transparent, revealing the RNAP active site Mg^{2+} (sand-colored sphere), TraR in the secondary channel, and either $\sigma^{70}_{1.1}$ (T-RPi1.5b and T-RPi2) or downstream duplex DNA (T-preRPo) in the RNAP channel. The βprotrusion (light blue) and βlobe-Si1 (slate blue) are outlined. **(A)** T-RPi1.5b: Downstream duplex DNA is accommodated in the gap between the βprotrusion and βlobe-Si1. The empty T_{-7}(nt) pocket in σ^{70}_{2} is denoted. **(B)** T-RPi2: The -10 element T_{-7}(nt) is engaged in its cognate σ^{70} pocket, the transcription bubble advances in the downstream direction, and the single-stranded nt-strand downstream to -4 is positioned in the complex much like RPo. The downstream edge of the transcription bubble and downstream duplex DNA are disordered and $\sigma^{70}_{1.1}$ occupies that RNAP channel. **(C)** T-preRPo: The transcription bubble is fully formed (-11 to +2). The downstream duplex DNA is accommodated in the RNAP channel in place of the ejected $\sigma^{70}_{1.1}$. **(D)** Conformational characteristics of the complexes (with RPo as reference): clamp rotation (magenta; positive rotation corresponds to clamp opening); βlobe-Si1 rotation (blue; positive rotation corresponds to rotation towards TraR).

P2 and increase the occupancy of intermediates earlier in the kinetic pathway (Gopalkrishnan et al., 2017; Rutherford et al., 2009). Analysis of the melting intermediates provides details about the mechanism of transcription initiation. Early intermediates reveal unanticipated transient events, including the melting of the -12 base pair and capture of the T_{-9}(t) base. We propose that the intermediate structures describe the steps in DNA opening at most if not all Eσ^{70} promoters and below I discuss the implications of this model for regulation. Finally, I outline how these complexes inform models of DNA opening.

6.8.1 The RNAP clamp

Clamp dynamics play an important role in promoter melting for all cellular RNAPs (Boyaci et al., 2019a; Chakraborty et al., 2012; Feklistov et al., 2017; He et al., 2013; Schulz et al., 2016). In **Chapter 5**, TraR-Eσ^{70} cryo-EM structures revealed TraR binding allosterically dampens RNAP clamp motions (Chen et al., 2019b). Analysis of RNAP clamp positions in the TraR-Eσ^{70}-promoter intermediates (compared to RPo as reference) revealed that the initial Eσ^{70}-promoter complex, T-RPc, has the most open clamp (7.2° open; **Figure 6.17D**). Transient closing of the clamp in T-RPc would pinch the DNA between the βprotrusion and σ^{70}_2 (**Figure 6.11A**), thereby initiating bubble nucleation (Feklistov et al., 2017). In the early intermediate complexes where A_{-11} capture is first detected (T-RPi1→T-RPi1.5b), the clamp is ~5° open. The clamp generally closes as the pathway approaches RPo, but not monotonically (**Figure 6.17D**).

6.8.2 RPo formation involves transient melting of the -12 base pair

The first intermediate visualizing bubble nucleation, T-RPi1, reveals that A_{-11}(nt) capture occurs before or concurrent with W-dyad isomerization and results in transient -12 base pair melting due to steric clash with W433 (**Figures 6.12C-D**). Subsequently, the W433 side chain rotates into the chair conformation, relieving the steric clash and stabilizing -12 base-pair formation by stacking on the exposed downstream face of the -12(nt) base (**Figures 6.12E-F**). We suggest this transient -12 base pair melting may occur at most promoters and could help explain conservation of the TA base pair at the -12 position (Shultzaberger et al., 2007).

6.8.3 T-RPi1.5b likely occurs during basal RPo formation and is stabilized by TraR

The *rpsT* P2* promoter was designed to stabilize the intermediate detected by $KMnO_4$ and DNase I footprinting on the *rpsT* P2 T-7A promoter (**Figure 6.1**, lanes 6, 7). T-RPi1.5b, the prominent intermediate observed by cryo-EM with *rpsT* P2* (**Figures 6.10A**, **6.13B**, **6.17A**), has a transcription bubble of 5 nucleotides (-11 to -7) and downstream DNA contacts that extend to +4. Since the size of the transcription bubble detected by $KMnO_4$ footprinting and the DNase I protection boundary of the *rpsT* P2 T-7A promoter (**Figure 6.1**, lanes 6, 7) are consistent with the properties of the T-RPi1.5b complex, we conclude that the structurally and biochemically detected intermediate are the same.

Several lines of evidence support the hypothesis that the T-RPi1.5b intermediate, or a similar complex, is a part of the normal RPo formation pathway, even in the absence of TraR: (1) At 23°C, the $KMnO_4$ and DNase I footprints on the *rpsT* P2 T-7A promoter were very similar with or without TraR (**Figure 6.1**, lanes 6, 7), indicating that TraR is not required for its formation; (2) substitutions of Eσ^{70} residues that interact with the DNA in T-RPi1.5b but not in RPo affect basal transcription without TraR as well as inhibition by TraR (**Figures 6.16A-B**); (3) the *rpsT* P2* promoter fragment containing mismatched base pairs within the -10 element (**Figure 6.6A**) has weaker transcription compared to WT *rpsT* P2 in the absence of TraR (**Figure**

6.6B). These results suggest that T-RPi1.5b intermediate occurs during promoter unwinding and its formation can slow the progression of the RPo pathway for this promoter, in the absence or presence of TraR.

6.8.4 T-RPi2, $\sigma^{70}_{1.1}$ ejection, and completion of the transcription bubble

The cryo-EM analysis of TraR-Eσ^{70} bound *rpsT* P2* [with the $T_{-7}A$(nt) substitution] yields two distinct classes, T-RPi1.5a and T-RPi1.5b, both of which have $\sigma^{70}_{1.1}$ occupying the RNAP channel. Image classification did not reveal RPo-like classes in this dataset. On the other hand, WT *rpsT* P2 yields RPo-like complexes in which the downstream duplex DNA displaces $\sigma^{70}_{1.1}$ (despite the presence of TraR), suggests that engagement of T_{-7}(nt) with its cognate σ^{70}_{2} pocket is an important determinant of $\sigma^{70}_{1.1}$ ejection. The pathway progresses from T-RPi1.5b, with a 5-nucleotide transcription bubble, to T-preRPo with a full, 13-nucleotide transcription bubble through a single intermediate (T-RPi2) where the extent of the transcription bubble and the path of the downstream duplex DNA is highly dynamic (**Figure 6.17**). This suggests that following nucleation of the transcription bubble and melting of the -10 element, the propagation of the transcription bubble downstream to the start site occurs rapidly, consistent with kinetic analyses of RPo formation (Hubin et al., 2017b; Ruff et al., 2015b; Saecker et al., 2011).

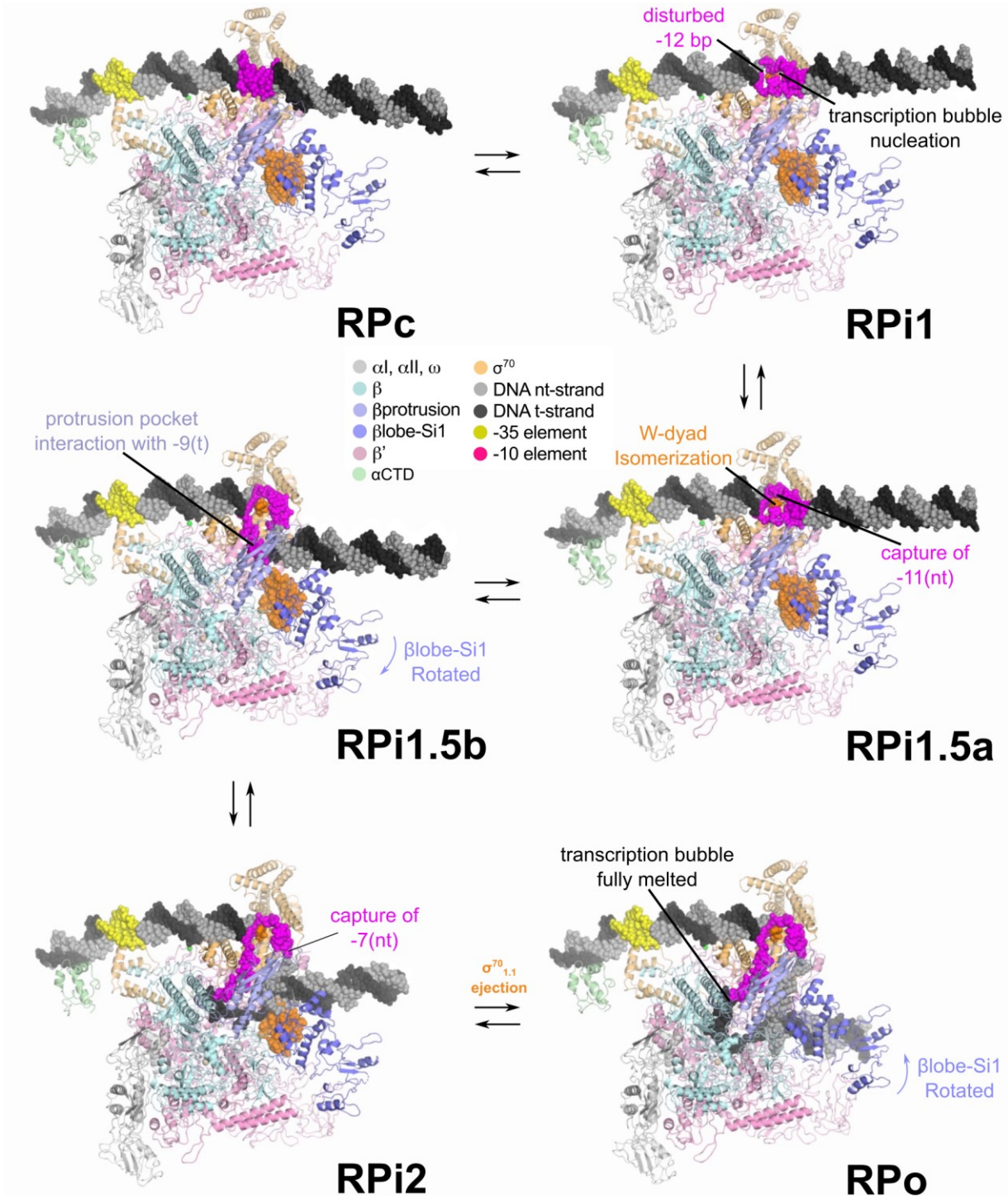
6.8.5 The complete RPo formation pathway and TraR binding

The βlobe-Si1 stays rotated throughout most of the pathway (until T-RPo), reflecting TraR binding, but how does this relate to basal RPo formation in the absence of TraR? In the early steps of the pathway preceding T-RPi1.5b (T-RPc ↔ T-RPi1 ↔ T-RPi1.5a), the promoter DNA is far from βlobe-Si1 so the rotation of this domain does not impact these steps directly. This suggests that the salient structural features of these intermediates (binding of duplex DNA in T-RPc, transcription bubble nucleation in T-RPi1 and subsequent W-dyad isomerization) reflect the RPo formation pathway with or without TraR. The βlobe-Si1 rotation stabilized by TraR likely has an indirect effect by helping to populate these early intermediates. Only one intermediate on the pathway, T-RPi1.5b, appears to require βlobe-Si1 rotation for its formation. We argue that T-RPi1.5b is on the basal pathway, and by extension, we suggest that βlobe-Si1 rotation occurs at this point during the basal pathway, but transiently. This proposed pathway is illustrated in **Figure 6.18**.

6.8.6 Five base-specific pockets in Eσ^{70} modulate RPo formation

RPo formation is controlled, in part, by base-specific pockets distributed throughout the Eσ^{70} structure. The cognate pockets for A_{-11}(nt) and T_{-7}(nt) in σ^{70} are essential for transcription bubble nucleation and -10 element melting (Feklistov and Darst, 2011), and these interactions are maintained in the final RPo. The protrusion pocket discovered here binds T_{-9}(t) transiently (**Figure 6.13C**), contributing to regulation by TraR (**Figure 6.13D**), and may play a role in RPo formation in the absence of factors (**Figure 6.1**, **lane 6**). A binding site for G_{-5}(nt) also plays a role in modulating RPo lifetime and regulation by guanosine pentaphosphate (ppGpp) and DksA (Haugen et al., 2006, 2008). Finally, G_{+2}(nt) binds in an RNAP β-subunit pocket (Zhang et al., 2012). The binding sites for these bases are shown in **Figure 6.19**. These five separate pockets and the myriad possible interactions with different promoter sequences give rise to a combinatorial effect that contributes to the 10,000-fold variation in initiation rates *in vivo* and *in vitro* (McClure, 1985; Ruff et al., 2015b).

Figure 6.18. Model of complete RPo formation pathway
Proposed model for RPo formation without TraR based on the cryo-EM structures in this chapter. Proteins are shown in cartoon format and color-coded based on key (αI, αII, ω, light gray; β, pale cyan; βprotrusion, tv blue; βlobe-Si1, slate; β', light pink; αCTD, pale green; σ^{70}, light orange). For clarity, $\sigma^{70}_{1.1}$ is shown as orange spheres. The promoter DNA is shown as spheres and color-coded based on key (nt-strand, gray; t-strand, dark gray; -35 element, yellow; -10 element, hot pink). The active site Mg^{2+} is shown as a sand-colored sphere. The pathway is composed of RPc ↔ RPi1 ↔ RPi1.5a ↔ RPi1.5b ↔ RPi2 ↔ RPo with key structural changes annotated. Note, the rotation of βlobe-Si1between RPi1.5a ↔ RPi1.5b (to accommodate the downstream duplex DNA) and between RPi2 ↔ RPo (to stabilize the nt-strand after $\sigma^{70}_{1.1}$ ejection).


disturbed
-12 bp
transcription bubble
nucleation
RPc
RPi1
αI, αII, ω
β
βprotrusion
βlobe-Si1
β'
αCTD
σ70
DNA nt-strand
DNA t-strand
-35 element
-10 element
protrusion pocket
interaction with -9(t)
W-dyad
Isomerization
capture of
-11(nt)
βlobe-Si1
Rotated
RPi1.5b
RPi1.5a
capture of
-7(nt)
σ70 1.1
ejection
transcription bubble
fully melted
βlobe-Si1
Rotated
RPi2
RPo

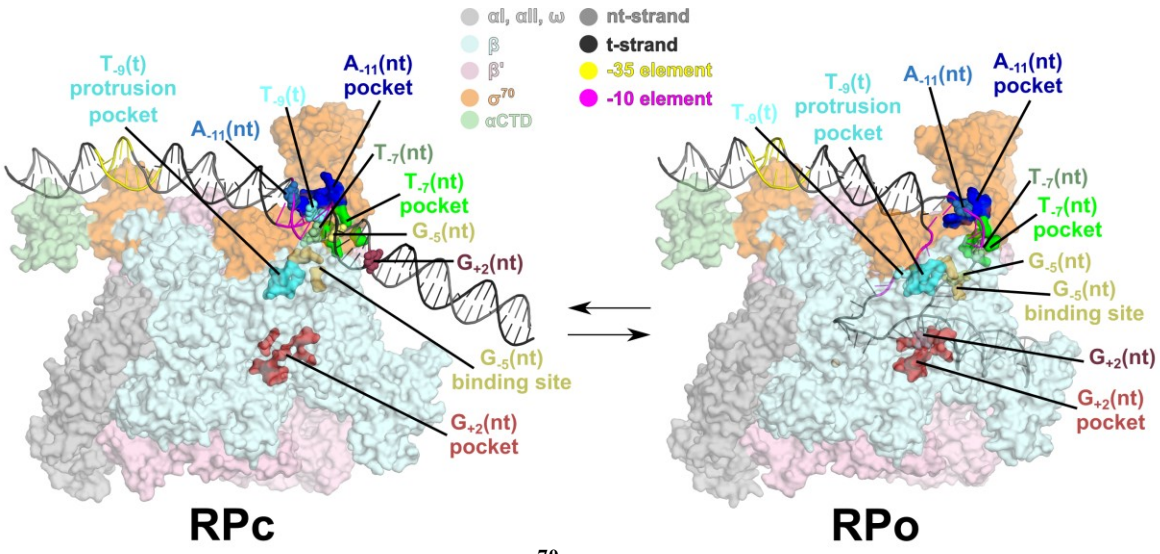


Figure 6.19. Five base-specific pockets in Eσ^{70}

Structural models were derived from the cryo-EM structures in this chapter. Proteins are shown as molecular surfaces and promoter DNA is shown in cartoon format. Protein and DNA are color-coded based on key. The five bases are shown as colored spheres [A_{-11}(nt), marine; T_{-9}(t), aquamarine; T_{-7}(nt), pale green; G_{-5}(nt), pale yellow; G_{+2}(nt), raspberry]. The location of the five base-specific pockets are shown as molecular surfaces and colored corresponding to the base they recognize [A_{-11}(nt) pocket, blue; T_{-9}(t) protrusion pocket, cyan; T_{-7}(nt) pocket, green; G_{-5}(nt) binding site, tv yellow; G_{+2}(nt) pocket, red]. (*left*) RPc. Promoter DNA is duplex and no base-specific interactions have been made with the -10 element. (*right*) RPo. Base-specific interactions are made: A_{-11}(nt) binds the A_{-11}(nt) pocket in σ^{70}_2, T_{-7}(nt) binds the T_{-7}(nt) pocket in σ^{70}_2, G_{-5}(nt) of the discriminator region interacts with G_{-5}(nt) binding site in $\sigma^{70}_{1.2}$, G_{+2}(nt) binds the G_{+2}(nt) pocket in β subunit. T_{-9}(t) interaction in the protrusion pocket is not maintained in RPo.

While A_{-11}(nt) and T_{-7}(nt) are present in nearly all $E\sigma^{70}$ promoters (Feklistov and Darst, 2011; Shultzaberger et al., 2007), T_{-9}(t) is less well conserved but is enriched for in promoters that are negatively regulated, and underrepresented in promoters that are positively regulated, by the DksA/ppGpp and TraR (Sanchez-Vazquez et al., 2019). Likewise, G_{-5}(nt) and G_{+2}(nt) are not conserved but modulate RPo formation when they are present.

6.8.7 Relationship to previously identified intermediates

It has long been appreciated that RPo formation is a multi-step process (Buc and McClure, 1985; Hawley and Mcclure, 1982; Kadesch et al., 1982; Roe et al., 1984; Rosenberg et al., 1982; Walter et al., 1967). RPo formation intermediates of *Eco* $E\sigma^{70}$ have been characterized on several promoters (Rogozina et al., 2009; Rutherford et al., 2009; Sclavi et al., 2005), none more extensively than λPR, where three 'kinetically significant' intermediates in RPo formation at λPR, I1, I2, and I3, have been identified (Ruff et al., 2015b). I1 is proposed to comprise an ensemble of closed complexes and conversion to I2 involves the melting of the entire transcription bubble, loading of the DNA into the RNAP cleft, and ejection of $\sigma^{70}_{1.1}$ (Ruff et al., 2015b). Thus, partially melted intermediates have not been observed at this promoter.

