Regulation of Mycobacterial Transcription Initiation

Jayan Rammohan

Washington University in St. Louis

Follow this and additional works at: https://openscholarship.wustl.edu/art_sci_etds

Recommended Citation
Rammohan, Jayan, "Regulation of Mycobacterial Transcription Initiation" (2016). Arts & Sciences Electronic Theses and Dissertations. 1004.
https://openscholarship.wustl.edu/art_sci_etds/1004
Regulation of Mycobacterial Transcription Initiation
by
Jayan Rammohan

A dissertation presented to
The Graduate School
of Washington University in
partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

December 2016
St. Louis, Missouri
Table of Contents

List of Figures ................................................................................................................................. iv
List of Tables ................................................................................................................................. ix
Acknowledgments ........................................................................................................................ x
Abstract of the Dissertation ......................................................................................................... xiii

Chapter 1: Introduction ................................................................................................................... 1
References ..................................................................................................................................... 24

Chapter 2: Binding studies of CarD, RNAP, and DNA ................................................................. 29
References ..................................................................................................................................... 47

Chapter 3: Mechanisms of CarD-regulated mycobactertial transcription initiation ................. 49
Introduction ................................................................................................................................... 50
References ..................................................................................................................................... 51
Published manuscript (Rammohan et al., Nucleic Acids Research, 2015, Vol. 43, No. 6, 3272-3285) ................................................................................................................................. 52
References ..................................................................................................................................... 64
Supplemental Information ............................................................................................................ 66
Epilogue ........................................................................................................................................ 76
References ..................................................................................................................................... 80

Chapter 4: Cooperative stabilization of Mycobacterium tuberculosis rrnAP3 promoter open complexes by RbpA and CarD ................................................................. 81
Introduction ................................................................................................................................... 82
Published manuscript (Rammohan et al., Nucleic Acids Research, 2015, Vol. 43, No. 6, 3272-3285) ................................................................................................................................. 83
References ..................................................................................................................................... 91
List of Figures

Chapter 1: Introduction................................................................................................................... 1
Figure 1: General kinetic mechanism of bacterial transcription initiation ................................. 3
Figure 2: Structure of CarD from *Thermus thermophilus* ........................................................... 13
Figure 3: Structure of RbpA ......................................................................................................... 14
Figure 4: How do the essential transcription factors CarD and RbpA affect the mechanism of mycobacterial transcription initiation? ......................................................................................... 15
Figure 5: Schematic of the *Mtb* rrnAP3 promoter used in this study ........................................... 18

Chapter 2: Binding studies of CarD, RNAP, and DNA................................................................. 29
Figure 1: EMSA for measuring MboRNAP holoenzyme's binding affinity for 150 bp promoter DNA .............................................................................................................................................. 37
Figure 2: EMSA for measuring CarD's binding affinity for 150 bp promoter DNA .................... 38
Figure 3: CarD increases the affinity of RNAP for DNA ............................................................. 39
Figure 4: Sub-µM CarD increases the affinity of promoter DNA only for MboRNAP holoenzyme with wild-type SigA ................................................................................................................................. 42
Figure 5: Fraction-bound calculated from double-filter binding RNAP holoenzyme titrations with increasing CarD ............................................................................................................................. 43

Chapter 3: Mechanisms of CarD-regulated transcription initiation.............................................. 49
Published manuscript (Rammohan et al., Nucleic Acids Research, 2015, Vol. 43, No. 6, 3272-3285) .............................................................................................................................................. 52
Figure 1: A real-time fluorescence assay for monitoring open-complex formation ..................... 56
Figure 2: RNAP concentration dependence of promoter opening ................................................ 57
Figure 3: CarD increases open-complex formation with MboRNAP ........................................... 58
Figure 4: MboRNAP titration at saturating CarD ......................................................................... 58
Figure 5: Kinetics of open-complex formation ............................................................................. 59
Figure 6: A CarD-dependent kinetic model of open-complex formation ..................................... 60
Figure 7: Mutants of CarD diminish the degree of open-complex stabilization ........................... 61
Figure 8: Energy diagram ............................................................................................................. 63
Figure S3: Concentration dependence of the observed rates with factor concentration plotted in log scale ........................................................................................................................................ 97
Figure S4: Models for factor mechanisms ................................................................................................................................. 98
Figure S5: Distribution of global fit parameters for cooperative binding model ............................................................ 99
Figure S6: Three nucleotide aborted transcription ............................................................................................................. 100
Figure S7: Fractional amplitude dependence on factor concentration in the presence of both factors .......................................................... 101
Epilogue ...................................................................................................................................... 102
Figure 1: Titration of RbpA in the presence of CarD-W85A ....................................................................................... 103
Figure 2: Titration of CarD in the presence of RbpA-R79A ....................................................................................... 104
Figure 3: General model for CarD-RbpA-RPc-RPo cooperativity ........................................................................ 105

Chapter 5: Kinetic models for factor-dependent mycobacterial open complex stabilization….. 106
Figure 1: Kinetic model of CarD stabilization of mycobacterial open complexes .................. 110
Figure 2: A single eigenvalue captures non-monotonic dependence of observed rate on concentrations of CarD and RNAP ........................................................................................................ 112
Figure 3: Constraints for qualitatively reproducing CarD-dependent observed kinetics .... 113
Figure 4: Pairs of on- and off-rates for CarD to/from the same state are poorly constrained .... 117
Figure 5: Example of well-constrained pairs of rate constants ........................................................................ 118
Figure 6: Example of well-constrained, non-unique pair of rate constants ........................ 119
Figure 7: Graphical summary of scaling all 28 pairs of rate constants in the thermodynamic cycle ........................................................................................................................................... 120
Figure 8: Summary of estimates of rate constants and uncertainty ........................................ 121
Figure 9: A 4-state thermodynamic cycle reproduces non-monotonic trend in observed rates, but cannot account for [RNAP] dependence ......................................................................................................................................... 123
Figure 10: A 4-state linear mechanism is unable to reproduce a non-monotonic trend in observed rate ................................................................................................................................................ 125
Figure 11: A 4-state linear mechanism which incorporates conformational selection and induced fit can account for RbpA's trend in observed rates ........................................................................ 126

Chapter 6: Promoter Escape .................................................................................................................................................. 129
Figure 1: Dilution-dependent dissociation/re-equilibration must be accounted for when studying dissociation and promoter escape by MboRNAP ................................................................. 132

Figure 2: Equal volume mixing in the presence and absence of CarD using dsDNA competitor under single-turnover conditions .......................................................................................... 134

Figure 3: Promoter escape by E. coli RNAP can be measured using equal volume mixing under single-turnover conditions ................................................................................................. 137

Figure 4: Calibration of 10:1 asymmetric mixing using Cy3-labeled DNA ........................................ 138

Figure 5: Dissociation and promoter escape by MboRNAP in the absence of CarD .............. 141

Figure 6: Promoter escape with CarD can be observed under single-turnover conditions...... 142

Figure 7: Promoter escape with CarD can be observed under multi-turnover conditions........ 143

Chapter 7: Potassium glutamate stabilizes mycobacterial open complexes ......................... 150

Figure 1: Experiments performed in KGlu buffer exhibited increased basal open-complex stability ........................................................................................................................................ 159

Figure 2: Fractional amplitudes as a function of [MboRNAP].............................................. 160

Figure 3: Observed rates as a function of [MboRNAP].......................................................... 161

Figure 4: Kinetic measurements as a function of [CarD] in KGlu ......................................... 162

Figure 5: Observed rate dependence on [CarD] ................................................................. 163

Chapter 8: Single molecule magnetic tweezers for real-time observation of CarD-dependent mycobacterial promoter unwinding ................................................................. 165

Figure 1: Single molecule magnetic tweezers for real-time detection of promoter unwinding.. 168

Figure 2: Rotation-extension curve for torsionally constrained DNA containing the Mtb rrnA P3 promoter ...................................................................................................................... 171

Figure 3: Signal for DNA only with positive and negative supercoils .................................... 173

Figure 4: Single molecule trace for MboRNAP on positively and negatively supercoiled DNA ....................................................................................................................................... 174

Figure 5: Observation of promoter-opening in the presence of CarD .................................... 176

Figure 6: DNA extension analysis ...................................................................................................... 177

Figure 7: Unwinding analysis ..................................................................................................... 178

Figure 8: Dwell-time distribution analysis ................................................................................. 179
List of Tables

Chapter 6: Promoter Escape........................................................................................................ 129
Table 1: Summary of observed rates of dissociation and promoter escape in the presence of CarD
..................................................................................................................................................... 144

Chapter 8: Single molecule magnetic tweezers for real-time observation of CarD-dependent
mycobacterial promoter unwinding............................................................................................ 165
Table 1: Single molecule magnetic tweezers file index.............................................................. 182

Appendix A: Sample preparation for Cy3-labeled promoter...................................................... 199
Table 1: Library of oligonucleotides ordered from Integrated DNA Technologies (IDT)....... 208
Acknowledgments

First and foremost, I would like to thank Eric Galburt for the opportunity to learn and grow as a scientist in his laboratory. It is highly unlikely that I will ever have so much fun working in any job for the rest of my life. Over the years, I have benefited from working closely with outstanding colleagues within the Galburt Lab including Eric “Inspector” Tomko, Drake Jensen, Haixin Ke, Tina Bradford, and many other talented scientists. Deserving of special mention is Ana Ruiz-Manzano, who is a constant force of optimism and practicality, and who I have enjoyed working closely with on this project.

I would like to thank Christina Stallings, who discovered CarD, and all members of her laboratory, especially Ashley Garner and Leslie Weiss, for showing me what is possible with regards to fruitful scientific collaboration. In addition to providing sage advice about bacterial transcription initiation, Tomasz Heyduk and members of his laboratory were kind enough to teach me how to perform an ensemble fluorescence assay that proved to be my primary research tool for this work. I am grateful to Barak Cohen, Roberto Galletto, Daved Fremont, Tim Lohman and Tom Ellenberger for their generosity with regards to scientific advice and equipment in their labs.

I am thankful for every staff member of the Department of Biochemistry and Molecular Biophysics and every staff member of the Division of Biology and Biomedical Sciences graduate
programs, including but not limited to Jayma Mikes, Anna Blanchard, Ginny Ribaudo, Linda Ketchens, Darrell Baldwin, Melissa Torres, Sally Vogt, Andrea Krussell, and John Russell, who have all operated like a well-oiled machine in support of my research and educational activities during my time at Washington University in St. Louis.

I am thankful for the friends that I have made during my time in St. Louis, including but not limited to Josh, Michi, Brian, Mitch, Joe, Stephanie and Emma. Without them, this journey would have been a lot less fun. Hopefully they have learned from me a fraction of what I have learned from them about friendship and life.

I would like to profoundly thank my parents Ram and Bindu for more than words can express. I thank my sister Kalyani and her energetic family for reminding me that one should never be afraid to take on the world. I thank all members of my immediate and extended family, near and far, past and present, for providing the current that keeps this leaf flowing down the river. Most of all, I am especially grateful for the unconditional love and support of my wife Adrienne, who has endured every challenge presented to us with patience, selflessness and grace. I hope that someday our son Naveen will read this and be as proud of his dad as I am of him.

Jayan Rammohan

Washington University in St. Louis

December 2016
Dedicated to my family.
Cellular life relies on gene expression, in which DNA is first transcribed into RNA, which is then translated into protein. Transcription is performed by the protein RNA polymerase (RNAP), which interacts with DNA in three sequential events known as initiation, elongation and termination. *Mycobacterium tuberculosis* (*Mtb*) represents a global burden to public health, and the transcription factors CarD and RbpA are both essential to *Mtb*. I have studied the effect of CarD and RbpA on transcription initiation *in vitro* at the *Mtb* *rrnA*P3 promoter. I have shown that CarD stabilizes unwinding of promoter DNA by RNAP using a two-tiered kinetic mechanism. I have also shown that RbpA stabilizes mycobacterial open complexes using a mechanism distinct from that of CarD. Furthermore, RbpA and CarD cooperatively stabilize mycobacterial open complexes, leading to increased transcription. Taken together, these findings lay the groundwork for a mechanistic understanding of gene regulation by essential transcription factors in a bacterium that kills millions of people each year.
Chapter 1

Introduction
1.1 Cells regulate transcription initiation to turn genes on and off

Cellular organisms respond to changes in their environment. They do this by choosing which RNA and protein molecules to manufacture based on environmental conditions. This process involves the transcription of DNA into RNA, which can then be translated into protein. This flow, or expression, of genetic information between different molecular classes is commonly known as the Central Dogma of Biology. Gene expression can be regulated at any of the steps of transcription and translation, and these regulatory mechanisms enable a cell to manufacture molecules as needed (1).

Transcription represents the beginning of gene expression, and consists of initiation, elongation, and termination (Figure 1). One of the most important opportunities for regulating gene expression is the beginning of the beginning: Transcription initiation. Transcription initiation consists of several reversible steps. RNA polymerase holoenzyme (RNAP) binds promoter DNA, unwinds the DNA, and begins transcription by incorporating nucleotides on the template strand. Transcription initiation is followed by an irreversible transition into transcription elongation, during which RNAP moves downstream along the gene and continues transcribing RNA. Transcription ends with termination, during which RNAP and the completed RNA transcript dissociate from each other and the template DNA. While all steps leading to the formation of a complete RNA transcript can in principle be used to regulate gene expression, regulation is most often observed at the step of transcription initiation.
R = RNA Polymerase, the enzyme that transcribes RNA from DNA
P = Promoter DNA sequence, where transcription starts at the beginning of a gene

R + P ⇄ RPe

Figure 1: General kinetic mechanism of bacterial transcription. Transcription begins with initiation, during which RNA Polymerase (R) binds Promoter DNA (P) to form the closed complex ($\text{RP}_c$), RNAP unwinds the DNA to form the open complex ($\text{RP}_o$), and RNA Polymerase incorporates the initial nucleotides of the RNA transcript onto the template strand to form the initial-transcribing complex ($\text{RP}_{\text{itc}}$). RNA Polymerase then transitions into an elongation complex ($\text{RP}_e$) in which the RNA transcript is extended. The end of transcription is known as termination, during which RNA polymerase and the RNA transcript dissociate from the DNA.

Transcription factors can be used to turn genes on and off depending on what stimuli the cell is experiencing. In a classic example from the model bacterial organism Escherichia coli (E.coli, or Eco), the lactose operon is regulated by the lac repressor such that the proteins for digesting lactose are not made unless lactose is present (1). The lac repressor binds to specific sequence of DNA at the beginning of the gene for beta-galactosidase, which is an enzyme that bacteria make for metabolizing lactose. By inhibiting the ability of RNAP to initiate transcription, the lac repressor acts to turn off the operon for lactose-metabolizing proteins when lactose is not present. When a sufficiently high concentration of lactose is present in the environment, lactose is able to bind the repressor, which leads to a reduction in the affinity of the repressor for promoter DNA. The reduction in the occupancy of the promoter by repressor allows the RNAP access to the gene and stimulates transcription. This removal of lac repressor by lactose turns the gene from “off” to “on”, which results in transcription of RNA and subsequent translation
of the protein beta-galactosidase. Thus, in response to an environmental stimulus (lactose), the cell regulates transcription initiation in order to turn this particular gene on.

The precise kinetic mechanism of lac repression has been a subject of study for decades (1-4). Initial studies posited that lac repressor functioned by blocking transcription initiation. Other studies suggested that lac repressor and RNAP could be bound to the DNA at the same time, and that lac repressor had an inhibitory effect on the activity of bound RNAP. Kinetic schemes to describe this process were hypothesized, and experiments were conducted to provide constraints and distinguish between models. This process was repeated iteratively over decades, and a minimal kinetic mechanism for lac repression has emerged (3). This model describes how lac repressor works by describing each of the most important states and the rates of conversion between them. Such a kinetic model quantitatively describes the thermodynamics and kinetics by which lac repressor competitively binds to the promoter DNA to block transcription. Complete kinetic characterization of a system represents a gold standard of understanding in biophysical mechanism since such a model can be used to quantitatively predict the pathways of equilibration and timing of relaxation when the system is subject to perturbation. Lac repressor continues to be the subject of mechanistic study – very recent work (published over 50 years after the system was first described) has shown how lac repressor bindings modes are linked to sequence properties such as spacer length (4). This serves to illustrate that kinetic mechanisms are never fully solved, and new investigations can always be performed to illuminate new mechanistic details.
1.2 Regulation of mycobacterial transcription initiation

One of the most perplexing mysteries of how an organism regulates transcription in response to its environment is *Mycobacterium tuberculosis* (*Mtb*), the obligate pathogenic bacteria responsible for the disease Tuberculosis. The disease was first described in 1882 by Robert Koch as the “tubercle bacillus” since the bacteria were rod-like in shape (bacilli), approximately 2-4 microns in length and 0.2-0.4 microns in width. In 1905, Koch received the Nobel Prize in Physiology and Medicine for his work describing Tuberculosis. Over a century later, we have learned much about Tuberculosis; however, in spite of these great advancements in our knowledge of this organism, Tuberculosis remains a highly successful pathogen. *Mtb* is typically transmitted by aerosol and infects the lungs of humans, where it is subject to acute and chronic attack by the immune system. *Mtb* is initially encapsulated by immune cells that form a granuloma where *Mtb* is subject to an onslaught of environmental stresses including attacks on the mycobacterial cell surface, acidification of phagosomes, reactive nitrogen intermediates, reactive oxygen species, hypoxia, nutrient starvation, phosphate deprivation and DNA damage (5). Amazingly, *Mtb* has adapted strategies to survive these stresses including prevention of fusion of the phagolysosome and the transition into a dormant (latent) state. Latent tuberculosis infects at least one-third of the world’s population, and in many cases it reactivates resulting in over 1 million deaths per year (World Health Organization). *Mycobacterium tuberculosis* has evolved within humans for over 50,000 years, and *Mtb* continues to evolve resulting in an alarming emergence of multidrug resistant strains. As such, *Mtb* poses a major global burden to public health.
One of the primary reasons that *Mtb* is able to respond to these stresses is that it has evolved sophisticated methods for regulating transcription initiation (6). The genome of the paradigm strain of *Mtb* known as H37Rv has over 4 million base pairs and 3959 known genes. The expression of these genes are carefully tuned by hundreds of transcription factors which form a regulatory network that is just now starting to be mapped using computational methods from systems biology (6). However, there have been relatively few mechanistic studies on how individual transcription factors from *Mtb* affect the kinetics of transcription initiation.

In fact, transcription by the basal system of mycobacterial RNAP alone is not yet understood. Our understanding of bacterial transcription initiation has been largely shaped by decades of studies of *E.coli*, which has served as a useful model system. There are many similarities between *Eco* and *Mtb* transcription systems. For instance, RNAP from both bacteria consist of the same homologous subunits α₂ββ’ω, and the promoter architecture is generally the same (-35 hexamer, spacer region, -10 hexamer, discriminator sequence, transcription start site) (7). However, there are differences between the two systems that could have important ramifications with regard to their respective mechanisms of transcription initiation. For example, mycobacterial promoters do not have a strong consensus -35 region, and they have a more well-defined “extended -10” motif (7). So far, the limited kinetic studies of mycobacterial RNAP have been performed using low-resolution studies and have been interpreted assuming that initiation proceeds through the same kinetic mechanism as *Eco*RNAP (8). Mycobacterial
promoters exhibit significantly reduced activity in *E.coli*, suggesting that *Eco*RNAP cannot initiate efficiently from mycobacterial promoters (9, 10). However, direct comparisons of *in vitro* transcription initiation kinetics between *Eco*RNAP and mycobacterial RNAP have not been performed. Such studies would provide valuable information on how transcription systems in the two organisms can be compared. Understanding the mechanisms of transcription initiation in *Mtb* and how this process is regulated by essential transcription factors represents an urgent biomedical research challenge, so that we can develop new strategies to treat tuberculosis.

Formation of the transcription “open complex” (RPₜ) has been recognized as a rate-limiting step in transcription initiation in several bacterial systems including the model system of *E.coli* (11-13) This is especially true at ribosomal promoters, which are known to have unstable open complexes in *E.coli* and *Mtb* (14, 8). Transcription of ribosomal RNA (rRNA) is tightly linked to the growth rate of bacteria, and under certain conditions represents the majority of all transcription within the cell (15). Open-complex stability of rRNA promoters in *E.coli* is determined by a number of factors including GC content, interactions with sigma factor, initial nucleotide concentration, (p)ppGpp, DksA and Fis (16, 17). However, DksA is not conserved across all bacteria, and mycobacteria have evolved different transcription factors to regulate transcription of rRNA (18). Thus, the unique mechanisms by which *Mtb* regulate open-complex formation at rRNA promoters represents an important research challenge not only for understanding the ability of *Mtb* to grow and survive in a range of environmental conditions, but also for understanding bacterial transcription in non-model organisms.
1.2.1 CarD

CarD is a recently discovered transcription factor that is essential to survival and pathogenesis of *Mycobacterium tuberculosis* (11). CarD is named for its similarity to a protein in *Myxococcus xanthus* that helps produce carotenoid pigments in response to blue light (Car\(^+\) phenotype, D locus) (19). CarD is highly conserved among the Mycobacteriacea family as well as the larger Actinobacteria phylum. The phylogenetic distribution of CarD also includes but is not limited to Proteobacteria, Spirochaetes, Deinococcus-Thermus, and Firmicutes (20). Importantly, CarD is not found in *E.coli*, nor is it found in humans (11, 13). Although structurally and evolutionarily unrelated, CarD’s role in transcription initiation has been functionally compared to TFIIE in eukaryotes and TFE in archaea (21). CarD is required for *Mtb* to survive under a range of stresses including reactive oxygen species, nutrient starvation, and DNA damage (5).

CarD consists of 162 amino acids that fold into two distinct structural domains. The N-terminal domain is known as the RNAP-Interacting Domain (RID), since it binds the \(\beta_1\) lobe of the \(\beta\) subunit of RNAP as shown by a bacterial two-hybrid assay (22). The C-terminal Domain (CTD) binds DNA non-specifically, and includes a tryptophan residue that is essential for functionality (22, 23). Interestingly, CarD has been directly linked to treatment of Tuberculosis because weakening CarD’s RID domain results in an increase in the efficacy of Rifampicin, an antibiotic that is used in the first line-of-defense (22). This underscores the possibility of using our understanding of CarD to improve clinical therapies.
In ChIP experiments, CarD is found at promoters throughout the mycobacterial genome, suggesting that CarD is a global regulator of transcription initiation (20). Furthermore, CarD has been shown to be important for regulating transcription of ribosomal RNA (rRNA), and is found at both promoters of the Mtb rRNA operon (20, 24, 11). CarD activates transcription from Mtb \textit{rrnAP3 in vitro}, however the mechanism of this activation is unknown (20). It is possible that CarD could affect any of the steps of transcription including initiation, elongation and termination.

Depletion of CarD leads to increased rRNA, which led to the initial hypothesis that CarD represses rRNA transcription in a manner that is functionally similar but structurally distinct from that of \textit{EcoDksA} (11). However, CarD mutants with weakened interactions with RNAP and DNA result in decreased levels of 16S rRNA (20), indicating that CarD may activate transcription. This \textit{in vivo} evidence that CarD activates transcription of rRNA is consistent with \textit{in vitro} studies in which CarD increases transcription from the \textit{Mtb rrnAP3} promoter (20). Importantly, CarD’s activation of \textit{in vitro} transcription was observed in the presence of competitor DNA, demonstrating that CarD stabilizes promoter-bound transcription competent complexes (20). CarD activation of transcription depends not only on residues in its RID domain known to interact with RNAP, but also on interactions between CarD’s CTD and DNA, including a highly conserved solvent-exposed tryptophan (\textit{TthCarD W86, MtbCarD W85}) (23, 20). Structural modeling of CarD with the transcription initiation open complex predicts that this tryptophan interacts with the upstream edge of the transcription bubble, leading to the
natural hypothesis that CarD stabilizes the mycobacterial transcription open-complex (20).

In my thesis work, I have studied the mechanism of CarD’s effect on transcription initiation at the \textit{Mtb rRNA}P3 promoter. Possible kinetic mechanisms by which CarD could activate transcription initiation include but are not limited to stabilization of bound complex, acceleration of promoter opening, and deceleration of promoter closing, and acceleration of promoter escape. The relative values of these rates ultimately give rise to transcriptional flux by RNAP from a promoter. As such, one of the primary goals of my doctoral work was to measure the effect of CarD on the energetics of promoter opening, in order to quantify thermodynamic and kinetic properties of CarD-regulated transcription initiation.

\section*{1.2.2 \textit{RbpA}}

\textit{RbpA} (\textit{RNAP-binding protein A}) is a small (14 kDa) RNAP-binding protein that was originally discovered in \textit{Streptomyces coelicolor} (\textit{Sco}), and identified as an important regulator of growth and Rifampicin resistance (25, 26). \textit{RbpA} is confined to and widespread within the Actinobacteria phylum, and it is essential for survival of \textit{Mtb} (27, 28). One important role of \textit{RbpA} that has been the subject of several recent studies is its ability to regulate the assembly of a subset of mycobacterial sigma factors with core RNAP. \textit{RbpA} binds group 1 (SigA) and certain group 2 (SigB) mycobacterial sigma factors, and increases their affinity for core RNAP (27, 29, 30, 28). In this manner, \textit{RbpA}
can selectively drive transcription from a subset of promoters in the mycobacterial genome by tuning the assembly of core RNAP with different sigma factors. RbpA has also been linked to Rifampicin activity, although the nature of this relationship is unclear. Initial studies indicated that \textit{Msm}RbpA could rescue SigA-dependent transcription from the \textit{rel} promoter in the presence of 100 \textmu{}M Rifampicin (27, 31). Combined with cross-linking studies that localized RbpA to a location in the active cleft of RNAP (\textit{Tth}RNAP beta-subunit R381), these results led to a hypothesis that RbpA competes with Rifampicin for binding within the active site of the transcription complex (the so-called “exclusion model”) (32). This model is consistent with \textit{in vitro} transcription assays performed by a different laboratory in which RbpA’s presence increases Rifampicin tolerance (27). However, the exclusion model is not consistent with the fact that RbpA does not affect the IC50 of Rifampicin (27). Furthermore, RbpA binding core RNAP was alternately mapped to the beta-subunit Sandwich-Barrel Hybrid Motif, which does not overlap with the Rifampicin binding site (33, 27). As a result, a different model has emerged in which RbpA and Rifampicin do not compete for binding, and any relationship between the two is indirect. Possible reasons for the apparent discrepancy in these reported results include the promoter being studied, the species of the reagents being used, and the conditions under which the assays were performed. Further studies are needed to elucidate the precise nature of RbpA’s interactions with core RNAP and how these may be related to linkage between RbpA and Rifampicin. Although RbpA binds to RNAP holoenzyme in the absence of DNA, RbpA’s activity is promoter-dependent (30).
Regulation of rRNA is hypothesized to be the driving force behind its link to growth rate (26). Similarly to CarD, *in vitro* transcription assays have demonstrated that RbpA activates transcription from the *Mtb* *rrn*AP3 promoter (27). Initial studies indicated that RbpA is important for stabilizing promoter-bound RNAP at the *Mtb* *rrn*AP3 promoter (27). Furthermore, using KMnO₄ footprinting, RbpA was required for the formation of open-complex by SigB-holoenzyme at the *sigA* promoter (30). However, the effect of RbpA on promoter-opening at the *Mtb* *rrn*AP3 promoter has yet to be studied.

Intriguingly, the fact that CarD and RbpA both activate transcription from the *Mtb* *rrn*AP3 promoter introduces the possibility that the two factors could potentially interact with the transcription initiation complex simultaneously. This possibility has been hypothesized, but never directly studied (27, 33). The interplay between CarD and RbpA could lead to a range of outcomes, not only with regard to open complex stability at the *Mtb* *rrn*AP3 promoter, but also on the overall flux of transcription from promoters throughout the mycobacterial genome.
Figure 2: Structure of CarD from *Thermus thermophilus* (20, 34). Top: Schematic of CarD showing the N-terminal RNAP-interacting domain (RID, pink) and the DNA-binding C-terminal Domain (CTD, green). Bottom: The N-terminal domain is almost entirely beta-sheet in structure, and the C-terminal domain is entirely alpha-helical.
**Figure 3: Structure of RbpA** (28, 33, 34). Top: Schematic of RbpA shows that it consists of an N-terminal “core” domain (white), connected by a basic linker (BL) to the C-terminal sigma-interacting domain (SID, blue). Bottom, left: An NMR structure of the N-terminal core domain indicates that it primarily consists of 4 anti-parallel beta strands. Bottom, right: A co-crystal structure of the alpha-helical RbpA-SID (purple) with subdomains of sigma-factor (orange and magenta).
Figure 4: How do the essential transcription factors CarD and RbpA affect the mechanism of mycobacterial transcription initiation?

Top: A general scheme for bacterial transcription initiation is shown on the top row, where RNA Polymerase (R) binds Promoter DNA (P) to form the closed complex (RPc) which isomerizes into the open complex (RPo), incorporates initial nucleotides to form the initial transcribing complex (RPn), and then irreversibly transitions into an elongation complex (RPe). Transcription factors such as CarD and RbpA (represented by X) could potentially bind to any of these states (represented by the box enclosed by the dashed line), affect their stabilities, and affect the rates of conversion between states.

Bottom: A free energy diagram aligned with the kinetic scheme (Top) in which states (numbered 1-5) and transition states (1* through 4*) are depicted as local minima and maxima along a reaction coordinate. A lower ΔG indicates a more stable state, and a lower activation barrier (difference between a state and a transition state) results in a faster rate of conversion between states. The dashed line represents the effect of a hypothetical transcription factor that stabilizes states 2 and 5, and stabilizes the transition states 3* and 4*. All of these effects change the kinetics of transcription initiation.
1.3 Biophysical methods: experimental and theoretical considerations

1.3.1 Ensemble and single-molecule approaches

Fluorescence spectroscopy using modified promoters has been used to study open-complex formation in many transcription systems including \textit{E.coli} (35-38), T7 RNAP (39, 40), and mitochondrial RNAP (41). Fluorescence offers a distinct advantage over alternate approaches such as potassium-permanganate footprinting because the readout provides a real-time signal for the equilibration of the system, and therefore includes information on the kinetics of open-complex formation and promoter escape (37). One common challenge of interpreting fluorescence data arises from assigning structural states to different phases of fluorescence enhancement. For example, Cy3 is an organic dye that is frequently used in protein-induced fluorescence enhancement (PIFE) experiments. The quantum yield of Cy3 can dramatically increase or decrease depending on its microenvironment.

A real-time fluorescence assay for measuring RP\textsubscript{o} formation and promoter escape was developed for the \textit{E.coli} transcription system that utilizes a Cy3 fluorophore attached to the +2 non-template dT (37). The signal largely reports on open complex formation in real-time, allowing for one to measure observed rates of open complex equilibration. The signal-over-background far exceeds that of other fluorophores (i.e. 2-aminopurine), which allows for assays to be performed at low concentrations of Cy3-promoter (i.e. 10 nM). I have used this approach to study how transcription factors essential to \textit{Mtb} affect
transcription initiation at the major promoter of the ribosomal operon in *Mtb* (Figure 5) (24, 37, 42, 34).

Single molecule approaches to biology offer complementary strengths to ensemble approaches, since they are able to provide information on sample heterogeneity that may be lost in ensemble averaging (43). Single molecule technique can leverage many types of spectroscopic signals. Perhaps the first earliest published single molecule measurement used electric current to detect discrete steps in “Excitability Inducing Materials”, which later turned out to be single channel recordings of ion channels (44). Fluorescence is commonly used for a range of single molecule techniques including but not limited to total-internal reflection fluorescence spectroscopy (TIRF), fluorescence correlation spectroscopy (FCS), and many super-resolution imaging techniques (45, 46). Force spectroscopy can be performed by many instruments including atomic force microscopes, optical tweezers, and magnetic tweezers (47-50). Notably, of all of the techniques mentioned here, magnetic tweezers uniquely allow for the study of promoter unwinding in real-time on supercoiled DNA tethers, and they have been used to study all phases of transcription initiation in the presence and absence of transcription factors (51-54).
Figure 5. Schematic of the *Mtb rrnAP3* promoter used in this study. Nontemplate and template strand sequences for the *Mtb rrnAP3* promoter are shown along with the predicted region of strand separation upon open-complex formation with RNAP. Promoter elements are shown including the -35 hexamer and -10 hexamer (both shown in red), 18 base pair spacer region, extended -10 (Ext -10, blue), +1 transcription start site (TSS, green), and the location of the Cy3 fluorophore which is covalently attached to the +2 nontemplate dT.
1.3.2 Energetic considerations of biophysical mechanism

*Thermodynamic measurements of binding and linkage*

State functions in thermodynamics describe path-independent properties of a closed system at equilibrium that do not depend on how the system arrived at that state. The Gibbs free energy is such a state function that describes the overall stability of a given state and for binding reactions, the Gibbs free energy is related to the binding affinity of the two reactants by $\Delta G = -RT \ln K_{eq}$, where $K_{eq} = \frac{k_{on}}{k_{off}}$. In this manner, binding affinities provide information about the concentration dependencies of molecular interactions and inform as to the relative stabilities of different states. At the time I began my doctoral research, there were no quantitative measurements of the binding affinities of CarD or RbpA to RNAP and DNA. Therefore, quantitative measurements of binding affinities represented an important goal towards characterizing the thermodynamic properties of these transcription factors.

Chemical linkage describes the notion that the presence or absence of a chemical component of the system directly affects the binding affinity between two molecules. Linkage is also known as cooperativity, and it can be negative (resulting in weaker affinity), positive (resulting in higher affinity), or zero (no effect on affinity). In its simplest case, linkage can be quantitatively described by a single quantity $\alpha$, which represents the ratio of affinities in the presence and absence of the third component. In the case of ternary complex formation, $\alpha$ represents the ratio of affinities of a ligand to receptor in the presence or absence of the third component.
**Kinetic signatures for inferring mechanism**

When attempting to dissect mechanisms of binding and conformational exchange, one must perform kinetic experiments to measure observed rates, and study how these observed rates depend on concentrations of the reactants. The dependence of observed rate on concentration produces a “kinetic signature” for a given reactant, which can be used to infer quantitative details about the underlying kinetic mechanism. Perhaps the most readily interpretable kinetic signature is for that of binding, which is a linear dependence of observed rate on reactant concentration. The slope of this line gives the on-rate (bimolecular rate of association, units M$^{-1}$s$^{-1}$), while the y-intercept gives the off-rate (unimolecular rate of dissociation, units s$^{-1}$).

When considering kinetic signatures that are the manifestation of bimolecular binding coupled to unimolecular isomerization, two limiting models are often considered. “Conformational selection” describes conformational exchange that occurs prior to ligand binding, whereas “induced fit” describes the case where ligand binding occurs prior to conformational exchange. Typically, the kinetic signature of induced fit mechanisms manifest as a hyperbolically increasing observed rate with ligand concentration. Conformational selection mechanisms can produce hyperbolically increasing or decreasing observed rates with ligand concentration, depending on the relative values of the rate-constants in the system (55).
The co-existence of conformational-selection and induced-fit in series is frequently used to explain the kinetic signatures of molecular interactions (38). Interestingly, recent analyses of thermodynamic cycles have demonstrated that they can also co-exist in parallel (39). However, kinetic signatures for the parallel coexistence of conformational selection and induced fit are rarely observed and have yet to be theoretically explored in a general sense. A thermodynamic cycle describes a closed loop of states, which can be connected by binding or conformational exchange. Several thermodynamic cycles can be found in the general mechanism of transcription initiation linked to transcription factor binding as shown in Figure 4. For example, the four states comprising RP_c, RP_o, X-RP_c and X-RP_o form a cycle that is connected by binding (+/- X) and conformational exchange (closed to open). Since free energy is a state function and therefore path-independent, the sum of free energy change around the cycle must equal zero. This property can be used to constrain possible combinations of kinetic rate constants that can be used to describe a thermodynamic cycle. For example, the product of the forward rates must be equal to the product of the reverse rates (56). As such, a thermodynamic cycle can be used to constrain kinetic parameters when exploring different models.
1.4 Scope of dissertation

Prior studies of the transcription factors CarD and RbpA have demonstrated that they are both essential in the deadly pathogen *Mycobacterium tuberculosis* (*Mtb*), and that they can activate transcription from promoters for ribosomal RNA. Therefore, the mechanisms by which CarD and RbpA affect the kinetics of mycobacterial transcription initiation have profound implications with respect to our ability to understand and treat this disease.

