LACK OF A VITAL RNA POLYMERASE RECOGNITION FEATURE RESULTS IN COORDINATED CONTROL OF THE MAJORITY OF *RHODOBACTER SPHAEROIDES* PROMOTERS

by

Kemardo Kitaro Henry

A dissertation submitted in partial fulfillment of

the requirement for the degree of

Doctor of Philosophy

(Microbiology)

at the

UNIVERSITY OF WISCONSIN-MADISON

2020

Date of final oral examination: 12/14/2020

The dissertation is approved by the following members of the Final Oral Committee:
Richard L. Gourse, Professor, Bacteriology (advisor)
Timothy J. Donohue, Professor, Bacteriology
James L. Keck, Professor, Biomolecular Chemistry
Robert Landick, Professor, Biochemistry and Bacteriology
Jue D. Wang, Professor, Bacteriology

Lack of a vital RNA polymerase recognition feature results in coordinated control of the majority of *Rhodobacter sphaeroides* promoters

Kemardo K. Henry

Under the supervision of Professor Richard L. Gourse

University of Wisconsin-Madison

Both the synthesis of ribosomes and translation itself are major consumers of cellular energy in bacteria. Consequently, rRNA synthesis is highly regulated in response to changes in nutritional and environmental conditions. Studies of *E. coli* rRNA transcription have provided us with a framework for understanding promoter recognition by RNA polymerase (RNAP) in general. However, rRNA promoters in most other bacteria, for example photosynthetic bacteria, have not been explored in great detail. Our goal was to investigate transcription from rRNA promoters in *Rhodobacter sphaeroides* to provide insights into how a photosynthetic bacterium regulates transcription and ribosome synthesis in response to light.

We first developed a method for purifying *R. sphaeroides* RNAP ($RNAP_{Rsph}$) using a strain encoding a C-terminal His_{10} tag on the β' subunit of RNAP for purification by affinity chromatography. We found that transcription from the promoters of all three *R. sphaeroides* rRNA operons with $RNAP_{Rsph}$ was extremely weak *in vitro*, although it was able to transcribe from some non-rRNA promoters and from an *E. coli* rRNA promoter. All 3 *R. sphaeroides* rRNA operons contain an A at the last position of the -10 hexamer rather than the almost invariant T at this positon found in *E. coli* promoters, yet the *R. sphaeroides* sigma factor responsible for recognizing housekeeping promoters contains the amino acid residues that recognize this base. Furthermore, $RNAP_{Rsph}$

transcribed from *R. sphaeroides* rRNA promoters very efficiently *in vitro* when the suboptimal A was replaced with a T. Therefore, we proposed that *R. sphaeroides* rRNA promoters might require a transcription factor to compensate for their suboptimal -10 elements.

We found that the transcription factor CarD strongly activates transcription from *R. sphaeroides* rRNA promoters *in vitro*. A subset of promoter mutations in the extended -10 and -10 elements bypassed the requirement for CarD, and amino acids in CarD that are essential for CarD to facilitate open complex formation in other species are required for $CarD_{Rsph}$ to activate *R. sphaeroides* rRNA promoters. Most surprisingly, in collaboration with the Donohue and Landick labs (UW-Madison) we found that many/most *R. sphaeroides* promoters contained a suboptimal base at the last position in the -10 hexamer and required CarD for activity. Consistent with a role in rRNA synthesis, Tn-seq experiments indicated that the *carD* gene appeared to be essential, and CarD levels remained constant during exponential growth but decreased upon entry into stationary phase under both aerobic and photosynthetic conditions. We found that CarD inhibited its own promoter by limiting RNAP escape from the promoter, indicating that CarD is not always an activator in *R. sphaeroides*. Using the in vitro transcription with CarD, we found that DksA and ppGpp inhibit transcription from *R. sphaeroides* rRNA and r-protein promoters.

DEDICATION

I'd like to dedicate my thesis to my family, who, despite having no idea what I work on and wondering why I love school so much, always root and care for me.

To all my mentors who believed in me, who guided me, and gave me a chance.

Finally, I'd like to dedicate my thesis to those who seek knowledge, especially to those, like myself, who grew up wanting to be scientists despite knowing little about what that journey entails.

ACKNOWLEDGEMENTS

I thank Richard Gourse for being my advisor. I thank him for his guidance, immense patience and for showing genuine interest in my development as a scientist and as a person. I'd like to thank him for introducing me to the world of Chemex brew, coffee hasn't been the same since.

I thank Wilma Ross for her intoxicating enthusiasm for science. I thank her for being willing to talk about a puzzling result, sharing love of birds and showing a sincere interest in my wellbeing. I thank her for helping me develop as a scientist.

I thank the Gourse lab members, past and present, for their friendship, constructive criticisms, and for being great people to work with and be around. Grad school was a rollercoaster, and I wouldn't pick another set of people to go with me on that ride. I will cherish all the memories we share.

I thank my committee members: Profs Timothy Donohue, James Keck, Robert Landick and Jade Wang for providing me with lots of advice and guidance. Particularly, I'd like to thank Timothy Donohue for making me an honorable Donohue lab member and carving out a workspace for me in your lab.

I'd like to thank the members of the Donohue lab for welcoming me into the world of *Rhodobacter*. I always felt welcomed. I thank you for helping me get set up in your lab, answering my naïve questions and providing me with a place to sleep during my long growth experiments.

I extend my gratitude to Cathy Davis Gray, Abbey Thompson and Christina Hull for creating an amazing space to learn and grow. I am thankful for their support and assistance. I also thank Prof. JD Sauer for his valuable time and scientific advice.

I thank the MSB staff to making sure that everything runs smoothly (Janet Newlands, Kari Straus, Kim Besta, Gary Pine, Kyle Boldon and the custodial staff).

To my family: I want to thank them for always believing me and encouraging me to pursue my passions. Never ending thanks to my mom and dad, Judith and Romeo, who instilled in me the value of an education, and who have worked so hard to ensure that I had access to a better life.

My partner, Loren: I am grateful for her constant love and support, and for challenging me every day to be a more reflective and diligent human being. And to our pets, Beaker and Georgette for keeping me company as I write my thesis.

And to you, for reading this.

TABLE OF CONTENTS

Abstract ii
Dedication iv
Acknowledgements v
Table of contents vi
List of figures ix
List of tables xi

Chapter One – Introduction 1
Bacterial RNA polymerase 2
Promoter regions 5
RNAP holoenzyme interaction with promoter 5
Transcription initiation: isomerization 12
Regulation of bacterial transcription in *Escherichia coli* 13
rRNA regulation 16
Regulation of transcription in gram-positive 21
rRNA transcription in Mycobacterium tuberculosis 22
Introduction of *Rhodobacter sphaeroides* 26
Stringent control in *R. sphaeroides* 28
Outline of thesis 29
References 33

Chapter Two – A majority of *R. sphaeroides* promoters lack a crucial RNA polymerase recognition feature, enabling coordinated transcription activation 41
Abstract 42
Significance 43
Introduction 44
Results 47
R. sphaeroides rRNA promoters initiate poorly in vitro with *R. sphaeroides* RNAP 47
R. sphaeroides CarD activates rRNA promoters 51
CarD directly activates other *R. sphaeroides* promoters 56
Promoter features that correlate with increased transcription by $CarD_{Rsp}$ 60
Context dependence of promoter sequences that bypass the effect of CarD on *rrnB* 63
The majority of *R. sphaeroides* promoters lack the T at the last position in the -10 element 66
Levels of $CarD_{Rsp}$ change in vivo 71
Discussion 74
The absence of -7T is a major determinant of the *R. sphaeroides* transcriptome 74
Role of -7T in promoter function 74
CarD facilitates a step in the mechanism of promoter opening that is affected by the RNAP interaction with promoter position -7T 75

Context affects promoter activity 76
The absence of a conserved RNAP recognition feature in a large set of promoters creates a “regulon” controlled by a transcription factor 78
The role of CarD and -7T in other α-proteobacteria and other bacterial phyla 78
Implications for regulation of synthesis of the *R. sphaeroides* translation apparatus 80
Materials and Methods 81
Acknowledgements 82
References 83
Supplemental Information 87
Expanded Materials and Methods 103
Tables 114
Dataset 1 125
Supplemental References 126

Chapter Three – *Rhodobacter sphaeroides* CarD Negatively Autoregulates Its Own Promoter 128
Abstract 129
Introduction 130
Results 133
CarD inhibits transcription from its own promoter 133
CarD substitutions that impair activation also impair inhibition 133
carD promoter sequence features important for high activity and inhibition by CarD 136
$CarD_{Rsp}$ inhibition of the *carD* promoter leads to an accumulation of abortive RNAs 145
Discussion 152
CarD is a transcription factor that is capable of activation and inhibition 152
Mechanism of inhibition by CarD 152
Potential role of autoregulation in determining levels of CarD during log phase growth 155
Supplemental Figures 157
Materials and Methods 163
Tables 167
References 170

Chapter Four – ppGpp/DksA inhibits CarD-activated transcription 174
Abstract 175
Introduction 176
Results 179
R. sphaeroides DksA does not inhibit transcription of the CarD activated *R. sphaeroides rrnB* promoter in the absence of ppGpp 179
ppGpp inhibits transcription from *R. sphaeroides rrnB* in the absence of DksA 182

DksA$_{Rsp}$ and ppGpp synergistically inhibit *rrnB* transcription *in vitro* by *R. sphaeroides* RNAP 182
DksA$_{Rsp}$ and ppGpp inhibit transcription from other CarD-activated promoters 185
DksA$_{Rsp}$ with ppGpp does not inhibit transcription from the *carD* promoter 189
DksA and CarD levels change similarly in vivo during aerobic growth 189
DksA and CarD levels change differentially in vivo during photosynthetic growth 192
Deletion of the *dksA* gene results in widespread changes in gene expression 196
Discussion 201
Regulatory features of *R. sphaeroides* rRNA promoters 202
Supplemental Figures 205
Materials and Methods 210
Tables 217
References 218

Chapter Five – Conclusions and Future Directions 222
Conclusions 223
Future Directions 227
Why did *R. sphaeroides* (and other alpha-proteobacteria) evolve to have a so many promoters without a thymine at the last position of the -10 element? 227
What regulates CarD protein levels during stationary phase? 228
When is ppGpp induced in *R. sphaeroides* cells? 229
Why does *R. sphaeroides* DksA by itself does not inhibit rRNA transcription? 229
What is the role of ppGpp and DksA in photosynthetic growth? 230
What genes are directly regulated by ppGpp and DksA in vivo? 230
Do DksA/ppGpp and CarD bind RNAP at the same time? 231
References 232

Appendix A – Measuring ppGpp in *R. sphaeroides* cells by LC-MS 235
Summary 236
Results 236
Materials and Methods 238
Reference 239

Appendix B – Transcriptional studies of CarD from *Caulobacter crescentus* on cognate RNA polymerase and *carD* promoter 240
Summary 241
Results 242
Materials and Methods 244
Tables 247
References 248

LIST OF FIGURES

Chapter One – Introduction

Figure 1.1 Crystal structure of *E. coli* RNA polymerase holoenzyme 3
Figure 1.2 Promoter elements recognized by modules in RNAP holoenzyme 6
Figure 1.3 The -11 and -7 bases flip into pockets within the σ subunit during open complex formation 10
Figure 1.4 Simplified schematic of the mechanism of transcription initiation 14
Figure 1.5 Crystal structure of RNAP showing the two ppGpp binding sites 19
Figure 1.6 Crystal structure of CArD bound to RNAP-DNA complex 24

Chapter Two – A majority of *Rhodobacter sphaeroides* promoters lack a crucial RNA polymerase recognition feature, enabling coordinated transcription activation

Figure 2.1 Activity of *R. sphaeroides* rRNA promoters with either *R. sphaeroides* RNAP or *E. coli* RNAP 48
Figure 2.2 Activities of the *R. sphaeroides* rRNA promoters in vitro in with or without purified *R. sphaeroides* CarD 52
Figure 2.3 Activation of *R. sphaeroides* non-rRNA promoters in vitro with $CarD_{Rsp}$ 58
Figure 2.4 Effects of individual *R. sphaeroides rrnB* promoter mutations on activation by CarD 61
Figure 2.5 Context dependence of effects of -7T or extended -10 mutants 64
Figure 2.6 Percentage of promoters in indicated bacterial species containing a thymine at the last position of the -10 hexamer (-7) or an extended -10 consensus (TG) at -14, -15, or both 68
Figure 2.7 $CarD_{Rsp}$ levels at different phases of growth 72
Figure 2.S1 RNAP purification 87
Figure 2.S2 Transcription of non-*R. sphaeroides* promoters with 20 nM $RNAP_{Rsp}$ 89
Figure 2.S3 *R. sphaeroides* rRNA promoter sequences 91
Figure 2.S4 Major σ factor and αCTD alignments 93
Figure 2.S5 Alignments of CarD, activation by wild-type or W91L mutant $CarD_{Rsp}$, and transcription from *R. sphaeroides rrnB* -11 mutant promoters 95
Figure 2.S6 In vitro transcription from *R. sphaeroides rrnA* wild-type and spacer mutant promoters 97
Figure 2.S7 *R. sphaeroides* promoter sequences 99
Figure 2.S8 Relative levels of $CarD_{Rsp}$ and CFU/ml in *R. sphaeroides* cultures during aerobic growth 101

Chapter Three – *Rhodobacter sphaeroides* CarD Negatively Autoregulates Its Own Promoter

Figure 3.1 CarD inhibits transcription from its own promoter 134

Figure 3.2 Inhibition of *carD* promoter transcription by CarD is reduced by substitutions in CarD 137
Figure 3.3 *carD* promoter sequences from various proteobacteria 140
Figure 3.4 Effects of individual *carD* promoter mutations on *carD* promoter activity and inhibition by CarD 143
Figure 3.5 Analysis of abortive transcripts from the carD promoter 147
Figure 3.6 Abortive probabilities of each initially transcribed sequence from the *carD* promoter 150
Figure 3.S1 Identification of lengths of abortive transcripts from the *carD* promoter 157
Figure 3.S2 CarD does not stabilize the *rrnB* promoter against competitor 159
Figure 3.S3 Clustal Omega alignment of CarD from *Rhodobacter sphaeroides* 2.4.1 and *Caulobacter crescentus* NA1000 161

Chapter Four – ppGpp/DksA inhibits CarD-activated transcription
Figure 4.1 $DksA_{Rsp}$ alone (without ppGpp) does not inhibit transcription from the *R. sphaeroides rrnB* promoter 180
Figure 4.2 ppGpp and DksA synergistically inhibit *rrnB* transcription 183
Figure 4.3 ppGpp and DksA synergistically inhibit transcription from some *R. sphaeroides* promoters, but do not inhibit transcription from the *carD* promoter 187
Figure 4.4 DksA and CarD levels at outgrowth, log and stationary phase in aerobically grown cells 190
Figure 4.5 DksA and CarD levels at outgrowth and log to stationary phase in photosynthetically growth cells 194
Figure 4.6: Global gene expression differences between Δ*dksA* and wild-type *R. sphaeroides* strains grown aerobically or photosynthetically 199
Figure 4.S1 Comparison of the amino acid sequence of portions of the *E. coli* and *R. sphaeroides* RNAP subunits or of full length *E. coli* or *R. sphaeroides* DksA containing residues involved in ppGpp binding 205
Figure 4.S2 *R. sphaeroides* RNAP binds radiolabeled ppGpp in DRaCALA assay 207
Figure 4.S3 Relative levels of $DksA_{Rsp}$ and CarD and CFU/mL in aerobically or photosynthetically grown *R. sphaeroides* cultures 208
Figure 4.S4 Heat-map showing expression changes for gene involved in translation that were upregulated in the *R. sphaeroides* Δ*dksA* vs wild-type strains 209

Appendix A – Measuring ppGpp in *Rhodobacter sphaeroides* cells by LC-MS
Figure AA Levels of ppGpp levels in aerobically or photosynthetically grown *R. sphaeroides* cells. ppGpp was measured by LC-MS 237

Appendix B – Transcriptional studies of CarD from *Caulobacter crescentus* on cognate RNA polymerase and *carD* promoter

Figure AB Activity of the *R. sphaeroides* or *C. crescentus carD* promoter with *C. crescentus* RNAP 243

LIST OF TABLES

Chapter Two – A majority of *Rhodoacter sphaeroides* promoters lack a crucial RNA polymerase recognition feature, enabling coordinated transcription activation

Table 2.S1 Strains and plasmids 114
Table 2.S2 Oligonucleotides used to construct plasmid templates for in vitro transcription 115
Table 2.S3 Plasmids used for in vitro transcription 119
Table 2.S4 Primers used to construct chromosomally-encoded C-terminal thrombin-His10 tagged β' 121
Table 2.S5 List of gene blocks 123
Table 2.S6 Primers used to construct pETSUMO-CarD 124

Chapter Three – *Rhodobacter sphaeroides* CarD Negatively Autoregulates Its Own Promoter

Table 3.1 Strain and Plasmid 167
Table 3.2 Oligonucleotides used for constructing *in vitro* transcription vectors 167
Table 3.3 Plasmids used for *in vitro* transcription 168
Table 3.4 Oligonucleotides used for generating *carD* promoter fragment for abortive transcription 169
Table 3.5 Sequence of *carD* promoter fragment in abortive transcription assay 169

Chapter Four – ppGpp/DksA inhibits CarD-activated transcription

Table 4.1 Strain and Plasmid 217
Table 4.2 Plasmids used for *in vitro* transcription 217

Appendix B – Transcriptional studies of CarD from *Caulobacter crescentus* on cognate RNA polymerase and *carD* promoter

Table AB.1 Strains and Plasmid 247
Table AB.2 Oligonucleotides used for constructing *in vitro* transcription vectors 247

CHAPTER ONE

INTRODUCTION

In this chapter, I provide an introduction to transcription initiation, how it is regulated in bacteria, and I introduce major topics discussed in this thesis.

Bacterial RNA polymerase

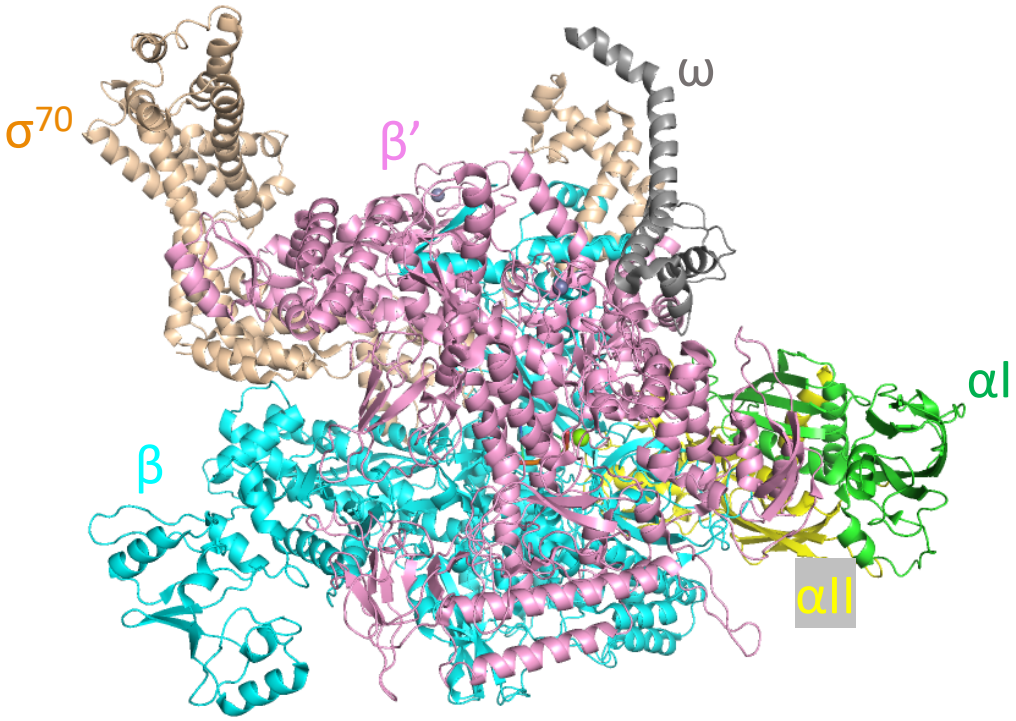
Transcription is the first step in gene expression, where genetic information encoded in DNA is converted to RNA. The process of transcription in bacteria is carried out by a multi-subunit DNA-dependent RNA polymerase (RNAP) (Figure 1), which is comprised of a core enzyme made of β, β', an α dimer, and an ω subunit (1). The core enzyme carries out all steps of transcription except for initiation for which it requires a dissociable initiation factor, σ, to bind recognition elements in template DNA called promoters (Figure 1.1) (1–3). The majority of gene expression in bacteria is initiated by the housekeeping σ (σ^{70} or σ^{A}) associated holoenzyme (core + σ) (4, 5).

The assembly pathway of RNAP core enzyme is sequential in that the β subunit binds to the α dimers and then the β' and ω subunits assembly onto the $\alpha_2\beta$ subassembly (6). Structural studies of RNA polymerase reveal a crab claw like structure where the two larger subunits, β and β', form the pincers (7–9). The cleft formed by the β and β' subunits houses the active site where template DNA enters to be transcribed into RNA (10, 11).

The process of rRNA transcription is highly regulated, as it presents an initial investment of a large amount cellular energy that will be dedicated to ribosome biogenesis and protein synthesis (12, 13). By regulating rRNA transcription, cells can fine tune gene expression in order to adapt to and survive changes in their environment such as nutritional depletion. In the next sections in this introduction, I will briefly review the current literature on promoter recognition and transcription initiation in general, and how it is regulated in the model organism *Escherichia coli*.

Figure 1.1. Crystal structure of *E. coli* RNA Polymerase holoenzyme

Crystal structure of *E. coli* RNAP holoenzyme containing σ^{70}. αI is colored in green, αII in yellow, β in cyan, β' in pink, ω in gray, and σ^{70} in wheat. (DNA was removed for clarity). This figure was generated from PDB4YLN deposited in ref. (7)


σ70
β’
ω
αI
β
αII

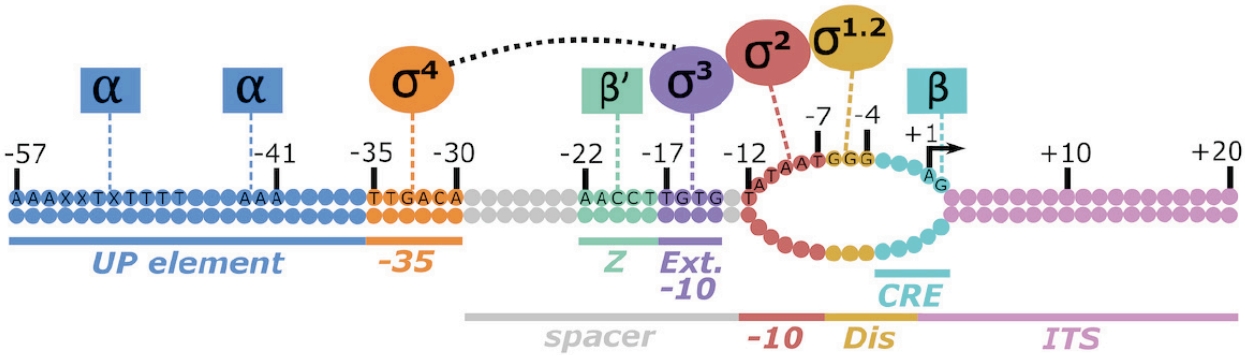
Promoter regions

Promoter DNA is generally comprised of -35 and -10 elements or hexamers which are numbered in relation to the transcription start site (TSS) (+1). The nucleic acid sequence of the -35 and -10 elements is highly conserved throughout different bacteria phyla and generally have the consensus sequence $T_{-35}TGACA_{-30}$ and $T_{-12}ATAAT_{-7}$ relative to the TSS respectively (14). A spacer sequence of varying length, generally 17bp, separates the -35 and -10 elements. Some promoters contain an extended -10 sequence (-15TG-14) that is upstream of the -10 element and/or an A+T rich region upstream of the -35 element at positions approximately -60 to -40 relative to the TSS, known as the UP element. The discriminator region is between the -10 element and the TSS. The sequence in each of these promoter regions can have a significant impact on the activity of a promoter. RNAP holoenzyme utilizes base specific and non-specific (DNA phosphate backbone) protein-DNA interactions involving distinct regions in σ^{70} and core enzyme to recognize promoter DNA and initiate transcription (Figure 1.2).

RNAP holoenzyme interaction with promoter

The initial step of transcription initiation is the formation of the closed complex where RNAP holoenzyme ($E\sigma^{70}$) binds to the promoter from approximately position -55 to +15 relative to the TSS (15, 16). The major holoenzyme binds promoter DNA utilizing a DNA binding domain with the α subunit (α-CTD) and distinct regions on the σ subunit (reviewed in (1, 17–19)). RNAP α-CTDs binds the minor groove of the A+T rich UP using a helix-hairpin-helix (HhH) motif (20, 21). UP elements have two subsites, proximal and distal, relative

Figure 1.2. Promoter elements recognized by modules in RNAP holoenzyme. The promoter regions and the respective interacting module in RNAP are colored the same. Most of the sequence-specific interactions occur between the σ subunit of RNAP and the promoter. These interactions are: σ4 with the -35 element (orange), σ3 with the extended -10 element (purple), σ2 contact the -10 element (red), and σ1.2 with the discriminator region (yellow). Other contacts with RNAP core include: α-CTD with the UP element (blue), β' with the Z-element (light green) and the β with the core recognition element (CRE, teal). The transcription start site (+1) is indicated with the black arrow and the initially transcribed sequence (ITS) is shown in pink). The figure is adapted from ref (16).


α
α
σ4
β'
σ3
σ2
σ1.2
β
-57
-41
-35
-30
-22
-17
-12
-7
-4
+1
+10
+20
AAAXXTXTTTT
AAA
TTGACA
AACCTTGTG
TATAATGGG
AG
UP element
-35
Z
Ext.
-10
CRE
spacer
-10
Dis
ITS

to the -35 (22). Some promoters have DNA binding sites for other transcription factors upstream of UP elements where transcription factors such as Fis can bind and help recruit RNAP to the promoter (18, 23, 24). A summary of RNAP holoenzyme interactions with promoter DNA is shown in Figure 1.2.

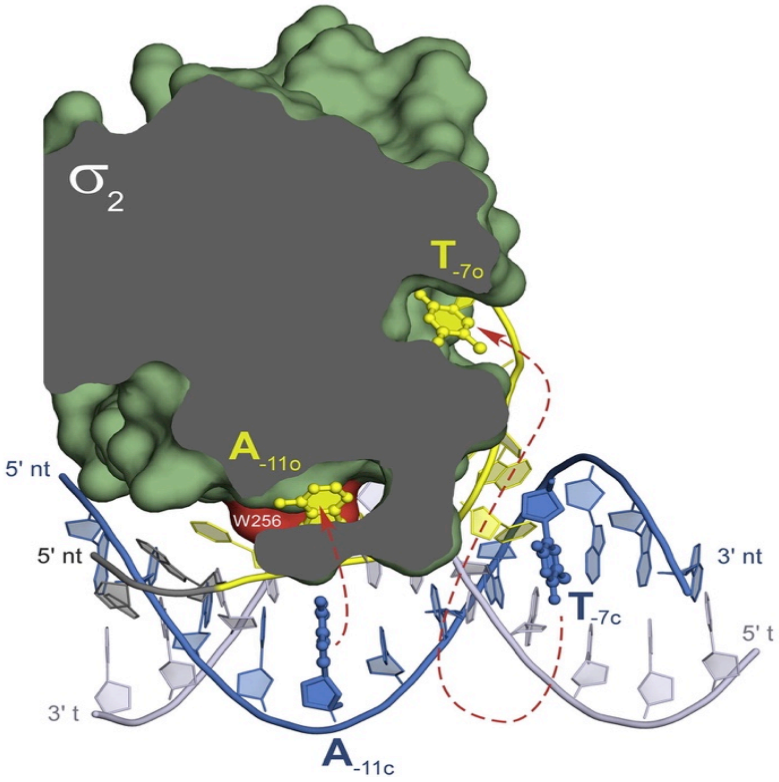
Immediately downstream of the UP element is the -35 element, which is recognized in a sequence-specific manor by σ subunit region 4 (σ_4). Early genetic and biochemical studies established the importance of the -35 element as an important contributor to RNAP binding to promoter DNA (25), and structural studies revealed that an helix-turn-helix (HTH) within σ_4 gets inserted into the major groove and makes non-specific interactions with the phosphate backbone and interacts specifically with bases of the -35 element, namely the TTG motif (26). Interactions of holoenzyme with the spacer sequence between the -35 and -10 elements, positions -29 to -17, are primarily non-specific or backbone interactions (11). The most common spacer length for promoters is 17bp. This length is structurally ideal as it allows σ_4 and σ_2 to make unstrained contacts with the -10 and -35 elements, respectively (8). Altering the spacer length of a promoter can significantly affect promoter activity or regulation (27–29). There have been reports proposing that Eσ can make sequence specific interactions with bases within the spacer region (30).

Present in some promoters is an extended -10 sequence ($T_{-15}G_{-14}n$) that lies upstream of the -10 element and makes sequence-specific interactions with σ_3 (8, 11). Two helices perpendicular to each other that lie between the σ_3 and σ_2 domains insert into the major groove of the extended -10 region and cause the promoter to bend towards σ_3. Bending of the DNA allows σ_3 to make sequence-specific interactions with

the TG sequence of the extended -10 (8, 11). Some promoters with the TGn motif appear to have weaker -35 elements, and introduction of an extended -10 sequence can significantly increase the activity of a promoter by increasing σ_3 contacts to DNA (5, 31, 31). Introduction of an extended -10 sequence onto test promoters has been shown to increase the lifetime of the open complex and decrease the rate of promoter escape (27, 32).

The -10 element is the most conserved region within promoters and has the consensus sequence of $T_{-12}ATAAT_{-7}$ in *E. coli*. It interacts with σ_2 and participates in binding and isomerization, i.e. the melting of promoter DNA (RPo). Substitutions within the -10 element towards a non-consensus sequence generally leads to a reduction in promoter activity. Particularly, mutations at the -11 and -7 positions have dramatic effects on open complex formation and promoter activity. During isomerization, the -11A and -7T positions are flipped into corresponding pockets within σ_2, with the capture of the -11A appearing to occur first (Figure 1.3) (10, 33). Although the -12T base is recognized as duplex DNA, recent structural work showed that it transiently melts during isomerization (7). Strand separation propagates from the -11 position downstream with bases from -11 to +2 being captured in specific pockets, namely the nontemplate strand A at -11 and T at -7, the template strand T at -9, the non-template G at -5 or -6, and the nontemplate G at +2, all of which appear to modulate RPo formation (7, 27, 33, 34). The discriminator region, which lies between the -10 element and the TSS makes a nontemplate strand contact with $\sigma_{1.2}$. Interaction of the base two positions downstream

Figure 1.3. The -11 and -7 bases flip into pockets within the σ subunit during open complex formation. σ region 2 has binding pockets that accommodate only an A at position -11 and a T at the -7 position. During isomerization/open complex formation, the flipping of these highly conserved bases into σ is important for promoter stability and activity. Figure was adapted from ref (33).


σ2
T-7o
A-11o
W256
5' nt
5' nt
3' nt
T-7c
5' t
3' t
A-11c

of the -10 element with $\sigma_{1.2}$ is sequence specific, and these interactions increases open complex lifetime (27). Lastly, the RNAP core β-subunit interact with the single-stranded DNA of the nontemplate strand positions -4 to +2, known as core recognition element (CRE), with the +2G being sequence specific. The presence of this interaction increases the lifetime of RPo on some promoters (34, 35).

RNAP holoenzyme makes contacts with promoter DNA that can be either specific or nonspecific. The basal activity of a promoter can be significantly increased by introducing sequences that improve specific interactions with $E\sigma^{70}$. Although this approach may produce a promoter variant with the desired activity, there is not a linear relationship between the number of specific interactions and promoter activity. For example, a promoter with a perfect -10 and -35 element can be weaker than one with a sequence one or two nt from consensus, or introducing an UP element to a promoter with a good -10 and -35 element can create a promoter that is crippled for transcription by inhibiting promoter escape (36). There are context-dependent effects that dictate the relationship of $E\sigma$ interactions with a promoter and its activity (reviewed in (17)).

Transcription initiation: isomerization

During transcription initiation, RNAP (R) binds promoter DNA (P) through interactions with the σ subunit and core enzyme that make specific and non-specific interactions with the promoter forming the closed complex (RPc), in which the DNA is duplex (1, 2). Binding of RNAP to the promoter is dependent on the concentration of RNAP and the presence of DNA bindings sites (16). The binding affinity of RNAP to promoter DNA can be increased by introducing consensus sequences for the UP element or the -35

element. Once formed, the RPc complex transitions to the open complex (RPo) through at least 6 additional steps (combined here as RP_I, where the duplex DNA in the -10 element gets unwound from -11 to +12 relative to the start site (Figure 1.4) (7).

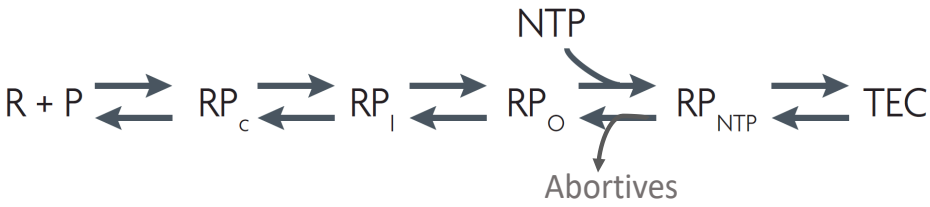
The isomerization from RPc to RPo involves the flipping of bases of the -10 element and discriminator into pockets within core enzyme or σ, the positioning of the template strand into the active clef where the TSS can be transcribed into RNA. NTPs enter RNAP and get incorporated into a nascent RNA chain that is complementary in sequence to the initial transcribing sequence (ITS) (Figure 1.2); however, as the chain elongates, RNAP has to disengage from the promoter in order to “escape”. If RNAP fails to escape the promoter, nascent RNAs are released as “abortives” and RNAP reattempts making nascent RNA (37, 38). For a particular promoter, any (or multiple) of these steps can be rate limiting or present a bottleneck for transcription. Transcription factors may affect any of these steps that are rate limiting to activate or inhibit transcription from a promoter.

Once the energy barrier is overcome, RNAP breaks contacts with and escapes from the promoter and transitions to the elongation complex (Figure 1.4). RNAP will then continue to elongate the RNA message using the DNA as template, until it terminates releasing the RNA (39).

Regulation of bacterial transcription in *E. coli*

Transcription is a highly regulated process, it is regulated at every step: initiation, elongation and termination. Transcription factors can either have a positive or negative effect on transcription and they vary dramatically in their mechanism of action.

Figure 1.4: Simplified schematic of the mechanism of transcription initiation. RNAP (R) interactions with the promoter (P) to form the closed complex (RPc) where the DNA is still duplex. Using thermal energy, the RPc complex transitions through multiple isomerization steps (abbreviated here as RPi) to form the open complex (RPo). Open complex formation involves the melting of promoter DNA and the positioning of the template strand into the active site. NTP then binds to the transcription start site (TSS) and RNAP attempts to incorporate incoming NTPs into a growing RNA chain. If RNAP is unable to break contacts with the promoter recognition elements, it will release the RNA as abortives and reattempt transcription. Once RNAP escapes the promoter, it converts to an elongation complex (TEC). Figure adapted from ref (1).


NTP
R + P
RP c
RP I
RP O
RP NTP
TEC
Abortives

Transcription factors such as Fis bind DNA specifically to DNA and most often activate transcription by recruitment of RNAP to the promoter while transcription factors such as DksA/ppGpp and TraR do not bind DNA but rather act on the rate limiting kinetics steps during isomerization to either activate or inhibit transcription (8). In *E. coli*, ribosomal RNA (rRNA) transcription can account for the majority of transcription within the cell and thus is the target of exquisite regulation. Our knowledge of bacterial transcription initiation has been significantly informed through studies of rRNA transcription. In the next section I will review literature on regulation of rRNA transcription in *E. coli* and other bacteria.

rRNA regulation

Ribosomes are the largest consumer of cellular energy, both because of their biogenesis and their role in translation of protein. Ribosomes are made up of 3 rRNAs and more than 50 r-proteins and accessory factors. As rRNA synthesis is a rate-limiting step in the production of ribosomes, its transcription is highly regulated (13, 40). The rate of rRNA synthesis is sensitive to the nutritional environment and consequently mirrors the growth rate of cells. Its level rapidly increases when cells are upshifted from poorer to richer medium. When cells enter stationary or are downshifted from richer to poor medium, rRNA synthesis decreases and in the latter case the rapid decrease is referred to as stringent control (12, 13, 40). Early studies found that increases or decreases in rRNA synthesis during shifts are tied to the concentration of two small molecules: initiating nucleotides (iNTPs) and the “alarmone” molecules ppGpp (and (p)ppGpp, which will be collectively referred to here as ppGpp) (41). During log growth

when the concentration of iNTPs is high, rRNA is not limited by the NTP concentration, however, when cells enter stationary phase rRNA transcription decreases in response to an increase in ppGpp levels. Once ppGpp levels decrease again in deep stationary phase, it is the low concentration of iNTPs that keeps transcription of rRNA low (41).

ppGpp can be synthesized by the RelA or SpoT proteins, but synthesis of ppGpp during nutritional starvations and shifts is primarily from RelA in response to the build-up of uncharged tRNA (amino acid starvation) and other nutritional stresses (1, 41). rRNA transcription decreases when ppGpp is induced, and in vitro studies revealed that ppGpp inhibits transcription directly (42). The magnitude of inhibition observed for rRNA transcription in vivo during nutrient downshifts could not be explained by the relatively small in vitro effects, and this led to the discovery of DksA as an important cofactor for ppGpp effects on transcription (43).