A study of the kinetics of RPo formation by *Mtb* RNAP on the *Mtb rrnA* P3 promoter identified a minimum of two significant intermediates, termed RP1 and RP2 (Hubin et al., 2017b). The structure of a partially melted intermediate, proposed to correspond to RP2, was revealed by cryo-EM (Boyaci et al., 2019a). The RP2 intermediate is not structurally similar to any of the intermediates observed here but would lie between T-RPi2 and T-preRPo. Given the differences in the nature of the N-terminal domains of *Mtb* σ^A and *Eco* σ^{70} (Hubin et al., 2017a), in lineage-specific inserts (LSI) in β and β' (Lane and Darst, 2010a), and the presence of *Mtb* general transcription factors (GTF) (RbpA and CarD) not found in *Eco* (Hubin et al., 2017b), it is unclear whether *Eco* $E\sigma^{70}$ would significantly populate an equivalent intermediate.

How can the intermediates characterized at *Mtb rrnA* P3 and *Eco* λPR be related to the intermediates observed here? We suggest that ensemble footprinting or fluorescence approaches do not have the sensitivity and/or temporal resolution to distinguish some of the intermediates identified in our structures. Thus, I1, I2, I3, RP1, RP2, and other intermediates described previously are likely ensembles of many intermediates that accumulate at kinetic bottlenecks along the RPo formation pathway.

Does DNA opening involve the same steps at every promoter or does the pathway depend on promoter sequence? We propose that RPo formation by $E\sigma^{70}$ proceeds through very similar conformational changes defined by these intermediates, whether assisted by transcription factors or not. Because binding free energy drives each interconversion, the overall net gain in $E\sigma^{70}$-DNA interactions versus the cost of duplex DNA disruption at each step determines the corresponding rate constants (Haugen et al., 2008; Ruff et al., 2015b). As DNA sequence dictates the significance of each step (i.e. whether a particular step is rate-limiting), not all intermediates are significantly populated at a given promoter, and additional intermediates not described here may be identifiable at other promoters. These kinetic differences allow regulators to alter rates at target promoters without significantly affecting others (Haugen et al., 2008).

6.8.8 Mechanism of promoter melting

Overall models for the mechanism of RPo formation by $E\sigma^{70}$ have been framed by two extremes that posit where duplex DNA unwinding occurs in RNAP (Mazumder and Kapanidis, 2019), either outside (melt-load model) or inside the RNAP cleft (load-melt). Based on observations of

a closed-clamp state in early crystal structures of bacterial RNAP holoenzymes, the melt-load model proposed that duplex DNA positioned outside the cleft could unwind and only single-stranded DNA would be allowed into the RNAP cleft (Vassylyev et al., 2002).

The load-melt model is consistent with footprinting and other kinetic studies, primarily at λPR, that suggest the ensemble of closed I1 complexes includes complexes in which the duplex DNA downstream of the -10 element is protected inside the RNAP cleft (Gries et al., 2010; Saecker et al., 2011). This model requires conformational changes in the RNAP to allow entry of duplex DNA into the cleft. Conformational changes in the mobile clamp domain (Boyaci et al., 2019a; Chakraborty et al., 2012; Chen et al., 2019b; Duchi et al., 2018; Gnatt et al., 2001) and the βlobe (Boyaci et al., 2019a; Chen et al., 2010a) play major roles in allowing DNA access to the RNAP cleft.

The results of this study, combined with other available evidence, support a combination of both models. Consistent with a melt-load model, the downstream duplex DNA in T-RPc (**Figure 6.11**) is located outside the cleft, and subsequent intermediates clearly show that transcription bubble nucleation occurs outside the cleft (T-RPi1, T-RPi1.5a; **Figures 6.12C**, **E**). However, consistent with a load-melt model, the effects of antibiotics suggest a role for clamp dynamics in RPo formation (Boyaci et al., 2019a; Feklistov et al., 2017; Lin et al., 2018; Srivastava et al., 2011). The RP2 intermediate observed with *Mtb* RNAP contains a partial bubble with the duplex DNA to be ultimately melted in RPo (including the transcription start site) enclosed in the RNAP cleft, indicating that final melting of the start site occurs within the RNAP cleft (Boyaci et al., 2019a). Further structural characterization of RPo formation intermediates, now enabled by advances in cryo-EM, on diverse promoters and with and without transcription factors, will be required to further delineate the promoter melting mechanism.

6.9 Acknowledgments

The work in this chapter was submitted as a manuscript that is currently in press at Molecular Cell. My role in this work involved: protein purification, cryo-EM sample preparation; data collection and processing; structural model building and analysis; conception of ideas and figures. Citation:

Chen, J., Chiu, C., Gopalkrishnan, S., Chen, A. Y., Olinares, P. D. B., Saecker, R. M., Winkelman, J. T., Maloney, M. F., Chait, B. T., Ross, W., Gourse, R. L., Campbell, E. A., Darst, S. A. (in press). Stepwise promoter melting by bacterial RNA polymerase. Mol. Cell.

I would like to thank the following people for their contributions to this chapter: A. Y. Chen, M. F. Maloney, J. T. Winkelman (University of Wisconsin-Madison) expressed and purified proteins and performed biochemical assays; B. T. Chait and P. D. B. Olinares (The Rockefeller University) performed native mass spectrometry analyses; C. Chiu (The Rockefeller University) assisted in cryo-EM preparation; E. A. Campbell and S. A. Darst (The Rockefeller University) provided mentorship and assisted in structural analyses and interpretations; S. A. Darst (The Rockefeller University) edited this chapter; M. Ebrahim and J. Sotiris (The Rockefeller University) helped with cryo-EM data collection; S. Gopalkrishnan (University of Wisconsin-Madison) performed footprinting and transcription assays; R. L. Gourse and W. Ross (University of Wisconsin-Madison) collaborated with our laboratory and supervised various aspects of the project, R. M. Saecker (The Rockefeller University) provided helpful insight and discussion.

CHAPTER 7. Conclusions and future directions

Before 1999, the only structural information for cellular RNA polymerases (RNAP) came from structural studies using electron microscopy (EM) to examine 2D crystals of RNAP formed on lipid bilayers. The studies provided the first 3D structures of RNAP: *Escherichia coli* (*Eco*) RNAP holoenzyme at 27 Å resolution (Darst et al., 1989), followed by yeast RNAP II at 16 Å resolution (Darst et al., 1991). Although these structures were low-resolution due to limitations of the technique, they revealed the overall architecture of RNAP. The structures showed that RNAP is shaped like a crab claw with a large internal cleft that could accommodate nucleic acids. The authors proposed that the active site of the enzyme likely resided in the cleft between the pincers (Darst et al., 1989, 1991). These results provided the first structural bases for mechanistic and functional hypotheses that were not possible previously.

In 1999, the first high-resolution (3.3 Å) structure of RNAP was solved using X-ray crystallography from 3D crystals of *Thermus aquaticus* (*Taq*) core RNAP. This structure revealed the location of the active site Mg^{2+} and identified many of the key structural features critical for the function of the enzyme (Zhang et al., 1999). Subsequently, X-ray crystallography became an enormously powerful structural tool that provided high-resolution snapshots of stable states of bacterial and eukaryotic transcription systems. In 2001, high-resolution structures of yeast RNAP II were solved, revealing striking structural similarities between the bacterial and eukaryotic RNAPs (Cramer et al., 2001; Gnatt et al., 2001). The culmination of this work led to the 2006 Nobel Prize in Chemistry, awarded to Roger D. Kornberg for his studies of the molecular basis of eukaryotic transcription (Kornberg, 2007; Landick, 2006; Svejstrup et al., 2006).

X-ray crystallography was instrumental in solving high-resolution structures of different forms of RNAP and stable states of the bacterial transcription cycle including: core RNAP (Zhang et al., 1999), holoenzyme (Murakami et al., 2002b; Vassylyev et al., 2002), open promoter complexes (RPo) (Bae et al., 2015a; Zuo and Steitz, 2015), and elongation complexes (Vassylyev et al., 2007). These structures provided the foundation for a plethora of genetic and biochemical studies to address mechanistic questions about bacterial transcription. Despite all this progress, X-ray crystallography had its limitations: (1) complexes of interest, even if stable, could not always be crystallized, and (2) the complexes that did crystallize were subject to conformations influenced by crystal packing and thereby affecting the interpretation of the structures. In addition, X-ray crystallography is ill-suited to study dynamic behavior, and transient, unstable states that characterize transitions between stable states.

The work presented in this thesis establishes cryo-electron microscopy (cryo-EM) as a powerful technique for the structural biology studies of bacterial transcription complexes. This technique overcomes the limitations of X-ray crystallography by allowing direct visualization of complexes in near native states without being impacted by crystal packing. Resolutions achieved in cryo-EM are comparable to those in X-ray crystallography, owing to advances in electron detectors and image processing software. The body of work in this thesis outlines how cryo-EM was used: (1) to solve high-resolution structures of complexes that were previously intractable to study by crystallography [e.g. *Eco* σ^{70}-holoenzyme ($E\sigma^{70}$) bound to *Eco* 6S RNA (6S-$E\sigma^{70}$) (Chen et al., 2017), Fidaxomicin (Fdx) bound to RNAP and σ^{N}-promoter DNA complexes (see **Appendix 1**)]; (2) to examine RNAP conformational changes induced by a small molecule [Fdx; (Boyaci et al., 2018)] and a transcription factor [TraR; (Chen et al., 2019b)], and (3) to observe promoter melting intermediates on the pathway to RPo formation.

Initial cryo-EM studies of bacterial RNAP revealed a severe particle orientation bias, a problem that needed to be solved before further progress could be achieved. I discovered that the zwitterionic detergent 3-([3-Cholamidopropyl]dimethylammonio)-2-hydroxy-1-propanesulfonate (CHAPSO) could be used to solve this bias by excluding polymerase particles from the air/water interface of cryo-EM grids, pushing them into the vitreous ice layer where they can freely adopt random orientations (Chen et al., 2019a). Interactions with the air/water interface have been proposed to not only cause preferred particle orientation but also to cause denaturation of the specimen (D'Imprima et al., 2019). Denatured particles can affect 3D classification, leading to structural classes that can be misinterpreted as holding biological relevance. The cryo-EM samples investigated in this thesis were all prepared using CHAPSO, and thus were unaffected by particle orientation bias or interactions with the air/water interface.

The cryo-EM structure of 6S-Eσ^{70} demonstrates how a non-coding RNA (ncRNA) can regulate a DNA-dependent RNAP by mimicking B-form DNA (Chen et al., 2017). Previous attempts to use X-ray crystallography to study this complex were unsuccessful. Cryo-EM was an advantageous technique to study this complex since sample preparation is rapid and does not require crystallization. Degradation of the RNA by contaminating RNase was minimized since sample preparation can be performed on the seconds to minutes timescale. Additionally, the inherent dynamics of the RNA structure (Sponer et al., 2018), which were detrimental to crystal packing, were bypassed using solution cryo-EM since crystallization was not required.

In this thesis, cryo-EM served as a robust technique for examining RNAP conformational changes induced by Fdx in *Mycobacterium tuberculosis* (*Mtb)* RNAP and by TraR in *Eco* RNAP. Fdx binding induced opening of the mobile RNAP clamp (Boyaci et al., 2018), was incompatible with crystal formation. The cryo-EM structures provided insight into the conformational changes in the clamp. These structures illustrated the extent and range of clamp motions as opposed to Förster resonance energy transfer (FRET) studies that only show that the clamp dynamics exist.

In *Mtb*, Fdx high affinity binding to RNAP required the essential *Mtb* transcription factor RbpA. Multi-component samples are easier to study by cryo-EM versus crystallography, allowing the capture of more 'complete' biological samples. This was important in linking the strength of Fdx inhibition to RbpA. Our laboratory has begun to use cryo-EM to characterize other RNAP inhibitors and their effects on RNAP conformations in the context of 'complete' transcription initiation complexes (Boyaci et al., 2019a).

In *Eco*, image classification of the cryo-EM data (Bai et al., 2015b) revealed a number of TraR-mediated RNAP conformational changes (Chen et al., 2019b) that were not seen in crystal structures due to crystal packing artifacts (Molodtsov et al., 2018). In addition, multibody refinement (MBR) and principal component analysis (PCA) (Nakane et al., 2018) deconvoluted RNAP clamp motions that were absent in the static structures. The cryo-EM structures reveal a complex interplay between TraR binding and RNAP conformation that allows TraR to modulate multiple features of the energy landscape of RPo formation to differentially regulate initiation without directly interacting with DNA (Chen et al., 2019b).

Structural characterization of promoter melting intermediates by crystallography has not been possible due to their transient nature and dynamic behavior. Using cryo-EM and image classification approaches, it was possible to capture melting intermediates of the *rpsT* P2 promoter bound to TraR-Eσ^{70} (Chen et al., in press). Image classification approaches (Bai et al., 2015b) were able to sort promoter/RNAP particles, frozen at equilibrium, into discrete structures that represent intermediate states spanning the RPo formation pathway, from RPc to RPo.

However, these intermediates only represent a subset of promoter melting states. Further structural characterization of RPo formation intermediates on other promoters, with and without transcription factors, will be required to further delineate the promoter melting mechanism.

The work in this thesis showcases the versatility of cryo-EM to study the molecular machine, RNAP. Cryo-EM has transformed structural biology (Bai et al., 2015a) and it continues to evolve with the advent of next-generation electron detectors, updated software, and novel sample preparation techniques. Next-generation detectors have improved detective quantum efficiency (DQE) that allows for shorter movie exposures and thus higher throughput (Mendez et al., 2019). Recent upgrades in software allow higher-order image correction (Zivanov et al., 2019) to further improve resolution as well as better ways to classify structural heterogeneity (Nakane et al., 2018; Punjani et al., 2017). In combination with the recent updates to processing programs, these large datasets can be used to obtain high-resolution structures of minimally populated states of a complex by brute force. The development of novel techniques in single-particle cryo-EM and cryo-electron tomography (cryo-ET) sample preparation offers new ways to look at biological questions. Time-resolved cryo-EM using microfluidics (Kaledhonkar et al., 2018) provides an approach to capture kinetically regulated intermediates, with recent implementation in the study of the late steps of bacterial translation initiation (Kaledhonkar et al., 2019). *In situ* cryo-ET and subtomogram averaging enable 3D visualization of macromolecular and subcellular biological structures in their cellular context (Koning et al., 2018).

These technical advancements will benefit the structural biology studies of bacterial RNAPs. Future structural studies will focus on other aspects of the bacterial transcription cycle such as promoter escape and termination as well as transcription factors and small molecules that regulate the cycle. Structural information about these transcription states can provide useful information for molecular dynamics (MD) simulations, enabling finer sampling of the conformational changes required to transition through the transcription cycle. As improvements to *in situ* cryo-ET are implemented, it will be possible to solve high-resolution 3D structures of transcription complexes and their higher-order assemblies in bacterial cells.

7.1 Acknowledgments

I would like to thank the following people for their contributions to this chapter: E. A. Campbell and S. A. Darst (The Rockefeller University) provided insight, edits, and revisions.

CHAPTER 8. Material and methods

8.1 Chapter 2

Protein expression and purification. Escherichia coli (*Eco*) RNA polymerase (RNAP) core lacking the α C-Terminal Domain (ΔαCTD-E) was purified using previously established protocols (Twist et al., 2011). *Eco* σ^{70} lacking region 1.1 ($\Delta1.1\sigma^{70}$) complex was prepared as described previously (outlined in S**ection 8.2**; Chen et al., 2017).

Preparation of 6S-Eσ^{70} for single-particle cryo-electron microscopy (cryo-EM). Frozen aliquots of ΔαCTD-E and $\Delta1.1\sigma^{70}$ were mixed in a 1:2 molar ratio and incubated for 15 mins at 37 °C, to form ΔαCTD-E-$\Delta1.1\sigma^{70}$ (termed Eσ^{70} in **Section 8.1** and **Section 8.2**). 6S RNA was added in a 1:1.5 molar ratio and incubated for 15 mins at room temperature, to form 6S-Eσ^{70}. The samples were diluted and detergents added (if used) immediately before grid preparation. 3.5 µL of sample were deposited on glow discharged Quantifoil R 1.2/1.3, 400 mesh, copper grids (EMS) and plunged frozen into liquid ethane using a Cryoplunge 3 system (Gatan).

Single-particle cryo-EM of 6S-Eσ^{70} complex. Cryo-EM grids of 6S-Eσ^{70} were imaged using a 300 kV Tecnai G2 Polara (FEI) equipped with a K2 Summit direct electron detector (Gatan). Dose-fractionated images were collected using UCSFImage4 (Li et al., 2015) in super-resolution mode with a nominal magnification of 31,000×, corresponding to a calibrated pixel size of 1.23 Å on the specimen level (0.615 Å for super-resolution). The dose rate on the camera was 8 counts/physical pixel/second using Digital Micrograph (Gatan). The exposure time per movie was 6 s (30 frames) leading to a total dose of 31.7 electrons/Å^2. Movies were collected using a defocus range from -1 µm to -2.5 µm. Movies were 2 × 2 binned using IMOD (Kremer et al., 1996) and then drift corrected using MotionCor (Li et al., 2013).

Preparation of 6S-Eσ^{70} for cryo-electron tomography (cryo-ET). For KGlu [10 mM Tris-HCl, pH 8.0, 150 mM K-glutamate, 1 mM $MgCl_2$, 2 mM dithiothreitol (DTT)] and KGlu + NP40S [10 mM Tris-HCl, pH 8.0, 150 mM K-glutamate, 1 mM $MgCl_2$, 2 mM DTT, 0.06 mM Nonidet P40 Substitute (NP40S)] conditions, Eσ^{70} was purified in KGlu buffer using a Superose6 Increase column (GE Healthcare). For KCl (10 mM Tris-HCl, pH 8.0, 150 mM KCl, 1 mM $MgCl_2$, 2 mM DTT) and KCl + CHAPSO [10 mM Tris-HCl, pH 8.0, 150 mM KCl, 1 mM $MgCl_2$, 2 mM DTT, 8 mM cholamidopropyl]dimethylammonio)-2-hydroxy-1-propanesulfonate (CHAPSO)], Eσ^{70} was purified in KCl buffer. Peak fractions corresponding to Eσ^{70} were pooled and concentrated by centrifugal filtration (VivaScience) to 4 mg/mL protein concentration. 6S RNA was added in 1.2 molar excess compared to holoenzyme and incubated at room temperature. Immediately prior to grid freezing, samples of KGlu or KCl were diluted 1:10 with their respective buffers while NP40S was added to the KGlu + NP40S to critical micelle concentration (CMC) (0.06 mM) and CHAPSO was added to the KCl + CHAPSO sample to CMC (8 mM). After centrifugation to remove aggregates, 3.5 µL of sample were deposited on glow discharged Quantifoil R 1.2/1.3, Au, 400 mesh grids (EMS) and plunged frozen into liquid ethane using a Vitrobot Mark IV (FEI).

Acquisition and analysis of 6S-Eσ^{70} cryo-ET data. Cryo-ET data was collected and analyzed by A. J. Noble (New York Structural Biology Center) (Chen et al., 2019a).

Sample preparation and single-particle cryo-EM of Eco transcription elongation complexes (TEC). Sample preparation and data collection of *Eco* TEC were performed by J. Y. Kang from the Darst-Campbell Lab (The Rockefeller University) (Chen et al., 2019a). *Eco* TEC was assembled and prepared using *Eco* RNAP core (ΔαCTD) as previously described (Kang et al., 2017).