I have studied the effect of CarD and RbpA on transcription initiation the *Mtb* *rrnAP3* promoter and my results are reported in the following chapters. In Chapter 2, I describe binding assays to characterize the concentration dependencies of transcription complex assembly. In Chapters 3 and 4, I detail my use of a real-time ensemble fluorescence assay to characterize the effect of CarD and RbpA on *Mtb* *rrnAP3* promoter open-complex formation *in vitro*. In Chapter 5, I detail my evaluation of multiple kinetic models that may be consistent with the data. In addition to pre-initiation studies, I have performed ensemble fluorescence studies in the presence of competitors and NTP’s to study both dissociation and promoter escape which are described in Chapter 6. In Chapter 7, experiments performed using glutamate instead of chloride as the primary anion are discussed which highlight the potentially dramatic effects of solution conditions on the mechanisms by which transcription initiation is regulated. Lastly, in Chapter 8, I describe magnetic-tweezers assays in which I observe promoter opening in real-time at the single-molecule level for the first time with the transcription initiation system from mycobacteria.
Taken together, this dissertation describes the energetics of CarD and RbpA interacting with the mycobacterial transcription initiation machinery, and poses mechanistic models that could explain how CarD and RbpA affect mycobacterial transcription open-complex stability. These models utilize fundamental thermodynamic and kinetic principles to explain measurements of binding, linkage, and observed rates. This work provides a quantitative framework for understanding the regulation of transcription initiation in *Mtb*, not only at the operon for ribosomal RNA but also genome-wide.
1.5 References


in mycobacteria. transcription, 2, 15–18.


Chapter 2

Binding studies of DNA, RNAP, and CarD
Introduction

Regulation of the housekeeping mycobacterial transcription initiation system depends on binding interactions between multiple protein and DNA components – Sigma factor A (SigA), core RNAP, promoter DNA, and transcription factors such as CarD and RbpA. Binding affinities provide information regarding the concentration-dependence of molecular interactions, and can be used to inform experimental design by providing predictable conditions for macromolecular assembly. Therefore, I sought to measure binding affinities between *Mycobacterium bovis* (*Mbo*) RNAP, CarD, and the *Mycobacterium tuberculosis* (*Mtb*) *rrnAP3* promoter. The experiments performed in this chapter are largely exploratory in nature. The trends presented here were reproducible across multiple experiments. However, since exact concentrations of DNA, RNAP and CarD were not consistently used across experiments, error estimates are not included, and estimates of affinities are based on individual data sets.

The specific DNA substrate I chose to study was a 150 bp promoter fragment containing the *Mtb rrnAP3* promoter fragment with an internal Cy3 label, since I was planning to use this substrate for measuring promoter opening and escape in real-time (1). To probe the assembly of the mycobacterial transcription initiation machinery on this promoter DNA template, I measured binding interactions between CarD, RNAP, and DNA using an electrophoretic mobility shift assay (EMSA) and a double-filter binding assay. CarD-DNA interactions have been previously studied. It has been shown that CarD prefers to shift longer pieces of DNA (Christina Stallings, personal communication). Therefore, it is possible that CarD’s binding affinity to a 150 bp promoter fragment represents a lower
bound on its affinity to genomic DNA. I measured the concentration dependence of CarD-binding to this specific promoter fragment in order to inform the design of ensemble fluorescence experiments that use this exact promoter template (2, 3).

To test how the affinity of mycobacterial RNAP holoenzyme for promoter DNA might be affected by residues in SigA important for open-complex formation, I wanted to test holoenzyme formed with a mutant SigA that could recognize promoter DNA but with reduced ability to open the DNA. In *Escherichia coli* (*E. coli* or *Eco*), residues in Sigma70 region 2.3 are thought to play a role in promoter opening by stabilizing unstacked bases in the transcription bubble. For example, tyrosine 430 is thought to stabilize the open conformation of the -11 non-template adenosine, where promoter opening is nucleated (4-6). The development of a mycobacterial holoenzyme capable of promoter-recognition and binding but deficient in promoter-opening will provide a unique ability to probe properties of the closed-complex. To this end, I conducted experiments using a holoenzyme formed with a mutant of SigA called “SigA-FYWW”, in which 4 highly conserved aromatic residues in SigA were mutated to alanines. Our working hypothesis is that holoenzyme formed with SigA-FYWW (*MboRNAP-FYWW*) is capable of promoter-specific binding, but is deficient in promoter opening.

**Methods**

*Preparation of Cy3-labeled promoter fragments*

Promoter fragments were made as previously described (see Appendix A for detailed protocol) (7).
**Preparation of proteins**

Preparation of *Mbo* core RNAP, *Mtb* SigA, were prepared as described (7). Preparation of *Mycobacterium smegmatis (Msm) CarD* and *Mtb CarD* were prepared as described (7). Cloning of *Mtb* SigA-FYWW was performed by Ashley Garner, and purification was performed by Ana Ruiz-Manzano.

**Electrophoretic Mobility Shift Assays (EMSA)**

Native Tris-Borate-EDTA (TBE) gels were used to test protein-nucleic acid binding interactions as previously described, with minimal modifications as follows (8). Precast 4-20% gradient non-denaturing TBE-PAGE gels (Invitrogen) were pre-run in 0.38x TBE buffer (100V / 90 min) to equilibrate the gel. The 0.38x TBE buffer was chosen in order to minimize the potentially destabilizing effect of salt on the binding-reaction while preserving the ability of the buffer to carry current. Promoter DNA (150 bp, containing *Mtb rrnAP3* promoter with Cy3 label) was mixed with *MboRNAP, CarD*, or both in 15 μl reactions, which were allowed to equilibrate for 1 hr at 25 C. The reactions were mixed with loading dye, each reaction was loaded into individual wells, and the gel was run at 100 V for 100 min.

Analysis of EMSA assays for estimating binding affinities were limited to those reactions performed in the absence of competing nucleic acid, which can affect measurements of CarD-DNA binding affinity and obscure the potential presence of nuclease contaminants in the reaction (see Appendix B). EMSA experiments with promoter DNA, MboRNAP
holoenzyme, and CarD were performed multiple times. However since each individual experiment did not necessarily use the exact same concentrations of reagents, these data sets were not combined and used to estimate error. Importantly, the trends from all of these experiments were consistent and sufficiently captured in the experimental data shown in Figures 1-3. There was significant difficulty in quantifying the data by densitometry due to material stuck in the wells of the gel and streaking of the gel. For this reason, estimates of apparent binding affinities are approximate.

**Double-filter binding assays**

Double-filter binding assays were performed as previously described with the minimal modifications as described below (9, 10). Nitrocellulose membranes were pre-treated by soaking for 10 min in 0.4 M KOH. After rinsing the nitrocellulose filters with water, both nitrocellulose and nylon filters were soaked in reaction buffer without bovine serum albumin (BSA) for 1 hr at 4°C.

**Reaction buffer:**

- 60% MQ H2O
- 10% 10x Mycobacterial transcription buffer +DTT/+BSA ("MTB +/+")
  ...also make MTB +/- (no BSA) for filter soaking and wash steps
- 20% RNAP Storage Buffer (10 mM Tris pH 8, 200 mM NaCl, 25% glycerol)
- 10% CarD Storage Buffer (20 mM Tris pH 8, 150 mM NaCl)

**Final glycerol = 5% (v/v)**

**Reaction volume:** 30 ul
Reaction mix:

6 ul RNAP Storage Buffer (up to 400 nM [RNAP]final)
3 ul CarD Storage Buffer (up to 11 uM [CarD]final)
3 ul 10x MTB +/+ 
18 ul MQ H2O/DNA for up to ~10nM [DNA]final, can use as low as 1nM

The concentration of DNA was chosen based on the minimal concentration that provided sufficient signal-to-noise for binding (Appendix B, Figure 2). Nitrocellulose and nylon membranes were loaded into a dot-blot apparatus and subjected to a 200 μl wash with reaction buffer without BSA (4x with 50 μl multichannel pipette). House vacuum was applied for ~10 s per load. Next, the 30 μl reaction was loaded into each well (1 x 30 μl multichannel pipette), followed by application of the vacuum for ~10 s until the reaction was pulled through the membranes. Next, each well was subject to another 200 μl wash with reaction buffer (4 x 50 μl, 10 s vacuum/load). The membranes were then removed from the dot-blot apparatus and placed in a glass tray. The filters were each gently dried with a hair drier, taking caution to hold the filters at their edges with a pair of tweezers. The filters were then dried for 1 hr in an oven set to 150 C. The dried filters were imaged using Typhoon Imager, operated with the Cy3 settings (Ex 532 nm / Em 560-600 nm). Densitometry was performed using ImageQuant, and the dot-blot grid feature was used to manually select spots. Plotting was performed in Matlab.

Interestingly, end-labeled DNA constructs stuck to the nitrocellulose, while internally labeled DNA did not. In addition, I found that pre-treatment of the nitrocellulose with 0.4
M KOH was required to observe higher signal on nylon than nitrocellulose (see Appendix B). For these reasons, analysis of double-filter binding assays for estimates of binding affinities were limited to experiments performed with internally-labeled DNA tested with nitrocellulose pre-treated with 0.4 M KOH. Filter-binding experiments with MboRNAP-FYWW were only done once. Filter-binding experiments with promoter DNA, core RNAP, MboRNAP holoenzyme, and CarD were performed multiple times. However since each individual experiment did not necessarily use the exact same concentrations of reagents, these data sets were not combined and used to estimate error. Importantly, the trends from all of these experiments were consistent are sufficiently captured in the single set of experimental data shown in Figure 4.

**Results and Conclusions**

*The apparent $K_d$ of MboRNAP for promoter DNA is at least 120 nM*

Titrations of MboRNAP holoenzyme were performed over promoter DNA in order to measure binding affinity. EMSA measurements qualitatively indicated that [MboRNAP] ~ 120 nM corresponded to ~50% bound (Figure 1, bottom). Densitometry of this experiment (Figure 1, top) was difficult to fit to a binding isotherm due to poor signal-to-noise. Contributing factors include material stuck in the well and streaking within the gel. Double-filter binding experiments indicated that even higher concentrations of MboRNAP were required to reach ~50% fraction bound, and the apparent $K_d$ of MboRNAP holoenzyme to promoter DNA estimated from these experiments is at least ~400 nM (Figure 4 column 1, Figure 5 blue). The large differences between these measurements likely arise from many factors. Perhaps the most important is the fact that
neither measurement is a solution-based measurement, and perturbation of the binding equilibrium is likely taking place in between the initial binding reaction and detection of binding in a gel or on a filter.

The apparent $K_d$ of CarD to DNA is at least 1 $\mu$M

To study the affinity of CarD to a 150 bp fragment of DNA containing the $Mtb$ $rrnA_P3$ promoter, titrations of CarD over DNA were performed. The ability of CarD to non-specifically bind DNA was already known, but my specific objective was to identify concentrations of CarD that bind this specific promoter fragment in order to inform future ensemble fluorescence experiments (2, 3). Experiments performed with $Msm$ CarD indicated that [CarD] < 1 $\mu$M does not bind DNA as shown by EMSA (Figure 2, Appendix B Figure 3) and double-filter binding (Figure 4). In some cases a super-shift can be observed at sufficiently high concentrations of CarD (Appendix B Figure 3).
Figure 1: EMSA for measuring *Mbo*RNAP holoenzyme’s binding affinity for 150 bp promoter DNA. Bottom: Native gel for probing *Mbo*RNAP-promoter binding. Lower band represents free DNA, which shifts to the top-half of the gel at higher concentrations of RNAP. Top: Densitometry for the top-half of the gel (RNAP-DNA) and bottom-half of the gel (DNA). These results provided qualitative estimates of *Mbo*RNAP binding affinity to promoter DNA (~120 nM), but were difficult to quantify and fit to a binding isotherm due to poor signal-to-noise.
Figure 2: EMSA for measuring CarD’s binding affinity for 150 bp promoter DNA. A 4-20% non-denaturing PAGE gel was used to measure CarD-DNA binding interactions. Lane 1 contained only the loading-dye. Lane 2 contained only 10 nM DNA. Lanes 3-12 are reactions containing 10 nM DNA and [CarD] increasing from 1 nM to 24 μM as labeled. CarD’s apparent K_d for this DNA substrate is estimated to be ~10 μM. No DNA-binding is observed at [CarD] = 1 μM.
Figure 3. CarD increases the affinity of RNAP for DNA. Titrations of RNAP in the absence of CarD only show shifted DNA when [RNAP] is increased from 140 nM to 280 nM (lanes 7 and 9). In the present of 1 uM CarD and 35 nM RNAP, most of the DNA has been bound (lane 4).
CarD increases the affinity of MboRNAP to promoter DNA

Titrations of MboRNAP over DNA indicated that the apparent affinity of MboRNAP for the *Mtb* *rrnP3* promoter fragment increases in the presence of CarD (Figures 3 and 4). This measurement of positive linkage between MboRNAP and CarD is perhaps not surprising, since CarD is known to bind both RNAP and DNA. Notably, CarD increases RNAP-promoter binding at sub-μM CarD concentrations (Figure 4 columns 1-6), which do not bind DNA in the absence of MboRNAP. This demonstration of positive linkage demonstrates that CarD has a higher affinity for MboRNAP-promoter complexes than for DNA alone.

The ability of sub-μM CarD to increase the affinity of MboRNAP holoenzyme for promoter DNA is linked to mutated SigA-FYWW

Interestingly, titrations of core RNAP (which binds DNA non-specifically) indicated that core RNAP could bind DNA, but these titrations were not responsive to sub-μM CarD (Figure 4, columns 7-9). These results demonstrate that CarD’s interactions with RNAP and DNA depend on the specific manner in which holoenzyme interacts with DNA. Furthermore, titrations of MboRNAP-FYWW indicated that MboRNAP-FYWW was able to bind DNA. Interestingly, more DNA is observed on the nylon filter at high concentrations of MboRNAP-FYWW compared to core RNAP (lanes 7-12, bottom filter). These results demonstrate that the presence of SigA-FYWW reduces the ability of RNAP to bind DNA compared to core alone. One possibility is that SigA-FYWW is forming holoenzyme with core RNAP, and that the promoter-specific binding of MboRNAP-FYWW binds DNA with lower affinity than core RNAP. Surprisingly, like
core RNAP, the titration of MboRNAP-FYWW was also not responsive to CarD (Figure 4, columns 10-12). These results suggest that CarD’s effect on holoenzyme-promoter complexes somehow depends on specific FYWW residues of wild-type SigA that were mutated. One possibility is that CarD directly interacts with these residues in the ternary complex of CarD-RNAP-DNA. Another possibility is that these residues promote a conformation of RNAP-DNA interaction that has a higher affinity for CarD, such as the RNAP-promoter open complex. This second possibility, which is also known as “conformational selection,” is explored and included in a model for CarD in which it has a higher affinity for open-complexes (Chapters 3 and 5).
Figure 4: Sub-μM CarD increases affinity of promoter DNA only for *Mbo*RNAP holoenzyme with wild-type SigA. Nitrocellulose filter (top) and nylon filter (bottom) are shown for a double-filter binding experiment using Cy3-labeled DNA, and titrations of RNAP and CarD. As CarD is increased from 0 to 1 μM (first 6 columns), the apparent binding affinity of *Mbo*RNAP drops from ~400 nM to ~50 nM, indicating positive linkage between CarD and *Mbo*RNAP. Titrations of Core RNAP (columns 7-9) and *Mbo*RNAP-FYWW (columns 10-12) each showed DNA-binding activity that was not sensitive to CarD.
**Figure 5: Fraction-bound calculated from double-filter binding RNAP holoenzyme titrations with increasing CarD.** Densitometry for Figure 4 columns 1-6 is shown above. These results suggest that in the absence of CarD (blue), only ~50% of the DNA is bound at 400 nM *Mbo*RNAP. However, with as little as 100 nM CarD, the apparent $K_d$ of *Mbo*RNAP for the promoter drops to ~50 nM (teal). This represents positive linkage, since the presence of CarD increases the apparent binding affinity of *Mbo*RNAP for promoter DNA.
Discussion and Future Directions

The results of these studies indicated that CarD does not shift DNA in an EMSA at sub-μM concentrations. These results also indicated that sub-μM CarD increases the affinity of *Mbo*RNAP for the *Mtb* rrnAP3 DNA, demonstrating that CarD’s affinity to RNAP-promoter complexes is higher than the affinity of CarD to DNA alone. Interestingly, titrations of core RNAP over DNA were not responsive to CarD. Furthermore, titrations of *Mbo*RNAP-FYWW were also not responsive to sub-μM CarD. These experiments suggested for the first time that interactions of sub-μM CarD with *Mbo*RNAP-promoter complexes are not only linked to holoenzyme but also to specific residues of SigA.

One possible explanation for this dependence could be that CarD’s ability to increase the affinity of *Mbo*RNAP holoenzyme for promoter DNA depends on the promoter-open complex that is formed with wild-type holoenzyme. The FYWW residues in SigA are highly conserved across all housekeeping bacterial sigma factors and are important for open-complex formation in *Eco*RNAP (4, 5). All four residues are aromatic and are thought to interact with unpaired bases that become flipped-open during the course of promoter opening (4-6). Assuming these residues perform the same function in mycobacterial RNAP holoenzyme, the lack of linkage between CarD and *Mbo*RNAP-FYWW suggests that CarD’s mechanism under these conditions is tightly linked to the open-complex. This model is explored further in the context of a thermodynamic cycle in Chapters 3 and 5.
Suggested future directions are to more comprehensively characterize the binding affinities within the mycobacterial transcription initiation complex, including all possible combinations of sigma, core, holoenzyme, DNA and transcription factors. The development of solution-based binding assays such as fluorescence anisotropy, fluorescence correlation spectroscopy, isothermal titration calorimetry and analytical ultracentrifugation may prevent complications encountered here using non-equilibrium methods such as EMSA and double-filter binding (13). Furthermore, such solution-based assays offer greater flexibility to probe how the thermodynamics and kinetics of binding depend on solution conditions such as pH, temperature and salt (13).

I hypothesize that CarD’s affinity for mycobacterial transcription components increases qualitatively in the following manner: Sigma < RNAP < DNA < RP<sub>c</sub> < RP<sub>o</sub>. The rationale for this hypothesis is as follows. Although CarD and Sigma are both found at promoters throughout the mycobacterial genome, there is no evidence that they interact directly in solution. Their overlapping ChIP-Seq profiles are likely to due to the fact that both dissociate from RNAP upon promoter escape (2). The CarD-RNAP interaction has been directly demonstrated using the bacterial two-hybrid approach, but has not yet been studied using purified proteins in solution (14). The CarD-DNA interaction has been characterized in solution, and has an apparent K<sub>d</sub> of at least 1 μM. CarD likely has a higher affinity to RNAP-promoter complexes than either RNAP or DNA alone, because these complexes provide multiple surfaces for interaction with CarD’s RID and CTD. Lastly, CarD must have a higher affinity for RP<sub>o</sub> than RP<sub>c</sub> since it stabilizes open-
complexes (see Chapters 3 and 5). Empirical measurements of these affinities will provide valuable constraints for the further development of mechanistic models.
References


Chapter 3

Mechanisms of CarD-regulated mycobacterial transcription initiation
Introduction

When I began my doctoral research, CarD’s effect on mycobacterial transcription open-complex stability was unknown. *In vitro* transcription assays indicated that CarD activated transcription at the Mtb *rrnAP3* promoter (1). Structural models predicted that CarD interacted with RNAP and the upstream edge of the transcription bubble (1). These results led me to hypothesize that CarD stabilizes mycobacterial transcription open complexes.

I used stopped-flow fluorescence spectroscopy to directly demonstrate that CarD stabilizes mycobacterial transcription open complexes at the *Mycobacterium tuberculosis* *rrnAP3* promoter. This work was published in Nucleic Acids Research in February of 2015 and is included here as published (2). My finding that CarD stabilizes mycobacterial open complexes matched results reached independently by another research group using a different method (potassium permanganate footprinting), which was published 2 months earlier in December of 2014 (3). Importantly, my experiments included titrations of CarD and RNAP using an assay that reports on open complex formation in real-time, thus I was able to report valuable information on the concentration dependencies and kinetics of CarD’s effect on mycobacterial open complex stability. I found that CarD’s effect on the kinetics of open complex equilibration is concentration-dependent, and my data is consistent with a 5-state model in which CarD stabilizes RPo using a conformational-selection pathway at low concentrations and further stabilizes RPo using an induced-fit pathway at higher concentrations. This work
was the first mechanistic study of CarD’s effect on open complex stability, and was chosen by the editors of Nucleic Acids Research as a “Breakthrough Article,” which is a designation for work that “present[s] high-impact studies answering long-standing questions in the field of nucleic acids research and/or opening up new areas and mechanistic hypotheses for investigation.”

References


CarD stabilizes mycobacterial open complexes via a two-tiered kinetic mechanism

Jayan Rammohan1, Ana Ruiz Manzano1, Ashley L. Garner2, Christina L. Stallings2 and Eric A. Galburt1,*

1Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine, St. Louis, MO 63110, USA and 2Department of Molecular Microbiology, Washington University School of Medicine, St. Louis, MO 63110, USA

Received December 8, 2014; Revised January 20, 2015; Accepted January 21, 2015

ABSTRACT

CarD is an essential and global transcriptional regulator in mycobacteria. While its biological role is unclear, CarD functions by interacting directly with RNA polymerase (RNAP) holoenzyme promoter complexes. Here, using a fluorescent reporter of open complex, we quantitate RPo formation in real time and show that Mycobacterium tuberculosis CarD has a dramatic effect on the energetics of RNAP bound complexes on the M. tuberculosis rrnAP3 ribosomal RNA promoter. The data reveal that Mycobacterium bovis RNAP exhibits an unstable RPo that is stabilized by CarD and suggest that CarD uses a two-tiered, concentration-dependent mechanism by associating with open and closed complexes with different affinities. Specifically, the kinetics of open-complex formation can be explained by a model where, at saturating concentrations of CarD, the rate of bubble collapse is slowed and the rate of opening is accelerated. The kinetics and open-complex stabilities of CarD mutants further clarify the roles played by the key residues W85, K90 and R25 previously shown to affect CarD-dependent gene regulation in vivo. In contrast to M. bovis RNAP, Escherichia coli RNAP efficiently forms RPo on rrnAP3, suggesting an important difference between the polymerases themselves and highlighting how transcriptional machinery can vary across bacterial genera.

INTRODUCTION

The regulation of gene expression via the control of DNA transcription allows all living organisms to adapt their cellular biochemistry to changes in their environment. In bacteria, the transcriptional machinery consists of RNA polymerase (RNAP) holoenzyme composed of the catalytic core enzyme (ββ′αα′ω) and a dissociable sigma factor (σ) subunit that directs promoter recognition. Transcription initiation when RNAP holoenzyme recognizes a promoter sequence and forms the RNAP-promoter closed complex (RPc) where the DNA strands are still annealed in the duplex. Through a series of conformational changes the RNAP-promoter open complex (RPo) forms where the DNA strands are melted from the −10 to +2 positions and the active site is accessible to the initiating nucleotide. In Escherichia coli, the mechanism of open-complex formation includes multiple kinetic intermediates between the initially bound complex and the stable open complex (1). While the core machinery and major intermediate states are conserved across bacteria, the detailed kinetics of transcription likely vary between bacterial species to accommodate different physiologies and niches. Therefore, a minimal kinetic scheme describing promoter binding and opening by RNAP in two reversible steps (R + P ↔ RPc ↔ RPo) provides a useful starting point when investigating open-complex formation in a non-model system.

Regulation of transcription initiation is achieved through the modulation of the stabilities of intermediate states and/or the rates of exchange between these states on the pathway to promoter escape. Transcription factors can mediate this regulation by directly affecting the polymerase-promoter interaction, manipulating the equilibrium between RPc and RPo, or affecting rates of promoter escape (2,3). Much of what has been studied in terms of the mechanisms of transcription initiation and its regulation has used E. coli as a model system. However, it has become evident that some bacteria require specialized factors to allow for efficient gene transcription. For example, many of the se-

*To whom correspondence should be addressed. Tel: +1 314 362 5201; Fax: +1 314 362 7183; Email: egalburt@biochem.wustl.edu

© The Author(s) 2015. Published by Oxford University Press on behalf of Nucleic Acids Research. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com
quences and proteins required to regulate transcription initiation in E. coli are absent in mycobacteria, including Fis (4), DksA (5), AT-rich upstream activating elements (6,7) and GC-rich discriminator sequences (8). The question remains how bacteria missing these factors confer efficient and regulable transcription.

CarD is a recently discovered RNAP-binding protein that is conserved in numerous bacterial species but not present in E. coli. CarD homologs are essential in Mycobacterium tuberculosis, Mycobacterium smegmatis (9) and Myxococcus xanthus (10) and knockouts were not attainable in Borrelia burgdorferi (11). Furthermore, CarD is essential for the response of M. tuberculosis to oxidative stress and certain antibiotics, as well as the acute and persistent infection of mice (9,12). Given the tremendous impact of M. tuberculosis on global health (1.3 million related deaths a year, WHO, 2013), it is of particular interest to understand unique biochemical pathways required by this pathogen. Yet, the reasons for CarD essentiality remain enigmatic primarily due to a lack of understanding of its molecular mechanisms.

CarD is associated with all RNAP-σ70 (mycobacterial housekeeping sigma) homologous to E. coli σ70 transcription initiation complexes within the M. smegmatis genome (13) and interacts directly with the β1 region of the RNAP β subunit through its N-terminal RNAP interaction domain (RID) (9–10,12). The current model for CarD activity is that this protein is directed to promoters via its interaction with RNAP β (13,14). At the promoter, a basic patch in the C-terminus of CarD contacts the DNA and the association of RNAP-bound CarD with the promoter DNA stabilizes the RNAP-promoter complex (13,14). Within this stabilized CarD-RNAP-promoter complex, a conserved tryptophan in CarD is required for proper transcriptional regulation, however its exact role is still unknown. Together, these three activities of CarD (RNAP binding, DNA binding and the conserved tryptophan) promote a gene expression pattern from a pAC27 plasmid pAC27 M (Vivaspin 20, MWCO 30 kDa, GE Healthcare), flash frozen and stored at −80°C.

**MATERIALS AND METHODS**

**Protein purification**

*Mycobacterium bovis* core RNAP and σ70 were over-expressed and purified using methods slightly modified from the literature (13). *Mycobacterium bovis* core RNAP holoenzyme is identical to *M. tuberculosis* RNAP except for the 69th residue of β which is a proline in *M. bovis* and an arginine in *M. tuberculosis*. The *M. bovis* core RNAP subunits were co-over-expressed from plasmid pAC27 in BL21 (ADE3) pRARE2. Cells were grown at 37°C until an OD₆₀₀ of 0.8. Cells were induced with 250 μM Isopropyl β-D-1-thiogalactopyranoside (IPTG) and grown 4 h at 20°C. Cell pellets were lysed and loaded onto a 10 ml Ni²⁺ affinity column (HP HiTrap, GE Healthcare) using a 5–1000 mM imidazole gradient. The eluted sample was dialyzed, concentrated and further purified by size exclusion chromatography (HiPrep Sephacryl 300, Pharmacia). The peak of core RNAP was dialyzed in storage buffer (50% glycerol, 10 mM Tris pH 7.9, 200 mM NaCl, 0.1 mM ethylenediaminetetraacetic acid (EDTA), 1 mM MgCl₂, 20 μM ZnCl₂, 2 mM Dithiothreitol (DTT)), concentrated to 4 μM (Vivaspin 20, MWCO 100 kDa, GE Healthcare), flash frozen in liquid nitrogen and stored at −80°C.

*Mycobacterium bovis* σ70 was over-expressed from plasmid pAC27 BL21 (ADE3) pRARE2. Cells were grown at 37°C until an OD₆₀₀ of 0.8. Cells were induced with 250 μM IPTG and grown 4 h at 20°C. Cell pellets were lysed and σ70 was purified by Ni²⁺ affinity chromatography (HP HiTrap, Pharmacia) using a 5–1000 μM imidazole gradient. Eluted protein was dialyzed into storage buffer (10 mM Tris pH 7.9, 250 mM NaCl, 0.1 mM EDTA, 1 mM MgCl₂, 20 μM ZnCl₂, 2 mM DTT), concentrated to 24 μM (Vivaspin 20, MWCO 30 kDa, GE Healthcare), flash frozen and stored at −80°C.

*Mycobacterium tuberculosis* CarD was over-expressed from a pETSUMO plasmid (13). Cells were grown at 37°C until OD₆₀₀ of 0.8, induced with 1 mM IPTG, and grown 3 h at 37°C. Cell pellets were lysed and His-SUMO-CarD was purified by Ni²⁺ affinity chromatography (Ni-NTA Agarose, Qiagen). Elution at 250 mM imidazole was quantified and His-Ulp1 protease added to cleave the His-SUMO tag from CarD. This mix was dialyzed overnight in 3 liter of 20 mM Tris pH 8.0, 150 mM NaCl and 1 mM β-ME and dialyzed again in fresh buffer for four more hours. Ni²⁺ resin
was added to dialyzed proteins and, after incubating for 1 h at 4 °C, CarD was collected in the flow through. Protein was concentrated to 100 μM (Vivaspin 20, MWCO 3 kDa, GE Healthcare), flash frozen and stored at −80 °C.

Preparation of fluorescent promoter DNA fragments

The DNA template contained the −41 to +4 bases of the rrrnA3 ribosomal RNA promoter (16) of M. tuberculosis centered in a 150 bp template (Supplemental Information). Fluorescently labeled promoter DNA was prepared as previously described (17). In short, two 85-mer oligonucleotides with a 20 bp overlapping sequence were synthesized and high performance liquid chromatography (HPLC) purified (IDT). The 85-mer corresponding to the +2 thymine for covalent attachment to Cy3-NHS (Lumiprobe). Labeled oligonucleotides were purified from unlabeled oligos and excess Cy3 using reverse-phase chromatography (C18 column, XTerra). Template and Cy3-non-template strands were annealed and extended with Taq polymerase (Invitrogen). The extension products were then purified by HPLC (DNA Swift Column, Dionex). Finally, the pure Cy3-labeled extension products were spin-concentrated and exchanged into Tris-EDTA buffer (Millipore Amicon Ultra, MWCO 30 kDa). Concentrations of ds-DNA (A260) and Cy3 (A280) were measured by spectrophotometer (Nanodrop) and indicated a equimolar ratio of dye to DNA template demonstrating 100% labeling. Labeled DNA template was stored at −20 °C.

Stopped-flow assay for real-time monitoring of open-complex formation

Stopped-flow experiments were performed as previously described (17,18) with minimal exceptions. All experiments were performed by mixing equal volumes of protein solution from one syringe with fluorescent promoter DNA from the other syringe. Accordingly, initial syringe concentrations of protein and DNA were prepared at twice the desired final concentration for the reaction. Including contributions from protein storage buffers, final reaction buffer conditions were as follows: 14 mM Tris pH 8.0, 120 mM NaCl, 10 mM MgCl2, 1 mM DTT, 0.1 mg/ml bovine serum albumin (BSA) and 10% glycerol by volume.

Protein solutions for stopped-flow experiments were prepared as follows. In the case of MboRNAP, core was mixed with saturating concentrations of σ^* and allowed to equilibrate for 15 min at room temperature. In the case of EcoRNAP, σ^* saturated holoenzyme was used (New England Biolabs). When required, CarD was added to RNAP holoenzyme and allowed to equilibrate for an additional 15 min at room temperature. Protein solutions were then diluted into a solution containing 10 mM Tris pH 8.0, 40 mM NaCl, 10 mM MgCl2, 1 mM DTT and 0.1 mg/ml BSA. Since MboRNAP core storage buffer contained 50% glycerol, the protein solution contained 10% glycerol by volume. For all experiments, any absence of protein volume was balanced by its appropriate storage buffer in order to preserve identical buffer conditions for different reactions.

Promoter DNA solutions were prepared for stopped-flow experiments as follows. Stock DNA was diluted to 20 nM into buffer containing 10 mM Tris pH 8.0, 40 mM NaCl, 10 mM MgCl2, 1 mM DTT, 0.1 mg/ml BSA and 10% glycerol by volume. Glycerol was added to the DNA solution to match the glycerol content of the protein solution, to minimize rapid mixing artifacts that can occur with asymmetric viscosities between syringes.

Experiments were performed on an SX-20 stopped-flow spectrophotometer (Applied Photophysics, dead-time < 2 ms) with a shot volume of 150 μL, using excitation at 515 nm (monochromator), and emission at 570+ nm (long-pass filter). Unless otherwise noted, data were collected for 20 min and at 25 °C by sampling 10,000 points in a logarithmic decay over the time course of the experiment. A circulating water bath with active feedback was used to maintain temperature within 0.1 °C. Multiple buffer shots as well as DNA only shots were performed before every experiment. At least two shots were collected for every protein condition, which were averaged before plotting as fold change (FC) where FC = (F − F0)/F0, where F = experimental signal − buffer signal in volts and F0 = DNA signal − buffer signal in volts. In all figures, error bars of fold change represent the standard error of the mean (SEM) for conditions that were repeated multiple times on different days. From these SEMs, average errors were then applied to conditions that were only repeated multiple times on the same day to better estimate the real error for these points.

Triple exponential fits of kinetic traces to extract observed rates

Fold changes as a function of time were fit to a sum of exponentials using the ProData Viewer software (Applied Photophysics). Single and double exponentials were insufficient to capture the entire shape of the curve without substantial systematic residual error and triple exponential fits of the entire time course yielded consistent assignment of phases (fast, medium and slow fits were interchangeable between a1, k1, a2, k2, a3 and k3). To fit the curves with three exponentials while preserving assignment of fast, medium and slow phases, 1–1200 s was fit to a double exponential and the amplitudes and observed rates for the medium (a2, k2) and slow (a3, k3) phases were fixed. Then, a third exponent was added and the time for fitting was expanded to include 0.1–1200 s in order to capture the fastest phase (a1, k1). In this manner, the curves could be fit with three exponentials from 0.1 to 1200 s, and assignments of amplitudes and observed rates for all phases between traces remained consistent. In all figures, error bars of observed rates represent the SEM for conditions that were repeated multiple times on different days. From these SEMs, average errors were then applied to conditions that were only repeated multiple times on the same day to better estimate the real error for these points.

Kinetic simulations

Simulation of the kinetic model of open-complex formation was carried out in MATLAB R2014b (Mathworks). Model parameters were determined by manually adjusting rate constants and comparing with the data. The model was adjusted until the trends of equilibrium fluorescence and
observed rate as functions of RNAP and CarD concentrations were best fit. By assuming that the apparent bimolecular association rates were all equal, fixing the equilibrium between closed and open complex based on the MboRNAP data in the absence of CarD, and requiring that the model satisfied detailed balance, only three adjustable rate constants remained. This allowed for the manual exploration of parameter space and yielded a model that is able to capture the observed trends. The rate constants used to generate Figure 6 are listed in the Supplemental Information.

RESULTS
A real-time fluorescence-based assay of open complex

To detect the formation of open complexes, a Cy3 label was incorporated on the +2 nucleotide of the non-template strand of a M. tuberculosis rRNA promoter rrnAP3 DNA template via a NHS-C6 amide linkage (Figure 1A). Cy3 exhibits a 2-fold enhancement in fluorescence intensity in the open complex and has been used to monitor DNA opening and promoter escape in E. coli (17). The rrnAP3 promoter is the major ribosomal RNA promoter in M. tuberculosis (16), has been used previously in studies of CarD, and is thought to be regulated by CarD in vivo (13,14). Mixing of protein components with labeled DNA templates was performed via stopped-flow spectrophotometry and fluorescence was monitored for 20 min. The addition of 21 nM E. coli RNAP (EcoRNAP) sigma factor 70 (σ70) holoenzyme to rrnAP3 resulted in robust enhancement of the fluorescence intensity as compared to DNA alone (>60% increase or an enhancement of 0.6, Figure 1B). In contrast, very small (≈~0.05) fluorescence enhancements are observed when the same concentration of EcoRNAP is added to templates lacking the rrnAP3 promoter sequence (Figure 1B), demonstrating that fluorescence enhancement is promoter dependent. Furthermore, an enhancement of approximately 0.3 is observed at low temperature where open-complex formation is inhibited (19–21) and the fluorescence enhancement increases with increasing temperature consistent with the known temperature-dependence of open-complex formation (Figure 1C). A possible structural rationale for the specific dependence on open complex exhibited by the observed fluorescence enhancement is that the +2 position is satisfied detailed balance, only three adjustable rate constants remained. This allowed for the manual exploration of parameter space and yielded a model that is able to capture the observed trends. The rate constants used to generate Figure 6 are listed in the Supplemental Information.