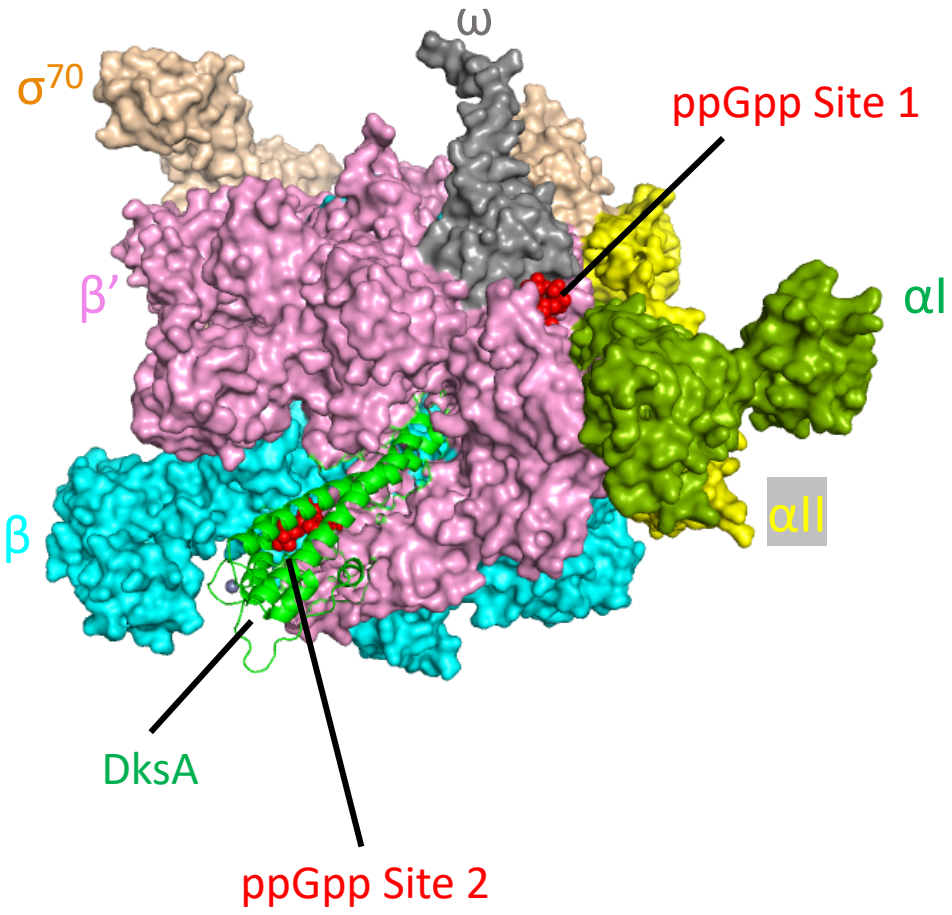
E. coli DksA is a 151 residue protein that binds to the secondary channel of RNAP. It has a globular domain that interacts with the β' rim helices which helps position the acidic tip of the DksA coiled-coil domain in close proximity to the RNAP active site (44, 45). While DksA can inhibit transcription alone at high concentrations, it works synergistically with ppGpp to allosterically inhibit transcription in vitro and in vivo, particularly rRNA transcription (43, 46). The combination of DksA and ppGpp also lead to an increase in amino acid biosynthesis gene transcription, directly and indirectly (42, 46). Cells unable to synthesize ppGpp (*ΔrelA ΔspoT*) or with a deletion of the *dksA* gene do not regulate rRNA synthesis during nutrient downshifts and are unable to grow in minimal medium without amino acids (43, 47). Genome-wide expression studies revealed that ppGpp and DksA directly regulate transcription of more than 700 genes in

E. coli (48). Included in these genes, ppGpp and DksA inhibit the expression of genes involved in ribosome biogenesis and metabolism and activate transcription of amino acid biosynthetic and stress response genes (48–50).

ppGpp regulates transcription by binding to two distinct sites on RNAP. The first identified ppGpp binding site (site 1) is at the interface of the ω and β' subunits of RNAP (51, 53) and accounts for the in vitro inhibition of rRNA promoters observed with ppGpp alone (42, 51). Cells lacking site 1 were partially defective in stringent control compared to *ΔrelA ΔspoT* or *ΔdksA* strains (51). Transcription with RNAP variants lacking site 1 could only account for a quarter or a third of the effect of ppGpp on transcription when DksA was included (46). This finding suggested that there was another ppGpp binding site (now referred to as site 2) that is formed when DksA is included. Through the use of genetic and biochemical assays, site 2 was determined to be formed at the interface of DksA and the β' rim helices (Figure 1.5) (46). Elimination of both ppGpp binding sites significantly reduced the ppGpp effect on transcription *in vitro*, and *E. coli* mutants lacking both sites had a similar phenotype to a *ΔrelA ΔspoT* strain (46, 48).

As mentioned above, ppGpp and DksA synergistically affect transcription. These factors do not bind DNA directly but instead bind RNAP and affect the rate-limiting step for a particular promoter to regulate transcription output (12, 54, 55). RNAP binds rapidly to the *E. coli rrnB* P1 promoter but forms an unstable open complex, which is in rapid equilibrium with earlier intermediates. ppGpp and DksA inhibit transcription from the *rrnB* P1 promoter by shifting the RNAP-complex to an intermediate that is incompetent for transcription (43, 54). In contrast, ppGpp and DksA appear to increase the rate of isomerization, which is rate limiting for activated promoters (12). The

Figure 1.5: Crystal structure of RNAP showing the two ppGpp bindings sites. X-ray crystallography was used to map the two ppGpp bindings sites on RNAP. ppGpp was soaked to a preformed RNAP-DksA crystal. The subunits of RNAP are indicated. DksA is shown in green and ppGpp is shown as red spheres. ppGpp binds at the interface of the ω and β' subunits (site 1) and at the interface of DksA and the β' subunit. Figure was taken from ref (51), PDB 5VSW.


ω
σ70
ppGpp Site 1
β′
αI
β
αII
DksA
ppGpp Site 2

secondary channel binding protein, TraR also inhibits transcription by a similar mechanism but unlike DksA it does not require ppGpp (55, 56).

TraR is functionally and structurally similar to DksA and binds to the secondary channel of RNAP. Biochemical and structural studies using either TraR or DksA revealed how the binding of these two factors induce conformational changes that ultimately affect transcription output (52, 55, 57). It appears that several sequence insertions (SI) in *E. coli* RNAP (58) undergo conformational changes when RNAP is bound by either TraR or DksA, or occupy a similar region on RNAP (52, 57). These sequence insertions are present in proteobacterial RNAPs but not found in RNAPs from gram-positive bacterial species (58). The binding of either TraR or DksA in the secondary channel places them in close proximity to β SI1 (also called βlobe SI1 or βi4 (58)) and β' SI3 (β'i6). The presence or absence of these sequence insertion regions affects the ability of TraR to regulate transcription (55). The distribution of β SI1 in proteobacteria correlates with the occurrence of DksA homologs (57). A DksA homolog has not been found in gram-positive bacteria (59).

Regulation of transcription in gram-positive bacteria

While almost all bacterial species appear to use ppGpp to regulate transcription (4), DksA and TraR are not predicted to be found in every species. *Bacillus subtilis,* a member of the phylum Firmicutes, uses ppGpp and pppGpp to indirectly regulate rRNA synthesis by reducing the intracellular concentration of the initiating NTP (iNTP), GTP. *B. subtilis* rRNA open complexes are short lived and require high iNTP concentrations (59), and by limiting the GTP pools these promoters are negatively regulated since they

all initiate with GTP. When the +1 NTP of the rRNA promoters was changed from G to A, rRNA promoter activity was insensitive to ppGpp levels (59). In particular, during nutrient downshifts, GTP pools were reduced by pppGpp; both by consuming GTP pools for its synthesis and from acting on proteins involved in GTP synthesis (60). pppGpp (and other hyper phosphorylated forms of GDP) was shown to bind to enzymes involved in purine biosynthetic and savage pathways (61–64).

rRNA transcription in *Mycobacterium tuberculosis*

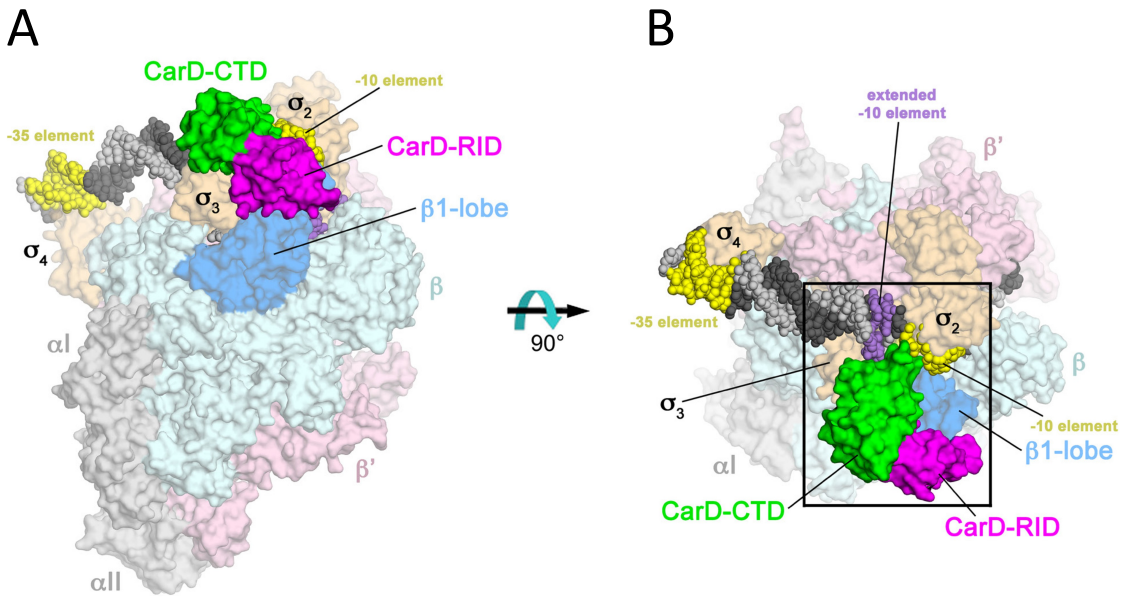
M. tuberculosis RNAP ($RNAP_{Mtb}$) forms short-lived open complexes on promoters on which *E. coli* RNAP forms longed lived open complexes. Biochemical characterization of the $RNAP_{Mtb}$ complex with test promoters found that the instability is an intrinsic property of $RNAP_{Mtb}$ rather than the promoter (65). This instability of $RNAP_{Mtb}$ is hypothesized in part to arise from the absence of lineage specific sequence insertions in RNAP β and β' subunits (66). $RNAP_{Eco}$ lacking β' SI3 (not found in $RNAP_{Mtb}$) forms unstable open complex compared to wild-type RNAP on the same promoter (67), supporting this hypothesis. Studies of transcription by mycobacterial RNAP revealed that *M. tuberculosis* uses transcription factors, not found in *E. coli*, to stabilize open complexes formed by $RNAP_{Mtb}$ and activate transcription (reviewed in (66)).

CarD is an essential protein found in many bacterial species, including *M. tuberculosis.* It was first identified as upregulated in response to DNA damage and was found to play a role in persistence and rRNA transcription (68). Although, it was first proposed to function similarly to DksA, it was later found to activate rather than inhibit

rRNA promoters (69). CarD has non-specific DNA binding properties and an N-terminal domain with amino acid similarity to the RNAP interacting domain (RID) of the transcription-repair coupling factor (TRCF or Mfd), which directly interacts with the β-subunit of RNAP (70). CarD homologs are widely distributed in bacteria; they are found in δ-proteobacteria, actinobacteria, some Firmicutes, Deinococcus/Thermus, Spirochaetes, and some α-proteobacteria lineages (69).

CarD binds directly to the β1 lobe of RNAP and colocalizes with the housekeeping sigma, σ^A, in promoter regions (Figure 1.6A) (69, 71, 72). In addition to *M. tuberculosis*, CarD activates rRNA transcription in *Thermus thermophilus* and *Myxococcus xanthus* (71, 73). CarD activates rRNA transcription in *M. tuberculosis* by stabilizing the open complex formed $RNAP_{Mtb}$ by inserting a highly conserved tryptophan residue into the minor groove at the upstream edge of the -10 element (Figure 1.6B) (65, 71, 74). Similar to DksA/ppGpp and TraR, CarD does not bind DNA in a sequence-specific manor but rather modulates transcription by affecting the kinetic properties of the promoter (75). CarD depletion or CarD substitutions were reported to affect mRNA levels and rRNA expression in *M. smegmatis* or *M. tuberculosis* (68, 76) and CarD has been proposed to be a bifunctional transcription factor, capable of activation and inhibition of transcription depending on the promoter kinetics (66, 75, 76).

Figure 1.6: Crystal structure of CarD bound to RNAP-DNA complex. (A) The CarD RNAP interaction domain (RID; magenta) binds to the beta1 lobe of RNAP (light blue). This interaction positions the DNA binding domain at CarD C-terminal (CTD; green) to interact with promoter DNA just upstream of the -10 element (yellow). (B) Same structure as in (A), rotated 90 degrees to visualize CarD interaction with DNA. Box indicates CarD-CTD interaction with DNA. The subunits of RNAP are indicated. Figure was taken from ref (71).


A
CarD-CTD
-10 element
σ2
-35 element
CarD-RID
σ3
β1-lobe
σ4
β
αI
β'
αII
90°
B
extended
-10 element
β'
σ4
-35 element
σ2
β
σ3
-10 element
β1-lobe
αI
CarD-CTD
CarD-RID

Although the studies of transcription in *E. coli*, *B. subtilis* and *M. tuberculosis* have provided us with a framework for understanding promoter recognition and transcription RNAP in general, we are limited in our understanding of how transcription is regulated in bacteria from different environmental niches. The basic features of promoter recognition are conserved throughout bacteria, and an assumption is often made that transcription in an uncharacterized bacterium will be similar to what is known from studying model organisms. Studies have often utilized RNAP purified from a model bacterium, such as *E. coli*, to study a noncognate gene or protein (77, 78). This approach has been powerful to understand gene expression, but some basic properties of the transcription system in different bacterial species may differ, invalidating the non-cognate approach.

Introduction of *Rhodobacter sphaeroides*

Rhodobacter sphaeroides is a purple non-sulfur α-proteobacterium, capable of growing aerobically, photosynthetically, or anaerobically (79, 80). It has a high GC genome (~65%), two chromosomes, and 3 rRNA operons: *rrnA* on chromosome 1 and *rrnB* and *rrnC* on chromosome 2. The promoters of the *rrnA*, *rrnB*, and *rrnC* operons were identified by primer extension (81, 82). The transcription apparatus in *R. sphaeroides* is made up of core enzyme ($\alpha_2\beta\beta'\omega$) and 17 sigma factors, with the housekeeping sigma designated σ^{93} due to its mobility as a 93 kDa polypeptide (83–85). When growing aerobically, *R. sphaeroides* is morphologically similar to other gram-negative bacteria, however, during photosynthetic growth (or growth in low oxygen) it creates specialized membrane invaginations called intracytoplasmic membranes (ICM)

where red pigmented protein complexes, which make up the photosystem, are inserted (86). Genetic analysis of the genes involved in *R. sphaeroides* photosynthetic growth showed that when component(s) of the photosynthesis apparatus are inhibited, *R. sphaeroides* colonies are pale and are unable to grow under strict photosynthetic conditions (87–89).

Most of the genes involved in photosynthesis (PS) are clustered in a ~67 kb DNA region in *R. sphaeroides* (79, 90). The expression of these genes is subjected to tight regulation by a network of light and oxygen-dependent factors, including PrrBA, a redox-responsive two-component system (87, 91). PrrA is an activator of photosynthesis genes in *R. sphaeroides* and *prrA* mutants are unable to grow photosynthetically (87). The cognate histidine kinase PrrB senses a decrease in oxygen tension in the cell membrane and activates PrrA by phosphorylation (92). PrrA has an N-terminal receiver domain that is phosphorylated by PrrB and has a C-terminal effector domain with a HTH motif that has structural homology to Fis, a DNA binding protein that increases *E. coli* rRNA promoter activity 3-8 fold, depending on the rRNA operon (93). Gene-reporter assays found that PrrA activates the promoters of photosynthetic genes in vivo (94).

In vitro transcription assays using RNAP holoenzyme purified from the photosynthetic bacteria *R. sphaeroides* (84, 94), *R. capsulatus* (95, 96), and *Rhodospirillum rubrum* (97) focused primarily on characterizing whether these RNAPs could recognize *E. coli*-derived promoters or function in the transcription of photosynthetic genes. The study of transcription in *R. sphaeroides* was advanced significantly when it was shown that a native rRNA promoter (*rrnB*) was transcribed by

R. sphaeroides holoenzyme ($E\sigma^{93}$) but not by *E. coli* RNAP (83). This suggested that there was a difference in promoter recognition by RNAP $E\sigma$ from *R. sphaeroides* and *E. coli*. However, the molecular basis for this difference was not explored further. In addition, the direct regulation of transcription, particular of rRNA transcription in response to the nutritional alarmone, ppGpp is poorly understand in *R. sphaeroides*.

Stringent control in *R. sphaeroides*

As stated earlier, *B. subtilis* and *E. coli* are evolutionarily far-removed organisms, but they employ different mechanisms for regulating rRNA promoters in response to amino acid availability. They both use ppGpp to do so, but it was reported that *R. sphaeroides* may use ppGpp for responses to light but not for responses to changes in amino acid concentrations (98). *R. sphaeroides* ppGpp levels increase with reduced light intensity, and the maximal concentration is obtained when cells are shifted completely into the dark (99). The increase in ppGpp levels is accompanied by inhibition of rRNA synthesis as in *E. coli* (100). Unlike *E. coli*, however, amino acid starvation did not affect ppGpp levels of aerobically and photosynthetically grown *R. sphaeroides* cells (98), but total RNA accumulation and cell turbidity were negatively affected. This suggests there is a ppGpp-independent mechanism for regulating rRNA transcription after amino acid starvation.

Our lab, in close collaboration with the Donohue lab (UW-Madison), recently characterized a DksA homolog in *R. sphaeroides.* The revised genome of *R. sphaeroides* 2.4.1 annotated a 158 amino acid protein (*RSP*_2654) as a possible DksA homolog. *RSP*_2654 is functionally similar to *E. coli* DksA in that it complements growth

on minimal medium without amino acids in a *dksA*$_{Eco}$ deletion strain and it inhibits *E. coli* rRNA transcription in vitro using *E. coli* components (101). *RSP_*2654 interacts with RNAP$_{Eco}$ and reduces the lifetime of RNA polymerase complexes like *DksA*$_{Eco}$ further justifying classifying it as DksA$_{Rsp}$ (101). Additionally, DksA$_{Rsp}$ like DksA$_{Eco}$, synergizes with ppGpp and reduced the half-life of the complex formed by *R. sphaeroides* Eσ^{93} and the *lacVU5* promoter *in vitro* (101). Further we found that *R. sphaeroides* cells harboring the chromosomal deletion of *dksA*$_{Rsp}$ produced paler colonies than wild-type and were unable to grow photosynthetically (101). Δ*dksA* cells produced components of the light harvesting red-pigmented complex but were unable to utilize this energy for growth and other cellular processes, which suggests that this protein is involved in photosynthetic growth.

Together, these results suggest that the stringent responsive molecules in *R. sphaeroides*, i.e., DksA and ppGpp, play a role in photosynthetic growth. *R. sphaeroides* is found in environmental niches distinct from *E. coli*, and it is our hope that by studying transcription from rRNA promoters in *R. sphaeroides* we can gain insights into how a photosynthetic bacterium regulates transcription and ribosome synthesis in response to light.

Outline of thesis:

The overall goal of this thesis was to examine the roles of DksA and ppGpp in rRNA transcription and photosynthetic growth in *R. sphaeroides*. Previous work done by a graduate student in the Gourse lab found that DksA$_{Rsp}$ and ppGpp had a similar effect on *R. sphaeroides* RNAP$_{Rsp}$ as reported for the analogous system in *E. coli*. Attempts to

transcribe an *R. sphaeroides* rRNA promoter were unsuccessful, however, and thus when I joined the project, my first goal was to determine why transcription from rRNA promoters was so weak.

In Chapter 2, I developed an *in vitro* transcription system using RNAP purified from *R. sphaeroides*. I engineered an *R. sphaeroides* strain to encode a C-terminal His_{10} tag on the β' subunit of RNAP and purified RNAP by affinity chromatography. Transcription from the promoters of all three *R. sphaeroides* rRNA operons with $RNAP_{Rsp}$ was extremely weak *in vitro*, although it was able to transcribe from some non-rRNA promoters and from an *E. coli* rRNA promoter. All 3 *R. sphaeroides* rRNA operons contain an A at the last position of the -10 hexamer rather than the almost invariant T at this position found in *E. coli* promoters, yet the *R. sphaeroides* sigma factor responsible for recognizing housekeeping promoters contains the amino acid residues that recognize this base. Furthermore, $RNAP_{Rsp}$ transcribed from *R. sphaeroides* rRNA promoters very efficiently *in vitro* when the suboptimal A was replaced with a T. I demonstrated that the *R. sphaeroides* CarD homolog activated transcription from *R. sphaeroides* rRNA promoters, thus compensating for their suboptimal -10 elements. One of the most surprising findings of Chapter 2 was that many/most *R. sphaeroides* promoters contained a suboptimal base at the last position in the -10 hexamer and required CarD for activity. Consistent with a role in rRNA synthesis, I found that CarD levels remained constant during exponential growth but decreased upon entry into stationary phase. These data provided an explanation for why *R. sphaeroides* rRNA promoters are not recognized by *E. coli* RNAP and identified CarD as an important

component of the *R. sphaeroides* transcription apparatus. This chapter was recently published (102).

In Chapter 3, I demonstrated that *R. sphaeroides* CarD autoregulates its own expression by inhibiting transcription of its own promoter. I found that the *carD* promoter synthesizes significantly more abortive transcripts when transcribed by $RNAP_{Rsp}$ in the presence of CarD, which supports the model that CarD inhibits transcription by slowing the rate of promoter escape.

In Chapter 4, I used the in vitro transcription developed in Chapter 2 to demonstrate the $DksA_{Rsp}$ and ppGpp inhibited *R. sphaeroides* rRNA transcription. In addition, $DksA_{Rsp}$ and ppGpp also regulate the expression of other genes involved in translation, inhibiting transcription from r-protein and tRNA genes. Utilizing RNA-seq, we found significant changes in gene expression when $DksA_{Rsp}$ is deleted in *R. sphaeroides*.

In summary, the major achievements of my work have been (1) the discovery that the majority of *R. sphaeroides* promoters lack the highly conserved T at the last position of the -10 element, (2) the identification and characterization of CarD as an important component of the *R. sphaeroides* transcription apparatus that compensates for the absence of -7T in many promoters, (3) the discovery that CarD inhibits transcription of its own promoter by preventing promoter escape, (4) the discovery that $DksA_{Rsp}$ and ppGpp inhibit rRNA transcription once it is activated by CarD, and (5) the finding that the promoters from other α-bacterial species also lack -7T (as reported also in (103)).

In Chapter 5, I discuss future directions for studying transcription in *R. sphaeroides* and present remaining questions that would advance our understanding of bacterial transcription in general.

REFERENCES:

1. Haugen SP, Ross W, Gourse RL. 2008. Advances in bacterial promoter recognition and its control by factors that do not bind DNA. Nat Rev Micro 6:507–519.
2. Ruff EF, Record MT, Artsimovitch I. 2015. Initial Events in Bacterial Transcription Initiation. 2. Biomolecules 5:1035–1062.
3. Feklístov A, Sharon BD, Darst SA, Gross CA. 2014. Bacterial Sigma Factors: A Historical, Structural, and Genomic Perspective. Annual Review of Microbiology 68:357–376.
4. Murakami KS, Darst SA. 2003. Bacterial RNA polymerases: the wholo story. Current Opinion in Structural Biology 13:31–39.
5. Gruber TM, Gross CA. 2003. Multiple Sigma Subunits and the Partitioning of Bacterial Transcription Space. Annual Review of Microbiology 57:441–466.
6. Ishihama A. 1981. Subunit of assembly of Escherichia coli RNA polymerase. Adv Biophys 14:1–35.
7. Chen J, Chiu C, Gopalkrishnan S, Chen AY, Olinares PDB, Saecker RM, Winkelman JT, Maloney MF, Chait BT, Ross W, Gourse RL, Campbell EA, Darst SA. 2020. Stepwise Promoter Melting by Bacterial RNA Polymerase. Molecular Cell https://doi.org/10.1016/j.molcel.2020.02.017.
8. Zuo Y, Steitz TA. 2015. Crystal Structures of the E. coli Transcription Initiation Complexes with a Complete Bubble. Molecular Cell 58:534–540.
9. Zhang G, Campbell EA, Minakhin L, Richter C, Severinov K, Darst SA. 1999. Crystal Structure of Thermus aquaticus Core RNA Polymerase at 3.3 Å Resolution. Cell 98:811–824.
10. Feklistov A, Bae B, Hauver J, Lass-Napiorkowska A, Kalesse M, Glaus F, Altmann K-H, Heyduk T, Landick R, Darst SA. 2017. RNA polymerase motions during promoter melting. Science 356:863–866.
11. Bae B, Feklistov A, Lass-Napiorkowska A, Landick R, Darst SA. 2015. Structure of a bacterial RNA polymerase holoenzyme open promoter complex. eLife 4:e08504.
12. Gourse RL, Chen AY, Gopalkrishnan S, Sanchez-Vazquez P, Myers A, Ross W. 2018. Transcriptional Responses to ppGpp and DksA. Annual Review of Microbiology 72:163–184.
13. Paul BJ, Ross W, Gaal T, Gourse RL. 2004. rRNA Transcription in Escherichia coli. Annual Review of Genetics 38:749–770.
14. Shultzaberger RK, Chen Z, Lewis KA, Schneider TD. 2007. Anatomy of Escherichia coli σ70 promoters. Nucl Acids Res 35:771–788.
15. Cook VM, deHaseth PL. 2007. Strand Opening-deficient Escherichia coli RNA Polymerase Facilitates Investigation of Closed Complexes with Promoter DNA effects of DNA sequence and temperature. J Biol Chem 282:21319–21326.

16. Saecker RM, Record MT, deHaseth PL. 2011. Mechanism of Bacterial Transcription Initiation: RNA Polymerase - Promoter Binding, Isomerization to Initiation-Competent Open Complexes, and Initiation of RNA Synthesis. J Mol Biol 412:754–771.

17. Jensen D, Galburt EA. 2020. The Context-Dependent Influence of Promoter Sequence Motifs on Transcription Initiation Kinetics and Regulation. Journal of Bacteriology https://doi.org/10.1128/JB.00512-20.

18. Gourse RL, Ross W, Gaal T. 2000. UPs and downs in bacterial transcription initiation: the role of the alpha subunit of RNA polymerase in promoter recognition. Molecular Microbiology 37:687–695.

19. Mejía-Almonte C, Busby SJW, Wade JT, van Helden J, Arkin AP, Stormo GD, Eilbeck K, Palsson BO, Galagan JE, Collado-Vides J. 2020. Redefining fundamental concepts of transcription initiation in bacteria. 11. Nature Reviews Genetics 21:699–714.

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

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

22. Estrem ST, Ross W, Gaal T, Chen ZWS, Niu W, Ebright RH, Gourse RL. 1999. Bacterial promoter architecture: subsite structure of UP elements and interactions with the carboxy-terminal domain of the RNA polymerase α subunit. Genes Dev 13:2134–2147.

23. Newlands JT, Ross W, Gosink KK, Gourse RL. 1991. Factor-independent activation of Escherichia coli rRNA transcription. Journal of Molecular Biology 220:569–583.

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

25. Kobayashi M, Nagata K, Ishihama A. 1990. Promoter selectivity of Escherichia coli RNA polymerase: effect of base substitutions in the promoter −35 region on promoter strength. Nucleic Acids Res 18:7367–7372.

26. Campbell EA, Muzzin O, Chlenov M, Sun JL, Olson CA, Weinman O, Trester-Zedlitz ML, Darst SA. 2002. Structure of the Bacterial RNA Polymerase Promoter Specificity σ Subunit. Molecular Cell 9:527–539.

27. Haugen SP, Berkmen MB, Ross W, Gaal T, Ward C, Gourse RL. 2006. rRNA Promoter Regulation by Nonoptimal Binding of σ Region 1.2: An Additional Recognition Element for RNA Polymerase. Cell 125:1069–1082.

28. Gaal T, Barkei J, Dickson RR, deBoer HA, deHaseth PL, Alavi H, Gourse RL. 1989. Saturation mutagenesis of an Escherichia coli rRNA promoter and initial characterization of promoter variants. J Bacteriol 171:4852–4861.

29. Auble DT, Allen TL, deHaseth PL. 1986. Promoter recognition by Escherichia coli RNA polymerase. Effects of substitutions in the spacer DNA separating the -10 and -35 regions. J Biol Chem 261:11202–11206.

30. Yuzenkova Y, Tadigotla VR, Severinov K, Zenkin N. 2011. A new basal promoter element recognized by RNA polymerase core enzyme: Novel class of bacterial promoters. The EMBO Journal 30:3766–3775.

31. Barne KA, Bown JA, Busby SJW, Minchin SD. 1997. Region 2.5 of the Escherichia coli RNA polymerase σ70 subunit is responsible for the recognition of the 'extended −10' motif at promoters. The EMBO Journal 16:4034–4040.

32. Voskuil MI, Chambliss GH. 2002. The TRTGn Motif Stabilizes the Transcription Initiation Open Complex. Journal of Molecular Biology 322:521–532.

33. Feklistov A, Darst SA. 2011. Structural Basis for Promoter −10 Element Recognition by the Bacterial RNA Polymerase σ Subunit. Cell 147:1257–1269.

34. Zhang Y, Feng Y, Chatterjee S, Tuske S, Ho MX, Arnold E, Ebright RH. 2012. Structural Basis of Transcription Initiation. Science 338:1076–1080.

35. Petushkov I, Pupov D, Bass I, Kulbachinskiy A. Mutations in the CRE pocket of bacterial RNA polymerase affect multiple steps of transcription 12.

36. Gaal T, Ross W, Estrem ST, Nguyen LH, Burgess RR, Gourse RL. 2001. Promoter recognition and discrimination by EσS RNA polymerase. Molecular Microbiology 42:939–954.

37. Hsu LM. 2002. Promoter clearance and escape in prokaryotes. Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression 1577:191–207.

38. Li L, Molodtsov V, Lin W, Ebright RH, Zhang Y. 2020. RNA extension drives a stepwise displacement of an initiation-factor structural module in initial transcription. PNAS 117:5801–5809.

39. Adhya S, Gottesman M. 1978. Control of Transcription Termination. Annu Rev Biochem 47:967–996.

40. Nomura M. 1999. Regulation of Ribosome Biosynthesis in Escherichia coli and Saccharomyces cerevisiae: Diversity and Common Principles. Journal of Bacteriology 181:6857–6864.

41. Murray HD, Schneider DA, Gourse RL. 2003. Control of rRNA Expression by Small Molecules Is Dynamic and Nonredundant. Molecular Cell 12:125–134.

42. Barker MM, Gaal T, Josaitis CA, Gourse RL. 2001. Mechanism of regulation of transcription initiation by ppGpp. I. Effects of ppGpp on transcription initiation in vivo and in vitro1. Journal of Molecular Biology 305:673–688.

43. Paul BJ, Barker MM, Ross W, Schneider DA, Webb C, Foster JW, Gourse RL. 2004. DksA: A Critical Component of the Transcription Initiation Machinery that Potentiates the Regulation of rRNA Promoters by ppGpp and the Initiating NTP. Cell 118:311–322.

44. Lennon CW, Gaal T, Ross W, Gourse RL. 2009. Escherichia coli DksA Binds to Free RNA Polymerase with Higher Affinity than to RNA Polymerase in an Open Complex. J Bacteriol 191:5854–5858.

45. Lee J-H, Lennon CW, Ross W, Gourse RL. 2012. Role of the Coiled-Coil Tip of Escherichia coli DksA in Promoter Control. Journal of Molecular Biology 416:503–517.

46. Ross W, Sanchez-Vazquez P, Chen AY, Lee J-H, Burgos HL, Gourse RL. 2016. ppGpp Binding to a Site at the RNAP-DksA Interface Accounts for Its Dramatic Effects on Transcription Initiation during the Stringent Response. Molecular Cell 62:811–823.

47. Xiao H, Kalman M, Ikehara K, Zemel S, Glaser G, Cashel M. 1991. Residual guanosine 3',5'-bispyrophosphate synthetic activity of relA null mutants can be eliminated by spoT null mutations. J Biol Chem 266:5980–5990.

48. Sanchez-Vazquez P, Dewey CN, Kitten N, Ross W, Gourse RL. 2019. Genome-wide effects on Escherichia coli transcription from ppGpp binding to its two sites on RNA polymerase. PNAS 201819682.

49. Durfee T, Hansen A-M, Zhi H, Blattner FR, Jin DJ. 2008. Transcription Profiling of the Stringent Response in Escherichia coli. Journal of Bacteriology 190:1084–1096.

50. Paul BJ, Berkmen MB, Gourse RL. 2005. DksA potentiates direct activation of amino acid promoters by ppGpp. PNAS 102:7823–7828.

51. Ross W, Vrentas CE, Sanchez-Vazquez P, Gaal T, Gourse RL. 2013. The Magic Spot: A ppGpp Binding Site on E. coli RNA Polymerase Responsible for Regulation of Transcription Initiation. Molecular Cell 50:420–429.

52. Molodtsov V, Sineva E, Zhang L, Huang X, Cashel M, Ades SE, Murakami KS. 2018. Allosteric Effector ppGpp Potentiates the Inhibition of Transcript Initiation by DksA. Molecular Cell 0.

53. Zuo Y, Wang Y, Steitz TA. 2013. The Mechanism of E. coli RNA Polymerase Regulation by ppGpp Is Suggested by the Structure of their Complex. Molecular Cell 50:430–436.

54. Rutherford ST, Villers CL, Lee J-H, Ross W, Gourse RL. 2009. Allosteric control of Escherichia coli rRNA promoter complexes by DksA. Genes Dev 23:236–248.

55. Chen J, Gopalkrishnan S, Chiu C, Chen AY, Campbell EA, Gourse RL, Ross W, Darst SA. 2019. E. coli TraR allosterically regulates transcription initiation by altering RNA polymerase conformation. eLife 8:e49375.

56. Gopalkrishnan S, Ross W, Chen AY, Gourse RL. 2017. TraR directly regulates transcription initiation by mimicking the combined effects of the global regulators DksA and ppGpp. PNAS 114:E5539–E5548.

57. Parshin A, Shiver AL, Lee J, Ozerova M, Schneidman-Duhovny D, Gross CA, Borukhov S. 2015. DksA regulates RNA polymerase in Escherichia coli through a

network of interactions in the secondary channel that includes Sequence Insertion 1. PNAS 112:E6862–E6871.

58. Lane WJ, Darst SA. 2010. Molecular Evolution of Multisubunit RNA Polymerases: Sequence Analysis. Journal of Molecular Biology 395:671–685.

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

60. Kriel A, Bittner AN, Kim SH, Liu K, Tehranchi AK, Zou WY, Rendon S, Chen R, Tu BP, Wang JD. 2012. Direct Regulation of GTP Homeostasis by (p)ppGpp: A Critical Component of Viability and Stress Resistance. Molecular Cell 48:231–241.

61. Liu K, Myers AR, Pisithkul T, Claas KR, Satyshur KA, Amador-Noguez D, Keck JL, Wang JD. 2015. Molecular Mechanism and Evolution of Guanylate Kinase Regulation by (p)ppGpp. Molecular Cell 57:735–749.

62. Anderson BW, Hao A, Satyshur KA, Keck JL, Wang JD. 2020. Molecular Mechanism of Regulation of the Purine Salvage Enzyme XPRT by the Alarmones pppGpp, ppGpp, and pGpp. Journal of Molecular Biology 432:4108–4126.

63. Anderson BW, Liu K, Wolak C, Dubiel K, She F, Satyshur KA, Keck JL, Wang JD. 2019. Evolution of (p)ppGpp-HPRT regulation through diversification of an allosteric oligomeric interaction. Elife 8.

64. Yang J, Anderson BW, Turdiev A, Turdiev H, Stevenson DM, Amador-Noguez D, Lee VT, Wang JD. 2020. The nucleotide pGpp acts as a third alarmone in Bacillus , with functions distinct from those of (p) ppGpp. 1. Nature Communications 11:5388.

65. Davis E, Chen J, Leon K, Darst SA, Campbell EA. 2015. Mycobacterial RNA polymerase forms unstable open promoter complexes that are stabilized by CarD. Nucleic Acids Res 43:433–445.

66. Boyaci H, Saecker RM, Campbell EA. 2019. Transcription initiation in mycobacteria: a biophysical perspective. Transcription 0:1–13.

67. Artsimovitch I, Svetlov V, Murakami KS, Landick R. 2003. Co-overexpression of Escherichia coliRNA Polymerase Subunits Allows Isolation and Analysis of Mutant Enzymes Lacking Lineage-specific Sequence Insertions. J Biol Chem 278:12344–12355.

68. Stallings CL, Stephanou NC, Chu L, Hochschild A, Nickels BE, Glickman MS. 2009. CarD is an essential regulator of rRNA transcription required for Mycobacterium tuberculosis persistence. Cell 138:146–159.

69. Srivastava DB, Leon K, Osmundson J, Garner AL, Weiss LA, Westblade LF, Glickman MS, Landick R, Darst SA, Stallings CL, Campbell EA. 2013. Structure and function of CarD, an essential mycobacterial transcription factor. PNAS 110:12619–12624.