Data analysis single-particle cryo-EM datasets. CTF estimations were calculated for each dataset using Gctf (Zhang, 2016). Particles were picked using Gautomatch (developed by K. Zhang, MRC Laboratory of Molecular Biology, Cambridge, UK, http://www.mrc-lmb.cam.ac.uk/kzhang/Gautomatch) without a 2D template. Picked particles were extracted from the dose-weighted images in RELION (Scheres, 2012). Particles were curated by 2D classification (N classes = 50) and 3D classification (N classes = 3) using a crystal structure of *Eco* RNAP (PDB ID 4LJZ; Bae et al., 2013) with σ^{70} removed. The crystal structure was converted to an EM map and low-pass filtered to 60 Å using EMAN2 (Tang et al., 2007) before classification and refinements. Particles were coarsely aligned to the 3D template using RELION (Scheres, 2012) 3D classification (N class = 1). Histograms of particle orientations (BILD format) were graphically represented as spheres. The PSFs and orientation distribution parameters (E_{od} and f_{empty}) were calculated using cryoEF (Naydenova and Russo, 2017).

8.2 Chapter 3

Preparation of RNase-free Eco RNAP core (ΔαCTD). ΔαCTD-E was prepared as described previously (Twist et al., 2011) but with one modification to obtain RNase-free RNAP. During Bio-rex 70 (Bio-Rad) cation exchange, the protein was washed extensively with TGED buffer [10 mM Tris-HCl, pH 8.0, 5% (v/v) glycerol, 0.1 mM ethylenediaminetetraacetic acid (EDTA), 1 mM DTT] + 100 mM NaCl for 10 column volumes (CV), then 10 CVs of TGED + 200 mM NaCl to remove contaminating RNases. The presence of RNases was detected using the RNaseAlert QC system (Ambion). The final purified RNAP was buffer-exchanged into storage buffer (10 mM Tris-HCl, pH 8.0, 150 mM KCl, 1 mM $MgCl_2$, 2 mM DTT) and concentrated to 20 mg/mL by centrifugal filtration (Millipore) and stored at -80°C until use.

Preparation of RNase-free Eco wild-type (WT) σ^{70} and Δ1.1σ^{70}. Full-length and Δ1.1σ^{70} were cloned from a pET21 plasmid into a pETsumo vector. The pETsumo expression plasmids were transformed into *Eco* BL21 (DE3) cells (Invitrogen) and grown overnight on LB-agar plates containing 50 μg/mL kanamycin (KAN). Single colonies were used to inoculate 50 mL LB medium containing 50 μg/mL KAN and grown overnight at 37°C. Each 2-L flask of Luria broth (LB) (50 μg/mL KAN) was inoculated with 10 mL of cells, induced at OD_{600} = 0.6-0.8 by the addition of IPTG to a final concentration of 1 mM, then grown for 1 hr at 30°C. Cells were collected by centrifugation, resuspended in lysis buffer [20 mM Tris-HCl, pH 8.0, 500 mM NaCl, 0.1 mM EDTA, 5% (v/v) glycerol, and 1 mM 2-mercaptoethanol (BME)] + 5 mM imidazole and lysed in a continuous-flow French press (Avestin). The lysate was cleared by centrifugation at 27,000 × g for 30 min, then loaded onto a Hitrap IMAC HP column (GE Healthcare), and eluted with lysis buffer + 250 mM imidazole. The fractions containing full-length or Δ1.1σ^{70} were pooled, dialyzed overnight into lysis buffer + 25 mM imidazole in the presence of ULP1 SUMO protease (1 mg protease/50 mg of full-length or Δ1.1σ^{70}). The cleaved full-length or Δ1.1σ^{70} were loaded onto Hitrap IMAC HP columns (GE Healthcare) to remove His_6-SUMO tag. Flow-through from a subtractive IMAC column containing σ^{70} was dialyzed

into TGED + 0.2 M NaCl. The protein was further purified on a Hitrap Heparin HP column (GE Healthcare), eluted using a gradient to TGED + 1 M NaCl to remove contaminating RNases. Fractions were assayed for RNase activity using RNaseAlert QC system (Ambion). Fractions containing no RNases were pooled and further purified by gel filtration chromatography on Superdex 200 (GE Healthcare; equilibrated with TGED + 0.5 M NaCl). Purified WT σ^{70} or $\Delta 1.1\sigma^{70}$ were buffer-exchanged into storage buffer and concentrated to 10 mg/mL by centrifugal filtration (Millipore) and stored at -80°C.

6S RNA synthesis and purification (for cryo-EM, footprinting, and transcription assays). A mutant T7 RNAP that dissociates more rapidly from the 3′ end of run-off transcription templates and therefore minimizes non-templated addition of nucleotides to the RNA transcript 3′ end (T7 RNAP∗) was purified from plasmid pRC9 [a generous gift from W.T. McAllister (SUNY Health Science Center at Brooklyn)] as previously described (He et al., 1997). DNA templates for *in vitro* T7 RNAP synthesis of *Eco* 6S RNA derivatives (6S RNAΔCS and 6S RNA∗) were generated by gene synthesis (GenScript) into a pUC57 vector with flanking EcoRI and BamHI restriction sites. Plasmids were maxi-prepped (QIAGEN), digested with PpuMI (NEB) and gel purified (QIAGEN) for *in vitro* T7 RNAP transcription. *In vitro* transcription of 6S RNA and derivatives was performed in 30 mM Tris-HCl, pH 8.0, 10 mM DTT, 2 mM spermidine, 12.5 mM $MgCl_2$, 20 mM each ribonucleotide (rNTP), 0.01 mg/mL T7 RNAP∗, 0.2 U/μL SUPERaseIn (Ambion), 1 mU/μL inorganic pyrophosphatase (Sigma), and 70 ng/μL linearized plasmid DNA. After incubating at 37°C for three hours, the reaction was directly applied to a HiTrap Q HP column (GE Healthcare Life Sciences) pre-equilibrated with 20 mM HEPES, pH 7.5, 300 mM NaCl (Koubek et al., 2013). The following salt gradients (with 20 mM HEPES, pH 7.5) were used to purify 6S RNAs from the Q column: 300 mM NaCl (5CV), 300 mM-500 mM NaCl (5CV), 500 mM-800 mM NaCl (25CV), 800 mM-1 M NaCl (5CV), 1 M NaCl (5CV). The 6S RNAs elute within the 500 mM-800 mM NaCl gradient. RNA was precipitated with two volumes of 1:1 ethanol:isopropanol and separated through centrifugation. The RNA pellet was resuspended in stabilization buffer (10 mM Tris-HCl, pH 8.0, 1 mM $MgCl_2$).

6S RNA construction, synthesis, and purification (for mutational analysis, crosslinking, binding experiments). WT *Eco* 6S RNA and 6S RNA mutants/derivatives for mutational analysis, BPA crosslinking, binding experiments were synthesized and purified, courtesy of K. M. Wassarman (University of Wisconsin-Madison) using previously established protocols (Chen et al., 2017; Trotochaud and Wassarman, 2005).

RNase and hydroxyl radical footprinting. Footprinting assays performed by A. Feklistov from the Darst-Campbell Lab (The Rockefeller University) (Chen et al., 2017).

Preparation of Eco 6S RNA-ΔαCTD-E-Δ1.1σ^{70} (6S-Eσ^{70}) complex for cryo-EM.* Purified ΔαCTD-E and $\Delta 1.1\sigma^{70}$ were mixed at a 1:1.5 molar ratio and incubated for 15 min at 37°C. The assembled holoenzyme was injected into a Superose6 Increase column (GE Healthcare) equilibrated with 10 mM Tris-HCl, pH 8.0, 150 mM KCl, 1 mM $MgCl_2$, 5 mM DTT. The eluted protein was concentrated by centrifugal filtration (VivaScience) to 4 mg/mL protein concentration. 6S RNA* was added in a 1.2 molar excess (compared with holoenzyme) and incubated at room temperature for 15 min. The sample was mixed with CHAPSO (final concentration 8 mM; Chen et al., 2019a) prior to grid preparation. C-flat CF-1.2/1.3 400 mesh

copper grids were glow-discharged for 10 s prior to the application of 3.5 µL of the sample (2.0 - 3.0 mg/ml protein concentration). The grids were plunge-frozen in liquid ethane using a CP3 (Gatan) with ~80% chamber humidity at 22°C.

Single-particle cryo-EM of Eco 6S-Eσ^{70}. The grids were imaged using a 300 keV Titan Krios (FEI) equipped with a K2 Summit direct electron detector (Gatan). Images were recorded with Serial EM (Mastronarde, 2005) in super-resolution counting mode with a super-resolution pixel size of 0.65 Å and a defocus range of -1 µm to -3.5 µm. Data were collected with a dose of 10 electrons/physical pixel/s (1.3 Å pixel size at the specimen). Images were recorded with a 15 s exposure and 0.3 s subframes (50 total frames) to give a total dose of 90 electrons/Å^2. Dose-fractionated subframes were 2 × 2 binned (giving a pixel size of 1.3 Å), aligned and summed using Unblur (Grant and Grigorieff, 2015). The contrast transfer function was estimated for each summed image using CTFFIND4 (Rohou and Grigorieff, 2015). From the summed images, approximately 5,000 particles were manually picked and subjected to 2D classification in RELION (Scheres, 2012). Projection averages of the most populated classes were used as templates for automated picking in RELION. Autopicked particles were manually inspected, then subjected to 2D classification in RELION specifying 100 classes. Poorly populated classes were removed, resulting in a dataset of 808,484 particles. These particles were individually aligned across movie frames using direct-detector-align_lmbfgs software (Rubinstein and Brubaker, 2015). Aligned particles were 3D auto-refined in RELION using a model of *Eco* Eσ^{70} (PDB ID 4LK1; Bae et al., 2013) low-pass filtered to 60 Å resolution using EMAN2 (Tang et al., 2007) as an initial 3D template. 3D classification into eight classes was performed on the particles using the refined model and alignment angles. Among the 3D classes, the best-resolved class, containing 362,926 particles, was 3D auto-refined and post-processed in RELION. Local resolution calculations were performed using blocres (Cardone et al., 2013).

Model building and refinement. To build an initial model of the protein components of the complex, *Eco* Eσ^{70} (PDB ID 4LK1; Bae et al., 2013) was manually fit into the cryo-EM density map using UCSF Chimera (Pettersen et al., 2004) and real-space refined using PHENIX (Adams et al., 2010). In the real-space refinement, domains of RNAP were rigid-body refined, then subsequently refined with secondary structure restraints. A model of 6S RNA was generated using a combination of manual building using Coot (Emsley and Cowtan, 2004) and Rcrane (Keating and Pyle, 2010) as well as predicted structures (for segments of the RNA upstream of UB2) from MCfold (Parisien and Major, 2008) that matched the cryo-EM density. The model was real-space refined with the assistance of ERRASER (Chou et al., 2013, 2016).

BPA Protein/RNA crosslinking and crosslink mapping. Eco RNAP/6S RNA BPA crosslink mapping performed by K. M. Wassarman (University of Wisconsin-Madison) using previously established protocols (Chen et al., 2017; Wassarman and Storz, 2000; Winkelman et al., 2015).

Native gel electrophoresis binding assay. 6S RNA binding assays performed by K. M. Wassarman (University of Wisconsin-Madison) using previously established protocols (Chen et al., 2017; Wassarman and Saecker, 2006).

Construction and purification of σ mutations. Using the WT *Eco* σ^{70} pETsumo construct previously described, point mutations of σ^{70} were generated using site-directed mutagenesis. The

mutant σ^{70} proteins were purified using the same procedure as WT σ^{70}. Point mutants of *Eco* σ^{S} were made using a pET29a construct containing WT σ^{S} and standard polymerase chain reaction (PCR) mutagenesis. WT and mutant σ^{S} proteins were purified using the refolding method previously described with the exclusion of gel filtration (Nguyen and Burgess, 1996).

Abortive transcription initiation assays. DNA promoter templates (T7A1 and DPS100; (Grainger et al., 2008) were amplified using standard PCR, then sodium acetate/isopropanol-precipitated and gel-purified (QIAGEN). *Eco* full-length or $\Delta 1.1\sigma^{70}$ or σ^{S} (2.5 pmol) was added to *Eco* ΔαCTD-E (0.5 pmol) in transcription buffer (40 mM HEPES, pH 7.5, 100 mM KCl, 0.5 mM $MgCl_2$, 1 mM DTT) and incubated at 37°C for 10 min to form holoenzyme. 6S RNA or T7A1 (1 pmol) templates were each added and incubated at 37°C for 5 min to allow for binding. To initiate transcription for T7A1 and 6S RNA, 0.1 mM ApU dinucleotide primer (Trilink), 40 μM CTP, and 0.2 μCi α-^{32}P-CTP were added together with 4.5 mM $MgCl_2$. To initiate transcription for DPS100, the substrates 0.1 mM GpU dinucleotide primer (Trilink), 40 μM UTP, and 0.2 μCi α-^{32}P-UTP were added together with 4.5 mM $MgCl_2$. The final reaction volumes were 10 μL. The reactions were incubated at 37°C for 5 min and then quenched with 2X stop buffer (1X TBE, pH 8.3, 7 M urea, 1% SDS, 10 mM EDTA, pH 8.0, 0.05% bromophenol blue). Samples were analyzed by denaturing polyacrylamide gel electrophoresis (PAGE) and quantified by phosphoimagery using the program ImageQuant (GE Healthcare).

8.3 Chapter 4

Protein expression and purification. Over-expression plasmids for WT *Mycobacterium tuberculosis* (*Mtb)* RNAP and *Mtb* RNAP mutants were provided through the courtesy of M. Palka and R.A. Mooney from the Landick Lab (University of Wisconsin–Madison) (Boyaci et al., 2018). *Mtb/*M*ycobacterium smegmatis* (*Msm*) RNAP, $\sigma^{A,}$ and RbpA were purified by H. Boyaci and M. Lilic from the Darst-Campbell Lab (The Rockefeller University) using previously established protocols (Davis et al., 2015; Hubin et al., 2015, 2017b). *Eco* core RNAP and *Eco* σ^{70} were purified using previously described methods (Bae et al., 2013; Twist et al., 2011).

In vitro transcription assays. Transcription assays were performed by H. Boyaci from the Darst-Campbell Lab (The Rockefeller University) (Boyaci et al., 2018).

Agar disk diffusion assay. Zone of inhibition assays were performed by E. A. Campbell from the Darst-Campbell Lab (The Rockefeller University) (Boyaci et al., 2018).

Preparation of Fdx/RbpA/σ^{A}-holoenzyme (σ^{A}-holo) complexes for cryo-EM. Mtb RbpA/σ^{A}-holo (0.5 ml of 5 mg/ml) was injected into a Superose 6 Increase column (GE Healthcare Life Sciences) equilibrated with 20 mM Tris-HCl pH 8.0, 150 mM K-Glutamate, 5 mM $MgCl_2$, 2.5 mM DTT. The peak fractions of the eluted protein were concentrated by centrifugal filtration (EMD-Millipore) to 6 mg/mL protein concentration. Fidaxomicin (Fdx) (when used) was added at 100 μM and upstream-fork DNA (us-fork; Integrated DNA Technologies) was added to 20 μM. The samples were incubated on ice for 15 min and then CHAPSO (Chen et al., 2019a) was added to the sample for a final concentration of 8 mM prior to grid preparation. C-flat CF-1.2/1.3-4Au 400 mesh gold grids (Protochips) were glow-discharged for 20 s prior to the application of 3.5 μl of the sample (4.0-6.0 mg/ml protein concentration). After blotting for 3–

4.5 s, the grids were plunge-frozen in liquid ethane using an FEI Vitrobot Mark IV (FEI) with 100% chamber humidity at 22°C.

Cryo-EM data acquisition and processing of Mtb RbpA/σ^A-holo. The grids were imaged using a 200 keV Talos Arctica (FEI) equipped with a K2 Summit direct electron detector (Gatan). Images were recorded with Serial EM (Mastronarde, 2005) in super-resolution counting mode with a super-resolution pixel size of 0.75 Å and a defocus range of 0.8 μm to 2.4 μm. Data were collected with a dose of 8 electrons/pixel/s. Images were recorded over a 15 s exposure using 0.3 s subframes (50 total frames) to give a total dose of 53 electrons/Å^2. Dose-fractionated subframes were 2 × 2 binned (giving a pixel size of 1.5 Å), aligned and summed using Unblur (Grant and Grigorieff, 2015). The contrast transfer function was estimated for each summed image using Gctf (Zhang, 2016). From the summed images, Gautomatch (developed by K. Zhang, MRC Laboratory of Molecular Biology, Cambridge, UK, http://www.mrc-lmb.cam.ac.uk/kzhang/Gautomatch) was used to pick particles with an auto-generated template. Autopicked particles were manually inspected, then subjected to 2D classification in RELION (Scheres, 2012) specifying 100 classes. Poorly populated classes were removed, resulting in a dataset of 289,154 particles. These particles were individually aligned across movie frames and dose-weighted using direct-detector-align_lmbfgs software to generate 'polished' particles (Rubinstein and Brubaker, 2015). A subset of the dataset was used to generate an initial model of the complex in cryoSPARC (ab-initio reconstruction). 'Polished' particles were 3D auto-refined in RELION using this ab-initio 3D template (low-pass filtered to 60 Å-resolution). RELION 3D classification into two classes was performed on the particles using the refined map and alignment angles. Among the 3D classes, the best-resolved class, containing 87,657 particles, was 3D auto-refined and post-processed in RELION. The overall resolution of this class was 6.9 Å (before post-processing) and 5.2 Å (after post-processing). Subsequent 3D classification did not improve the resolution of this class.

Cryo-EM data acquisition and processing of Fdx/RbpA/σ^A-holo. The same procedure as described above for *Mtb* RbpA/σ^A-holo was used. After RELION 2D classification, poorly populated classes were removed, resulting in a dataset of 63,839 particles. In the end, the best-resolved 3D class, containing 21,115 particles, was 3D auto-refined and post-processed in RELION. The overall resolution of this class was 8.1 Å (before post-processing) and 6.5 Å (after post-processing).