M. bovis RNAP forms open complex less efficiently than E. coli RNAP on rrnAP3

We first studied the concentration dependence of EcoRNAP (σ70) and MboRNAP (σ70) holoenzymes on open-complex formation on the M. tuberculosis rrnAP3 promoter at 25°C. Fluorescence enhancements after the addition of 2–283 nM EcoRNAP (Figure 2A) or 37.5–600 nM MboRNAP (Figure 2B) were monitored for 20 min. Final fold enhancements were plotted versus RNAP holoenzyme concentration and fit to extract a concentration at which the enhancement is half-maximal (Keff; Figure 2C). As expected, the amount of open complex increases as concentration increases for both forms of RNAP, however EcoRNAP exhibits a Keff almost 10-fold smaller than MboRNAP (23 ± 5 nM versus 212 ± 43 nM, 95% confidence bounds). More strikingly, at saturating concentrations of the respective polymerases, the fluorescence enhancement for EcoRNAP is 1.44 ± 0.1 while MboRNAP only reaches an enhancement of 0.30 ± 0.1. The enhancement for EcoRNAP is comparable to that described previously for EcoRNAP on a consensus promoter (~1.4) (17), suggesting that this signal is indicative of fully open DNA (i.e. 100% open complex) on rrnAP3. By comparison, MboRNAP σ70 holoenzyme, even when fully occupying the DNA template at saturating concentrations of polymerase, is not capable of opening a large percentage of the promoters leaving the majority of bound complexes in the closed state. This demonstrates that MboRNAP forms a significantly less stable open complex as compared to EcoRNAP, even on its own mycobacterial promoter.

CarD stabilizes the open complex of MboRNAP

We next looked at the effect of CarD on open-complex formation. Holoenzyme was incubated with CarD for 15 min prior to mixing with the DNA template and recording changes in fluorescence intensity over time. While CarD shows no fluorescence enhancement when added to DNA alone and has little effect on the fluorescence signal from EcoRNAP (Supplementary Figure S1), its presence leads to a dramatic increase in the magnitude of fluorescence enhancements observed when incubated with MboRNAP (Figure 3A). As fluorescence enhancement serves as a reporter of open complex, we conclude that CarD significantly stabilizes open complex specifically for MboRNAP. This is consistent with known differences in the β1 lobe of EcoRNAP and MboRNAP on the CarD-binding interface that would predict tighter binding of M. tuberculosis CarD to MboRNAP than to EcoRNAP (23). In the presence of 225 nM MboRNAP, a rise in fluorescence fold change can be seen as the concentration of CarD is increased from 0 - 1.1 μM. The concentration of CarD that exhibits half-maximal effect (77 ± 35 nM) and the fold fluorescence enhancement over 0 μM CarD at saturation (6.2-fold) was determined by normalizing the equilibrium (final) fluorescence enhancements relative to 0 μM CarD and fitting to a binding isotherm (Figure 3B). Assuming that EcoRNAP generates 100% open complex at saturating concentrations (283 nM, Figure 2A) and that MboRNAP generates the same enhancement of signal per open complex, at saturating concentrations of CarD (≥ 1 μM) the percent of DNA
that are in open complexes at equilibrium in the presence of 225 nM $Mbo$RNAP increases from approximately 15% to 93% (Figure 3A).

To compare the $Mbo$RNAP holoenzyme concentration dependence in the presence and absence of CarD, a holoenzyme titration (37.5–600 nM) was performed in the presence of saturating CarD (1 μM) at 25°C (Figure 4A). Comparing RNAP titrations in the presence and absence of CarD shows that 1 μM CarD reduces $K_{eq}$ from 212 ± 43 nM to 106 ± 3 nM and increases the amount of open complex 4.3-fold at saturation (Figure 4B).

The ability of CarD to stabilize open complexes is also temperature dependent. In experiments with 100 nM $Mbo$RNAP in the presence and absence of 1 μM CarD, traces at different temperatures show different amounts of CarD-dependent fold change (Figure 4C). Specifically, the CarD effect is significantly larger at 25°C (4.2-fold) than it is at either 10°C (1.5-fold) or 37°C (1.9-fold). These observations fit well with the known temperature-dependence of open complex (19–21). At low temperature, opening is severely inhibited and the binding energy from CarD is not sufficient to tilt the landscape enough toward opening to have a large effect. At 25°C, the energy landscape is more balanced between closed and open and the extra binding energy supplied by CarD is able to significantly affect the equilibria between the two states. Lastly, at 37°C, open complex forms more readily than at lower temperatures and thus there is an appreciable amount of open complex to begin with (i.e. 50%). In this case, even if CarD is capable of stabilizing all bound complexes in the open form, it will exhibit a lower fold change in open complex equilibrium concentration (i.e. 2-fold).

**CarD pushes the equilibria of polymerase-bound states toward open complex**

In the presence of a saturating concentration of $Mbo$RNAP (450 nM) where the DNA templates are fully occupied by holoenzyme, the amount of open complex observed in the presence of CarD is dramatically increased in the presence of 1 μM CarD (Figures 2B and 4A). This can only be the case if CarD stabilizes the open complex relative to closed complex either by increasing the effective rate of opening or decreasing the effective rate of closing or both.

Furthermore, the observation that CarD stabilizes open complexes relative to closed complexes suggests that CarD has a higher affinity to open complex than it does to closed
complex. In a thermodynamic cycle (Figure 4D) where binding of CarD is coupled to the open/closed equilibrium, the sum of free energies around the cycle must be zero. The data show that RP _o_ is more stable relative to RP _c_ in the presence of CarD (ΔG _o_  <  ΔG _c_). Since the sum of free energies along paths that begin and end in the same state must be the same, ΔG _o_  + ΔG _c_ = ΔG _o_  + ΔG _c_. This then requires the free energy differences between RP and RP CarD in the open and closed states to satisfy ΔG _c_  <  ΔG _o_. Thus, as the free energy difference determines the affinity (K_d = e^(-ΔG/R^T)), we predict that CarD binds more tightly to open complex than to closed complex.

The kinetics of open-complex formation can be monitored via the slowest observed rate

The time-dependent traces allow for an analysis of the kinetics of open-complex formation. The traces are well fit by a triple exponential resulting in three observed rates, consistent with the well-known multi-state kinetic complexity of open-complex formation (1) (Figure 5A, Materials and Methods). The three observed rates differ by orders of magnitude with a fast (k^1_obs_ ∼ 1 s^{-1}), intermediate (k^2_obs_ ∼ 0.1 s^{-1}) and slow (k^3_obs_ ∼ 0.005 s^{-1}) observed rates. In traces exhibiting robust opening (i.e. high polymerase or high CarD concentrations), the amplitude of the curve is dominated by the amplitude of the slowest observed rate (Supplementary Figure S2). Furthermore, although promoter-less control traces (Figure 1B and D) do exhibit a modest amount of fluorescence enhancement, this enhancement occurs very rapidly (< 1 s) and the slow phase (k^3_obs_) is not observed. Therefore, we reason that, to a first-approximation, the fastest observed rates report on a protein–DNA interaction not related to opening (i.e. binding) and the slowest observed rate (k^3_obs_) specifically reports on the approach to equilibration of open complex for MboRNAP.

Observed rates typically represent combinations of microscopic rate constants. In the simplest case, the two states A and B are connected by a forward and a reverse rate constant. In this case, if one starts with the system entirely in A and monitors the approach to equilibrium, the observed rate will simply be the sum of the two rate constants. There are two mechanisms by which B could be stabilized relative to A: (i) the forward rate constant could be accelerated or (ii) the reverse rate constant could be slowed. In the first case, B would be stabilized and the observed rate would increase while in the second case, B would be stabilized and the observed rate would decrease. In more complicated kinetic mechanisms, the interpretation may not be so straightforward, but nonetheless, these two limiting cases serve as a useful backdrop for extracting mechanistic information from trends in an observed rate (Supplementary Figure S3).

CarD exhibits a concentration-dependent effect on the kinetics of open-complex formation

To extract information regarding the effect of CarD on the kinetics of open-complex formation, we first analyzed the CarD-dependence of k^3_obs_ using data collected at 150 nM MboRNAP (Figure 5B). Here, k^3_obs_ decreases with increasing CarD concentration suggesting that CarD slows a re-
Figure 3. CarD increases open-complex formation with MboRNAP. (A) Fluorescence enhancement as a function of time for 225 nM MboRNAP with increasing concentrations of CarD (0–5.55 μM) at 25°C. (B) Equilibrium values of fluorescence enhancement are plotted as a function of CarD concentration. A fit (solid line) of the data reveals a concentration of half-maximal effect $K_{eff} = 77 \pm 35$ nM and an amplitude of 6.2-fold over 225 nM MboRNAP alone.

Figure 4. MboRNAP titration at saturating CarD. (A) Fluorescence enhancement as a function of time for 1 μM CarD with increasing concentrations of MboRNAP (0–450 nM) at 25°C. (B) Equilibrium values of fluorescence enhancement fold change are plotted in the presence (red) and absence (green) of 1 μM CarD. The data were best fit with amplitudes of 1.3 and 0.28 and $K_{eff}$ of 106 ± 3 nM and 212 ± 43 nM in the presence and absence of CarD, respectively. (C) Observed traces for 100 nM MboRNAP in the presence or absence of 1 μM CarD at 10°C (blue), 25°C (green) and 37°C (red). The bar graph shows the fold increase in equilibrium fluorescence enhancement for each temperature. (D) The thermodynamic cycle linking closed complex (RPc), open complex (RPo), closed complex bound to CarD (RPcCarD) and open complex bound to CarD (RPoCarD). The equality shown below follows since the sum of the free energy differences along the two paths from RPc to RPoCarD must be equal.
Kinetics of open-complex formation. (A) Triple exponential fits are required to obtain fits with good residuals for traces where DNA opening is occurring. A typical trace, the fits, and the resulting residuals are shown for with double (red) and triple exponential (blue) fits. (B) The CarD concentration dependence of $k_{3}^{obs}$ in the presence of 150 nM (black) or 225 nM (red) MboRNAP. (C) The MboRNAP concentration dependence of $k_{3}^{obs}$ in the presence (black) and absence (red) of 1 µM CarD.

A kinetic model with CarD binding to both open and closed complexes

To understand the concentration-dependent effects of CarD, a kinetic model of open-complex formation in the presence of CarD was constructed (Figure 6A). The model consists of five states including unbound DNA ($R + P$), closed complex ($RP_c$), open complex ($RP_o$), CarD-bound closed complex ($RP_cCarD$) and CarD-bound open complex ($RP_oCarD$). In principle, CarD may interact both with DNA alone and unbound RNAP. However, CarD has been found by ChIP-seq only in genomic positions where RNAP holoenzyme is also found (13). In addition, the interaction between CarD and free RNAP is likely much weaker than
Figure 6. A CarD-dependent kinetic model of open-complex formation. (A) The model consists of five states: unbound DNA (R + P), closed complex (RPc), open complex (RPo), closed complex bound to CarD (RPcCarD) and open complex bound to CarD (RPsCarD). Values indicate the ratio of rate constants between each pair of states in the model. For example, at 250 nM, the forward and reverse rates between R + P and RPc are equal. Furthermore, the ratio of forward to reverse rates between RPc and RPo are 0.125 and 40 in the absence and presence of CarD, respectively. The rates used to generate the figures below can be found in Supplementary Information. (B) Simulations of the kinetic model titrating CarD concentration with 150 nM (black) and 225 nM (red) MboRNAP. (C) Simulations of the kinetic model titrating RNAP concentration with 0 nM (black) and 1 μM (red) CarD.

The association of CarD to DNA-bound RNAP complexes where it may interact with both polymerase and DNA cooperatively. For these reasons, CarD-DNA and CarD-RNAP states are left out of the current model.

Initially models where CarD only associated with open complex were attempted, but these models are unable to capture the biphasic dependence of $k_{obs}$ on CarD concentration. However, a model where CarD also interacts with closed complex is able to account for the concentration dependencies of both the polymerase and CarD with a single set of rate constants (Figure 6, Supplemental Methods). The model has three key features: (i) CarD has a lower affinity to closed complex than it does to open complex; (ii) in the open complex, CarD inhibits the rate of bubble collapse; and (iii) the rates between CarD in the closed complex, CarD accelerates the rate of DNA opening. Importantly, the model was required to satisfy detailed balance so that the energetics linking the affinities of CarD to the equilibria between open and closed complexes do not violate thermodynamics (Supplemental Methods).

The model was used to simulate time-dependent traces of open-complex formation and the resulting curves were fit to extract an observed rate of the approach to equilibrium ($k_{obs}$). The model captures the dependence of the observed rate on CarD concentration at different fixed MboRNAP concentrations (Figures 5B and 6B). Specifically, the observed rate decreases initially at both polymerase concentrations and then increases more at 225 nM RNAP than it does 150 nM RNAP. The model also captures the dependence of the observed rate on RNAP concentration in the presence and absence of 1 μM CarD (Figures 5C and 6C).

CarD mutants are deficient in open-complex stabilization

Single amino acid mutants of CarD have previously been used in vivo and in vitro to understand the roles played by different domains or regions of the protein (12,14). Three groups of mutants have been identified according to their distinct phenotypic effects in vivo, namely those with DNA-binding defects, RNAP interaction defects and with an alanine substituted for a conserved tryptophan residue (W85A). Although W85 is part of the DNA-binding domain, it appears to play a more specific role in the stabi-
lization of open complex, perhaps by interacting with the upstream end of the DNA bubble (13, 14).

CarD mutants that weaken the interaction with RNAP (R25E, (12)), weaken the interaction with DNA (K90A, (14)) and W85A (13, 14) were used as representative mutants from the three classes and were tested in the open-complex formation assay. Fluorescence enhancement traces using MboRNAP in the presence of 1 μM CarD (Figure 7A) show that all three mutants are partially deficient in stabilizing open complex relative to WT, as evidenced by their lower equilibrium (final) fluorescence enhancements. Increasing the concentration of mutant CarDs showed more enhancement, but the final enhancements at the highest protein levels tested were still less than 50% that of wild type (WT) (Figure 7B, Supplementary Figure S6). Fits of the mutant CarD titrations result in estimates of open-complex stabilization and concentrations of half-maximal effect. K90A and W85A both reach a maximum of 3-fold fluorescence enhancement as opposed to the 6.2-fold enhancement seen with WT CarD. The half-maximal concentrations of K90A and W85A are 400 ± 48 nM and 381 ± 118 nM respectively as compared to the 77 ± 35 nM observed with WT CarD. The data for R25E were unable to be fit as we were unable to reach saturation due to limitations of protein concentration, but the activity of the RNAP-interacting domain mutant is even further decreased relative to the other mutants. Taken together, the data demonstrate that all three residues, and by extension all three activities (polymerase binding, DNA binding and W85 activity), are required for full CarD activity.

Fits of the time-dependent fluorescence enhancements again showed three phases and the slowest rate (k_{3obs}) in the presence of CarD mutants depended on CarD concentration as with WT CarD (Figure 7C, Supplementary Figure S6). However, the degree to which k_{3obs} was decreased was less in all the mutants compared to WT. Based on the model that the decrease in k_{3obs} is due to CarD’s ability to inhibit DNA closing, this suggests that the mutants are unable to prevent bubble collapse to the same extent as WT. More specifically, the initial decrease of k_{3obs} observed as a function of CarD concentration showed a similar concentration dependence as WT for the W85A and K90A mutants, but the observed rate decreased less (Figure 7C). In contrast, R25E exhibited a much slower decrease suggesting a lower affinity to open complex (Figure 7C) consistent with previous work showing that this mutant associates less strongly with MboRNAP (12). At higher concentrations, W85A and K90A show signs of an increasing k_{3obs} suggesting that they are able to interact with closed complex to accelerate DNA opening, but not to the same extent of WT. Even though the observed rate does not decrease as much as with WT CarD at low concentrations, it is not as fast at higher CarD concentrations either (i.e. at 1 – 2 μM CarD). In contrast, R25E shows no signs of the increasing phase of k_{3obs}, suggesting, perhaps, that is unable to effectively interact with closed complex at these concentrations. These observations coupled with the reduced fluorescence enhancement observed at saturating levels of mutant CarD, suggest that W85A and K90A mutants are able to associate with both open and closed complexes as in WT, but are deficient in both inhibiting the rate of bubble collapse and increasing rate of opening. In contrast, it appears that the association of R25E with the complexes occurs with a much lower affinity and that it may only be able to interact weakly with open complex.

Figure 7. Mutants of CarD diminish the degree of open-complex stabilization. (A) Fluorescence enhancement as a function of time are shown for 225 nM RNAP + 1 μM WT and mutant CarDs at 25°C. W85A (green) and K90A (red) traces generate less than half the open complex at equilibrium as compared to WT (blue) and R25E (purple) is even further compromised. (B) Equilibrium fluorescence enhancement as a function of CarD concentration taken from titrations of mutant CarDs (Supplementary Figure S6). (C) The CarD concentration dependence of k_{3obs} for WT (blue), W85A (green), K90A (red) and R25E (purple) mutant proteins.

61
DISCUSSION

A better understanding of the role of *M. tuberculosis* CarD in transcription initiation will not only be relevant to mycobacteria, but also to the diverse bacterial species that encode CarD homologs (9,13). The result that *Mbo*RNAP is less able to form *rrnAP3* open complexes as compared to *Eco*RNAP rationalizes the essentiality of CarD in *M. tuberculosis* (and its absence from *E. coli*) and suggests that polymerases from organisms in which CarD is found may also be deficient in open-complex formation. This observation also stresses the importance of directly studying transcription systems in a wide spectrum of bacteria, especially in pathogen like *M. tuberculosis*, that may diverge from traditional model systems.

Previous studies of the effect of CarD on transcription initiation focused on open complex lifetimes in the presence of competitor DNA (13–14,25). Specifically, proteins and DNA were allowed to come to equilibrium and open-complex stability was assayed by the time-dependent decay of reporters of open complex (i.e. the production of transcripts upon the addition of nucleoside triphosphate (NTPs)). These valuable assays have led to the model where CarD stabilizes open complex. However, in the cell, there is no initial equilibration time and transcriptional regulation must be enacted in real time. That is to say, the flux of transcript production must be attenuated or augmented. Therefore, both the rate at which open complexes are formed and the rate at which they decay are crucial for gene regulation in vivo. With this mindset, the rates at which DNA opens and closes in DNA-bound holoenzyme complexes and their dependence on CarD likely have important regulatory consequences.

In this study, we adapt a real-time, fluorescence-based assay to gain insights on the mechanisms used by *M. tuberculosis* CarD to stabilize mycobacterial open complexes. We show that CarD stabilizes RPo, for *Mbo*RNAP on *rrnAP3* in a concentration-dependent manner that saturates at 1 μM and approaches the level of stability displayed by *Eco*RNAP on the same promoter. The fact that CarD increases the concentration of open complex on RNAP-saturated DNA templates demonstrates that CarD specifically affects the equilibrium between promoter-bound complexes and stabilizes RPo with respect to RPc. This result, in combination with a thermodynamic cycle coupling CarD binding and the equilibrium between RPo and RPc, shows that CarD has a higher affinity to open complex as compared to closed complex. The temperature dependence of CarD activity also matches well with the known temperature dependence of open complex (Figure 4C) (19–21). Furthermore, the kinetics of the approach to equilibrium display interesting trends as a function of both RNAP and CarD concentrations. Specifically, the biphasic behavior of the slowest observed rate (*k^1_{obs}*), as a function of CarD concentration at high RNAP concentrations suggests that CarD uses more than one mechanism to stabilize open complex (Figure 5B).

Importantly, *k^1_{obs}* also depends on RNAP concentration at high concentrations of CarD (Figure 5C). Since the concentration of the polymerase affects the rate at which it encounters and binds the template, this suggests that the effect of CarD on the kinetics of open-complex formation depends on the rate of association of RNAP. One possibility is that CarD interacts with free polymerase and accelerates its association with DNA. However, while we do not exclude this interaction as part of CarD’s overall mechanism, it is unlikely to play a major role as *k^1_{obs}* shows no RNAP concentration dependence in the absence of CarD (Figure 5C, black) suggesting that only increasing the rate of association of holoenzyme has no effect. Instead, to account for the above observations, we hypothesize that CarD is able to associate with closed complex (Figure 6A). The addition of this state to the mechanism has the potential to account for all the data as the formation of a CarD-bound closed complex will naturally depend on the concentrations of both RNAP and CarD.

The full model postulates that at low concentrations, CarD associates with pre-formed open complex and stabilizes it, thus inhibiting its isomerization to closed complex. At higher concentrations, CarD associates with closed complex as well and destabilizes it, thus promoting its isomerization to open complex (Figures 6 and 8). The model satisfies detailed balance and is also consistent with the observation that CarD must have a higher affinity to RPc as compared to RPo. We caution that we do not interpret the presented model parameters as measurements of affinities or equilibria. These parameters will need to be determined via future experiments, however the ability of the model to capture all the trends we observe with a single parameter set suggests that the model topology is correct and is able to capture the critical interactions between *Mbo*RNAP, CarD and DNA. At saturating concentrations of both RNAP and CarD, the model predicts that CarD interacts with both open and closed complexes and tilts the equilibrium toward open complexes. As both proteins are present in high concentrations in the cell (> 1 μM, Supplemental Methods, Supplementary Figure S4 and (24)), both activities of CarD are apt to play important roles in transcriptional regulation in vivo.

Recently, DNA footprinting and transcription assays were used to measure the effect of CarD on the stability of the open complex and transcript production (25). This work nicely shows that although mycobacterial RNAP makes the same contacts with the promoter as *E. coli* RNAP, it forms an unstable open complex on the *rrnAP3* promoter. Furthermore, the addition of CarD at high concentrations led to the stabilization of the open complex and stimulated single rounds of transcript production. Our results are consistent with these data with regard to CarD’s ability to inhibit bubble collapse in the open complex. Additionally, the concentration-dependent kinetic analysis of the formation of open complex significantly expands our understanding of the mechanism of CarD and shows that it also functions by accelerating bubble formation.

The two-step binding and opening model is able to capture all of the trends in observed rates presented in this work. The model has served as a useful tool for analyz-
ing open-complex formation of *E. coli* RNAP on a range of promoters (18, 26) and to analyze open-complex formation for mycobacterial RNAP in the absence of transcription factors (27). Given that the mechanism of initiation in *E. coli* is known to be significantly more complex (1), we expect that the two-step model will serve as a starting point for more detailed studies of CarD’s kinetic mechanism of regulating transcription initiation in mycobacteria. Promoter-specific initiation in *E. coli* involves transitions through multiple closed and open complexes of varying stabilities which are defined by conformational changes in both the DNA as well as RNAP (1). One subtle possibility with respect to the interpretation of the data presented is that early open intermediates in the *E. coli* mechanism (i.e. I2) also exist in the *M. bovis* mechanism and do not generate fluorescence enhancement. In this case, the data could be explained via the CarD-dependent acceleration of the formation of RPc via its association with these intermediates instead of with RPc. However, as transitions between intermediates prior to melting of the transcription bubble are thought to be rate limiting, transcription factors which act on these intermediates (i.e. CRP (28) and DksA (29)) can have more dramatic effects on open-complex kinetics. Thus, while we favor the model presented where CarD interacts directly with RPc, we exclude neither the possibility that multiple closed and open intermediate states exist in the mechanism of mycobacterial open-complex formation, nor the possibility that CarD may act on more than one of these states.

To dissect the roles played by different domains of CarD within the context of the proposed kinetic mechanism, real-time fluorescence enhancements were acquired using point-mutants of CarD. Mutants of CarD that weaken its interactions with DNA or the polymerase, or lack the critical tryptophan W85 all show defects in open-complex stabilization. The equilibrium concentration of open complex and the relative changes of the slowest observed rate with the DNA-binding (K90A) and W85A mutants are consistent with a model in which W85 performs both of these roles. While it is not clear what interactions would lead to a different affinity between open and closed complex in WT CarD, one possibility is the hypothesized steric clash between W85 and closed complex DNA (13). This clash would result in non-optimal binding between the C-terminal domain of CarD and DNA and could lead to a destabilization of CarD-bound closed complex. This destabilization would naturally lead to an increase in all rates leading out of the CarD-bound closed complex. This would increase the rate of CarD dissociation,
explaining the decreased affinity of CarD to closed complex and would also lead to an increase in the rate of isomerization to open complex resulting in an acceleration of DNA opening.

In conclusion, the kinetic model of CarD-dependent open-complex stabilization proposed here might have important consequences in terms of the gene-specific effects of CarD. While CarD is localized in vivo at every σ70-dependent promoter (13), its role at each promoter is unclear. The ability of CarD to accelerate opening and inhibit bubble collapse will likely lead to complex dependencies on both promoter sequence and the presence of additional transcriptional regulators. More specifically, CarD could either inhibit or promote transcription, depending on the basal kinetics of a given promoter and the basal rates of DNA opening, bubble collapse and promoter escape. Future research along these lines will be required to elucidate the role of CarD across the genome.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

The authors would like to thank Dr Bob Landick for assistance with protocols and reagents related to the isolation of MboRNAP. We also thank Dr Tim Lohman for training on and the use of his stopped-flow spectrophotometer and Tom Ellenberger for training on and the use of the HPLC and FPLC in his laboratory. We also thank Dr Lohman, Dr Roberto Galletto and Dr Tom Ellenberger for their scientific input. Lastly, we thank Dr Tomasz Heyduk for his assistance with setting up the fluorescence assay and for critically reading the manuscript while in preparation.

Author contributions: J.R. performed the experiments. E.A.G. and C.L.S. conceived the research. E.A.G. and J.R. planned the experiments, analyzed the data and interpreted the results. A.R.M. purified MboRNAP components and WT CarD. A.L.G. constructed and purified CarD mutants and measured protein levels in the cell. E.A.G., J.R. and C.L.S. wrote the paper.

FUNDING

National Institutes of Health (NIH) [R01GM107544 to E.A.G. and C.L.S.]; Sigma-Aldrich Predoctoral Fellowship [to J.R.]; NIGMS Cell and Molecular Biology Training [GM07067 to A.L.G.]; Stephen I. Morse Graduate Fellowship [to A.L.G.]; Funding for open access charge: NIH [R01GM107544]. Conflict of interest statement. None declared.

REFERENCES


Supplemental Information

CarD stabilizes mycobacterial open complexes via a two-tiered mechanism.

Jayan Rammohan\textsuperscript{1}, Ana Ruiz Manzano\textsuperscript{1}, Ashley L. Garner\textsuperscript{2}, Christina L. Stallings\textsuperscript{2}, and Eric A. Galburt\textsuperscript{1}\textsuperscript{*}

\textsuperscript{*} To whom correspondence should be addressed. Tel: (314)362-5201 Email: egalburt@biochem.wustl.edu

\textsuperscript{1} Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine, St. Louis, MO 63110, USA

\textsuperscript{2} Department of Molecular Microbiology, Washington University School of Medicine, St. Louis, MO 63110, USA
SUPPLEMENTAL METHODS

Template sequence

The full non-template strand sequence of the 150 bp \textit{rrnA} P3 DNA template used in the fluorescent assay. The italics/underline indicate genomic sequence and the green, blue, red, and orange type indicate the -35 box, -10 box, start-site, and Cy3 label position respectively.

AAAATAGGCGTATACGAGGCCCTTTCGTCTTCAAAGAATTCTGTCTTAAGACGTCCCATAGCCGG
GTTTGTTTACGGGCAGGGTTGAAGCCTATATGTACCATGCATGATGAACTC
AGAAGTGAAAACGTATATCCGATGGTAGTG

Kinetic simulations

The model as presented in the paper was constructed with five states and seven independent rates. The three effective bimolecular on rates were linearly dependent on the concentrations of the binding partners and were all set to be equal. Furthermore, the dissociation rate of CarD from open complex was calculated to satisfy detailed balance. Specifically, \( k_6 = \frac{(k_3 k_8 k_5)}{(k_4 k_7)} \).

\[
\begin{align*}
R + P & \xrightleftharpoons[k_2]{k_1} RP_c \xrightleftharpoons[k_4]{k_3} RP_o \\
RP_c & \xrightleftharpoons[k_5]{k_1} RP_c \text{CarD} \xrightleftharpoons[k_8]{k_7} RP_o \text{CarD}
\end{align*}
\]

The values of the rates used to generate the plots in the main text are as follows: \( k_1 = 0.3 \text{ s}^{-1} \text{nM}^{-1} \), \( k_2 = 75 \text{ s}^{-1} \), \( k_3 = 0.0005 \text{ s}^{-1} \), \( k_4 = 0.004 \text{ s}^{-1} \), \( k_5 = 400 \text{ s}^{-1} \), \( k_6 = 2.0 \text{ s}^{-1} \), \( k_7 = 0.0125 \text{ s}^{-1} \), \( k_8 = 0.0005 \text{ s}^{-1} \).
Measurements of protein concentration in vivo

Five mL of 4 replicate cultures of *Mycobacterium smegmatis* Mc^2^155 in exponential growth phase were collected and the cells were pelleted and lysed in 500 µl of NP-40 (10 mM sodium phosphate, pH 8.0, 150 mM NaCl, and 1% Nonidet® P-40) by bead beating (FastPrep; MP Bio). Before lysing, the OD(λ=600) was measured for each culture and the conversion of OD (λ=600) of 1 corresponds to 5 x 10^8 cells/ml was used to determine the number of cells in each sample. The total protein concentrations of the lysates were also measured by BCA assay (Pierce). Undiluted, 1:5 dilution, and 1:10 dilution of cell lysates were run on a denaturing polyacrylamide gel adjacent to lanes containing 0.05, 0.1, 0.5, or 1.0 pmol each of *M. smegmatis* CarD, *M. bovis* RNAP core, and *M. bovis* RNAP σ^A^ purified protein. Following SDS PAGE, western blot analyses were performed and CarD, RNAP σ, and RNAP β were detected using mouse monoclonal antibodies specific for CarD (clone 10F05; Memorial Sloan-Kettering Cancer Center Monoclonal Antibody Core Facility), RNAP σ (clone 2G10, Neoclonne, Madison, WI), and RNAP β (clone 8RB13; Neocloone, Madison, WI), respectively. Western blots were imaged and quantified using an Odyssey CLX imaging system (LI-COR). Standard curves for CarD, RNAP σ^A^, and RNAP β were generated and used to determine the pmol/lane for each protein in the lysates. The molecules/cell calculation assumes 100% cell lysis and a cell count given by OD. The conversion of molecules/cell to protein concentration assumes a cell volume of 4 µM^3^.
Figure S1: CarD-dependent fluorescent enhancement depends on MboRNAP. (A) The trace of 1 µM CarD in the absence of RNAP shows no fluorescent enhancement. (B) In the presence of 21 nM EcoRNAP, the addition of 1 µM CarD shows only a slight fold change relative to EcoRNAP alone. (C) The magnitude of the fluorescence fold change when CarD is added to concentrations of EcoRNAP and MboRNAP in the middle of their respective concentration dependences. Specifically, 21 nM EcoRNAP is compared to 225 nM MboRNAP. This comparison shows the effect of CarD on the respectively polymerases under conditions where they alone generate about half open complex. This demonstrates that CarD has a much greater effect on MboRNAP.
Figure S2: Fractional amplitudes of the three observed rates. Each observed rate ($k_{1\text{obs}}$, $k_{2\text{obs}}$, $k_{3\text{obs}}$) obtained from the triple exponential fit carries an amplitude ($a_1$, $a_2$, $a_3$) which describes that phase’s contribution to the total fluorescent enhancement signal. The majority of the increase of the overall opening signal (fluorescent enhancement) as a function of CarD concentration is mostly due to increases in the amplitude of slowest observed rate ($a_3$). This is shown by plotting the fractional amplitude of the rates (i.e. $a_1/A$ where $A$ is the total amplitude of the trace) as a function of CarD concentration. The fraction of the total signal accounted for by the amplitude of the slow phase (red) increases while it decreases for the faster phases (blue and green). This is true for both concentrations of $Mbo$RNAP tested (150 nM and 225 nM).
Figure S3: Observed rates in a three state example. (A) A simpler model is used to demonstrate the effect of microscopic rate constants on observed rate. In this model $k_1$ and $k_{-1}$ describe the association and dissociation of RNAP with DNA respectively. Similarly, $k_2$ and $k_{-2}$ describe the opening and closing rates respectively. (B) The CarD effect is captured simply by running the simulation with different combinations of $k_2$ and $k_{-2}$ to represent different concentrations of CarD. The dependence on CarD concentration for $k_2$ (red) and $k_{-2}$ (green) show the values input into the simulation used to generate time-dependent curves of the equilibration of open complex concentration (not shown). These curves are fit by an exponential to extract a measure of the observed rate (black). The concentration dependence of the measured observed rate (black) closely follows the trend of the sum of the two rates controlling the equilibrium between open and closed complex, $k_2 + k_{-2}$ (blue). This example demonstrates how combinations of microscopic rate constants lead to observed rates and how trends in observed rates can be analyzed to extract mechanistic information.
Figure S4: Concentrations of RNAP and CarD in vivo. (A) Western blot used for quantitation of protein levels found in lysates of M. smegmatis during logarithmic growth. Dilutions of lysate are shown next to lanes loaded with known amounts of each protein. (B) Bar graph shows the mean ± SEM of for each protein in the 4 replicates. The number of molecules/cell and the equivalent concentration are shown on the 2 y-axes. The conversion of the raw data to molecules/cell assumes 100% lysis of the collected cells and the conversion to concentration assumes a cell volume of 4 µm³.
Figure S5: Simulated equilibrium fold enhancements. (A) Using the kinetic model, equilibrium fold changes were simulated and plotted as a function of RNAP concentration in the presence (red) and absence (green) of 1 µM CarD. (B) Experimental equilibrium fold changes for the same conditions from the main text (Figure 4B).
Figure S6: Concentration titrations of mutant CarDs. Fluorescent fold change as a function of time for the (A) W85A, (B) K90A, and (C) R25E mutant CarDs in the presence of 225 nM MboRNAP.
Figure S7: Comparison of mutant CarD-dependent fold-changes. The fluorescent fold changes for 21 nM EcoRNAP or 225 nM MboRNAP with WT or mutant CarD. The largest effect is seen for MboRNAP and WT CarD. Both W85A and K90A show about half the effect and R25E with MboRNAP shows an effect almost as low as EcoRNAP with WT CarD.
Epilogue

The work presented in this chapter was continued by testing point mutants of CarD-I27 (I27W and I27F) with higher affinity for RNA polymerase. This data is part of a larger manuscript in submission titled “Effects of Increasing the Affinity of CarD for RNA Polymerase on Open-Complex Formation at the Mycobacterium tuberculosis rrnAP3 Promoter” by Ashley Garner et al. The work presented here represents the collective work of Eric Galburt and members of his laboratory (Ana Ruiz-Manzano and Drake Jensen) as well as Christina Stallings and members of her laboratory (Ashley Garner, Jeremy Huynh and Leslie Weiss). Importantly, the results with these point mutants are also consistent with the 5-state model of CarD stabilization of RPo presented in this chapter, in the sense that open complex stabilization and acceleration both occur at lower concentrations of CarD. Interestingly, as CarD-I27 approaches saturating concentrations, the equilibrium approaches approximately the same fold change as wild-type CarD, suggesting that the higher affinity CarD does not lead to more open complex at saturation.

Results and Conclusions

CarDI27F and CarDI27W mutants stabilize open complexes at lower concentrations than CarDWT.

CarD regulates transcription initiation by binding to and stabilizing RNAP-promoter open complexes (1-5). We have previously developed a fluorescence assay for CarD activity that reports on open complex formation in realtime. In this assay, a Cy3 label, which exhibits a 2-fold fluorescence enhancement in open complex, is incorporated at the +2 position on the non-template strand of the Mtb rrnAP3 promoter within a linear fragment.
of Mtb genomic DNA containing nucleotides 1,470,151 to 1,470,300 (8). Mbo RNAP-σA holoenzyme with or without Mtb CarD is then mixed with the labeled DNA fragment via stopped-flow spectrophotometry and fluorescence is monitored for 20 minutes. The Cy3 label increases fluorescence in the open complex conformation and thus the rate of open complex formation can be monitored as a change in fluorescence. The amplitude of the fluorescence intensity curve correlates to the equilibrium amount of open complex in a given condition (4). To monitor the effect of increasing the affinity of CarD for RNAP on the formation and stability of open complexes, we performed this assay with concentrations of CarD I27F and CarD I27W ranging for 0 to 1800 nM and fixed concentrations of RNAP (225 nM) and promoter DNA (10 nM) (Fig. 2A). We found that the CarD concentration necessary to reach half of the maximum level of open complex at saturation (half-maximal C3 concentration) was lower for CarD I27F (17 ± 2 nM) and CarD I27W (23 ± 3 nM) than for CarD WT (59 ± 10 nM), consistent with a higher affinity of the CarD I27 mutants for RNAP in initiation complexes compared to CarD WT (Fig. 2A).

CarD I27F and CarD I27W mutants accelerate promoter opening at lower concentrations than CarD WT.