70. Stallings CL, Glickman MS. 2011. CarD. Transcription 2:15–18.

71. Bae B, Chen J, Davis E, Leon K, Darst SA, Campbell EA. 2015. CarD uses a minor groove wedge mechanism to stabilize the RNA polymerase open promoter complex. eLife Sciences e08505.

72. Landick R, Krek A, Glickman MS, Socci ND, Stallings CL. 2014. Genome-wide mapping of the distribution of CarD, RNAP σA, and RNAP β on the Mycobacterium smegmatis chromosome using chromatin immunoprecipitation sequencing. Genom Data 2:110–113.

73. Gallego-García A, Mirassou Y, García-Moreno D, Elías-Arnanz M, Jiménez MA, Padmanabhan S. 2014. Structural Insights into RNA Polymerase Recognition and Essential Function of Myxococcus xanthus CdnL. PLOS ONE 9:e108946.

74. Rammohan J, Ruiz Manzano A, Garner AL, Stallings CL, Galburt EA. 2015. CarD stabilizes mycobacterial open complexes via a two-tiered kinetic mechanism. Nucleic Acids Res 43:3272–3285.

75. Galburt EA. 2018. The calculation of transcript flux ratios reveals single regulatory mechanisms capable of activation and repression. Proc Natl Acad Sci USA 115:E11604–E11613.

76. Zhu DX, Garner AL, Galburt EA, Stallings CL. 2019. CarD contributes to diverse gene expression outcomes throughout the genome of *Mycobacterium tuberculosis*. Proc Natl Acad Sci USA 116:13573–13581.

77. Furman R, Biswas T, Danhart EM, Foster MP, Tsodikov OV, Artsimovitch I. 2013. DksA2, a zinc-independent structural analog of the transcription factor DksA. FEBS Letters 587:614–619.

78. Sherlock ME, Sudarsan N, Breaker RR. 2018. Riboswitches for the alarmone ppGpp expand the collection of RNA-based signaling systems. PNAS 115:6052–6057.

79. Kaplan S. 2002. Photosynthesis genes and their expression in Rhodobacter sphaeroides 2.4.1: a tribute to my students and associates. Photosynthesis Research 73:95–108.

80. Mackenzie C, Eraso JM, Choudhary M, Roh JH, Zeng X, Bruscella P, Puskás Á, Kaplan S. 2007. Postgenomic Adventures with Rhodobacter sphaeroides. Annual Review of Microbiology 61:283–307.

81. Dryden SC, Kaplan S. 1990. Localization and structural analysis of the ribosomal RNA operons of Rhodobacter sphaeroides. Nucleic Acids Res 18:7267–7277.

82. Dryden SC, Kaplan S. 1993. Identification of cis-acting regulatory regions upstream of the rRNA operons of Rhodobacter sphaeroides. J Bacteriol 175:6392–6402.

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

84. Kansy JW, Kaplan S. 1989. Purification, characterization, and transcriptional analyses of RNA polymerases from Rhodobacter sphaeroides cells grown

chemoheterotrophically and photoheterotrophically. Journal of Biological Chemistry 264:13751.

85. Mackenzie C, Choudhary M, Larimer FW, Predki PF, Armitage JP, Barber RD, Donohue TJ, Hosler JP, Newman JE, Shapleigh JP, Sockett RE, Zeilstra-Ryalls J. The home stretch, a first analysis of the nearly completed genome of Rhodobacter sphaeroides 2.4. 23.

86. Zeilstra-Ryalls J, Gomelsky M, Eraso JM, Yeliseev A, O'Gara J, Kaplan S. 1998. Control of Photosystem Formation in Rhodobacter sphaeroides. J Bacteriol 180:2801–2809.

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

88. Eraso JM, Kaplan S. 1996. Complex regulatory activities associated with the histidine kinase PrrB in expression of photosynthesis genes in Rhodobacter sphaeroides 2.4.1. J Bacteriol 178:7037–7046.

89. Pemberton JM, Horne IM, McEwan AG. 1998. Regulation of photosynthetic gene expression in purple bacteria. Microbiology 144:267–278.

90. Choudhary M, Kaplan S. 2000. DNA sequence analysis of the photosynthesis region of Rhodobacter sphaeroides 2.4.1T. Nucleic Acids Res 28:862–867.

91. Metz S, Jäger A, Klug G. 2009. In Vivo Sensitivity of Blue-Light-Dependent Signaling Mediated by AppA/PpsR or PrrB/PrrA in Rhodobacter sphaeroides. J Bacteriol 191:4473–4477.

92. Eraso JM, Kaplan S. 2009. Half-Site DNA Sequence and Spacing Length Contributions to PrrA Binding to PrrA Site 2 of RSP3361 in Rhodobacter sphaeroides 2.4.1. J Bacteriol 191:4353–4364.

93. Hirvonen CA, Ross W, Wozniak CE, Marasco E, Anthony JR, Aiyar SE, Newburn VH, Gourse RL. 2001. Contributions of UP Elements and the Transcription Factor FIS to Expression from the Seven rrn P1 Promoters in Escherichia coli. J Bacteriol 183:6305–6314.

94. Karls RK, Wolf JR, Donohue TJ. 1999. Activation of the cycA P2 promoter for the Rhodobacter sphaeroides cytochrome c2 gene by the photosynthesis response regulator. Molecular Microbiology 34:822–835.

95. Cullen PJ, Kaufman CK, Bowman WC, Kranz RG. 1997. Characterization of the Rhodobacter capsulatus Housekeeping RNA Polymerase in vitro Transcription of Photosynthesis and Other Genes. J Biol Chem 272:27266–27273.

96. Forrest ME, Beatty JT. 1987. Purification of Rhodobacter capsulatus RNA polymerase and its use for in vitro transcription. FEBS Letters 212:28–34.

97. Janey Lee W, Bush JW, Chamberlin MJ. 1979. Utilization of promoter and terminator sites on bacteriophage T7 DNA by RNA polymerases from a variety of bacterial orders. Cell 16:97–109.

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

99. Campbell TB, Lueking DR. 1983. Light-mediated regulation of phospholipid synthesis in Rhodopseudomonas sphaeroides. J Bacteriol 155:806–816.

100. Eccleston Jr. ED, Gray ED. 1973. Variations in ppGpp levels in Rhodopseudomonas spheroides during adaptation to decreased light intensity. Biochemical and Biophysical Research Communications 54:1370–1376.

101. Lennon CW, Lemmer KC, Irons JL, Sellman MI, Donohue TJ, Gourse RL, Ross W. 2014. A Rhodobacter sphaeroides Protein Mechanistically Similar to Escherichia coli DksA Regulates Photosynthetic Growth. mBio 5.

102. Henry KK, Ross W, Myers KS, Lemmer KC, Vera JM, Landick R, Donohue TJ, Gourse RL. 2020. A majority of Rhodobacter sphaeroides promoters lack a crucial RNA polymerase recognition feature, enabling coordinated transcription activation. PNAS 117:29658–29668.

103. Vera JM, Ghosh IN, Zhang Y, Hebert AS, Coon JJ, Landick R. 2020. Genome-Scale Transcription-Translation Mapping Reveals Features of Zymomonas mobilis Transcription Units and Promoters. mSystems 5.

CHAPTER TWO

A majority of *Rhodoacter sphaeroides* promoters lack a crucial RNA polymerase recognition feature, enabling coordinated transcription activation

This chapter has been published (Kemardo K. Henry, Wilma Ross, Kevin S. Myers, Kimberly C. Lemmer, Jessica M. Vera, Robert Landick, Timothy J. Donohue, and Richard L. Gourse. *Proceedings of the National* Academy of Sciences. 117 (47) 29658-29668; DOI: 10.1073/pnas.2010087117.

I performed all the experiments except for Figure 6, which was contributed by KSM, KCL and JMV. The manuscript was written by WR, KSM, TJD, RLG and me and revised together with the other authors. Dataset 1, which can be accessed online (www.pnas.org/cgi/doi/10.1073/pnas.2010087117) is omitted for brevity.

Abstract

Using an in vitro transcription system with purified RNA polymerase (RNAP) to investigate rRNA synthesis in the photoheterotrophic α-proteobacterium *Rhodobacter sphaeroides*, we identified a surprising feature of promoters recognized by the major holoenzyme. Transcription from *R. sphaeroides* rRNA promoters was unexpectedly weak, correlating with the absence of -7T, the very highly conserved thymine found at the last position in -10 elements of promoters in most bacterial species. Thymine substitutions for adenine at position -7 in the 3 rRNA promoters strongly increased intrinsic promoter activity, indicating that *R. sphaeroides* RNAP can utilize -7T when present. rRNA promoters were activated by purified *R. sphaeroides* CarD, a transcription factor found in many bacterial species but not in β and γ proteobacteria. Overall, CarD increased the activity of 15 of the 16 native *R. sphaeroides* promoters tested in vitro that lacked -7T, whereas it had no effect on 3 of the 4 native promoters that contained -7T. Genome-wide bioinformatic analysis of promoters from *R. sphaeroides* and two other α-proteobacterial species indicated that 30-43% contained -7T, whereas 90-99% of promoters from non-α-proteobacteria contained -7T. Thus, promoters lacking -7T appear to be widespread in α-proteobacteria and may have evolved away from consensus to enable their coordinated regulation by transcription factors like CarD. We observed a strong reduction in *R. sphaeroides* CarD levels when cells enter stationary phase, suggesting that reduced activation by CarD may contribute to inhibition of rRNA transcription when cells enter stationary phase, the stage of growth when bacterial ribosome synthesis declines.

SIGNIFICANCE

Bacterial promoters are often predicted by similarity to the *Escherichia coli* -10 and -35 consensus elements. Although these elements are highly conserved in diverse bacterial phyla, only 30 to 43% of promoters we analyzed from *Rhodobacter sphaeroides* and two other α-proteobacteria contained -7T, a base in the -10 element present in 90-99% of promoters from non-α-proteobacteria. Expression from 15 of 16 native *R. sphaeroides* promoters we tested in vitro that lacked -7T was very weak, but we identified an essential transcription factor, CarD, that activated all 15 of these promoters. We suggest that promoters lacking a critical base in a consensus element can shape the transcriptome by coordinating expression of large numbers of genes by a single transcription factor.

INTRODUCTION

The α-proteobacteria are a gram-negative, metabolically diverse, biotechnologically important class of bacteria. Although regulation of transcription has been well-characterized in the γ-proteobacterium *Escherichia coli*, much less is known about transcription and its regulation in α-proteobacteria or in most other bacterial classes and phyla. *Rhodobacter sphaeroides* is a purple nonsulfur α-proteobacterium capable of aerobic and anaerobic respiration, photosynthesis, CO_2 and N_2, as well as H_2 and polyhydroxybutyrate production (1, 2). Aerobic growth is similar to that of other gram-negative bacteria. However, when oxygen levels decline, *R. sphaeroides* switches to photosynthetic growth, completely remodeling its intracellular membrane and creating pigments necessary to capture light energy (3). In addition to its use as a photosynthetic α-proteobacterial model system, *R. sphaeroides* has also been studied extensively for its potential in bio-based production of fuels and chemicals (1, 4, 5). A better understanding of its basic transcription properties would improve the usefulness of *R. sphaeroides* for bioproduction as well as our understanding of α-proteobacterial gene regulation in general.

In *E. coli*, the synthesis of the translation machinery is tightly regulated at the level of rRNA transcription initiation in order to ensure a sufficient number of ribosomes to support the cellular growth rate. Study of rRNA transcription from the promoters for the 7 *E. coli* rRNA operons has led to many general insights about transcription and its regulation, including the discovery that the nucleoid protein Fis is an important transcription factor (6), that the α subunit of RNAP is a DNA binding protein that

contributes to specific promoter recognition (7), and that DksA is a transcription factor that functions in conjunction with the second messenger ppGpp to regulate large numbers of bacterial promoters (8, 9). Therefore, we focused on studying the properties of *R. sphaeroides* rRNA promoters to obtain insights about the mechanism of transcription initiation and its control in α-proteobacteria.

In contrast to *E. coli*, *R. sphaeroides* 2.4.1 has two chromosomes and only 3 rRNA operons, *rrnA* on chromosome 1 and *rrnB* and *rrnC* on chromosome 2. The transcription start sites for the *rrnA*, *rrnB,* and *rrnC* promoters were mapped previously by primer extension (10, 11). In several bacterial species, including *Mycobacterium tuberculosis*, *Thermus thermophilus*, and *Myxococcus xanthus*, rRNA promoters are activated in vitro by CarD, a small protein that binds to the lobe of the RNAP β subunit and interacts with promoter DNA just upstream of the -10 hexamer (12-16). In the α-proteobacterium *Caulobacter crescentus*, the CarD homolog localizes to rRNA promoters and depletion decreases rRNA transcription (17). (The CarD homologs in *M. xanthus* and *C. crescentus* are sometimes referred to as CdnL.)

CarD family members are found in the *Actinomycetes, Firmicutes*, *Deinococcus / Thermus*, *Spirochaetes*, *δ-proteobacteria*, and in most classes of α-proteobacteria, but are not found in β- and γ-proteobacteria (13, 14, 17). Although *carD* is essential in some species and its depletion or deletion affects expression of many genes in *M. tuberculosis* and *C. crescentus* (12, 17-19), a direct, widespread role for CarD in transcription from non-rRNA promoters has not been demonstrated previously with purified components in vitro.

We report here that *R. sphaeroides* rRNA promoters are activated by $CarD_{Rsp}$ in vitro, at least in part because these promoters lack a critical thymine at the last position on the non-template strand of the -10 hexamer (-7T) that is present in almost all *E. coli* promoters recognized by the major holoenzyme. The T at -7 in *E. coli* promoters fits tightly in a pocket in σ^{70}, an interaction that is critical for transcription initiation (20). Not only do *R. sphaeroides* rRNA promoters utilize $CarD_{Rsp}$ to compensate, at least in part, for the absence of this -7T interaction with σ, but we also show here that many other *R. sphaeroides* promoters are activated by $CarD_{Rsp}$. In fact, bioinformatic analyses indicate that the majority of promoters in *R. sphaeroides*, as well as in two other α-proteobacterial species analyzed, lack a T at the last position in the -10 element, suggesting that many promoters in α-proteobacteria may also utilize CarD to increase transcription initiation.

RESULTS

R. sphaeroides rRNA promoters initiate poorly in vitro with *R. sphaeroides* RNAP.

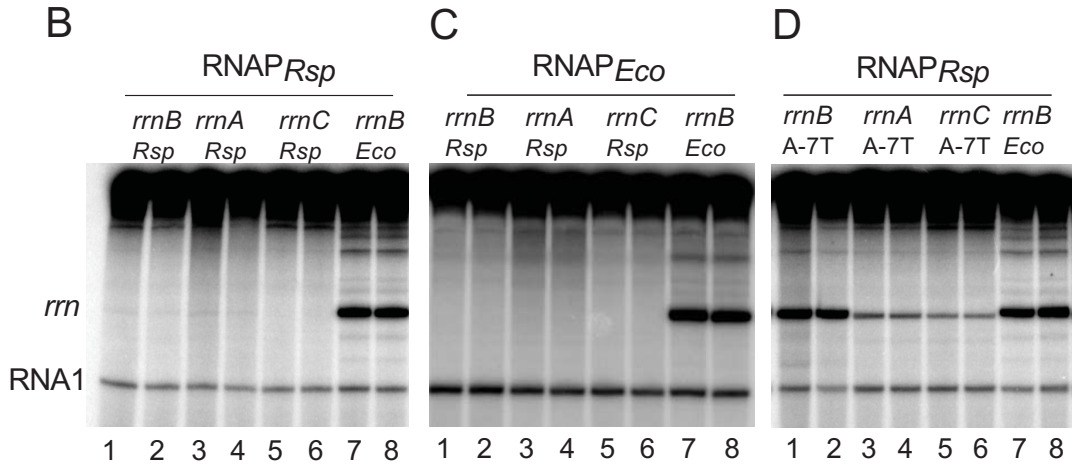
To purify *R. sphaeroides* RNAP for analysis of regulation of *R. sphaeroides* rRNA promoters in vitro, we inserted sequences coding for a C-terminal 10X histidine tag into the *rpoC* gene in the *R. sphaeroides* chromosome by homologous recombination, then purified *R. sphaeroides* RNAP ($RNAP_{Rsp}$) by nickel-affinity chromatography (*SI Appendix*, Figure 2.S1*A*; *Expanded Materials and Methods*). The resulting purified $RNAP_{Rsp}$ preparation contained proteins of the sizes expected for the β, β', α, and ω subunits, as well as the major σ subunit, σ^{93}. The identity of σ^{93} was confirmed by Western blotting with a polyclonal anti-*E. coli* σ^{70} antibody that cross-reacted with *R. sphaeroides* σ^{93} (*SI Appendix*, Figure 2.S1*B*). Unless otherwise indicated, we use $RNAP_{Rsp}$ to refer to the holoenzyme containing σ^{93}.

In vitro transcription with the 3 *R. sphaeroides* rRNA promoters was carried out with supercoiled plasmids containing rRNA promoter fragments (non-template strand sequences shown from -57 to +1, with respect to the transcription start sites, in Figure 2.1*A*) inserted into plasmid pRLG770 (6). Transcript lengths of ~200 nt were predicted, based on termination at the *E. coli rrnB* T1 terminator sequence downstream of the promoter fragment insertion site in the plasmid. Few transcripts of this size were detected from the *R. sphaeroides rrnA, rrnB*, *and rrnC* promoters with $RNAP_{Rsp}$ (Figure 2.1*B*, lanes 1-6) or $RNAP_{Eco}$ (Figure 2.1*C*, lanes 1-6) under conditions where

Figure 2.1: Activity of *R. sphaeroides* rRNA promoters with either *R. sphaeroides* RNAP or *E. coli* RNAP. (*A*) Sequences of the 3 *R. sphaeroides* rRNA promoters, *rrnA*, *rrnB*, *rrnC* from -57 to the transcription start site, +1, and the *E. coli rrnB* P1 promoter. The Discriminator (Dis) region, T-tract sequence, and UP element are indicated, in addition to the -10 and -35 RNAP recognition hexamers and the transcription start site which are in bold. The last bp in the -10 element is referred to as the "-7" position (boxed) although it is 8 or 9 bp rather than 7 bp upstream from the transcription start site due to the larger than consensus number of bp between the -10 element and the TSS in these promoters. (*B*) In vitro transcription of the indicated rRNA promoters from plasmid templates with *R. sphaeroides* RNAP in buffer containing 170 mM NaCl (see *SI Appendix, Expanded Materials and Methods*). Duplicate lanes are shown for each promoter. The RNA I promoter and transcript are part of the plasmid replication control system (see *SI Appendix*, *Expanded Materials and Methods*). (*C*) In vitro transcription of the indicated promoters as in (*B*), but with *E. coli* RNAP. (D) In vitro transcription with *R. sphaeroides* RNAP as in (*B*), but with the A-7T promoter variants of the 3 *R. sphaeroides* rRNA promoters or with the wild type *E. coli rrnB* P1 promoter. A higher concentration of *R. sphaeroides* RNAP (50 nM) than *E. coli* RNAP (10 nM) was used to ensure that the absence of transcription from the *R. sphaeroides* rRNA promoters was not a result of limiting $RNAP_{Rsp}$. Robust, approximately equivalent transcription was observed from the *E. coli rrnB* promoter at 50 nM $RNAP_{Rsp}$ and 10 nM $RNAP_{Eco}$ (Figures 2.1B-D).

A

-35
-10
UP Element
T-tract
"-7"
Dis
+1
GTCTCTTCGTCATTTTTCCTCTTGCGGGTTTTTTTGCGGTTCCCTAGATAGCGCCTCA Rsp rrnA
TACGGAGCCCAAAAAATCCGCTTGCGCCCGGGGCCGTCTGCTCCTAGAAACCGCTTCA Rsp rrnB
ATGCGGCTGTCATTTTTCCTCTTGCGGGTTTTTTTGCGGTTCCCTAGATAGCGCCTCA Rsp rrnC
AAAATTATTTTAAATTTCCTCTTGTCAG-GCCGGAATAACTCCCTATAATGCGCCACCA Eco rrnB P1


B
C
D
RNAP Rsp
RNAP Eco
RNAP Rsp
rrnB Rsp
rrnA Rsp
rrnC Rsp
rrnB Eco
rrnB A-7T
rrnA A-7T
rrnC A-7T
rrnB Eco
rrn
RNA1
1 2 3 4 5 6 7 8
1 2 3 4 5 6 7 8
1 2 3 4 5 6 7 8

transcription was detected from the RNA I promoter, part of the replication control system encoded by the plasmid templates (Figure 2.1*B* and *C*). In contrast, an abundant transcript of the expected size was detected in reactions containing the *E. coli rrnB* P1 promoter and either *R. sphaeroides or E. coli* RNAP (Figure 2.1*B* and *C*, lanes 7-8). $RNAP_{Rsp}$ also produced transcripts from two mutant *E. coli* or phage promoters that were tested, *lacUV5* and T7A1*lacO34* (*SI Appendix*, Figure 2.S2), consistent with the conclusion that the lack of transcription from the *R. sphaeroides* rRNA promoters was not a result of inactivity of the purified $RNAP_{Rsp}$ holoenzyme.

The inactivity of the *R. sphaeroides* rRNA promoters was unexpected given that rRNA promoters are very active in moderate to fast growing bacterial cells (11, 21). In addition, the *R. sphaeroides* rRNA promoters contain several features characteristic of *E. coli* rRNA promoters (11), including a TTG sequence (non-template strand) at the upstream end of a putative -35 element, a TA sequence at the upstream end of a likely -10 element, an A+T-rich UP element-like sequence upstream of the -35 element (22), and a G+C-rich sequence immediately downstream from the -10 element, corresponding to the discriminator sequence in *E. coli* rRNA promoters (23) (Figure 2.1*A*). The high A+T content of the UP element-like sequence and the 7 nt T-tract in the spacer region of *rrnA* and *rrnC* were especially striking given the ~70% G+C content of the *R. sphaeroides* genome (1).

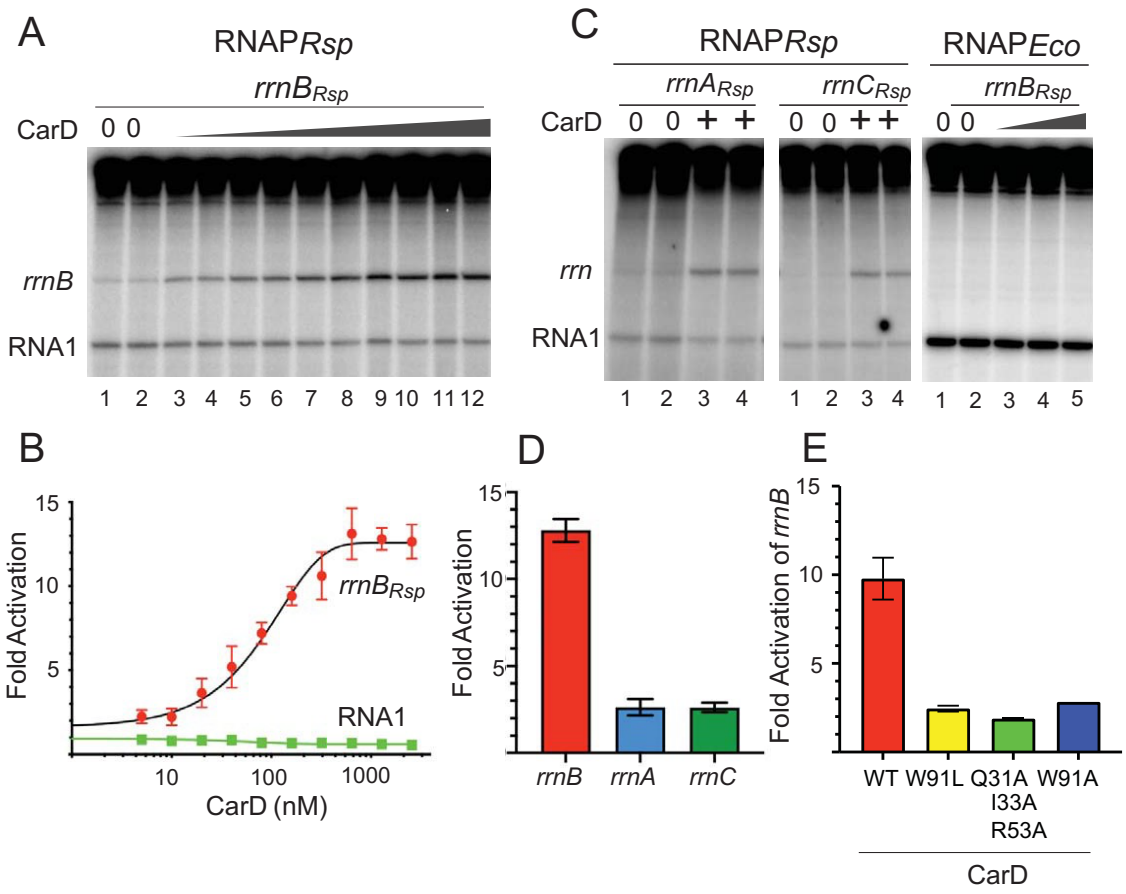
In contrast to *E. coli* rRNA promoters (and most other $E\sigma^{70}$-dependent *E. coli* promoters), each of the *R. sphaeroides* rRNA promoters contains an A rather than the highly conserved T at the last position in the -10 element (designated here as position -7) (Figure 2.1*A*). In fact, the predicted rRNA promoters from other sequenced *R.*

sphaeroides strains also lack a T at -7 (Figure 2.S3). The significance of the conserved T at -7 for Eσ^{70}-dependent promoter activity in *E. coli* was demonstrated long ago by genetic analyses and by structural studies showing that only the thymine, and not other bases at position -7, is compatible with binding in a pocket in σ^{70} (20, 24-26). The σ residues that contribute to the -7 pocket are conserved in the major $RNAP_{Rsp}$, consistent with the ability of $RNAP_{Rsp}$ to recognize *E. coli* promoters (*SI Appendix*, Figure 2.S4*A*). Also consistent with the absence of -7T as a determinant of the low activity of the *R. sphaeroides* rRNA promoters in vitro, we found that a -7T substitution increased transcription from each of the rRNA promoters with $RNAP_{Rsp.}$ (Figure 2.1*D*), although the mutated *rrnA*(A-7T) and *rrnC*(A-7T) promoters were much weaker than *rrnB*(A-7T) or *E. coli rrnB* P1. Taken together, these results suggested that *R. sphaeroides* rRNA promoters might require a transcription activator for function.

***R. sphaeroides* CarD activates rRNA promoters.** Previous studies have reported that the CarD protein from *M. tuberculosis, M. smegmatis, T. thermophilus*, and *M. xanthus* stimulates rRNA transcription in vitro (12, 13, 27, 28). Among the proteobacteria, CarD is found in the α and δ, but not in the β and γ classes (i.e. not in *E. coli*) (13, 14). The *R. sphaeroides* CarD homolog (*rsp*_2425) is similar in size (169 amino acids), is 25-60% identical to CarD proteins from the other bacteria (*SI Appendix*, Figure 2.S5), and is essential (5), consistent with a role in one or more critical functions including rRNA transcription.

To test this idea, we purified *R. sphaeroides* CarD after overexpression in *E. coli* and measured its effects on *R. sphaeroides* rRNA promoters in vitro (Figure 2.2*A*).

Figure 2.2: **Activities of the *R. sphaeroides* rRNA promoters in vitro in with or without purified *R. sphaeroides* CarD.** (*A*) In vitro transcription of the *R. sphaeroides rrnB* promoter with 20 nM *R. sphaeroides* RNAP with or without $CarD_{Rsp}$ (wedge indicates $CarD_{Rsp}$ range of 5-2,560 nM) in buffer with 170 mM NaCl. (*B*) Fold-activation of the *R. sphaeroides rrnB* promoter by $CarD_{Rsp}$ (with/without the indicated concentration of $CarD_{Rsp}$) from experiments like that in (*A*). The RNA I transcript is from a plasmid-encoded promoter. Error bars indicate standard deviation from 3 separate assays. (*C*) In vitro transcription of the *R. sphaeroides rrnA* or *rrnC* promoters with *R. sphaeroides* RNAP as in (*A*) with or without 1,280 nM CarD or of the *R. sphaeroides rrnB* promoter with 10 nM *E. coli* RNAP with or without 20, 40, or 80 nM CarD. (D) Fold-activation of the *R. sphaeroides rrnA, rrnB,* and *rrnC* promoters by 1,280 nM CarD with standard deviation from 3 assays for each promoter. (*E*) Fold-activation of the *R. sphaeroides rrnB* promoter by $RNAP_{Rsp}$ and 1,280 nM wild-type CarD or W91A, W91L, or Q31A/I33A/R53A triple mutant CarD (standard deviation from 3 assays). Wild-type CarD used in *(B*) and (*E*) was from different protein preparations, so differences in specific activities could account for the slight differences in fold-activation.


A
RNAP*Rsp*
rrnB$_{Rsp}$
CarD 0 0
rrnB
RNA1
1 2 3 4 5 6 7 8 9 10 11 12
C
RNAP*Rsp*
RNAP*Eco*
rrnA$_{Rsp}$
rrnC$_{Rsp}$
rrnB$_{Rsp}$
CarD 0 0 + + 0 0 + + 0 0
rrn
RNA1
1 2 3 4 1 2 3 4 1 2 3 4 5
B
Fold Activation
15
10
5
rrnB$_{Rsp}$
RNA1
10 100 1000
CarD (nM)
D
Fold Activation
15
10
5
rrnB *rrnA* *rrnC*
E
Fold Activation of *rrnB*
15
10
5
WT W91L Q31A I33A R53A W91A
CarD

$CarD_{Rsp}$ activated transcription from the *R. sphaeroides rrnB* promoter at least 10-fold in vitro, with a half-maximal stimulatory effect (EC_{50}) at 85 nM (Figure 2.2*B*). When added to $RNAP_{Rsp}$, $CarD_{Rsp}$ also activated the *R. sphaeroides rrnA* and *rrnC* promoters, although transcription from *rrnA* and *rrnC* was weaker than from *rrnB*, and the fold-activation by $CarD_{Rsp}$ was smaller (Figure 2.2*C* and *D*). These results are consistent with a previous report that the *R. sphaeroides rrnB* promoter was stronger than the *rrnA* promoter in vivo, based on measurements of promoter-*xylE* fusions (11).

To determine whether the 7 nt T-tract found in the spacer region of *R. sphaeroides rrnA* (and *rrnC*, which is identical to *rrnA*), but not in *rrnB*, could account for the promoter activity difference, we constructed a derivative of *rrnA* containing a replacement of the T-tract with the corresponding spacer sequence of *rrnB*. Replacement of the T-tract resulted in a 3-fold increase in basal *rrnA* promoter activity but did not alter the extent of activation by CarD (2- to 3-fold) (Figure 2.2 and *SI Appendix*, Figure 2.S6). Thus, the T-tract accounts, in part, for the reduced activities of *rrnA* and *rrnC* but does not account for the difference in activation by CarD (2- to 3-fold for *rrnA* vs 10 to 12-fold for *rrnB*).

$CarD_{Rsp}$ did not increase transcription from the plasmid-encoded RNA I promoter when added to $RNAP_{Rsp}$ (Figure 2.2*A*-*C*) or from *R. sphaeroides rrnB* using *E. coli* RNAP (Figure 2.2*C*). The failure of $CarD_{Rsp}$ to stimulate *E. coli* RNAP is consistent with previous reports that effects of CarD are species-specific because the interacting residues in the CarD RID domain and in the β subunit of RNAP are not highly conserved (28, 29).

We note that a low level of *R. sphaeroides rrnB* transcript was detected in a previous study with *R. sphaeroides* RNAP (30). Based on the results reported here, we suggest that the observed activity may have resulted from the presence of small amounts of CarD in the RNAP used in the previous study. In fact, using an antibody raised against $CarD_{Rsp}$ (see below), we detected a very low level of CarD in our purified *R. sphaeroides* RNAP (~0.01 mol $CarD_{Rsp}$ per mol RNAP). In our in vitro transcription experiments, the concentration of CarD introduced from $RNAP_{Rsp}$ was well below the EC_{50} determined for CarD (85 nM) but perhaps sufficient to account for the previously reported transcription from *rrnB* (*SI Appendix*, Figure 2.S1*C*, lanes 7-8).

To address whether the mechanism of activation by $CarD_{Rsp}$ is similar to that by $CarD_{Mtb,}$ we created substitutions in $CarD_{Rsp}$ at positions analogous to those previously shown to affect the DNA interaction or RNAP interaction domain (RID) activities of $CarD_{Mtb}$ (13, 14). $CarD_{Mtb}$ interacts with DNA at the upstream edge of the -10 hexamer using residue W85 to help wedge open the -10 hexamer. Alignment of the amino acid sequences from *Rhodobacter*, *Mycobacterium*, *Thermus*, *Myxococcus*, and *Caulobacter* CarD homologs indicated that $CarD_{Rsp}$-W91 corresponds to $CarD_{Mtb}$-W85 (14, 31) (*SI Appendix,* Figure 2.S5). Substitutions for $CarD_{Rsp}$-W91 ($CarD_{Rsp}$-W91L or W91A) greatly reduced activation of *R. sphaeroides rrnB* transcription, decreasing the 10-fold effect observed with wild-type $CarD_{Rsp}$ to only ~2.5-fold (Figure 2.2*E* and *SI Appendix* Figure 2.S5), a partial defect similar to that observed for $CarD_{Mtb}$ W85A and $CarD_{Tth}$ W86A (14). Conserved basic residues in CarD also interact with the DNA near W86 in the *T. thermophilus* CarD-RNAP-promoter DNA complex structure (13), suggesting that the partial activity of the W91A or W91L $CarD_{Rsp}$ proteins can be attributed to DNA

interactions that are weakened but not eliminated, rather than to general folding defects. A $CarD_{Rsp}$ variant with 3 substitutions, Q31A, I33A, R53A, corresponding to *T. thermophilus* CarD RID domain residues R25, I27, and R47, activated $rrnB_{Rsp}$ transcription only ~2-fold (Figure 2.2*E* and *SI Appendix*, Figure 2.S5). We tentatively conclude that $CarD_{Rsp}$ uses a mechanism similar to that of previously characterized CarD proteins.

CarD directly activates other *R. sphaeroides* promoters. Depletion of CarD from *M. smegmatis* and *M. tuberculosis* in vivo (12, 18), or deletion of the *C. crescentus* homolog (19), suggested that CarD affects gene expression broadly, although the approaches did not distinguish direct from indirect effects. CarD co-localized with *M. smegmatis* RNAP in ChIP-seq experiments at all promoters recognized by the primary σ factor in vivo (14, 32). Two *M. tuberculosis* promoters, the AP3 rRNA promoter and the VapB promoter, were also shown to be activated by CarD in vitro (16).

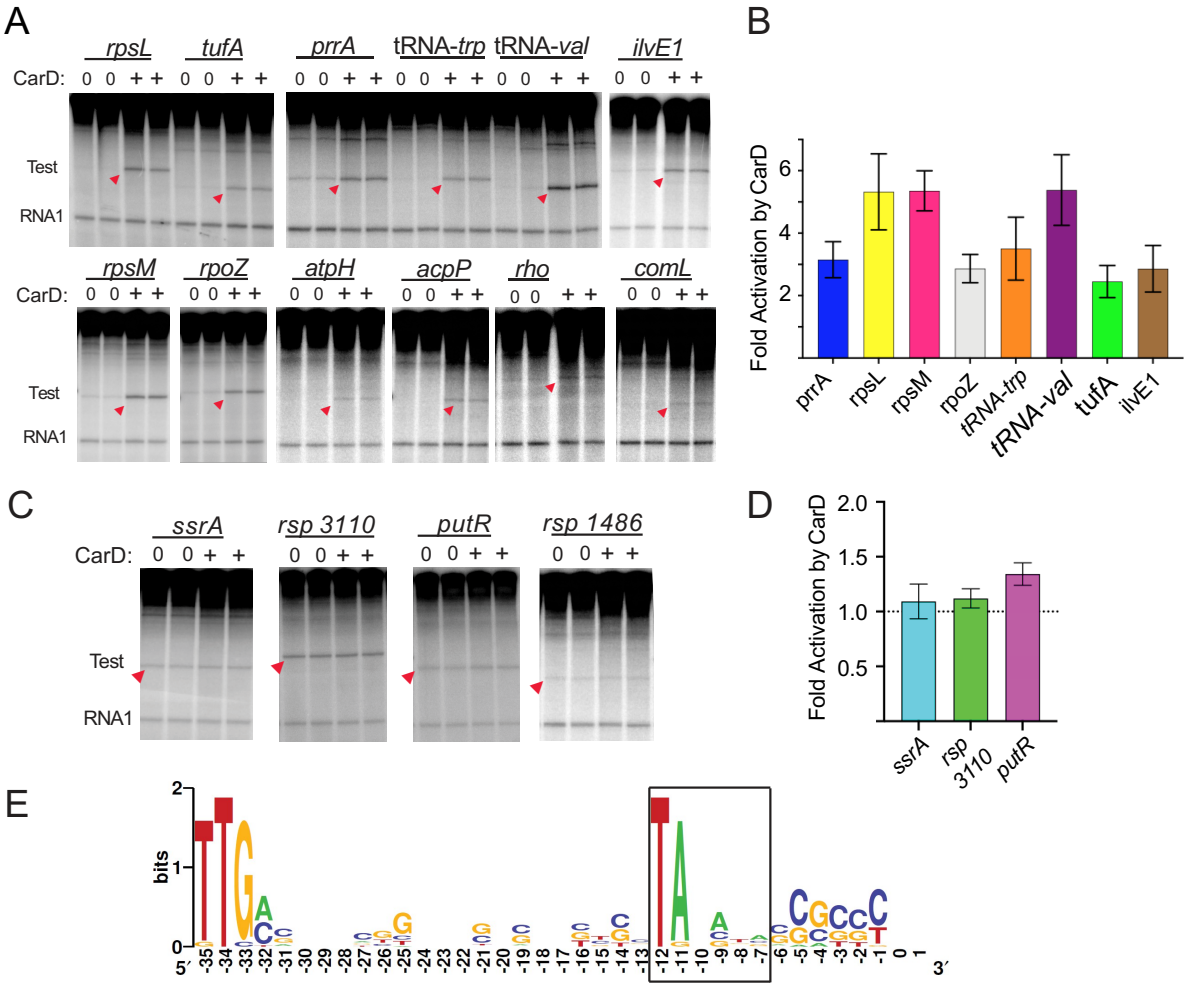
To evaluate whether CarD has a more general role in *R. sphaeroides* gene expression, we tested its direct effects on transcription in vitro from a set of previously uncharacterized *R. sphaeroides* promoters. Sixteen additional *R. sphaeroide*s promoters were chosen based on several criteria: genome-wide ChIP signals for σ^{93} from an *R. sphaeroides* ChIP-seq dataset (3), proximity to mapped *R. sphaeroides* transcription start sites (TSSs) (33), and sequence similarity to elements of an *E. coli* $E\sigma^{70}$ promoter consensus (25). We note that 13 of the 16 non-rRNA promoters lacked -7T (*SI Appendix*, Figure 2.S7); even though they were not chosen based on that criterion. PCR-derived fragments containing sequences from ~150 bp upstream to 30-

50 bp downstream of the proposed TSSs were cloned into pRLG770, the same plasmid described above for measuring transcription from rRNA promoters, and promoter activities were determined with and without 1280 nM purified $CarD_{Rsp}$.