Cryo-EM data acquisition and processing of Fdx/RbpA/σ^A-holo/us-fork. The grids were imaged using a 300 keV Titan Krios (FEI) equipped with a K2 Summit direct electron detector (Gatan, Warrendale, PA). Images were recorded with Leginon (Nicholson et al., 2010) in counting mode with a pixel size of 1.1 Å and a defocus range of 0.8 μm to 1.8 μm. Data were collected with a dose of 8 electrons/px/s. Images were recorded over a 10 s exposure with 0.2 s frames (50 total frames) to give a total dose of 66 electrons/Å^2. Dose-fractionated subframes were aligned and summed using MotionCor2 (Zheng et al., 2017) and subsequent dose-weighting was applied to each image. The contrast transfer function was estimated for each summed image using Gctf (Zhang, 2016). From the summed images, Gautomatch (developed by K. Zhang, MRC Laboratory of Molecular Biology, Cambridge, UK, http://www.mrc-lmb.cam.ac.uk/kzhang/Gautomatch) was used to pick particles with an auto-generated template. Autopicked particles were manually inspected, then subjected to 2D classification in cryoSPARC

(Punjani et al., 2017) specifying 50 classes. Poorly populated and dimer classes were removed, resulting in a dataset of 582,169 particles. A subset of the dataset was used to generate an initial model of the complex in cryoSPARC (ab-initio reconstruction). Using the ab-initio model (low-pass filtered to 30 Å-resolution), particles were 3D classified into two classes using cryoSPARC heterogenous refinement. CryoSPARC homogenous refinement was performed for each class using the class map and corresponding particles, yielding two structures with different clamp conformations: open (Fdx/RbpA/σ^A-holo/us-fork) and closed [RbpA/σ^A-holo/(us-fork)$_2$]. Two rounds of heterogenous/homogeneous refinements were performed for each class to achieve the highest resolution. The open class (Fdx/RbpA/σ^A-holo/us-fork) contained 173,509 particles with an overall resolution of 3.38 Å while the closed class [*Mtb* RNAP/σ^A/RbpA/(us-fork)$_2$] contained 171,547 particles with an overall resolution of 3.27 Å. Particle orientations of each class were plotted in cryoSPARC. FSC calculations were performed in cryoSPARC and the half-map FSC was calculated using EMAN2 (Tang et al., 2007). Local resolution calculations were performed using blocres (Cardone et al., 2013).

Model building and refinement. To build initial models of the protein components of the complex, *Msm* RbpA/σ^A-holo/us-fork structure (PDB ID 5TW1, Hubin et al., 2017b) was manually fit into the cryo-EM density maps using UCSF Chimera (Pettersen et al., 2004) and real-space refined using PHENIX (Adams et al., 2010). In the real-space refinement, domains of RNAP were rigid-body refined. For the high-resolution structures, the rigid-body refined models were subsequently refined with secondary structure restraints. A model of Fdx was generated from a crystal structure (Serra et al., 2017), edited in PHENIX REEL, and refined into the cryo-EM density. Refined models were inspected and modified in Coot (Emsley and Cowtan, 2004) according to cryo-EM maps, followed by further real-space refinement with PHENIX.

8.4 Chapter 5

TraR expression and purification. TraR was made by cloning the traR gene (Gopalkrishnan et al., 2017) in a pET28-based His_{10}-SUMO vector which allowed removal of the cleavable N-terminal His_{10}-SUMO tag with Ulp1 protease. The His_{10}-SUMO-TraR plasmid was transformed into competent *Eco* BL21(DE3) by heat shock. The cells were grown in the presence of 25 μg/mL KAN to an OD_{600} of 0.5 in a 37°C shaker. TraR expression was induced with a final concentration of 1 mM IPTG for 3 hr at 37°C. Cells were harvested by centrifugation and resuspended in 50 mM Tris-HCl, pH 8.0, 250 mM NaCl, 5 mM imidazole, 10% glycerol (v/v), 2.5 mM DTT, 10 μM $ZnCl_2$, 1 mM PMSF (Sigma-Aldrich), 1x protease inhibitor cocktail (Sigma-Aldrich). Cells were homogenized using a continuous-flow French Press (Avestin, Ottawa, ON, Canada) at 4°C and the resulting lysate was centrifuged to isolate the soluble fraction. The supernatant was loaded onto two 5 mL HiTrap IMAC HP columns (GE Healthcare) for a total CV of 10 mL. His_{10}-SUMO-TraR was eluted at 300 mM imidazole in Ni-column buffer [50 mM Tris-HCl, pH 8.0, 500 mM NaCl, 10% glycerol (v/v), 10 μM $ZnCl_2$, 2.5 mM DTT]. Peak fractions were combined, treated with UlP1 SUMO-protease overnight, and dialyzed against 20 mM Tris-HCl, pH 8.0, 5% glycerol (v/v), 0.1 mM EDTA, 500 mM NaCl, 10 μM $ZnCl_2$, 2.5 mM DTT, resulting in a final imidazole concentration of 25 mM. The ULP1-cleaved sample was loaded onto one 5 mL HiTrap IMAC HP column to remove His_{10}-SUMO-tag along with any remaining uncut TraR. Tagless TraR was collected in the flowthrough and concentrated by centrifugal filtration (Amicon Ultra, EMD Millipore). The sample was purified in a final step

on a HiLoad 16/60 Superdex 200 column (GE Healthcare). Purified TraR was concentrated to 16 mg/mL by centrifugal filtration, flash-frozen in liquid N_2, and stored at -80°C.

Eco His_{10}-PPX-RNAP core (with full-length α-subunits) expression and purification. A pET-based plasmid overexpressing each subunit of RNAP (full-length α, β, ω) as well as β'-PPX-His_{10} (PPX; PreScission protease site, LEVLFQGP, GE Healthcare) was co-transformed with a pACYCDuet-1 plasmid containing *Eco* rpoZ into *Eco* BL21(DE3). Protein expression and purification of this RNAP construct are similar to those of RNase-free ΔαCTD-E **(Section 8.2)** with a few modifications: (1) The PPX site was not cleaved and (2) a second HiTrap IMAC was not used since the His_{10}-PPX tag was not removed (Chen et al., 2019b).

Eco His_6-SUMO-σ^{70} expression and purification. Eco His_6-SUMO-σ^{70} was expressed and purified as previously described in **Section 8.2** with one modification: (1) the ion-exchange step using Hitrap Heparin HP was omitted (Chen et al., 2019b).

Preparation of $E\sigma^{70}$ for cryo-EM. $E\sigma^{70}$ was formed by mixing purified RNAP and 2.5-fold molar excess of σ^{70} and incubating for 20 min at 37°C. $E\sigma^{70}$ was purified on a Superose 6 Increase 10/300 GL column (GE Healthcare) in gel filtration buffer (10 mM Tris-HCl, pH 8.0, 200 mM KCl, 5 mM $MgCl_2$, 10 μM $ZnCl_2$, 2.5 mM DTT). The eluted $E\sigma^{70}$ was concentrated to ~10 mg/mL (~21 μM) by centrifugal filtration (Amicon Ultra).

Preparation of TraR-$E\sigma^{70}$ for cryo-EM. $E\sigma^{70}$ was formed by mixing purified RNAP and a 2-fold molar excess of σ^{70} and incubating for 15 min at room temperature. $E\sigma^{70}$ was purified over a Superose 6 Increase 10/300 GL column in gel filtration buffer. The eluted $E\sigma^{70}$ was concentrated to ~5.0 mg/mL (~10 μM) by centrifugal filtration. Purified TraR was added (5-fold molar excess over RNAP) and the sample was incubated for 15 min at room temperature. An rrnB P1 promoter fragment (Integrated DNA Technologies) was added (2-fold molar excess over RNAP) and the sample was incubated for a further 15 min at room temperature. The *rrnB* P1 promoter fragment did not bind to TraR-$E\sigma^{70}$ under the cryo-EM grid preparation conditions - the subsequent structural analyses did not reveal any evidence of promoter binding.

Preparation of rpsT P2-RPo for cryo-EM. $E\sigma^{70}$ was prepared as described for TraR-$E\sigma^{70}$, but after the size exclusion chromatography, the complex was concentrated to ~10 mg/mL (~20 μM) by centrifugal filtration. Duplex *rpsT* P2 promoter fragment (-60 to +25, IDT) was added to the concentrated $E\sigma^{70}$ to 3-fold molar excess. The sample was incubated for 20 mins at room temperature prior to cryo-EM grid preparation.

Cryo-EM grid preparation. CHAPSO (Anatrace) was added to the samples to a final concentration of 8 mM (Chen et al., 2019a). The final buffer condition for all the cryo-EM samples was 10 mM Tris-HCl, pH 8.0, 100 mM KCl, 5 mM $MgCl_2$, 10 μM $ZnCl_2$, 2.5 mM DTT, 8 mM CHAPSO. C-flat holey carbon grids (CF-1.2/1.3-4Au) were glow-discharged for 20 s prior to the application of 3.5 μL of the samples. Using a Vitrobot Mark IV (FEI), grids were blotted and plunge-froze into liquid ethane with 100% chamber humidity at 22°C.

Acquisition and processing of TraR-$E\sigma^{70}$ cryo-EM dataset. Grids were imaged using a 300 keV Krios (FEI) equipped with a K2 Summit direct electron detector (Gatan, Pleasanton, CA).

Datasets were recorded with Serial EM (Mastronarde, 2005) with a pixel size of 1.3 Å over a defocus range of 0.8 μm to 2.4 μm. Movies were recorded in counting mode at eight electrons/physical pixel/second in dose-fractionation mode with subframes of 0.3 s over a 15 s exposure (50 frames) to give a total dose of 120 electrons/physical pixel. Dose-fractionated movies were gain-normalized, drifted-corrected, summed, and dose-weighted using MotionCor2 (Grant and Grigorieff, 2015; Zheng et al., 2017). CTFFIND4 (Rohou and Grigorieff, 2015) was used for contrast transfer function estimation. Particles were picked using Gautomatch (http://www.mrc-lmb.cam.ac.uk/kzhang/) using a 2D template. Picked particles were extracted from the dose-weighted images with RELION (Zivanov et al., 2018) using a box size of 256 pixels. Two TraR-Eσ^{70} datasets were collected: dataset 1 consisted of 1546 motion-corrected images with 631,880 particles and dataset 2 consisted of 2132 motion-corrected images with 378,987 particles. The particles from each dataset were curated using RELION 3D classification (N = 3) using a cryoSPARC ab-initio reconstruction (Punjani et al., 2017) generated from a subset of the particles. The highest resolution classes from each dataset were subjected to RELION 3D auto-refinement resulting in a 4.69 Å resolution map from dataset 1 and a 4.38 Å resolution map from dataset 2. Refinement metadata and post-processing were used as inputs for RELION CTF refinement and Bayesian Polishing (Zivanov et al., 2018). The polished particles from both datasets were combined, resulting in 372,670 particles. The particles were aligned using RELION 3D auto-refinement resulting in a consensus map with a nominal resolution of 3.62 Å. Using the refinement parameters, subtractive 3D classification (N = 3) was performed on the particles by subtracting density outside of β'Si3 and classifying in a mask around β'Si3 (Bai et al., 2015b). Classification revealed three distinct β'Si3 dispositions. Local refinement metadata for TraR-Eσ^{70}(I) and TraR-Eσ^{70}(II) was used for RELION multi-body refinements to examine clamp motions (Nakane et al., 2018). Local resolution calculations were performed using blocres and blocfilt from the Bsoft package (Cardone et al., 2013).

Acquisition and processing of Eσ^{70} cryo-EM dataset. The Eσ^{70} image acquisition and processing were the same as for TraR-Eσ^{70} except with the following differences. Grids were imaged using a 200 keV Talos Arctica (FEI) equipped with a K2 Summit direct electron detector. Datasets were recorded with a pixel size of 1.3 Å over a defocus range of -1.0 μm to -2.5 μm. Movies were recorded in counting mode at 8.4 electrons/physical pixel/second in dose-fractionation mode with subframes of 0.2 s over a 10 s exposure (50 frames) to give a total dose of 84 electrons/physical pixel. Picked particles were extracted from the dose-weighted images in RELION (Scheres, 2012) using a box size of 200 pixels. The Eσ^{70} dataset consisted of 3,548 motion-corrected images with 1,387,166 particles. A subset of the particles was subjected to cryoSPARC ab-initio reconstruction (Punjani et al., 2017) to generate a 3D template for 3D classifications in cryoSPARC and 3D refinements in RELION (Scheres, 2012). Particles were split into two groups (1st group: particles from images 1-2,000; 2nd group: particles from images 2001–3548. Particles from each group were curated using cryoSPARC heterogeneous refinement (N = 3) resulting in a subset of 479,601 particles for the first group and 329,293 particles for the second group. Curated particles were combined and a consensus refinement was performed in RELION using the cryoSPARC generated initial model resulting in a map with a nominal resolution of 4.54 Å (without post-processing). Particles from this refinement were further analyzed using RELION multi-body refinement (Nakane et al., 2018). Additionally, particles were further curated using RELION 3D classification (N = 3) without alignment. Classification revealed two lower resolution class and a higher resolution class. The higher resolution class

containing 358,725 particles was RELION 3D auto-refined and subjected to RELION CTF refinement and RELION Bayesian Polishing (Zivanov et al., 2018). After polishing, particles were refined to a nominal resolution of 4.05 Å after RELION post-processing.

Acquisition and processing of rpsT P2-RPo cryo-EM dataset. The *rpsT* P2-RPo cryo-EM image acquisition and processing were the same as for TraR-Eσ^{70} except with the following differences. The imaging defocus range was -0.5 μm to -2.5 μm. Movies were recorded in super-resolution mode at eight electrons/physical pixel/second in dose-fractionation mode with subframes of 0.2 s over a 10 s exposure (50 frames) to give a total dose of 80 electrons/physical pixel. The *rpsT* P2-RPo dataset consisted of 6912 motion-corrected images with 973,481 particles. In RELION (Scheres, 2012), a consensus refinement was performed using the extracted particles and a cryoSPARC-generated (Punjani et al., 2017) initial model resulting in a 4.62 Å resolution map. Using the refinement parameters, 3D classification (N = 2) was performed on the particles without alignment. Classification revealed a lower resolution class and a higher resolution class, the latter with 370,965 particles with a nominal resolution of 4.38 Å after RELION 3D auto-refinement. Refinement metadata and post-processing were used as inputs for RELION CTF refinement and RELION Bayesian Polishing (Zivanov et al., 2018). Local refinement metadata for these particles was used for RELION multi-body refinements to examine clamp motions (Nakane et al., 2018). Subsequent 3D classification (N = 3) was used to further classify the polished particles resulting in one junk class and two high-resolution classes. The highest resolution reconstruction (3.43 Å) contained 289,679 particles.

Model building and refinement of cryo-EM structures. To build initial models of the protein components of the complexes, a crystal structure of *Eco* Eσ^{70} [PDB ID 4LJZ, with σ^{70}1.1 from 4LK1; Bae et al., 2013)] was manually fit into the cryo-EM density maps using UCSF Chimera (Pettersen et al., 2004) and manually adjusted using Coot (Emsley and Cowtan, 2004). For TraR-Eσ^{70}, σ^{70}1.1 from PDB ID 4LK1 (Bae et al., 2013) and TraR from PDB ID 5W1S (Molodtsov et al., 2018) were also added. For *rpsT* P2-RPo, the promoter DNA was manually added. Appropriate domains of each complex were rigid-body refined, then subsequently refined with secondary structure and nucleic acid restraints using PHENIX real-space refinement (Adams et al., 2010).

In vitro transcription assays. TraR and variants were purified as described previously (Gopalkrishnan et al., 2017). WT and variant RNAPs were purified as described previously (Ross et al., 2016). The *Eco* Δ1.1σ^{70} was expressed and purified as described previously (**Section 8.2**; Chen et al., 2017). *In vitro* transcription assays, site-directed mutagenesis, and TraR-RNAP binding assays were performed by S. Gopalkrishnan and A. Y. Chen from the Gourse Lab (University of Wisconsin–Madison) using procedures that were carried out exactly as previously described (Chen et al., 2019b; Gopalkrishnan et al., 2017).

Flux calculator calculations. S. A. Darst (The Rockefeller University) modeled effects of RNAP mutants (Δ1.1σ^{70}) or TraR on an energy diagram for transcription initiation using a kinetic flux calculator (Galburt, 2018). Details for these calculations are described in Chen et al., 2019b.

8.5 Chapter 6

Protein expression and purification. Eco RNAP (harboring full-length α-subunits), σ^{70}, and TraR were purified as described previously (**Sections 8.2, 8.4**; Chen et al., 2019b).

Native mass spectrometry analysis. Mass spectrometry experiments were performed by P. D. B. Olinares from the Chait Lab (The Rockefeller University) (Chen et al., in press).

KMnO4 and DNase I footprinting. Footprinting data were performed by A. Y. Chen, S. Gopalkrishnan, M. F. Maloney, W. Ross and J. T. Winkelman from the Gourse Lab (University of Wisconsin-Madison) (Chen et al., in press).

In vitro transcription assays. Transcription assays were performed courtesy of S. Gopalkrishnan from the Gourse Lab (University of Wisconsin–Madison) (Chen et al., in press). *In vitro* transcription was carried out on supercoiled templates as described (Gopalkrishnan et al., 2017).

Preparation of TraR-Eσ^{70}-DNA complexes for cryo-EM. Eσ^{70} was formed by mixing RNAP core and a 2-fold molar excess of σ^{70} and incubating for 15 minutes at room temperature. RNAP holo was purified over a Superose 6 Increase 10/300 GL column (GE Healthcare) in gel filtration buffer (10 mM Tris-HCl, pH 8.0, 200 mM KCl, 5 mM $MgCl_2$, 10 μM $ZnCl_2$, 2.5 mM DTT). The eluted RNAP holo was concentrated to ~10.0 mg/mL (~20 μM) by centrifugal filtration (Amicon Ultra). TraR was added (5-fold molar excess over RNAP) and the sample was incubated for 15 min at room temperature. Duplex *rpsT* P2 promoter fragment (-60 to +25, Integrated DNA Technologies), either WT or *rpsT* P2*, was added to the concentrated TraR-Eσ^{70} to 3-fold molar excess. The sample was incubated for 20 min at RT prior to cryo-EM grid preparation. CHAPSO (Anatrace) was added to the samples to a final concentration of 8 mM (Chen et al., 2019a). The final buffer condition for all the cryo-EM samples was 10 mM Tris-HCl, pH 8.0, 100 mM KCl, 5 mM $MgCl_2$, 10 μM $ZnCl_2$, 2.5 mM DTT, 8 mM CHAPSO. C-flat holey carbon grids (CF-1.2/1.3-4Au, Protochips) were glow-discharged for 20 sec prior to the application of 3.5 μL of the samples. Using a Vitrobot Mark IV (Thermo Fisher Scientific Electron Microscopy, Hillsboro, OR), grids were blotted and plunge-froze into liquid ethane with 100% chamber humidity at 22ºC.