We have previously reported modeling based on the trends of the slowest observed rate in the real-time fluorescent traces (kobs C3) that suggests that CarD stabilizes open complex through a two-tiered concentration-dependent mechanism where CarD associates with both open and closed complexes with different affinities (4). More specifically, the model predicts that at low concentrations (i.e. < 100 nM), CarD WT binds to open complex and
prevents bubble collapse, resulting in more open complex and a slower observed rate. The model further predicts that at higher concentrations, CarD WT binds to the closed complex and accelerates the rate of opening, resulting in still more open complex and an acceleration in the observed rate. Analysis of kobs 3 as a function of CarD concentration was performed for the CarD I27 mutants as previously described (4) (Fig. 2B). We found that the I27 mutants begin to accelerate kobs 3 at lower concentrations (< 50 nM) than WT CarD (100 nM). In fact, for CarDI27W we did not observe a deceleration in kobs 3 even at the lowest concentration tested (16 nM). The two-tiered kinetic model predicts that acceleration in kobs 3 arises from CarD binding to closed complex, which increases the rate of promoter opening. Thus, our working model predicts that, even at 16 nM, CarDI27W is associating with closed complex and accelerating opening. The acceleration of the observed rate at lower concentrations coupled with the lower half-maximal concentration for open complex stabilization is consistent with a model where the CarD I27 mutants have higher affinities to both closed and open RNAP-promoter complexes as compared to CarD WT.
Figure 1. Effects of increasing the affinity of CarD for RNAP on open complex stability. (A) Equilibrium fluorescence fold-change normalized to the fold-change at high CarD concentrations (saturation) for each mutant of CarD with 225 nM Mbo RNAP-σA and 10 nM Cy3-labeled Mtb rrnAP3 promoter. Both Mtb CarDI27F (red triangles) and CarDI27W (blue squares) mutants achieve a half-maximal effect (dashed line) at lower concentrations (17 ± 2 nM for CarDI27F, and 23 ± 3 nM for CarDI27W) than Mtb CarDWT (black circles, 59 ± 10 nM). (B) kobs,3, calculated using ProData Viewer software from Applied Photophysics, of open complex formation as a function of CarD concentration for Mtb CarDWT (black circles), CarDR25E, CarDI27F, and CarDI27W (blue squares). The legend for B is shared with A. (C) In vitro transcription assay showing a representative gel and a graph of the ratio of the amount of 3nt initial transcript formed by 200 nM Mbo RNAP-σA from 10 nM of a linear DNA fragment in the presence of 2 μM CarD versus in the absence of CarD for reactions containing no CarD, Mtb CarDWT, CarDR25E, CarDI27F, or CarDI27W. The graph shows the mean ± SEM of data from at least four replicates. Statistical significance was analyzed by ANOVA and Tukey’s multiple comparison test. *, p ≤ 0.05; **, p ≤ 0.01; ***, p ≤ 0.001; or ****, p ≤ 0.0001.
References


Chapter 4

Cooperative stabilization of *Mycobacterium tuberculosis* rrnAP3 promoter open complexes by RbpA and CarD
**Introduction**

Like CarD, the transcription factor RbpA is known to activate transcription from the *Mtb* *rrnAP3* promoter, however the effect of RbpA on open complex kinetics is unknown. Structural modeling predicts that both transcription factors interact with the transcription initiation complex at different sites (Figure 1 below), but no studies have been performed to test this hypothesis directly. I studied the effect of RbpA on the kinetics of *rrnAP3* promoter-opening in the presence and absence of CarD. I found that RbpA stabilizes open complexes using a distinct kinetic mechanism from that of CarD. Furthermore, the two factors cooperatively stabilize RPo and this cooperativity can be captured by a simple model of positive linkage. The work presented here represents a step towards understanding how multiple essential transcription factors can work in concert for regulating transcription initiation in *Mtb*.

![Figure 1. Structural model of CarD and RbpA simultaneously bound to the transcription open complex. (PyMol image courtesy of Ashley Garner).](image-url)
Cooperative stabilization of *Mycobacterium tuberculosis* rrnAP3 promoter open complexes by RbpA and CarD

Jayan Rammohan1, Ana Ruiz Manzano1, Ashley L. Garner2, Jerome Prusa2, Christina L. Stallings2 and Eric A. Galburt1,*

1Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine, St. Louis, MO 63110, USA and 2Department of Molecular Microbiology, Washington University School of Medicine, St. Louis, MO 63110, USA

Received April 27, 2016; Revised June 14, 2016; Accepted June 16, 2016

**ABSTRACT**

The essential mycobacterial transcriptional regulators RbpA and CarD act to modulate transcription by associating to the initiation complex and increasing the flux of transcript production. Each of these factors interacts directly with the promoter DNA template and with RNA polymerase (RNAP) holoenzyme. We recently reported on the energetics of CarD-mediated open complex stabilization on the *Mycobacterium tuberculosis* rrnAP3 ribosomal promoter using a stopped-flow fluorescence assay. Here, we apply this approach to RbpA and show that RbpA stabilizes RNAP-promoter open complexes (RPo) via a distinct mechanism from that of CarD. Furthermore, concentration-dependent stopped-flow experiments with both factors reveal positive linkage (cooperativity) between RbpA and CarD with regard to their ability to stabilize RPo. The observation of positive linkage between RbpA and CarD demonstrates that the two factors can act on the same transcription initiation complex simultaneously. Lastly, with both factors present, the kinetics of open complex formation is significantly faster than in the presence of either factor alone and approaches that of *E. coli* RNAP on the same promoter. This work provides a quantitative framework for the molecular mechanisms of these two essential transcription factors and the critical roles they play in the biology and pathology of mycobacteria.

**INTRODUCTION**

The regulation of gene expression serves as a gateway between genotype and phenotype. By modulating the output of specific genes, cells tune their molecular makeup to best suit environmental conditions. Much of this regulation is enacted by modulating the rates of transcription initiation to control the flux of RNA production. In bacteria, the basal transcriptional machinery is composed of RNA polymerase holoenzyme which consists of a catalytic core enzyme (ββ′α2σ) and a dissociable sigma factor (σ) that directs promoter recognition. Transcription initiation proceeds via the binding of RNAP holoenzyme to promoter DNA to form closed complex (RPc) followed by the spontaneous unwinding of approximately a turn of DNA to form open complex (RPo). In RPo, the single-stranded DNA template is correctly positioned in the polymerase active site, incoming ribonucleotides may bind and RNA polymerization may ensue. The polymerization of the initial ribonucleotides leads to promoter escape, the formation of a stable elongation complex and the production of transcript. In *Escherichia coli*, transcription initiation mechanisms have been well-studied and multiple kinetic intermediates have been identified between the initial DNA-holoenzyme closed complex and open complex (1). These details are of crucial importance in understanding the structural transitions that the complexes go through during the isomerization to open complex. However, a minimal kinetic scheme that describes promoter binding and opening in two reversible steps (R+P ↔ RPc ↔ RPo) has proven to be a useful and practical starting point when investigating mechanisms of open-complex formation and regulation (2).

The regulation of transcription initiation involves factor-dependent tuning of the stabilities of transcription intermediates and the rates of interconversion between these states. Examples include modulation of polymerase-promoter affinity, changing the equilibrium between RPc and RPo and influencing the rate of NTP-dependent promoter escape (3–5). It has recently become apparent that initiation in *Mycobacterium tuberculosis* (*Mtb*) is controlled via mechanisms that are distinct from those found in *E. coli*. For example, mycobacteria lack Fis (6), DksA (7) and AT-rich up-
stream activating elements (8,9). Furthermore, mycobacteria possess unique transcription factors not found in E. coli including CarD (10–12) and RbpA (13–15).

CarD is an RNAP- and DNA-binding protein that is essential for growth in M. tuberculosis, other mycobacteria (10–12). Furthermore, CarD is required for the response of Mtb to oxidative stress, some antibiotics and infection of mice (10). In vitro, CarD binds to initiation complexes and stimulates the production of RNA by stabilizing the relatively unstable RPo generated by mycobacterial RNAP holoenzyme on ribosomal promoters (11,12,16). In initiation complexes, it makes direct molecular interactions with the β1 lobe of the RNAP β subunit through its N-terminal RNAP-interaction domain (RID) (10,17,18) (Figure 1A and B). CarD also binds DNA non-specifically through its C-terminal domain (CTD) (19, Srivastava: 2013ga;20), which contains a tryptophan residue that interacts with the upstream edge of the transcription bubble to stabilize RPo (18). Mutations in the RID, CTD or tryptophan lead to distinct in vivo and in vitro effects suggesting that full CarD activity requires each of these three functional modules (12,17,20). Based on these studies and previous analyses by our group, our working model for how CarD stabilizes open complex and stimulates transcription consists of a two-tiered, concentration-dependent mechanism. The model predicts that at low concentrations (i.e. <100 nM), CarD binds RPo and slows the rate of bubble collapse by conformational selection, while at higher concentrations, CarD binds RPs and accelerates the rate of DNA unwinding by induced fit (12). We expect both of these kinetic effects to play roles in vivo where the concentration of CarD is well in excess of 100 nM (12).

RbpA is also an essential RNAP- and DNA-binding protein found in M. tuberculosis and other Actinobacteria, but not in E. coli (13,15). RbpA consists of an unstructured N-terminal tail, a central core β-barrel domain, a 15 amino acid basic linker (BL) and a C-terminal sigma-interacting domain (SID) (15,21) (Figure 1A). The SID binds to the second domain of σ (σ2, domains 1,2, 2.3 and the non-conserved region) in both the presence and absence of the core RNAP enzyme and has specificity to the housekeeping sigma factor σ^A and the stress-response sigma factor σ^S (15,21,22). A crystal structure of a BL/SID construct bound to σ^A has been solved (15) and can be used to position RbpA in the initiation complex (Figure 1B). RbpA’s sigma specificity has led to proposals that RbpA plays a role in the competition of sigma factors for RNAP core (14,23). RbpA potentially stabilizes RNAP holoenzyme by binding both sigma and either β2 or another region of the core subunits (14,15). Furthermore, RbpA can increase the affinity of holoenzyme to the promoter, presumably via the interaction of the BL with DNA (15,24). An arginine residue (R79) in the BL is thought to play a role in RbpA’s ability to bind DNA, although it is unknown whether this interaction may contribute to RbpA’s promoter specificity (15). RbpA stimulates transcription in vitro and the BL/SID region of the protein are sufficient for partial stimulation (15). However, the mechanism of transcriptional stimulation by RbpA remains unclear.

Here, we perform a mechanistic analysis of the function of Mtb RbpA during open complex formation on the Mtb rnrAP3 promoter using a real-time fluorescence assay (25) and show that RbpA stabilizes open complex. Furthermore, RbpA accelerates the approach to equilibrium providing insight into the kinetic mechanism at play. Interestingly, this trend is qualitatively distinct from that of CarD, suggesting that the two factors function in fundamentally different ways.

Furthermore, stopped-flow fluorescence experiments performed in the presence of both factors reveal positive linkage between the effect of CarD and RbpA on RPo stability at the Mtb rnrAP3 promoter. The observation that the two factors bind cooperatively demonstrates for the first time that RbpA and CarD can act on the same initiation complex simultaneously. This is consistent with the prediction that CarD and RbpA have distinct binding sites on initiation complexes (15,18). We present a thermodynamic analysis to quantitatively describe the positive linkage between the two factors. Lastly, we observe a dramatic acceleration in observed rates when both factors are present, and the approach to equilibrium under these conditions resembles that of E. coli RNAP acting on the same promoter.

The date and analysis presented here reveal important details regarding the mechanistic differences between CarD and RbpA and provide a kinetic framework for the function of these two essential mycobacterial transcription factors both independently and cooperatively. This work brings us a step closer to understanding the functional logic of transcription regulation in a pathogenic bacteria that represents a significant burden to human health worldwide.

**MATERIALS AND METHODS**

**Protein purification**

*Mbo* RNAP, *Mbo* σ^A and *Mtb* CarD were prepared as previously described (12). *Mtb* RbpA was cloned into a pET-SUMO-His6 vector and introduced into E. coli BL21(DE3) cells. After growth at 37°C to an OD_{600} of 0.8, protein overexpression was induced by the addition of 1 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG). The tagged protein was purified by nickel affinity chromatography (HP HiTrap, GE Healthcare) and the SUMO tag was cleaved overnight with His-tagged Ulp1 protease. Pure untagged RbpA was collected in the flow through of a second nickel affinity column. The protein was stored in 20 mM Tris (pH 8.0), 150 mM NaCl and 1 mM beta-mercaptoethanol at −80°C.

**Promoter DNA**

A total of 150 base-pair *Mtb* rnrAP3 promoter fragments with a Cy3 label on the +2 non-template dT were prepared as previously described (12), with one notable exception: residues flanking the promoter sequence were replaced with those native to the *Mtb* genome (H37Rv coordinates 1471 S77–1 471 727, Supplemental Sequence) (26,27). Control experiments indicated that this change had no effect on previously published results describing the effect of CarD on open-complex stabilization on the same promoter with random flanking sequence (12).
**Figure 1.** Mycobacterial transcription factors and a fluorescence assay for following open complex formation. (A) CarD is made up of an N-terminal RNAP-interaction domain (RID, light blue) and a C-terminal DNA-binding domain (dark blue). RbpA consists of a core domain (white), a basic linker domain (BL) and a sigma-interacting domain (SID, red). (B) Ribbon representations of both factors are shown in the context of an open complex structure. Both domains of CarD and the SID of RbpA are modeled together based on the structure of CarD bound to open complex (4XLS (18)) and the structure of the RbpA SID domain bound to domain 2 of σ^H (4X8K (15)). (C) Fluorescence assay for following the formation of open complexes. A Cy3 fluorophore attached to the +2 non-template dT undergoes fluorescence enhancement upon open-complex formation.

**Stopped-flow fluorescence assay**

The stopped-flow assay was performed as previously described (12). In brief, Mbo RNAP core was mixed with 3- to 6-fold molar excess Mbo σ^H at 25 °C to form holoenzyme. Holoenzyme was mixed with transcription factor(s) or an equivalent volume of transcription factor storage buffer such that the proteins were initially at twice the desired final reaction concentrations. Promoter DNA was also prepared at twice the desired final reaction concentration. Equal volume mixing was performed in a stopped-flow apparatus (Applied Photophysics SX-20, total shot volume 150 μl, dead time < 2 ms), so the initial protein and DNA solutions were each diluted by half in order to reach their final reaction concentrations. All experiments were performed using a final MboRNAP holoenzyme concentration of 225 nM and a DNA concentration of 10 nM unless otherwise noted. Excitation light was provided by a 510 nm LED light source, as opposed to a 515 nm light source from an arc lamp passed through a monochromator. Emission was collected at 570+ nm using a long-pass filter. Control experiments indicated that this subtle change in excitation wavelength had no effect on previously published results describing the effect of CarD on mycobacterial open-complex stabilization. All experiments were performed at 25 °C in the following final solution conditions: 14 mM Tris pH 8.0, 120 mM NaCl, 10 mM MgCl_2, 1 mM DTT, 0.1 mg/ml BSA and 10% glycerol by volume.

Two to three traces were collected per condition. Traces for each condition were averaged and plotted as fold-change over DNA by first subtracting the buffer signal from all traces and then plotting the data as (F−F_0)/F_0, where F_0 is the signal for DNA alone and F is the signal for DNA mixed with protein. The fold-change traces were fit to a triple exponential from 0.1–1200 s using the ProData Viewer software from Applied Photophysics:

\[ F(t) = \sum_{i=1}^{3} A_{obs,i} \cdot e^{-k_{obs,i}t} \]  

where \( A_{obs,i} \) and \( k_{obs,i} \) are the amplitude and observed rate of the \( i \)th kinetic phase. To facilitate consistency in assignment of fast, intermediate and slow phases, traces were anchored to the intermediate and slow phases using either a single or double exponential before fitting the fast phase of the trace. Fractional amplitudes were calculated according to:

\[ A_{frc,i} = \frac{A_{obs,i}}{\sum_{i=1}^{N} A_{obs,i}} \]  

In order to estimate an overall rate for the approach to equilibrium, an amplitude-averaged rate was calculated using the intermediate and slow phases according to:

\[ k_{obs} = \frac{(A_{obs,2} \cdot k_{obs,2}) + (A_{obs,3} \cdot k_{obs,3})}{A_{obs,2} + A_{obs,3}} \]  

Conditions that were repeated multiple times on different days were used to estimate standard error of the mean (SEM). An average SEM was used to estimate uncertainty for specific conditions that were only repeated multiple times on the same day to better estimate the actual error for these points.
In vitro aborted transcription assay

CarD and RbpA used in this assay were diluted into 1 × dialysis buffer (20 mM Tris pH 8.0, 150 mM NaCl and 1 mM BME). Mbo σ^A was mixed in 8-fold molar excess with core RNAP to reconstitute the RNAP holoenzyme. A total of 85 bp overlapping primers (IDT) were annealed and extended to prepare a linear fragment of dsDNA Mtb Erdman strain genomic DNA containing nucleotides 1 470 and extended to prepare a linear fragment of dsDNA A total of 85 bp overlapping primers (IDT) were annealed with core RNAP to reconstitute the RNAP holoenzyme.

RESULTS

RbpA stabilizes rrnAP3 open complexes

To study the effect of RbpA on transcription open complexes (RPo) we used a fluorescence assay that reports on RPo formation in real-time as previously reported (12,25). The final reaction conditions were: 225 nM Mbo core and σ^A were incubated for 10 min at room temperature. CarD, RbpA and/or dialysis buffer were added to the polymerase and the proteins were incubated for 10 more minutes at room temperature. The DNA template was added and the reactions were diluted to 17.5 μL followed by an additional 10 min at room temperature. Reactions were initiated with addition of a 2.5 μl mixture containing GpU, UTP and the radiolabeled UTP. After 20 min incubation at room temperature, the reactions were stopped with 2x formamide buffer [98% (vol/vol) formamide, 5 mM EDTA] and run on a 22% urea PAGE gel.

Figure 2. RbpA and CarD stabilize open complex. (A) Fluorescent fold-change is plotted over time from mixing Mbo/RNAP and Mbo factors with a +2 T Cy3-labeled Mtb rrnAP3 promoter template. DNA alone mixed with buffer (black), a titration of RbpA from 0–2 μM (red) and 1 μM CarD (blue) are shown. (B) The equilibrium fluorescence fold change over 225 nM RNAF alone is plotted versus the concentration of either RbpA (red) or CarD (blue). The data are fit with the binding model (factor/[factor]+K_{eff}) to extract amplitudes (A_{CarD} = 6.1 ± 0.1 and A_{RbpA} = 2.9 ± 0.1) and effective binding constants (K_{eff,RbpA} = 177 ± 23 nM and K_{eff,CarD} = 73 ± 4 nM). Two to three shots were collected for each condition and error bars represent standard error of the mean.

Figure 2. RbpA and CarD stabilize open complex. (A) Fluorescent fold-change is plotted over time from mixing Mbo/RNAP and Mbo factors with a +2 T Cy3-labeled Mtb rrnAP3 promoter template. DNA alone mixed with buffer (black), a titration of RbpA from 0–2 μM (red) and 1 μM CarD (blue) are shown. (B) The equilibrium fluorescence fold change over 225 nM RNAF alone is plotted versus the concentration of either RbpA (red) or CarD (blue). The data are fit with the binding model (factor/[factor]+K_{eff}) to extract amplitudes (A_{CarD} = 6.1 ± 0.1 and A_{RbpA} = 2.9 ± 0.1) and effective binding constants (K_{eff,RbpA} = 177 ± 23 nM and K_{eff,CarD} = 73 ± 4 nM). Two to three shots were collected for each condition and error bars represent standard error of the mean.

In incubating MboRNAP with increasing concentrations of wild-type RbpA results in increasing equilibrium fluorescence, demonstrating that RbpA stabilizes RPo at the rrnAP3 promoter (Figure 2A). The R79A mutant of RbpA, which is known to play an important role in transcriptional activation (15), results in approximately half the increase in equilibrium fluorescence when added at a concentration at which WT RbpA saturates (2 μM, Supplementary Figure S1). A fit of the equilibrium fluorescence fold-change generated by WT RbpA normalized to RNAF alone to the binding isotherm 1 + A*[RbpA]/([RbpA]+K_{eff}) gives a value of the concentration of half-maximal effect, K_{eff} = 177 ± 23 nM. This K_{eff} is approximately 2-fold weaker than that of CarD (73 ± 4 nM). At saturation, RbpA produces ~3 times (A = 2.9) more fluorescence enhancement than MboRNAP alone, compared to CarD which saturates at ~6-fold (A =
over RNAP (Figure 2B). Assuming equilibrium fluorescence enhancement correlates with the extent of RPo stabilization, the lower equilibrium fluorescence generated by RbpA at saturation suggests that RbpA stabilizes RPo to a lesser extent than CarD. However, given that Cy3 fluorescence enhancement is sensitive to local changes in the dye’s microenvironment, we do not exclude the possibility that the assay reports on multiple open complexes, each with subtle conformational differences leading to different fluorescence enhancements. With this in mind, it is possible that the RbpA-stabilized RPo produces lower fluorescence enhancement than the RPo stabilized by CarD. If this were the case, the extent of fluorescence enhancement would not correspond to the extent of open complex stabilization, but instead would report on the average fluorescence enhancement of an ensemble of open complexes. Clearly, the difference in fluorescence enhancements contains information regarding the mechanisms of open complex stabilization by the two factors. Here, we can only speculate on the molecular details that lead to the observed enhancements, however, these details do not impact the conclusions drawn from the analyses that follow.

**RbpA stabilizes open complexes through a different kinetic mechanism than CarD**

Although RbpA and CarD are unique proteins that bind the transcription initiation complex at distinct sites, the two factors share several important structure-function properties. Specifically, both factors bind RNAP-holoenzyme and DNA, and both factors stabilize RPo. For these reasons, we considered the possibility that RbpA stabilizes open complexes using a similar kinetic mechanism as CarD (12). Curves in the presence of either RpbA or CarD were well-fit by a triple exponential (see Materials and Methods) and the phases were separable by approximately an order of magnitude. The low fractional amplitude of $k_{obs,1}$ made it challenging to measure, so we cannot exclude the possibility that RbpA or CarD influence the kinetics of the fastest observed phase.

We observed dramatic differences between RbpA and CarD in the intermediate and slow phases. CarD traces are dominated by the slowest phase ($k_{obs,3}$), whereas RbpA traces contain significant contributions from both the intermediate and slowest phases ($k_{obs,2}$ and $k_{obs,3}$, respectively) (Supplementary Figure S2). Analysis of the trends in the observed rates themselves also indicated differences between RbpA and CarD (Figure 3B and C). We observe monotonic and saturable increases in $k_{obs,2}$ and $k_{obs,3}$ with increasing RbpA concentration (Figure 3B). Conversely, we were unable to discern a systematic trend in $k_{obs,2}$ as a function of CarD concentration (Figure 3B). Furthermore, $k_{obs,3}$ displays a non-monotonic trend with increasing CarD concentration (Figure 3C and Supplementary Figure S3). Thus, RbpA accelerates the approach to equilibrium at all concentrations tested while CarD decelerates equilibration at low concentrations and accelerates equilibration at high concentrations.

We considered the mechanistic implications of the distinct kinetic trends observed for RbpA and CarD. CarD's non-monotonic trend in $k_{obs,3}$ can be explained in the text of a two-step reversible model coupled to factor binding (12) (Supplementary Figure S4A). However, the observation of multiple saturable observed rates in the presence of RbpA demonstrates that the simple two-step reversible mechanism of open complex formation cannot account for the observed kinetics. Therefore, our data suggest that the mechanism of RpbA open complex stabilization must involve more states. Since RbpA is known to bind sigA, RNAP core and holoenzyme, one possibility is that RbpA's protein-binding interactions contribute to the observed kinetics (Supplementary Figure S4B). Another possibility is that RbpA interacts with additional intermediates on pathway to RPo (Supplementary Figure S4B). These possibilities are not mutually exclusive, and a detailed anal-
RbpA and CarD cooperatively stabilize open complex

As each factor stabilizes RP₀ and structural modeling suggests that they could both bind RP complexes concurrently (15), we tested the effect of RbpA and CarD together on open-complex stability. Since RbpA generates a lower fluorescence enhancement compared to CarD, we reasoned that if RbpA competed with CarD for binding to the initiation complex, the addition of both factors would lead to an intermediate enhancement. However, in the presence of both factors at saturating concentrations, the equilibrium fluorescence fold change is actually greater than that for saturating CarD alone (Figure 4A).

If RbpA and CarD bind concurrently to the rrrAP3 promoter-RNAP complex, one expects to measure positive linkage or cooperativity between the two factors as both stabilize the same conformation, namely open complex. To test this model, we performed titrations of each factor in the presence of the other factor at saturation and asked whether we observed a shift in $K_{\text{eff}}$ relative to each factor alone. Indeed, titrating RbpA in the presence of 1 μM CarD results in a $K_{\text{eff}}$ (48 ± 10 nM) that is lower than that obtained from a titration of RbpA alone (177 ± 23 nM), suggesting that the presence of CarD increases the binding affinity of RbpA to transcription initiation complexes (Figure 4B). Likewise, a CarD titration in the presence of 2 μM RbpA results in a $K_{\text{eff}}$ (16 ± 2 nM) that is lower than a titration of CarD alone (73 ± 4 nM), indicating that the presence of RbpA allows CarD to interact with the complex at lower concentrations (Figure 4C). This observation of heterotropic, positive linkage between RbpA and CarD demonstrates that they act cooperatively on transcription initiation complexes at the Mtb rrrAP3 promoter.

To provide an overall quantification of the observed linkage, we consider the thermodynamic cycle involving four states: (i) RP, (ii) RP-CarD, (iii) RP-RbpA and (iv) RP-RbpA-CarD, where RP represents promoter-bound RNAP and includes an ensemble of states (i.e. RP₀, RP₁ and all intermediates) (Figure 4D). We globally fit the four experimental binding curves (CarD titrations ± 2 μM RbpA and RbpA titrations ± 1 μM CarD) to a model where each factor has an effective affinity ($K_{\text{eff}}$) and linkage is captured by the cooperativity factor $\alpha$ (Supplementary Figure S5).

In the context of this model, the effective affinity of a factor in the presence on the other is given by $K_{\text{eff}}/\alpha$. A fit of the data with this three-parameter model results in values of $K_{\text{eff,CarD}} = 66 ± 7$ nM, $K_{\text{eff,RbpA}} = 175 ± 20$ nM and $\alpha = 3.8 ± 0.4$. All four $K_{\text{eff}}$ values are within error of the values obtained from fits of the individual titration traces alone. Thus, a simple model of cooperativity captures the positive linkage between RbpA and CarD at initiation complexes.

Given that RbpA and CarD cooperatively stabilize open complex, we hypothesized that they would also stimulate transcription above the level of either factor acting alone. To test this hypothesis, we performed an aborted transcription assay in which we measured the production of a three nucleotide transcript from the rrrAP3 using a dinucleotide primer (GpU) and a radio-labeled UTP. RbpA and CarD acting individually at saturating concentrations stimulate transcription of the 3-nt product over RNAP alone (Supplementary Figure S6). The presence of both factors results in a further increase in 3-nt production over either factor acting alone. These results are consistent with a model in which CarD and RbpA cooperatively stabilize a transcription-competent open complex.

RbpA and CarD act together to accelerate the approach to RP₀ equilibrium

To identify any cooperative kinetic effects stemming from the presence of both RbpA and CarD on open complex equilibration, we used a triple exponential to fit the raw data traces and examined the manner in which the observed rates (Figure 5B and C) and the fractional amplitudes (Supplementary Figure S7) depended on transcription factor concentration with both factors present. Compared to RbpA and CarD individually, we found that $k_{\text{obs,2}}$ and $k_{\text{obs,3}}$ were faster when both factors were present. While the presence of CarD did not affect the dependence of the observed rates on RbpA concentration, the presence of RbpA did affect the dependence of the observed rates on CarD concentration. Titrations of CarD performed in the presence of RbpA at saturation indicated that both $k_{\text{obs,2}}$ (Figure 5B) and $k_{\text{obs,3}}$ (Figure 5C) now increase in a saturable and monotonic manner. This trend is quite different than what is observed for CarD titrations performed in the absence of RbpA, in which $k_{\text{obs,2}}$ does not exhibit a discernible trend, and $k_{\text{obs,3}}$ exhibits non-monotonic behavior.

To facilitate a general kinetic comparison of RbpA and CarD acting individually and together, we calculated an amplitude-averaged rate as a means to quantify an overall, apparent rate of open-complex equilibration (Materials and Methods). This analysis supports the conclusion that open-complex equilibration is faster in the presence of both factors (Figure 6). Specifically, the averaged rate ($k_{\text{obs}}$) observed with saturating concentrations of both factors is at least 3 times faster than in the presence of either factor alone. Interestingly, the presence of both factors leads to traces with similar kinetics and equilibrium fluorescence to those obtained using the RNAP holoenzyme from E. coli (Figure 7).

DISCUSSION

We have presented a study of RbpA and CarD acting individually and together to stabilize mycobacterial transcription open complexes. Like CarD, we found that RbpA stabilizes RP₀ at the Mtb rrrAP3 promoter, albeit via a different kinetic mechanism. Using a stopped-flow fluorescence assay that reports on open-complex formation in real time, we observed that, compared to CarD, RbpA generates a lower equilibrium fluorescence at saturating concentrations and has a weaker apparent affinity for initiation complexes. In addition, RpbA exhibits two observed rates ($k_{\text{obs,2}}$ and $k_{\text{obs,3}}$) with appreciable fractional amplitudes in contrast to the one dominant observed rate ($k_{\text{obs,1}}$) detected in the presence of CarD. The magnitudes of both of these rates show monotonic and saturable increasing with increasing RbpA concentration. These results all suggest that the ki-
Figure 4. RbpA and CarD bind to initiation complexes cooperatively. (A) Fluorescent fold change as a function of time for DNA alone (black), MboRNAP alone (green), 2 μM RbpA (red), 1 μM CarD and 2 μM RbpA + 1 μM CarD (magenta). (B) Normalized fluorescence fold-change as a function of RbpA concentration in the absence (red) and presence (magenta) of 1 μM CarD. (C) Normalized fluorescence fold-change as a function of CarD concentration in the absence (red) and presence (magenta) of 2 μM RbpA. (D) A thermodynamic cycle between apo-RP, CarD-bound RP, RbpA-bound RP and the doubly-bound RP. Here, RP represents the total population of DNA-bound polymerase and includes RPo and RPs. The four titrations shown in (B) and (C) were globally fit with three parameters: Keff,CarD, Keff,RbpA and a cooperativity factor which relates the ratio of Keff of a factor in the presence and absence of the other. The fits are shown in (B) and (C) along with the values for Keff. The fit cooperativity factor, α = 3.8 ± 0.4, corresponds to a ∆ΔG = −RTln(α) = −3.3 kJ/mol (−0.8 kcal/mol). Two to three shots were collected for each condition and error bars represent standard error of the mean.

netic mechanism by which RbpA stabilizes mycobacterial open complexes is distinct from that of CarD.

A two-step reversible model (R+P ↔ RPc ↔ RPo) can capture the kinetics of open complex formation under certain conditions. However, this model is almost certainly an over-simplification of the real mechanism. For example, we have shown that kinetic traces of open complex formation by MboRNAP alone exhibit three observed rates which cannot be accounted for in the context of the two-step reversible model. Furthermore, the kinetic mechanism of open-complex formation by E. coli RNAP includes multiple closed and open intermediates (1,28–30) and transcription with mycobacterial RNAP has previously been analyzed with E. coli derived models (31). Therefore, the kinetic mechanism of mycobacterial open complex formation likely involves intermediates between the initial closed complex and the final transcription-competent open complex. We hypothesize that the distinct kinetic signatures of RbpA and CarD are due, in part, to their differential affinities with these intermediates. Furthermore, intermediate open complexes may have different fluorescence properties leading to the differential enhancements observed between RbpA and CarD. Another possible way to expand the two-step reversible model to account for the complexity of the observed kinetics is to add the interaction of these factors with free polymerase in the absence of DNA. RbpA is known to bind sigma factor, core RNAP and holoenzyme (14,15,22) and CarD binds to RNAP (10,18,19). In light of these possibilities, analysis using kinetic models that include multiple intermediates and DNA-independent assembly states remains an ongoing research direction (Supplementary Figure S4B).

In addition to studying each factor individually, we performed experiments using both RbpA and CarD together. Stopped-flow titrations indicated that the presence of one factor increases the apparent affinity of the other factor to the RNAP-promoter complex, demonstrating positive linkage between the two factors. The observation of positive linkage demonstrates that not only can CarD and RbpA bind the initiation complex simultaneously, but also that they do so cooperatively. Cooperativity between transcription factors has been observed for transcription initiation
Figure 5. Kinetic effects of cooperative stabilization. (A) The intermediate phase ($k_{obs,2}$) increases monotonically with concentration for either RbpA titrated in the presence of CarD (red) or CarD titrated in the presence of RbpA (blue). (B) The slowest phase ($k_{obs,3}$) increases monotonically with concentration for either RbpA titrated in the presence of CarD (red) or CarD titrated in the presence of RbpA (blue). Two to three shots were collected for each condition and error bars represent standard error of the mean.

This cooperativity can lead to activation or repression and affords complexity and tunability to gene regulation (37). Furthermore, cooperativity amongst transcription factors can occur between the same factors (homotropic) or different factors (heterotropic) (32–34). In *Mtb* specifically, cooperative binding of DevR factors to their DNA sites plays a role in the activation of a regulon required for the induction of dormancy in response to hypoxic conditions (38,39). However, to our knowledge the cooperative association of RbpA and CarD is the first example of heterotropic cooperativity between transcription factors that directly bind initiation complexes to stabilize open complex.

When speculating about the mechanism of cooperativity between RbpA and CarD, we consider two general possibilities that are not mutually exclusive: (i) a ‘direct’ mechanism in which RbpA and CarD physically interact, so that the presence of one factor provides an additional binding surface for the other; and (ii) an ‘allosteric’ mechanism where the binding of one factor results in conformational changes in the initiation complex that lead to higher affinity binding of the second factor (i.e. conformational selection). Although a ‘direct’ interaction between RbpA and CarD is not predicted by structural modeling, we do not exclude the possibility that direct interactions between the two factors contribute to the observed positive linkage. In particular, the location of RbpA’s core domain in the initiation complex is unknown and predicted to be in close proximity to the C-terminal domain of CarD. The allosteric mechanism, however, must contribute to the observed positive linkage through conformational selection of $R_P^o$. RbpA and CarD each must have a higher affinity to $R_P^o$ than to $R_P$, as they both stabilize $R_P^o$. Thus, the binding of one factor shifts the population of initiation complexes to the higher affinity open-complex conformation(s), resulting in a lower $K_{eff}$ for the other factor. Whether the allosteric mechanism of conformational selection accounts for all or just some of the positive linkage remains an active area of investigation.

When studying the effect of both factors acting together on the kinetics of open-complex equilibration, we noticed a substantial acceleration in $k_{obs,2}$ and $k_{obs,3}$, demonstrating that the overall approach to equilibrium is faster compared to either factor acting alone (Figures 5 and 6). Based on this observation, we hypothesize that the interactions of both factors with the initiation complex accelerate distinct forward rates on the pathway to open-complex formation. The pathway to open complex determined with the *E. coli* system involves many structural intermediates including DNA bending and wrapping, DNA unwinding and loading of the template strand, assembly of the clamp/jaw domains and closing of the clamp/jaw domains around the downstream DNA (1,30). More specifically, the binding of downstream duplex into the polymerase cleft has been linked with the
nucleation of DNA melting (30). The basic linker of RbpA interacts with the extended −10 motif (15) and tryptophan 85 in CarD's DNA-binding domain interacts with the upstream edge of the transcription bubble at the junction between the double-stranded −12 base-pair and the single-stranded −11 position (11,18), precisely where nucleation of promoter melting occurs (Figure 1B). Taken together, these structural considerations suggest the possibility that RbpA and CarD cooperate to facilitate promoter bending and nucleation of promoter-melting. This combined mechanism of RbpA and CarD may be similar to the induced-fit mechanism of bending and opening used to describe the effect of the transcription factor MlfI on mitochondrial open-complex kinetics (40). Importantly, this model is also compatible with the proposed conformational-selection mechanism in which CarD acts to prevent bubble collapse after promoter DNA has been opened (11,12,18).

We also show that RbpA and CarD jointly lead to transcriptional dynamics for MboRNAP similar to E. coli RNAP (Figure 7). This result leads us to speculate that RbpA and CarD play the role of general transcription factors for Mtb, at least at housekeeping ρ^−-dependent promoters. In light of the cooperativity between the two factors, the preference of RbpA (14,22) for σ^A and σ^B holoenzymes and the presumed sigma-independence of CarD raise interesting questions regarding the regulatory logic behind each of these essential factors. For example, the cooperativity may provide a mechanism for the preferential recruitment of CarD to RbpA-dependent promoters. Alternatively, the association of CarD with other sigma factors may lead to RbpA binding at non-σ^A or σ^B promoters. In addition, our data predict that the overall effect on open complex equilibrium at initiation complexes containing both factors will be greater than those containing either factor alone. This effect could lead to the different regulatory outcomes depending on which factors are present at specific promoters.