The sizes of the transcripts produced in these assays relative to standards of known length confirmed the predicted locations of the promoters (Figure 2.3 and *SI Appendix,* Figure 2.S7). As with the rRNA promoters, $CarD_{Rsp}$ increased the activities of 12 of the 16 non-rRNA promoters. The 12 non-rRNA promoters that were activated by $CarD_{Rsp}$ included two for ribosomal protein operons (*rpsM* and *rpsL*), two for tRNAs ($tRNA^{trp}$ and $tRNA^{val}$), *rpoZ* (ω subunit of RNAP), *tufA* (translation factor EF-Tu), *prrA* (transcription factor PrrA), *rho* (transcription termination factor Rho), *ilvE1* (a putative aminotransferase), *acpP* (acyl carrier protein), *comL* (branched chain amino acid BamD subunit), and *atpH* (ATP synthase F1 δ subunit) (Figure 2.3*A*). For the promoters whose basal activities were high enough to quantify, activation by $CarD_{Rsp}$ ranged from 2- to almost 6-fold (Figure 2.3*B*).

Four of the 16 non-rRNA promoters, including *ssrA* (tmRNA), *putR* (a putative transcription factor), *rsp_3110* (a putative glutathione S-transferase), and *rsp_1486* (tetR family homolog) were unaffected by $CarD_{Rsp}$ (Figure 2.3*C*). $CarD_{Rsp}$ increased transcription less than 1.4-fold from three promoters whose basal activities were high enough to quantify the fold-effect (Figure 2.3*D*). In addition, three of the four *E. coli* promoters tested with $RNAP_{Rsp}$ were unaffected by $CarD_{Rsp}$ (*lacUV5, T7A1lacO34*, RNA I; *SI Appendix,* Figure 2.S2). Thus, positive regulation by $CarD_{Rsp}$ is promoter-specific, not limited to rRNA promoters, and not all promoters are dependent on $CarD_{Rsp}$ for activity.

Figure 2.3: Activation of *R. sphaeroides* non-rRNA promoters in vitro with $CarD_{Rsp}$. (*A*) Transcription in vitro with $RNAP_{Rsp}$ (20 nM) and $CarD_{Rsp}$ (1,280 nM) in buffer with 100 mM NaCl (see *SI Appendix*, *Expanded Materials and Methods*). Duplicate lanes with or without CarD are shown for each promoter. Transcripts derived from test promoters are indicated with red arrowheads. Gel images also show position of RNA 1 transcript from the plasmid promoter. See *SI Appendix*, Figure 2.S7 for promoter sequences. (*B*) Fold-activation by CarD (+CarD/no CarD) for promoters shown in (*A*) whose activities were high enough to quantify accurately. Standard deviations are shown from 3-6 assays except for *ilvE1* (range from 2 assays). *(C*) Transcription in vitro of 4 *R. sphaeroides* promoters whose activities were unaffected by CarD under the same conditions as in (*B*). See *SI Appendix*, Figure 2.S7 for promoter sequences. (*D*) Fold-activation of promoters in (*C*) with or without CarD. The standard deviations are shown are from 5 assays for *ssrA* or the ranges are shown from 2 assays for *rsp_3110* and *putR*. (*E*) WebLogo (34) representation of consensus sequence for 15 promoters whose transcription was activated by CarD in vitro (3 rRNA promoters and the 12 promoters shown in (*A*). -10 hexamer (-12 to -7) is boxed.

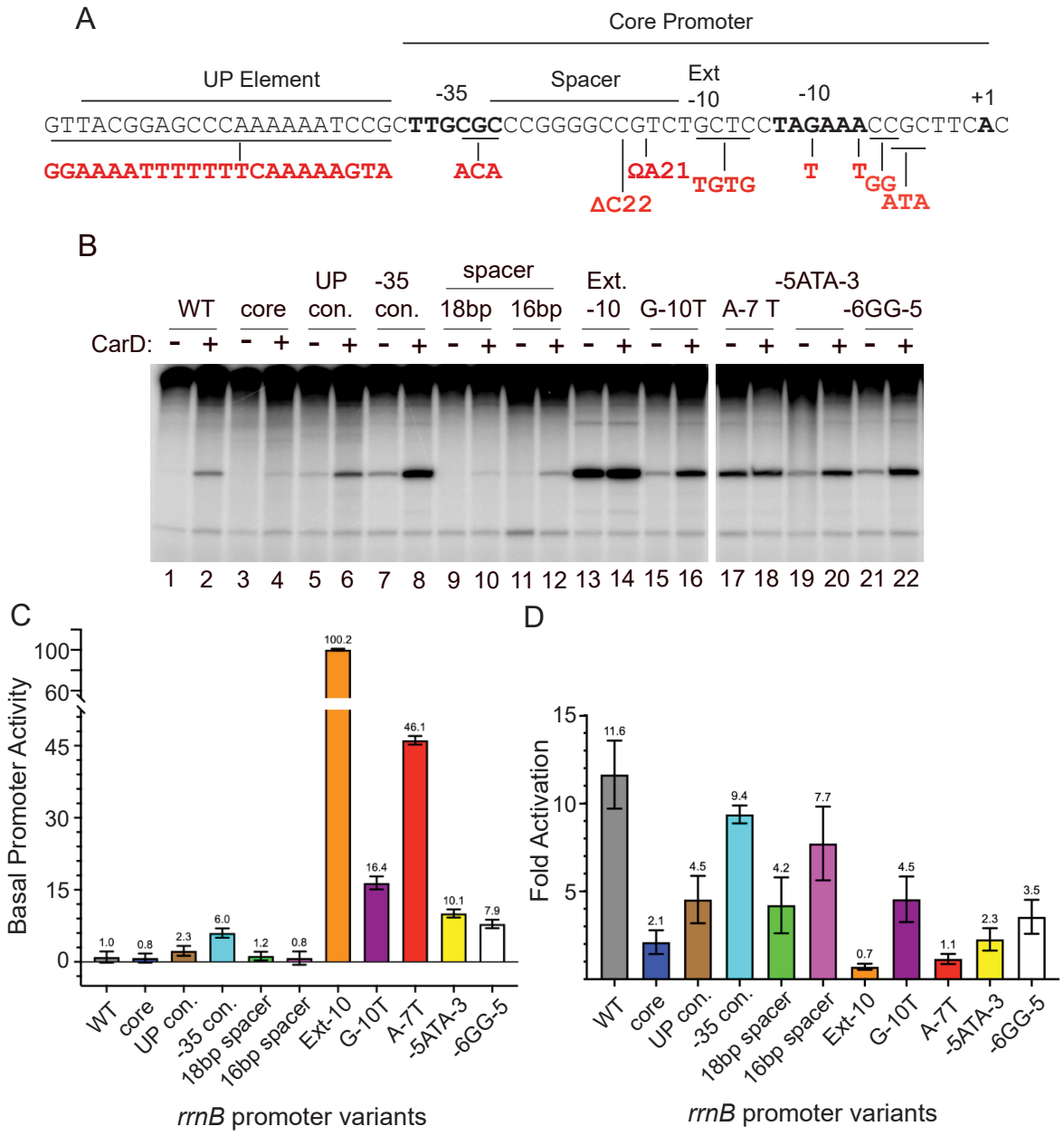

A
rpsL
tufA
prrA
tRNA-trp
tRNA-val
ilvE1
CarD:
0 0 + + 0 0 + + 0 0 + + 0 0 + + 0 0 + + 0 0 + +
Test
RNA1
rpsM
rpoZ
atpH
acpP
rho
comL
CarD:
0 0 + + 0 0 + + 0 0 + + 0 0 + + 0 0 + + 0 0 + +
Test
RNA1
B
Fold Activation by CarD
6
4
2
prrA
rpsL
rpsM
rpoZ
tRNA-trp
tRNA-val
tufA
ilvE1
C
ssrA
rsp 3110
putR
rsp 1486
CarD:
0 0 + + 0 0 + + 0 0 + + 0 0 + +
Test
RNA1
D
Fold Activation by CarD
2.0
1.5
1.0
0.5
ssrA
rsp 3110
putR
E
bits
2
1
0
5′
-35 -34 -33 -32 -31 -30 -29 -28 -27 -26 -25 -24 -23 -22 -21 -20 -19 -18 -17 -16 -15 -14 -13 -12 -11 -10 -9 -8 -7 -6 -5 -4 -3 -2 -1 0 1
3′

Promoter features that correlate with increased transcription by CarD$_{Rsp}$. A logo depicting the degree of sequence conservation at each promoter position in the 15 *R. sphaeroides* promoters activated by CarD$_{Rsp}$ (the 12 above plus the 3 rRNA promoters) is shown in Figure 2.3*E* and promoter sequences are shown in *SI Appendix*, Figure 2.S7. None of the 15 promoters activated by CarD$_{Rsp}$ contained a T at -7. Each of the other three bases at this position (A, G, and C) was represented in the activated promoter population, suggesting that the absence of -7T rather than the presence of another specific base was the relevant determinant for activation by CarD$_{Rsp}$. In contrast, three of the four *R. sphaeroides* promoters unaffected by CarD$_{Rsp}$ contained a T at -7 (Figure 2.3*C* and *SI Appendix*, Figure 2.S7).

To address whether sequence features in addition to the absence of -7T might also contribute to activation by CarD$_{Rsp}$, we constructed *R. sphaeroides rrnB* promoter variants with increased or decreased similarity to the *E. coli* consensus for the UP element, -35 hexamer, extended -10 element, -10 hexamer, discriminator region, and -10/-35 spacer length (35) (Figure 2.4*A*). Transcription activities of the variant promoters were compared to the wild-type promoter in the absence and presence of 1280 nM CarD$_{Rsp}$ in vitro (Figure 2.4*B*).

We did not detect changes in basal promoter activity resulting from replacement of the *rrnB* UP element region with either a non-UP element-like sequence or with the *E. coli* UP element consensus, because transcription was too weak for quantitation (Figure 2.4*B* and *C*). However, in the presence of CarD$_{Rsp}$, it was apparent that either the native UP element sequence or the consensus UP element increased transcription (Figure

Figure 2.4: Effects of individual *R. sphaeroides rrnB* promoter mutations on activation by CarD. (*A*) Promoter recognition elements (UP element, -35 hexamer, spacer, extended -10, -10 hexamer, and core promoter) are indicated above the *R. sphaeroides rrnB* promoter sequence. Mutations to create promoter variants are indicated in red (see also *SI Appendix*, Tables S2 and S3). Each mutation creates the *E. coli* σ^{70} consensus element sequence at the indicated position, except the spacer insertion or deletion for which the wild-type *R. sphaeroides* promoter has the *E. coli* 17 bp consensus spacing. Triple mutation -5 ATA -3 alters the base composition in the discriminator region. (*B*) In vitro transcription of each promoter mutant with $RNAP_{Rsp}$ (20 nM) with or without CarD (1,280 nM) in buffer with 170 mM NaCl. (*C*) Basal activities of mutant *rrnB* promoters relative to transcription from the wild-type promoter, with standard deviations from 3-6 assays for each promoter. (*D*). Fold-activation of each promoter by CarD (+CarD/no CarD). Standard deviations are from 3-6 assays for each promoter.

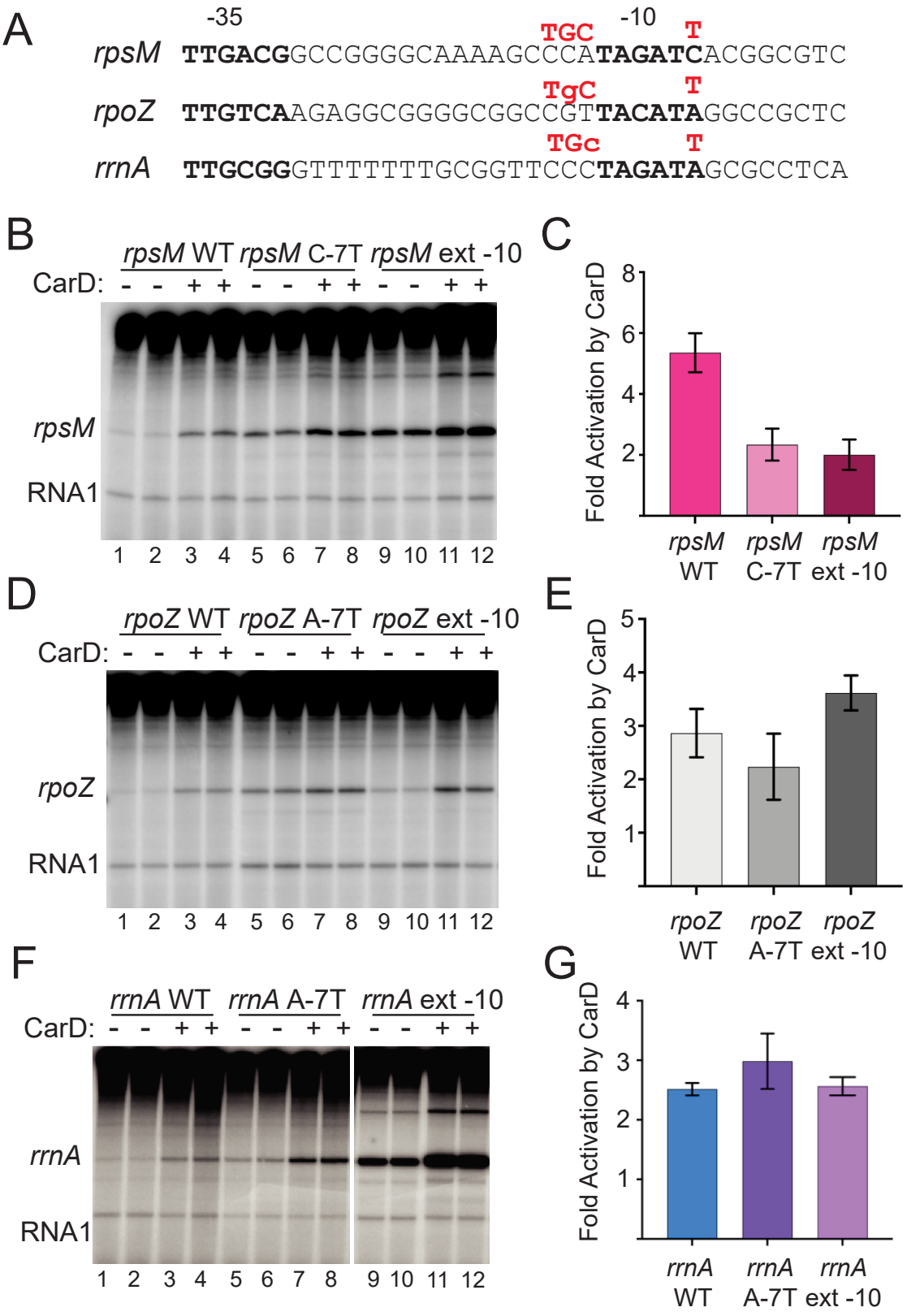

A
Core Promoter
UP Element
-35
Spacer
Ext -10
-10
+1
GTTACGGAGCCCAAAAAATCCGCTTGCGCCCGGGGCCGTCTGCTCCTAGAAACCGCTTCAC
GGAAAATTTTTTTCAAAAAGTA
ACA
ΩA21
ΔC22
TGTG
T
T
GG
ATA
B
WT
core
UP con.
-35 con.
spacer
18bp
16bp
Ext. -10
G-10T
-5ATA-3
A-7 T
-6GG-5
CarD: - + - + - + - + - + - + - + - + - + - + - + - +
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22
C
Basal Promoter Activity
100
60
45
30
15
0
1.0
0.8
2.3
6.0
1.2
0.8
100.2
16.4
46.1
10.1
7.9
WT
core
UP con.
-35 con.
18bp spacer
16bp spacer
Ext-10
G-10T
A-7T
-5ATA-3
-6GG-5
rrnB promoter variants
D
Fold Activation
15
10
5
11.6
2.1
4.5
9.4
4.2
7.7
0.7
4.5
1.1
2.3
3.5
WT
core
UP con.
-35 con.
18bp spacer
16bp spacer
Ext-10
G-10T
A-7T
-5ATA-3
-6GG-5
rrnB promoter variants

2.4*B* and *D*), consistent with conservation of the DNA binding residues in the *R. sphaeroides* and *E. coli* α subunit C-terminal domains (22) (*SI Appendix*, Figure 2.S4*B*).

Substitutions toward consensus for recognition by RNAP in the -35 hexamer, -10 hexamer, and discriminator region increased basal *R. sphaeroides rrnB* promoter activity (Figure 2.4*B* and *C*; compare lane 1 with lanes 7, 13, 16, 17, 19, and 21). Not surprisingly since the wild-type *R. sphaeroides rrnB* promoter already has consensus -10/-35 spacing (17 bp), changing the spacing to either 16 bp or 18 bp did not increase basal promoter activity (compare lanes 2 with 10 or 12). Each of these *R. sphaeroides rrnB* promoter variants was activated by $CarD_{Rsp}$ (Figure 2.4*B* and *D*), suggesting that they altered transcription without fully bypassing the step(s) affected by $CarD_{Rsp}$. In contrast, the improved extended -10 element and the A-7T mutations increased basal promoter activity to a much greater extent than any of the other substitutions (compare lanes 1 with 13 and 17), and these promoter variants were no longer activated at all by $CarD_{Rsp}$ (Figure 2.4*B* and *C*). We conclude that the absence of -7T or an extended -10 element can create a barrier to transcription initiation, and that $CarD_{Rsp}$ or mutations to consensus or both can alleviate this barrier.

Context dependence of promoter sequences that bypass the effect of CarD on *rrnB*. We next addressed whether the -7T or extended -10 substitutions would bypass the requirement for CarD in three other contexts, the *rpsM*, *rpoZ*, and *rrnA* promoters (Figure 2.5*A*). Both substitutions increased the basal activity of the *rpsM* promoter substantially, and although the substitutions reduced the extent of activation by $CarD_{Rsp}$

Figure 2.5: Context dependence of effects of -7T or extended -10 mutants. (*A*) Positions in three CarD-activated *R. sphaeroides* promoters where mutations were made are indicated in red above each promoter sequence, either a T at -7 or a TGC at -15, -14, -13. *(B)* In vitro transcription with *rpsM*, *rpsM* C-7T, and *rpsM* extended -10 with or without 1280 nM CarD with 20 nM *R. sphaeroides* RNAP in buffer with 100 mM NaCl. Duplicate samples are shown for each. (*C)* Fold-activation (+CarD/no CarD) with standard deviation from 3 assays like that shown in (*B*). (*D*) Same as (*B*) except the *rpoZ*, *rpoZ* A-7T, and *rpoZ* extended -10 promoter plasmids were used as templates. (*E*) Fold-activation (+CarD/no CarD) with standard deviation from 3 assays like that shown in (*D*). (*F*) Same as (*B*) and (*D*) except the *rrnA*, *rrnA* A-7T, and *rrnA* extended -10 promoter plasmids were used as templates, and the buffer contained 170 mM NaCl. All lanes are from the same gel, with intervening lanes removed for clarity. (*G*) Fold-activation (+CarD/no CarD) with range from 2 assays like that shown in (*F*).


A
-35
-10
TGC
T
rpsM TTGACGGCCGGGGCAAAAGCCCATAGATCACGGCGTC
TgC
T
rpoZ TTGTCAAGAGGCGGGGCGGCCGTTACATAGGCCGCTC
TGc
T
rrnA TTGCGGGTTTTTTTGCGGTTCCCTAGATAGCGCCTCA
B
rpsM WT rpsM C-7T rpsM ext -10
CarD: - - + + - - + + - - + +
rpsM
RNA1
1 2 3 4 5 6 7 8 9 10 11 12
C
Fold Activation by CarD
8
6
4
2
rpsM WT
rpsM C-7T
rpsM ext -10
D
rpoZ WT rpoZ A-7T rpoZ ext -10
CarD: - - + + - - + + - - + +
rpoZ
RNA1
1 2 3 4 5 6 7 8 9 10 11 12
E
Fold Activation by CarD
5
4
3
2
1
rpoZ WT
rpoZ A-7T
rpoZ ext -10
F
rrnA WT rrnA A-7T rrnA ext -10
CarD: - - + + - - + + - - + +
rrnA
RNA1
1 2 3 4 5 6 7 8 9 10 11 12
G
Fold Activation by CarD
4
3
2
1
rrnA WT
rrnA A-7T
rrnA ext -10

from ~5-fold to ~2-fold (Figure 2.5*B* and *C*), they did not completely eliminate the effect of $CarD_{Rsp}$. As with the *rpsM* and *rrnB* promoters, both substitutions increased the basal activities of the *rrnA* and *rpoZ* promoters. However, unlike *rpsM* and *rrnB*, the *rpoZ* and *rrnA* mutant promoters containing the extended -10 element or -7T substitutions were still activated by $CarD_{Rsp}$ almost to the same extent as with the wild-type promoters (Figures 2.4*B, C,* and *D*).

Thus, there are promoters lacking -7T (i.e. *rsp_3110*, above) that are not activated by $CarD_{Rsp}$, and there are mutant promoters containing -7T that are still activated, at least to some extent, by $CarD_{Rsp}$. These "context effects" indicate that $CarD_{Rsp}$ can compensate for rate-determining steps in promoter complex formation that derive from other promoter sequence features in addition to or instead of -7T. Such context effects have long been recognized in studies of effects of promoter substitutions on transcription (24, 36). The mechanistic explanations for the context effects reported here are described more extensively in the Discussion.

The majority of *R. sphaeroides* promoters lack the T at the last position in the -10 element. To identify likely -10 elements used by *R. sphaeroides* RNAP genome-wide, we analyzed available *R. sphaeroides* transcription start site (TSS) data obtained from cells in exponential growth where most transcription is likely by $E\sigma^{93}$, the major holoenzyme (33). This analysis was followed by a bioinformatic search for a TA sequence corresponding to the highly conserved TA at positions -12 and -11 in *E. coli* -10 hexamers, located at an appropriate distance upstream of each TSS (*SI Appendix*, Dataset 1). The A at position -11 is almost universally conserved in *E. coli* promoters

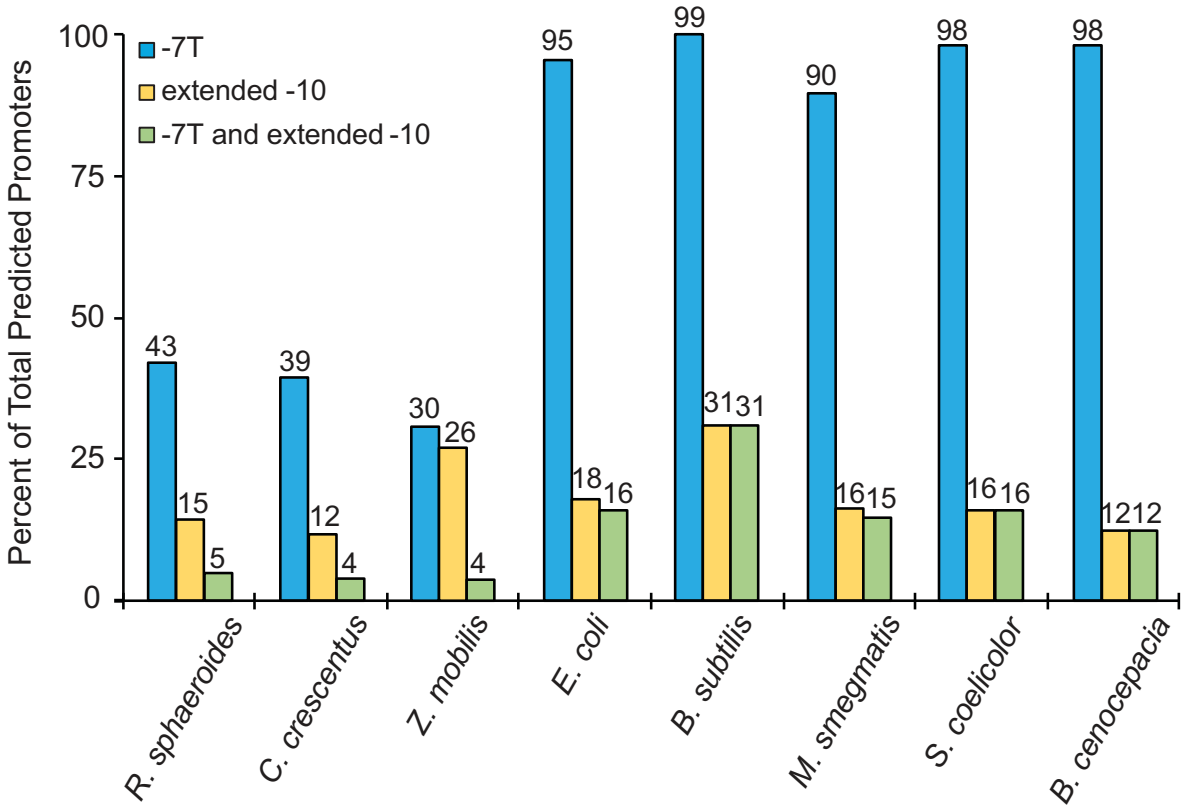
utilizing the major σ factor (25) and is essential for promoter activity with $RNAP_{Eco}$ (24, 37). Residues involved in recognition of A-11 are conserved in the *R. sphaeroides* primary sigma factor (σ^{93}) (*SI Appendix*, Figure 2.S4A). We confirmed that -11A is essential for activity of the *R. sphaeroides rrnB* promoter with $RNAP_{Rsp}$ by showing that substitutions for -11A eliminated transcription (*SI Appendix*, Figure 2.S5*C*).

In Figure 2.6 we analyze the promoters in *R. sphaeroides* that contain -7T genome-wide, based on the presence of the TA motif appropriately positioned upstream of the experimentally defined TSSs, as described above. In support of the interpretation that this promoter collection consists primarily of σ^{93}-dependent promoters, the genome of *R. sphaeroides* does not contain a gene coding for a σ^{S} homolog (11), and other holoenzymes containing alternative sigma factors (at least in *E. coli*) do not contain a TA motif at the upstream end of their -10 elements. The Bioinformatic analysis section in *SI Appendix*, *Expanded Materials and Methods*, contains further details about the promoters included in the genome-wide promoter analysis.

In contrast to *E. coli* promoters where 95% contain a T at -7, only 43% of *R. sphaeroides* promoters contained -7T (Figure 2.6). As with the *R. sphaeroides* promoters lacking -7T that we analyzed in vitro, each of the other 3 bases at position -7 (A, G, and C) was represented in the genome-wide collection of *R. sphaeroides* promoters lacking -7T, indicating the importance of the absence of -7T rather than the presence of another specific base (*SI Appendix*, Dataset 1 and Bioinformatic analysis section in *SI Appendix*, *Expanded Materials and Methods*).

We also used available TSS data to analyze the predicted -10 elements from a selection of other bacterial species, including the α-proteobacteria *C. crescentus* (38)

Figure 2.6: Percentage of promoters in indicated bacterial species containing a thymine at the last position of the -10 hexamer (-7) or an extended -10 consensus (TG) at -14, -15, or both. Percentages of promoters with identifiable -10 elements are shown on the Y-axis. Percentages of those promoters with -7T or extended -10 elements or both are indicated at the top of each bar and are relative to all promoters with identified -10 elements for that species. Promoters were identified based on published TSS data (cited in the main text) and sequence similarity to consensus *E. coli* -10 element as described in *SI Appendix*, *Expanded Materials and Methods*.


Percent of Total Predicted Promoters
100
75
50
25
0
-7T
extended -10
-7T and extended -10
43
15
5
39
12
4
30
26
4
95
18
16
99
31
31
90
16
15
98
16
16
98
12
12
R. sphaeroides
C. crescentus
Z. mobilis
E. coli
B. subtilis
M. smegmatis
S. coelicolor
B. cenocepacia

and *Zymomonas mobilis* (39), the γ-protobacterium *E. coli* (40)**,** the firmicute *Bacillus subtilis* (41). the actinobacteria *Streptomyces coelicolor* and *M. smegma*tis (42, 43), and the β-proteobacterium *Burkholderia cenocepacia* (44). This analysis showed that ~39% of the predicted promoters in *C. crescentus* and 30% in *Z. mobilis* contain -7T, whereas 95%, 99%, 90%, 98%, and 98% of the predicted promoters from *E. coli, B. subtilis*, *M. smegmatis*, *S. coelicolor*, and *B. cenocepacia*, respectively, contain -7T (Figure 2.6). Thus, although the number of species that we analyzed is limited, our results strongly suggest that the -7T is much less conserved in α-proteobacterial promoters than in promoters from other bacterial phyla or classes.

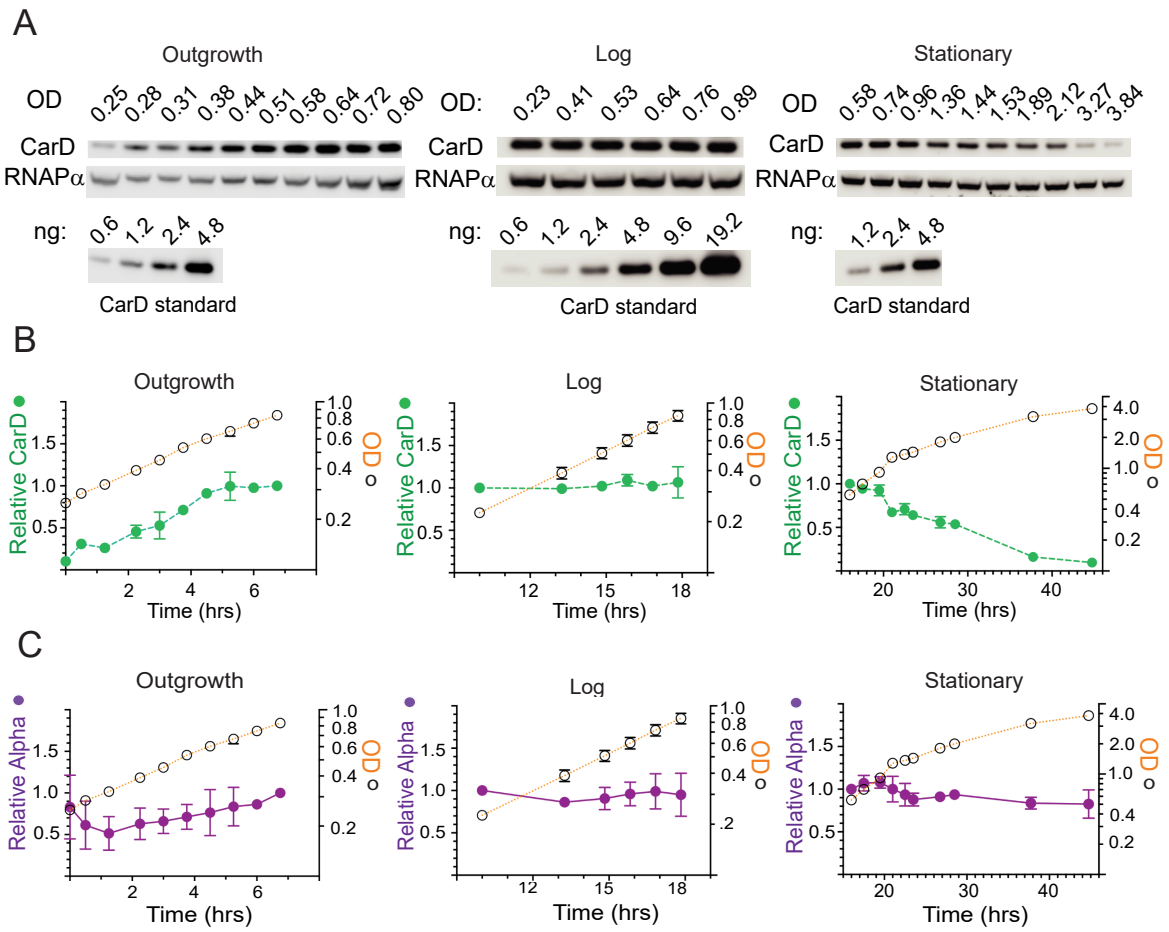
Because the presence of an extended -10 element eliminated activation of the *R. sphaeroides rrnB* promoter by $CarD_{Rsp}$ (Figure 2.4), we also analyzed the percentage of promoters with this element in the genomes of *R. sphaeroides* and the other bacterial species analyzed above (Figure 2.6). Unlike -7T, the extended -10 motif was found in about the same percentage of promoters from both α-proteobacteria and non-α-proteobacteria. The number of promoters with an extended -10 element was much smaller than the percentage with a -7T in most of the species analyzed (Figure 2.6).

Although an extended -10 element compensated for the absence of -7T in an *rrnB* mutant promoter variant, and this promoter was no longer activated by $CarD_{Rsp}$ (Figure 2.4), an extended -10 element did not correlate well with activation by $CarD_{Rsp}$ in the 19 native promoters tested in vitro, in stark contrast to -7T (Figure 2.3 and *SI Appendix* Figure 2.S7). Together with our results that the majority of putative promoters in α-proteobacteria lack an extended -10 element (Figure 2.6), that the presence of an extended -10 element in 6 other native *R. sphaeroides* promoters tested in vitro did not

correlate with activation by CarD$_{Rsp}$ (Figure 2.S7), and that 85% of *R. sphaeroides* promoters lack an extended -10 element, we conclude that the absence of an extended -10 element is not predictive of activation by CarD$_{Rsp}$.

Levels of CarD$_{Rsp}$ change in vivo. One prediction of our data is that promoters lacking -7T may have evolved in α-proteobacteria to be regulated by changing concentrations of CarD. To test whether CarD levels change in vivo, we used purified CarD$_{Rsp}$ to develop a CarD-specific antibody (see *SI Appendix*, *Expanded Materials and Methods*), and CarD$_{Rsp}$ levels were analyzed by Western blot during aerobic growth in minimal medium as cells emerged from stationary phase (outgrowth), grew throughout log phase (log), and went into and remained in stationary phase (stationary) (Figure 2.7*A*-*C*). During outgrowth, the CarD$_{Rsp}$ concentration rose gradually to the level present in log phase, whereas the level of the α subunit of RNAP changed little. The levels of CarD and α remained constant per OD_{600} (an estimate of total protein concentration) during log phase. However, the CarD concentration dropped more than 10-fold during stationary phase, whereas the RNAP α subunit concentration decreased only slightly. We analyzed the cellular CarD concentration relative to colony forming units and found that the decrease in CarD concentration did not reduce cell viability (*SI Appendix*, Figure 2.S8). Based on the purified CarD$_{Rsp}$ standards run in parallel (Figure 2.7*A*) and a volume of 0.87 μm^3 for aerobically grown log phase cells (45), we estimate that the concentration of CarD is ~1.1 μM in log phase and <100 nM in stationary phase (see *SI Appendix, Extended Materials and Methods* for calculation), in the range of the concentrations of CarD that affected transcription in vitro. Taken together, our data

Figure 2.7: CarD$_{Rsp}$ levels at different phases of growth. (*A*) Western blots were performed from *R. sphaeroides* aerobic cultures grown at 30°C, diluted from stationary phase into fresh minimal medium at time zero, and sampled at the indicated times. Separate experiments were conducted for each growth phase (outgrowth, log, and stationary; see *SI Appendix*, *Expanded Materials and Methods*, for details). For each sample, the OD_{600} was measured, CarD levels were determined by Western blotting with anti-CarD$_{Rsp}$ antibody, and RNAP$_{Rsp}$ levels were determined with an anti-α subunit antibody. Purified CarD standards were analyzed in parallel for quantitation of CarD levels. (*B*) Relative CarD levels and OD_{600} values for cultures sampled during outgrowth, log, and stationary phase. (C) Relative RNAP$_{Rsp}$ α subunit levels in the same samples analyzed for CarD levels. See *SI Appendix*, *Expanded Materials and Methods* for details. Values for CarD$_{Rsp}$, RNAP$_{Rsp}$ α-subunit, and OD_{600} were plotted on the same scale by normalizing to the CarD level at an overlapping $OD_{600.}$ Error ranges were determined from 2 experiments for each.