Cryo-EM data acquisition and processing of TraR-Eσ^{70}-WT rpsT P2 complexes. Grids were imaged using a 300 keV Titan Krios (Thermo Fisher Scientific Electron Microscopy) equipped with a K2 Summit direct electron detector (Gatan, Pleasanton, CA). Images were recorded with Serial EM (Mastronarde, 2005) with a pixel size of 1.3 Å over a defocus range of -0.5 μm to -3.0 μm. Movies were recorded in super-resolution mode at 8 electrons/physical pixel/s in dose-fractionation mode with subframes of 0.2 s over a 10 s exposure (50 frames) to give a total dose of 80 electrons/physical pixel. Dose-fractionated movies were gain-normalized, drift-corrected, binned, summed, and dose-weighted using MotionCor2 (Zheng et al., 2017). The contrast transfer function was estimated for each summed image using Gctf (Zhang, 2016). Gautomatch (developed by K. Zhang, MRC Laboratory of Molecular Biology, Cambridge, UK, http://www.mrc-lmb.cam.ac.uk/kzhang/Gautomatch) was used to pick particles with an auto-generated template. Picked particles were extracted from the dose-weighted images in RELION (Scheres, 2012) using a box size of 256 pixels. The TraR-Eσ^{70}-WT *rpsT* P2 dataset consisted of 5,330 motion-corrected images with 1,189,185 particles. A subset of the particles was used to

generate an initial model of the complex in cryoSPARC (ab initio reconstruction) (Punjani et al., 2017) to generate a 3D template for RELION. In RELION, a consensus refinement was performed using the extracted particles and the cryoSPARC-generated initial model resulting in a 5.5 Å map. Using the refinement parameters, 3D classification (N=2) was performed on the particles without alignment, revealing a high-resolution class with 370,441 particles (nominal resolution 3.9 Å) after RELION 3D auto-refinement and a low-resolution 'junk' class that could not be classified further. Using the refinement parameters, a subsequent 3D classification (N=2) was performed on the high-resolution particles without alignment, revealing distinct classes with different DNA configurations: Class 1a contained duplex DNA bound to RNAP while class 1b contained a transcription bubble. Subsequent 3D masked classification (N=2, without alignment) was performed on particles from class 1a using a mask around the downstream DNA, βprotrusion, and $\sigma^{70}2$. Classification revealed two distinct classes: TRPc and TRPi1. Using the refinement parameters, subtractive 3D classification (N=3) was performed on the particles from class 1b by subtracting density outside of TraR, βlobe-Si1, β'Si3, and the downstream channel, followed by classifying the remaining density with a mask (Bai et al., 2015b). Classification revealed three distinct classes: TRPi2, TpreRPo, and TRPo. After 3D classifications, the particles within each class were further processed using RELION CTF refinement and Bayesian Polishing (Zivanov et al., 2018). RELION 3D auto-refinement and post-processing of the polished particles resulted in structures with the following nominal resolutions: TRPc (3.4 Å), TRPi1 (3.4 Å), TRPi2 (3.9 Å), TpreRPo (3.5 Å), TRPo (3.7 Å). Local resolution calculations were generated using blocres and blocfilt from the Bsoft package (Cardone et al., 2013).

Cryo-EM data acquisition and processing of TraR-Eσ^{70}-rpsT P2 complexes.* Grids were imaged as for the TraR-Eσ^{70}-WT *rpsT* P2 dataset with the following exceptions: 1) The defocus range was -0.5 µm to -2.0 µm. Data were collected with a dose of 5.6 electrons/pixel/s. Images were recorded over a 15 s exposure using 0.3 s subframes (50 total frames) to give a total dose of 84 electrons/physical pixel. Dose-fractionated subframes were gain-normalized, drift-corrected, binned, summed, and dose-weighted using MotionCor2 (Zheng et al., 2017) in RELION 3.0 (Zivanov et al., 2018). The contrast transfer function was estimated for each summed image using CTFFIND4 (Rohou and Grigorieff, 2015). The TraR-Eσ^{70}-*rpsT* P2* dataset consisted of 1,500 motion-corrected images with 523,503 particles. A subset of the particles was subjected to cryoSPARC ab initio reconstruction (Punjani et al., 2017) generate a 3D template for RELION refinements and classifications. In RELION, 3D classification (N=2) was performed on the extracted particles with alignment to the cryoSPARC ab initio reconstruction. Classification revealed a low-resolution class and a high-resolution class containing 150,387 particles with a nominal resolution of 4.6 Å after RELION 3D auto-refinement. Refinement metadata and post-processing were used as inputs for RELION CTF refinement and RELION Bayesian Polishing (Zivanov et al., 2018). Polishing improved the map to a nominal resolution of 3.1 Å after RELION 3D auto-refinement. Using the refinement parameters, subtractive 3D classification (N=3) was performed on the polished particles by subtracting density outside of $\sigma^{70}1.1$, $\sigma^{70}2$, β-lobe, β-protrusion, and downstream DNA, followed by a 3D classification of the remaining density with a mask (Bai et al., 2015b). This classification revealed two distinct classes: TRPi1.5a and TRPi1.5b. Particles from the TRPi1.5a class were further processed using RELION CTF refinement and RELION Bayesian Polishing, resulting in an improved map with a nominal resolution of 3.5 Å after RELION 3D auto-refinement and post-processing. RELION

CTF refinement and RELION Bayesian Polishing did not improve the resolution of the TRPi1.5b class (nominal resolution of 3.0 Å after RELION 3D auto-refinement and post-processing).

Model building and refinement. For initial models of the complexes, the TraR-Eσ^{70} structure (PDB ID 6N57; Chen et al., 2019b) was manually fit into the cryo-EM density maps using UCSF Chimera (Pettersen et al., 2004) and real-space refined using PHENIX (Adams et al., 2010). The DNAs were mostly built de novo based on the density maps. For real-space refinement, rigid body refinement with sixteen manually-defined mobile domains was followed by all-atom and B-factor refinement with Ramachandran and secondary structure restraints. Refined models were inspected and modified in Coot (Emsley and Cowtan, 2004).

Appendix 1. Cryo-EM studies of $E\sigma^N$-*dhsU*36

A1.1. Introduction

In bacteria, binding of σ factors confer promoter specificity to the RNA polymerase (RNAP). In addition to the housekeeping σ factor, bacteria have many alternative σ's that control gene expression in response to environmental or physiological cues (Gruber and Gross, 2003). σ^N, the alternative σ factor involved in nitrogen limiting conditions, is evolutionary divergent compared to the σ^{70} family of σ's and functions using a unique mechanism to initiate transcription (Buck et al., 2000). In *Eco*, the σ^N-holoenzyme ($E\sigma^N$) recognizes distinct promoters comprised of the -24 and -12 DNA elements [e.g. *dhsU, glnB*, and *nirB (Studholme et al., 2000)*] to form transcriptionally inactive promoter complexes (Barrios et al., 1999). These closed promoter complexes require eukaryotic-like enhancer-binding proteins and ATPases in order to melt the duplex promoter DNA for transcription (Wedel and Kustu, 1995).

We wanted to investigate $E\sigma^N$ transcription initiation by first, solving a structure of a closed complex of *Eco* $E\sigma^N$ bound to the *dhsU* promoter. A recent publication outlined a 3.8 Å cryo-electron microscopy (cryo-EM) structure of $E\sigma^N$ bound to promoter DNA (Glyde et al., 2017). However, this structure has severe particle orientation bias and protein domains along with the promoter DNA are modeled erroneously due to low-resolution (Glyde et al., 2017). With new software [RELION (Zivanov et al., 2018) and cryoSPARC (Punjani et al., 2017)], I wanted to reanalyzed the cryo-EM data collected by M. Wu and C. Chiu from the Darst-Campbell Lab (The Rockefeller University) on a complex of $E\sigma^N$-*dhsU*36.

A1.2. Results

A 36 base pair fragment of the *dhsU* promoter with a two-base pair mismatch (-12 to -11) was used to capture a stable $E\sigma^N$-*dhsU*36 closed complex (Cannon et al., 2000) **(Table A1.1).** Initial binding assays and cryo-EM experiments were performed by M. Wu from the Darst-Campbell Lab (The Rockefeller University). Images were collected on a 300kV Titan Krios (FEI) in super-resolution mode (0.65 Å) with a dose of 10 electrons/pixel/second over 15 seconds. 2D and 3D reconstructions show severe particle orientation bias, consistent with the published cryo-EM structure (Glyde et al., 2017).

In order to counteract this bias, cryo-EM data was collected in the presence of cholamidopropyl]dimethylammonio)-2-hydroxy-1-propanesulfonate (CHAPSO) (Chen et al., 2019a). Although there was an improvement in the orientation distribution using CHAPSO, the majority of complexes did not have DNA bound. In addition to this class, an opened clamp class and a closed clamp class (due non-specific DNA binding in the RNAP cleft) were observed from these datasets. Glutaraldehyde cross-linking increased DNA binding to $E\sigma^N$ in the presence of CHAPSO, as seen in gel shifts and subsequently cryo-EM experiments performed by C. Chiu from the Darst-Campbell Lab (The Rockefeller University).

Using updated programs of RELION (Zivanov et al., 2018) and cryoSPARC (Punjani et al., 2017), I combined all the datasets of $E\sigma^N$-*dhsU*36 **(Table A1.2)** and processed them using the scheme outlined in **Figure A1.3A-B**. After 3D classification and refinement **(Figure A1.3B)**, the final class (closed clamp) contained 267,097 particles with an overall resolution of 3.71 Å using the Gold-standard FSC cutoff of 0.143 in RELION (Zivanov et al., 2018) after post-processing **(Figure A1.4C)**. Local resolution calculations for the final structure indicated that the central core of the structure was determined to 3.0-4.0 Å resolution (Cardone et al., 2013) **(Figure A1.4A-B)**.

Table A1.1. Sequences of *dhsU* promoters used in cryo-EM studies

	Sequences
dhsU27am_blunt (top)	5'-GGAATTGGCACGAAAATTGC**C**T**T**AACG-3'
dhsU27a_blunt (bot)	5'-CGTTA**T**T**T**GCAATTTTCGTGCCAATTCC-3'
dhsU36am_blunt (top)	5'-CCAGAAATTGGCACGAAAATTGC**C**T**T**AAATACAACG-3'
dhsU36a_blunt (bot)	5'-CGTTGTATTTA**T**T**T**GCAATTTTCGTGCCAATTTCTGG-3'

† Highlighted bases belong to the *dhsU* promoter DNA elements (Studholme et al., 2000) that are recognized by σ^N (yellow, -24 element; pink, -12 element). Bolded and underlined bases are mismatched (Cannon et al., 2000).

Table A1.2. Buffer conditions for cryo-EM datasets of Eσ^N-*dhsU*36

Date	# of images	Conditions
28SEP2015[a]	1734	20 mM Tris pH 8, 100 mM NaCl, 2 mM DTT, 5 mM $MgCl_2$
05NOV2015[a]	1636	20 mM Tris pH 8, 100 mM NaCl, 2 mM DTT, 5 mM $MgCl_2$
02DEC2015[a]	1464	20 mM Tris pH 8, 100 mM NaCl, 2 mM DTT, 5 mM $MgCl_2$, 8mM CHAPSO
29FEB2016[a]	2560	20 mM Tris pH 8, 100 mM NaCl, 2 mM DTT, 5 mM $MgCl_2$, 8mM CHAPSO
08JUN2016[a]	1207	20 mM Tris pH 8, 100 mM NaCl, 2 mM DTT, 5 mM $MgCl_2$, 8mM CHAPSO, 0.005% glutaraldehyde
14SEP2016[b]	1235	20 mM Tris pH 8, 100 mM NaCl, 2 mM DTT, 5 mM $MgCl_2$, 8mM CHAPSO, 0.005% glutaraldehyde
14APR2017	2570	20 mM Tris pH 8, 100 mM NaCl, 2 mM DTT, 5 mM $MgCl_2$, 8mM CHAPSO, 0.005% glutaraldehyde

[a] Cryo-EM sample preparation and data collection, courtesy of M. Wu from the Darst-Campbell Lab (The Rockefeller University)
[b] Cryo-EM sample preparation and data collection, courtesy of C. Chiu from the Darst-Campbell Lab (The Rockefeller University)

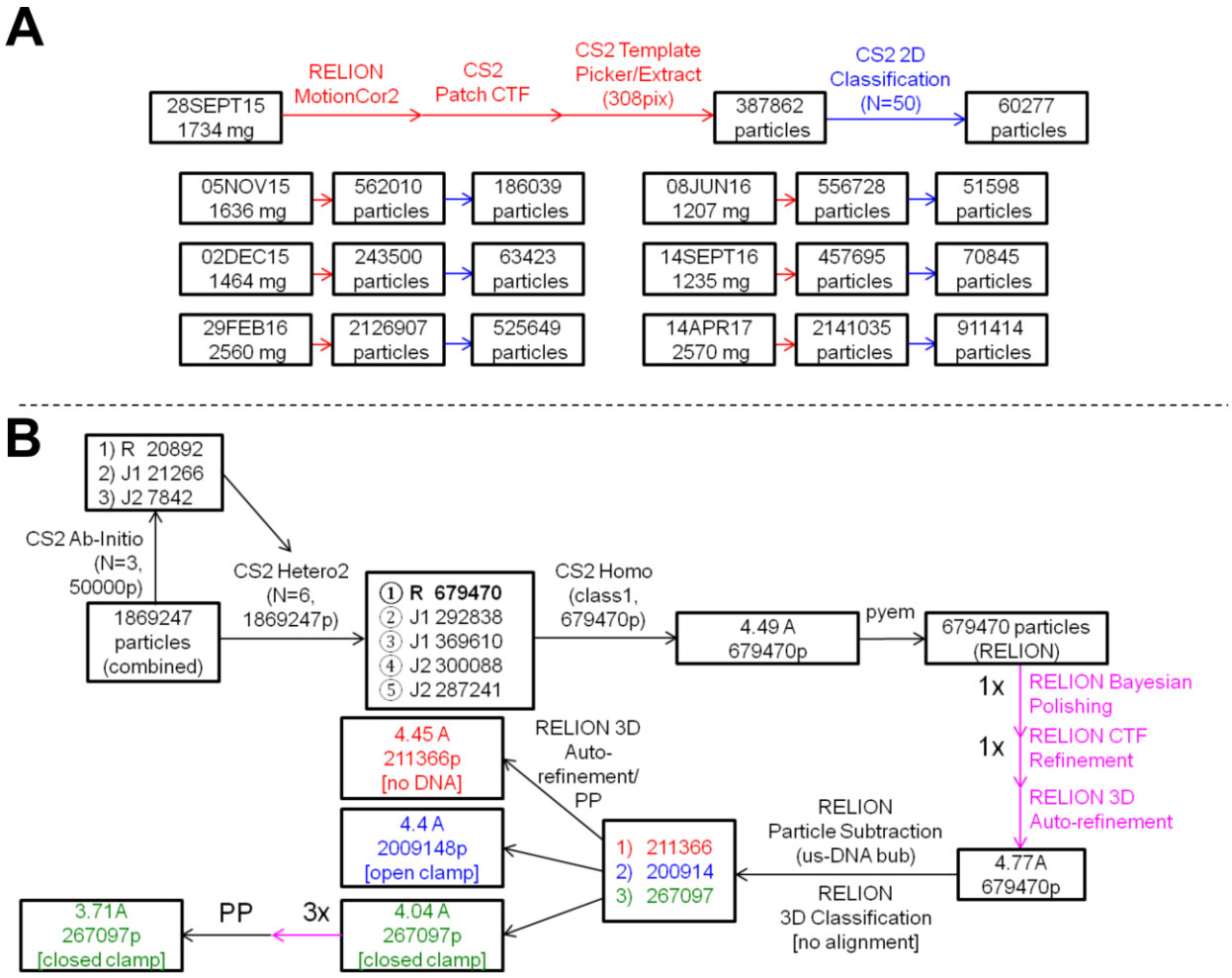


Figure A1.3. Eσ^{N}-*dhsU*36 data processing pipeline

(A-B) Flowchart showing the image processing pipeline for the cryo-EM data of *Eco* Eσ^{N}-*dhsU*36. **(A)** Dose-fractionated movies were aligned into summed images using MotionCor2 (Zheng et al., 2017), CTF-corrected using cryoSPARC2 (CS2) Patch CTF, and template picked using CS2 (Punjani et al., 2017). These steps are summarized as red arrows. To curate particles, 2D classification in CS2 was performed (blue arrows, Punjani et al., 2017). These initial processing steps were performed on all datasets **(Table A1.2)**. **(B)** 3D templates [RNAP class (R), junk class 1 (J1), junk class 2 (J2)] were generated using CS2 ab-initio reconstruction (Punjani et al., 2017). Particles were further curated using CS2 heterogeneous refinement and further processed using RELION Bayesian Polishing and CTF Refinement (Zivanov et al., 2018). RELION 3D classification (Zivanov et al., 2018) sorted three classes of Eσ^{N}-*dhsU*36: (1, red) no DNA bound; (2, blue) DNA bound with an opened clamp; (3, green) DNA bound with a closed clamp (end-bound). Further polishing and CTF refinements were performed on the closed clamp class, resulting in a final structure containing 267,097 particles with a 3.71 Å resolution after auto-refinement and post-processing in RELION (Zivanov et al., 2018).

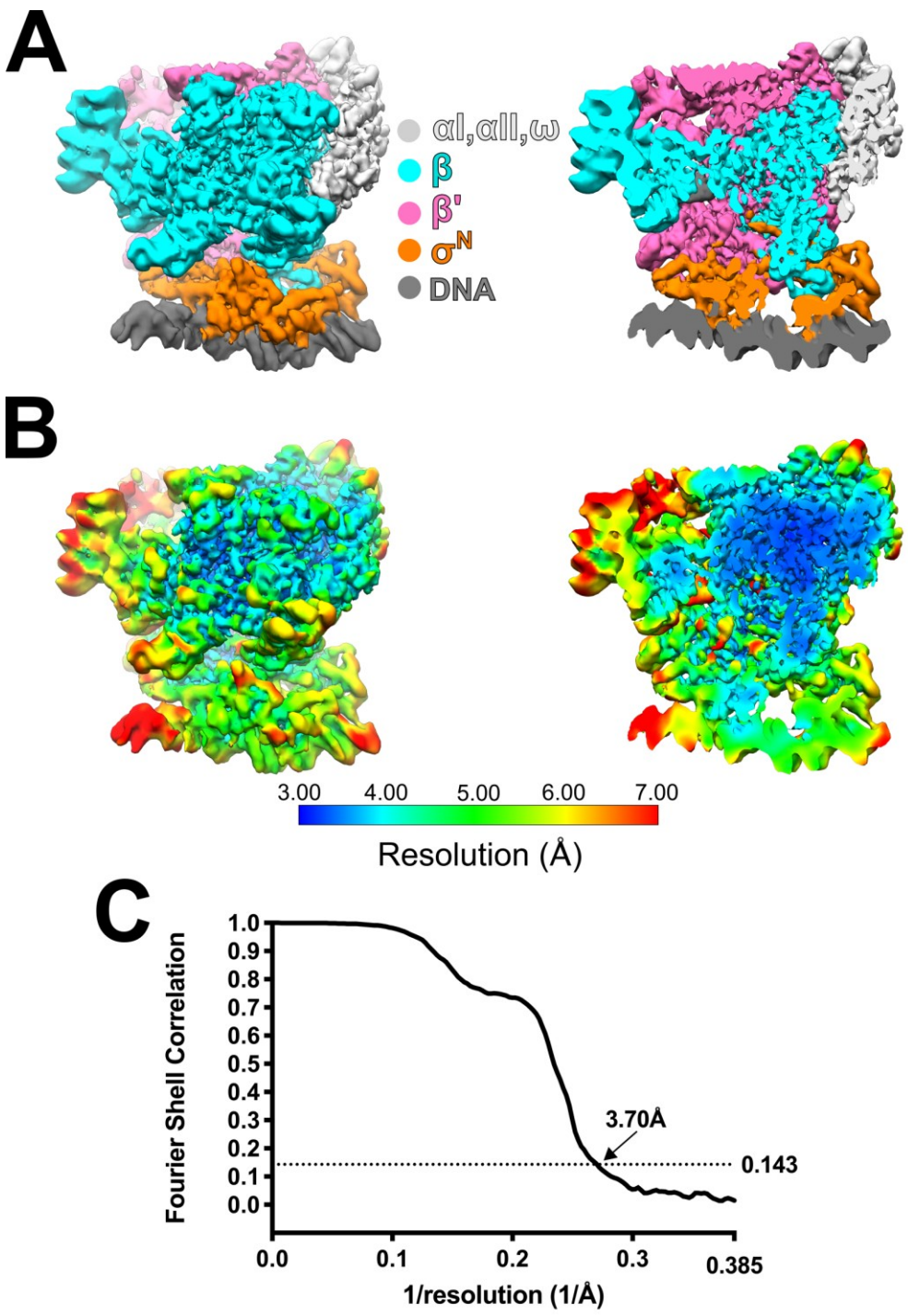


Figure A1.4. Cryo-EM structure of Eσ^{N}-*dhsU*36 (closed clamp)
(A) Two views of the cryo-EM density map colored according to the key. The right view is a cross-section through the left view. **(B)** Same views as **(A)** but colored by local resolution (Cardone et al., 2013). **(C)** Gold-standard FSC, calculated by comparing the two independently determined half-maps from RELION (Zivanov et al., 2018). The dotted line represents the 0.143 FSC cutoff.