In summary, we describe the kinetics and concentration dependencies of RbpA and CarD acting individually and cooperatively on open-complex formation at the Mtb rrnA P3 promoter. We expect the work presented here to provide a quantitative framework that can be used to develop mechanistic models of RbpA and CarD. We hypothesize that the concurrent binding and positive linkage between these essential transcription factors play important roles in mycobacterial gene regulation in that they result in (i) the more efficient recruitment of transcription factors to initiation complexes and (ii) the rapid formation of a more stable open complex. Further studies are needed to describe the detailed kinetics of both factors on all phases of transcription initiation to understand their overall effect on transcriptional flux at promoters throughout the mycobacterial genome.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

The authors would like to thank Dr Tim Lohman for the use of his stopped-flow spectrophotometer and Dr Tom Ellenberger for the use of the HPLC and FPLC in his laboratory. We also thank Dr Roberto Galletto, Dr Lohman, Dr Ellenberger and Dr Tomasz Heyduk for scientific input at various stages of this work.


FUNDING

National Institutes of Health (NIH) [R01GM107544 to E.A.G. and C.L.S.]; Sigma Aldrich Predoctoral Fellowship [to J.R.]; NIGMS Cell and Molecular Biology Training Grant [GM007067 to J.P. and A.L.G.]; Stephen I. Morse Graduate Fellowship [to A.L.G.]; Funding for open access charge: NIH [R01GM107544].

Conflict of interest statement. None declared.

REFERENCES


Supplemental Information

Cooperative stabilization of *Mycobacterium tuberculosis* *rrnAP3* promoter open complexes by RbpA and CarD.

Jayan Rammohan\(^1\), Ana Ruiz Manzano\(^1\), Ashley Garner\(^2\), Jerome Prusa\(^2\), Christina L. Stallings\(^2\), and Eric A. Galburt\(^1\)*

* To whom correspondence should be addressed. Tel: (314)362-5201 Email: egalburt@biochem.wustl.edu

\(^1\) Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine, St. Louis, MO 63110, USA

\(^2\) Department of Molecular Microbiology, Washington University School of Medicine, St. Louis, MO 63110, USA
Supplemental Sequence

Non-template strand of 150 bp Mtb rrnA P3 promoter fragment used in this study (genome coordinates 1471577 – 1471727). The -10 and -35 boxes (blue), +1 start site (red), and +2 T modified with a C6-amine and labeled with Cy3-NHS (green) are indicated [GonzalezYMerchand:1996-va].

GGCGACGTCACCTATGGATATCTATGGATGACCGAACCTGGTCTTGACTCCATTGCCGGATTTGTATTAGACTGGCAGGGTTGCCCCGAAGCGGGCGGAAACAAGCAAGCGTGTTGTTTGA-GAACTCAATAGTGTTTGGTGGTTTCA
Supplemental Figure 1: RbpA R79A mutant is defective in open complex stabilization. The fluorescence fold change over RNAP alone is shown for 2 μM WT RbpA, 2 μM R79A RbpA, and 1 μM WT CarD. Error bars represent standard error of the mean. The inset shows the raw time traces for 10 nM DNA alone (black dashed), 225 nM RNAP and 10 nM DNA (black), and 225 nM RNAP, 10 nM DNA, and 2 μM R79A RbpA (red).
Supplemental Figure 2: Fractional amplitude dependence on factor concentration for each factor individually. Triple exponential fits to RbpA and CarD time courses exhibit different fractional amplitudes. For both RbpA (dashed lines) and CarD (solid lines), the amplitude of the fastest phase ($k_{\text{obs},1}$, blue) drops below 10% as the concentration of each factor approaches 1 µM. Unlike CarD, the fractional amplitudes of RbpA fits are not dominated by the slowest phase ($k_{\text{obs},3}$, red); instead, the intermediate phase ($k_{\text{obs},2}$, green) exhibits a more significant contribution.
Supplemental Figure 3: Concentration dependence of the observed rates with factor concentration plotted in log scale. (A) The second observed rate ($k_{obs,2}$) as a function of the concentration of CarD (blue) and RbpA (red). (B) The third observed rate ($k_{obs,3}$) as a function of the concentration of CarD (blue) and RbpA (red). (C) The second observed rate ($k_{obs,2}$) as a function of the concentration of CarD in the presence of 2 µM RbpA (blue) and of the concentration of RbpA in the presence of 1 µM CarD (red). (D) The third observed rate ($k_{obs,3}$) as a function of the concentration of CarD in the presence of 2 µM RbpA (blue) and of the concentration of RbpA in the presence of 1 µM CarD (red).
Supplemental Figure 4: Models for factor mechanisms. (A) The binding of RNAP polymerase (R) to the promoter (P) creates the closed complex (RP_c) which can isomerize to open complex (RP_o). A ligand (L) that may represent CarD or RbpA or any other factor, may bind to either RP_c or RP_o. Isomerizations between RP_c and RP_o may also occur in their ligand-bound states (RP_c·L and RP_o·L). (B) As described in the text, the observed kinetics dictate that a more complicated kinetic scheme be considered to account for the data. Here, we illustrate two possible additions to the scheme to consider. The first is the addition of the state (R·L + P) which represents the association of either CarD or RbpA to free, promoter-unbound holoenzyme subunits. The R here can be further expanded to specifically include contributions from core-, sigma-, and holoenzyme-bound complexes. The second is the explicit representation of intermediates (RP_i) between RP_c and RP_o as well as the possible complexes between these intermediates and CarD or RbpA (RP_i·L).
Supplemental Figure 5: Distributions of global fit parameters for cooperative binding model. Data points were randomly selected with replacement and then fit with a cooperative binding model with 3 parameters. The distributions of the affinity ($K_{eff}$) of (A) CarD and (B) RbpA alone to polymerase bound promoter complexes are shown along with (C) the cooperativity factor alpha. These distributions were fit with a gaussian to determine the both the mean and uncertainty for each parameter based on the data.
**Supplemental Figure 6: Three nucleotide aborted transcription.** The relative amount of 3-nt product compared to that generated by MboRNAP holoenzyme alone is shown as a function of the addition of CarD (2 µM), RbpA (2 µM), or both factors. Error bars represent standard errors of the mean from 8 replicates. The inset shows a gel from one of the replicates and includes lanes for each factor at both 1 µM and 2 µM as well as the combination of both factors at saturating concentrations.
Supplemental Figure 7: Fractional amplitude dependence on factor concentration in the presence of both factors. The titration of RbpA in the presence of 1 µM CarD (dashed lines) and the titration of CarD in the presence of 2 µM RbpA (solid lines) reveal similar trends in the fractional amplitudes of the fastest ($k_{obs,1}$, blue), intermediate ($k_{obs,2}$, green), and slowest ($k_{obs,3}$, red) phases. Interestingly, CarD titrations performed in the presence of RbpA no longer show a dominance of the slowest phase ($k_{obs,3}$).
Epilogue

The work presented in the chapter was extended to include mutants of CarD and RbpA. Titrations of RbpA were performed in the presence of 1 µM CarD-W85A, and titrations of CarD were performed in the presence of 2 µM RbpA-R79A. Less positive linkage was observed when RbpA was titrated with 1 µM CarD-W85A (Figure 1). This result is intuitive, since CarD-W85A exhibits reduced ability to stabilize open complexes and RbpA’s linkage must in part arise from conformational selection to the open complex. Surprisingly, titrations of CarD performed in the presence of RbpA-R79A showed negative linkage, indicating that the presence of RbpA-R79A reduced the affinity of CarD to transcription initiation complexes (Figure 2). These results suggest the possibility that RbpA-R79A interferes with CarD binding, whereas wild-type RbpA does not. These titrations were only performed once, and more repetitions are needed to verify these preliminary observations.

One future direction that remains unexplored is a more general model of cooperativity. I have analyzed the positive linkage between CarD and RbpA in terms of ternary complex formation. This linkage must arise in part from conformational selection, since each factor stabilizes the open complex, and each factor has a higher affinity for the open complex. However, there may be additional components that contribute to the observed linkage, for instance direct interactions between CarD and RbpA within the complex. This more complex model of cooperativity can be visualized as a cube (Figure 3), and represents a more general kinetic scheme which describes factor-dependent open-complex stabilization in mycobacteria.
Figure 1. Titration of RbpA in the presence of CarD-W85A. These traces indicate a half-maximal effect for RbpA ~ 50 nM, indicating reduced positive linkage compared to titrations of RbpA performed in the presence and absence of wild-type CarD.
Figure 2. Titration of CarD in the presence of RbpA-R79A. These traces indicate a half-maximal effect for CarD ~ 100 nM, indicating negative linkage compared to titrations of CarD performed in the presence and absence of wild-type RbpA.
Figure 3: General model for CarD-RbpA-RPc-RPo cooperativity. CarD and RbpA binding and linkage to closed and open complexes can be modeled as a cube. This model includes cooperativity factors $\alpha$, $\beta$, and $\gamma$ which describe various types of linkage between binding and conformational exchange as indicated above.
Chapter 5

Kinetic Models for Factor-Dependent Stabilization of Mycobacterial Open-Complexes
**Introduction and Background**

We have studied the kinetics of mycobacterial transcription open-complex formation using Cy3-labeled DNA and stopped-flow fluorescence spectroscopy. We have shown that CarD stabilizes mycobacterial open complexes using a two-tiered concentration-dependent kinetic mechanism, which can be semi-quantitatively described by a 5-state model (1). We have also shown that RbpA stabilizes mycobacterial open complexes using a kinetic mechanism distinct from that of CarD, and that RbpA exhibits two saturable observed rates for the intermediate and slow phases (2). In order to aid our interpretation of the experimental data, I have evaluated kinetic models that could potentially be used to account for the trends in observed rates. I have focused this evaluation on two objectives as described below.

**Objective 1:** Evaluate the sensitivity and uncertainty in the published 5-state model used to describe CarD stabilization of mycobacterial open complexes. I have performed a detailed, local exploration of the rate constants published for the 5-state model. I have identified rate constants that are well constrained as well as pairs of rate constants which are poorly constrained. For rate constants that are well constrained, I have quantitatively determined the ranges of values that qualitatively fit the experimental data.

**Objective 2:** Evaluate alternate kinetic models that could potentially reproduce observed trends for CarD and RbpA. For CarD, I have evaluated and excluded two 4-state schemes that are unable to capture the experimental trends. For RbpA, I have identified a kinetic
scheme that is able to qualitatively reproduce the monotonic and saturable increase in intermediate and slowest rates, with semi-quantitative agreement to experimental data.

**Objective 1: Estimate sensitivity and uncertainty in rate constants of the reported 5-state kinetic scheme describing CarD stabilization of open complex.**

We recently reported a working model for CarD stabilization of open-complexes that consists of a 5-state kinetic scheme (Figure 1). We estimated values for all 10 rate constants in this scheme by inspecting their ability to qualitatively capture empirical trends. However, we do not know the relative importance of each of these parameters, and the extent to which they can vary.

(1) We do not know if the set of 10 rate constants originally proposed *uniquely* describes the trend in observed rates. In some cases, kinetic schemes with very different sets of rate constants are capable of producing the same concentration-dependent trends in observed rates. To this end, I have investigated whether other combinations of rate constants can reproduce the experimentally observed trends in observed rates.

(2) We do not know *which rates are more important* than others with regard to recapitulating the trends in observed rates. In certain kinetic models, rate constants can vary over wide ranges without affecting the ability of the model to describe the data. In other cases, particular rate constants may be tightly constrained and even slight variation can dramatically affect the ability of the
model to describe the data. To this end, I have identified which rate constants in
the 5-state model can vary over wide ranges, and which rate constants are well
constrained.

My approach is to compare eigenvalues to observed rates to quantify uncertainty in our
estimation of model parameters, and characterize the manner in which rates (or pairs of
rates) affect the concentration-dependent relaxations of the system. An alternative
approach is to choose rate constants and simulate timecourses assuming fluorescence
enhancement for each state. Programs such as KinTek Explorer can perform global
fitting of raw data in order to identify ranges of rate constants that are able to reproduce
the observed trends (3-5). Currently, I have been unable to reproduce the raw
fluorescence timecourses using a 3-step reversible model coupled to factor-binding,
therefore those results are not presented here.
Figure 1: Kinetic model of CarD stabilization of mycobacterial open complexes. The values of the rates used to generate the plots in the main text are as follows: $k_1 = 0.3 \text{nM}^{-1}\text{s}^{-1}$, $k_2 = 75 \text{s}^{-1}$, $k_3 = 0.0005 \text{s}^{-1}$, $k_4 = 0.004 \text{s}^{-1}$, $k_5 = 0.3 \text{nM}^{-1}\text{s}^{-1}$, $k_6 = 450 \text{s}^{-1}$, $k_7 = 0.3 \text{nM}^{-1}\text{s}^{-1}$, $k_8 = 1.4 \text{s}^{-1}$, $k_9 = 0.02 \text{s}^{-1}$, $k_{10} = 0.0005 \text{s}^{-1}$. 
METHODS

A system of differential equations was used to describe the manner in which the populations of each of the 5-states change with time and concentration. A Matlab program was used to calculate eigenvalues of this matrix for a given set of rate constants and concentrations, where each nonzero eigenvalue represents a (potentially) observable rate.

\[
\begin{align*}
\lambda = & \begin{bmatrix}
\frac{d[R]}{dt} \\
\frac{d[R_{PC}]}{dt} \\
\frac{d[R_{Po}]}{dt} \\
\frac{d[X_{RPc}]}{dt} \\
\frac{d[X_{RPo}]}{dt}
\end{bmatrix} \\
= & \begin{bmatrix}
-k[R]_0 & k_2 & 0 & 0 & 0 \\
-k[R]_0 + (k_{1} + k_{2} + k_{5} \cdot [X]) & k_{4} & k_{6} & 0 & 0 \\
0 & k_{1} & -(k_{4} + k_{7} \cdot [X]) & 0 & k_{5} \\
0 & k_{5} \cdot [X] & 0 & -(k_{6} + k_{3}) & 0 \\
0 & 0 & k_{7} \cdot [X] & k_{5} & -(k_{3} + k_{10})
\end{bmatrix} \begin{bmatrix}
[R] \\
[R_{PC}] \\
[R_{Po}] \\
[X_{RPc}] \\
[X_{RPo}]
\end{bmatrix}
\end{align*}
\]

Using the reported rate constants, I plotted all 5 eigenvalues of the system for 0 < [CarD] < 10 µM, with a fixed [RNAP] = 225 nM. Each of the 4 non-zero eigenvalues showed unique dependencies on [CarD]. Interestingly, I found that a single eigenvalue (eig2) was approximately the same magnitude as \(k_{obs,3}\). Additionally, eig2 exhibited the same non-monotonic dependence on [CarD] as \(k_{obs,3}\), namely deceleration at low [CarD] and acceleration at higher [CarD].

Furthermore, calculations of eig2 vs. [CarD] performed at different [RNAP] qualitatively reproduced empirical dependence of \(k_{obs,3}\) on [RNAP]. Specifically, at lower [RNAP],
we observe less acceleration with increasing \([\text{CarD}]\) (see below). Taken together, these results suggest that eig2 can be compared to \(k_{\text{obs,3}}\) in order to evaluate the ability of model parameters to describe empirical trends.

**Figure 2:** A single eigenvalue captures non-monotonic dependence of observed rate on concentrations of CarD and RNAP. **(Top)** Experimental data for \(k_{\text{obs,3}}\) vs. \([\text{CarD}]\) measured at two different concentrations of \([\text{RNAP}]\) (1). **(Bottom)** Eig2 vs. \([\text{CarD}]\) calculated at different concentrations of \([\text{RNAP}]\).
I applied several constraints when considering whether or not a parameter set could qualitatively account for the experimental data:

1) For [CarD] = 0, 0.003 s⁻¹ \(\leq k_{obs,3} \leq 0.006\) s⁻¹

2) 25 nM < [CarD] for minimum \(k_{obs,3} < 150\) nM

3) For [CarD] = 10 µM, 0.003 s⁻¹ \(\leq k_{obs,3} \leq 0.015\) s⁻¹

**Figure 3:** Constraints for qualitatively reproducing CarD-dependent observed kinetics. First, for [CarD] = 0, 0.003 s⁻¹ \(\leq k_{obs,3} \leq 0.006\) s⁻¹. Second, 25 nM < [CarD] for minimum \(k_{obs,3} < 150\) nM. Third, for [CarD] = 10 µM, 0.003 s⁻¹ \(\leq k_{obs,3} \leq 0.015\) s⁻¹.
To examine the sensitivity and uncertainty of parameters within the 5-state model, I evaluated the effect of varying individual rate constants on an eigenvalue vs. [CarD]. To do this, I apply a scaling factor to one of the rates, and plot the eigenvalue vs. [CarD] overlaid with the experimentally-measured $k_{obs,3}$ vs. [CarD] in order to visually inspect how the model compares to the data. To quantify goodness-of-fit, I calculate the sum-squared residual error (SSE) as a function of the magnitude of the scaled rate constant. I applied scaling factors to vary a rate constant over 5 orders of magnitude (i.e. 0.001*k to 100*k), as well as finer scaling (i.e. 0.1-2.5) to examine how global versus local variance relates to SSE. For instances in which the SSE was “low” ($\leq 1 \times 10^{-5}$), the fit was considered to quantitatively describe the trend within error, and the SSE cutoff was used to estimate uncertainty in the rates as shown below.

In addition to $k_1$ and $k_2$, which are the on- and off- rates for RNAP-promoter binding, there are 28 pairs of rates that need to be studied in order to fully characterize all of the parameters of the 5-state model. This is because for any of the 8 rates ($k_{3-10}$) that are part of the scheme’s thermodynamic cycle, when one rate constant is varied another rate must also vary in order to maintain detailed balance. To study how pairs of rates may be related, I evaluated all 28 pairs of rate constants twice (scaling one over 5 orders of magnitude and calculating the other member of the pair, as well as the converse), and identified parameter sets that were able to reasonably describe the experimental data.
Results

The ability of the model to reproduce experimental trends depends on both $k_1$ and $k_2$

I have studied the effect of independently varying $k_1$ (on-rate of RNAP to promoter) and $k_2$ (off-rate of RNAP from promoter) over 5 orders of magnitude. The plots look the same as what is shown for the titration of [RNAP] (Figure 2), indicating that for this particular trend, the fraction of bound promoter is more important than the magnitudes of the on- and off-rates themselves.

Pairs of on- and off-rates of CarD binding to the same state are poorly-constrained

When scaling/calculating a pair of rate constants over a large range, in some instances I have observed almost no change in the plot of eigenvalue vs. [CarD]. In these instances, the rate constants themselves can vary widely over a large range with virtually no effect on the goodness-of-fit. These rates correspond to the on- and off-rates of CarD to RP$_c$ ($k_5$ and $k_6$) and to RP$_o$ ($k_7$ and $k_8$). For these pairs of rates, it is not their absolute magnitude that matters, but rather their ratio.

One example of this can be seen when scaling $k_5$ (on-rate of CarD to RP$_c$) and calculating $k_6$ (off-rate of CarD from RP$_c$) in order to maintain detailed balance. In this case, the plots of eig2 vs. [CarD] for each parameter set look virtually the same even though the rates are being varied over 5 orders of magnitude (Figure 4). Interestingly, the SSE indicates an improvement in fit as the rate is increased, suggesting that the rate constants themselves do actually affect the shape of the curve. The improvement in SSE is modest at best when compared to the improvement in SSE observed for other pairs of

115
rate constants, and is only bounded on the lower end. However, this improvement is imperceptible in the plot of eig2 vs. [CarD], demonstrating that the observation of a decreasing SSE needs to drop below a meaningful number that is representative of a good fit to the data.

All other pairs of rate constants in the cycle (besides pairs of on- and off-rates of CarD to the same state) are well constrained

An example of a well-constrained rate pair is the result of scaling $k_5$ over 5 orders of magnitude and calculating $k_3$ to maintain detailed balance (Figure 5, top). In this case, the plot of eig2 vs. [CarD] changes dramatically as the rates are scaled. Some of the curves overlay with the experimental data within error. These curves have an SSE < $10^{-5}$, and this cutoff can be used to estimate uncertainty in a given rate (Figure 5, bottom).

One pair of rate constants ($k_3$ and $k_{10}$) has at least two non-unique, well-constrained values that provide good fits

I found that using $1/$SSE instead of SSE made it easier to visualize the error, and in this case a “good fit” is $1/$SSE > $10^5$ (instead of SSE < $10^{-5}$). When $k_3$ and $k_{10}$ were varied together, there are multiple, separable values that each provide good fits to the data. This can easily be seen when plotting $1/$SSE versus $k_3$ (Figure 6, top) and $1/$SSE versus $k_{10}$ (Figure 6, bottom).
**Figure 4:** Pairs of on- and off-rates for CarD to/from the same state are poorly constrained. (Top) When $k_5$ is varied with $k_6$, the rates can change over 5 orders of magnitude without any perceptible change in eig2 vs. [CarD]. (Bottom) Plot of SSE vs. $k_5$ shows that the fit slightly improves as $k_5$ increases, however the improvement is imperceptible in the plot of eig2 vs. [CarD].
Figure 5: Example of well-constrained pairs of rate constants. (Top) When $k_5$ is scaled over 5 orders of magnitude and $k_3$ is calculated to maintain detailed balance, the shape of $e_{ig2}$ vs. [CarD] changes dramatically. Some of these curves describe the experimental data within error, and others do not. (Bottom) The plot of SSE vs. $k_5$ shows that an SSE $< 10^{-5}$ provides a reasonable range for $k_5$, where the optimal value is determined at the minimum SSE.
Figure 6: Example of a well-constrained, non-unique pair of rate constants. When $k_3$ and $k_{10}$ are scaled together, two different combinations of rate pairs fit the experimental data very well. $1/$SSE is used to better visualize error in the data, and in this case a peak over $10^5$ indicates a good fit. **(Top)** $1/$SSE vs. $k_3$ indicates 2 separable peaks, each of which represent sets of rate constants that fit the data well. **(Bottom)** $1/$SSE vs. $k_{10}$ indicates 2 separable peaks, each of which represent sets of rate constants that fit the data well. The two different combinations of $k_3$ and $k_{10}$ represented by each set of peaks are shown in boxes in black and blue text.
Summary and Future Directions

The results are summarized below (Figure 7). In some cases, the pair of rates is poorly constrained (red squares, below), since they can be varied over several orders of magnitude without affecting the goodness-of-fit. In these cases, the ratio of the rates is more important than their absolute values. In all other cases, the pair of rates were well constrained qualitatively according to criteria described above (yellow squares, below) and in some cases described the experimental data very well (SSE $\leq 10^{-5}$, green squares, below).

**Figure 7:** Graphical summary of scaling all 28 pairs of rate constants in the thermodynamic cycle. Scaled rates are shown on the top row, and calculated rates are shown on the left column. Red squares indicate poorly constrained rate pairs, yellow squares indicate well-constrained rate pairs, and green squares indicate good fits (SSE $< 10^{-5}$). The bold outlines are for $k_3$ and $k_{10}$, which have 2 well constrained, non-unique sets.
The model below summarizes ranges of each rate constant that can be used to qualitatively describe the CarD trend (presented as mean and standard deviation). These were calculated by averaging the mean values of an individual rate when calculated from every other scaled rate pair that was well constrained.

**Figure 8:** Summary of estimates of rate constants and uncertainty. Best estimates for each of the rate constants are shown above along with their estimated error. For $k_3$ (opening in the absence of CarD) and $k_{10}$ (closing in the presence of CarD), two different pairs of rate sets could each fit the data equally well (corresponding pairs are shown in black or blue).

For future work, it would be efficient to utilize a program that could explore a wider range of rate space so that combinations of rate constants far outside of the starting rate set can be evaluated in an efficient manner. Such approaches have been used to evaluate kinetic models using Mathematica programs (6).
**Objective 2: Evaluate alternative kinetic schemes for CarD and RbpA.**

To this point, we have only considered in detail one kinetic scheme to describe the mechanism of RPr stabilization by CarD. It is possible that other schemes of equal or lesser complexity could potentially be used to capture empirical trends. Furthermore, we have not proposed a kinetic scheme that can potentially be used to account for RbpA. To this end, I have evaluated alternative kinetic schemes with regard to their ability to reproduce experimentally observed trends.

My approach had been to (1) identify alternative kinetic schemes other than the 5-state model that could also potentially be used to describe factor-dependent stabilization of RPr, and (2) calculate eigenvalues for these schemes across a range of concentrations and rate constants to determine which, if any, could be used to reproduce observed trends. These models are evaluated using Matlab to calculate eigenvalues as a function of ligand concentration. Priority has been given to testing models with fewer than 5 states to see if they are capable of reproducing experimental trends.
Results

Non-monotonic CarD data can be reproduced using a 4-state cycle, but this cannot account for [RNAP] dependence.

Figure 9: A 4-state thermodynamic cycle reproduces non-monotonic trend in observed rates, but cannot account for [RNAP] dependence.

I have been able to qualitatively reproduce the trend in non-monotonic observed rates using a kinetic scheme simpler than the 5-state model. Specifically, the non-monotonic trend can be reproduced using only a 4-state cycle in which conformational exchange is linked to ligand binding. This classic cycle represents two general pathways of ligand binding to substrate, namely conformational-selection and induced-fit. However, while this scheme can reproduce the non-monotonic dependence of $k_{\text{obs,}3}$ on [CarD], it needs more states in order to account for the [RNAP]-dependence, and thus cannot reproduce all of the experimentally-observed trends.
A linear 4-state mechanism cannot account for CarD data

An induced-fit model is capable only of increasing rates with ligand concentration, therefore it cannot account for CarD data. Conversely, conformational selection is capable of producing increasing or decreasing observed rates, depending on the relative values of the rate constants in the model. Since a linear 4-state conformational-selection kinetic scheme has the potential to exhibit deceleration and acceleration, I have evaluated this model with regard to its ability to exhibit both (a non-monotonic trend) with increasing [CarD]. I have explored ranges of rates that exhibit both increasing and decreasing rates, however so far using this scheme I cannot reproduce the non-monotonic dependence of $k_{obs,3}$ on [CarD] (Figure 10).

A 4-state mechanism of conformational-selection and induced-fit can potentially account for RbpA data

Using a 4-state scheme in which conformational-selection is followed by induced-fit, I have identified a kinetic scheme that semi-quantitatively reproduces the trend in multiple observed rates with increasing [RbpA] (Figure 11). This scheme involves a pre-equilibrium of the RNAP-promoter complex, followed by preferential RbpA binding (conformational-selection), followed by conformational exchange in the presence of ligand (induced-fit). Interestingly, this scheme suggests the possibility that RbpA associates with more than one state. This scheme can serve as a starting point from which to build our understanding of how RbpA stabilizes mycobacterial open complexes.
Figure 10: A 4-state linear mechanism is unable to reproduce a non-monotonic trend in observed rate. (A) 4-state linear kinetic scheme. (B) With a slow pre-equilibrium and fast binding, only deceleration is observed with increasing [CarD]. (C) With a fast pre-equilibrium and slow binding, either acceleration or deceleration can be observed with increasing [CarD], depending on [RNAP].
Figure 11: A 4-state linear scheme which incorporates conformational selection and induced fit can account for RbpA’s trend in observed rates. (A) Kinetic scheme. (B) Eig2 and Eig3 vs. [RbpA] semi-quantitatively agree with experimentally-observed results.
**Summary and Future Directions**

In summary, I have presented a quantitative evaluation of multiple kinetic schemes with regard to their ability to reproduce empirical trends in CarD and RbpA data. Schemes with more states (5 or more) should be evaluated with regard to their ability to reproduce trends. For these studies it will be useful to use a more sophisticated computational approach that explores a larger rate space in a more efficient manner (6).
References


Chapter 6

Promoter Escape
Introduction

The flux of transcription from a given promoter depends on all phases of transcription initiation, including promoter-binding, promoter-opening, nucleotide incorporation, and promoter escape. The stability of these states and the rates of conversion between them represent quantitative biophysical parameters that can be used to understand transcriptional flux at a mechanistic level.

We have used an ensemble fluorescence assay to study CarD stabilization of mycobacterial transcription open complexes. These studies showed that CarD’s kinetic mechanism of open-complex stabilization is concentration-dependent and is consistent with a model in which at low concentrations, CarD stabilizes $R_{P_0}$ by conformational-selection and at high concentrations, CarD stabilizes $R_{P_0}$ by induced-fit. These studies provide a useful quantitative framework to understanding the initial stages of transcription initiation. To use this model to understand CarD’s overall effect on transcriptional flux, I sought to measure the effect of CarD on NTP-dependent processes such as nucleotide incorporation and promoter escape by using the ensemble fluorescence assay in the presence of NTPs.
Results and Conclusions

Dilution due to symmetric mixing has different effects on pre-assembled MboRNAP- and EcoRNAP-promoter complexes

Initial experiments were designed using standard equal-volume mixing in which one syringe contained pre-equilibrated MboRNAP-promoter complexes, and the other syringe contained buffer, competitor (dsDNA), or competitor and NTP’s (Figure 1). When the F-syringe contained pre-mixed MboRNAP/CarD/Cy3-promoter DNA, and it was mixed with excess dsDNA (100 bp consensus promoter), the fluorescence level gradually decreased over the course of 20 min (Figure 1, green trace). When this F syringe was mixed with dsDNA/NTP’s, a slight increase in fluorescence was observed in the first 5 seconds, followed by a dramatic multi-exponential decrease in fluorescence intensity that equilibrated over the course of 20 minutes (Figure 1, red trace). At first glance, one might be compelled to interpret the green dsDNA trace as a measure of the rate of dissociation, and the red dsDNA/NTP trace as a measure of dissociation combined with promoter escape. However, when this same F syringe was mixed with buffer alone, the fluorescence intensity showed an increase over time that equilibrated over the course of 20 min (Fig 1, blue trace). This result suggests that when the protein/DNA mix is diluted two-fold, the proteins re-equilibrate with the DNA over the course of 20 min. Therefore, it is difficult to interpret the traces with competitor and/or NTP’s as solely reporting on dissociation and/or promoter escape, because these traces likely include some re-equilibration of the system.
Figure 1: Dilution-dependent dissociation/re-equilibration must be accounted for when studying dissociation and promoter escape by *Mbo*RNAP. Pre-equilibrating *Mbo*RNAP, CarD, and promoter DNA and then performing an equal-volume mix with buffer (blue), dsDNA competitor (green) or dsDNA and NTPs (red) each show different time courses over 20 minutes. Importantly, the buffer trace indicates that the *Mbo*RNAP-CarD-promoter complexes must re-equilibrate after dilution due to mixing, which is an effect that must be accounted for when analyzing dissociation and single-turnover promoter escape.
Using symmetric mixing, pre-equilibrated RNAP-promoter complexes were mixed with dsDNA +/- NTPs (Figure 2). Final reaction concentrations were 10 nM Cy3-DNA, 225 nM MboRNAP, 1 µM CarD, 400 nM dsDNA competitor, and 150 µM NTP. In the absence of CarD, complexes mixed with dsDNA appear to be competitor-resistant over 10 minutes (Figure 2, blue trace), and in the presence of dsDNA and NTPs (Figure 2, green trace) there appears to be a slight decay over time but almost no difference from the trace with dsDNA only. In the presence of CarD, complexes mixed with dsDNA (Figure 2, red trace) appear competitor-resistant but a gradual decay can be observed over the course of 10 minutes ($k_{obs} = 0.0013 \text{ s}^{-1}$). When mixed with dsDNA and NTPs (Figure 2, yellow trace), a decay over the course of 10 minutes can be clearly distinguished from the red trace. The rate of this decay can be fit to a single-exponential decay from 1-600 ($k_{obs} = 0.0043 \text{ s}^{-1}$) or from 20-600 seconds to avoid the contribution of the initial part of the trace ($k_{obs} = 0.0056 \text{ s}^{-1}$). Alternately, the decay can be fit to a double exponential from 20-600 s ($k_{obs,1} = 0.0163 \text{ s}^{-1}$, $k_{obs,2} = 0.0036 \text{ s}^{-1}$) which captures a fast and slow component of the decay. Interestingly, the dissociation kinetics of CarD-MboRNAP-promoter complexes in the presence of competitor DNA and NTP’s has been probed using runoff transcription as a readout and was also described using a double exponential ($k_{obs,1} = 0.0290 \text{ s}^{-1}$, $k_{obs,2} = 0.00013 \text{ s}^{-1}$) (1). Differences in the observed rates may be due to many factors including their use of a different competitor (AC50 bubble promoter), solution conditions (glutamate instead of chloride) and time-resolution of the assay (8 time points collected over the course of 200 minutes).
Figure 2: Equal volume mixing in the presence and absence of CarD using dsDNA competitor under single-turnover conditions. The traces shown above indicate fluorescence fold change over 10 min, and include traces for Cy3-promoter alone (black), MboRNAP-promoter complexes mixed with dsDNA (blue) and dsDNA+NTPs (green), and CarD-MboRNAP-promoter complexes mixed with dsDNA (red) and dsDNA+NTPs (yellow).
The observation of an NTP-dependent decay that is distinguishable from the dsDNA-dependent decay indicates that the signal is providing information on nucleotide incorporation and promoter escape. Interestingly, a small increase in signal can be observed over the first few seconds of the trace before the decay begins. This lag preceding the signal decay was observed in previous experiments, and could be due to fluorescence arising from intermediate states during initial nucleotide incorporation, which precede the transition into elongation and subsequent reannealing of the transcription bubble at the +2 to become double-stranded once again. Importantly, this experiment did not include mixing the RNAP/promoter complexes with buffer alone, which is an important negative control. Previous experiments (described above) suggest that there is a dilution-induced dissociation that occurs with this system that must be accounted for when measuring rates of dissociation and escape. It is therefore useful to develop a method of performing these experiments that minimizes dilution-induced dissociation.

Symmetric mixing performed using *Eco*RNAP indicate that *Eco*RNAP binds this promoter with much higher affinity than *Mbo*RNAP, and forms a more stable RPo (2). As a result, the two-fold dilution may not perturb the system enough to cause dissociation and subsequent re-equilibration of *Eco*RNAP over time. To examine this system, I studied reactions containing final concentrations of 10 nM Cy3-DNA, 42.5 nM *Eco*RNAP, 400 µg/ml Heparin and NTP ranging from 25 µM to 1 mM. First, I mixed *Eco*RNAP in the F syringe with DNA in the C syringe and collected a typical “melting” curve and measure fluorescence increase over time (Figure 3, green trace). When pre-
mixed EcoRNAP/promoter DNA was mixed with buffer, I found that the trace was relatively flat and almost exactly as the same fluorescence level of where the melting curve ended (Figure 3, black trace). This is consistent with almost no dilution/re-equilibration effect for the EcoRNAP system, which is in stark contrast the MboRNAP, which re-equilibrates over the course of 20 min. When pre-mixed EcoRNAP/promoter DNA complexes were mixed with Heparin (Figure 3, blue trace), the trace almost overlays with that of the EcoRNAP/DNA mixed with buffer trace, suggesting that the EcoRNAP transcription complexes are Heparin-resistant. This is consistent with previous results indicating that EcoRNAP has a higher affinity and forms much more stable open complexes at Mtb rrnAP3 than MboRNAP (2). I tested the effect of mixing RNAP/DNA with Heparin/NTPs using [NTP] = 25, 250, and 1000 µM (Figure 3, yellow, orange, and red traces respectively) and was able to observe faster decay with increasing NTP concentration. The effect seems to saturate a concentration of [NTP]’s that is comparable with previous studies using EcoRNAP at a different promoter (3). Plots of the observed rate vs. [NTP] could be used to constrain kinetic models of EcoRNAP promoter escape from the Mtb rrnAP3 promoter.
Figure 3: Promoter escape by *E. coli* RNAP can be measured using equal volume mixing under single-turnover conditions. *Eco*RNAP was mixed with Cy3-promoter to collect a melting curve (green). A control traces indicated that *Eco*RNAP-promoter complexes in one syringe mixed with buffer equilibrated to approximately the same fluorescence fold change value (black). *Eco*RNAP-promoter complexes mixed with Heparin (blue) almost overlay with buffer traces, demonstrating Heparin resistance that is typical of stable open-complexes. In addition to Heparin, NTP’s were included at concentrations of 25 (yellow), 250 (orange) and 1000 (red) µM, and these traces exhibited decay that reports on the kinetics of promoter escape under single-turnover conditions.
Figure 4: Calibration of 10:1 asymmetric mixing using Cy3-labeled DNA. Traces shown are 10 s of raw fluorescence intensity resulting from 10:1 mixes of buffer-buffer (blue), buffer-DNA (teal), DNA-buffer (red) and DNA-DNA (green).
To measure promoter escape using the mycobacterial transcription system, I had to identify a strategy to observe competitor/NTP-dependent fluorescence decay even though the MboRNAP’s affinity to promoter DNA will result in dissociation due to dilution upon 1:1 mixing. One way to mitigate the dilution effect, and thus minimize re-equilibration of MboRNAP-promoter complexes upon mixing, is to perform asymmetric 10:1 mixing experiments. This can easily be accomplished by replacing one of the 2.5 mL syringes in the Applied Photophysics SX-20 with a 250 µl syringe. Care should be taken to increase the total shot volume to at least 250 µl (For example, 230 µl F syringe + 23 µl C syringe = 253 µl total shot volume), to ensure that enough fluid from the smaller syringe enters the reaction cell of the instrument. Also, the pressure should be reduced to < 2 bar using the control knob behind the instrument to avoid damaging the smaller syringe.