A
Outgrowth
OD 0.25 0.28 0.31 0.38 0.44 0.51 0.58 0.64 0.72 0.80
CarD
RNAPα
ng: 0.6 1.2 2.4 4.8
CarD standard
Log
OD: 0.23 0.41 0.53 0.64 0.76 0.89
CarD
RNAPα
ng: 0.6 1.2 2.4 4.8 9.6 19.2
CarD standard
Stationary
OD 0.58 0.74 0.96 1.36 1.44 1.53 1.89 2.12 3.27 3.84
CarD
RNAPα
ng: 1.2 2.4 4.8
CarD standard
B
Outgrowth
Relative CarD
OD
Time (hrs)
Log
Relative CarD
OD
Time (hrs)
Stationary
Relative CarD
OD
Time (hrs)
C
Outgrowth
Relative Alpha
OD
Time (hrs)
Log
Relative Alpha
OD
Time (hrs)
Stationary
Relative Alpha
OD
Time (hrs)

indicate that CarD concentrations change in vivo and thus have the potential to regulate transcription in cells. Future studies will investigate the mechanisms responsible for regulating the concentrations (or activities) of CarD under different conditions.

DISCUSSION

The absence of -7T is a major determinant of the *R. sphaeroides* transcriptome. We discovered an unexpected feature of *R. sphaeroides* promoters and the promoters of two other α-proteobacterial species, *C. crescentus* (38) and *Z. mobilis* (39). Only 30-43% of the promoters from these 3 α-proteobacterial species contain a thymine base at promoter position -7, the most downstream position in the -10 element, whereas this base is 90-99% conserved in promoters in other bacterial phyla (Figure 2.6). Most of the 19 *R. sphaeroides* promoters tested in vitro were very poorly transcribed by *R. sphaeroides* RNAP and lacked -7T, suggesting an activator might be required to compensate for the absence of the -7T interaction with σ. We suggest that activation of the large number of promoters lacking -7T by transcription factor(s) makes a major contribution to shaping the transcriptome of α-proteobacteria.

Role of -7T in promoter function. Promoters recognized by the major holoenzyme, $E\sigma^{70}$ in *E. coli* or $E\sigma^{93}$ in *R. sphaeroides*, have sequences that vary except at 6 very highly conserved positions, 3 each in the -35 and -10 hexamers (35), one of which is the thymine at -7. Transcription initiation is characterized by a series of conformational changes in the promoter and in RNAP, driven by binding free energy, in which RNAP first binds to the promoter to form a closed complex, followed by a series of steps in

which RNAP melts the DNA strands to form an open complex (35, 46-48). -7T plays an important role in this multistep process, fitting base-specifically into a conserved pocket formed by the major σ factor in promoter complexes in very diverse bacterial phyla (e.g. *E. coli*, *T. thermophilus*, and *M. smegmatis)* (20, 26, 49). This interaction with σ occurs subsequent to the formation of an initial partially opened complex containing a 5 nt transcription bubble (47). The -7T interaction with σ facilitates further DNA strand separation, displacement of the N-terminal domain of σ ($\sigma_{1.1}$) from the main DNA channel, and stabilization of the open complex to yield a 13 nt bubble (47, 48). The absence of the -7T interaction in the majority of *R. sphaeroides* promoters would be expected to reduce the rate of formation and stability of the open complex, creating the need for a transcription factor to enhance the rate-limiting kinetic step(s) at these promoters and thereby enhance promoter activity.

CarD facilitates a step in the mechanism of promoter opening that is affected by the RNAP interaction with promoter position -7T. We identified a strong correlation between the absence of -7T and activation of native *R. sphaeroides* promoters by the transcription factor, CarD. CarD increased the activity of 15 of the 16 promoters that lacked -7T, whereas it had no effect on 3 of the 4 promoters that contained -7T (Figure 2.3, and *SI Appendix,* Figure 2.S7). *R. sphaeroides* CarD shares significant similarity with other CarD homologs from different phyla (e.g. *Mycobacteria* and *Thermus).* Transcription activation by $CarD_{Rsp}$ required a highly conserved tryptophan residue (W91) in the predicted DNA-binding domain, corresponding to W85 in *Mycobacteria* and W86 in *Thermus*, suggesting that $CarD_{Rsp}$ has mechanistic features in common with

these previously characterized CarD proteins (Figure 2.2) (13, 14, 27). Previous structural studies indicated that CarD$_{Tth}$ interacts sequence non-specifically with promoter DNA at the junction of double-stranded and single-stranded DNA at the upstream end of the -10 hexamer (12-16, 27), adjacent to position -11 where strand separation initiates. Kinetic studies with the *Mycobacterial* and *Thermus* CarDs indicated that they affect isomerization step(s) during open complex formation, stimulating formation of a partially melted intermediate and stabilizing open complexes (13, 16, 27, 50).

Thus, CarD and -7T both act on isomerization steps involving extension of a partially open to a fully open transcription bubble and stabilization of the open complex, consistent with the ability of CarD$_{Rsp}$ to compensate, at least in part, for the absence of -7T in many promoters.

Context affects promoter activity. Our in vitro experiments established a strong correlation between promoters lacking -7T and activation by CarD$_{Rsp}$. The high percentage of *R. sphaeroides* promoters genome-wide lacking -7T is consistent with the relative ease with which we identified specific promoters activated directly by CarD$_{Rsp}$, even though the promoters were not chosen for study because they lacked -7T. However, as described below, the absence of -7T is not a perfect predictor of activation by CarD in all promoter contexts.

As a result of the multi-step nature of promoter complex formation and the influence of multiple RNAP interactions with other sequences in the promoter, substitutions at the same position in different promoters have long been known to

exhibit different effects on transcription output (“context effects”, e.g. 24, 36). Promoter interactions with RNAP containing the primary sigma factor involve not only the specific base interactions at the critical conserved positions in the consensus elements, but also phosphate backbone interactions within and outside of the consensus hexamers and base-specific interactions in regions of the promoter that may only be present in a subset or promoters. These interactions nevertheless collectively contribute to the rate of open complex formation and can lead to differences in the rate-determining steps that limit promoter activity at different promoters (48).

Transcription can only be activated by a factor that acts on a rate-determining kinetic step. In some contexts, an interaction involving a particular promoter position may be the major determinant of the step affected by the activator, but in other contexts another promoter interaction, or more than one interaction, may limit promoter activity, and the activator may have little or no effect on transcriptional output (35, 46). The one promoter in our cohort of 16 native *R. sphaeroides* promoters examined in vitro that lacked -7T but was not activated by $CarD_{Rsp}$ is an example of such a context effect. This promoter, *rsp_3110*, contains an extended -10 element that we suggest bypasses the requirement for $CarD_{Rsp}$ by facilitating strand opening, just as a mutation to -7T in some promoter contexts (e.g. *rpoZ* and *rrnA*) does not fully eliminate activation by $CarD_{Rsp}$.

Taken together, our observations suggest that the absence of -7T is a major contributor to the dependence of many *R. sphaeroides* promoters on an activator for transcription. $CarD_{Rsp}$ may not be the transcription factor responsible for regulating every promoter lacking -7T that requires an activator, nor is the presence/absence of -7T likely to be the only determinant of a response to the regulator(s) in all promoter

contexts. Thus, prediction of the promoters activated by $CarD_{Rsp}$ genome-wide, as well as prediction of the magnitude of the effect of $CarD_{Rsp}$ on a specific promoter, will be more complex than simply identifying promoters lacking -7T.

The absence of a conserved RNAP recognition feature in a large set of promoters creates a "regulon" controlled by a transcription factor. The unexpectedly low percentage of *R. sphaeroides* promoters containing -7T (43%) stands in stark contrast to the very high percentage (90-99%) containing -7T in other bacterial phyla (Figure 2.6). Together, the strong correlation between the absence of -7T and activation by $CarD_{Rsp}$ in our in vitro experiments (Figures 2.2, 2.3, and *SI Appendix*, Figure 2.S7), along with the observations that CarD is essential in *R. sphaeroides* (5) and that CarD concentrations vary with growth phase (Figure 2.7), suggest that CarD plays a role in activating many promoters lacking -7T. Most bacterial phyla (e.g. β- and γ-Proteobacteria, Firmicutes, Actinobacteria, Thermus-Deinococcus) have promoters containing -7T. How such a large subset of promoters lacking -7T evolved in α-proteobacteria whereas diverse bacterial phyla contain promoters with -7T remains a question for future exploration. We suggest that regulons in other bacterial phyla that derive from utilization of a transcription factor that compensates for the absence of a crucial base in the core promoter, not necessarily -7T, could be a common mechanism contributing to regulation.

The role of CarD and -7T in other α-proteobacteria and other bacterial phyla. Our bioinformatic analysis of published genome-scale TSS data indicated that the absence

of -7T is a feature of a majority of promoters in *R. sphaeroides* and the two other α-proteobacterial species for which we were able to predict -10 element sequences genome-wide. Consensus sequences for promoters in the α-proteobacteria *C. crescentus* and *Sinorhizobium melliloti* have been proposed (38, 51-53). These sequence predictions included a -10 element consensus element consisting only of positions matching bases near the upstream end of the *E. coli* -10 element. Furthermore, the *rrnA* promoter from the δ-proteobacterium *M. xanthus* lacks -7T and is activated by CarD (14). Together, these studies are consistent with our proposal that CarD activates transcription in many other species. However, the previous studies did not correlate the absence of -7T with activation by CarD. Deletion of *C. crescentus* CarD (CdnL) was reported to affect transcription from many promoters in vivo, but it was proposed that the number of promoters regulated directly by CarD was small and that most effects of CarD on transcription were likely to be indirect (19). In contrast, our data suggest that α-proteobacterial CarD is likely to affect transcription initiation from many promoters directly**.**

In *Mycobacterial* systems, where CarD has been most thoroughly investigated in vitro, most promoters including the rRNA promoter AP3 contain -7T, and thus a role for CarD in compensating for the -7T interaction with sigma was not proposed. Other interactions must be rate-limiting in *Mycobacterial* promoters as a result of DNA sequence differences in the promoter and/or amino acid sequence variation in RNAP among *Mycobacteria*, *Thermus*, and proteobacteria (13, 14, 27, 54). The effects of CarD on the *M. tuberculosis* rRNA promoter are also amplified by a second transcription factor, RbpA, not found in *R. sphaeroides* (16, 50).

Implications for regulation of synthesis of the *R. sphaeroides* translation apparatus. The *R. sphaeroides* promoters activated by CarD include a variety of housekeeping genes, including several involved in synthesis and assembly of the translation apparatus (rRNAs, ribosomal proteins, tRNAs, transcription factors; Figure 2.3 and *SI Appendix*, Figure 2.S7). Consistent with its critical role in translation, *carD* is an essential gene in *R. sphaeroides* (5*).* The decrease in CarD levels in stationary phase may contribute to the decrease in expression of rRNA and other translation-related gene products at this stage in growth.

Additional transcription factors are also likely to contribute to rRNA transcription regulation in *R. sphaeroides*. In *E. coli*, the stringent response factors DksA and ppGpp, like CarD in *R. sphaeroides*, interact directly with RNAP and affect specific promoters because of their specific kinetic properties, in contrast to classical transcription factors whose promoter specificity results from DNA binding sites adjacent to individual promoters (8, 35, 46, 50). *R. sphaeroides* DksA was previously shown to regulate transcription by *E. coli* RNAP in vitro (55), suggesting that *R. sphaeroides* uses ppGpp/DksA as well as CarD to regulate rRNA transcription. In *E. coli*, the Fis protein activates rRNA promoters by binding to sites upstream of their -35 elements, contributing to their regulation (6). *R. sphaeroides* rRNA promoters could also respond to transcription factor(s) yet to be identified that bind upstream of the core promoter, since deletions upstream of the -35 element reduce rRNA promoter activity *in vivo* (11). The integration of effects of CarD and other regulators of *R. sphaeroides* rRNA transcription will be subjects of a separate study.

MATERIALS AND METHODS

Further details for each section are provided in *SI Appendix, Expanded Materials and Methods*.

Bacterial strains: *E. coli* and *R. sphaeroides* strains are listed in *SI Appendix*, Table 2.S1.

Bacterial Growth: *R. sphaeroides* was grown aerobically in a succinate-based minimal medium (56).

Purification of *R. sphaeroides* RNAP. An *R. sphaeroides* 2.4.1 derivative for purification of RNAP was constructed by creating a His10 tag fused to the C-terminus of the β' subunit and introducing it into the *R. sphaeroides* chromosome using the non-replicative plasmid pk18mobsacB (57) and a two-step recombination method. RNAP holoenzyme containing σ^{93} was purified by Ni-affinity chromatography from aerobically grown cells.

Construction and purification of *R. sphaeroides* Sumo-tagged CarD. A codon optimized *carD* gene was inserted into pETSUMO for purification of an N-terminally tagged His10-SUMO-tagged CarD. Wild type and variant CarD proteins were overexpressed and purified from *E. coli* BL21 (DE3) pLysS cells (14).

in vitro transcription. Test promoter fragments were PCR amplified from *R. sphaeroides* 2.4.1 chromosomal DNA and inserted into pRLG770 (6), and promoter variants were constructed using Multi Site Lightning Quick Change Mutagenesis (Agilent) using primers listed in *SI Appendix*, Table 2.S2. Transcription was performed with either *R. sphaeroides* RNAP holoenzyme containing the major sigma factor σ^{93} or *E. coli* Eσ^{70}, as indicated.

Western blot analysis of CarD. CarD concentrations in aerobically grown *R. sphaeroides* cells were determined throughout a growth curve using a polyclonal antibody raised against CarD without a SUMO tag. Polyclonal antibody raised against the α subunit of *R. sphaeroides* RNAP was used for comparison.

Bioinformatic analysis of promoter elements. We used available transcription start site (TSS) information in the literature to determine the likely -10 elements of promoters from the bacterial species indicated in Figure 6, first identifying the bases most likely corresponding to -11A and -12T and then the DNA sequences most likely corresponding to -7T and the extended -10 element, as described in detail in *SI Appendix*, *Expanded Materials and Methods*. The TSSs for *R. sphaeroides* were determined as described (33).

ACKNOWLEDGMENTS

We thank A. Mehle for access to the Li-Cor Fc instrument used for imaging Western blots and D. Grainger for providing genome-wide *B. subtilis* transcription start site information prior to publication. We also thank F. Alberge for measurements of *R. sphaeroides* cell volume. Work in the Gourse lab is supported by NIH R01 GM37048. K.K.H. was also supported by fellowships from the Jack Kent Cooke Foundation, an NIH Molecular Biosciences Training Grant (T32GM007215), and the Department of Bacteriology. Work in the Donohue and Landick labs is supported by U.S. Department of Energy, Office of Biological and Environmental Research under Award Number DE-SC0018409.

REFEREENCES

1. C. Mackenzie *et al*, Postgenomic Adventures with *Rhodobacter sphaeroides, Annu Rev Microbiol* **61**:283–307 (2007).
2. S. Imam, D.R. Noguera, T.J. Donohue, Global insights into energetic and metabolic networks in *Rhodobacter sphaeroides*, *BMC Syst Biol* **7**:89 (2013).
3. S. Imam, D.R. Noguera, T.J. Donohue, Global analysis of photosynthesis transcriptional regulatory networks, *PLoS Genet* **10**: e1004837 (2014).
4. J.H. Zeilstra-Ryalls, M. Gomelsky, A.A. Yeliseev, J.M. Eraso, S. Kaplan, Transcriptional regulation of photosynthesis operons in *Rhodobacter sphaeroides* 2.4.1, *Meth Enzymol* **297**: 151-166 (1998).
5. B.T. Burger, S. Imam, M.J. Scarborough, D.R. Noguera, T.J. Donohue, Combining genome-scale experimental and computational methods to identify essential genes in *Rhodobacter sphaeroides. mSystems* **2**:e00015-17 (2017).
6. W. Ross, J.F. Thompson, J.T. Newlands, R.L. Gourse, Fis protein activates ribosomal RNA transcription in vitro and in vivo. *EMBO J* **9**:3733-42 (1990).
7. W. Ross *et al*, A third recognition element in bacterial promoters: DNA binding by the alpha subunit of RNA polymerase, *Science* **262**:1407-1413 (1993).
8. R.L. Gourse, A.Y. Chen, S. Gopalkrishnan, P. Sanchez-Vazquez, W. Ross, Transcriptional Responses to ppGpp and DksA, *Annu Rev Microbiol* **72**:163-184 (2018).
9. P. Sanchez-Vazquez, C.N. Dewey, N. Kitten, W. Ross, R.L. Gourse, Genome-wide effects on *Escherichia coli* transcription from ppGpp binding to its two sites on RNA polymerase, *Proc Natl Acad Sci USA* **116**:8310-8319 (2019).
10. S.C. Dryden, S. Kaplan, Localization and structural analysis of the ribosomal RNA operons of *Rhodobacter sphaeroides*, *Nucleic Acids Res* **18**:7267–7277(1990).
11. S.C. Dryden, S. Kaplan, Identification of cis-acting regulatory regions upstream of the rRNA operons of *Rhodobacter sphaeroides*, *J. Bacteriol* **175**:6392-6402 (1993).
12. C. L. Stallings *et al*, CarD is an essential regulator of rRNA transcription required for Mycobacterium tuberculosis persistence. *Cell* 138:146-59 (2009).
13. B. Bae *et al*, CarD uses a minor groove wedge mechanism to stabilize the RNA polymerase open promoter complex. *eLife* **4:**08505 (2015).
14. D.B. Srivastava *et al*, Structure and function of CarD, an essential mycobacterial transcription factor. *Proc Natl Acad Sci USA* **110**:12619-12624 (2013).
15. A. Gallego-García *et al*, Structural insights into RNA polymerase recognition and essential function of *Myxococcus xanthus* CdnL. *PLoS One* **9**:e108946 (2014).
16. E.A. Hubin *et al*, Structure and function of the mycobacterial transcription initiation complex with the essential regulator RbpA. *eLife* **6**:e22520 (2017).

17. A. Gallego-García *et al*, *Caulobacter crescentus* CdnL is a non-essential RNA polymerase-binding protein whose depletion impairs normal growth and rRNA transcription, *Sci Rep* **7**:43240 (2017).

18. D.X. Zhu, A.L. Garner, E.A. Galburt, C.L. Stallings, CarD contributes to diverse gene expression outcomes throughout the genome of *Mycobacterium tuberculosis, Proc Natl Acad Sci U S A* **116**:13573-13581 (2019).

19. S.A. Woldemeskel *et al*, The conserved transcriptional regulator CdnL is required for metabolic homeostasis and morphogenesis in *Caulobacter*, *PLoS Genet* **16**:e1008591 (2020).

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

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

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

23. S. P. Haugen *et al*, rRNA promoter regulation by nonoptimal binding of sigma region 1.2: an additional recognition element for RNA polymerase, *Cell* **125**:1069-1082 (2006).

24. H. Moyle, C. Waldburger, M.M. Susskind, Hierarchies of base pair preferences in the P22 ant promoter. *J Bacteriol* **173**:1944-1950 (1991).

25. R.K. Shultzaberger, Z. Chen, K.A. Lewis, T.D. Schneider, Anatomy of *Escherichia coli* sigma70 promoters, *Nucleic Acids Res* **35**:771-788 (2007).

26. Y. Zhang *et al*, Structural basis of transcription initiation, *Science* **338**:1076-1080 (2012).

27. E. Davis, J. Chen, K. Leon, S.A. Darst, E. A. Campbell, Mycobacterial RNA polymerase forms unstable open promoter complexes that are stabilized by CarD, *Nucleic Acids Res* **43**:433-45 (2014).

28. D. García-Moreno *et al*, CdnL, a member of the large CarD-like family of bacterial proteins, is vital for *Myxococcus xanthus* and differs functionally from the global transcriptional regulator CarD, *Nucleic Acids Res* **38**:4586-4598 (2010).

29. L.F. Westblade *et al*, Structural basis for the bacterial transcription-repair coupling factor/RNA polymerase interaction, *Nucleic Acids Res* **38**:8357-8369 (2010).

30. R.K. Karls, D.J. Jin, T.J. Donohue, Transcription properties of RNA polymerase holoenzymes isolated from the purple nonsulfur bacterium *Rhodobacter sphaeroides*, *J Bacteriol* **175**:7629-7638 (1993).

31. L.A. Weiss *et al*, Interaction of CarD with RNA polymerase mediates *Mycobacterium tuberculosis* viability, rifampin resistance, and pathogenesis, *J Bacteriol* 194:5621-31 (2012).

32. R. Landick, A. Krek, M.S. Glickman, N.D. Socci, C.L. Stallings, Genome-wide mapping of the distribution of CarD, RNAP σ^A, and RNAP β on the *Mycobacterium smegmatis* chromosome using chromatin immunoprecipitation sequencing, *Genom Data* **2**:110-113 (2014).

33. K.S. Myers *et al*, Genome-wide Identification of Transcription Start Sites for *Rhodobacter sphaeroides* Grown Aerobically and Photosynthetically, *ASM Microbiology Resource Announcements* **9**:e00880-20 (2020).

34. G.E. Crooks, G. Hon, J.M. Chandonia, S.E. Brenner, WebLogo: a sequence logo generator, *Genome Research* **14**:1188-1190 (2004).

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

36. D. Graña, T. Gardella, M.M. Susskind, The effects of mutations in the ant promoter of phage P22 depend on context, *Genetics* **120**:319-327 (1988).

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

38. B. Zhou et al, The global regulatory architecture of transcription during the *Caulobacter* cell cycle, *PLoS Genet*. **11**:e1004831 (2015).

39. J.M. Vera *et al*, Genome-scale transcription–translation mapping reveals features of *Zymomonas mobilis* transcription units and promoters, *mSystems* **5**:e00250-20 (2020).

40. M.K. Thomason *et al*, Global transcriptional start site mapping using differential RNA sequencing reveals novel antisense RNAs in *Escherichia coli*, *J Bacteriol* 1**97**:18–28 (2015).

41. E. Warman, D. Forrest, J.T. Wade, D.C. Grainger, Widespread divergent transcription from prokaryotic promoters, doi.org/10.1101/2020.01.31.928960doi: bioRxiv preprint first posted online Feb. 2, 2020.

42. Y. Jeong *et al*, The dynamic transcriptional and translational landscape of the model antibiotic producer *Streptomyces coelicolor* A3(2), *Nat Commun* **7**:11605 (2016).

43. X. Li *et al*, Transcriptome Landscape of *Mycobacterium smegmatis, Front Microbiol* **8**:2505 (2017).

44. A.M. Sass *et al*, Genome-wide transcription start site profiling in biofilm-grown *Burkholderia cenocepacia* J2315, *BMC Genomics* **16**:775 (2015).

45. K. Lemmer *et al*, The NtrYX two-component system regulates the gram-negative cell envelope. *mBio* **11**:e00957-20 (2020).

46. E. Galburt, The calculation of transcript flux ratios reveals single regulatory mechanisms capable of activation and repression. *Proc Natl Acad Sci USA* **115**: E11604-E11613 (2018).

47. J. Chen *et al*, Stepwise promoter melting by bacterial RNA polymerase. *Mol Cell* **78**:275-288 (2020).

48. E.F. Ruff, M.T. Record Jr, I. Artsimovitch, Initial events in bacterial transcription initiation, *Biomolecules* **5**:1035-62 (2015).

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

50. J. Rammohan, A. Ruiz Manzano, A.L. Garner, C.L. Stallings, E.A. Galburt, CarD stabilizes mycobacterial open complexes via a two-tiered kinetic mechanism. *Nucleic Acids Res* **43**:3272–3285 (2015).

51. J. Malakooti, S.P. Wang, B. Ely, A consensus promoter sequence for Caulobacter crescentus genes involved in biosynthetic and housekeeping functions, *J Bacteriol* **177**:4372-4376 (1995).

52. S.R. MacLellan, A.M. MacLean, T.M. Finan, Promoter prediction in the rhizobia, *Microbiology* **152**:1751-1763 (2006).

53. J. P. Schlüter *et al*, Global mapping of transcription start sites and promoter motifs in the symbiotic α-proteobacterium *Sinorhizobium meliloti* 1021, *BMC Genomics* **14**:156 (2013).

54. T. Kouba *et al*, The core and holoenzyme forms of RNA polymerase from *Mycobacterium smegmatis*. *J Bacteriol* **201**:e00583-18 (2019).

55. C.W. Lennon *et al*, A Rhodobacter sphaeroides protein mechanistically similar to *Escherichia coli* DksA regulates photosynthetic growth. *mBio* **5**:e01105-14. (2014).

56. W. R. Sistrom, A requirement for sodium in the growth of *Rhodopseudomonas spheroides*, *J Gen Microbiol* **22**:778–785 (1960).

57. A. Schäfer *et al*, Small mobilizable multi-purpose cloning vectors derived from the *Escherichia coli* plasmids pK18 and pK19: selection of defined deletions in the chromosome of *Corynebacterium glutamicum*, *Gene* **145**:69–73 (1994).

SUPPLEMENTAL INFORMATION

Supplemental Figures

Figure 2.S1: RNAP purification. (*A*) 1.0 µg purified *E. coli* RNAP (lane 1) or *R. sphaeroid*es RNAP (lane 3) run on a 4-12% Invitrogen NuPAGE Bis-Tris gel with MES buffer. MW (kDa) of See Blue 2 size markers (Invitrogen, lane 2) are indicated above bands. Intervening lanes were removed for clarity. (*B*) All lanes are from the same gel (12% NuPAGE Bis-Tris/MES). Lanes 1- 3 were Coomassie stained. Lanes 1 and 2, 1 µg of the indicated RNAP. Lane 3, Size marker as in (*A*). Lanes 4-6, Western blot of purified $RNAP_{Rsp}$ to detect σ and α subunits. Lane 4, 10 ng $RNAP_{Rsp}$. Lane 5, 20 ng $RNAP_{Rsp}$, Lane 6, marker as in (*A*). Top portion of blot (above dotted arrow) was probed with anti-σ^{70} antibody WI168. Lower portion of blot was probed with anti-α antibody (WI515). All lanes (1-6) were visualized by Li-COR imaging. (*C*) Western blot with anti-$CarD_{Rsp}$ antibody, using indicated amounts of purified $CarD_{Rsp}$ or purified $RNAP_{Rsp}$. The RNAP:CarD molar ratio was 126:1. Marker (M) is the same as in (*A*).

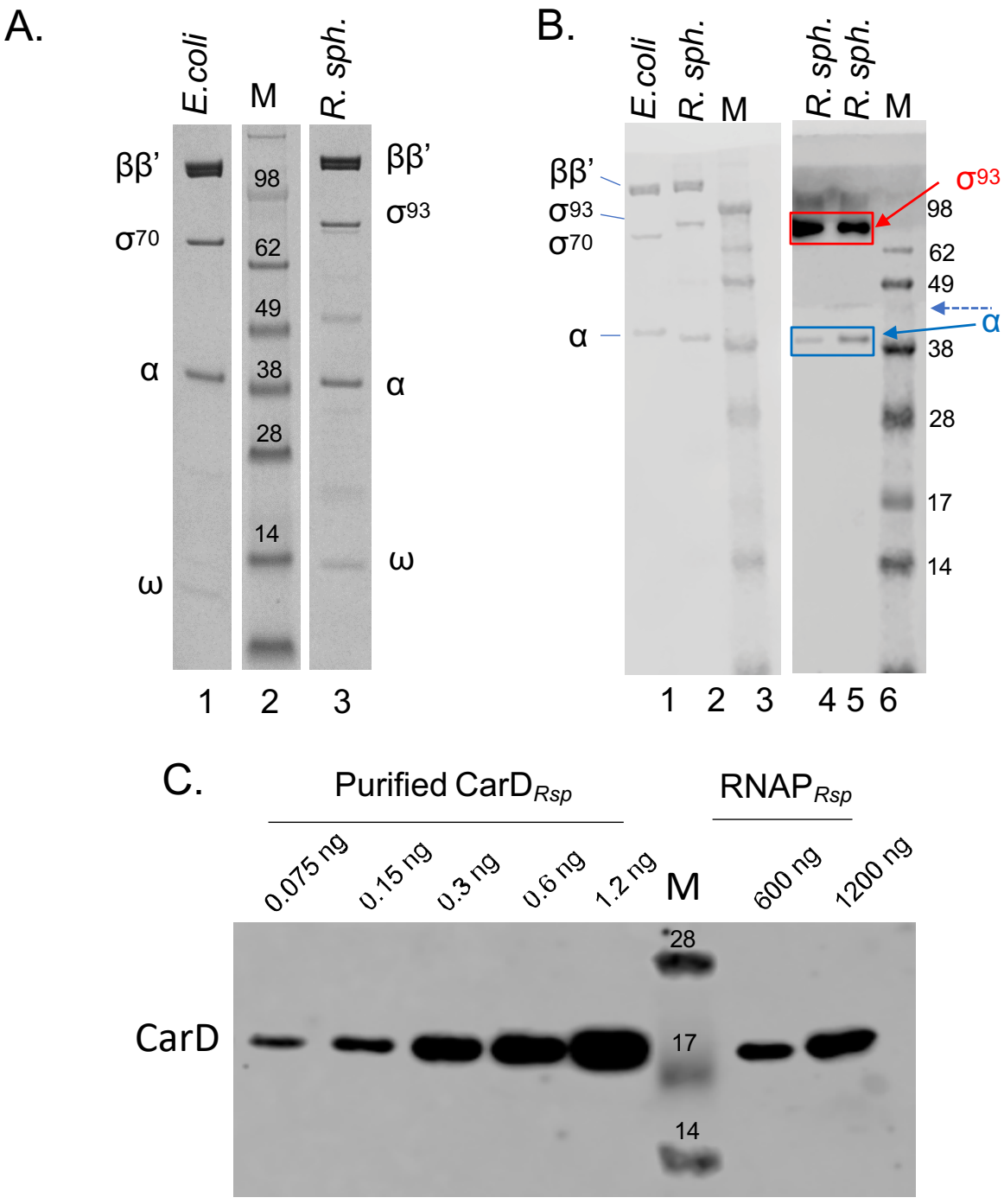

A.
E.coli
M
R. sph.
ββ'
σ70
α
ω
98
62
49
38
28
14
ββ'
σ93
α
ω
1 2 3
B.
E.coli
R. sph.
M
R. sph.
R. sph.
M
ββ'
σ93
σ70
α
σ93
α
98
62
49
38
28
17
14
1 2 3 4 5 6
C.
Purified CarD Rsp
RNAP Rsp
0.075 ng
0.15 ng
0.3 ng
0.6 ng
1.2 ng
M
600 ng
1200 ng
CarD
28
17
14

Figure 2.S2: Transcription of non-*R. sphaeroides* promoters with 20 nM RNAP$_{Rsp}$. Buffer contained 170 mM NaCl. Only *E. coli rrnB* P1 was activated by CarD$_{Rsp}$. Promoter fragments were inserted upstream of the *rrnB* P1 terminator to produce the transcripts indicated by the red arrowheads. RNA 1 is encoded by the plasmid template. Where indicated the CarD concentration was 1280 nM (see *Expanded Materials and Methods*). The *lacUV5* promoter is a variant of the *E. coli lac* promoter. The *T7A1lacO34* promoter is a variant of the phage T7A1 promoter (see *Expanded Materials and Methods*). In lanes 13-16, the template contains the RNA I promoter inserted upstream of the *rrnB* T1 terminator, so that RNA I is transcribed from both its native location in the plasmid and also from the site of fragment insertion.

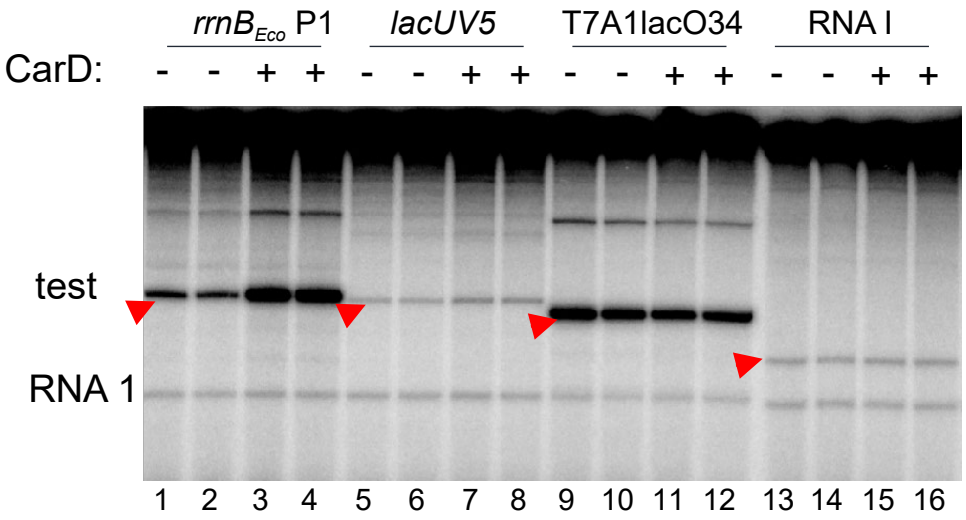

rrnB Eco P1
lacUV5
T7A1lacO34
RNA I
CarD: - - + + - - + + - - + + - - + +
test
RNA 1
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16

Figure 2.S3: *R. sphaeroides* rRNA promoter sequences. An A at position -7 (6th position of the -10 hexamer; highlighted in red) is present in each of the rRNA promoters (*rrnA, rrnB, rrnC*, and where present, *rrnD*) in six different *R. sphaeroid*es strains. Genome sequences were obtained from NCBI. Strain accession numbers for Rsph 2.4.1: GCA_003324715.1; Rsph_ATCC_17025: GCA_000016405.1; Rsph_ATCC_17029: GCA_000015985.1; Rsph_KD131: GCA_000021005.1; Rsph_WS8N: GCA_000212605.1; Rsph_MBTJL-8: GCA_001576595.1; Rsph_MBTJL-13: GCA_001685625.1.

```
CTTGCGGGTTTTTTTGCGGTTCCCTAGATAGCGCCTCACCG          Rsph_2.4.1_rrnA
CTTGCGCCCGGGGCCGTCTGCTCCTAGAAACCGCTTCACCG          Rsph_2.4.1_rrnB
CTTGCGGGTTTTTTTGCGGTTCCCTAGATAGCGCCTCACCG          Rsph_2.4.1_rrnC

CTTGCGGTGCCCCGAGAGAGCGCCTAGATAGCCGCTCCACG          Rsph_ATCC_17025_rrnA
CTTGCGGGTTTTTTTGTGGTTCCCTAGATAGCGCTTCACCG          Rsph_ATCC_17025_rrnB
CTTGCAGACCGGGGCGACTGCCCCTAAAAACCGCTTCACCG          Rsph_ATCC_17025_rrnC
CTTGCGGGTTTTTTTGTGGTTCCCTAGATAGCGCCTCACCG          Rsph_ATCC_17025_rrnD

CTTGCGGGTTTTTTTGCGGTTCCCTAGATAGCGCCTCACCG          Rsph_ATCC_17029_rrnA
CTTGCGGGTTTTTTTGTGGTTCCCTAGATAGCGCCTCACCG          Rsph_ATCC_17029_rrnB
CTTGCGCGGGCCGTCTGCTCCCCGTAGAAACCGCTTCACCG          Rsph_ATCC_17029_rrnC
CTTGCGGGTTTTTTTGCGGTTCCCTAGATAGCGCCTCACCG          Rsph_ATCC_17029_rrnD

CTTGCGGGTTTTTTTGCGGTTCCCTAGATAGCGCCTCACCG          Rsph_KD131_rrnA
CTTGCGGGTTTTTTTGCGGTTCCCTAGATAGCGCCTCACCG          Rsph_KD131_rrnB
CTTGCGCCCGGGGCCGTCTGCTCCTAGAAACCGCTTCACCG          Rsph_KD131_rrnC
CTTGCGGGTTTTTTTGTGGTTCCCTAGATAGCGCCTCACCG          Rsph_KD131_rrnD

CTTGCGGGTTTTTTTGCGGTTCCCTAGATAGCGCCTCACCG          Rsph_WS8N_rrnA
CTTGCGCCCGGGGCCGTCTGCTCCTAGAAACCGCTTCACCG          Rsph_WS8N_rrnB
CTTGCGGGTTTTTTTGCGGTTCCCTAGATAGCGCCTCACCG          Rsph_WS8N_rrnC

CTTGCGGGTTTTTTTGCGGTTCCCTAGATAGCGCCTCACC           Rsph_MBTJL-8_rrnA
CTTGCGGGTTTTTTTTGCGGTTCCCTAGATAGCGCCTCACC          Rsph_MBTJL-8_rrnB
CTTGCGCCCGGGGCCGTCTGCTCCTAGAAACCGCTTCACC           Rsph_MBTJL-8_rrnC

CTTGCGGGTTTTTTTTGCGGTTCCCTAGATAGCGCCTCACC          Rsph_MBTJL-13_rrnA
CTTGCGGGTTTTTTTTGCGGTTCCCTAGATAGCGCCTCACC          Rsph_MBTJL-13_rrnB
CTTGCGCCCGGGGCCGTCTGCTCCTAGAAACCGCTTCACC           Rsph_MBTJL-13_rrnC
```

Figure 2.S4: **Major σ factor and αCTD alignments showing in (*A*) conservation of residues involved in promoter -10 recognition or in (*B*) residues involved in recognition of UP elements**. (*A*) Alignment of part of major σ factor amino acid sequences from *E. coli* (E), *R. sphaeroides* (R), *Thermus thermophilus* (T), and *Mycobacterium tuberculosis* (M). Conserved residues that contribute directly to the -7 pocket are in red and are boxed, the -11 pocket residues are in blue, and residues that interact with the extended -10 element are in green.

Residue numbers are indicated at the end of each line. Alignments were made using Clustal Omega. The residues comprising the -7T binding pocket in *E. coli* σ^{70} include N383, R385, L386, and S428, which are conserved in *R. sphaeroides* σ^{93} (N499, R501, L502, S543) and in σ factors from species as evolutionarily diverse as *E. coli* (23, 24), *T. thermophilus* / *T.aquaticus* (25-27), and *M. smegmatis / M.tuberculosis* (28). (*B*) Alignment of the alpha C-terminal domains from *E. coli* (E), *R. sphaeroides* (R), and *C. crescentus* (C) RNAP, showing DNA binding residues for UP element recognition (red and boxed) (29).