A1.3. Future Directions

Using the 3.71 Å cryo-EM reconstruction of $E\sigma^N$-*dhsU*36, an atomic model will be built, followed by a structure-driven mutational analysis of $E\sigma^N$ DNA binding determinants. Cryo-EM structures of the activated $E\sigma^N$/promoter DNA/ATPase complexes and their intermediates have recently been published (Glyde et al., 2018). Like the published $E\sigma^N$ closed complex (PDB ID 5NSR; Glyde et al., 2017), these cryo-EM structures have significant particle orientation bias that affects the interpretability of the maps and the reliability of the built models (Glyde et al., 2018).

A1.4. Acknowledgments

I would like to thank the following people for their contributions to this appendix: C. Chiu (The Rockefeller University) assisted in cryo-EM experiments; E. A. Campbell (The Rockefeller University) provided mentorship and developed protein purification and biochemical protocols; M. Ebrahim and J. Sotiris (The Rockefeller University) helped with cryo-EM data collection; S. A. Darst (The Rockefeller University) provided mentorship and assisted in structural analyses and interpretations; T. Nixon and J. Du (Penn State University) collaborated on the ATPase aspects of this project; M. Wu (The Rockefeller University) performed binding assays and initial cryo-EM experiments.

REFERENCES.

Abascal-Palacios, G., Ramsay, E.P., Beuron, F., Morris, E., and Vannini, A. (2018). Structural basis of RNA polymerase III transcription initiation. Nature *553*, 301–306.

Adams, P.D., Afonine, P. V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., et al. (2010). PHENIX: A comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. Sect. D Biol. Crystallogr. *66*, 213–221.

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

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

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

Bae, B., Chen, J., Davis, E., Leon, K., Darst, S.A., and Campbell, E.A. (2015b). CarD uses a minor groove wedge mechanism to stabilize the RNA polymerase open promoter complex. Elife *4*, 1–19.

Bai, X.C., McMullan, G., and Scheres, S.H.W. (2015a). How cryo-EM is revolutionizing structural biology. Trends Biochem. Sci. *40*, 49–57.

Bai, X.C., Rajendra, E., Yang, G., Shi, Y., and Scheres, S.H.W. (2015b). Sampling the conformational space of the catalytic subunit of human g-secretase. Elife *4*, e11182.

Baker, N.A., Sept, D., Joseph, S., Holst, M.J., and McCammon, J.A. (2001). Electrostatics of nanosystems: application to microtubules and the ribosome. Proc. Natl. Acad. Sci. U. S. A. *98*, 10037–10041.

Barker, M.M., and Gourse, R.L. (2001). Regulation of rRNA transcription correlates with nucleoside triphosphate sensing. J. Bacteriol. *183*, 6315–6323.

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

Barne, K.A., Bown, J.A., Busby, S.J.W., and 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 J. *16*, 4034–4040.

Barrick, J.E., Sudarsan, N., Weinberg, Z., Ruzzo, W.L., and Breaker, R.R. (2005). 6S RNA is a widespread regulator of eubacterial RNA polymerase that resembles an open promoter. RNA *11*, 774–784.

Barrios, H., Valderrama, B., and Morett, E. (1999). Compilation and analysis of σ54-dependent promoter sequences. Nucleic Acids Res. *27*, 4305–4313.

Barth, M., Bryan, R.K.K., and Hegerl, R. (1989). Approximation of missing-cone data in 3D electron microscopy. Ultramicroscopy *31*, 365–378.

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

Beckmann, B.M., Burenina, O.Y., Hoch, P.G., Kubareva, E.A., Sharma, C.M., and Hartmann, R.K. (2011). In vivo and in vitro analysis of 6S RNA-templated short transcripts in Bacillus subtilis. RNA Biol. *8*, 839–849.

Beckmann, B.M., Hoch, P.G., Marz, M., Willkomm, D.K., Salas, M., and Hartmann, R.K. (2012). A pRNA-induced structural rearrangement triggers 6S-1 RNA release from RNA polymerase in Bacillus subtilis. EMBO J. *31*, 1727–1738.

Belogurov, G.A., Vassylyeva, M.N., Sevostyanova, A., Appleman, J.R., Xiang, A.X., Lira, R., Webber, S.E., Klyuyev, S., Nudler, E., Artsimovitch, I., et al. (2009). Transcription inactivation through local refolding of the RNA polymerase structure. Nature *457*, 332–335.

Benoff, B., Yang, H., Lawson, C.L., Parkinson, G., Liu, J., Blatter, E., Ebright, Y.W., Berman, H.M., and Ebright, R.H. (2002). Structural basis of transcription activation: The CAP-αCTD-DNA complex. Science *297*, 1562–1566.

Bernecky, C., Herzog, F., Baumeister, W., Plitzko, J.M., and Cramer, P. (2016). Structure of transcribing mammalian RNA polymerase II. Nature *529*, 551–554.

Blankschien, M.D., Potrykus, K., Grace, E., Choudhary, A., Vinella, D., Cashel, M., and Herman, C. (2009). TraR, a homolog of a RNAP secondary channel interactor, modulates transcription. PLoS Genet. *5*, e1000345–e1000345.

Blatter, E.E., Ross, W., Tang, H., Gourse, R.L., and Ebright, R.H. (1994). Domain organization of RNA polymerase α subunit: C-terminal 85 amino acids constitute a domain capable of dimerization and DNA binding. Cell *78*, 889–896.

Bortoluzzi, A., Muskett, F.W., Waters, L.C., Addis, P.W., Rieck, B., Munder, T., Schleier, S., Forti, F., Ghisotti, D., Carr, M.D., et al. (2013). Mycobacterium tuberculosis RNA polymerase-binding protein A (RbpA) and its interactions with sigma factors. J. Biol. Chem. *288*, 14438–14450.

Boyaci, H., Chen, J., Lilic, M., Palka, M., Mooney, R.A., Landick, R., Darst, S.A., and Campbell, E.A. (2018). Fidaxomicin jams Mycobacterium tuberculosis RNA polymerase motions needed for initiation via RbpA contacts. Elife *7*, 1–18.

Boyaci, H., Chen, J., Jansen, R., Darst, S.A., and Campbell, E.A. (2019a). Structures of an RNA polymerase promoter melting intermediate elucidate DNA unwinding. Nature *565*, 382–385.

Boyaci, H., Saecker, R.M., and Campbell, E.A. (2019b). Transcription initiation in mycobacteria: a biophysical perspective. Transcription 1–13.

Brown, A., Fernández, I.S., Gordiyenko, Y., and Ramakrishnan, V. (2016). Ribosome-dependent activation of stringent control. Nature *534*, 277–280.

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

Browning, D.F., and Busby, S.J.W. (2016). Local and global regulation of transcription initiation in bacteria. Nat. Rev. Microbiol. *14*, 638–650.

Buc, H., and McClure, W.R. (1985). Kinetics of Open Complex Formation between Escherichia coli RNA Polymerase and the lac UV5 Promoter. Evidence for a Sequential Mechanism Involving Three Steps. Biochemistry *24*, 2712–2723.

Buck, M., Gallegos, M.T., Studholme, D.J., Guo, Y., and Gralla, J.D. (2000). The bacterial enhancer-dependent sigma(54) (sigma(N)) transcription factor. J. Bacteriol. *182*, 4129–4136.

Bunge, J., Willis, A., and Walsh, F. (2014). Estimating the Number of Species in Microbial Diversity Studies. Annu. Rev. Stat. Its Appl. *1*, 427–445.

Burgess, R.R. (1969). Separation and characterization of the subunits of ribonucleic acid polymerase. J. Biol. Chem. *244*, 6168–6176.

Burgess, R.R., and Travers, A.A. (1970). Escherichia coli RNA polymerase: purification, subunit structure, and factor requirements. Fed. Proc. *29*, 1164–1169.

Burgess, R.R., Travers, A.A., Dunn, J.J., and Bautz, E.K.F. (1969). Factor stimulating transcription by RNA polymerase. Nature *221*, 43–46.

Callaci, S., Heyduk, E., and Heyduk, T. (1998). Conformational changes of Escherichia coli RNA polymerase σ70 factor induced by binding to the core enzyme. J. Biol. Chem. *273*, 32995–33001.

Callaci, S., Heyduk, E., Heyduk, T., and Doisy, E.A. (1999). Core RNA polymerase from E. coli induces a major change in the domain arrangement of the σ70 subunit. Mol. Cell *3*, 229–238.

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

Cannon, W. V., Gallegos, M.T., and Buck, M. (2000). Isomerization of a binary sigma-promoter DNA complex by transcription activators. Nat. Struct. Biol. *7*, 594–601.

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

Cartagena, A.J., Banta, A.B., Sathyan, N., Ross, W., Gourse, R.L., Campbell, E.A., and Darst, S.A. (2019). Structural basis for transcription activation by Crl through tethering of σS and RNA polymerase. Proc. Natl. Acad. Sci. U. S. A. *116*, 18923–18927.

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

Cashel, M., Hsu, L.M., and Hernandez, V.J. (2003). Changes in conserved region 3 of Escherichia coli σ70 reduce abortive transcription and enhance promoter escape. J. Biol. Chem. *278*, 5539–5547.

Cavanagh, A.T., and Wassarman, K.M. (2014). 6S RNA, a Global Regulator of Transcription in Escherichia coli, Bacillus subtilis, and Beyond. Annu. Rev. Microbiol. *68*, 45–60.

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

Cavanagh, A.T., Sperger, J.M., and Wassarman, K.M. (2012). Regulation of 6S RNA by pRNA synthesis is required for efficient recovery from stationary phase in E. coli and B. subtilis. Nucleic Acids Res. *40*, 2234–2246.

Chakraborty, S., and Rhee, K.Y. (2015). Tuberculosis drug development: History and evolution of the mechanism-based paradigm. Cold Spring Harb. Perspect. Med. *5*, 1–11.

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

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

Chen, J., Chiu, C., Gopalkrishnan, S., Chen, A.Y., Olinares, P.D.B., Saecker, R.M., Winkelman, J.T., Maloney, M.F., Chait, B.T., Ross, W., et al. (in press). Stepwise promoter melting by bacterial RNA polymerase. Mol. Cell.

Chen, J., Darst, S.A., and Thirumalai, D. (2010a). Promoter melting triggered by bacterial RNA polymerase occurs in three steps. Proc. Natl. Acad. Sci. U. S. A. *107*, 12523–12528.

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

Chen, J., Noble, A.J., Kang, J.Y., and Darst, S.A. (2019a). Eliminating effects of particle adsorption to the air/water interface in single-particle cryo-electron microscopy: Bacterial RNA polymerase and CHAPSO. J. Struct. Biol. X *1*, 100005.

Chen, J., Gopalkrishnan, S., Chiu, C., Chen, A.Y., Campbell, E.A., Gourse, R.L., Ross, W., and Darst, S.A. (2019b). E . coli F element-encoded TraR protein allosterically regulates transcription initiation by altering RNA polymerase conformation and conformational dynamics. Elife *8*, 1–81.

Chen, V.B., Arendall, W.B., Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., and Richardson, D.C. (2010b). MolProbity: All-atom structure validation for macromolecular crystallography. Acta Crystallogr. Sect. D Biol. Crystallogr. *66*, 12–21.

Chen, Z.A., Jawhari, A., Fischer, L., Buchen, C., Tahir, S., Kamenski, T., Rasmussen, M., Lariviere, L., Bukowski-Wills, J.C., Nilges, M., et al. (2010c). Architecture of the RNA polymerase II-TFIIF complex revealed by cross-linking and mass spectrometry. EMBO J. *29*, 717–726.

Cheng, Y. (2018). Single-particle cryo-EM-How did it get here and where will it go. Science *361*, 876–880.

Chlenov, M., Masuda, S., Murakami, K.S., Nikiforov, V., Darst, S.A., and Mustaev, A. (2005). Structure and function of lineage-specific sequence insertions in the bacterial RNA polymerase β′ subunit. J. Mol. Biol. *353*, 138–154.

Chou, F.C., Sripakdeevong, P., Dibrov, S.M., Hermann, T., and Das, R. (2013). Correcting pervasive errors in RNA crystallography through enumerative structure prediction. Nat. Methods *10*, 74–76.

Chou, F.C., Echols, N., Terwilliger, T.C., and Das, R. (2016). RNA structure refinement using the ERRASER-Phenix pipeline. Methods Mol. Biol. *1320*, 269–282.

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

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

D'Imprima, E., Floris, D., Joppe, M., Sánchez, R., Grininger, M., and Kühlbrandt, W. (2019). Protein denaturation at the air-water interface and how to prevent it. Elife *8*, 1–18.

Danev, R., Yanagisawa, H., and Kikkawa, M. (2019). Cryo-Electron Microscopy Methodology: Current Aspects and Future Directions. Trends Biochem. Sci. *44*, 837–848.

Darst, S.A., Kubalek, E.W., and Kornberg, R.D. (1989). Three-dimensional structure of Escherichia coli RNA polymerase holoenzyme determined by electron crystallography. Nature *340*, 730–732.

Darst, S.A., Edwards, A.M., Kubalek, E.W., and Kornberg, R.D. (1991). Three-dimensional structure of yeast RNA polymerase II at 16 Å resolution. Cell *66*, 121–128.

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

Ding Jun Jin, Burgess, R.R., Richardson, J.P., and Gross, C.A. (1992). Termination efficiency at rho-dependent terminators depends on kinetic coupling between RNA polymerase and rho. Proc. Natl. Acad. Sci. U. S. A. *89*, 1453–1457.

Dobro, M.J., Melanson, L.A., Jensen, G.J., and McDowall, A.W. (2010). Plunge freezing for electron cryomicroscopy. In Methods in Enzymology, pp. 63–82.

Dombroski, A.J., Walter, W.A., and Gross, C.A. (1993). Amino-terminal amino acids modulate σ-factor DNA-binding activity. Genes Dev. *7*, 2446–2455.

Duchi, D., Mazumder, A., Malinen, A.M., Ebright, R.H., and Kapanidis, A.N. (2018). The RNA polymerase clamp interconverts dynamically among three states and is stabilized in a partly closed state by ppGpp. Nucleic Acids Res. *46*, 7284–7295.

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

Estrem, S.T., Gaal, T., Ross, W., and Gourse, R.L. (1998). Identification of an UP element consensus sequence for bacterial promoters. Proc. Natl. Acad. Sci. U. S. A. *95*, 9761–9766.

Estrem, S.T., Ross, W., Gaal, T., Chen, Z.W.S., Niu, W., Ebright, R.H., and Gourse, R.L. (1999). Bacterial promoter architecture: Subsite structure of UP elements and interactions with the carboxy-terminal domain of the RNA polymerase α subunit. Genes Dev. *13*, 2134–2147.

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

Feklistov, A., Barinova, N., Sevostyanova, A., Heyduk, E., Bass, I., Vvedenskaya, I., Kuznedelov, K., Merkiene, E., Stavrovskaya, E., Klimašauskas, S., et al. (2006). A Basal Promoter Element Recognized by Free RNA Polymerase σ Subunit Determines Promoter Recognition by RNA Polymerase Holoenzyme. Mol. Cell *23*, 97–107.

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

Feklístov, A., Sharon, B.D., Darst, S.A., and Gross, C.A. (2014). Bacterial Sigma Factors: A Historical, Structural, and Genomic Perspective. Annu. Rev. Microbiol. *68*, 357–376.

Frost, L.S., Ippen-Ihler, K., and Skurray, R.A. (1994). Analysis of the sequence and gene products of the transfer region of the F sex factor. Microbiol. Rev. *58*, 162–210.

Gaal, T., Ross, W., Estrem, S.T., Nguyen, L.H., Burgess, R.R., and Gourse, R.L. (2001). Promoter recognition and discrimination by EσS RNA polymerase. Mol. Microbiol. *42*, 939–954.

Galburt, E.A. (2018). The calculation of transcript flux ratios reveals single regulatory mechanisms capable of activation and repression. Proc. Natl. Acad. Sci. U. S. A. *115*, E11604–E11613.

Glaeser, R.M. (2015). How good can cryo-EM become? Nat. Methods *13*, 28–32.

Glaeser, R.M., and Han, B.-G. (2017). Opinion: hazards faced by macromolecules when confined to thin aqueous films. Biophys. Reports *3*, 1–7.

Glyde, R., Ye, F., Darbari, V.C., Zhang, N., Buck, M., and Zhang, X. (2017). Structures of RNA Polymerase Closed and Intermediate Complexes Reveal Mechanisms of DNA Opening and Transcription Initiation. Mol. Cell *67*, 106-116.e4.

Glyde, R., Ye, F., Jovanovic, M., Kotta-Loizou, I., Buck, M., and Zhang, X. (2018). Structures of Bacterial RNA Polymerase Complexes Reveal the Mechanism of DNA Loading and Transcription Initiation. Mol. Cell *70*, 1111-1120.e3.

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

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

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

Grainger, D.C., Goldberg, M.D., Lee, D.J., and Busby, S.J.W. (2008). Selective repression by Fis and H-NS at the Escherichia coli dps promoter. Mol. Microbiol. *68*, 1366–1377.

Grant, T., and Grigorieff, N. (2015). Measuring the optimal exposure for single particle cryo-EM using a 2.6 Å reconstruction of rotavirus VP6. Elife *4*.

Gries, T.J., Kontur, W.S., Capp, M.W., Saecker, R.M., and Record, M.T. (2010). One-step DNA melting in the RNA polymerase cleft opens the initiation bubble to form an unstable open complex. Proc. Natl. Acad. Sci. U. S. A. *107*, 10418–10423.

Grigorieff, N. (2007). FREALIGN: High-resolution refinement of single particle structures. J. Struct. Biol. *157*, 117–125.

Gruber, T.M., and Bryant, D.A. (1997). Molecular systematic studies of eubacteria, using sigma70-type sigma factors of group 1 and group 2. J. Bacteriol. *179*, 1734–1747.

Gruber, T.M., and Gross, C.A. (2003). Multiple Sigma Subunits and the Partitioning of Bacterial Transcription Space. Annu. Rev. Microbiol. *57*, 441–466.

Gualtieri, M., Villain-Guillot, P., Latouche, J., Leonetti, J.P., and Bastide, L. (2006). Mutation in the Bacillus subtilis RNA polymerase β′ subunit confers resistance to lipiarmycin. Antimicrob. Agents Chemother. *50*, 401–402.