Calibration experiments using a 2.5 ml F syringe and a 250 µl C syringe were performed to measure dilution effects. Four samples were tested (Figure 4), and the results are summarized below:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>Raw Signal (V)</th>
<th>Signal-Buffer % max (predicted)</th>
</tr>
</thead>
<tbody>
<tr>
<td>F</td>
<td>C</td>
<td>Buffer  Buffer</td>
<td>0.358563</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Buffer  DNA</td>
<td>0.843319</td>
</tr>
<tr>
<td>DNA</td>
<td>Buffer</td>
<td>5.516577</td>
<td>5.158014</td>
</tr>
<tr>
<td>DNA</td>
<td>DNA</td>
<td>6.139301</td>
<td>5.780738</td>
</tr>
</tbody>
</table>

These control experiments demonstrate that for the given experimental setup (2.5 ml F-syringe, 250 µl C-syringe, shot volume ~ 250 µl, pressure < 2 bar), the experimentally measured dilution of fluorescence signal almost exactly matches the predicted dilution effects. Therefore, asymmetric 10:1 mixing offers a potential way to mix pre-equilibrated...
RNAP-promoter complexes in the large volume syringe with either buffer/competitor/NTPs in order to collect measurements of dissociation and promoter escape, while minimizing effects of dissociation due to dilution alone.

Dissociation and promoter escape in the absence of CarD

Asymmetric mixing performed with pre-mixed MboRNAP/promoter complexes in the F-syringe and buffer in the C-syringe dramatically reduced the level of signal change upon mixing compared to 1:1 mixing tests (Figure 5, black trace vs. Figure 1, blue trace). There is still a slight increase in signal over time, which suggests that even a dilution of ~9% is enough to cause dissociation and re-equilibration of MboRNAP/promoter complexes. However, the magnitude of this much smaller than what is observed for equal volume mixing, so this approach minimizes dilution-induced dissociation. In these experiments, I tested Heparin as a competitor for RNAP. Heparin is a highly-negatively charged molecule that binds RNAP with high affinity. It is considered a stronger competitor than dsDNA for RNAP because it can potentially “strip” RNAP from promoter-closed complexes. However, as a stronger competitor, the chances of RNAP re-binding the fluorescent promoter are less than if dsDNA were used as a trap.

Asymmetric mixing of MboRNAP/promoter DNA with 205 µg/ml and 818 µg/ml Heparin resulted in rapid decay of fluorescence over the course of 10 minutes (Figure 5, green and red traces). The rate of decay for 205 µg/ml and 818 µg/ml were very close, indicating that 205 µg/ml was a saturating amount of competitor under these reaction conditions.
Figure 5: Dissociation and promoter escape by MboRNAP in the absence of CarD. Using asymmetric 10:1 mixing with MboRNAP-promoter complexes in the larger volume syringe, traces were collect mixing the complexes with buffer (black) to control for dilution effects. Mixing MboRNAP-promoter complexes with Heparin at 205 µg/ml (green) and 818 µg/ml (red) provide information on dissociation and also possibly stripping of MboRNAP closed complexes. Mixing MboRNAP-promoter complexes with Heparin and NTP’s (purple trace) provides information on nucleotide incorporation and promoter escape, although aside from the initial lag it is difficult to distinguish this information from the Heparin traces. Mixing MboRNAP-promoter complexes with NTP’s in the absence of competitor (blue) provides information on nucleotide incorporation and promoter escape in multi-turnover conditions, because this non-monotonic and multiphasic curve is qualitatively discernable from the buffer control (black).
Figure 6: Promoter escape with CarD can be observed under single-turnover conditions. Pre-equilibrated CarD-MboRNAP-promoter complexes were mixed with buffer (black) as well as Heparin at 205 µg/ml and 818 µg/ml in order to study dissociation. CarD-MboRNAP-promoter complexes were mixed with Heparin and NTP’s at 45 µM (teal) and 455 µM (purple) to study nucleotide incorporation and promoter escape in single-turnover conditions.
Figure 7: Promoter escape with CarD can be observed under multi-turnover conditions. CarD-MboRNAP-promoter complexes were mixed with buffer (black), and NTP’s at 45 µM (yellow) and 455 µM (blue) in order to study nucleotide incorporation and promoter escape in multi-turnover conditions.
Table 1: Summary of observed rates of dissociation and promoter escape in the presence of CarD.

<table>
<thead>
<tr>
<th>ft</th>
<th>F</th>
<th>C</th>
<th>c</th>
<th>a1</th>
<th>k1</th>
<th>a2</th>
<th>k2</th>
<th>a3</th>
<th>k3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear 1-10s</td>
<td>Buffer</td>
<td>0.344692</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Linear 1-600s</td>
<td>DNA-Buffer</td>
<td>4.027980</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1exp 0.1-200s</td>
<td>DNA</td>
<td>4.601136</td>
<td>-0.19801</td>
<td>0.104399</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1exp 1-600s avg0</td>
<td>MbcRCD Buffer</td>
<td>1.061541</td>
<td>-0.09917</td>
<td>0.001996</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3exp 1-600s avg1</td>
<td>MbcRCD Hep205</td>
<td>0.2176</td>
<td>0.024648</td>
<td>2.014882</td>
<td>0.044291</td>
<td>0.05276</td>
<td>0.594767</td>
<td>0.032787</td>
<td></td>
</tr>
<tr>
<td>2exp 1-600s avg2</td>
<td>MbcRCD Hep118</td>
<td>0.217811</td>
<td>0.044291</td>
<td>0.05276</td>
<td>0.594767</td>
<td>0.032787</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2exp 1-600s avg3</td>
<td>MbcRCD Hep205/NTPH5</td>
<td>0.2112429</td>
<td>0.045244</td>
<td>0.05841</td>
<td>0.599269</td>
<td>0.03152</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1exp 1-600s avg4</td>
<td>MbcRCD Hep205/NTPH5</td>
<td>0.218531</td>
<td>0.498423</td>
<td>0.004418</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2exp 1-600s avg5</td>
<td>MbcRCD NTP45</td>
<td>0.2127923</td>
<td>0.501986</td>
<td>0.00429</td>
<td>0.027961</td>
<td>0.585136</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>avg4</td>
<td>MbcRCD NTP45</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>avg5</td>
<td>MbcRCD NTP495</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
I performed experiments using pre-formed \textit{Mbo}RNAP/promoter complexes mixed with 455 $\mu$M NTP’s in the presence and absence of a saturating amount of Heparin (205 $\mu$g/ml). In the absence of Heparin, I observed a trace that almost overlaid with the buffer trace, however this trace was multiphasic (Figure 5, blue). The trace increased in fluorescence intensity for the first 30-60 seconds, decreased over the next 30-60 seconds, and finally increased for the remaining 8 minutes of the trace. This NTP-dependent signal is qualitatively distinguishable from the buffer trace, which indicates that promoter escape can be measured in multi-turnover conditions. Mixing pre-formed \textit{Mbo}RNAP/promoter complexes with NTP’s and Heparin resulted in rapid decay of fluorescence signal almost matching traces collected with Heparin alone (Figure 5, purple). Interestingly, however, these traces include a lag before the decay. The observation of the NTP-dependent lag indicates that these traces can potentially provide information regarding nucleotide incorporation and promoter escape under single-turnover conditions.

\textit{Dissociation and promoter escape in the presence of CarD}

I performed experiments using pre-formed CarD-\textit{Mbo}RNAP-promoter complexes in single- and multi-turnover conditions to study the effect of CarD on transcription initiation (Summary of observed rates shown in Table 1). Single-turnover conditions were studied using Heparin at 205 and 818 $\mu$g/ml, which resulted in almost identical rates and extents of signal decay (Figure 6, red and green traces). These results indicated that under the conditions studied, 205 $\mu$g/ml Heparin has a saturating effect with regard to its ability to serve as a trap. Mixing CarD-\textit{Mbo}RNAP-promoter complexes with 205 $\mu$g/ml
Heparin and 45 µM NTPs resulted in a trace that was almost identical to those of Heparin alone (Figure 6, teal trace), indicating that an NTP-dependence could not be obviously discerned at this concentration and therefore the observed signal is most likely arising entirely from Heparin-dependent stripping/dissociation. However, when the NTP concentration was increased to 455 µM, an initial increase in fluorescence for the first few seconds was followed by a biphasic decay (Figure 6, pink trace) that was qualitatively faster than Heparin-only decays, indicating that under these conditions the NTP-dependent trace can be used to extract information on initial nucleotide incorporation and promoter escape.

I examined the effect of mixing RNAP/CarD/promoter with NTPs in the absence of competitor in order to study this process under multi-turnover conditions. At low concentrations of NTP (45 µM) I observed a non-monotonic multi-phasic trace that slightly increased for ~1 min and then slightly decayed over the remaining 9 minutes (Figure 7, yellow trace). Interestingly, the trace collected at 455 µM NTP was also non-monotonic, exhibiting a slight initial increase in fluorescence intensity for the first few seconds followed by a multiphasic decay for the remainder of the 10 minute reaction (Figure 7, blue trace). These results were both distinguishable from the buffer-only trace, indicating that this process can be studied under multi-turnover conditions. The transition from initial increase in signal to eventual decay seems to depend on [NTP] concentration, in the sense that the transition takes longer at 45 µM NTP (> 1 minute) compared to 455 NTP (< 10 seconds). A titration of NTPs under these conditions would likely provide a quantitative measure of how the time of this inflection point depends on NTP
concentration, and the amplitudes and rates of these non-monotonic multiphasic curves could potentially be used to provide kinetic constraints for a model of multi-round transcription initiation on the presence of CarD.

The small bump in the initial part of the NTP curves, seen in both single- and multi-turnover conditions, could be due to initial nucleotide incorporation. I explored this possibility by adding initial nucleotides to pre-equilibrated EcoRNAP/AP3 complexes as well as MboRNAP/CarD/AP3 complexes using manual mixing in the fluorometer. EcoRNAP traces showed NTP-dependent trends, however it was difficult to tell if MboRNAP transcription complexes changed fluorescence due to poor signal-to-noise (data not shown). I continued exploring the effect of initial nucleotide incorporation on fluorescence by performing asymmetric mixing experiments in the stopped-flow. For these experiments, I used pre-equilibrated MboRNAP/AP3 mixed with subsets of initiating nucleotides, including dinucleotides, and saw no increase in fluorescence (data not shown).

**Discussion and Future Directions**

These experiments demonstrate the utility of asymmetric mixing for studying RNAP-promoter dissociation and NTP-dependent processes such as nucleotide incorporation and promoter escape. It is expected that these experiments will provide a foundation for studying the processes in a more comprehensive and controlled manner so that we may learn about the effect of CarD and RbpA on all phases of transcription initiation.
One specific area of investigation that may prove fruitful is the comparison of dsDNA and Heparin as competitors. The fact that Heparin can potentially strip RNAP from closed complexes, while dsDNA cannot, could potentially be leveraged to use these competitors to probe closed-complex stability (4). In other words, if one were to mix RNAP/promoter complexes with a saturating amount of dsDNA and compare the rate of decay to a reaction in which RNAP/promoter complexes are mixed with a saturating amount of Heparin, the decay with Heparin should in principle be faster and the difference may be attributable to active stripping of RNAP from closed complexes. This information in turn could be used to probe open-complex stability, since unstable open-complexes spend more time in the strippable closed-complex state.

Perhaps the most useful way to advance this work is to titrate [NTP] to characterize the concentration-dependence of NTP-dependent rates (i.e. initial nucleotide incorporation and promoter escape). These kinetic signatures will allow the development of kinetic models that can be used to describe the effect of transcription factors on NTP-dependent phases of transcription initiation in both single- and multi-turnover conditions. In this manner, the model for rRNA transcription initiation can be expanded and the regulation of transcriptional flux arising from this operon can be understood at a mechanistic level. Furthermore, the same strategy can be generally applied to understand transcriptional flux at promoters throughout the mycobacterial genome.
References


Chapter 7

Potassium glutamate increases mycobacterial open complex stability
Introduction

The ionic environment of *Mtb* is dynamic and it experience high fluxes of chloride as part of the host immune response (1). To this point, *in vitro* studies of mycobacterial transcription initiation have primarily been performed in NaCl or KCl based reaction buffers (2-6). These studies provide an important link between *in vitro* and *in vivo* conditions. However, chloride is not always the dominant anion in bacterial transcription. In *E.coli* cells, the most physiologically relevant anion is not chloride but rather glutamate, which can vary between 30-260 mM depending on the media (7). Estimates suggest that intracellular glutamate concentration is even higher in *Mtb* than in *E.coli* (8). Studies have shown that at certain promoters, increasing concentrations of potassium glutamate have stabilized *Eco*RNAP-promoter interactions compared to KCl or NaCl (9). Interestingly, sensitivity to chloride vs glutamate is promoter specific and has been demonstrated for the *Eco rrnB* promoter, suggesting that anion species and concentration can regulate of rRNA transcription in *E.coli* (10).

Mycobacterial RPo stability at the *Mtb rrnAP3* promoter has been studied in chloride and glutamate buffers using a 3-nt transcription assay (11). This study demonstrated that CarD-regulated transcription from the *Mtb rrnAP3* promoter was sensitively inhibited from 10 mM < [Cl⁻] < 100 mM, where at 100 mM NaCl almost no transcription could be observed even in the presence of CarD. In contrast, the reaction was able to easily tolerate up to 150 mM [Glu⁻] without reduction in transcriptional activity, and a dramatic effect of CarD over RNAP could be seen (11).
I have investigated the effect of chloride versus glutamate on mycobacterial open complexes in the ensemble fluorescence assay by changing the buffer conditions from a “NaCl buffer” to a “KGlLu buffer” (final reaction buffer compositions below).

**NaCl buffer:** 14 mM Tris pH 8.0, **100 mM NaCl**, 10.2 mM MgCl₂, 4 uM ZnCl₂, 10% glycerol (v/v), 0.1 mg/ml BSA, 1 mM DTT

**KGlLu buffer:** 14 mM Tris pH 8.0, **40 mM NaCl**, 1.2 mM MgCl₂, **75 mM K-Glu**, 4 uM ZnCl₂, 10% glycerol (v/v), 0.1 mg/ml BSA, 1 mM DTT

**Methods**

Proteins were prepared as previously described, with notable exceptions as follows (5). RNAP core was dialyzed into a KGlLu storage buffer (10 mM Tris pH 8.0, 250 mM KGlLu, 1 mM MgCl₂, 20 µM ZnCl₂, 1 mM EDTA, 50% glycerol by volume) instead of the standard NaCl-based storage buffer. When designing the experiment, the “10X MTB +/-” buffer was prepared without any NaCl and 10-fold less MgCl₂. This resulted in a total [Cl⁻] = 40 mM in the final reaction conditions, compared to 100 mM in the NaCl-based reaction conditions. Importantly, this approach reduced the total reaction concentration of [Mg²⁺] from 10.2 mM in the NaCl buffer to 1.2 mM in the KGlLu buffer.
## Results and Conclusions

**KGlue increases** \( R_{P_o} \) **stability compared to NaCl**

Titrations of \( MboRNAP \) indicated that in KGlue buffer produced dramatically more equilibrium fluorescence compared to titrations of \( [MboRNAP] \) performed in NaCl (Fig 1A). A fit of the equilibrium binding isotherm yielded a \( K_{\text{eff}} \) of \( [RNAP] = 187 \text{ nM} \) (95% confidence intervals: \( A = 2.238 +/- 0.573, K = 186.9 +/- 107.67 \text{ nM} \), which is slightly less but within error of the NaCl \( K_{\text{eff}} \) of \( [RNAP] = 212 \text{ nM} \) (\( A = 0.3, K_{\text{eff}} = 212 +/- 43 \text{ nM} \)) (5). These results indicate that the increased fluorescence enhancement is not due to increased binding of RNAP to the promoter.

Possible explanations for the increased fluorescence fold change in KGlue compared to NaCl include (1) a photophysical effect of the buffer conditions on the quantum yield of the Cy3 (resulting in greater fluorescence from the same amount of open complex), or (2) an increased amount of open complex. To distinguish between these two possibilities, I performed the experiment using a fixed amount of \( [MboRNAP] = 240 \text{ nM} \) and including

```latex
\begin{tabular}{|c|c|c|c|c|c|c|c|c|}
\hline
& Tris 8.0 & NaCl & MgCl2 & KGlue & ZnCl2 & Glycerol & BSA & DTT & EDTA & (v/v) & (mg/ml) & (%)final \\
\hline
Core & 10 & 200 & 1 & 0 & 0.02 & 50 & 0 & 0 & 0 & 1 & 10 & 0.2 & 100 \\
SigA & 10 & 250 & 1 & 0 & 0.02 & 50 & 0 & 0 & 0 & 1 & 10 & 0.2 & 100 \\
CarD & 20 & 150 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 10 & 0.2 & 100 \\
10x MTB++ & 100 & 400 & 100 & 0 & 0 & 0 & 0 & 1 & 10 & 0 & 10 & 0.2 & 60 \\
MQH2O & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\hline
TOTAL & 14 & 100 & 10.2 & 0 & 0.004 & 5 & 0.1 & 1 & 0.2 & 100 & \\
\hline
\end{tabular}
```

```latex
\begin{tabular}{|c|c|c|c|c|c|c|c|c|}
\hline
& Tris 8.0 & NaCl & MgCl2 & KGlue & ZnCl2 & Glycerol & BSA & DTT & EDTA & (v/v) & (mg/ml) & (%)final \\
\hline
Core & 10 & 0 & 1 & 250 & 0.02 & 50 & 0 & 0 & 1 & 10 & 0.2 & 100 \\
SigA & 10 & 250 & 1 & 0 & 0.02 & 50 & 0 & 0 & 1 & 10 & 0.2 & 100 \\
CarD & 20 & 150 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 10 & 0.2 & 100 \\
10x MTB++ & 100 & 0 & 10 & 500 & 0 & 0 & 0 & 1 & 10 & 0 & 10 & 0.2 & 60 \\
MQH2O & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\hline
TOTAL & 14 & 40 & 1.2 & 75 & 0.004 & 5 & 0.1 & 1 & 0.2 & 100 & \\
\hline
\end{tabular}
```
CarD at 100 nM and 5 µM. If KGlu simply increased the quantum yield of Cy3 without affecting RP$_o$ stability, one would expect the same fold effect of CarD over RNAP that is observed in NaCl. I found that compared to NaCl, CarD has a lower fold effect over RNAP in K-Glu (22% in K-Glu versus ~600% in NaCl), and saturates at a lower concentration (100 nM, versus 1 µM in NaCl) (Figure 1A, Figure 3 bottom). These results are consistent with an increased basal amount of RP$_o$ for the following reasons. First, a higher basal RP$_o$ means that CarD would have a smaller relative effect on RP$_o$ stability. Second, we have shown that since CarD stabilizes RP$_o$ it must have a higher affinity for RP$_o$ than RP$_c$, therefore the half-maximal effect of CarD would decrease with increasing basal RP$_o$ stability.

**Increased RP$_o$ stability is due to changing the anionic species from Cl$^-$ to Glu$^-$**

Transcription assays performed using equimolar concentrations of KCl and KGlu indicated that addition of [Glu$^-$] to the NaCl buffer was not able to increase transcriptional activity as observed in a 3-nt assay. Reduction of [Cl$^-$] accompanied by addition of [Glu$^-$] were both required to observe increased transcriptional activity (results not shown). These results are consistent with an explanation for the increased observed fluorescence resulting from changing the primary anionic species of the buffer from Cl$^-$ to Glu$^-$.

**KGlu changes the kinetics of RP$_o$ formation**
Analysis of the kinetics indicated that traces collected in KGlu buffer fit well to a triple exponential, just like traces in NaCl. However, there were several important differences between the observed rates in KGlu compared to NaCl. The fastest phase ($k_{obs,1}$) was fit by fixing to a 1-1200 s double-exponential (as previously done for traces collected in NaCl buffer), but interestingly in KGlu $k_{obs,1}$ can be fit starting from 0.01 s (as opposed to 0.1 s for NaCl), since burst phase has a stronger signal (larger amplitude). All three phases contributed significantly to the fit (Figure 2).

The [RNAP]-dependence of the fastest phase ($k_{obs,1}$) could be reasonably approximated with a line ($y = 0.009327x + 2.355$, R-square = 0.9603) (Figure 3). This slope of this fit could potentially be reporting on a bimolecular rate of association of $\sim 10^7$ M$^{-1}$s$^{-1}$ (9.327 x $10^6$ M$^{-1}$s$^{-1}$). In this case, the off-rate can be estimated from y-intercept of linear fit: 2.355 ± 0.489 s$^{-1}$. If these numbers represented the actual rate constants, $K_d$ can be estimated from $k_{off}/k_{on} \sim 25$ nM. It is not surprising that this is so different than $K_{eff}$ from fit of the equilibrium titration (189 nM), since $K_{eff}$ also includes signal from isomerizations of the bound complex and not just burst phase of the traces. Interestingly, the intermediate ($k_{obs,2}$) and slow ($k_{obs,3}$) phases also exhibited a dependence on [RNAP] that could be approximated by a linear fit (Figure 3). It is unlikely that these observed rates are reporting on binding, since their slopes would yield unusually slow bimolecular rates of association ($9.2 \times 10^4$ M$^{-1}$s$^{-1}$ and $6.7 \times 10^4$ M$^{-1}$s$^{-1}$, for intermediate and slow phases, respectively). It possible that these observed rates have a hyperbolic dependence on [RNAP] which will only be observed at concentrations higher than what was tested. Interestingly, in all three cases the observed rate for 50 nM RNAP is faster than 100 nM
RNAP, which could potentially result from binding under conditions that deviate from the pseudo-first-order approximation (10-fold molar excess RNAP over 5 nM DNA is the lower limit for this approximation). Collectively, all of these kinetic signatures are in sharp contrast to titrations in NaCl, in which \( k_{\text{obs},3} \) did not depend on [RNAP], and fractional amplitude for \( k_{\text{obs},2} \) was too low to interpret.

*CarD has different qualitative effects on each of the 3 observed rates*

When comparing no CarD, low CarD (100 nM) and high CarD (5 µM), one can observe qualitatively different effects on each of the fast, intermediate and slow phases. With increasing CarD, \( k_{\text{obs},1} \) decreases, while \( k_{\text{obs},2} \) and \( k_{\text{obs},3} \) both increase (Figure 4). The net result of kinetic changes result in more RP\(_o\) and faster equilibration of RP\(_o\). However, interpreting the manner in which rate constants produce these observed effects will require detailed kinetic modeling.

**Discussion and Future Directions**

These experiments demonstrate that in the KGlu buffer, mycobacterial RP\(_o\) is more stable than in the NaCl buffer. This results in a smaller fold effect of CarD on open complex stability, and a lower \( K_{\text{eff}} \), and a faster overall approach to equilibrium.

The most logical next steps to utilize the data for future experiments are as follows:

1. **Control for cation species and concentrations.** When switching between NaCl and KGlu, KCl can be used to isolate the effect of the anion from effects that might be
associated with switching from Na\(^+\) to K\(^+\). Additionally, the concentration of Mg\(^{2+}\) was reduced from 10.2 mM in the NaCl buffer to 1.2 mM in the KGlu buffer. Mg\(^{2+}\) is required for enzymatic activity of transcription. Although Mg\(^{2+}\) is not required for RPo formation, controls should be performed to quantitatively test the dependence of mycobacterial RPo stability on concentration of Mg\(^{2+}\).

2. Choose a fixed [MboRNAP] for upcoming transcription factor titration experiments. Maximizing the dynamic range for CarD’s effect would be ideal, so try [MboRNAP] = 60 nM and perform a CarD titration. This combination will show a much larger effect of CarD, since the equilibrium fluorescence fold change of MboRNAP at 60 nM is lower (~0.4) than equilibrium fluorescence of MboRNAP at 240 nM (~1.3).

3. Perform titrations of CarD mutants and RbpA at this same fixed [MboRNAP]. These experiments will provide kinetic data that will allow direct comparison between transcription factors, in a manner similar to comparisons made in in NaCl-based buffers (6).

4. Consider how to interpret these curves since the amplitude for \(k_{\text{obs,3}}\) no longer dominates the fits. The data suggests that the mechanism for RPo-formation in KGlu is different than in NaCl-based reaction buffer. This could be due to (1) changes in rate between the same states, or (2) new intermediates that are formed in one buffer but not the other. The equilibrium fluorescence can still be interpreted as the relative concentration of RPo. An amplitude-averaged rate could potentially be used to quantify the overall approach to equilibrium as a crude measure of the overall reaction kinetics.
The implications of characterizing factor-dependent mechanisms of RP₀ stability (and all phases of transcription) in KGlu are numerous and of great value to this project. Depending on the energetics of these interactions, transcription factors may have a dramatically different effect on transcriptional flux in KGlu compared to NaCl. For instance, it is possible that in KGlu CarD and RbpA together hyperstabilize RP₀ at *Mtb* *rrnAP3* to the extent that promoter escape is inhibited, thus reducing the amount of full-length rRNA transcription. This is analogous to the effect of the *E.coli* transcription factor H-NS, which is able to stabilize RP₀ and increase transcription of abortive transcripts, but decreases overall transcription due to hyperstabilization of RP₀ (12). Considering the dynamic nature of the ionic environment experienced by *Mtb*, it is important to compare mycobacterial transcription under conditions where either chloride or glutamate are the dominant anion. By the same rational, studies performed at 37°C will be informative with regards to linking *in vitro* mechanisms to *in vivo* phenotypes.
Figure 1: Experiments performed in K glu buffer exhibit increased basal open-
complex stability.  (A) In the absence of transcription factors, titration of MboRNAP in 
K glu buffer saturates at a higher equilibrium level of fluorescence (~1.5 fold over DNA) 
compared to NaCl buffer (~0.4 fold over DNA, not shown) indicating a higher level of 
RPo stability in the absence of transcription factors.  (B) In K glu buffer, saturating CarD 
(purple and red traces) has a lower effect over RNAP alone (blue trace) compared to that 
in NaCl-based buffer.  CarD’s effect has already saturated at 100 nM (purple trace).  All 
data collected at 25°C.
Figure 2: Fractional amplitudes as a function of [MboRNAP]. Fractional amplitudes for the fast (a1, red), intermediate (a2, green) and slow (a3, red) phases are shown as a function of [MboRNAP].
Figure 3. Observed rates as a function of $[\text{MboRNAP}]$. Observed rates for the fast (top), intermediate (middle) and slow (bottom) phases are shown as a function of $[\text{MboRNAP}]$. If the lowest concentration were excluded (red x in each plot), each titration can be approximated to a line (blue, in each plot).
Figure 3: Kinetic measurements as a function of [CarD] in KGlu. (Top) Fractional amplitudes of the slow (a1, blue), intermediate (a2, green) and fast (a3, red) phases are shown as a function of [CarD]. (Bottom) Equilibrium fold change over DNA is shown for CarD = 0, 100, and 5000 nM.
Figure 4: Observed rate dependence on [CarD]. Observed rates for the fast (top), intermediate (middle) and slow (bottom) phases are shown as a function of [CarD].
References


Chapter 8

Single Molecule Magnetic Tweezers for Real-Time Detection of Mycobacterial Promoter Unwinding
**Introduction**

Transcription initiation involves RNA polymerase (RNAP)-promoter binding, promoter-opening during which 10-12 bases of DNA are unwound, nucleotide incorporation and promoter escape. Understanding the mechanisms of regulating transcription initiation in *Mycobacterium tuberculosis* (*Mtb*) represents a clinically relevant challenge to biomedical research. CarD and RbpA are essential transcription factors that are known to regulate mycobacterial transcription initiation at the *Mtb* *rrnA*P3 promoter, which is the major promoter for ribosomal RNA (rRNA) in *Mtb* (1-3). We have shown that CarD and RbpA stabilize the RNAP-promoter open complex using an ensemble fluorescence assay (4, 5). These studies have provided valuable information about the thermodynamics and kinetics of factor-dependent regulation of mycobacterial transcription initiation. However, ensemble measurements arise from the average behavior of many molecules, and the stochastic and unsynchronized transitions that underlie molecular mechanism can only be accessed indirectly in these ensembles (6).

Single-molecule magnetic-tweezers is an established approach to measure all phases of bacterial transcription initiation in real-time (6-8). In addition, the assay allows for the study of topological variables such as super-helicity, which are known to have dramatic effects on the process of transcription. Briefly, individual DNA tethers are attached at one end to a flow cell surface and at the other end to a superparamagnetic bead. A magnet that is positioned above the surface can be used to control the tension in the DNA tether by adjusting its height above the surface. Furthermore, for nick-free torsionally constrained DNA, supercoils can be introduced in the DNA tether by rotating the magnet.
The height of the bead above the surface is monitored using video microscopy. On supercoiled DNA, promoter unwinding by RNAP results in a detectable change in bead height above the surface (6-8).

This approach offers several strengths that are complementary to ensemble fluorescence studies of promoter-opening. First, the signal arises directly from the nanomechanical unwinding of DNA, and therefore interpretation of the signal is unambiguous and does not depend on the oft-mysterious photophysical basis of a fluorophore’s quantum yield. Second, the assay is performed on supercoiled DNA, which is more representative of the in vivo state of DNA and can have a dramatic effect on the ability of RNAP to unwind promoter DNA (negatively supercoiled DNA favors unwinding, while positively supercoiled DNA inhibits unwinding). Third, the relationship between rotation and bead height enables detection of promoter opening with near-bp resolution (7). This offers the possibility to observe important transcription initiation intermediates such as DNA “scrunching”, which involve bubbles larger than the initial 10-12 basepairs that are typically unwound in the RNAP-promoter open complex (7, 8). Scrunching can have dramatic effects on transcriptional flux, especially at rRNA promoters, and it is unknown whether this process occurs in mycobacteria.

There are also limitations to using single-molecule magnetic-tweezers to study transcription initiation. First, the response time of the bead’s movement is limited by hydrodynamic drag, which results in a time-resolution of the assay on the order of ~ 1 second. This is useful for observing states with lifetimes longer than 1 second, but it is
Figure 1: Single molecule magnetic tweezers for real-time detection of promoter unwinding (adapted from Revyakin et al.). Rotation-extension curve shows how DNA extension is related to supercoiling. Linear regimes on the positive and negative sides allow conversion between changes in bead height and loss of twist. Promoter unwinding by RNAP removes 1 twist of DNA, which results in a gain of 1 writhe due to conservation of linking number. Example data traces show raw data (green) and averaged data (red) for promoter opening events on positively supercoiled DNA by EcoRNAP.
challenging to use the assay to detect transient opening events that last less than 1 second (i.e. from unstable open complexes). Secondly, magnetic tweezers (and single molecule force spectroscopy approaches in general) suffer from low throughput, since it is difficult to measure many individual molecules simultaneously. Approaches exist for microcontact printing of surfaces in order to enable multiplexing, but these methods are still experimental and under development (9). As a result, the assay is most useful when it can be reproducibly used to collect many measurements of individual molecules, and for measuring states that have lifetimes greater than 1 second.

I have used this assay to study mycobacterial promoter binding and opening at the Mtb rrnAP3 promoter using MboRNAP and MtbCarD. I have observed stable promoter opening in the presence of CarD on positively supercoiled DNA templates. The extent of promoter unwinding was larger than expected which could be due to (1) a larger transcription bubble in the open complex, (2) scrunching, or (3) some combination of the two. These measurements represent the first single molecule observations of mycobacterial transcription initiation.

Methods

DNA tether construction

4 kb DNA tether fragments containing the Mtb rrnAP3 promoter were prepared as previously described. Briefly, PCR was used to amplify a 2 kb “core” fragment that contained the promoter as well as unique restriction sites at either end. Two 1 kb “handles” were prepared by PCR-amplification using either digoxigenenin- or biotin-
modified dUTP. The core and handles were digested and ligated to form a 4-kb DNA tether with unique 1 kb handles. The handles were used to attach the tether to anti-digoxigenin (flow cell surface) and streptavidin (1 μm magnetic bead).

**Flow cell preparation**

Several methods have been used to construct flow cells with surface chemistry that is amenable to DNA tether attachment, yet also passivated to sufficiently prevent non-specific adsorption of molecules to the glass surface. Perhaps the most frequent strategy employs modified PEG surfaces (10). I used a nitrocellulose-coated flow cell decorated with anti-digoxigenin and passivated with BSA. I found that this surface chemistry could more reproducibly allow analysis of many torsionally constrained tethers compared to the alternate PEG surface chemistry.

**Results and Conclusions**

*Collection of symmetric rotation-extension curves on torsionally constrained DNA*

DNA tethers containing the *Mtb* *rrnAP3* promoter were constructed with 1 kb handles and a 2 kb stretch of DNA between the flow-cell surface and a magnetic bead. Force-calibrations indicated that at a magnet height of ~2.5 mm produced a force of approximately 0.3 pN in the DNA tether. A typical rotation extension curve collected at this force is shown in Figure 1. The symmetric curve is centered around 21 turns (arbitrary units). Subtraction of 10 turns (from 21 to 11) results in a decrease in bead height from 0.48 um to 0.15 um, which is due to the incorporation of negative supercoils in the DNA. Likewise, addition of 10 turns (from 21 to 31 on the x-axis) results in a
Figure 2: Rotation-extension curve for torsionally constrained DNA containing the *Mtb* *rrnAP3* promoter. Relaxed DNA is centered at approximately 21 turns (arbitrary units). Removal of 10 turns (21 to 11 turns) results in a decrease in bead height as negative supercoils are introduced. Likewise, addition of 10 turns (21 to 31 turns) results in an increase in bead height as positive supercoils are introduced. The slope of each side in the linear regime is 50 nm/turn.
decrease in bead height from 0.48 µm to 0.11 µm, which is due to the incorporation of positive supercoils in the DNA. The linear regime on the negative supercoiled side has a slope of 50 nm/turn, and the linear regime on the positive supercoiled side also has slope of 50 nm/turn. Thus, removal of 1 twist of DNA (i.e. due to promoter-unwinding) would result in a gain of 1 writhe of DNA that would be equivalent to a 50 nm change. On the negatively supercoiled side, this would result in the bead moving up by 50 nm, and on the positively supercoiled side this would result in the bead moving down by 50 nm.

_DNA-only traces collected in the absence of protein are steady over the course of minutes_ As a negative control, data was collected for relaxed, positively supercoiled and negatively-supercoiled DNA in the absence of any protein. An example trace is shown in Figure 2. For this particular tether, the height of the bead above the surface is 0.53 µm, and the addition or removal of 6 supercoils results in a bead height of approximately 0.45 µm. Over the course of minutes, the supercoiled DNA has a peak-to-peak noise of approximately 60 nm.

_Positively supercoiled DNA shows increased noise in the presence of MboRNAP_ When _MboRNAP_ was introduced into the flow cell at either 2 nM (Figure 4) or 20 nM, the peak-to-peak noise observed over the course of minutes for positively supercoiled DNA approximately doubled that of DNA alone (~100 nm). However, the height of the bead appeared to be relatively constant on both positively and negatively supercoiled DNA, and the height of the bead appeared to be relatively symmetric. It is possible that increased noise on the positively-supercoiled DNA is the result of _MboRNAP_ binding the
Figure 3: Signal for DNA only with positive and negative supercoils. Raw data (gray) and averaged data (red) are shown for a trace of DNA in the absence of protein. The trace starts out with relaxed DNA. 6 positive turns (going from 0 to +6 supercoils) were introduced at ~100 s, and 12 negative turns (going from +6 to -6 supercoils) were introduced at ~ 250 s. At 400 s, 6 positive turns were introduced (going from -6 to 0 supercoils).
Figure 4: Single molecule trace for MboRNAP on positively and negatively supercoiled DNA. Raw data (gray) and averaged data (red) are shown for a single DNA tether in the presence of 2 nM MboRNAP. At ~100 s, +5 turns are introduced (going from 0 to +5 supercoils). Some possible transient activity can be seen from 100 to 750 s. At ~750 s, -10 turns were introduced (going from +5 to -5 supercoils). At ~1350 s, +5 turns were introduced (going from -5 to 0 supercoils).
promoter and forming unstable open-complexes that do not persist long enough to be detected by this assay.

**Positively supercoiled DNA shows stable promoter opening in the presence of MboRNAP and CarD**

The introduction of MboRNAP (2 nM or 20 nM) with CarD (1 µM) resulted in clear hopping between two states for positively-supercoiled DNA (Figure 5). These states were qualitatively easily distinguishable by eye and differed in bead height by approximately 70 nm (Figure 6). The lifetime of these states lasted approximately tens of seconds to minutes (Figure 8). Mapping the change in height of the bead to the rotation-extension curve predicts that if the entire change in bead height were due to unwinding, then approximately 1.5 turns of DNA were being unwound (Figure 7). Alternately, the observed change in bead height could be the result of a combination of unwinding and wrapping. Measurements of promoter unwinding on negatively supercoiled DNA would allow for this distinction, however no changes in bead height were observed on negatively supercoiled DNA.