A Major σ

```
E       WSEKLHDVSEEVHRALQKLQQIEEETGLTIEQVKDINRRMSIGEAKARRAKKEMVEANLR  385
R       WQTLMEKSRDRVEDLRGEMAQVGQYVGVDISEFRRIVNQVQKGEKEARQAKKEMVEANLR  440
T       LKETLDP--KTVEEIDQKLK------SLP-KEHKR----YLHIAREGEAARQHLIEANLR  193
M       ------------------E------KLPAAQRRD----MMWICRDGDRAKNHLLEANLR  301
                                 :   : :           ..  *::.::******

E       LVISIAKKYTNRGLQFLDLIQEGNIGLMKAVDKFEYRRGYKFSTYATWWIRQAITRSIAD  445
R       LVISIAKKYTNRGLQFLDLIQEGNIGLMKAVDKFEYRRGYKFSTYATWWIRQAITRSIAD  500
T       LVVSIAKKYTGRGLSFLDLIQEGNQGLIRAVEKFEYKRRFKFSTYATWWIRQAINRAIAD  253
M       LVVSLAKRYTGRGMAFLDLIQEGNLGLIRAVEKFDYTKGYKFSTYATWWIRQAITRAMAD  361
        **:*:**:**.**: ********* **::**:**:* : :**************.*::**

E       QARTIRIPVHMIETINKLNRISRQMLQEMGREPTPEELAERMLM--PEDKIRKVLKIAKE  503
R       QARTIRIPVHMIETINKLVRTGRQMLHEIGREPTPEELAEKLQM--PLEKVRKVMKIAKE  558
T       QARTIRIPVHMVETINKLSRTARQLQQELGREPTYEEIAEAMGPGWDAKRVEETLKIAQE  313
M       QARTIRIPVHMVEVINKLGRIQRELLQDLGREPTPEELAKEMDI--TPEKVLEIQQYARE  419
        ***********:*.**** *  *:: :::***** **:*: :      .:: :  : *:*
         *:  .   **:*:*  :.********** *:: **: * **:. **.:**

E       RQPEVK--EEKPEFDPILLRPVDDLELTVRSANCLKAEAIHYIGDLVQRTEVELLKTPNL  295
R       ESATRHDVEDGLEFNPLLLKKVDELELSVRSANCLKNDNIVYIGDLIQKTEAEMLRTPNF  296
C       KAKSADESKPELPFNPALLKKVDELELSVRSANCLKNDNIVYIGDLIQKTEAEMLRTPNF  296
        .   .  :    *:* **: **:***:******** : * *****:*:**.*:*:***:

E       GKKSLTEIKDVLASRGLSLGMRLENWPPASIADE--------  329
R       GRKSLNEIKEVLSGMGLHLGMDVEDWPPENIEDLAKRFEDQF  338
C       GRKSLNEIKEVLAGMGLHLGMDVPNWPPENIEDLAKKFEDQI  338
        *:***.***:**:. ** *** : :*** .* *
```

Figure 2.S5: Alignments of CarD, activation by wild-type or W91L mutant CarD$_{Rsp}$, and transcription from *R. sphaeroides rrnB* -11 mutant promoters. (A) Alignments of CarD from *Thermus thermophilus* HB8, *Myxococcus xanthus* DK1622 (also referred to as CdnL), *Mycobacterium tuberculosis* H37Rv, Rhodobacter sphaeroides 2.4.1 and *Caulobacter crescentus* NA1000 (CdnL). Substitutions introduced into CarD$_{Rsp}$ are boxed (Q31A/I33A/R53A; W91A; W91L).

W91 is conserved among the CarD homologs (4). (B) Activation of *rrnB*$_{Rsp}$ transcription by 20 nM RNAP$_{Rsp}$ with 20 nM to 1280 nM CarD$_{Rsp}$ WT or CarD$_{Rsp}$ W91L. (C) In vitro transcription of the *rrnB* wild type, A-11T or A-11G mutant promoters by 20 nM RNAP$_{Rsp}$ ± 1280 nM CarD$_{Rsp}$. Buffer in (*B*) and (*C*) contained 170 mM NaCl.

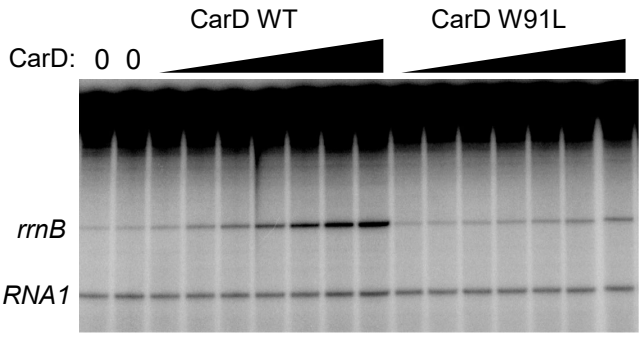
A

```
Tth      -----MKEFRPGDKVVLPPYGVGVVAGIAQRSVSGVSRAYYQVDFPGSRSKAYVPVEAPH       55
Myx      ----MQTSFKTGDKAVYPGQGVGEVMGIEHTEVAGQRQSFYVLRILENGMRIMIPINKVG       56
Mtb      ------MIFKVGDTVVYPHHGAALVEAIETRTIKGEQKEYLVLKVAQGDLTVRVPAENAE       54
Rsp      MTKTKKPEFRPNEFVVYPAHGVGRIISIEEQEIAGIRLELFVISFEKDKMTIRVPTHKAT       60
Ccr      -MSKTGLEFSVGDHVVYPAHGVGSIQAIETQEVAGMSLEVYVITFDHEKMTIRVPTKKAK       59
                  *  .: .* *  *.. : .*    : *      : .       :* .

Tth      SVGLRKALAPEEVPVILDLLKNGR-MPLPKQWAARHRKTSEILADGNPYRIAQMAGQLRA       114
Myx      SVGLREIISEEDVKQVYSILKEKDISVDSTTWNRRYREYMEKIKTGSVFEIAEVLRDLYL       116
Mtb      YVGVRDVVGQEGLDKVFQVLRAPH-TEEPTNWSRRYKANLEKLASGDVNKVAEVVRDLWR       113
Rsp      EVGMRSLSTPDVVTKALDTLKGKA-RVKRAMWSRRAQEYEQKINSGDLMSIAEVVRDLHR       119
Ccr      TAGLRPLAEGGTVSQALTTLKGRA-RVKRTMWSRRAQEYEAKINSGDLISIAEVVRDLHR       118
          .*:*       :      *:          *  * :     :  *.   :*::  :*

Tth      WEVERGLPDLDRQALRRAIHLLAEEVAQSLEITVQEAKRLFEEAWGEELN  164
Myx      LKGDKDLSFGERKMLDTARSLLIKELSLAKDCSEDEIESDLKKIFNLA--  164
Mtb      RDQERGLSAGEKRMLAKARQILVGELALAESTDDAKAETILDEVLAAAS-  162
Rsp      TDDQREQSYSERQLYEAALERLTREVAAVSGVDEAGAQKAVDAVLVSRAA  169
Ccr      AENQPEQSYSERQLYESALDRMAREVAAIERIDREAAIGILTKSLVKAA-  167
```

B



C

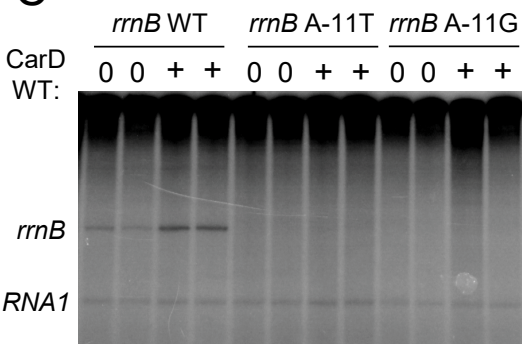


Figure 2.S6: In vitro transcription from *R. sphaeroides rrnA* wild-type and spacer mutant promoters. (*A*) Sequence of the WT *rrnA* promoter (-59 to +1), or spacer mutant with the sequence in red (CCGGGGCC) derived from the corresponding region of the *R. sphaeroides rrnB* promoter replacing the sequence GTTTTTTT. (*B*) Representative gel showing transcripts produced with 20 nM $RNAP_{Rsp}$ and1280 nM CarD where indicated in buffer with 170mM NaCl. The RNA I transcript is from a plasmid-encoded promoter (See *Expanded Materials and Methods*). (*C*) Quantitation of *rrnA* transcripts from gels like that in (B). Values were normalized to that of the wild type promoter in the absence of CarD. Error bars indicate range from two assays. (*D*) Quantitation of fold activation by CarD for either wild-type *rrnA* (lanes 3,4 vs lanes 1,2) or the *rrnA* spacer mutant (lanes 7,8 vs lanes 5,6) from two assays like that in (*B*).

A

rrnA WT

-35 -10 +1

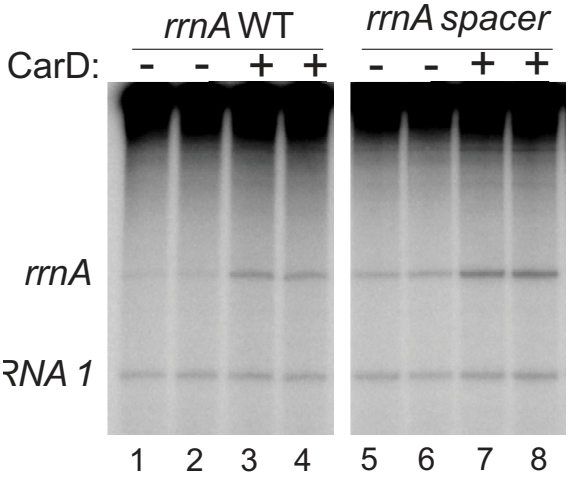
ATCGTCTCTTCGTCATTTTTCCTC**TTGCGG**GTTTTTTTGCGGTTCCC**TAGATA**GCGCCTC**A**

rrnA spacer

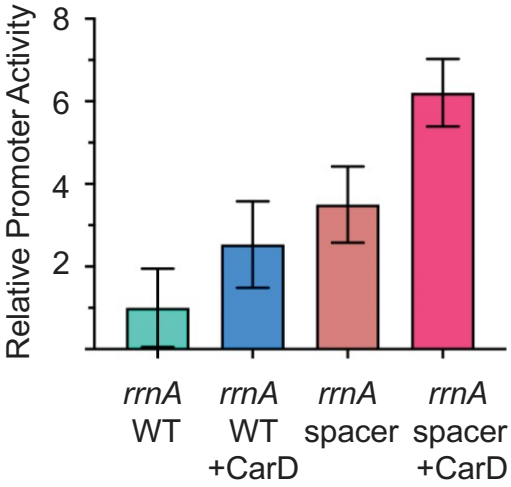
-35 -10 +1

ATCGTCTCTTCGTCATTTTTCCTC**TTGCGG**CCGGGGCCGCGGTTCCC**TAGATA**GCGCCTC**A**

B



C



D

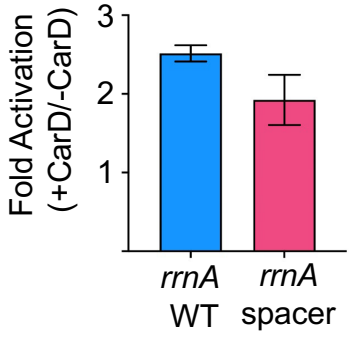
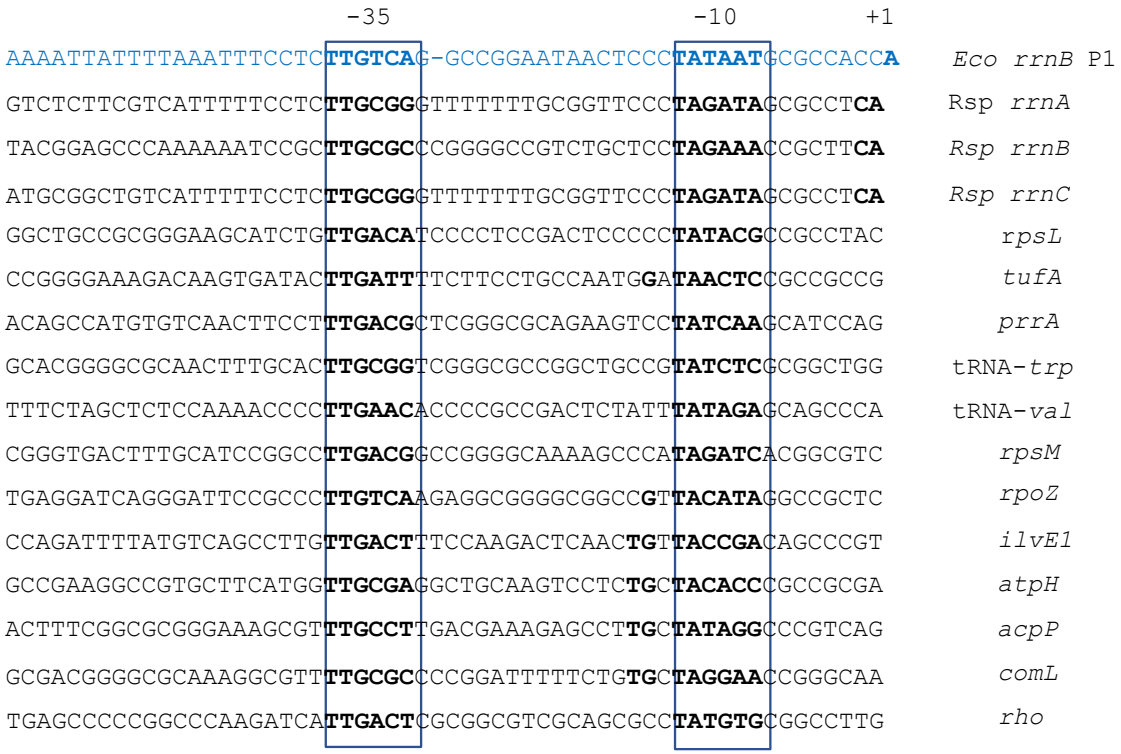


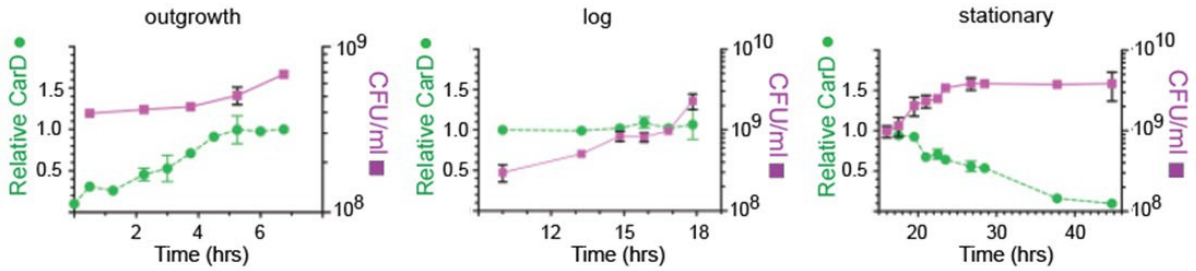
Figure 2.S7: *R. sphaeroides* promoter sequences. (*A*) Promoter sequences (~-60 to +1) from 15 CarD-activated *R. sphaeroides* promoters and *E. coli rrnB* P1, which is activated by CarD when transcribed by *R. sphaeroides* RNAP but not by *E. coli* RNAP (see also Figure 2.S2). (*B*) Promoter sequences of 4 *R. sphaeroides* promoters and 3 *E. coli* promoters (in blue) whose activities are not altered by CarD. Positions corresponding to the -35 and -10 recognition motifs are in bold and boxed. For promoters with 18 bp spacing between the -35 and -10, the sequences are aligned with the -10 hexamer resulting in a shift of the -35 hexamer 1 bp further upstream.

A. Promoters Activated by CarD



B. Promoters Not Activated by CarD

Figure 2.S8: Relative levels of CarD$_{Rsp}$ and CFU/ml in *R. sphaeroides* cultures during aerobic growth. CarD levels and CFU/ml were determined in samples from outgrowth, log phase, and stationary phase cultures and plotted as a function of time after inoculation of the culture. Data in each panel represent average and range from 2 experiments. Outgrowth, log, and stationary phase samples were obtained in separate experiments. Representative Western blots and OD_{600} values for these experiments are shown in Figure 2.7*A*.


outgrowth
log
stationary
Relative CarD
CFU/ml
Time (hrs)
1.5
1.0
0.5
2
4
6
12
15
18
20
30
40
10^9
10^8
10^{10}

EXPANDED MATERIALS AND METHODS

Bacterial strains and growth conditions. Bacterial strains are listed in Table 2.S1. *E. coli* strains were grown at 37°C in either LB Lennox media or on LB agar plates supplemented with kanamycin (30 µg/mL) or chloramphenicol (20 µg/mL) when required. *R. sphaeroides* 2.4.1 strains were grown in succinate-based minimal medium (SIS) (1) at 30°C with kanamycin 25 µg/mL or sucrose (10%) when appropriate. For liquid growth conditions, *R. sphaeroides* 2.4.1 was grown in 500 mL cultures with constant sparging with 69% N_2, 30% O_2 and 1% CO_2 at 30°C for aerobic growth. For aerobic growth curves, samples were removed from the main culture under positive pressure, optical density was determined at OD_{600}, and viable cell counts were determined by plating on solid minimal medium.

Construction of chromosomally encoded *R. sphaeroides* His10-β'. For purification of RNAP by Ni-NTA chromatography, a thrombin-His10 tag was constructed in frame with the C-terminus of the RNAP β' subunit and introduced into wild-type *R. sphaeroides* 2.4.1 using the non-replicative plasmid pk18mobsacB (2) and a two-step recombination method, and the primers are listed in Table 2.S4.

First, a 3 kb fragment containing 1.5 kb of sequence flanking either side of the stop codon of *rpoC* was PCR amplified from chromosomal DNA with primers RLG7576 and RLG7577. The pk18mobsacB vector was PCR linearized with oligonucleotides RLG7578 and RLG7579 that contained sequence overlapping the 3 kb fragment. The 3 kb fragment was then joined to linearized pk18mobsacB by HiFi Assembly (NEB) to form pRLG14688. Insertion of the 3 kb fragment into pk18mobsacB was confirmed with primers RLG7584 and RLG7585.

Next, pRLG14688 was linearized by performing PCR with outward directed primers RLG7594 and RLG7589 lacking the *rpoC* stop codon. The thrombin-His10 tag was produced by annealing complementary oligonucleotides RLG7588 and RLG7589, which were then extended by PCR to create a 130 bp fragment with sequence overlapping the arms of linearized pRLG14688. This fragment was then joined to linearized RLG14688 by Hi-Fi Assembly to create pRLG14713. The resulting plasmid was verified by PCR with primers that annealed to the thrombin-His10 coding sequence (RLG7598) and within the *rpoC* gene (RLG7431) and then sequenced using the oligonucleotides listed in Table 2.S4.

pRLG14713 was conjugated into *R. sphaeroides* 2.4.1 using *E. coli* S17-1 as a donor. Strains containing single crossovers were selected for kanamycin resistance, indicating integration of the pk18mobsacB into the *rpoC* locus. Double crossovers were then selected for the loss of sucrose sensitivity on agar plates containing sucrose (10%), and strains containing the thrombin-His10 coding sequence were identified by PCR using primers RLG7598 and RLG7431. The genomic region flanking the site of homologous recombination was PCR amplified using oligonucleotides RLG7511 and RLG7654, and the PCR fragment was sequenced. Glycerol stocks of the resulting strain RLG14765 were made from single colonies and stored at -80°C.

Purification of *R. sphaeroides* RNAP Holoenzyme. *R. sphaeroides* holoenzyme was purified using Ni-agarose affinity chromatography from aerobically grown cells containing the chromosomally encoded His10-β' (RLG14765). Cells were harvested at OD_{600} ~0.5-0.6 (log phase), and cell pellets were resuspended in Bugbuster (Novagen), 5 mL of Bugbuster per gram of wet cell, supplemented with 5 µL per gram Lysonase

(Novagen), 1X HALT protease inhibitor (ThermoFisher Scientific), and 23 µg/mL phenylmethylsulfonyl fluoride (PMSF). After incubation at room temperature with gentle rocking for 1 hour, cell debris was removed by centrifugation at 14,000 RPM for 20 min at 4°C, and the cleared lysate was then added to a 1 mL Ni-NTA resin column (Qiagen) pre-equilibrated with RNAP wash buffer (50 mM NaH_2PO_4 [pH 8], 300 mM NaCl, 20 mM imidazole, 5% glycerol, 0.1% Tween 20, and 5 mM β-mercaptoethanol). After washing the column with RNAP wash buffer, RNAP was eluted with wash buffer containing 300 mM imidazole, dialyzed twice into 2X RNAP storage buffer (10 mM Tris-Cl [pH 8.0], 100 mM KCl, 10 mM MgCl2, 0.1 mM EDTA, 1 mM DTT, 20 µM ZnCl2) at 4°C, and concentrated using a 5 mL spin concentrator (Sartorius) with a 5 kDa molecular weight cutoff. An equal volume of 100% glycerol was added, and the protein concentration was assayed using the Bradford reagent (Bio-Rad) using BSA as a standard. Proteins were stored at -20°C. Native *E. coli* holo-RNAP was purified as described (3). *R. sphaeroides* σ^{93} was a gift from K. Wassarman (UW-Madison).

As is often observed with RNAP preparations made this way, a small number of other proteins co-purified with the *R. sphaeroides* RNAP. These proteins are present in very low, substoichiometric amounts compared to the other subunits (see Figure S1) and could have co-purified with the RNAP due to weak interactions with RNAP itself (e.g. they could represent true RNAP-associated factors), or they could have co-purified because of affinity for the N-NTA agarose column matrix. Attempts to purify the *R. sphaeroides* RNAP further by heparin-sepharose chromatography were not successful. Although we cannot exclude the possibility that these proteins could influence transcription, they should not affect the interpretation of the results reported here, which

in almost all cases compare the effects of the same RNAP preparation on different promoters or the effects of CarD on the same RNAP preparation.

Construction, overexpression, and purification of SUMO-tagged CarD wild-type and variant proteins. The *R. sphaeroides carD* gene was cloned into the pETSUMO vector for overexpression and purification from *E. coli* BL21 using its IPTG-inducible T7 promoter. The coding sequence of CarD (*rsp_2425*) was produced as a codon optimized gene block (Table 2.2.S5) from GeneArt (Life Technologies). CarD was PCR amplified from the gene block with primers RLG7660 and RLG7661 (Table 2.S6) containing sequence homology to pETSUMO that had been linearized with oligonucleotides RLG7369 and RLG7659. The two DNA fragments were joined by Hi-Fi Assembly (NEB), electroporated into *E. coli* XL1-blue cells, and grown on LB supplemented with kanamycin (30 μg/mL). Plasmids with the correct insert were identified by PCR with oligonucleotides RLG4583 and RLG7386 and then sequenced with the oligonucleotides listed in Table 2.S6. The resulting plasmid (pRLG15433) created an Ulp1 cleavable His10-SUMO tag on the N-terminus of CarD. Substitutions were introduced by Multi-site Lightning QuikChange (Agilent) using oligonucleotides listed in Table 2.S6 and verified by sequencing.

His10-SUMO-tagged CarD was overexpressed from BL21 (DE3) pLysS cells transformed with pETSUMO coding for CarD or CarD variants. Transformed cells were grown in 500 mL of LB containing kanamycin and chloramphenicol inoculated from an overnight culture to a starting OD_{600} of 0.05. Cultures were grown at 37°C to OD_{600} of 0.5-0.6, induced with 1 mM IPTG for 3-4 hr, and purified by Ni-NTA chromatography as described (4). Cell pellets were resuspended in 25 mL of lysis buffer (20 mM Tris-Cl pH

8, 500 mM NaCl, 10% glycerol (vol/vol), 0.5 mM β-mercaptoethanol), PMSF (23 μg/ml), 0.08% sodium deoxycholate and 1X HALT protease inhibitor cocktail (ThermoFisher Scientific), and lysed by sonication. Cell debris was pelleted by centrifugation at 15,000 RPM for 30 min at 4°C, and the cleared lysate was filtered through a 0.45 μm filter added to a 0.5 mL Ni-NTA agarose resin (Qiagen) pre-equilibrated with lysis buffer. The protein-Ni-NTA column was washed with lysis buffer containing 25 mM imidazole, and the protein was eluted with lysis buffer containing 300 mM imidazole. The concentration of the eluted His10-SUMO-tagged CarD was determined using the Bradford assay reagent with bovine serum albumen (BSA) as a standard. The His10-SUMO tag was removed by incubating the eluted His10-SUMO-tagged CarD with His8-Ulp1 protease at a mass ratio of 400:1 (His10-SUMO-tagged CarD to Ulp1 protease) and dialyzed against Ulp1 cleavage buffer (20 mM Tris (pH 8), 150 mM NaCl, and 1 mM β-mercaptoethanol) at 4°C overnight. Free His10-SUMO and His8-Ulp1 protease were removed by passage through a ~300 μL Ni-NTA column. The flow-through containing untagged CarD was dialyzed overnight against CarD storage buffer (20 mM Tris-HCl [pH 8.0], 100 mM NaCl, 2 mM DTT, 50% glycerol). Protein concentration was assayed by the Bradford assay. Proteins were stored at -20°C. For in vitro transcription assays, CarD was diluted into CarD storage buffer to the desired concentration.

We cannot eliminate the possibility that the reduction in activity of the CarD mutant proteins in our studies (and in previous studies on CarD) resulted from general folding defects. We also note that we could not compare the effects of the triple substitution, $CarD_{Rsp}$ variant Q31A, I33A, R53A, corresponding to *T. thermophilus* $CarD_{Mtb}$ RID domain residues R25, I27, and R47, to those of an analogous triple RID

substitution in *T. thermophilus* CarD, because a triple substitution variant was not constructed in previous studies.

Construction of in vitro transcription templates. *R. sphaeroides* promoter fragments were PCR amplified from *R. sphaeroides* 2.4.1 chromosomal DNA and inserted into pRLG770 (5) using the oligonucleotides listed in Table 2.S2. Each oligonucleotide set created promoter fragments with the described endpoints relative to the predicted transcription start site based on the presence of predicted -10 and -35 elements. The resulting promoter fragments contained sequences for ligation to linearized pRLG770 (amplified using oligonucleotides RLG7057 and RLG7056) by Hi-Fi Assembly (NEB). Plasmids were electroporated into *E. coli* XL1-blue, and those confirmed to have the correct insert size on gels were sequenced with oligonucleotides RLG4252 and RLG4253. Promoter variants were created by Multi-site Lightning QuikChange (Agilent) as described above using the mutagenic primers listed in Table 2.S2. Promoter insertions in pRLG15507 and pRLG15543 (Table 2.S5) were gene blocks from IDT that were ligated to linearized pRLG770. pRLG770 and its derivatives all contain the promoter and gene coding for the RNA I transcript, a naturally occurring product of ColE1 plasmids from which these plasmids derive. The RNA I promoter is located within the plasmid replication control region, and makes RNA I, an antisense regulatory RNA that controls the primer for DNA replication, RNA II (6). pRLG15474 containing the *T7A1lacO34* promoter was amplified from pIND4 (a replicable plasmid used for protein overexpression in *R. sphaeroides*) (7). *T7A1lacO34* is sometimes referred to as *T7A1lacO-1* (8). Plasmid DNA was purified from *E. coli* XL1-blue by miniprep (Qiagen), phenol extracted, diluted to 200 ng/µL in 10 mM Tris-HCl (pH 8.0), and stored at -20°C.

in vitro transcription. Multiple round transcription assays were carried out on supercoiled plasmid templates containing the promoters listed in Table 2.S3 using *R. sphaeroides* RNAP (20-50 nM) or *E. coli* RNAP (10 nM) in 25 µL reactions. Template DNA (5 pmol or 400 ng) was incubated in transcription buffer (20 mM Tris-HCl [pH 7.9], 10 mM MgCl2, 1 mM DTT, 0.1 mg/ml BSA, and either 170 or 100 mM NaCl, as indicated, containing NTPs (200 µM ATP, GTP, CTP, 10 µM UTP, and 1.2 µCi α-^{32}P UTP [Perkin Elmer]) with CarD or CarD storage buffer as indicated at room temperature for 10 min. Transcription was initiated by addition of RNAP and reactions were incubated at 30°C for 20 min. The reactions were terminated by addition of 25 µL water-saturated phenol, aqueous and phenol phases were separated by brief centrifugation, and 16 µL of the aqueous phase containing the transcript was transferred to 16 µL 2X loading solution (8 M urea, 10 mM EDTA, 2X TBE, 1% SDS and 0.02% bromophenol blue). Transcripts were denatured by heating the samples at 95°C for 1 min, resolved on 6% polyacrylamide 7M urea gels, imaged using a Typhoon phosphorimager, and quantified with ImageQuant 5.2 software.

Western blot analysis. *R. sphaeroides* cells were harvested from aerobically grown cultures at the OD_{600} indicated. Separate cultures were grown for each of the three growth phases, outgrowth, log, and stationary (Figures 2.7 and 2.S8). For outgrowth cultures, an aliquot of a stationary phase culture (OD_{600} ~3.5) was diluted 1:12.5 into fresh minimal medium to a starting OD_{600} of 0.25 at time 0, and samples were removed at the indicated times (from t=0 to 7 h). For the log phase culture, an aliquot of a stationary phase culture was diluted 1:50 to a starting OD_{600} of 0.02 into fresh medium at time 0, the culture was grown for 10 hr to an OD_{600} of ~0.2 before sampling, and

samples were then collected at intervals from 10 to 18 hr, during which time the culture grew exponentially. For the stationary culture experiment, the culture was inoculated from a stationary phase culture to a starting OD_{600} of ~0.013 at t=0 and was sampled at times between 15 to 45 h, during which it entered and continued through stationary phase.

At the indicated OD_{600}, 2 OD equivalents of cells were pelleted, resuspended in 200 µL Laemmli loading buffer (75 mM Tris-HCl [pH 8], 2% SDS, 0.5 M β-mercaptoethanol, 10% glycerol, 0.2% Triton X100, 0.02% bromophenol blue), and heated at 95°C for 10 min. The whole cell lysates were cleared by brief centrifugation, and aliquots (10 µL/lane or 0.1 OD_{600} equivalent) and purified standards were separated on 12% NuPAGE Bis-Tris gels (Invitrogen) in MES buffer. Proteins were then transferred onto a nitrocellulose membrane (GE Healthcare). The section of the membrane containing the CarD band was incubated for 2 hr with a primary antibody against CarD (rabbit polyclonal antibody raised against untagged CarD; Covance Research Products), and the section containing the α subunit of RNAP was incubated for 1 hr with a primary antibody against α at a 1:10,000 dilution at room temperature in blocking buffer (phosphate buffered saline [pH 7.4], 0.1% Tween20 (PBST) and 5% dried milk). In Figure 2.S1*B*, sections of the membrane were incubated with antibodies against the σ^{70} or the α subunits of RNAP to determine if the *R. sphaeroides* RNAP contained σ^{93} (with which the *E. coli* anti-σ^{70} cross reacts). The membranes were subsequently washed 3 times with PBST and then incubated with secondary antibody (goat anti-rabbit IgG horseradish peroxidase conjugate [Promega]; added at 1:10,000) for 1 hr in blocking buffer. The membrane was then washed 3 times with PBST again,

incubated in SuperSignal West Dura Extended Duration Substrate (Thermo Scientific), and bands were detected with a LI-COR Odyssey Fc Imaging System. The intensity of each band was quantified using Image Studio Lite software (LI-COR). Polyclonal antibodies to the α subunit of *R. sphaeroides* RNAP (WI515) and σ^{70} (WI168) were gifts from Karen Wassarman (University of Wisconsin-Madison). CarD protein concentrations in the whole cell lysate were determined from a standard curve.

Viable cell counts. *R. sphaeroides* colony forming units (CFU) were determined by 10-fold serially diluting cultures into 1X M9 salts at the OD_{600} indicated and plating 10 uL (in triplicate) onto 1X SIS agar plate. Viable cells were scored after 48-96 hr at 30°C. The number of viable cells loaded/lane (CFU/lane) was determined by dividing the CFU by the respective OD_{600} value multiplied by 0.1 (the OD_{600} equivalent loaded per lane).

Determination of CarD concentration in cells. The molar concentration of CarD in log phase cells was calculated from the number of moles of CarD in the samples, as determined from Western blots with purified CarD standards using a molecular weight of 19.045 g/mol for CarD, the number of cells per sample (CFU) determined from viable cell counts, and the volume of a log phase aerobically grown *R. sphaeroides* cell (0.87 μm^3 per CFU) (9).

Bioinformatic analysis of transcription start sites and promoter elements.

Total *R. sphaeroides* RNA was extracted from separate cultures grown to mid-log phase aerobically or anaerobically, as described in the “Bacterial strains and growth conditions” section above and in reference (9). The RNA was then processed as described in reference (10), and transcription start sites (TSSs) were identified as described in reference (11). Both the anaerobic and aerobic

sets of TSSs were then combined for promoter analysis. All TSSs were included for identification of promoter elements; the search did not exclude promoters used by holoenzymes containing sigma factors other than the major sigma, σ^{93} in *R. sphaeroides*. However, since most alternative sigmas are produced or activated only under specific stress conditions, most of promoters identified under the mid-log growth conditions we used would be expected to correspond to those utilizing σ^{93} holoenzyme.

For Figure 2.6, we searched for probable -10 elements by analyzing the DNA sequences from -19 to -5 relative to each TSS (+1) using MEME version 5.1.0 (12) ("zoops" method, minimum width of 9 bp, maximum width of 10 bp, no palindromic sequence). TA motifs (corresponding to positions most frequently found at -12 and -11 at the upstream end of *E. coli* -10 elements) were identified in 763 putative promoters. The TA motifs were not required to be positioned at -12 and -11 relative to the TSS, since the spacing between the -10 hexamer and the TSS can vary. The base at position -7 in *R. sphaeroides* promoters was identified relative to the position of the TA motif. In cases where there were multiple closely-spaced TSSs associated with the same -10 element (~9% of genes), only one -10 element was used for quantitation of the percentage of -7T and extended -10 elements in the population. In cases where there was more than one TSS associated with the same gene, but they were not close together, the corresponding -10 elements for each putative promoter were included in the analysis. Since the genome of *R. sphaeroides* does not contain a gene coding for a σ^{S} homolog (13), and alternative holoenzymes (at least in *E. coli*) do not

contain a TA motif at the upstream end of their -10 elements, the promoter collection analyzed in Figure 2.6 for -7T and extended -10 elements likely consists primarily of σ^{93}-dependent promoters.

The -10 elements in our promoter library were separated into two groups based on the identity of the base at position -7, one group with -7T and the other with -7A/G/C combined. The percentage of promoters containing -7T, extended -10 elements, or both was determined for each species in Figure 2.6 using published TSS information. The list of all *R. sphaeroides* TSSs, indicating those with probable -7T and extended -10 sequences, is provided as Dataset 1 (www.pnas.org/cgi/doi/10.1073/pnas.2010087117). We note that the number of *R. sphaeroides* promoters identified by MEME in our analysis is an underestimate of the true number of *R. sphaeroides* promoters, because we specified -10 elements with both a T at -12 and an A at -11. The same methods were used to calculate the percentages of promoters with -7T, extended -10 elements, or both for each bacterial species shown in Figure 2.6. The sources of the TSS data used for this purpose were *R. sphaeroides* (11), *C. crescentus* (14), *Z. mobilis* (10), *E. coli* (15), *B. subtilis* (16), *M. smegmatis* (17), *S. coelicolor* (18), and *B. cenocepacia* (19). Using our criteria, we estimate that ~65% of the TSSs were included in the analyses of -10 elements from each species.