Gualtieri, M., Tupin, A., Brodolin, K., and Leonetti, J.P. (2009). Frequency and characterisation of spontaneous lipiarmycin-resistant Enterococcus faecalis mutants selected in vitro. Int. J. Antimicrob. Agents *34*, 605–606.

Gusarov, I., and Nudler, E. (1999). The mechanism of intrinsic transcription termination. Mol. Cell *3*, 495–504.

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

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

Hawley, D.K., and Mcclure, W.R. (1982). Mechanism of activation of transcription initiation from the λPRM promoter. J. Mol. Biol. *157*, 493–525.

Hawley, D.K., and Mcclure, W.R. (1983). Compilation and analysis of Escherichia coli promoter DNA sequences. Nucleic Acids Res. *11*, 2237–2255.

He, B., Rong, M., Lyakhov, D., Gartenstein, H., Diaz, G., Castagna, R., McAllister, W.T., and Durbin, R.K. (1997). Rapid mutagenesis and purification of phage RNA polymerases. Protein Expr. Purif. *9*, 142–151.

He, Y., Fang, J., Taatjes, D.J., and Nogales, E. (2013). Structural visualization of key steps in human transcription initiation. Nature *495*, 481–486.

He, Y., Yan, C., Fang, J., Inouye, C., Tjian, R., Ivanov, I., and Nogales, E. (2016). Near-atomic resolution visualization of human transcription promoter opening. Nature *533*, 359–365.

Henderson, R. (1995). The Potential and Limitations of Neutrons, Electrons and X-Rays for Atomic Resolution Microscopy of Unstained Biological Molecules. Q. Rev. Biophys. *28*, 171–193.

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

Heyduk, E., Kuznedelov, K., Severinov, K., and Heyduk, T. (2006). A Consensus Adenine at Position –11 of the Nontemplate Strand of Bacterial Promoter Is Important for Nucleation of Promoter Melting. J. Biol. Chem. *281*, 12362–12369.

Hoffmann, N.A., Jakobi, A.J., Moreno-Morcillo, M., Glatt, S., Kosinski, J., Hagen, W.J.H., Sachse, C., and Müller, C.W. (2015). Molecular structures of unbound and transcribing RNA polymerase III. Nature *528*, 231–236.

Hook-Barnard, I.G., and Hinton, D.M. (2009). The promoter spacer influences transcription initiation via sigma70 region 1.1 of Escherichia coli RNA polymerase. Proc. Natl. Acad. Sci. U. S. A. *106*, 737–742.

Hsu, L.M., Zagorski, J., Wang, Z., and Fournier, M.J. (1985). Escherichia coli 6S RNA gene is part of a dual-function transcription unit. J. Bacteriol. *161*, 1162–1170.

Hubin, E.A., Tabib-Salazar, A., Humphrey, L.J., Flack, J.E., Olinares, P.D.B., Darst, S.A., Campbell, E.A., and Paget, M.S. (2015). Structural, functional, and genetic analyses of the actinobacterial transcription factor RbpA. Proc. Natl. Acad. Sci. U. S. A. *112*, 7171–7176.

Hubin, E.A., Lilic, M., Darst, S.A., and Campbell, E.A. (2017a). Structural insights into the mycobacteria transcription initiation complex from analysis of X-ray crystal structures. Nat. Commun. *8*, 16072.

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

Ishihama, A. (2000). Functional Modulation of Escherichia Coli RNA Polymerase . Annu. Rev. Microbiol. *54*, 499–518.

Iyer, L.M., Koonin, E. V, and Aravind, L. (2003). Evolutionary connection between the catalytic subunits of DNA-dependent RNA polymerases and eukaryotic RNA-dependent RNA polymerases and the origin of RNA polymerases. BMC Struct. Biol. *3*, 1.

Jensen, D., Manzano, A.R., Rammohan, J., Stallings, C.L., and Galburt, E.A. (2019). CarD and RbpA modify the kinetics of initial transcription and slow promoter escape of the Mycobacterium tuberculosis RNA polymerase. Nucleic Acids Res. *47*, 6685–6698.

Jeon, Y.H., Negishi, T., Shirakawa, M., Yamazaki, T., Fujita, N., Ishihama, A., and Kyogoku, Y. (1995). Solution structure of the activator contact domain of the RNA polymerase α subunit. Science *270*, 1495–1497.

Kadesch, T.R., and Chamberlin, M.J. (1982). Studies of in vitro transcription by calf thymus RNA polymerase II using a novel duplex DNA template. J. Biol. Chem. *257*, 5286–5295.

Kadesch, T.R., Rosenberg, S., and Chamberlin, M.J. (1982). Binding of Escherichia coli RNA polymerase holoenzyme to bacteriophage T7 DNA. J. Mol. Biol. *155*, 1–29.

Kaledhonkar, S., Fu, Z., White, H., and Frank, J. (2018). Time-resolved cryo-electron microscopy using a microfluidic chip. In Methods in Molecular Biology, pp. 59–71.

Kaledhonkar, S., Fu, Z., Caban, K., Li, W., Chen, B., Sun, M., Gonzalez, R.L., and Frank, J. (2019). Late steps in bacterial translation initiation visualized using time-resolved cryo-EM. Nature *570*, 400–404.

Kang, J.Y., Olinares, P.D.B., Chen, J., Campbell, E.A., Mustaev, A., Chait, B.T., Gottesman, M.E., and Darst, S.A. (2017). Structural basis of transcription arrest by coliphage HK022 nun in an Escherichia coli RNA polymerase elongation complex. Elife *6*, 1–20.

Kang, J.Y., Mishanina, T. V., Bellecourt, M.J., Mooney, R.A., Darst, S.A., and Landick, R. (2018a). RNA Polymerase Accommodates a Pause RNA Hairpin by Global Conformational Rearrangements that Prolong Pausing. Mol. Cell *69*, 802-815.e1.

Kang, J.Y., Mooney, R.A., Nedialkov, Y., Saba, J., Mishanina, T. V., Artsimovitch, I., Landick, R., and Darst, S.A. (2018b). Structural Basis for Transcript Elongation Control by NusG Family Universal Regulators. Cell *173*, 1650-1662.e14.

Keating, K.S., and Pyle, A.M. (2010). Semiautomated model building for RNA crystallography using a directed rotameric approach. Proc. Natl. Acad. Sci. U. S. A. *107*, 8177–8182.

Keilty, S., and Rosenberg, M. (1987). Constitutive function of a positively regulated promoter reveals new sequences essential for activity. J. Biol. Chem. *262*, 6389–6395.

Kettenberger, H., Armache, K.J., and Cramer, P. (2004). Complete RNA polymerase II elongation complex structure and its interactions with NTP and TFIIS. Mol. Cell *16*, 955–965.

Klocko, A.D., and Wassarman, K.M. (2009). 6S RNA binding to Eσ70 requires a positively charged surface of σ70 region 4.2. Mol. Microbiol. *73*, 152–164.

Koning, R.I., Koster, A.J., and Sharp, T.H. (2018). Advances in cryo-electron tomography for biology and medicine. Ann. Anat. *217*, 82–96.

Kontur, W.S., Saecker, R.M., Capp, M.W., and Record, M.T. (2008). Late Steps in the Formation of E. coli RNA Polymerase-λPR Promoter Open Complexes: Characterization of Conformational Changes by Rapid [Perturbant] Upshift Experiments. J. Mol. Biol. *376*, 1034–1047.

Kontur, W.S., Capp, M.W., Gries, T.J., Saecker, R.M., and Record, M.T. (2010). Probing DNA binding, DNA opening, and assembly of a downstream clamp/jaw in Escherichia coli RNA polymerase-λPR promoter complexes using salt and the physiological anion glutamate. Biochemistry *49*, 4361–4373.

Kornberg, R.D. (2007). The molecular basis of eukaryotic transcription. Proc. Natl. Acad. Sci. U. S. A. *104*, 12955–12961.

Kors, C.A., Wallace, E., Davies, D.R., Li, L., Laible, P.D., and Nollert, P. (2009). Effects of impurities on membrane-protein crystallization in different systems. Acta Crystallogr. Sect. D Biol. Crystallogr. *65*, 1062–1073.

Koubek, J., Lin, K.F., Chen, Y.R., Cheng, R.P., and Huang, J.J.T. (2013). Strong anion-exchange fast performance liquid chromatography as a versatile tool for preparation and purification of RNA produced by in vitro transcription. RNA *19*, 1449–1459.

Kovacic, R.T. (1987). The 0 degree C closed complexes between Escherichia coli RNA polymerase and two promoters, T7-A3 and lacUV5. J. Biol. Chem. *262*, 13654–13661.

Kremer, J.R., Mastronarde, D.N., and McIntosh, J.R. (1996). Computer Visualization of Three-Dimensional Image Data Using IMOD. J. Struct. Biol. *116*, 71–76.

Krissinel, E., and Henrick, K. (2007). Inference of Macromolecular Assemblies from Crystalline State. J. Mol. Biol. *372*, 774–797.

Krummel, B., and Chamberlin, M.J. (1989). RNA Chain Initiation by Escherichia coli RNA Polymerase. Structural Transitions of the Enzyme in Early Ternary Complexes. Biochemistry *28*, 7829–7842.

Kuehne, S.A., Dempster, A.W., Collery, M.M., Joshi, N., Jowett, J., Kelly, M.L., Cave, R., Longshaw, C.M., and Minton, N.P. (2018). Characterization of the impact of rpoB mutations on the in vitro and in vivo competitive fitness of Clostridium difficile and susceptibility to fidaxomicin. J. Antimicrob. Chemother. *73*, 973–980.

Kulbachinskiy, A., and Mustaev, A. (2006). Region 3.2 of the σ subunit contributes to the binding of the 3′-initiating nucleotide in the RNA polymerase active center and facilitates promoter clearance during initiation. J. Biol. Chem. *281*, 18273–18276.

Kurabachew, M., Lu, S.H.J., Krastel, P., Schmitt, E.K., Suresh, B.L., Goh, A., Knox, J.E., Ma, N.L., Jiricek, J., Beer, D., et al. (2008). Lipiarmycin targets RNA polymerase and has good activity against multidrug-resistant strains of Mycobacterium tuberculosis. J. Antimicrob. Chemother. *62*, 713–719.

Landick, R. (2006). A Long Time in the Making-The Nobel Prize for RNA Polymerase. Cell *127*, 1087–1090.

Lane, W.J., and Darst, S.A. (2010a). Molecular Evolution of Multisubunit RNA Polymerases: Sequence Analysis. J. Mol. Biol. *395*, 671–685.

Lane, W.J., and Darst, S.A. (2010b). Molecular Evolution of Multisubunit RNA Polymerases: Structural Analysis. J. Mol. Biol. *395*, 686–704.

Laskowski, R.A., MacArthur, M.W., Moss, D.S., and Thornton, J.M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. *26*, 283–291.

Lee, C. a, Fournier, M.J., and Beckwith, J. (1985). Escherichia coli 6S RNA is not essential for growth or protein secretion. J. Bacteriol. *161*, 1156–1161.

Lee, D.J., Minchin, S.D., and Busby, S.J.W. (2012). Activating Transcription in Bacteria. Annu. Rev. Microbiol. *66*, 125–152.

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

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

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

Li, X., Mooney, P., Zheng, S., Booth, C., Braunfeld, M.B., Gubbens, S., Agard, D.A., and Cheng, Y. (2013). Electron counting and beam-induced motion correction enable near atomic resolution single particle cryoEM. Nat. Methods *10*, 584–590.

Li, X., Zheng, S., Agard, D.A., and Cheng, Y. (2015). Asynchronous data acquisition and on-the-fly analysis of dose fractionated cryoEM images by UCSFImage. J. Struct. Biol. *192*, 174–178.

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

Lin, W., Mandal, S., Degen, D., Liu, Y., Ebright, Y.W., Li, S., Feng, Y., Zhang, Y., Mandal, S., Jiang, Y., et al. (2017). Structural Basis of Mycobacterium tuberculosis Transcription and Transcription Inhibition. Mol. Cell *66*, 169-179.e8.

Lin, W., Das, K., Degen, D., Mazumder, A., Duchi, D., Wang, D., Ebright, Y.W., Ebright, R.Y., Sineva, E., Gigliotti, M., et al. (2018). Structural Basis of Transcription Inhibition by Fidaxomicin (Lipiarmycin A3). Mol. Cell *70*, 60-71.e15.

Liu, B., Hong, C., Huang, R.K., Yu, Z., and Steitz, T.A. (2017). Structural basis of bacterial transcription activation. Science *358*, 947–951.

Lonetto, M., Gribskov, M., and Gross, C.A. (1992). The σ70 family: Sequence conservation and evolutionary relationships. J. Bacteriol. *174*, 3843–3849.

López-Maury, L., Marguerat, S., and Bähler, J. (2008). Tuning gene expression to changing environments: From rapid responses to evolutionary adaptation. Nat. Rev. Genet. *9*, 583–593.

Lu, X.J., and Olson, W.K. (2008). 3DNA: A versatile, integrated software system for the analysis, rebuilding and visualization of three-dimensional nucleic-acid structures. Nat. Protoc. *3*, 1213–1227.

Luo, J., Liu, Z., Guo, Y., and Li, M. (2015). A structural dissection of large protein-protein crystal packing contacts. Sci. Rep. *5*, 14214.

Malhotra, A., Severinova, E., and Darst, S.A. (1996). Crystal structure of a σ70 subunit fragment from E. coli RNA polymerase. Cell *87*, 127–136.

Malinen, A.M., Turtola, M., Parthiban, M., Vainonen, L., Johnson, M.S., and Belogurov, G.A. (2012). Active site opening and closure control translocation of Multisubunit RNA polymerase. Nucleic Acids Res. *40*, 7442–7451.

Maneewannakul, K., and Ippen-Ihler, K. (1993). Construction and analysis of F plasmid traR, trbJ, and trbH mutants. J. Bacteriol. *175*, 1528–1531.

Mastronarde, D.N. (2005). Automated electron microscope tomography using robust prediction of specimen movements. J. Struct. Biol. *152*, 36–51.

Mazumder, A., and Kapanidis, A.N. (2019). Recent Advances in Understanding σ70-Dependent Transcription Initiation Mechanisms. J. Mol. Biol. *431*, 3947–3959.

McClure, W.R. (1985). Mechanism and Control of Transcription Initiation in Prokaryotes. Annu. Rev. Biochem. *54*, 171–204.

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

Mendez, J.H., Mehrani, A., Randolph, P., and Staggb, S. (2019). Throughput and resolution with a next-generation direct electron detector. IUCrJ *6*, 1007–1013.

Minakhin, L., Bhagat, S., Brunning, A., Campbell, E.A., Darst, S.A., Ebright, R.H., and Severinov, K. (2001). Bacterial RNA polymerase subunit ω and eukaryotic polymerase subunit RPB6 are sequence, structural, and functional homologs and promote RNA polymerase assembly. Proc. Natl. Acad. Sci. U. S. A. *98*, 892–897.

Miropolskaya, N., Artsimovitch, I., Klimašauskas, S., Nikiforov, V., and Kulbachinskiy, A. (2009). Allosteric control of catalysis by the F loop of RNA polymerase. Proc. Natl. Acad. Sci. U. S. A. *106*, 18942–18947.

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

Moreno, A., Théobald-Dietrich, A., Lorber, B., Sauter, C., and Giegé, R. (2005). Effects of macromolecular impurities and of crystallization method on the quality of eubacterial aspartyl-tRNA synthetase crystals. Acta Crystallogr. Sect. D Biol. Crystallogr. *61*, 789–792.

Morichaud, Z., Chaloin, L., and 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. J. Mol. Biol. *428*, 463–476.

Mukhopadhyay, J., Das, K., Ismail, S., Koppstein, D., Jang, M., Hudson, B., Sarafianos, S., Tuske, S., Patel, J., Jansen, R., et al. (2008). The RNA Polymerase "Switch Region" Is a Target for Inhibitors. Cell *135*, 295–307.

Munson, L.M., and Reznikoff, W.S. (1981). Abortive Initiation and Long Ribonucleic Acid Synthesis. Biochemistry *20*, 2081–2085.

Murakami, K.S. (2013). X-ray crystal structure of Escherichia coli RNA polymerase σ70 holoenzyme. J. Biol. Chem. *288*, 9126–9134.

Murakami, K.S. (2015). Structural biology of bacterial RNA polymerase. Biomolecules *5*, 848–864.

Murakami, K.S., and Darst, S.A. (2003). Bacterial RNA polymerases: The wholo story. Curr. Opin. Struct. Biol. *13*, 31–39.

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

Murakami, K.S., Masuda, S., and Darst, S.A. (2002b). Structural basis of transcription initiation: RNA polymerase holoenzyme at 4 Å resolution. Science *296*, 1280–1284.

Murata, K., and Wolf, M. (2018). Cryo-electron microscopy for structural analysis of dynamic biological macromolecules. Biochim. Biophys. Acta - Gen. Subj. *1862*, 324–334.

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

Nagy, J., Grohmann, D., Cheung, A.C.M., Schulz, S., Smollett, K., Werner, F., and Michaelis, J. (2015). Complete architecture of the archaeal RNA polymerase open complex from single-molecule FRET and NPS. Nat. Commun. *6*, 6161.

Nakane, T., Kimanius, D., Lindahl, E., and Scheres, S.H.W.W. (2018). Characterisation of molecular motions in cryo-EM single-particle data by multi-body refinement in RELION. Elife *7*, e36861.

Narayanan, A., Vago, F.S., Li, K., Qayyum, M.Z., Yernool, D., Jiang, W., and Murakami, K.S. (2018). Cryo-EM structure of Escherichia coli 70 RNA polymerase and promoter DNA complex revealed a role of non-conserved region during the open complex formation. J. Biol. Chem. *293*, 7367–7375.

Naydenova, K., and Russo, C.J. (2017). Measuring the effects of particle orientation to improve the efficiency of electron cryomicroscopy. Nat. Commun. *8*, 8–12.

Newell, K. V., Thomas, D.P., Brekasis, D., and Paget, M.S.B. (2006). The RNA polymerase-binding protein RbpA confers basal levels of rifampicin resistance on Streptomyces coelicolor. Mol. Microbiol. *60*, 687–696.

Neyer, S., Kunz, M., Geiss, C., Hantsche, M., Hodirnau, V.V., Seybert, A., Engel, C., Scheffer, M.P., Cramer, P., and Frangakis, A.S. (2016). Structure of RNA polymerase I transcribing ribosomal DNA genes. Nature *540*, 607–610.

Nguyen, L.H., and Burgess, R.R. (1996). Overproduction and purification of σ(s), the Escherichia coli stationary phase specific sigma transcription factor. Protein Expr. Purif. *8*, 17–22.

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

Noble, A.J., and Stagg, S.M. (2015). Automated batch fiducial-less tilt-series alignment in Appion using Protomo. J. Struct. Biol. *192*, 270–278.

Noble, A.J., Dandey, V.P., Wei, H., Brasch, J., Chase, J., Acharya, P., Tan, Y.Z., Zhang, Z., Kim, L.Y., Scapin, G., et al. (2018a). Routine single particle CryoEM sample and grid characterization by tomography. Elife *7*, e34257.

Noble, A.J., Wei, H., Dandey, V.P., Zhang, Z., Tan, Y.Z., Potter, C.S., and Carragher, B. (2018b). Reducing effects of particle adsorption to the air–water interface in cryo-EM. Nat. Methods *15*, 793–795.