Puzzlingly, the observed change in bead height appears to increase from the baseline position on the positively supercoiled DNA, which is the opposite direction from the predicted direction. The same phenomenon was observed for multiple experiments performed on three different days. One remote possibility is that the DNA was not B-form (right-handed double-helix) but rather Z-form (left-handed double-helix), which would have resulted in the bead moving up instead of down upon promoter unwinding of
Figure 5. Observation of promoter-opening in the presence of CarD. Raw data (gray) and averaged data (red) are shown for traces collected with 1 µM CarD and MboRNAP. (A) For 2 nM MboRNAP, the trace begins with relaxed DNA and at ~600 s +5 turns were introduced (going from 0 to positive 5 supercoils). From 600 – 1200s, the signal hops between 2 clearly distinguishable states, which is presumably the result of CarD-mediated promoter unwinding by MboRNAP. At 1200 s, -10 turns were introduced (going from +5 to -5 supercoils). No activity is observed, and at ~1800s +5 turns were introduced to return the DNA back to 0 supercoils. (B) For 20 nM MboRNAP, +6 supercoils were introduced at the beginning of the trace and hopping between two clearly distinguishable states is observed until ~800 s, when -12 turns were introduced (going from +6 to -6 supercoils). A transient drop in signal is observed followed by a spike, and then no further activity is observed until +6 turns were introduced at ~1500 s (going from -6 to 0 supercoils).
Figure 6: DNA extension analysis. Plotting a histogram of DNA extension states over time for 20 nM $Mbo$RNAP and 1 µM CarD shows a clear bimodal distribution, which is indicative of 2 states. These states are centered around DNA extensions of approximately 0.35 µm and 0.42 µm.
Figure 7: Unwinding analysis. For 20 nM MboRNAP and 1 µM CarD, DNA extension can be mapped onto the rotation-extension curve in order to measure the number of turns that would lead to the observed change in bead height. In this case, the change of ~70 nm is approximately equal to 1.5 turns of DNA.
Figure 8: Dwell-time distribution analysis. The amount of time spent in the “lower” state (A) and the “higher” state (B) are shown as histograms. Building up such histograms can provide information on the relaxation times of each state.
positively supercoiled DNA. To confirm rotational direction before every experiment, I recommend that rotation-extension curves on DNA only should be collected at high force in order to observe asymmetry on the negative supercoiling side.

**Discussion and Future Directions**

These results are the first single molecule observations of mycobacterial promoter unwinding by single molecule methods. Stable promoter unwinding was only observed in the presence of CarD and only on what is assumed to be positively supercoiled DNA. Lifetimes of the states lasted tens of seconds to minutes, and the change in bead height (~70 nm) predicts ~1.5 turns of unwound DNA. Previous measurements of promoter unwinding by *Eco*RNAP suggest that closed to 1 turn of B-form DNA is unwound by RNAP. The observation that more than 1 turn of DNA is being unwound raises 2 possibilities. First, the bubble size formed by *Mbo*RNAP/CarD could be bigger than that formed by *Eco*RNAP on this same promoter. Potassium permanganate footprinting studies of *Mbo*RNAP +/- CarD compared to *Eco*RNAP on the *Mtb* *rrnA*P3 promoter predicted that the bubble sizes are the same, however the resolution of this technique is limited to single-stranded thymines and it is possible that a larger bubble formed by *Mbo*RNAP/CarD would not be detectable in that assay.

The potential observation of 1.5 turns of unwound DNA also raises the possibility that scrunching is taking place in the absence of NTP’s, which has recently been reported for *Eco*RNAP (11). This is an especially intriguing possibility, since mycobacterial RNAP may use scrunching as a mechanism to store energy for increased efficiency of promoter
escape from ribosomal promoters which are known to have unstable open-complexes (12).

Future directions for this project would be greatly enhanced by observation of promoter unwinding in the absence of transcription factors, so that factor-dependent promoter unwinding can be compared to the basal system. Possible ways to increase the stability of RP₀ for MboRNAP acting along include changing the buffer (Glu⁻ instead of Cl⁻) (see Chapter 7) and increasing the temperature form 25 C to 37 C (4).

In the event that the magnetic tweezers approach proves to be low throughput or not reproducible, an attractive alternative single molecule approach is TIRF microscopy. This approach could be used with a single fluorophore or with multiple fluorophores (13-15). A single Cy3 fluorophore has been used at the -4 nontemplate promoter position to detect promoter opening in real time for EcoRNAP. These studies found that fluorescence lifetime provided better signal-to-noise than fluorescence intensity, and this signal could be used to detect not only promoter binding and opening but also DNA scrunching (13). The use of multiple fluorophores can be used for colocalization single molecule spectroscopy (CoSMoS) or for fluorescence resonance energy transfer (FRET), and these approaches can be used to measure all phases of transcription initiation including binding, opening, scrunching, abortive initiation and promoter escape (14, 15).
Table 1: Single molecule magnetic tweezers file index. This spreadsheet details different experiments in which promoter opening was observed, and includes information on [MboRNAP], [CarD], date, raw file information, supercoiling, and number of events per trace.
References


Chapter 9

Summary
and
Recommendations for Future Work
Summary of research

When this research was initiated in 2012, CarD was recently discovered as an essential transcription factor in *Mycobacterium tuberculosis* (*Mtb*) important for global regulation of transcription (1). CarD was known to bind RNAP and DNA, and was found at promoters throughout the mycobacterial genome including both promoters in the *Mtb* rRNA operon known as *rrnA*P1 and *rrnA*P3. RbpA was also a recently discovered transcription factor that is essential in *Mtb* (2, 3). CarD and RbpA were each known to activate transcription from the *rrnA*P3 promoter, but the mechanisms by which each factor regulated transcriptional flux were unknown (4, 3). All phases of bacterial transcription initiation were possible opportunities for factor-dependent regulation including RNAP-promoter binding, promoter opening, nucleotide incorporation and promoter escape.

I have studied the effect of CarD on all phases of transcription initiation. Using EMSA and double-filter binding assays, I found that CarD increased the apparent affinity of *Mbo*RNAP for promoter DNA. I found that this positive linkage was dependent not only on holoenzyme formation, but also on specific residues within SigA (FYWW), which are predicted to be important for promoter opening. Using ensemble fluorescence, I studied the effect of CarD on RNAP-promoter binding and opening and found that CarD stabilizes mycobacterial transcription open complexes using a concentration-dependent two-tiered kinetic mechanism (5). Experiments performed with CarD point mutants demonstrated that the ability of CarD to stabilize RPo depends on all three of its functional domains (RID, DNA-binding, and tryptophan). The data is consistent with a
model in which CarD stabilizes $RP_0$ by conformational selection (at low concentrations of CarD) and induced fit (at high concentrations of CarD) acting in parallel (5). I established an asymmetric mixing approach to observe transcription in the presence of competitors and NTP’s, so that this work can be extended to measure the effect of transcription factors on nucleotide incorporation and promoter escape. Considering the dynamic buffer environment experienced by $Mtb$ in vivo, I performed ensemble fluorescence experiments in a glutamate-based buffer and found that $RP_0$ is more stable when compared to chloride-based buffers (6). To study the system in more detail, I collected single-molecule magnetic-tweezers traces to measure the size of the transcription bubble in real time (7). These experiments indicated that stable promoter opening could only be observed in the presence of CarD, and estimates of the size of the transcription bubble suggest that the possibility that 1.5 turns of DNA are being unwound. These experiments have established a working model to explain the effect of CarD on mycobacterial transcription initiation kinetics, and will provide a foundation for a comprehensive description of CarD’s mechanisms of transcription regulation.

I have also studied a second transcription factor RbpA, which was also known to increase transcription from $rrnAP3$ and stabilize open-complexes at some mycobacterial promoters. Using ensemble fluorescence, I found that RbpA stabilizes the $rrnAP3$ promoter using a kinetic mechanism distinct from that of CarD (8). These results are consistent with a model in which RbpA utilizes conformational selection and induced fit in series to stabilize $RP_0$ (Chapter 5). Experiments performed with both transcription factors revealed positive linkage between RbpA and CarD with regard to mycobacterial open complex stabilization, demonstrating that the two transcription factors work
together at the *rrnA*P3 promoter. The cooperative stabilization of *rrnA*P3 open complexes led to increased fluorescence and faster kinetics, and resembled time courses resulting from *Eco*RNAP on the *rrn*AP3 promoter. Taken together, these data demonstrate that RbpA and CarD dramatically increase the amount of open complex and accelerate the approach to open complex equilibrium compared to either factor acting alone (8).

**Figure 1. Summary of CarD and RbpA regulation of mycobacterial transcription initiation in NaCl-based buffer.** (Top) General kinetic mechanism in the absence of factors includes an intermediate (RPi) in between closed and open complexes. Transcription factors (X) could potentially interact with any of these states (dashed box). (Bottom) Free energy diagram. In the absence of factors (black), opening is rate-limiting. In the presence of low concentrations of CarD (short blue dashed line), all promoter-bound states are stabilized, and RPo is further stabilizes relative to other states compared to no factor. In the presence of high concentrations of CarD (long blue dashed line), the activation barrier to opening is lowered leading to acceleration of open complex equilibration. In the presence of RbpA (dashed red line), all states are stabilized relative to no factor and the activation barrier to opening is lowered. In the presence of both factors (green dashed line), all states are stabilized and all activation barriers are lowered. In each case, the end of each curve qualitatively represents the amount of transcriptional flux, where lower free energy represents more transcription.
Recommendations for future work

Characterization of the basal mechanism of mycobacterial transcription initiation

Mechanistic studies of transcription factors in *E. coli* have been greatly facilitated by knowledge of the kinetic mechanism of transcription initiation in the absence of factors (9-12). Characterization of multiple closed and open intermediate states as well as the rates of conversion between them has required decades of study, and new states continue to be revealed (13). The basal mechanism of mycobacterial transcription initiation is often assumed to take place through similar mechanisms, however the work presented here demonstrates that there are fundamental and dramatic differences between how *Eco*RNAP and *Mbo*RNAP initiate transcription from the same promoter. Therefore, the assumption that the kinetic mechanism of mycobacterial transcription includes the same intermediates as *E. coli* should be used with caution. Studies presented here suggest the presence of at least one intermediate in between closed and open complexes. Detailing the structural properties of such intermediates will allow mechanisms of transcription factors to be studied with greater clarity.

Solution-based binding assays

I have studied binding interactions using EMSA and double-filter binding, but these approaches provide measurements under non-equilibrium conditions and thus may not be representative of true solution binding equilibria (14). Furthermore, they cannot be used to measure the kinetics of binding, which is often diffusion limited and beyond the time-resolution of manual mixing approaches. The development of solution-based binding assays would greatly aid this research because such assays allow for one to control for
solution conditions such as pH, salt and temperature. Given the dynamic solution environment experienced by *Mtb in vivo*, the effects of chloride and glutamate should specifically be tested so that the effects of anion concentration on transcription complex assembly can be characterized (6). Furthermore, solution-based binding assays can be used to measure kinetics of association and dissociation in a controlled manner. Direct measurements of the rates of binding will provide additional constraints to kinetic models.

Examples of especially valuable binding measurements to establish include sigma-saturation of core RNAP, holoenzyme saturation of promoter DNA, and affinities of CarD and RbpA to all combinations of RNAP and DNA. Measurements of the oligomeric states of the proteins should also be performed, so that analysis of data can take into account the stoichiometry of interactions. This is important because studies of both CarD and RbpA have been reported as dimeric structures (15, 2). Fluorescence anisotropy and fluorescence correlation spectroscopy are particularly attractive approaches, since they could potentially be used to probe the affinity of transcription factors to RPo using holoenzyme-saturated bubble DNA templates.
Fluorescence studies of transcription initiation

There are many useful controls that can be performed in order to aid interpretation of the fluorescence time courses presented in this work. The use of the \textit{MboRNAP-FYWW} (Chapter 2) would enable measurements of fluorescence enhancement resulting from a holoenzyme expected to be deficient in promoter opening. Measurement of any fluorescence enhancement arising from closed complexes may enable global fitting of time courses, which has thus far proven elusive (Chapter 5). 2-Aminopurine is an alternative fluorophore that can also be used to measure promoter opening in real-time, although the incorporation of multiple fluorophores is often necessary to obtain sufficient signal-to-noise (16). Thus, these studies must be interpreted carefully because they involve the perturbation of multiple bases in the promoter sequence. Additionally, fluorophores could be conjugated to downstream regions of DNA templates to study the arrival of RNAP during transcription (17). Dependence of these signals on NTP concentration and distance to fluorophore will provide information on the kinetics of nucleotide incorporation, promoter escape and elongation.

Quench flow studies for transcription initiation

Many techniques can elucidate intermediates on pathway to productive mycobacterial transcription initiation, but of particular interest are those techniques that leverage a quench flow apparatus. This rapid mixing technique is able to provide millisecond time resolution for footprinting techniques that can measure binding and opening (DNAseI, KMnO$_4$, hydroxyl radical) as well as time-resolved production of RNA transcripts in the presence of NTPs (18). The primary advantage of such an approach is that the signal
provides a complementary measurement to fluorescence-based studies of transcription initiation.

**Computational approaches for global evaluation of kinetic schemes**

Global exploration of rate space is essential for the development of mechanistic models that can account for experimental data. Developing a global fitting program (i.e. through the use of published Mathematica notebooks) will greatly facilitate exploration of many kinetic models over a wide range of rate space (19). This will be useful not only for examining the relationship between eigenvalues and observed rates (as described in Chapter 5), but also for examining the relationship between eigenvectors and amplitudes, which are thus far an unused and valuable constraint (14). Furthermore, programs such as KinTek Explorer should be used to simulate time courses (20-22). In principle, a kinetic model that sufficiently captures the mechanism can be used to reproduce experimental time courses, observed rates, and amplitudes. This remains the goal of future computational work for evaluation of kinetic schemes.

**Promoter-dependence of transcription factor mechanisms**

Studying *rrnAP1* by itself as well as together with *rrnAP3* will provide great insight to how the rRNA operon works in *Mtb*. Preliminary data indicates that RNAP-promoter complexes at *rrnAP3* have a longer half-life in the presence of competitor when *rrnAP1* is present upstream (as in the genomic sequence of the *Mtb* rRNA operon), as opposed to when *rrnAP3* is present alone. These results indicate that there is a link between *rrnAP1* and RNAP’s affinity for *rrnAP3*. Studies of lac repressor and T7 RNAP indicate that DNA-binding proteins can affect RNAP-promoter affinity at a distance, suggesting the
possibility that RNAP-AP1 interactions could be increasing RNAP-AP3 affinity through the DNA. This effect could be dependent on the spacing of the two binding sites, which could be tested by creating templates with varied spacing between the two promoter sites (23).

Since CarD and RbpA are known to affect many promoters throughout the mycobacterial genome, the development of an \textit{in vitro} transcription assay that utilizes purified mycobacterial genomic DNA offers a way to probe the effect of these transcription factors when many promoters are present. To further bridge the gap between \textit{in vitro} and \textit{in vivo} phenotypes, fractionated cell lysate could be used to test transcription factor effects on genomic DNA. Such experiments offer the potential to elucidate pleiotropic effects of these transcription factors that may involves molecules that are not as yet known (1).

\textit{Single-molecule studies of mycobacterial transcription initiation}

Single-molecule magnetic-tweezers can be used to measure promoter unwinding with near-base-pair resolution, and therefore represent a powerful way to probe the potential effects of transcription factors during transcription initiation. Of particular interest is the possibility of “scunching” during transcription initiation, which can take place in \textit{E.coli} (in the presence and absence of initiating nucleotides) but has not yet been demonstrated in the mycobacterial transcription system (24, 13). The mechanistic implications of DNA scrunching may be linked to efficiency of promoter escape, especially from ribosomal promoters (13). However, some measurements suggest that the energy stored in scrunched DNA is modest and would not have a major effect on promoter escape (17).
Nevertheless, such measurements may be dependent on species and promoter and therefore scrunching intermediates may play a major role in mycobacterial transcription initiation in the presence of regulators.

Single molecule TIRF can be used with multiple fluorophores to probe the effects of multiple factors acting simultaneously during transcription initiation (25). Such an approach could be used with labeled DNA and proteins in order to study the energetics of transcription initiation in the presence of many factors under a range of conditions.

*Links between transcription factor mechanisms and treatments for TB*

Weakening CarD’s interactions with RNAP led to increased efficacy of Rifampicin, suggesting a direct link between transcription factor mechanism and treatment of Tuberculosis (26). Many treatment strategies of Tuberculosis are aimed at preventing growth, but these strategies are less effective at treating latent Tuberculosis. Thus, the development of small molecules that target CarD and RbpA offers a way forward for treating this disease. The development of the genome-editing technology of CRISPR-Cas9 has advanced to the point where it can be reliably used in the mycobacterial system (27, 28). Furthermore, nanoparticle delivery systems can be developed to introduce the CRISPR system to *Mtb* within a granuloma (29). Therefore, CRISPR targeted to the genes of CarD and RbpA could soon be a new way to eliminate these essential genes *in vivo*, and in doing so kill the bacterium.
Conclusion

The work presented here represents the first thermodynamic and kinetic measurements of two transcription factors essential in *Mycobacterium tuberculosis*. The working models, recommendations, and hypotheses that have emerged from this dissertation will serve to inform future studies of these factors not only at the *Mtb* *rrnAP3* promoter but also throughout the mycobacterial genome. These data are tangible steps towards understanding the complex mechanisms of gene regulation in a deadly pathogenic bacterium that continues to mystify and intrigue biomedical researchers around the world. I trust that future researchers will look back on this work with a feeling of justified superiority.
References


Appendix A

Sample preparation of Cy3-labeled promoter DNA
This protocol was originally taught to me in March 2013 by Tomasz Heyduk and a member of his laboratory, Agnieszka Lass-Napiorkowska. I adapted the protocol as needed in order to account for differences in equipment between Saint Louis University and the Department of Biochemistry and Molecular Biophysics at Washington University in St. Louis. Any differences between the two protocols are minor and have been noted.

**Oligo design**

The nontemplate oligo is 85 nt long from -80 to +5, and the T at +2 is modified with an amine attached to a C6 linker. In IDT, when typing the oligo into the field from 5' to 3', the T is replaced by /iAmMC6T/ like this: 5'-...G/iAmMC6T/TTGC-3'

For the modification, you can either type it in manually, or insert an "Internal Mod" using the tools on the website, under the category "Attachment Chemistry / Linkers". This is code for "Int Amino Modifier C6 dT", which will be used to conjugate a Cy3-NHS Ester (Lumiprobe Cat#21020) to the base. The modified base not be counted as a base, so the length will be displayed as 84 bases even though it's really 84 + 1 modified base = 85.

The template oligo is also 85 nt long and overlaps the nontemplate oligo by 20bp, and extends downstream. So, it is typically from -15 to +60. Remember to enter both oligos 5'-3'.

Using IDT's Oligo Analyzer Tool, each oligo should be checked for secondary structure propensity (Hairpin) as well as self-assembly (Self-Dimer) and any structures with (delta)G < -20 kcal/mol should be considered disruptive to the prep. Lastly, both oligos should be checked together for their annealing stability (Hetero-Dimer) to make sure the 20 bp overlap is the lowest-energy hetero-dimer between the two primers.

Order both oligos on a scale of 100 nmol, which is enough for several preps of fluorescent promoter. The default purification is PAGE, but manually select HPLC instead because it is better.
Promoter preparation

GENERAL TIMELINE FOR PREP
Day 1 – Label amine-modified oligo with Cy3-NHS-Ester overnight
Day 2 – Ethanol precipitate oligos, HPLC to purify labeled oligo from unlabeled oligo
Day 3 – Extension reaction and FPLC to purify double-stranded product

DAY 1: Label amine-modified non-template strand (NTS) oligo with Cy3-NHS

1. Measure concentration of oligo. Oligos typically arrive dry from IDT.

Resuspend amine-oligo in 50 ul freshly-prepared 0.1 M NaHCO3 (sodium bicarbonate) on the same day that it will be labeled overnight. Measure Abs(260) by diluting 1 ul of the sample ~100 fold. This material is not recovered for further use.

The template strand pellet can remain dry until Day 3, at which point it can be resuspended in 50 ul MQ H2O (or Buffer EB) prior to its use in the extension reaction.

2. Calculate molar amount of amine-oligo.

A “100 nmol” order usually produces around 10 nmol of oligo at most. This is because the scale of the order does not refer to the final product, it refers to the amount of phosphoramidites that are used in the synthesis process.

3. Prepare a 10x molar excess of Cy3-NHS-ester for labeling.

Usually, Cy3-NHS-ester arrives dry, is resuspended in DMSO, aliquoted such that each tube contains 100 nmoles of dye, and frozen at -80C. These aliquots are made assuming that they will be combined with 10 nmoles of amine-oligo (10x molar excess dye over oligo). It is fine to have greater than a 10-fold molar excess dye over oligo.

4. Add the DNA in 0.1 M NaHCO3 onto an aliquot of the dry Cy3.

Gently pipette up and down to mix. You may need to use the pipette tip to gently dislodge the pellet of dye, but do not let the dye get stuck in the hole of the pipette tip otherwise you will risk removing material from the reaction. Let the DNA/Cy3 mix remain at room temperature overnight protected from light (covered in foil or in a drawer is fine).
DAY 2: Purification of Cy3-oligo

Ethanol precipitate Cy3-oligo

1. Add 2 M NaCl to DNA-Cy3 mix to a [NaCl]_{final} = 0.25 M. Add 2.5x volume 100% EtOH from -20C.

Example:
50 ul DNA + 7.15 ul of 2 M NaCl = 57.15 ul
57.15 ul x 2.5 = 142.9 ul EtOH
142.9 ul EtOH + 57.15 ul = 200.05 ul total

2. Balance the reaction tube with another tubes and spin at 15,000 rpm for 30 min at +4C.

3. Remove and save supernatant using a protein-loading tip without disturbing pellet. This is a precautionary measure in case the pellet accidentally gets removed with the supernatant.

4. Wash with 300 ul 70% EtOH (-20C), add gently to side of tube by the pellet. Remove and save the 70% EtOH, leaving the pellet behind. Once again, saving the wash is a precautionary measure in case the pellet gets removed with the wash.

5. A Vacufuge can be used to evaporate the residual ethanol, 5 min should be enough. Be careful when removing tubes from the centrifuge, pellet may have dislodged from wall of tube. Alternately, you can place the eppendorf tube with the cap open in heat block (set to at least 60 C) and wait a few minutes for residual ethanol to evaporate.

6. Resuspend the Cy3-NTS pellet in 50 ul H2O, pipette up and down until you see a uniform pink color. Alternately, the pellet can be resuspended in Buffer A as described below.

Reverse-phase HPLC-purification of Cy3-labeled oligo from unlabeled oligo

Column: XTerra C18
Reference: Thang Ho, Lohman Lab

C18 Column storage conditions:
40% H2O
60% CH₃CN

HPLC buffers:
<table>
<thead>
<tr>
<th>Component</th>
<th>Buffer A (1 L)</th>
<th>Buffer B (1 L)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TEAA (0.1 M stock)</td>
<td>[final] = 0.098 M</td>
<td>[final] = 0.01 M</td>
</tr>
<tr>
<td>CH₃CN</td>
<td>5% (v/v)</td>
<td>50% (v/v)</td>
</tr>
</tbody>
</table>
Waters HPLC Directions:

Install one of the C18 columns (either small or large will work) and equilibrate in Buffer A.

In software program:

Click the top left button, and select “Make single injections”.

Increase flow for Buffer A to 1 ml/min. If the flow rate is very low or zero, ramp the rate up over the course of 1 min. Turn the injection valve to “Inject” to make sure the loop is rinsed and equilibrated with Buffer A. Turn the injection valve back to “Load”.

Select the program titled “Analytical 260 549 Cy3”.

Click the button to “Prepare”. Inject the sample into the loop (usually a 250 ul loop is installed), be careful to not inject any air bubbles.

Turn the injection valve to “Inject”, and immediately after, click the button for “Inject” in the software to start the program.

The program will go from 0 to 100% B linearly over the course of 10 min. At this flow rate and time, it is useful to note that each minute represents a 10% increase in %B.

Collect fractions by hand. Elution will occur in the following order:
1. Non-labeled NTS oligo – fractions will look clear, will elute approximately halfway through run (~50% B, corresponding to ~25% CH3CN v/v).
2. Cy3-labeled NTS-oligo – fractions will appear pink, will elute towards the end of the run (~75% B, corresponding to ~37.5% CH3CN v/v).

Eluted volume will be at least 1 ml. Concentrate the Cy3-NTS products using a spin-concentration column. Alternately, centrifugation under vacuum or ethanol precipitation can be used. Pall Microsep spin concentrators (10K or 3K MWCO) work well. Concentrate down to <500 ul. Perform a buffer exchange by resuspending the concentrated solution in 1-2 ml of Buffer EB, and concentrating down to <500 ul. Perform 2-3 washes with Buffer EB, ending with as low a volume as possible (~100 ul).

Spec the final, concentrated, buffer-exchanged Cy3-labeled oligo using an EB blank. Typical concentrations are in the range of 10-100 uM.
DAY 3: Anneal, extend and purify Cy3-labeled, double-stranded promoter

Extension reaction of oligos with Taq Polymerase

Objective: Anneal template strand (TS) and Non-template strand (NTS-Cy3) (in this case, 20 bp overlap) and extend in both directions to create a double-stranded promoter (in this case, 150 bp total length).
Spec DNA using Tris buffer to calculate volumes of NTS and TS needed for reaction:

Final concentrations/volumes of NTS-Cy3 and TS for extension reaction:

\[
\text{[NTS-Cy3]}_{\text{final}} = 0.25 \text{ uM} \\
\text{[TS]}_{\text{final}} = 0.275 \text{ uM (use excess of TS to extend all of NTS-Cy3)}
\]

Example
Non-template strand
\[
(0.25 \text{ uM} / 54.07 \text{ uM}) \times (1000 \text{ ul} / 1 \text{ ml total rxn vol}) = 4.62 \text{ ul}
\]
Template strand:
\[
(0.275 \text{ uM} / 48.57 \text{ uM}) \times (1000 \text{ ul} / 1 \text{ ml total rxn vol}) = 5.66 \text{ ul}
\]

<table>
<thead>
<tr>
<th>Component</th>
<th>[final]</th>
<th>Vol (ul)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oligo 1 – labeled</td>
<td>0.25 uM</td>
<td>4.62</td>
</tr>
<tr>
<td>Oligo 2</td>
<td>0.275 uM</td>
<td>5.66</td>
</tr>
<tr>
<td>10X PCR Buffer</td>
<td>1 X</td>
<td>100</td>
</tr>
<tr>
<td>dNTP</td>
<td>0.2 mM</td>
<td>20</td>
</tr>
<tr>
<td>Taq Polymerase</td>
<td>1:100</td>
<td>10</td>
</tr>
<tr>
<td>ddH2O (nuclease-free)</td>
<td>859.73</td>
<td></td>
</tr>
<tr>
<td>TOTAL</td>
<td></td>
<td>1000 ul</td>
</tr>
</tbody>
</table>

Add components to a 1 ml master mix, starting with ddH2O and ending with enzyme. Mix well with pipette then aliquot into PCR tubes (20 tubes, 50 ul each). (20 tubes) x (50 ul/tube) = 1000 ul.

Thermocycler Program:
*Choose Annealing Temperature to be 5C less than the lowest melting temperature (Tm) of the two oligos.

95C for 4 min (initial melt)
Repeat 2X:
95C for 1 min (melt)
60C* for 2 min (anneal)
72C for 3 min (extend)
72C for 2 min (final extend)
4C hold
**FPLC purification of labeled duplex promoter DNA**

Objective: Purify the extended, double-stranded Cy-3 labeled promoter.

FPLC Column: DNA Swift  
Stored in 20% EtOH.  
Set pressure limit for Pumps A and B to 1.5 MPa.

FPLC Buffers (2L each):

<table>
<thead>
<tr>
<th>Component</th>
<th>Buffer A</th>
<th>Buffer B</th>
</tr>
</thead>
<tbody>
<tr>
<td>NaCl</td>
<td>100 mM</td>
<td>1.5 M</td>
</tr>
<tr>
<td>Tris</td>
<td>20 mM</td>
<td>20 mM</td>
</tr>
<tr>
<td>Acetonitrile</td>
<td>5%</td>
<td>5%</td>
</tr>
</tbody>
</table>

Both buffers are vacuum-filtered and degassed before use.

Prepare finished extension reactions for FPLC:  
Consolidate all 20 rxns into one 1.5 ul Epp tube.

Using Dionex HPLC, choose the “Primer_Extension” program and start a new batch.

1. Create sequence using “primer-extension-preparative” program.  
2. Load sequence  
3. Start Batch

Load sample loop and turn valve to inject to start program.

Watch elution. Initial peaks soon after the loading injections (~ 5 min) are simply components of the extension reaction (dNTP’s, etc) that are eluting early.

Usually, the extended, labeled promoter DNA will elute between 10-12 minutes.

Concentrate elution fractions using an Amicon30 spin column, spinning max speed on tabletop centrifuge for no more than 3 min at a time. Resuspend in TE (or Buffer EB from Qiagen kit) 3 times to wash and ensure full buffer exchange. Last, elute by flipping the column upside down and placing it into a fresh Eppendorf tube. Spin for 2 min to elute.

Determine concentration using Nanodrop.

A typical yield for a 100 nmol scale order used for a 1 ml extension reaction is ~100 µl at ~1 µM.
Sample Name: 130904 JR tbAP3 P1
Vial Number: 1
Sample Type: unknown
Control Program: primer-extension-preparative
Quantif. Method: Test_QNT
Recording Time: 9/4/2013 15:49
Run Time (min): 22.01
Injection Volume: 2000.0
Channel: UV_VIS_1
Wavelength: 260
Bandwidth: n.a.
Dilution Factor: 1.0000
Sample Weight: 1.0000
Sample Amount: 1.0000

<table>
<thead>
<tr>
<th>No.</th>
<th>Ret.Time (min)</th>
<th>Peak Name</th>
<th>Height (mAU)</th>
<th>Area (mAU*min)</th>
<th>Rel.Area (%)</th>
<th>Amount</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.72</td>
<td>n.a.</td>
<td>0.002</td>
<td>0.000</td>
<td>0.00</td>
<td>n.a.</td>
<td>BMB</td>
</tr>
<tr>
<td>2</td>
<td>0.79</td>
<td>n.a.</td>
<td>0.002</td>
<td>0.000</td>
<td>0.00</td>
<td>n.a.</td>
<td>BMB</td>
</tr>
<tr>
<td>3</td>
<td>2.03</td>
<td>n.a.</td>
<td>7.314</td>
<td>7.812</td>
<td>0.09</td>
<td>n.a.</td>
<td>BM</td>
</tr>
<tr>
<td>4</td>
<td>2.76</td>
<td>n.a.</td>
<td>1.526</td>
<td>0.377</td>
<td>0.00</td>
<td>n.a.</td>
<td>Ru</td>
</tr>
<tr>
<td>5</td>
<td>2.97</td>
<td>n.a.</td>
<td>12.294</td>
<td>4.773</td>
<td>0.05</td>
<td>n.a.</td>
<td>M</td>
</tr>
<tr>
<td>6</td>
<td>3.29</td>
<td>n.a.</td>
<td>8.881</td>
<td>3.014</td>
<td>0.03</td>
<td>n.a.</td>
<td>M</td>
</tr>
<tr>
<td>7</td>
<td>4.18</td>
<td>n.a.</td>
<td>4986.418</td>
<td>2826.299</td>
<td>31.17</td>
<td>n.a.</td>
<td>M</td>
</tr>
<tr>
<td>8</td>
<td>5.57</td>
<td>n.a.</td>
<td>4993.129</td>
<td>3476.617</td>
<td>38.34</td>
<td>n.a.</td>
<td>M</td>
</tr>
<tr>
<td>9</td>
<td>6.84</td>
<td>n.a.</td>
<td>4991.155</td>
<td>2163.045</td>
<td>23.86</td>
<td>n.a.</td>
<td>M</td>
</tr>
<tr>
<td>10</td>
<td>7.14</td>
<td>n.a.</td>
<td>192.772</td>
<td>22.545</td>
<td>0.25</td>
<td>n.a.</td>
<td>Mbo</td>
</tr>
<tr>
<td>11</td>
<td>7.54</td>
<td>n.a.</td>
<td>44.210</td>
<td>3.073</td>
<td>0.03</td>
<td>n.a.</td>
<td>Ru</td>
</tr>
</tbody>
</table>

Operator: Washington University
Timebase: Dionex_iocs-5000
Sequence: 130913 JR tbAP3 P1
Page 1-2
9/4/2013 4:13 PM