Table 2.S1. Strains and plasmids used in this study

Strains	**Genotype**	**Source**
BL21 (DE3) pLysS	*E. coli* str. B F^- *ompT gal dcm lon hsdS*$_B$(r_B^- m_B^-) λ(DE3 [*lacI lacUV5-T7p07 ind1 sam7 nin5*]) [*malB*$^+$]$_{K\text{-}12}$(λ^S) pLysS[*T7p20 ori*$_{p15A}$](Cm^R)	Novagen
E. coli S17-1	TpR SmR recA, thi, pro, hsdR-M+RP4: 2-Tc:Mu: Km Tn7 λpir	(20)
E. coli XL1-Blue	Δ(*mcrA*)*183* Δ(*mcrCB-hsdSMR-mrr*)*173 endA1 supE44 thi-1 recA1* gyrA96 relA1 lac *[F'* proAB $lacI^qZ\Delta M15$ *Tn*10 *(Tet^r)]*	Stratagene
E. coli XL10-Gold	*endA1 glnV44 recA1 thi-1 gyrA96 relA1 lac Hte Δ(mcrA)183 Δ(mcrCB-hsdSMR-mrr)173 tet^R F'[proAB lacI^{q}ZΔM15 Tn10(Tet^R Amy Cm^R)]*	Stratagene
RLG14946	*Rhodobacter sphaeroides* 2.4.1 (Wild type lab strain)	
RLG14765	RLG14946 with a C-terminal thrombin-His_{10} chromosomally tagged β' (*rpoC*)	This study

Plasmids	**Description**	**Source**
pRLG770	Transcription vector	(21)
pETSUMO	pET derived plasmid that encodes a Ulp1 cleavable His_{10}SUMO tag. ORFs are cloned in frame with the coding sequence of the SUMO tag.	Invitrogen (4)
pK18mobsacB	Mobilization plasmid used for making chromosomal insertion into *R. sphaeroides* chromosome	(2)

Table 2.S2. Oligonucleotides used to construct plasmid templates for in vitro transcription

Primers for linearizing RLG770

Oligonucleotide	Sequence
7057	5'-AAGCTTGGGTCCCACCTGACC-3'
7058	5'-GAATTCTTGAAGACGAAAGGGCC-3'

Primers for screening and sequencing promoters inserted into RLG770

Oligonucleotide	Sequence
4252	5'-CCCGAAAAGTGCCACCTGACG-3'
4253	5'-CAGTTCCCTACTCTCGCATGG-3'

Primers to construct *in vitro* transcription plasmids
(both forward and reverse primers)

Plasmid	Oligo	Sequence	Description
pRLG14562	7411	GGCCCTTTCGTCTTCAAGAATTC GATCCGGCGGATGGGAGCGGC	*R. sphaeroides rrnA* promoter, -320 to +50
	7412	GGTCAGGTGGGACCCAAGCTTG GCAGCACCGCCGCCTCTCAAC	
pRLG14564	7413	GGCCCTTTCGTCTTCAAGAATTC CCGCGAGCGGCTGACCGG	*R. sphaeroides rrnB* promoter, -230 to +50
	7415	GGTCAGGTGGGACCCAAGCTTCT GTCATCCGCTCCCTCGTCTCCGT CC	
pRLG14566	7416	GGCCCTTTCGTCTTCAAGAATTC GCCGTCTCGAGGGGGATCTGGA G	*R. sphaeroides rrnC* promoter, -270 to +50
	7412	GGTCAGGTGGGACCCAAGCTTG GCAGCACCGCCGCCTCTCAAC	
pRLG15054	8008	GGCCCTTTCGTCTTCAAGAATTC CATCCTGACGGAATGGGGCATC	*R. sphaeroides prrA* promoter, -379 to +51
	8009	GGTCAGGTGGGACCCAAGCTTA GTTCGAATACCAGATCCTCAGCC	
pRLG15224	8026	GGCCCTTTCGTCTTCAAGAATTC GGGCTGGCCCTTCTGGGATTTG	*R. sphaeroides rpsL* promoter, -132 to +65
	8136	GGTCAGGTGGGACCCAAGCTTG CCCGGATCATGACCACTACAG	
pRLG15228	8028	GGCCCTTTCGTCTTCAAGAATTCT CATCGGCTACTACTATGCCAAGG	*R. sphaeroides rpsM* promoter, -156 to +50
	8143	GGTCAGGTGGGACCCAAGCTTG ACTGTTCGCCGCAGGACAAAACC	
pRLG15227	8141	GGCCCTTTCGTCTTCAAGAATTC GTTGATGGCGGGCGATCTCGAG	*R. sphaeroides* $tRNA^{trp}$ RSP_4298 promoter, -106 to +56
	8142	GGTCAGGTGGGACCCAAGCTTCA CGACCCTCGGTTTTGGAGACC	

pRLG15229	8144	GGCCCTTTCGTCTTCAAGAATTC CCGCTCGAGGCTTGAGGATC	*R. sphaeroides rpoZ* promoter, -68 to +62
	8145	GGTCAGGTGGGACCCAAGCTTCA ACGCAGTCTTCAACCGTCACG	
pRLG15230	8146	GGCCCTTTCGTCTTCAAGAATTC GCGCCGGGGAAAGACAAGTGAT AC	*R. sphaeroides tufA* promoter, -60 to +25
	8147	GGTCAGGTGGGACCCAAGCTTC GACACCGGACAGCCCGTTC	
pRLG15472	8303	GGCCCTTTCGTCTTCAAGAATTCT CACCGCCCGATAGTGCTTC	*R. sphaeroides tRNA-val* (*RSP*_4309) promoter, -79 to +47
	8304	GGTCAGGTGGGACCCAAGCTTCA CACCTGTTCCGCGCGAG	
pRLG15474	8309	GGCCCTTTCGTCTTCAAGAATTC CCTATAAAAATAGGCGTATCACG AGG	Inducible pT7A1lacO34 promoter, -102 to +31 IPTG inducible promoter in pIND4 (5) (for overexpression in *R. sphaeroides*)
	8310	GGTCAGGTGGGACCCAAGCTTG ATGTGTGAAATTGTTATCCGCTC	
pRLG15601	8569	GGCCCTTTCGTCTTCAAGAATTC GATGCTGTGAAAAAGGCCGATGG	*R. sphaeroides acpP* promoter, -89 to +46
	8570	GGTCAGGTGGGACCCAAGCTTCA GGCAACTGGTGCAGTTCAG	
pRLG15602	8573	GGCCCTTTCGTCTTCAAGAATTC GGTCGGACGTCCCTGCACTG	*R. sphaeroides comL* promoter, -88 to +35
	8574	GGTCAGGTGGGACCCAAGCTTC CTCTGCCCCGTCAGACCATAG	
pRLG15604	8581	GGCCCTTTCGTCTTCAAGAATTC GGTGCTCCTTGCGACATGATCTT C	*R. sphaeroides ilvE1* promoter, -95 to +80
	8582	GGTCAGGTGGGACCCAAGCTTCA TGTCACCCTCCTGCGCATAG	
pRLG15606	8589	GGCCCTTTCGTCTTCAAGAATTC GTCCGGGGGCTCTACAGGATTC	*R. sphaeroides RSP*_1486 promoter, -78 to +30
	8590	GGTCAGGTGGGACCCAAGCTTGT CCGTCGCCTGTCGTCATTG	
pRLG15607	8591	GGCCCTTTCGTCTTCAAGAATTC CAGACAGCGGTCGCGGATGAG	*R. sphaeroides RSP*_3110 promoter, -90 to +72
	8592	GGTCAGGTGGGACCCAAGCTTC GAGCATGTCGGCGACCTTTCC	

pRLG15611	8603	GGCCCTTTCGTCTTCAAGAATTC GTCTTCTCCGCTGTGGCGAAATG	*R. sphaeroides rho* promoter, -94 to +110
	8604	GGTCAGGTGGGACCCAAGCTTG AGGTTGAGACGTTCGTTCATGGC	
pRLG15613	8571	GGCCCTTTCGTCTTCAAGAATTC GCGGAGGACCAAGGAAACAATG G	*R. sphaeroides atpH* promoter, -91 to +49
	8572	GGTCAGGTGGGACCCAAGCTTCA GAGGCATTGTATGGCGCGAG	
pRLG15619	8597	GGCCCTTTCGTCTTCAAGAATTC GATTGCGTGCGTCGACGGTTTTC	*R. sphaeroides putR* promoter, -100 to +40
	8598	GGTCAGGTGGGACCCAAGCTTG GTCGATGGGGTCCAGTTTCGATC	

Mutagenic primers for construction of transcription templates

Plasmid	**Oligo**	**Sequence** (mutations are in lowercase)	**Description**
pRLG14563	7364	TGCGGTTCCCTAGATtGC GCCTCACCGAAGC	*R. sphaeroides rrnA* A-7T promoter
pRLG14565	7365	CCGTCTGCTCCTAGAAtC CGCTTCACCGAGACG	*R. sphaeroides rrnB* A-7T promoter
pRLG14567	7364	TGCGGTTCCCTAGATtGC GCCTCACCGAAGC	*R. sphaeroides rrnC* A-7T promoter
pRLG15180	8148	CGCCCGGGGCCGTCTTg tgCTAGAAACCGCTTCAC	*R. sphaeroides rrnB* ext-10 promoter (to -17 to -14 TGTG)
pRLG15181	8149	CCAAAAAATCCGCTTGac aCCGGGGCCGTCTGCTC	*R. sphaeroides rrnB* -35 consensus promoter
pRLG15182	8150	GTCTGCTCCTAGAAACat aTTCACCGAGACGAAGA C	*R. sphaeroides rrnB* discriminator mutant promoter (-5 to -3, CCG to ATA)
pRLG15183	8151	GTCTGCTCCTAGAAAggG CTTCACCGAGACGAAG	*R. sphaeroides rrnB* discriminator mutant promoter (-6 to -5, CC to GG)
pRLG15185	8153	CTTGCGCCCGGGGCCG aTCTGCTCCTAGAAACC G	*R. sphaeroides rrnB* promoter with 18 bp spacer. insertion of A at -21
pRLG15186	8154	CCGCTTGCGCCCGGGG CGTCTGCTCCTAGAAAC C	*R. sphaeroides rrnB* promoter with 16 bp spacer, deletion of C at -21

pRLG15350	8287	GGCCGTCTGCTCCTtGAAACCGCTTCACC	*R. sphaeroides rrnB* promoter mutant, A-11T
pRLG15351	8288	GGCCGTCTGCTCCTgGAAACCGCTTCACC	*R. sphaeroides rrnB* promoter mutant, A-11G
pRLG15352	8289	GCCGTCTGCTCCTAtAAACCGCTTCACCG	*R. sphaeroides rrnB* promoter mutant, G-10T
pRLG15485	8383	TGCGGGTTTTTTTGCGGTgtgCTAGATAGCGCCTCACCG	*R. sphaeroides rrnA*, ext-10 promoter mutant GTG from -16 to -14
pRLG15488	8376	GGGCGGCCGTTACATtGGCCGCTC	*R. sphaeroides rpoZ* promoter mutant, A-7T
pRLG15489	8377	GAGGCGGGGCGGCtGcTACATAGGCCGCT	*R. sphaeroides rpoZ* ext-10 consensus TGC from -15 to -13
pRLG15490	8378	GGGGCAAAAGCCCATAGATtACGGCGTCCC	*R. sphaeroides rpsM* promoter mutant, C-7T
pRLG15491	8379	GGCCGGGGCAAAAGCtgcTAGATCACGGCGTCC	*R. sphaeroides rpsM* ext-10 consensus, TGC from -15 to -13
pRLG15505	8381	TCTCTTCGTCATTTTTCCTCTTGCGGccggggccGCGGTTCCCTAGATAGCGCCTCACCG	*R. sphaeroides rrnA* promoter, CCGGGGCC spacer substitution from -29 to -22

Table 2.S3. Plasmids used for in vitro transcription

Plasmids	Promoter[a]	Endpoints[b]	Source
pRLG770	No promoter		(21)
pRLG1616	*E. coli rrnB* (*rrnB* P1)	-88 to +50	(21)
pRLG593	*lacUV5*	-59 to +38	(22)
pRLG3265	*RNA I*	-61 to +1	(22)
pRLG14562	*R. sphaeroides rrnA*	-320 to +50	This study
pRLG14563	*R. sphaeroides rrnA* A-7T	-320 to +50	This study
pRLG14564	*R. sphaeroides rrnB*	-230 to +50	This study
pRLG14565	*R. sphaeroides rrnB* A-7T	-230 to +50	This study
pRLG14566	*R. sphaeroides rrnC*	-270 to +50	This study
pRLG14567	*R. sphaeroides rrnC* A-7T	-270 to +50	This study
pRLG15054	*R. sphaeroides prrA*	-379 to +51	This study
pRLG15224	*R. sphaeroides rpsL*	-132 to +65	This study
pRLG15227	*R. sphaeroides tRNA^{trp}* (*RSP*_4298)	-106 to +56	This study
pRLG15228	*R. sphaeroides rpsM*	-156 to +50	This study
pRLG15229	*R. sphaeroides rpoZ*	-68 to +62	This study
pRLG15230	*R. sphaeroides tufA*	-60 to +25	This study
pRLG15350	*R. sphaeroides rrnB* A-11T	-230 to +50	This study
pRLG15351	*R. sphaeroides rrnB* A-11G	-230 to +50	This study
pRLG15352	*R. sphaeroides rrnB* G-10T	-230 to +50	This study
pRLG15472	*R. sphaeroides tRNA^{val}* (*RSP*_4309)	-79 to +47	This study
pRLG15474	*T7A1lacO34 =A1lacO-1*	-102 to +31	(8)
pRLG15485	*R. sphaeroides rrnA*, ext-10 GTG -16 to -14	-320 to +50	This study
pRLG15488	*R. sphaeroides rpoZ* A-7T	-68 to +62	This study
pRLG15489	*R. sphaeroides rpoZ* ext-10 TGC -15 to -13	-68 to +62	This study
pRLG15490	*R. sphaeroides rpsM* C-7T	-156 to +50	This study
pRLG15491	*R. sphaeroides rpsM* ext-10 TGC -15 to -13	-156 to +50	This study
pRLG15505	*R. sphaeroides rrnA* CCGGGGCC from -29 to -22	-320 to +50	This study
pRLG15507	*R. sphaeroides ssrA*	-71 to +55	This study
pRLG15543	*R. sphaeroides rrnB* full consensus UP element	-230 to +50	This study
pRLG15601	*R. sphaeroides acpP*	-89 to +46	This study
pRLG15602	*R. sphaeroides comL*	-88 to +35	This study
pRLG15604	*R. sphaeroides ilvE1*	-95 to +80	This study
pRLG15606	*R. sphaeroides RSP*_1486 (TetR transcription regulator family)	-78 to +30	This study
pRLG15607	*R. sphaeroides RSP*_3110 (putative Glutathione S-transferase)	-90 to +72	This study
pRLG15611	*R. sphaeroides rho*	-94 to +110	This study

pRLG15613	*R. sphaeroides atpH*	-91 to +49	This study
pRLG15619	*R. sphaeroides putR*	-100 to +40	This study

[a] Promoters are named for the first gene in the operon
[b] endpoints are relative to the predicted transcription start site (unless previously published otherwise in the indicated source)

Table 2.S4. Primers used to construct chromosomally-encoded C-terminal thrombin-His10 tagged β'

1. Insertion of *rpoC* chromosomal fragment into pk18mobsacB (pRLG14688)

Primers to isolate *rpoC* chromosomal fragment

Oligo	Sequence
7576	GCCGCCCAGTCGATCGG
7577	AGGCGAACCTTGGCGACCTTAC

Primers to linearized pK18mobsacB

Oligo	Sequence
7578	CGCTATGCGTAAGGTCGCCAAGGTTCGCCTAAGCTTGG CACTGGCCGTC
7579	CGTACCCGGTTCGCCGATCGACTGGGCGGCGGATCCC CGGGTACCGAG

Primers to confirm 3 kb insert in pK18mobsacB

Oligo	Sequence
7584	CACTCATTAGGCACCCCAGGC
7585	CGCCAGGGTTTTCCCAGTCAC

2. Creation of C-terminally tagged *rpoC* with thrombin-His_{10} in pK18mobsacB (pRLG14713)

Primers to linearize pK18mobsacB with *rpoC* chromosomal fragment

Oligonucleotide	Sequence
7594	CCTGACCATGGGCCCGC
7595	CAGCCCCGAGTCCTCGG

Primers to introduce thrombin-His_{10} into *rpoC* C-terminus (Region of complementarity is underlined)

Oligonucleotide	Sequence
7588	CCGAGGACTCGGGGCTGGTGGAAACGGTGGAGAACCGCGAG GAGCTGGTGCCGCGCGGCA<u>GCCATCACCACCATCATCAC</u>
7589	TGCGCGGGCCCATGGTCAGGCGAGAGAGGGGAGGATTAATG GTGGTGATG<u>GTGATGATGGTGGTGATGGC</u>

Primers to confirm insertion of thrombin-His_{10} tag onto *rpoC* C-terminus

Oligo	Sequence
7598	ATGATGGTGGTGATGGCTGCC
7431	GCGTCAAGAAGATCGCGCAC

Primers to sequence RLG14713 and RLG14765

Oligo	Sequence
7431	GCGTCAAGAAGATCGCGCAC
7432	GTGATCCCGAGCAGGATCGGTTC
7512	TGCCGATCATCGCCGAGAAGG
7513	GAACCGATCCTGCTCGGGATC
7515	CGGGTTGCCGTCCATGATG
7516	GCCAGTCGGACACGATCTTCTG
7584	CACTCATTAGGCACCCCAGGC
7594	CCTGACCATGGGCCCGC
7602	CGAGAGTCGCACGAAATCCTCC
7612	GGGCCTCTTCGCTATTACGCC
7615	CGTGCTGAAGGAGCCACTGG
7616	CCCTCCGACTCCCCCTATACG
7617	GCGCCAGATCGTGAAGAGC
7618	CGTACGATGAGCAGGACGCC

Primers to amplify crossover region in RLG14765

Oligo		Sequence
7511		GCTCACCTGTGAAGCGGAAGAG
7654		GATCTTTGACACCCTGGGTATCGAGG

Table 2.S5. Gene blocks used in this study

Plasmid	Sequence[a]	Description
pRLG15507	**CACGAGGCCCTTTCGTCTTCAAGAATTC**TG CCCCACGGATCCGCCCCGGCACGCCTGCG GCCGCTTGCGGAGGCCGCCCCTCCCGCCT ATATTCCGCCTGTCCGGGGCAACCCGGACT ATGGACATAAACGCGCTCGTCATAAGCGGA TCGGAC**AAGCTTGGGTCCCACCTGACCCC ATG**	*R. sphaeroides ssrA* promoter -71 to +55
pRLG15543	**GGCCCTTTCGTCTTCAAGAATTCG**gaaaatttttt tcaaaaagtaCTTGCGCCCGGGGCCGTCTGCT CCTAGAAACCGCTTCACCGAGACGAAGACC GGCAGCGCCGGACGGAGACGAGGGAGCG GATGACAG**AAGCTTGGGTCCCACCTGACC**	*R. sphaeroides rrnB* promoter with full consensus UP element (-36 to +50)

Gene block for CarD (*RSP*_2425)

Plasmid	Sequence	Description
pRLG15433	**ATG**ACCAAAACCAAAAAACCGGAATTTCGT CCGAATGAGTTTGTTGTTTATCCGGCACAT GGTGTGGGTCGTATTATTAGCATTGAAGAA CAAGAAATTGCCGGTATTCGTCTGGAACTG TTTGTGATCAGCTTCGAGAAAGATAAAATGA CCCTGCGTGTTCCGACCCATAAAGCAACCG AAGTTGGTATGCGTAGCCTGAGCACACCGG ATGTTGTTACCAAAGCACTGGATACCCTGA AAGGTAAAGCACGTGTGAAACGTGCAATGT GGTCACGTCGTGCACAAGAATATGAACAGA AAATTAACAGCGGTGATCTGATGAGCATTG CCGAAGTTGTTCGTGATCTGCATCGTACCG ATGATCAGCGTGAACAGAGCTATAGCGAAC GTCAGCTGTATGAAGCAGCACTGGAACGTC TGACCCGTGAAGTTGCAGCAGTTAGCGGTG TTGATGAAGCCGGTGCACAGAAAGCAGTTG ATGCAGTTCTGGTTAGCCGTGCAGCA**TAA**	Sequence of codon optimized (GeneArt, Life Technologies) *R. sphaeroides* CarD (*RSP*_2425)

[a] Promoter sequences in these plasmids were gene blocks (IDT) inserted into RLG770 by HiFi Assembly (NEB). Sequence homology to RLG770 is in boldface. -35 element to -10 element is underlined. Consensus UP element introduced into *R. sphaeroides rrnB* promoter is shown in lowercase. Start and stop codons of CarD are in boldface.

Table 2.S6. Primers used to construct pETSUMO-CarD

Primers to linearize pETSUMO (pRLG14641)

Oligo	Sequence
7369	TCGTAACAAAGCCCGAAAGGAAGCTG
7659	TCCACCAATCTGTTCTCTGTGAGCC

Primers to PCR amplify CarD with sequence overlapping pETSUMO

Plasmid	Oligo	Sequence	Description
pRLG15433	7660	ACAGAGAACAGATTGGTGGAATGACCAAAACCAAAAAACCGGAATTTCG	Primers to amplify codon optimized CarD, which generates overlapping sequence to pETSUMO (pRLG14641)
	7661	CCTTTCGGGCTTTGTTACGATTATGCTGCACGGCTAACCAGAAC	

Primers to sequence pETSUMO-CarD (pRLG15433)

Oligo	Sequence
7386	ACGCTGCGCGTAACCACC
7837	GACAGGGTAAGGAAATGGACTCC
7839	GCGTGAACAGAGCTATAGCGAACG
4583	GGTGATGTCGGCGATATAGGC
7840	CATCCGGTGTGCTCAGGCTAC
7838	GGTGTCCGATGGATCTTCAGAGATC

Mutagenic primers to construct CarD variants

Plasmid	Oligo	Sequence	Description
pRLG15053	8058	GTGTGAAACGTGCAATGGCGTCACGTCGTGCACAAG	CarD W91A substitution
pRLG15434	8195	TGTGAAACGTGCAATGTTGTCACGTCGTGCACAAG	CarD W91L substitution
pRLG15435	8190	TGGTGTGGGTCGTATTATTAGCATTGAAGAAGCAGAAGCTGCCGGTATTCGTCTG	CarD Q31A/I33A/ R53A substitution
	8191	GAAAGATAAAATGACCCTGGCTGTTCCGACCCATAAAGCA	

Legend for Dataset 1

The spreadsheet contains the chromosome location (A), locus tag (B), gene name if known (C), proposed product (D), position of the transcription start site (E), whether the start site is on the + or – strand (F), promoter sequence from -20 to +1 corresponding to the TSS (G), the sequence of the predicted extended -10 and -10 hexamer (H), and the identity of the base at position -7 (I). Detailed information about the bioinformatic analysis is in *Expanded Materials and Methods*. The dataset was omitted for brevity as it is published. It can be accessed online: www.pnas.org/cgi/doi/10.1073/pnas.2010087117

SUPPLEMENTAL REFERENCES

1. W.R. Sistrom, A requirement for sodium in the growth of R*hodopseudomonas spheroides*, *J Gen Microbiol* **22**:778–785 (1960).
2. A. Schäfer *et al*, Small mobilizable multi-purpose cloning vectors derived from the *Escherichia coli* plasmids pK18 and pK19: selection of defined deletions in the chromosome of Corynebacterium glutamicum, *Gene* **145**:69–73 (1994).
3. R.R. Burgess, J.J. Jendrisak, Procedure for the rapid, large-scale purification of Escherichia coli DNA-dependent RNA polymerase involving polymin P precipitation and DNA-cellulose chromatography. *Biochemistry* **14**:4634–4638 (1975).
4. D.B. Srivastava *et al*, Structure and function of CarD, an essential mycobacterial transcription factor, *Proc Natl Acad Sci USA* **110**:12619-12624 (2013).
5. W. Ross, J.F. Thompson, J.T. Newlands, R.L. Gourse, *E.coli* Fis protein activates ribosomal RNA transcription in vitro and in vivo, *EMBO J* **9**:3733–3742 (1990).
6. Davison J., Mechanism of control of DNA replication and incompatibility in ColE1-type plasmids--a review. *Gene* **28**:1-15 (1984).
7. A.C. Ind *et al*, Inducible-Expression Plasmid for *Rhodobacter sphaeroides* and *Paracoccus denitrificans*, *Appl Environ Microbiol* **75**:6613–6615 (2009).
8. R. Lutz H. Bujard, Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements. *Nucleic Acids Res* **25**:1203-1210 (1997).
9. K. Lemmer *et al*, The NtrYX two-component system regulates the gram-negative cell envelope, *mBio* **11**:e00957-20.
10. J.M. Vera *et al*, Genome-scale transcription–translation mapping reveals features of *Zymomonas mobilis* transcription units and promoters, *mSystems* **5**:e00250-20 (2020).
11. K.S. Myers *et al*, Genome-wide Identification of Transcription Start Sites for *Rhodobacter sphaeroides* Grown Aerobically and Photosynthetically, *ASM Microbiology Resource Announcements* (2020), in press.
12. T.L. Bailey *et al*, MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res* **37**:W202-8 (2009).
13. C. Mackenzie *et al*, The home stretch, a first analysis of the nearly completed genome of *Rhodobacter sphaeroides* 2.4.1. *Photosynth Res* **70**:19-41 (2001).
14. B. Zhou *et al*, The global regulatory architecture of transcription during the *Caulobacter* cell cycle, *PLoS Genet* **11**:e1004831 (2015).
15. M.K. Thomason *et al*, Global transcriptional start site mapping using differential RNA sequencing reveals novel antisense RNAs in *Escherichia coli*, *J Bacteriol* **197**:18–28 (2015).
16. E. Warman, D. Forrest, J.T. Wade, D.C. Grainger, Widespread divergent transcription from prokaryotic promoters, doi.org/10.1101/2020.01.31.928960doi:

bioRxiv preprint first posted online Feb. 2, 2020.

17. X. Li *et al*, Transcriptome landscape of *Mycobacterium smegmatis*, *Front Microbiol* **8**:2505 (2017).
18. Y. Jeong *et al*, The dynamic transcriptional and translational landscape of the model antibiotic producer *Streptomyces coelicolor* A3(2), *Nat Commun* **7**:11605 (2016).
19. A.M. Sass *et al*, Genome-wide transcription start site profiling in biofilm-grown *Burkholderia cenocepacia* J2315, *BMC Genomics* **16**:775 (2015).
20. R. Simon, U. Priefer, A. Pühler, A broad host range mobilization system for in vivo genetic engineering: transposon mutagenesis in gram negative bacteria, *Nat Biotech* **1**:784–791 (1983).
21. W. Ross, S.E. Aiyar, J. Salomon, R.L. Gourse, *Escherichia coli* Promoters with UP elements of different strengths: modular structure of bacterial promoters, *J Bacteriol* **180**:5375–5383 (1998).
22. T. Gaal, M.S. Bartlett, W. Ross, C.L. Turnbough, R.L. Gourse, Transcription regulation by initiating NTP concentration: rRNA synthesis in bacteria, *Science* **278**: 2092–2097 (1997).
23. Y. Zuo, T.A. Steitz, Crystal structures of the *E. coli* transcription initiation complexes with a complete bubble, *Mol Cell* **58**:534-540 (2015).
24. A. Narayanan *et al*, Cryo-EM structure of *Escherichia coli* σ^{70} RNA polymerase and promoter DNA complex revealed a role of σ non-conserved region during the open complex formation, *J Biol Chem* **293**:7367-7375 (2018).
25. A. Feklistov, S.A. Darst, Structural basis for promoter -10 element recognition by the bacterial RNA polymerase σ subunit, *Cell* **147**:1257-1269 (2011).
26. Y. Zhang *et al*, Structural basis of transcription initiation, *Science* **338**:1076-1080 (2012).
27. B. Bae, A. Feklistov, A. Lass-Napiorkowska, R. Landick, S.A. Darst, Structure of a bacterial RNA polymerase holoenzyme open promoter complex, *eLife* **4**:e08504 (2015).
28. E.A. Hubin, M. Lilic, S.A. Darst, E.A. Campbell, Structural insights into the mycobacteria transcription initiation complex from analysis of X-ray crystal structures, *Nat Commun* **8**:16072 (2017).
29. B. Benoff *et al*, H. Yang, C.L. Lawson, G. Parkinson, J. Liu, E. Blatter, Y.W. Ebright, H.M. Berman, R.H. Ebright, Structural basis of transcription activation: the CAP-alpha CTD-DNA complex, *Science* **297**:1562-1566 (2002).

CHAPTER THREE
Rhodobacter sphaeroides CarD Negatively Autoregulates Its Own Promoter

This chapter is a draft of a manuscript with an expected submission of Spring 2021. I performed all the experiments. Wilma Ross, Richard Gourse and I are writing the manuscript. It focuses primarily on CarD autoregulation of its own promoter.

Abstract

rRNA transcription in the photoheterotrophic α-proteobacterium *Rhodobacter sphaeroides* is dependent on the transcription factor CarD for activity. Bioinformatic analysis showed that a majority of *R. sphaeroides* promoters lack the highly conserved -7T at the last position of the -10 element. This absence of the -7T correlated with activation by CarD in vitro. Here we show that the basal activity of the previously uncharacterized *R. sphaeroides carD* promoter, which lacks a -7T, is nevertheless strong relative to many other *R. sphaeroides* promoters and that CarD inhibits rather than activates transcription from its own promoter. The *carD* promoter produces abortive transcription products in the presence of CarD which reduces the production of full length RNA products, suggesting that CarD inhibits transcription from its promoter by limiting promoter escape.

Introduction

Transcription in bacterial cells is carried out by a multi-subunit DNA-dependent RNA polymerase (1). Bacterial RNA polymerase (RNAP) is comprised of an α dimer, β, β', and ω subunits, as well as a dissociable σ subunit that directs RNAP to recognize specific elements in promoter DNA. When σ-associated RNAP (holoenzyme) binds promoter DNA, it forms a closed complex where promoter DNA is duplex (RPc). Through a series of isomerization steps, promoter DNA is unwound, forming the open complex (RPo). The template strand of the unwound DNA is then positioned into the active site of RNAP, where NTPs are catalyzed to make a growing RNA message (2–4). As the RNA elongates, RNAP breaks its contacts with promoter DNA, σ dissociates, and RNAP escapes the promoter.

Transcription factors have different effects on different promoters. Transcription factors like Fis or CRP can activate or inhibit transcription by utilizing DNA binding sites just upstream of or within the RNAP binding site, respectively (5–7). Alternatively, they can work by binding only to RNAP and not to DNA and exerting their specificity based only on the kinetic features of the promoters themselves (8–11).

DksA (with its cofactor ppGpp) and TraR are examples of the latter class, activating or inhibiting transcription depending on the kinetic features of different promoters (8–11). DksA activates or inhibits multiple genes in *Escherichia coli*, and it also feedback regulates its own promoter in order to keep its concentration constant (12). Although DksA is widespread in proteobacteria, it is not present in all proteobacteria, and it does not appear to be present in firmicutes, thermophiles, mycobacteria, or bacteroidetes.

CarD is a small protein that binds the β lobe of the RNAP β subunit and activates rRNA transcription in vitro in Mycobacterial species, *Thermus thermophilus*, *Myxococcus xanthus*, and *Rhodobacter sphaeroides* (13-21). It binds upstream of the -10 element in promoter DNA and favors open complex formation by stabilizing RNAP-promoter complexes (13–21). The majority of *R. sphaeroides* promoters lack the highly conserved critical thymine nucleotide at the last position of the -10 element, and the *R. sphaeroides* CarD homolog can compensate for the absence of -7T in many promoters, including rRNA and at least some r-protein promoters, in order to activate transcription (21).

carD is an essential gene in Mycobacteria, M*yxococcus xanthus and R. sphaeroides* (14, 18, 22)**,** and its deletion affects several growth properties in *C. crescentus* (23, 24). Depletion of CarD was reported to increase rRNA expression in *M. smegmatis* and *M. tuberculosis* (18), although CarD activates the Mycobacterial rRNA promoter in vitro (13). *carD* depletion or CarD substitutions in *M. tuberculosis* were reported to differentially alter a majority of transcripts in vivo, increasing some and decreasing others, leading to the proposal that CarD activated some promoters and inhibited others (17, 18, 20)**.** However, it has not been shown previously that CarD directly inhibits the activity of a promoter below its unactivated level (i.e., that it doesn't simply interfere with activation), nor that it inhibits transcription in a purified system, where an observed reduction in transcription cannot be ascribed to indirect effects.

CarD levels are constant during exponential growth in *R. sphaeroides* and decrease during stationary phase (21). Even though CarD has been identified as an important transcription factor in various bacterial species, little work has been done to

investigate how CarD expression is regulated. Here we show that the CarD promoter has high basal activity in vitro and does not require activation by CarD, even though it lacks a -7T, and that CarD negatively regulates its own promoter. We then use in vitro analysis to determine how CarD functions as a negative regulator of transcription.

Results

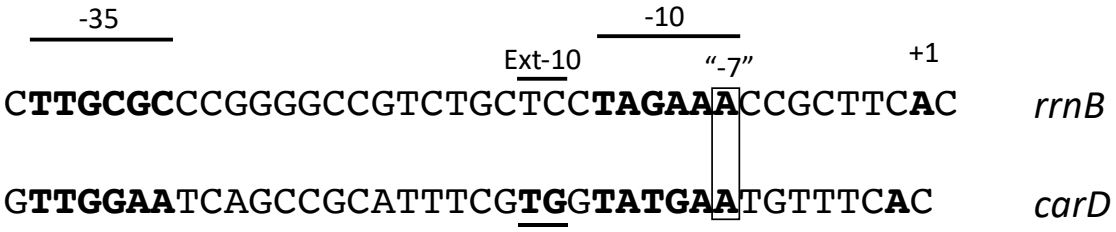
CarD inhibits transcription from its own promoter. Many of the *R. sphaeroides* promoters tested previously, like *rrnB*, were inactive or had very low activity in vitro. A majority of the *R. sphaeroides* promoters tested lacked a thymine at the last position of their putative -10 element (-7T) but $CarD_{Rsp}$ activated their transcription (21). The *carD* promoter lacked -7T (Figure 3.1A), so we tested whether the promoter was active with $RNAP_{Rsp}$ and whether $CarD_{Rsp}$ would further increase its transcription.

We found that the basal activity of the *carD* promoter (the activity in the absence of CarD) was relatively strong compared to that of most other *R. sphaeroides* promoters we tested (21) (Figure 3.1B, lanes 3-4 and 5-6), and $CarD_{Rsp}$ inhibited its own transcription ~3-4-fold at saturation (Figures 3.1B, lanes 7 to 15, and Figure 3.1C). The IC_{50} for inhibition of the *carD* promoter (120 nM) was similar to the EC_{50} for activation of $rrnB_{Rsp}$ (85nM) (21), implying it was working through the same interactions with RNAP. In contrast, there was no effect of CarD on transcription from the *RNA*-1 promoter encoded by the plasmid vector containing the *R. sphaeroides* promoters (Figure 3.1B and Figure 3.1C). The IC_{50} for CarD was within the range of the protein concentrations present in vivo, ~1 μM in exponential phase and <100 nM in stationary phase cells (21).

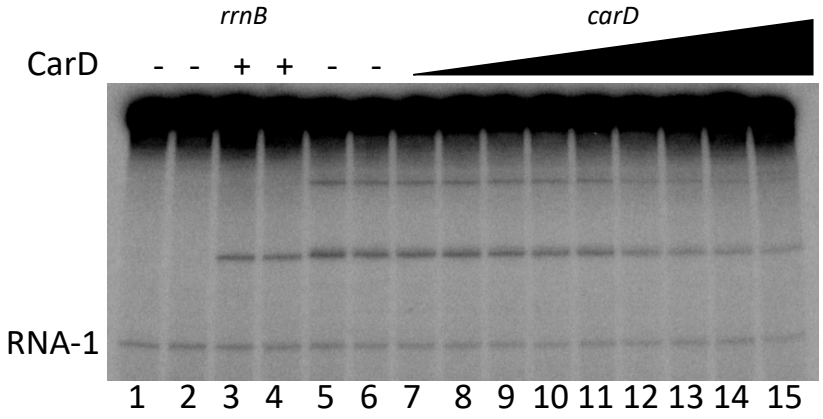
CarD substitutions that impair activation also impair inhibition. Next, we tested whether residues in CarD that are important for its ability to activate transcription from the *rrnB* promoter are also needed for inhibition of the *carD* promoter. Variants with substitutions in the highly conserved tryptophan residue of CarD (W91), a key residue in

Figure 3.1: CarD inhibits transcription from its own promoter. (A) Sequences of the *R. sphaeroides carD* promoter from -36 to +2 and the *R. sphaeroides rrnB* promoter. The -35, -10 elements and transcription start site (+1) are indicated, in addition to the extended -10 sequence within the *carD* promoter. The last position of the -10 element is indicated as the -7 position, even though for the *rrnB* promoter it is 8bp upstream of the transcription start site. (B) in vitro transcription of the $rrnB_{Rsp}$ promoter with $RNAP_{Rsp}$ (20nM) with or without CarD (1280nM) and of the *carD* promoter with and without CarD (wedge indicates CarD concentration range of 10-2560nM) in 170mM NaCl. (C) Relative transcription of the *R. sphaeroides carD* promoter by CarD from experiments as in (B). CarD activates transcription from the *rrnB* promoter ~13-fold while it inhibits transcription from the *carD* promoter ~4-fold (with an IC50 of 120nM). The best fit curve was calculated with 0nM CarD set to a relative transcription of 1. Error bars indicate standard deviation from 5 separate assays. CarD inhibits transcription from the *carD* promoter but not from the *RNA-1* promoter.