O'Neill, A., Oliva, B., Storey, C., Hoyle, A., Fishwick, C., and Chopra, I. (2000). RNA polymerase inhibitors with activity against rifampin-resistant mutants of Staphylococcus aureus. Antimicrob. Agents Chemother. *44*, 3163–3166.

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

Paget, M.S.B., Molle, V., Cohen, G., Aharonowitz, Y., and Buttner, M.J. (2001). Defining the disulphide stress response in Streptomyces coelicolor A3(2): Identification of the σR regulon. Mol. Microbiol. *42*, 1007–1020.

Panchapakesan, S.S.S., and Unrau, P.J. (2012). E. coli 6S RNA release from RNA polymerase requires σ70 ejection by scrunching and is orchestrated by a conserved RNA hairpin. RNA *18*, 2251–2259.

Parisien, M., and Major, F. (2008). The MC-Fold and MC-Sym pipeline infers RNA structure from sequence data. Nature *452*, 51–55.

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

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

Paul, B.J., Ross, W., Gaal, T., and Gourse, R.L. (2004b). rRNA Transcription in Escherichia coli . Annu. Rev. Genet. *38*, 749–770.

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

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

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

Plaschka, C., Hantsche, M., Dienemann, C., Burzinski, C., Plitzko, J., and Cramer, P. (2016). Transcription initiation complex structures elucidate DNA opening. Nature *533*, 353–358.

Polyakov, A., Severinova, E., and Darst, S.A. (1995). Three-dimensional structure of E. coil core RNA polymerase: Promoter binding and elongation conformations of the enzyme. Cell *83*, 365–373.

Puglisi, J.D., Chen, L., Blanchard, S., and Frankel, A.D. (1995). Solution Structure of a Bovine Immunodeficiency Virus Tat-TAR Peptide-RNA Complex. Science *270*, 1200–1203.

Punjani, A., Rubinstein, J.L., Fleet, D.J., and Brubaker, M.A. (2017). CryoSPARC: Algorithms for rapid unsupervised cryo-EM structure determination. Nat. Methods *14*, 290–296.

Rao, L., Ross, W., Appleman, J.A., Gaal, T., Leirmo, S., Schlax, P.J., Record, M.T., and Gourse, R.L. (1994). Factor independent activation of rrnB p1. An "Extended" promoter with an upstream element that dramatically increases promoter strength. J. Mol. Biol. *235*, 1421–1435.

Richardson, J.S., Schneider, B., Murray, L.W., Kapral, G.J., Immormino, R.M., Headd, J.J., Richardson, D.C., Ham, D., Hershkovits, E., Williams, L.D., et al. (2008). RNA backbone: Consensus all-angle conformers and modular string nomenclature (an RNA Ontology Consortium contribution). RNA *14*, 465–481.

Roe, J.H., Burgess, R.R., and Record, M.T. (1984). Kinetics and mechanism of the interaction of Escherichia coli RNA polymerase with the λPR promoter. J. Mol. Biol. *176*, 495–522.

Rogozina, A., Zaychikov, E., Buckle, M., Heumann, H., and Sclavi, B. (2009). DNA melting by RNA polymerase at the T7A1 promoter precedes the rate-limiting step at 37 degrees C and results in the accumulation of an off-pathway intermediate. Nucleic Acids Res. *37*, 5390–5404.

Rohou, A., and Grigorieff, N. (2015). CTFFIND4: Fast and accurate defocus estimation from electron micrographs. J. Struct. Biol. *192*, 216–221.

Rosenberg, S., Kadesch, T.R., and Chamberlin, M.J. (1982). Binding of Escherichia coli RNA polymerase holoenzyme to bacteriophage T7 DNA. J. Mol. Biol. *155*, 31–51.

Rosenthal, P.B., and Henderson, R. (2003). Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. J. Mol. Biol. *333*, 721–745.

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

Ross, W., and Gourse, R.L. (2009). Analysis of RNA polymerase-promoter complex formation. Methods *47*, 13–24.

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

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

Ross, W., Schneider, D.A., Paul, B.J., Mertens, A., and Gourse, R.L. (2003). An intersubunit contact stimulating transcription initiation by E. coli RNA polymerase: Interaction of the α C-terminal domain and σ region 4. Genes Dev. *17*, 1293–1307.

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

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

Roszak, D.B., and Colwell, R.R. (1987). Survival strategies of bacteria in the natural environment. Microbiol. Rev. *51*, 365–379.

Rubinstein, J.L., and Brubaker, M. a. (2015). Alignment of cryo-EM movies of individual particles by optimization of image translations. J. Struct. Biol. *192*, 188–195.

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

Ruff, E.F., Record, M.T., and Artsimovitch, I. (2015b). Initial events in bacterial transcription initiation. Biomolecules *5*, 1035–1062.

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

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

Saecker, R.M., Tsodikov, O. V., McQuade, K.L., Schlax, P.E., Capp, M.W., and Record, M.T. (2002). Kinetic studies and structural models of the association of E. coli σ70 RNA polymerase with the λPR promoter: Large scale conformational changes in forming the kinetically significant intermediates. J. Mol. Biol. *319*, 649–671.

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

Saenger, W. (1984). Principles of Nucleic Acid Structure. Springer Adv. Texts Chem.

Sanchez-Vazquez, P., Dewey, C.N., Kitten, N., Ross, W., and Gourse, R.L. (2019). Genome-wide effects on Escherichia coli transcription from ppGpp binding to its two sites on RNA polymerase. Proc. Natl. Acad. Sci. U. S. A. *116*, 8310–8319.

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

Schickor, P., Metzger, W., Werel, W., Lederer, H., and Heumann, H. (1990). Topography of intermediates in transcription initiation of E.coli. EMBO J. *9*, 2215–2220.

Schneider, T.D., and Stephens, R.M. (1990). Sequence logos: A new way to display consensus sequences. Nucleic Acids Res. *18*, 6097–6100.

Schulz, S., Gietl, A., Smollett, K., Tinnefeld, P., Werner, F., Grohmann, D., and Geiduschek, E.P. (2016). TFE and Spt4/5 open and close the RNA polymerase clamp during the transcription cycle. Proc. Natl. Acad. Sci. U. S. A. *113*, E1816–E1825.

Schwartz, E.C., Shekhtman, A., Dutta, K., Pratt, M.R., Cowburn, D., Darst, S., and Muir, T.W. (2008). A Full-Length Group 1 Bacterial Sigma Factor Adopts a Compact Structure Incompatible with DNA Binding. Chem. Biol. *15*, 1091–1103.

Sclavi, B., Zaychikov, E., Rogozina, A., Walther, F., Buckle, M., and Heumann, H. (2005). Real-time characterization of intermediates in the pathway to open complex formation by Escherichia coli RNA polymerase at the T7A1 promoter. Proc. Natl. Acad. Sci. U. S. A. *102*, 4706 LP – 4711.

Serra, S., Malpezzi, L., Bedeschi, A., Fuganti, C., and Fonte, P. (2017). Final demonstration of the co-identity of lipiarmycin A3 and tiacumicin B(Fidaxomicin) through single crystal X-ray analysis. Antibiotics *6*.

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

Shephard, L., Dobson, N., and Unrau, P.J. (2010). Binding and release of the 6S transcriptional control RNA. RNA *16*, 885–892.

Shi, Y. (2014). A glimpse of structural biology through X-ray crystallography. Cell *159*, 995–1014.

Shimamoto, N., Kamigochi, T., and Utiyama, H. (1986). Release of the σ subunit of Escherichia coli DNA-dependent RNA polymerase depends mainly on time elapsed after the start of initiation, not on length of product RNA. J. Biol. Chem. *261*, 11859–11865.

Shultzaberger, R.K., Chen, Z., Lewis, K.A., and Schneider, T.D. (2007). Anatomy of Escherichia coli σ70 promoters. Nucleic Acids Res. *35*, 771–788.

Simpson, R.B. (1979). The molecular topography of rna polymerase-promoter interaction. Cell *18*, 277–285.

Sponer, J., Bussi, G., Krepl, M., Banas, P., Bottaro, S., Cunha, R.A., Gil-Ley, A., Pinamonti, G., Poblete, S., Jurečka, P., et al. (2018). RNA structural dynamics as captured by molecular simulations: A comprehensive overview. Chem. Rev. *118*, 4177–4338.

Srivastava, A., Talaue, M., Liu, S., Degen, D., Ebright, R.Y., Sineva, E., Chakraborty, A., Druzhinin, S.Y., Chatterjee, S., Mukhopadhyay, J., et al. (2011). New target for inhibition of bacterial RNA polymerase: "switch region." Curr. Opin. Microbiol. *14*, 532–543.

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

Stallings, C.L., Stephanou, N.C., Chu, L., Hochschild, A., Nickels, B.E., and Glickman, M.S. (2009). CarD Is an Essential Regulator of rRNA Transcription Required for Mycobacterium tuberculosis Persistence. Cell *138*, 146–159.

Stefano, J.E., and Gralla, J.D. (1982). Spacer mutations in the lac p(S) promoter. Proc. Natl. Acad. Sci. U. S. A. *79*, 1069–1072.

Steuten, B., Setny, P., Zacharias, M., and Wagner, R. (2013). Mapping the spatial neighborhood of the regulatory 6S RNA bound to Escherichia coli RNA polymerase holoenzyme. J. Mol. Biol. *425*, 3649–3661.

Steuten, B., Hoch, P.G., Damm, K., Schneider, S., Köhler, K., Wagner, R., and Hartmann, R.K. (2014). Regulation of transcription by 6S RNAs: Insights from the Escherichia coli and Bacillus subtilis model systems. RNA Biol. *11*, 1.

Studholme, D.J., Wigneshwereraraj, S.R., Gallegos, M.T., and Buck, M. (2000). Functionality of purified σ(N) (σ54) and a NifA-like protein from the hyperthermophile Aquifex aeolicus. J. Bacteriol. *182*, 1616–1623.

Stumper, S.K., Ravi, H., Friedman, L.J., Mooney, R.A., Corrêa, I.R., Gershenson, A., Landick, R., and Gelles, J. (2019). Delayed inhibition mechanism for secondary channel factor regulation of ribosomal RNA transcription. Elife *8*, e40576.

Svejstrup, J.Q., Conaway, R.C., and Conaway, J.W. (2006). RNA Polymerase II: A "Nobel" Enzyme Demystified. Mol. Cell *24*, 637–642.

Tabib-Salazar, A., Liu, B., Doughty, P., Lewis, R.A., Ghosh, S., Parsy, M.L., Simpson, P.J., O'Dwyer, K., Matthews, S.J., and Paget, M.S. (2013). The actinobacterial transcription factor RbpA binds to the principal sigma subunit of RNA polymerase. Nucleic Acids Res. *41*, 5679–5691.

Tafur, L., Sadian, Y., Hoffmann, N.A., Jakobi, A.J., Wetzel, R., Hagen, W.J.H., Sachse, C., and Müller, C.W. (2016). Molecular Structures of Transcribing RNA Polymerase I. Mol. Cell *64*, 1135–1143.

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

Tagami, S., Sekine, S. ichi, and Yokoyama, S. (2011). A novel conformation of RNA polymerase sheds light on the mechanism of transcription. Transcription *2*, 162–167.

Talpaert, M., Campagnari, F., and Clerici, L. (1975). Lipiarmycin: An antibiotic inhibiting nucleic acid polymerases. Biochem. Biophys. Res. Commun. *63*, 328–334.

Tan, Y.Z., Baldwin, P.R., Davis, J.H., Williamson, J.R., Potter, C.S., Carragher, B., and Lyumkis, D. (2017). Addressing preferred specimen orientation in single-particle cryo-EM through tilting. Nat. Methods *14*, 793–796.

Tang, G., Peng, L., Baldwin, P.R., Mann, D.S., Jiang, W., Rees, I., and Ludtke, S.J. (2007). EMAN2: An extensible image processing suite for electron microscopy. J. Struct. Biol. *157*, 38–46.

Taylor, K.A., and Glaeser, R.M. (2008). Retrospective on the early development of cryoelectron microscopy of macromolecules and a prospective on opportunities for the future. J. Struct. Biol. *163*, 214–223.

Tomsic, M., Tsujikawa, L., Panaghie, G., Wang, Y., Azok, J., and DeHaseth, P.L. (2001). Different Roles for Basic and Aromatic Amino Acids in Conserved Region 2 of Escherichia coli σ70 in the Nucleation and Maintenance of the Single-stranded DNA Bubble in Open RNA Polymerase-Promoter Complexes. J. Biol. Chem. *276*, 31891–31896.

Trotochaud, A.E., and Wassarman, K.M. (2004). 6S RNA function enhances long-term cell survival. J. Bacteriol. *186*, 4978–4985.

Trotochaud, A.E., and Wassarman, K.M. (2005). A highly conserved 6S RNA structure is required for regulation of transcription. Nat. Struct. Mol. Biol. *12*, 313–319.

Tupin, A., Gualtieri, M., Leonetti, J.P., and Brodolin, K. (2010). The transcription inhibitor lipiarmycin blocks DNA fitting into the RNA polymerase catalytic site. EMBO J. *29*, 2527–2537.

Twist, K.-A., Husnain, S.I., Franke, J.D., Jain, D., Campbell, E.A., Nickels, B.E., Thomas, M.S., Darst, S.A., and 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 Sci. *20*, 986–995.

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

Vassylyev, D.G., Vassylyeva, M.N., Perederina, A., Tahirov, T.H., and Artsimovitch, I. (2007). Structural basis for transcription elongation by bacterial RNA polymerase. Nature *448*, 157–162.

Venugopal, A.A., and Johnson, S. (2012). Fidaxomicin: A novel macrocyclic antibiotic approved for treatment of clostridium difficile infection. Clin. Infect. Dis. *54*, 568–574.

Vo, N. V., Hsu, L.M., Kane, C.M., and Chamberlin, M.J. (2003). In vitro studies of transcript initiation by Escherichia coli RNA polymerase. 3. Influences of individual DNA elements within the promoter recognition region on abortive initiation and promoter escape. Biochemistry *42*, 3798–3811.

Vorländer, M.K., Khatter, H., Wetzel, R., Hagen, W.J.H., and Müller, C.W. (2018). Molecular mechanism of promoter opening by RNA polymerase III. Nature *553*, 295–300.

Vuthoori, S., Bowers, C.W., McCracken, A., Dombroski, A.J., and Hinton, D.M. (2001). Domain 1.1 of the σ70 subunit of Escherichia coli RNA polymerase modulates the formation of stable polymerase/promoter complexes. J. Mol. Biol. *309*, 561–572.

Wagner, S.D., Yakovchuk, P., Gilman, B., Ponicsan, S.L., Drullinger, L.F., Kugel, J.F., and Goodrich, J.A. (2013). RNA polymerase II acts as an RNA-dependent RNA polymerase to extend and destabilize a non-coding RNA. EMBO J. *32*, 781–790.

Walter, G., Zillig, W., Palm, P., and Fuchs, E. (1967). Initiation of DNA-Dependent RNA Synthesis and the Effect of Heparin on RNA Polymerase. Eur. J. Biochem. *3*, 194–201.

Wang, D., Bushnell, D.A., Westover, K.D., Kaplan, C.D., and Kornberg, R.D. (2006). Structural basis of transcription: role of the trigger loop in substrate specificity and catalysis. Cell *127*, 941–954.

Wassarman, K.M., and Saecker, R.M. (2006). Synthesis-mediated release of a small RNA inhibitor of RNA polymerase. Science *314*, 1601–1603.

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

Wedel, A., and Kustu, S. (1995). The bacterial enhancer-binding protein NTRC is a molecular machine: ATP hydrolysis is coupled to transcriptional activation. Genes Dev. *9*, 2042–2052.

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

Willkomm, D.K., Minnerup, J., Hüttenhofer, A., and Hartmann, R.K. (2005). Experimental RNomics in Aquifex aeolicus: Identification of small non-coding RNAs and the putative 6S RNA homolog. Nucleic Acids Res. *33*, 1949–1960.

Wilson, C., and Dombroski, A.J. (1997). Region 1 of σ70 is required for efficient isomerization and initiation of transcription by Escherichia coli RNA polymerase. J. Mol. Biol. *267*, 60–74.

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

Winkelman, J.T.T., Winkelman, B.T.T., Boyce, J., Maloney, M.F.F., Chen, A.Y.Y., Ross, W., and Gourse, R.L.L. (2015). Crosslink Mapping at Amino Acid-Base Resolution Reveals the Path of Scrunched DNA in Initial Transcribing Complexes. Mol. Cell *59*, 1–13.

World Health Organization (2018). Global TUBERCULOSIS Report 2018- Executive summary (Geneva PP - Geneva: World Health Organization).

Yakovchuk, P., Goodrich, J.A., and Kugel, J.F. (2009). B2 RNA and Alu RNA repress transcription by disrupting contacts between RNA polymerase II and promoter DNA within assembled complexes. Proc. Natl. Acad. Sci. *106*, 5569–5574.

Ye, X., Kumar, R.A., and Patel, D.J. (1995). Molecular recognition in the bovine immunodeficiency virus Tat peptide-TAR RNA complex. Chem. Biol. *2*, 827–840.

Yee, D., Armstrong, V.W., and Eckstein, F. (1979). Mechanistic Studies on Deoxyribonucleic Acid Dependent Ribonucleic Acid Polymerase from Escherichia coli Using Phosphorothioate Analogues. 1. Initiation and Pyrophosphate Exchange Reactions. Biochemistry *18*, 4116–4120.

Yuzenkova, Y., Tadigotla, V.R., Severinov, K., and Zenkin, N. (2011). A new basal promoter element recognized by RNA polymerase core enzyme. EMBO J. *30*, 3766–3775.

Zakharova, N., Bass, I., Arsenieva, E., Nikiforov, V., and Severinov, K. (1998). Mutations in and monoclonal antibody binding to evolutionary hypervariable region of Escherichia coli RNA polymerase β' subunit inhibit transcript cleavage and transcript elongation. J. Biol. Chem. *273*, 24912–24920.

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

Zhang, G., Campbell, E.A., Minakhin, L., Richter, C., Severinov, K., and Darst, S.A. (1999). Crystal structure of thermus aquaticus core RNA polymerase at 3.3 Å resolution. Cell *98*, 811–824.

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

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

Zivanov, J., Nakane, T., Forsberg, B.O., Kimanius, D., Hagen, W.J.H., Lindahl, E., and Scheres, S.H.W. (2018). New tools for automated high-resolution cryo-EM structure determination in RELION-3. Elife *7*, e42166.

Zivanov, J., Nakane, T., and Scheres, S.H.W. (2019). Estimation of High-Order Aberrations and Anisotropic Magnification from Cryo-EM Datasets in RELION-3.1. BioRxiv 798066.

Zumla, A., George, A., Sharma, V., Herbert, R.H.N., Oxley, A., and Oliver, M. (2015). The WHO 2014 Global tuberculosis report-further to go. Lancet Glob. Heal. *3*, e10–e12.

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

Zuo, Y., Wang, Y., and Steitz, T.A. (2013). The Mechanism of E. coli RNA Polymerase Regulation by ppGpp is suggested by the structure of their complex. Mol. Cell *50*, 430–436.