Chromeleon (c) Dionex 1996-2006
Version 6.80 SR11 Build 3161 (184582)
<table>
<thead>
<tr>
<th>Order Date</th>
<th>Oligo name</th>
<th>Length</th>
<th>Notes</th>
<th>5' - 3' sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>2/20/2013</td>
<td>tbAP3-aminoNTS-85</td>
<td>85</td>
<td>Genomic from -40 to +4, random flanking sequence</td>
<td>AAAATAGGCGTATCACGAGGGCCCTTT CAGCTTTCAAGAAATTCGTTCTGTGACTCCA TGGGACGGATTTGTATTAGACTGGCAGGG/iAmMC6T/TGA</td>
</tr>
<tr>
<td>2/20/2013</td>
<td>tbAP3-TS-85</td>
<td>85</td>
<td>Genomic from -40 to +4, random flanking sequence</td>
<td>CACTACCCAATCGGCGTACGAGGGCCCTTT ACTTCTAGTGGCATGACTGGG TGGGACGGCAAGCTTCAACCCTTACCCAGTCCTCTAATA</td>
</tr>
<tr>
<td>4/23/2013</td>
<td>tbAP3 NTS +2mod</td>
<td>85</td>
<td>Genomic from -40 to +4, random flanking sequence</td>
<td>AAAATAGGCGTATCACGAGGGCCCTTT CAGCTTTCAAGAAATTCGTTCTGTGACTCCA TGGGACGGATTTGTATTAGACTGGCAGGG</td>
</tr>
<tr>
<td>4/23/2013</td>
<td>tbAP3 TS biotin</td>
<td>85</td>
<td>Genomic from -40 to +4, random flanking sequence</td>
<td>/5BiosG/CACTACCCAATCGGCGTACGAGGGCCCTTT ACTTCTAGTGGCATGACTGGG TGGGACGGCAAGCTTCAACCCTTACCCAGTCCTCTAATA</td>
</tr>
<tr>
<td>4/23/2013</td>
<td>tbAP3 NTS no mod</td>
<td>85</td>
<td>Genomic from -40 to +4, random flanking sequence</td>
<td>AAAATAGGCGTATCACGAGGGCCCTTT CAGCTTTCAAGAAATTCGTTCTGTGACTCCA TGGGACGGATTTGTATTAGACTGGCAGGGGTTGA</td>
</tr>
<tr>
<td>4/23/2013</td>
<td>tbAP3 TS</td>
<td>85</td>
<td>Genomic from -40 to +4, random flanking sequence</td>
<td>CACTACCAATGCGCGTACGAGGGCCCTTT ACTTCTAGTGGCATGACTGGG TGGGACGGCAAGCTTCAACCCTTACCCAGTCCTCTAATA</td>
</tr>
<tr>
<td>4/23/2013</td>
<td>tbAP3 TS lessGC</td>
<td>85</td>
<td>GC content reduced downstream</td>
<td>CACTACCAATGCGGATATACGTTAGTGGCATGACTGGG ATACATAAGCCTCTCAACCCTGCGGAGTCTATA</td>
</tr>
<tr>
<td>Date</td>
<td>Description</td>
<td>Temperature</td>
<td>Template Information</td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>---------------------------------------</td>
<td>-------------</td>
<td>--------------------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>4/23/2013</td>
<td>empty NTS +2mod</td>
<td>85</td>
<td>Empty promoter</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>AAAATAGGCGTATCAGAGGCGGCTTT CGTCTTTCAAGAATTTCGTCGTTACGCC ATTGCCGGGATTGTATGCTAAGGCA GGG</td>
<td></td>
</tr>
<tr>
<td>4/23/2013</td>
<td>empty TS lessGC</td>
<td>85</td>
<td>Empty promoter</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CACTACCATCGGATATCATGTAAGTTTCA TTCTGAGTTCAATCATGACGTCAAGCG ATACATAAGCTTCAACCTGAGGATA</td>
<td></td>
</tr>
<tr>
<td>4/23/2013</td>
<td>empty TS less GC biotin</td>
<td>85</td>
<td>Empty promoter, GC content reduced downstream</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>/5BGosG/CACTACCATCGGATATACGT AGTTCACCTTCTGAGTTCATCATGACG TCAGGTGATACATAAGCTTCAACCTGAGG</td>
<td></td>
</tr>
<tr>
<td>4/23/2013</td>
<td>tbAP3 TS less GC biotin</td>
<td>85</td>
<td>GC content reduced downstream</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>/5BGosG/CACTACCATCGGATATACGT AGTTCACCTTCTGAGTTCATCATGACG TCAGGTGATACATAAGCTTCAACCTGAGG</td>
<td></td>
</tr>
<tr>
<td>8/15/2013</td>
<td>tbAP3-NT-amino-85</td>
<td>85</td>
<td>Anneal with either FullCon-T-100 or BubbleCon-T-100</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>GTATACAGACGCTTTGCTTCAAG AATCGGTCTTCAGGTCACATTGCGGATT TGATTATACTGCAAGGTTGAAAGCTTCAACCTGAGG</td>
<td></td>
</tr>
<tr>
<td>8/15/2013</td>
<td>FullCon-NT-100</td>
<td>100</td>
<td>Anneal with FullCon NT 100 for dsDNA competitor</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ATGACGTCAGGTGATAAGGCAGTTTC AAGCTGTTGCTCAGGTCAGAATTGTCAGATTGAAGCTTCAACCTGAGG</td>
<td></td>
</tr>
<tr>
<td>8/15/2013</td>
<td>FullCon-T-100</td>
<td>100</td>
<td>Anneal with FullCon NT 100 for dsDNA competitor</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ATGACGTCAGGTGATAAGGCAGTTTC AAGCTGTTGCTCAGGTCAGAATTGTCAGATTGAAGCTTCAACCTGAGG</td>
<td></td>
</tr>
<tr>
<td>8/15/2013</td>
<td>BubbleCon-T-100</td>
<td>100</td>
<td>Anneal with FullCon NT 100 for Bubble DNA competitor</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ATGACGTCAGGTGATAAGGCAGTTTC AAGCTGTTGCTCAGGTCAGAATTGTCAGATTGAAGCTTCAACCTGAGG</td>
<td></td>
</tr>
<tr>
<td>1/27/2014</td>
<td>tbAP3-NT85_minus3mod-2AP</td>
<td>85</td>
<td>2-Aminopurine at - 3 non-template</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>AAAATAGGCGTATCAGAGGCGGCTTT CGTCTTTCAAGAATTTCGTCGTTACGCC ATTGCCGGGATTGTATGCTAAGGCA GGG</td>
<td></td>
</tr>
<tr>
<td>Date</td>
<td>Sample Description</td>
<td>Length</td>
<td>Notes</td>
<td></td>
</tr>
<tr>
<td>------------</td>
<td>-------------------------------------------</td>
<td>--------</td>
<td>-----------------------------------------------------------------------------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>1/27/2014</td>
<td>tbAP3-T85-unmodified</td>
<td>85</td>
<td>CACTACCATCGGATATACGTAGTTTCA CTTCCTGAGTTTACATGACGTCAGGTG ATACATAAGCTTCAACCCTGCCAGTCTAATA</td>
<td></td>
</tr>
<tr>
<td>1/27/2014</td>
<td>tbAP3-T85_plus2mod-2AP</td>
<td>85</td>
<td>2-Aminopurine at +2 template</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CACTACCATCGGATATACGTAGTTTCA CTTCCTGAGTTTACATGACGTCAGGTG ATACATAAGCTTCA/iAmMC6T/CCCTGCCAGTCTAATA</td>
<td></td>
</tr>
<tr>
<td>1/27/2014</td>
<td>FullCon-NT100_plus2mod-amin</td>
<td>100</td>
<td>FullCon 100 promoter with +2 label, use with Full/BubbleCon-T-100</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>GTATCAGCAGGCCCCCTTTCTCATTTGATAGGTG AATTCGTCTTTGACACATTGCCGGATT TGTATTATACTGGAGG/iAmMC6T/TGAAGCTTATGTACATCATGACGTCAT</td>
<td></td>
</tr>
<tr>
<td>3/20/2014</td>
<td>t3_NT85_minus3-2AP</td>
<td>85</td>
<td>2-Aminopurine at -3 nontemplate</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>AAAATAGGCGTTATCAGCAGGCCCTTTTGCTTCTCAAGAATTCGTCTTGACTCCA TCGGCCGATTTTGTTAGAAGACTGCG/i2AmPr/GGGTTGAG</td>
<td></td>
</tr>
<tr>
<td>3/20/2014</td>
<td>t3_T85_plus2-2AP</td>
<td>85</td>
<td>2-Aminopurine at +2 template</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CACTACCATCGGATATACGTAGTTTCA CTTCCTGAGTTTACATGACGTCAGGTG ATACATAAGCTTCA/i2AmPr/CCCTGCCAGTCTAATA</td>
<td></td>
</tr>
<tr>
<td>4/22/2014</td>
<td>FullCon-NT100_17sp</td>
<td>100</td>
<td>FullCon with 17 bp spacer</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>GTATCAGCAGGCCCCCTTTCTCATTTGATAGGTG AATTCGTCTTTGACACATTGCCGGATT TGTATCATACTGGCAAGGGTTGAAGCTTATGTACATCATGACGTCAT</td>
<td></td>
</tr>
<tr>
<td>4/22/2014</td>
<td>FullCon-T100_17sp</td>
<td>100</td>
<td>FullCon with 17 bp spacer</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CATGACGTAGGGTAGATACATAAGCTT CAACCTGGCAGTATAGCACAAATTAAA CACTCCTGTCAGCAGAATTTCAAGCTTCAAGAGGAAGGGCCCTGCTGAT</td>
<td></td>
</tr>
<tr>
<td>7/22/2014</td>
<td>tbAP3-86t-nomod</td>
<td>86</td>
<td>CACTACCATCGGATATACGTAGTTTCA CTTCCTGAGTTTACATGACGTCAGGTG ATACATAAGCTTCAACCCTGCCAGTCTAATA</td>
<td></td>
</tr>
<tr>
<td>Date</td>
<td>Reagent Name</td>
<td>Length</td>
<td>Description</td>
<td>Sequence</td>
</tr>
<tr>
<td>-----------</td>
<td>-----------------------</td>
<td>--------</td>
<td>--------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>7/22/2014</td>
<td>tbAP3-86t-5'amine</td>
<td>86</td>
<td>End-label</td>
<td>/5AmMC6/GCTCTAGAGCCGGATATACTGATTTCCTCGAGTTTACATGAGTCAAGTAACATTTCAACCTGCGAATGCTCTAATA</td>
</tr>
<tr>
<td>7/22/2014</td>
<td>tbAP3-86nt-nomod</td>
<td>86</td>
<td></td>
<td>CGGGATCCCGGTATACAGGGCCCTTCTGTATCTTCAAGATATTTCGTCCGGACTCCATTGCCGGATTTGATTAGACTGGAATCGGGGGTGA</td>
</tr>
<tr>
<td>7/22/2014</td>
<td>tbAP3-86nt-5'amine</td>
<td>86</td>
<td>End-label</td>
<td>/5AmMC6/CGGGATCCCGGTATACAGGGCCCTTCTGTATCTTCAAGATATTTCGTCCGGACTCCATTGCCGGATTTGATTAGACTGGAATCGGGGGTGA</td>
</tr>
<tr>
<td>7/22/2014</td>
<td>770rev-nomod</td>
<td>22</td>
<td>Primer</td>
<td>CCCACACTACATCGCAGCTGGGTCTCACTAC</td>
</tr>
<tr>
<td>7/22/2014</td>
<td>770rev-5'amine</td>
<td>22</td>
<td>Primer for labeling</td>
<td>/5AmMC6/CCCACACTACATCGCAGCTGGGTCTCACTAC</td>
</tr>
<tr>
<td>7/22/2014</td>
<td>770fwd-nomod</td>
<td>21</td>
<td>Primer</td>
<td>AATAGGCTATGACAGGAGGCCC</td>
</tr>
<tr>
<td>7/22/2014</td>
<td>770fwd-5'amine</td>
<td>21</td>
<td>Primer for labeling</td>
<td>/5AmMC6/AATAGGCGTATGACAGGAGGCCC</td>
</tr>
<tr>
<td>3/26/2015</td>
<td>Gen_tbAP3_NT85_5'bt_+2Am</td>
<td>85</td>
<td>Genomic flanking sequence around Mtb rnap promoter</td>
<td>/5BiosG/GGCCACGGTACCTATGGAATCTATGGAATACCGAACCTGCTTTGACTGCAACCCATGACCCCTGTCGGGCAACCCTGCAGACCTTGGAAGG/iAmMC6T/TGC</td>
</tr>
<tr>
<td>3/26/2015</td>
<td>Gen_tbAP3_T85</td>
<td>85</td>
<td>Genomic flanking sequence around Mtb rnap promoter</td>
<td>TGCAACCCAAACACACTTATGAGCTCAAACCCACCTGTCGGGCAACCCTGCAGACCTTGGAAGG/iAmMC6T/TGC</td>
</tr>
<tr>
<td>7/23/2015</td>
<td>Gen_tbAP3_NT85_5'bt_+2Am</td>
<td>85</td>
<td>Genomic flanking sequence around Mtb rnap promoter</td>
<td>/5BiosG/GGCCACGGTACCTATGGAATCTATGGAATACCGAACCTGCTTTGACTGCAACCCATGACCCCTGTCGGGCAACCCTGCAGACCTTGGAAGG/iAmMC6T/TGC</td>
</tr>
<tr>
<td>7/23/2015</td>
<td>Gen_tbAP3_T85</td>
<td>85</td>
<td>Genomic flanking sequence around Mtb rnap promoter</td>
<td>TGCAACCCAAACACACTTATGAGCTCAAACCCACCTGTCGGGCAACCCTGCAGACCTTGCAAAACACCTGTCGGGCAACCCTGCAGACCTTGGAAGG/iAmMC6T/TGC</td>
</tr>
</tbody>
</table>

Note: The sequences are given in 5' to 3' orientation.
<table>
<thead>
<tr>
<th>Date</th>
<th>Sample ID</th>
<th>Genomic Flanking Sequence Around M. tuberculosis rRNA AP3 Promoter</th>
</tr>
</thead>
<tbody>
<tr>
<td>8/21/2015</td>
<td>Gen_tbAP3_NT85_+2Am</td>
<td>Genomic flanking sequence around Mtb rrnAP3 promoter</td>
</tr>
<tr>
<td></td>
<td>85</td>
<td>/5BiosG/GGCACGGTCACCTATGGATATCTATGGATGACCGAACCTGGTCTTGACTCCATTGGCGATTTGTATTAGAC/TGGCAGGG/iAmMC6T/TGC</td>
</tr>
<tr>
<td>8/21/2015</td>
<td>Gen_tbAP3_T85</td>
<td>Genomic flanking sequence around Mtb rrnAP3 promoter</td>
</tr>
<tr>
<td></td>
<td>85</td>
<td>TGAAACCACCAAACACACTATTGAGTCTCAAACCACACGCTTGCTTGTTTCCGCCCCCTCGGGGGCAACCCTGCCAGTTAATA</td>
</tr>
</tbody>
</table>
Appendix B

Binding protocols and controls
Introduction

I have studied binding interactions between CarD, RNAP, and DNA using double-filter binding assays and electrophoretic mobility shift assays (EMSA) and as described in Chapter 2.

The development of the double-filter binding assays required control experiments testing the effects of the positions of the Cy3 dye (end versus internally labeled) as well as different instruments (STORM vs. Typhoon) for maximizing signal for DNA-only on the nylon over the nitrocellulose. The protocol was developed taking these factors into consideration, and is presented here along with data control experiments.

The development of the EMSA assay was initially difficult due to the presence of a nuclease contaminant in a preparation of Mtb CarD, which caused degradation of the Cy3-DNA substrate. Identifying the cause of the Cy3-DNA degradation was confounded by the use of competitor DNA (i.e. Poly dI-dC) in the assay, as recommended by Hellman and Fried. The use of competitor DNA can not only affect measurements of CarD-DNA binding (which is non-specific and therefore susceptible to competitor DNA), but the presence of competitor DNA can also “soak up” nuclease in the reaction and prevent a researcher from knowing about the contaminant. I have included EMSA results from troubleshooting this issue in the hopes that they can inform future experiments.
Protocol: Nitrocellulose and Nylon Filter-Binding for Quantitative Protein-DNA Binding Interactions

Jayan Rammohan
Galburt Lab

References:

Buffers

<table>
<thead>
<tr>
<th>Reaction buffer:</th>
</tr>
</thead>
<tbody>
<tr>
<td>60% MQ H2O</td>
</tr>
<tr>
<td>10% 10x Mycobacterial transcription buffer +DTT/+BSA (&quot;MTB +/+&quot;)</td>
</tr>
<tr>
<td>...also make MTB +/- (no BSA) for filter soaking and wash steps</td>
</tr>
<tr>
<td>20% RNAP Storage Buffer (10 mM Tris pH 8, 200 mM NaCl, 25% glycerol)</td>
</tr>
<tr>
<td>10% CarD Storage Buffer (20 mM Tris pH 8, 150 mM NaCl)</td>
</tr>
<tr>
<td>Final glycerol = 5% (v/v)</td>
</tr>
</tbody>
</table>

| Reaction volume: 30 ul |
| Reaction mix: |
| 6 ul RNAP Stg Buffer (up to 400 nM [RNAP]final ) |
| 3 ul CarD Stg Buffer (up to 11 uM [CarD]final ) |
| 3 ul 10x MTB +/- |
| 18 ul MQ H2O/DNA for up to ~ 10nM [DNA]final, can use as low as 1nM |

Nitrocellulose and nylon filters

Whatman Protran Nitrocellulose Transfer Membrane (pore size = 0.1 um)

| Treatment: |
| 10 min soak in 0.4 M KOH |
| 3 x 5 min washes in distilled water (bring pH to neutral) |
| 1 hour soak in Rxn Buffer (-BSA) at 4 deg C |

Millipore Immobilon-Ny+ Transfer Membrane

| Treatment: |
| 1 hour soak in Rxn Buffer (-BSA) |

Reaction procedure

Binding assays will be performed at minimal fixed [DNA]. Perform reactions in a sterile, low-protein-binding polystyrene 96-well assay plate, or in PCR tubes.
1. Mix each component of the reaction and incubate for 30 min at desired temperature.
2. Disassemble filter-binding apparatus, place Nitrocellulose filter (notch-cut, top right corner) on top of Nylon filter (diagonal cut, top-right corner), reassemble apparatus.
3. Immediately before filtering each set of samples, wells of the vacuum-filtering apparatus are flushed with 200 ul of Rxn Buffer (-BSA) at desired temperature using a multi-channel pipette, and vacuum-filtered (house vacuum). Avoid overdrying membrane.
4. Load 30 ul reactions into the wells using a multi-channel pipette and vacuum filter. Immediately after filtering each set of samples, the wells of the vacuum-filtering apparatus are washed once with 200 ul of Rxn Buffer (-BSA) at the desired temperature using a multi-channel pipette.
5. The filters are then removed from the apparatus and placed in a heated vacuum dessicator for 20 min (up to 45 min) at 80 deg C under constant vacuum. Overbaking (>1 hr) results in browning and fragmentation of the membrane.
6. Detection of Cy3-DNA was performed using Typhoon with Cy3-fluorescence settings, and “press-sample” was included to reduce warping or wrinkling of the filters.

Note: inclusion of BSA in the soaking and washing buffers can potentially “hyperblock” the nitrocellulose membrane leading to inverse data.

Chemiluminescent Detection (haven’t tried yet, notes from Mascotti et. al.)

1. Soak the nitrocellulose membrane in blocking buffer for 30 min. Blocking Buffer: TBST (+BSA) (10 mM Tris-Cl, pH 8, 100 mM NaCl, 0.1 % Tween 20 containing 1% BSA)
2. Discard blocking buffer and replace with identical blocking buffer containing 40 ul of AvidX streptavidin (PE Biosystems), or 1 ul of streptavidin-alkaline phosphatase conjugate (Ambion).
3. 3 x 5 min washes with blocking buffer.
4. 2 x 2 m in washes with reaction buffer (include 5% solution of Nitroblock (Ambion Nitro-Block II, Cat. No. T2184, 20 ml, $127.00) in 1 x Ambion Brightstar assay buffer)
5. Substrate: CPDSTAR (Ambion, T2146, 100 mL, $315.00) is reacted with the alkaline phosphatase moiety of the membrane-bound AvidX for 5 min before a digital picture was taken with a GDS-8000 cooled CCD camera (UVP, Palo Alto, CA). Exposure times ranged from 4-10 min for quantitative data collection, although band intensity was typically visible within 1-2 min. If no cooled CCD camera is available, one could, alternatively, expose the membrane to X-ray film and quantitate by densitometry.
140612 notes:

Membranes were cut at top right corner – diagonally for nitro, notch for nylon.
Nitro: 0.4M KOH (10 min), 3 washes with MQ H2O, 2 hr 4C soak in Rxn Buff no BSA
Nylon: 1 min wash with 0.5 M NaOH, 3 washes with MQ H2O, 2 hr soak in MQ H2O
Assembled vacuum manifold as described in Wong and Lohman 1993.

Binding reactions were mixed in pcr tubes at room temp, held in metal block for stability.
All additions were performed with a multichannel pipette (5-50ul volume).
96 half-well plate was pre-loaded with rxn buffer (no bsa) for easy loading of multichannel pipette for wash steps.
For each 8-well row:
   50 ul wash (rxn buff no BSA) – vacuum on 5-10s, off
   30 ul load of rxn – vacuum on 5-10s, off
   50 ul wash (rxn buff no BSA) – vacuum on 5-10s, off
Disassembled manifold and placed membranes in a large glass tray. Held down edges and corners with small weights (sharpies and plastic pipette trays) so a gentle stream of hot air could be applied by a hair drier. Moved glass tray to 60C oven (Frieden Lab) containing dessicant. Allowed to dry for 30 min before imaging in STORM.

STORM imaging – “Blue fluorescence/chemiluminescence” mode. Excites at 450 nm, detects ~520nm(?). Best images were collected when filter was face-down, no saran wrap, and “press sample” option was chosen (***without pressing the sample, there were large areas of the filter not visible due to poor contact with the imaging surface***). PMT was set to 800V.
Stored filters in the dark (in a plastic tray in a drawer) over the weekend.
Repeated imaging on 140616 am/pm. Afternoon settings, changed PMT to 650V.
Less saturation and better image contrast.
Initial conclusions and next steps:
High background signal for RNAP with no DNA.
Need at least 10-100nM DNA for detection.
RNAP-dependent signal over background could be detected on Nitrocellulose (100 nM DNA), but not Nylon.

Typhoon Imaging – 1-10nM of Cy3 DNA can be easily detected, with far excess signal on nylon compared to nitrocellulose. Use this for all future experiments.
End-labeled Cy3-DNA substrates cannot be used in double-filter binding assays

I hypothesized that the location of the Cy3 label within the promoter fragment would have no effect on the ability of free Cy3-DNA to pass through the nitrocellulose filter and bind the nylon filter. To test this hypothesis, I constructed promoter DNA fragments with Cy3 labels on the ends of the DNA. I tested these end-labeled fragments against internally labeled DNA using a double-filter binding assay (Figure 1). Ideal signal-over-background in this assay is achieved when DNA-only samples have the least signal on nitrocellulose (which in theory specifically traps protein-bound DNA samples), and the most signal on nylon (which in theory traps free DNA). Interestingly, end-labeled DNA constructs stuck to the nitrocellulose, while internally labeled DNA did not. In addition, I found that pre-treatment of the nitrocellulose with 0.4 M KOH was required to observe higher signal on nylon than nitrocellulose. For these reasons, analysis of double-filter binding assays for estimates of binding affinities were limited to experiments performed with internally-labeled DNA tested with nitrocellulose pre-treated with 0.4 M KOH.
Appendix B, Figure 1: Effects of label position and pre-treatment of nitrocellulose on binding signal. End-labeled DNA (first 6 columns) was compared to internally labeled DNA (last 3 columns) to test whether it would provide sufficient signal-to-noise for double-filter binding experiments. Also tested was pre-treatment of the nitrocellulose filter with 0.4 M KOH. The only sample which provided sufficient nylon signal over nitrocellulose (expected for free DNA) was internally labeled DNA used with 0.4 M KOH-pre-treated nitrocellulose (green boxes).
Appendix B, Figure 2: Effect of DNA concentration on signal-to-noise in double-filter binding experiments. Titrations of [MboRNAP] were performed with DNA ranging from 0 to 100 nM to identify the minimum concentration of DNA that could provide sufficient signal-to-noise for binding experiments. Based on these results, 10 nM DNA was used for all subsequent binding experiments.
Presence of a competing nucleic acid can affect measurements of binding affinity and observation of degraded DNA

It is not uncommon for EMSA protocols to include a competing nucleic acid in order to improve specific detection of protein-DNA interactions. From Hellman and Fried, subsection titled “Competing nucleic acid”:

Often a protein sample will contain more than one nucleic acid binding activity. When secondary binding activities obscure the one of interest, the addition of unlabeled competing nucleic acid to the reaction mixture can reduce the binding of secondary proteins to the labeled target (1). This strategy works when the protein of interest binds the target nucleic acid with greater affinity than it binds the competitor and when the secondary binding activities do not discriminate between competitor and target sequences. Since competing nucleic acids also reduce the amount of specific binding, even under favorable conditions, it is best to test a range of competitor concentrations to optimize discrimination of specific and nonspecific binding. Commonly used competitors include genomic DNAs, poly d(A–T) and poly d(I–C) (2).

I tested the effect of genomic DNA by measuring the binding of *Msm* CarD and *Mtb* CarD to DNA in the presence and absence of Salmon Sperm (SaSp) at 100 ng/μl (75 nM) (Supplemental Figure 3). These experiments indicated that \([\text{CarD}] > 1 \mu\text{M}\) could shift DNA in the absence of SaSp DNA, but not in the presence of SaSp DNA, demonstrating that the presence of competing nucleic acid at this concentration can effect measurements of CarD-DNA interactions. Furthermore, the observation of degraded DNA substrate was possible only in the presence of *Mtb* CarD and the absence of SaSp DNA. This is important because degraded DNA can be symptomatic
of nucleases in the reaction, which could have important consequences for future experiments performed with the same reagents and buffers.

I tested the effect of including Poly(dI-dC) in the EMSA assay by performing titrations of CarD over Cy3-DNA in the presence and absence of 83 ng/ml (63 nM) Poly(dI-dC) (Supplemental Figures 4 and 5). I observed that 11.6 μM Msm CarD was able to shift DNA in the absence but not presence of Poly(dI-dC). These results demonstrate that the presence of Poly(dI-dC) at this concentration can affect the measured binding affinity of CarD to DNA. I also found degraded DNA could be observed in the absence but not presence of Poly(dI-dC). These results further demonstrate that the observation of degraded DNA can be affected by the presence of a competing nucleic acid. Taken together with the results from the Salmon Sperm experiments, they eliminate the possibility that these effects are unique to a specific type of competing nucleic acid, or a particular batch or lot of reagent.

One possible reason for the effect of competing nucleic acids to prevent observation of degraded DNA is that if a nuclease is present, excess competing nucleic acid could potentially “soak up” the nuclease and prevent its activity on the labeled nucleic acid. If nuclease presence is suspected, possible courses of action are (1) to treat buffers with diethyl pyrocarbonate, or (2) re-prepare the buffers and reagents in question (3). For these reasons, analysis of EMSA assays for estimating binding affinities were limited to those reactions performed in the absence of competing nucleic acid.
Appendix B, Figure 3: EMSA of CarD-DNA binding in the presence and absence of Salmon Sperm DNA. Shown is a 4-20% non-denaturing PAGE gel for probing MsmCarD-DNA and MtbCarD-DNA binding interactions in the presence and absence of competing genomic DNA. Lane 10 shows DNA-only. Lanes 1-5 show MsmCarD-shifted DNA. A super-shift can be seen in lane 3, which is no longer visible when the same [MsmCarD] is used in the presence of SaSp DNA (lanes 4). Lanes 6-9 show MtbCarD-shifted DNA. Lanes 6 and 7 show degraded DNA, which are no longer visible when the same [MtbCarD] is used in the presence of SaSp DNA (lanes 8 and 9).
Appendix B, Figure 4: Poly(dI-dC) can affect measurements of CarD-DNA binding affinity in an EMSA. Titrations of MsmCarD over 10 nM Cy3-labeled promoter DNA were performed in the presence (lanes 3-6) and absence (lanes 7-10) of Poly(dl-dC). The blue boxes demonstrate that MsmCarD at 11.6 μM shifts DNA in the absence of Poly(dl-dC) (lane 10) but not in the presence of Poly(dl-dC).
Appendix B, Figure 5: Poly(dI-dC) can mask observation of degraded DNA in an EMSA. Titrations of \textit{Mtb} CarD were performed over 10 nM Cy3-labeled DNA in the presence (lanes 3-6) and absence (lanes 7-10) of Poly(dI-dC). The red boxes indicate that when \([\text{CarD}] > 1 \ \mu\text{M}\), no degraded DNA can be seen in the presence of Poly(dI-dC) (lanes 5 and 6). However, in the absence of Poly(dI-dC), degraded DNA can be seen (lanes 9 and 10). The degradation of DNA is partially-rescued by EDTA, which chelates metal ions and in doing so helps to prevent nuclease activity (lane 11). In the absence of nuclease, Proteinase K would degrade CarD restore the DNA-only band, which is not observed in lane 12. This suggests the presence of a nuclease contaminant in the \textit{Mtb} CarD prep.
References


Appendix C

Stopped-flow fluorescence protocol
**Experiment design and execution**

This section will contain a protocol for experiment design and execution using an Applied Photophysics SX-20 spectrophotometer.

1. **Instrument initialization**

   Clean instrument with Nitric Acid and NaOH (if necessary)
   
   Change emission filter to 570+ (if necessary)
   
   Power off all local electronics
   
   Ignite Lamp – allow ~1 hr for warmup and stabilization
   
   Turn on lamp power supply
   
   Adjust knob to 8 A for ignition
   
   Press ignition button
   
   Adjust knob to 150 W
   
   Turn on water bath, adjust setpoint to desired temp (default = 25 C)
   
   Add water to appropriate level if necessary
   
   Open valve for gas, pressure should be ~120 psi
   
   Power on CPU
   
   Power on electronic control box for SX-20
   
   Load and advance at least 10 ml MQ H2O through the

2. **Prepare Buffers**

   Fill 50 ml conical with MQ H2O
Make 10 ml of 10X MTB +/-

9450 ul 10X MTB -/

500 ul stock NEB BSA (20 mg/ml), [10x BSA]i = 1 mg/ml

50 ul stock DTT (2M), [10x DTT]i = 10 mM

Make 25 ml Protein BLANK buffer:

5 ml RNAP Core storage buffer

5 ml SigA storage buffer

5 ml CarD storage buffer

2.5 ml 10x MTB +/+ 
7.5 ml MQ H2O

Make 25 ml DNA BLANK buffer:

2.5 ml 10x MTB +/+ 

2.5 ml glycerol

20 ml MQ H2O

3. Collect MQ H2O and Buffer Shots

Instrument settings

Fluorescence

Excitation wavelength: 515 nm

Load and advance at least 5 ml of Protein BLANK and DNA BLANK buffers into the stopped-flow. Convention: Protein in Syringe F (on left), DNA in Syringe C (on right).
Set the Detector High Voltage such that the live reading from the PMT is between 0.5-1 V. Typical DHV values range between 500-750 depending on instrument variables (i.e. lamp intensity).

Collect “Buffer” traces.

Collect 3 x 10x/1000pt(log) shots, all 3 shots should overlay. If they do not, collect 3 more. These will serve as the baseline which will be subtracted from all traces.

4. Prepare DNA

Calculate total volume of DNA needed for entire day

*shots/condition = 3 prime + 2 data = 5 total

75 ul / shot

Volume / condition = 5 shots * 75 ul / shot = 425 ul (round up to 500 ul)

Total volume = # conditions * 500 ul

Example: 6 conditions (DNA initial, RNAP, RNAP-CarD1, RNAP-CarD2, RNAP-CarD3, DNA final) would require 6 x 500ul = 3000 ul total DNA at 20 nM.

For equal volume mixing, this results in [DNA]final = 10 nM in the reaction.

5. Collect DNA only traces

Load DNA into Syringe C (on the right). Reload more Protein BLANK into Syringe F (left) if necessary.

Collect “DNA prime” traces.
3 x 10s/1000pts(log).

Collect “DNA only” traces.

2 x 1200s/10,000pts(log).

6. Prepare protein
7. Collect protein traces
8. Shut down instrument

A.3 Data analysis

This

1. Create Fold Change Plots

Open the buffer traces collected at the same DHV as the rest of the experiment.

Using the trace-selection dropdown menu, average the traces

Fit the averaged trace from 0.1-10s using a linear fit.

Record the value. Copy the exact number from the fitting screen using CTRL-C.

Open the DNA and Protein traces (in order if possible) on a single plot by drag/drop. Make sure that each of the traces shows up for a given condition, if the initial file opened only has one trace, subsequent “dropped” files will only display one trace even if multiple traces were collected:

DNA only
MbR225
MbR225-1
MbR225-2
...

231
Trace selection – dropdown menu:

Average the replicate traces for each condition (Shift+mouse click to select multiple traces, then click “Average”). This will result in filenames “Average0, Average1, Average2, … AverageN). Do each condition individually if this is easier, in order to keep the trace numbering in register with sample types.

Select all of the averaged traces, and click “Simple Maths” and then subtract the buffer only value (CTRL-V) to buffer-correct all averaged traces.

Select the buffer-corrected and averaged DNA-only trace, and fit it with a line from 0.1-1200s. Use the y-intercept (“c”) as the signal for the DNA. Again, copying the exact number from the fitting screen using CTRL-C is useful. This value is referred to as “F0”, or “F naught”.

For all of the data traces, calculate fold change as \((F - F0) / F0\). This can be done by selecting all of the traces, and using Simple Maths to first subtract the DNA-only value (CTRL-V), and then again using Simple Maths to divide by the DNA-only value (CTRL-V again).

To clean up the oversampled noise at the beginning of the traces, Smooth the data using a 20-pt window.

Save the finished Fold-Change plot!

2. Measure equilibrium fluorescence

Create a spreadsheet where each row represents each protein condition, and the columns are as follows:
Open the “Fold Change” plot.

Measure the final point of each trace and record this as the EQBM fluorescence value.

3. Fit Protein data to a 3exp

Fit the DNA only (probably called “Smooth0”) to a line from 0.1-1200s, and after fit make sure the “save residual on graph” box is checked and then click “OK”. This will create 2 more traces, one for the fit and one for the residual error.

Fit each protein trace first by fitting to a 2exp from 1-1200s. Then expand the fit to a 3exp and extend the start time earlier from 1s to 0.1s. Copy and past the \([a1, k1, a2, k2]\) to \([a2, k2, a3, k3]\), respectively. Lock these numbers. Do not re-click Estimate. Click Fit until the numbers converge. Record C, a1, k1, a2, k2, a3, and k3 in the table.
Appendix D

Matlab code for eigenvalue calculation
Introduction

Shown here is Matlab code written for evaluating eigenvalues calculated from a 5-state model for factor-dependent stabilization of open-complexes.

```matlab
function SSE_k3_eig2()
%UNTITLED Summary of this function goes here
%   Detailed explanation goes here

% k1 k3
% DNA + R + X    <----> RPc <-----> RPo
% k2 ^     k4 ^
% k5 | k6     k7 | k8
% v     k9 v
% XRc <----> XRPo
% k10

%% FIRST, plot the raw data of k(obs)3 vs. [CarD].
% This section uses C to represent [CarD].
C = [0 5 11 16 25 32 55 105.5 250 527.5 1027.5 1800 2000];
kobs3 = [0.00381 0.00418 0.00186 0.00235 0.00278 0.00212 0.00148
0.00201 0.00457 0.00389 0.00632 0.00812 0.00837];
sem=[
0.000257292
0.000987779
0.000987779
0.000987779
0.000987779
0.000987779
0.000987779
0.00442219
0.000987779
0.00149522
0.001210826
0.000987779
0.000987779];
figure(30);hold on;
errorbar(C,kobs3,sem,'k');

%% SECOND, calculate and plot the eigenvalues vs. [CarD] in a semi-
% continuous fashion
% This section uses "scale" to titrate a rate constant as an outer loop.
% This section uses X to titrate [CarD] as an inner loop, from 0-10,000nM.
% This section uses k#a to designate the original rate value, and k#(i) is
```
defined as a vector that titrates the original value across a range of values defined by "scale". \(k_8(i)\) is calculated accordingly to maintain detailed balance within the thermodynamic cycle.

\[k1 = 0.3; \quad \text{% RNAP on rate}\]
\[k2 = 75; \quad \text{% RNAP off rate}\]
\[k3 \text{ calculated below} \quad \text{% intrinsic opening rate}\]
\[k4 = 0.004; \quad \text{% intrinsic closing rate}\]
\[k5 = 0.3; \quad \text{% CarD on rate to RPC}\]
\[k6 = 400; \quad \text{% CarD off rate from RPC}\]
\[k7 = 0.3; \quad \text{% CarD on rate to RPo}\]
\[k8 = 2.0; \quad \text{% CarD off rate from RPo}\]
\[k9 = 0.0125; \quad \text{% CarD opening rate}\]
\[k10a = 0.0005; \quad \text{% CarD closing rate}\]

\[k3a = (k5*k9*k8*k4)/(k7*k10a*k6) \quad \text{% intrinsic opening rate calculated from detailed balance.}\]

\[%R = [125:50:525];\]
\[R=225;\]
\[\text{scale}=[0.001 0.01 0.05 0.1 0.5 1 1.5 2 2.5 3 4 5 10 100];\]
\[%scale is a scaling factor which is used to titrate a rate constant\]
\[X = [0:0.5:10000]; \quad \text{%X represents a semi-continuous vector of CarD concentrations}\]

\begin{verbatim}
for i = 1:length(scale)
    k10(i) = k10a*scale(i);  \% create a vector for k#
    k3(i) = (k5*k9*k8*k4)/(k7*k10(i)*k6);  \% calculate a vector for k3
    end
\end{verbatim}

\begin{verbatim}
%% PLOT Eigenvalue vs. [CarD] across the range of rate constants defined by "scale"

for i = 1:length(scale)
    for j = 1:length(X)
        %R=2^r;
        %X=[1:10000];
        rates = [ -k1*R k2 0 0 0; k1*R -(k2+k3(i)+k5*X(j)) k4 k6 0; 0 k3(i) -(k4+k7*X(j)) 0 k8; 0 k5*X(j) 0 - (k6+k9) k10(i); 0 0 k7*X(j) k9 -(k8+k10(i)); ];
        E(:,j) = eigs(rates);
    end

% for(j=1:5)      \% plots a fig for each eig
\end{verbatim}
% for(j=2)
%      %figure(19+j);hold on;   %plots a fig for each eig
figure(30);hold on;     %overlay on top of raw data
using same fig# as above

%      set(gca,'ColorOrder',parula);
plot(X,-E(2,:));        %just plot second
eigenvalue
%      lstring{n} = num2str(R(n));
end

%% THIRD, calculate SSE using eigenvalues for each of the experimental
%   values of [CarD]
%   This section uses the same vector R for [RNAP] as the outer loop.
%   This section now uses C to represent [CarD] in the inner loop,
%   where C
%   is the concentrations of CarD tested experimentally.

for i = 1:length(scale)
   for k = 1:length(C)
      rates = [  -k1*R    k2               0              0
                 0;     k1*R     -(k2+k3(i)+k5*C(k)) k4            k6
                 0;     0           k3(i)              -(k4+k7*C(k))  0
                 k8;     0           k5*C(k)         0               -
                 (k6+k9) k10(i);    0               0           k7*C(k)         k9
                -(k8+k10(i))];

      E2(:,k) = eigs(rates); % E2 represents a vector of eigenvalues
      corresponding to each experimental CarD concentration
   end

   % calculate squared error (SE) for each value of experimental CarD
   % (C)
   for a=1:length(C)
      SE(a) = (kobs3(a) - (-E2(2,a)))^2;  % squared error (SE) of
      prediction taking into account every experimental concentration
   end

figure(31);hold on;
plot(k10(i),sum(SE),'o')  % plot titration of rates vs. sum squared
error (SSE) of prediction
figure(32);hold on;
plot(k3(i),sum(SE),'o')
figure(33);hold on;
plot(k10(i)/k3(i),sum(SE),'o')
end

figure(30); hold on; % overlay on top of raw data using same fig# as above
    title('k10 and k3 titration');
    xlabel('[CarD](nM)');
    ylabel('eig2');
    savefig('k10_w_k3_ttrn.fig');

figure(31); hold on;
    title('k10, based on k10 and k3 titration');
    xlabel('k10');
    ylabel('SSE');
    savefig('k10_w_k3_SSE.fig');

figure(32); hold on;
    title('k3, based on k10 and k3 titration');
    xlabel('k3');
    ylabel('SSE');
    savefig('k3_SSE.fig');

figure(33); hold on;
    title('k10/k3, based on k10 and k3 titration');
    xlabel('k10/k3');
    ylabel('SSE');
    savefig('k10_w_k3_ratio.fig');