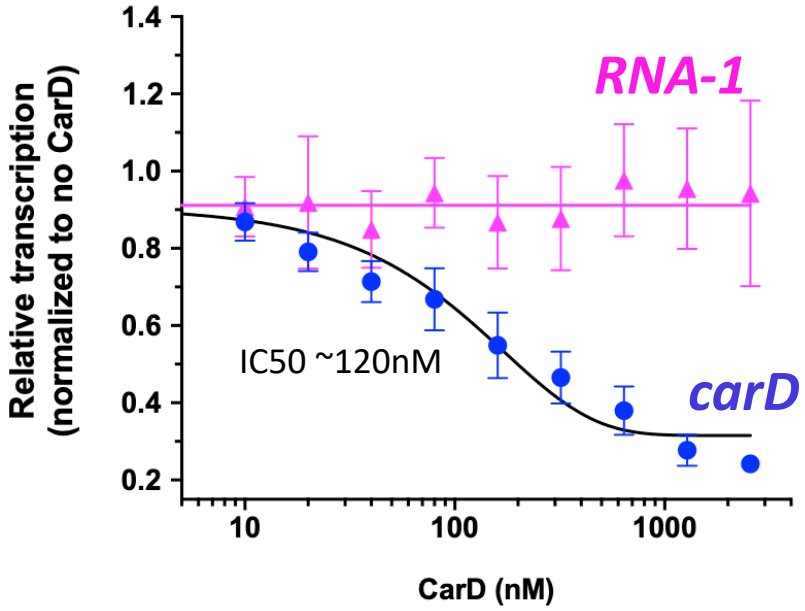
A


-35
-10
Ext-10
"-7"
+1
CTTGCGCCCGGGGCCGTCTGCTCCTAGAAACCGCTTCAC
rrnB
GTTGGAATCAGCCGCATTTCGTGGTATGAATGTTTCAC
carD

B


rrnB
carD
CarD
- - + + - -
RNA-1
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15

C

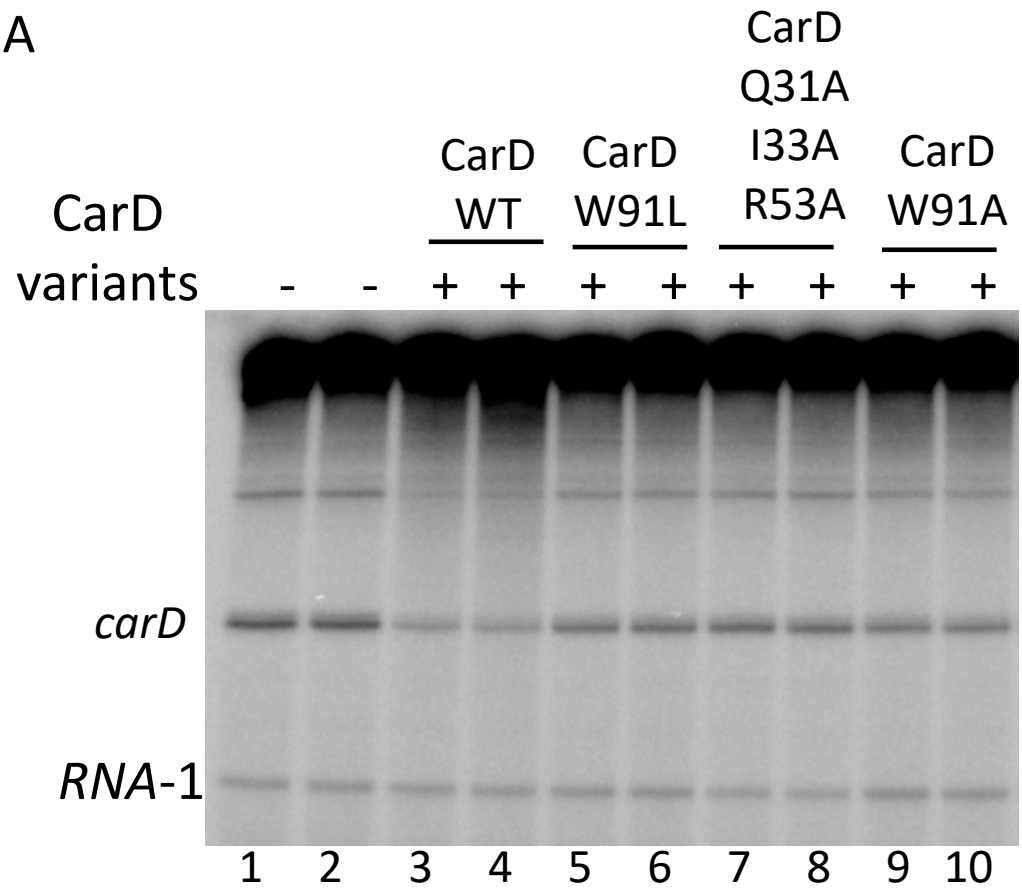

Relative transcription
(normalized to no CarD)
1.4
1.2
1.0
0.8
0.6
0.4
0.2
RNA-1
IC50 ~120nM
carD
10
100
1000
CarD (nM)

the DNA interaction domain that reduced the ability of CarD to activate transcription in *M. tuberculosis* (13), also reduced activation in *R. sphaeroides* (21). These substitution mutants were also partially defective for inhibition of transcription from the *R. sphaeroides carD* promoter in vitro (Figure 3.2A, Lanes 5-6 and 9-10, vs Lanes 3-4; Figure 3.2B, 1.5- and ~1.6-fold inhibition versus ~3.6-fold for CarD-WT). Furthermore, a CarD variant with a triple substitution in the RNAP interaction domain (RID), (Q31A, I33A, R53A) that was partially defective in activation (20) inhibited transcription only ~1.3-fold (Figure 3.2A, Lanes 7 and 8; Figure 3.2B). Thus, the same CarD variants that retained partial function in activation (21) also retained partial function in inhibition.

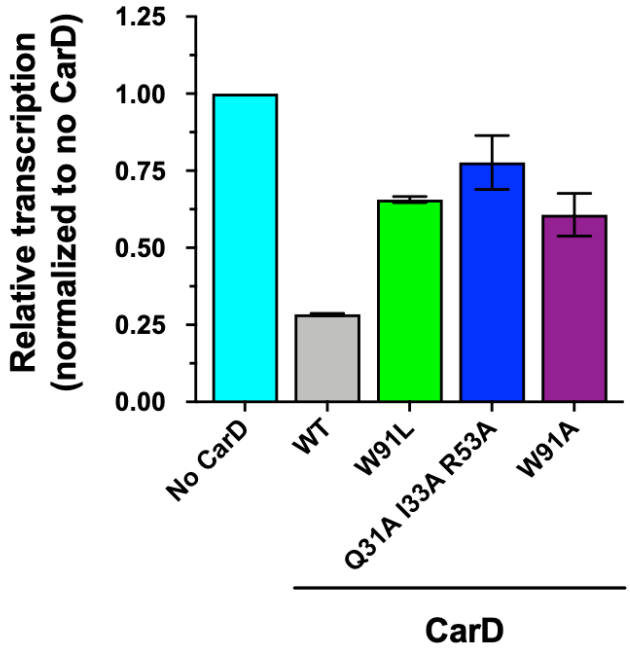
***carD* promoter sequence features important for high activity and inhibition by CarD.** To understand more about the properties of the *carD* promoter that make it subject to inhibition by CarD, the *carD* promoter sequence was compared with that of other *R. sphaeroides* promoters and with *carD* promoters from other α-proteobacteria species (Figure 3.3). Like most *R. sphaeroides* promoters, including the *rrnB* and *rpsM* promoter, the *carD* promoter lacked a thymine at the last position of the -10 element (Figure 3.1A, and Figure 3.3). It shares many promoter features in common with other *R. sphaeroides* promoters, i.e., a TTG sequence at the beginning of the -35 element and a 17 bp spacer between the -10 and -35 elements (Figure 3-3). The *carD* promoter contained an extended -10 sequence which is only found in about 15% of identified *R. sphaeroides* (21, 25). Two *R. sphaeroides* promoters that contained extended the -10 sequence (*RSP*_3110 and *RSP*_1486) were active in vitro and were not activated by CarD. In addition, introduction

Figure 3.2: Inhibition of *carD* promoter transcription by CarD is reduced by substitutions in CarD. (A) in vitro transcription of *carD* promoter with 20nM RNAP with and without 1280nM WT CarD or W91A, W91L or Q31A/I33A/R53A triple mutant CarD. (B) Fold-inhibition of *carD* promoter by WT CarD or mutant CarD (standard deviation from 2 assays).

A


CarD variants
CarD WT
CarD W91L
CarD Q31A I33A R53A
CarD W91A
- - + + + + + + + +
carD
RNA-1
1 2 3 4 5 6 7 8 9 10

B


Relative transcription (normalized to no CarD)
1.25
1.00
0.75
0.50
0.25
0.00
No CarD
WT
W91L
Q31A I33A R53A
W91A
CarD

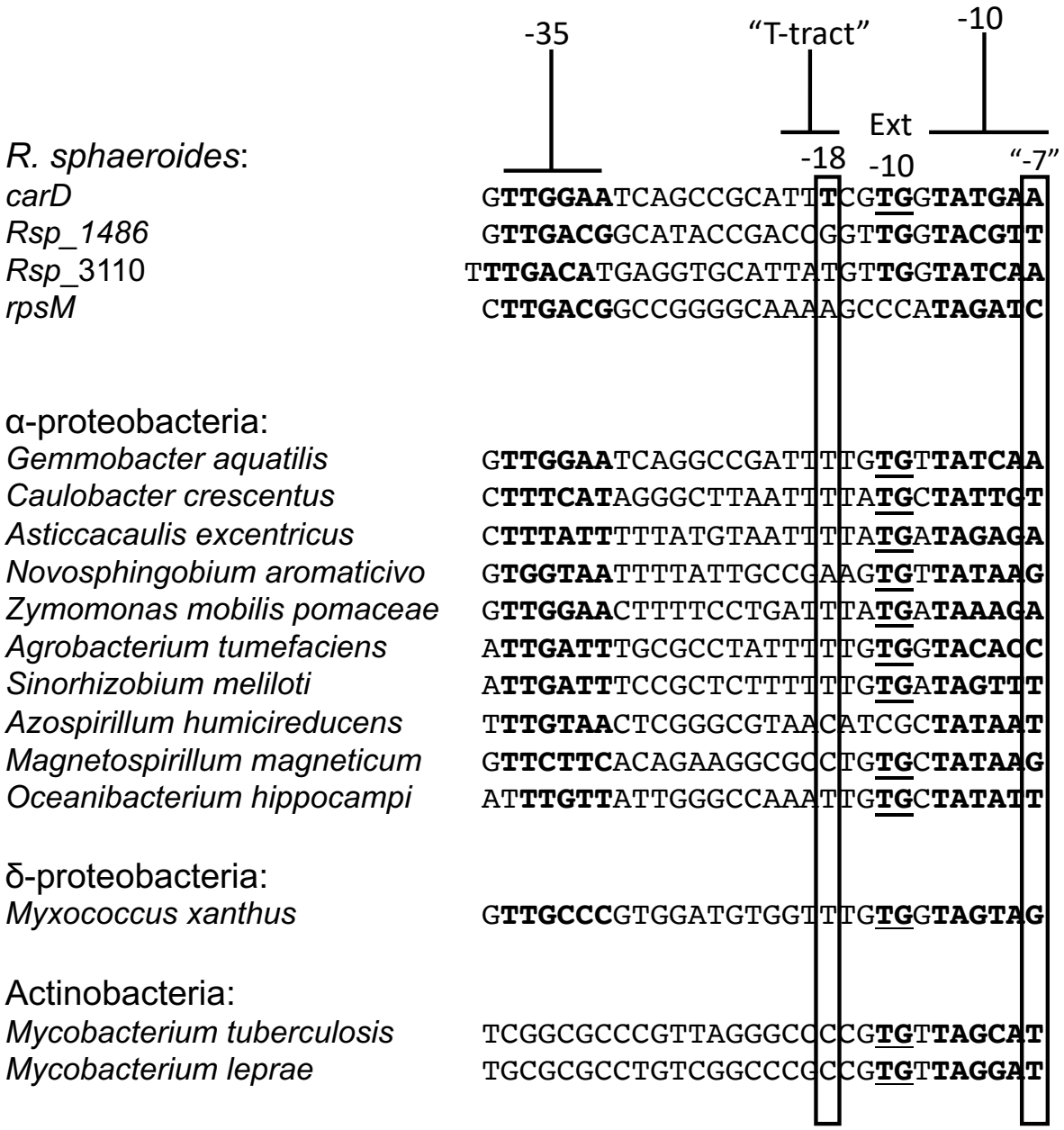
of an extended -10 element into the *rrnB* promoter increased basal transcription and made this promoter insensitive to CarD (21).

The extended -10 sequence is also present in the *carD* promoters from *Caulobacter crescentus* and *Myxococcus xanthus* (14, 23). To determine how widespread the extended -10 sequence is, we also aligned the *carD* promoter sequence from a δ-proteobacterium and two actinobacterial species. In addition, we utilized available transcription start site (TSS) data to identify the *carD* promoter in *Novosphingobium aromaticivo*, *Zymomonas mobilis*, *Sinorhizobium meliloti* and *Mycobacterium tuberculosis* (25–28), and for other bacteria, we identified the *carD* promoter by comparison with the promoter sequence of *R. sphaeroides carD* promoter.

Interestingly, the extended -10 sequence appears to be a widely conserved feature of *carD* promoters from different bacterial species including *M. tuberculosis* (Figure 3.3). There is also some conservation of a 3-nt T-tract upstream of the extended -10 in the spacer region, most notably, there is a T at -18T, which has been shown to be important for transcription for some previously studied promoters (29). Although the *R. sphaeroides carD* promoter lacked -7T, the *carD* promoter in some other species contains -7T, indicating that the identity of the -7 position sequence is not a conserved feature.

To determine whether the promoter features identified from the alignment of *carD* promoters contribute to the high basal activity of the *carD* promoter or to its inhibition by CarD, we constructed *R. sphaeroides carD* promoter variants with substitutions in the extended -10, the -10 element, or in the spacer region upstream of the extended -10, thereby disrupting the 3-nt T-tract (Figure 3.4A). Transcription activities of the promoter

Figure 3.3: *carD* promoter sequences from various proteobacteria. Alignment of the *R. sphaeroides carD* promoter sequence with the *R. sphaeroides RSP*_3110, *RSP*_1486 and *rpsM* promoters and with the *carD* promoter from various α-proteobacteria, one δ-proteobacteria and two actinobacterial species. An extended -10 ("TG") is present in almost all of the *carD* promoters in bacteria from α-proteobacteria, δ-proteobacteria and actinobacteria. The α-proteobacteria are from the order: **Rhodobacterales**: *R. sphaeroides* and *G. aquatilis*; **Caulobacterales**: *C. crescentus* and *A. excentricus*; **Sphingomonodales**: *N. aromaticivo* and *Z. mobilis*; **Rhizobiales**: *A. tumefaciens* and *Sinorhizobium meliloti*; **Rhodospirillales**: *A. humicireducens* and *M. magneticum*; **Sneathiellales**: *O. hippocampi.* Promoters are aligned with the -7 position and promoters with 18 spacers are shifted 1nt to the left.

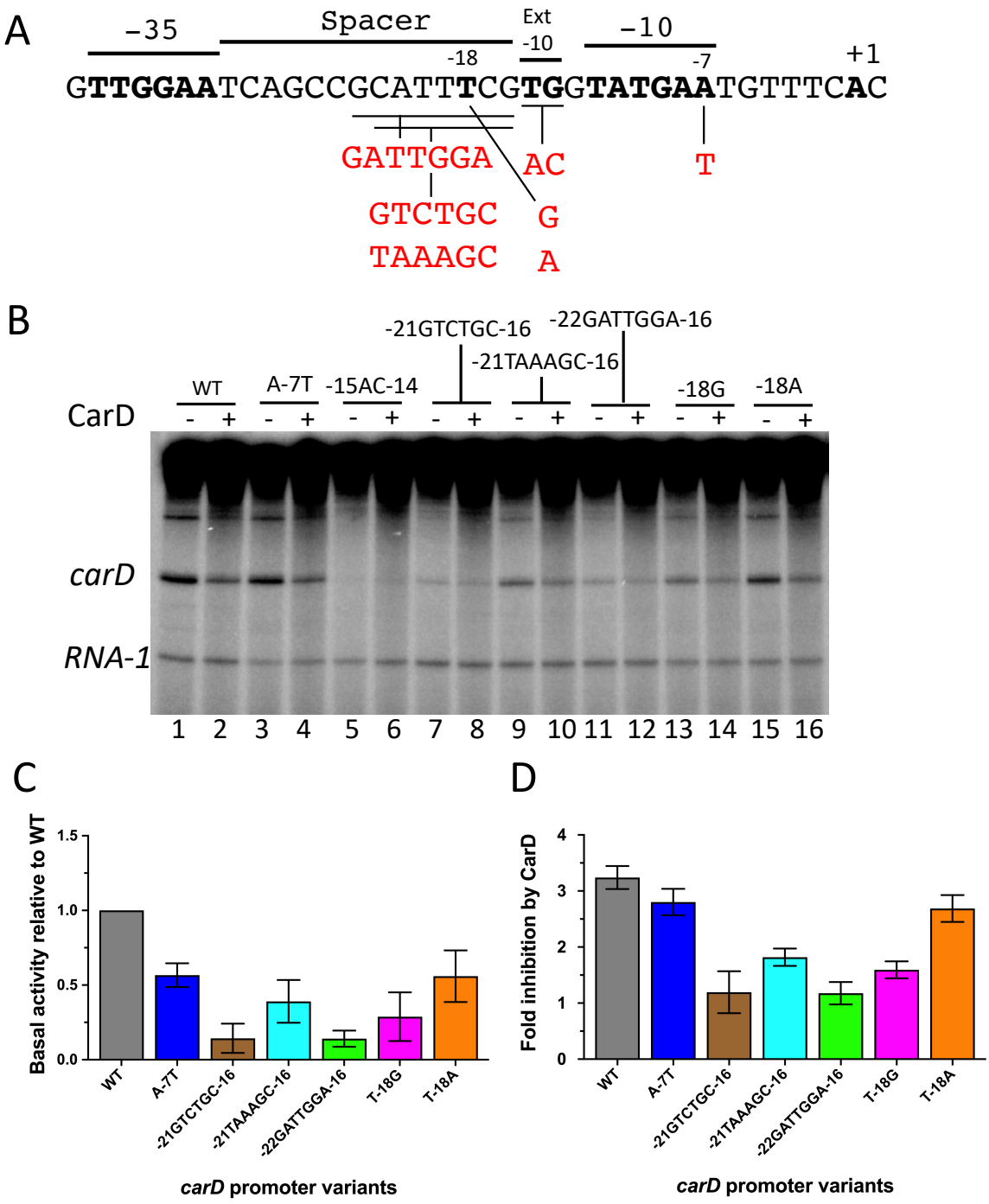

-35
"T-tract"
-10
Ext
-18
-10
"-7"
R. sphaeroides:
carD GTTGGAATCAGCCGCATTTCGTGGTATGAA
Rsp_1486 GTTGACGGCATACCGACCGGTTGGTACGTT
Rsp_3110 TTTGACATGAGGTGCATTATGTTGGTATCAA
rpsM CTTGACGGCCGGGGCAAAAGCCCATAGATC
α-proteobacteria:
Gemmobacter aquatilis GTTGGAATCAGGCCGATTTTGTGTTATCAA
Caulobacter crescentus CTTTCATAGGGCTTAATTTTATGCTATTGT
Asticcacaulis excentricus CTTTATTTTTATGTAATTTTATGATAGAGA
Novosphingobium aromaticivo GTGGTAATTTTATTGCCGAAGTGTTATAAG
Zymomonas mobilis pomaceae GTTGGAACTTTTCCTGATTTATGATAAAGA
Agrobacterium tumefaciens ATTGATTTGCGCCTATTTTTGTGGTACACC
Sinorhizobium meliloti ATTGATTTCCGCTCTTTTTTGTGATAGTTT
Azospirillum humicireducens TTTGTAACTCGGGCGTAACATCGCTATAAT
Magnetospirillum magneticum GTTCTTCACAGAAGGCGCCTGTGCTATAAG
Oceanibacterium hippocampi ATTTGTTATTGGGCCAAATTGTGCTATATT
δ-proteobacteria:
Myxococcus xanthus GTTGCCCGTGGATGTGGTTTGTGGTAGTAG
Actinobacteria:
Mycobacterium tuberculosis TCGGCGCCCGTTAGGGCCCCGTGTTAGCAT
Mycobacterium leprae TGCGCGCCTGTCGGCCCGCCGTGTTAGGAT

variants were compared to the wild-type promoter in the absence (basal) and presence of saturating CarD in vitro (1280 nM; Figure 3.4B).

First, the effects of these substitutions on the basal activity of the promoter will be discussed (Figure 3.4C). Introduction of the consensus -7T into the *carD* promoter did not increase basal transcription, in contrast to what was observed previously for other *R. sphaeroides* promoters (21). Transcription activity from this promoter was similar to that of the WT *carD* promoter (decreased ~1.8 fold relative to that of the WT promoter; Figure 3-4C). Substitutions within the extended -10 sequence virtually abolished the basal activity of the *carD* promoter (Figure 3.4B, lane 5 and Figure 3-4C). These results indicate that the *carD* promoter activity is dependent on the presence of an extended -10 element and that activity is not dependent on the identity of the base at the -7 position.

To determine whether the 3-nt T-tract found just upstream of the extended -10 sequence in the spacer region of the *carD* promoter contributed to basal promoter activity, we constructed three variants: a *carD* spacer sequence substituted with the corresponding sequence from the *rrnB* promoter (-21GTCTGC-16, *rrnB* spacer), an inversion of the *carD* promoter spacer sequence from -21 to -16 (TAAAGC), or a substitution with a sequence previously reported to reduce the activity of the T7A1 promoter (-22GATTGGA-16, in ref. (30)) (Figure 3.4A). The replacement of the T-tract with the *rrnB* spacer resulted in a ~7-fold reduction in basal promoter activity (Figure 3.4B, lane 7; Figure 3.4C), and inversion of the *carD* promoter spacer sequence reduced basal promoter activity ~2.5-fold relative to the WT *carD* promoter (Figure 3.4B, lane 9; Figure 3.4C). Changing the *carD* promoter spacer regions to -22TATTGGA-16

Figure 3.4: Effects of individual *carD* promoter mutations on *carD* promoter activity and inhibition by CarD. (A) Key features of the carD promoter (-35 element, spacer, position -18, extended -10 element, -10 element and start site (+1)) are indicated above the promoter sequence. Mutations introduced into the *carD* promoter are shown in red. (B) in vitro transcription of each promoter mutant with RNAP (20nM) either without CarD (even numbered lanes) or with 1280nM CarD (even numbered lanes) in buffer with 170mM NaCl. (C) Basal activity of each promoter mutant (without CarD) relative to the *carD* WT promoter (standard deviation from 3-5 assays). The *carD* -15AC-14 mutant was not included in graph as the transcriptional signal was undetectable (D) Fold inhibition by CarD (ratio of activity with vs without CarD) for each *carD* mutant promoter (standard deviation from 3-5 assays). Basal activity of each promoter without CarD defined as 1.


A
-35
Spacer
Ext
-10
-10
-18
-7
+1
GTTGGAATCAGCCGCATTTCGTGGTATGAATGTTTCAC
GATTGGA
AC
T
GTCTGC
G
TAAAGC
A
B
-21GTCTGC-16
-22GATTGGA-16
-21TAAAGC-16
WT
A-7T
-15AC-14
-18G
-18A
CarD
- + - + - + - + - + - + - + - +
carD
RNA-1
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
C
Basal activity relative to WT
1.5
1.0
0.5
0.0
WT
A-7T
-21GTCTGC-16
-21TAAAGC-16
-22GATTGGA-16
T-18G
T-18A
carD promoter variants
D
Fold inhibition by CarD
4
3
2
1
0
WT
A-7T
-21GTCTGC-16
-21TAAAGC-16
-22GATTGGA-16
T-18G
T-18A
carD promoter variants

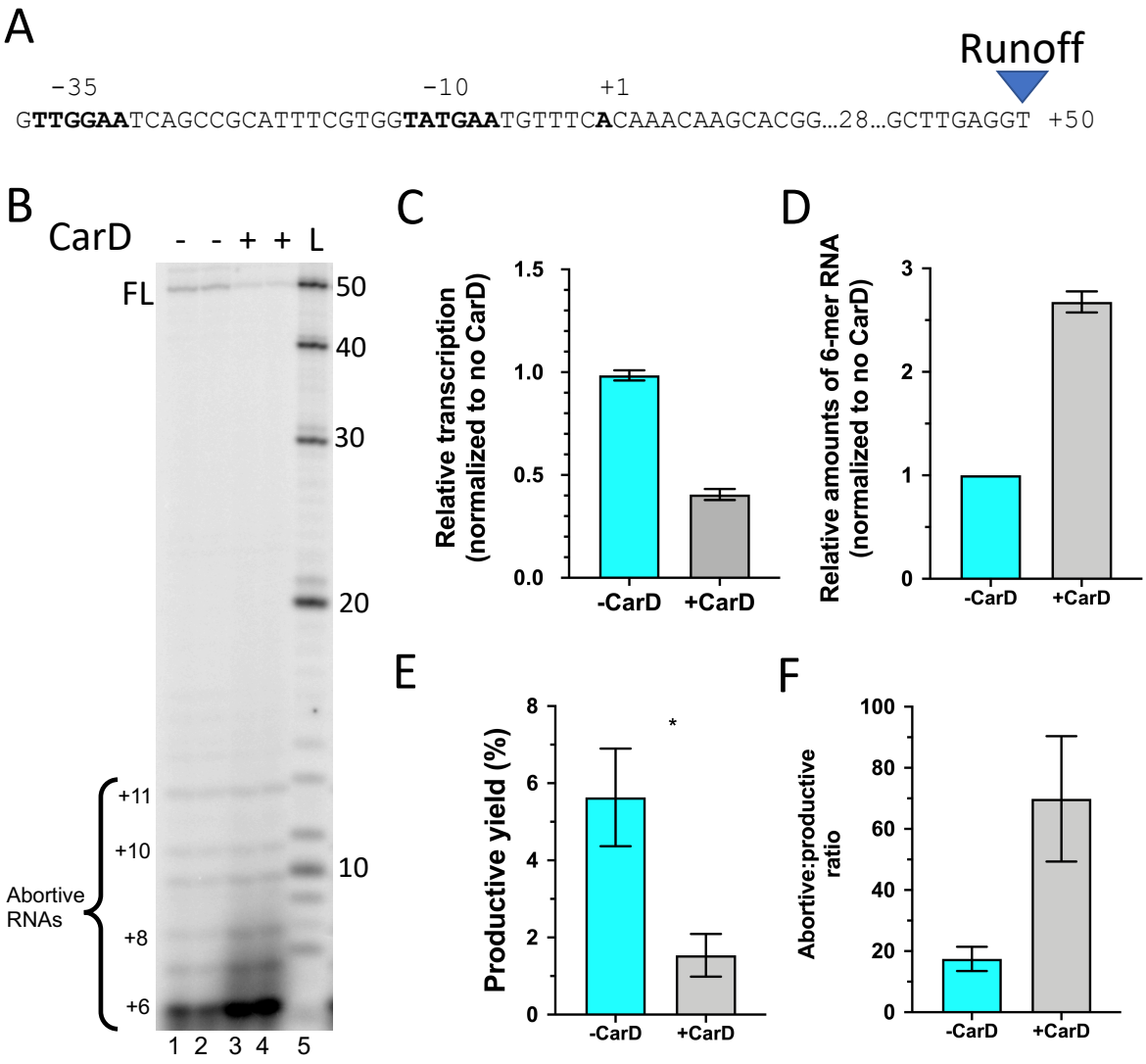
reduced the transcription from this promoter ~7-fold relative to the WT promoter, like replacement with the *rrnB* sequence (Figure 3.4B, lane 11; Figure 3.4C). Two *carD* promoter variants, T-18G or T-18A, had lower activity, 3.6- and 1.8-fold respectively relative to the WT (Figure 3.4B, lanes 13 and 14; Figure 3.4C), consistent with the previously reported effect of substitutions at the -18 position in the *E. coli cbpA* promoter (29). Taken together, these results indicate that the spacer sequence just upstream of the extended -10 in the *carD* promoter is an important contributor to the activity of the promoter.

We also tested whether substitutions in the spacer of the *carD* promoter affected inhibition of transcription by CarD (Figure 3.4B and 3.4D; compare activities of each promoter with vs. without CarD). CarD inhibited transcription from the A-7T substituted promoter and the T-18A promoter to about the same extent as the WT promoter (2.5-3-fold). However, inhibition of four of the spacer substitution variants was either reduced (for the TAAAGC or the T-18G substitutions) or nearly abolished (for the *rrnB* spacer substitution or the *carD* spacer inversion substitution) (Figure 3.4D), suggesting that some aspect of this part of the promoter sequence influences whether the promoter is inhibited by CarD. We could not detect transcription from the extended -10 variant (-15AC-14) in the absence (or presence) of CarD, so we could not determine whether this sequence is also important for regulation of the *carD* promoter by CarD. Nevertheless, these results indicate that the extended -10 element is crucial for both *carD* promoter activity and regulation.

CarD$_{Rsp}$ inhibition of the *carD* promoter leads to an accumulation of abortive RNAs. *M. tuberculosis* CarD was proposed to inhibit transcription from some promoters by slowing the rate of promoter escape by RNAP (17, 31). If the mechanism of inhibition by CarD involved slowing the rate of promoter escape, there might be an observable accumulation of abortive RNA products as RNAP attempted to cycle to productive transcription (as shown previously on promoters with *E. coli* RNAP) (32, 33). To test whether *R. sphaeroides* CarD inhibited transcription from its own promoter by a similar mechanism, we designed an *in vitro* transcription assay to monitor abortive transcription and full-length transcript (50nt) from a short DNA fragment (~122 nt) containing the *carD* promoter (Figure 3.5A). This approach allowed us to visualize and quantify both abortive and productive transcripts in the presence and absence of CarD.

In addition to the full-length transcript, RNAP produced abortive RNAs that ranged from 6-mers to 11-mers from the *carD* promoter (Figure 3.5B, Lanes 1 and 2). The sizes of the abortive products were determined by performing transcription using a subset of nucleotides (see *Materials and Methods*; Figure 3.S1B). CarD inhibited full-length transcript formation from the promoter-containing fragment about ~2.5-fold (Figure 3.5B, Lanes 3, 4 vs. 1,2 and Figure 3.5C), relatively similar to the observed inhibition from a supercoiled template (Figure 3.1B). Correlating with the inhibition of productive RNA transcript, CarD caused an increase in abortive RNAs 6-11 nt in length (Figure 3.5B), a ~2.5-fold increase in the level of the 6-mer RNA (Figure 3.5D), and a ~3-4 fold decrease in the productive yield (full-length transcript as a percentage of total transcripts) (Figure 3.5E), The abortive-to-productive RNA ratio was much higher in the

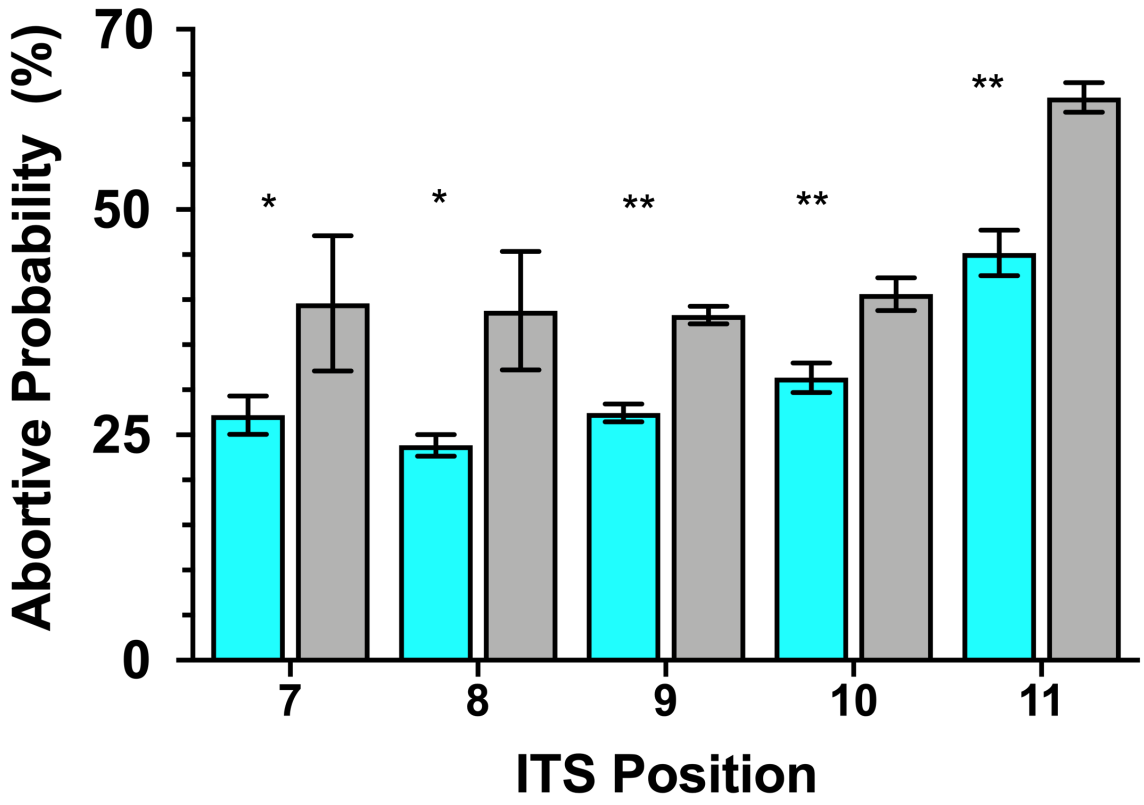
Figure 3.5: Analysis of abortive transcripts from the *carD* promoter. (A) *carD* promoter sequence from -36 to +50 from the 122 nt linear fragment used in the abortive transcription analysis (entire fragment contains *carD* promoter sequence from -72 to +50). The -35 and -10 elements are shown in bold. The wedge indicates the downstream end of the linear template. The runoff (full length, FL) transcript is 50nt. (B) Gel profile of the abortive and productive transcripts produced from the *carD* promoter fragment with RNAP (120nM) and with or without CarD (4800nM) in transcription buffer with 50mM NaCl. (C) CarD inhibits full length transcription 2.5-fold. (D) Accumulation of the levels of the 6-mer abortive RNA from the *carD* promoter in the presence of CarD (gray) relative to that in the absence of CarD. (E) Comparison of the productive yield (FL transcript) with (cyan) and without CarD (gray) as a percentage of total products (abortive plus FL). The yield percentages were plotted as average values and standard deviation was determined from 5 assays. (* = p-value <0.005, determined from t-test: Two-Sample Assuming Unequal Variances). (F) Abortive yield to productive yield ratio with (cyan) and without (gray) CarD. On the *carD* promoter, RNAP aborts transcription 17.5 times per productive cycle without CarD compared to 69.8 times in the presence of CarD.


A
Runoff
-35
-10
+1
GTTGGAATCAGCCGCATTTCGTGGTATGAATGTTTCACAAACAAGCACGG…28…GCTTGAGGT +50
B
CarD - - + + L
FL
50
40
30
20
10
+11
+10
+8
+6
Abortive RNAs
1 2 3 4 5
C
Relative transcription (normalized to no CarD)
1.5
1.0
0.5
0.0
-CarD
+CarD
D
Relative amounts of 6-mer RNA (normalized to no CarD)
3
2
1
0
-CarD
+CarD
E
Productive yield (%)
8
6
4
2
0
*
-CarD
+CarD
F
Abortive:productive ratio
100
80
60
40
20
0
-CarD
+CarD

presence of CarD (Figure 3.5F), i.e. RNAP aborted 17.8 times per productive cycle in the absence of CarD and 64.9 times in the presence of CarD.

On some promoters, RNAP will transcribe abortive RNAs that are identical in sequence to the initaly transcribed sequence (ITS) of the promoter (positions +1 to +20) (33) on its path to promoter escape. RNAP produced abortive RNAs correlating with ITS positions +7 to +11 on the *carD* promoter. To determine whether CarD presents a barrier to productive transcription at each ITS position, we expressed the ratio of abortive transcripts as probabilities. CarD increased the probability of RNAP aborting at each ITS position (7-mer to 11-mer, Figure 3.6), suggesting that CarD presents a barrier to promoter clearance at each ITS position.

Figure 3.6: Abortive probabilities of each initially transcribed sequence (ITS) from the *carD* promoter. Abortive probability was calculated as described previously (Hsu et al. 1996) (also see *Materials and Methods*) and analyzed using two tailed t-test (assuming unequal variances). P-value: * <0.01, ** <0.0001. CarD increases the abortive probability at each ITS position when RNAP is transcribing the *carD* promoter. For ease of comparison, the abortive probabilities are given for each position without (cyan) and with (gray) CarD. The abortive probabilities are plotted as average values with standard deviations from 5 assays.


Abortive Probability (%)
70
50
25
0
*
*
**
**
**
7
8
9
10
11
ITS Position

DISCUSSION:

CarD is a transcription factor that is capable of activation and inhibition

We previously showed that *R. sphaeroides* CarD activates several *R. sphaeroides* promoters, including *rrnB*, correlating (albeit not perfectly, see reference (21)) with the absence of -10 element consensus position -7T in these promoters. CarD protein levels varied with growth phase; they were constant during exponential growth and declined over 10-fold as cells entered stationary phase, suggesting a correlation between transcription regulation of these promoters and the level of CarD (21). The mechanism of regulation of CarD protein levels was not known. Therefore, we characterized the *R. sphaeroides carD* promoter to determine whether CarD affected transcription of its own promoter. We report here, in contrast to the activation of many promoters by CarD reported previously, CarD inhibits transcription from its own promoter ~4-fold (Figure 3-1B), with an IC_{50} similar to that for activation, using the same critical CarD residue (W91) for inhibition and activation. This represents the first report of direct inhibition of a promoter by CarD in vitro, although RNA-seq data from a *M. tuberculosis* strains with functional variants of its essential CarD protein implicated CarD in negative regulation of some *M. tuberculosis* promoters (17).

Mechanism of inhibition by CarD

We have shown here that CarD not only activates transcription from some promoters, but it also inhibits transcription of its own promoter. It was proposed previously that CarD might differentially affect promoters, depending upon their specific promoter kinetics (19, 31), by stabilizing the open complex, leading either to activation

of promoters that are rate limited by complex instability, or inhibition of promoters that form stable complexes by further stabilizing them and preventing promoter escape. Here we show that this hypothesis is consistent with our data for the *R. sphaeroides rrnB* and *carD* promoters. This mechanism differs from that of DksA which also activates or inhibits transcription (in conjunction with ppGpp), depending upon the kinetic properties of the promoter (8, 9, 12, 34, 35). DksA/ppGpp activates many promoters, including those for amino acid biosynthesis, and it negatively autoregulates its own promoter (12) as well as many other promoters including rRNA promoters (35), but it appears to do so by inhibiting a step prior to open complex formation (36). CarD's mechanism of negative regulation also differs from that of most repressive transcription factors, which compete with RNAP for overlapping binding site on promoters (37).

We found that promoters activated by CarD display very low basal activity that correlated with the lack of the highly conserved -7T position in the -10 hexamer (21), a position associated with formation and stabilization of open complexes (38). Although we proposed that CarD stabilized the open complex at the *R. sphaeroides rrnB* promoter, thereby activating its transcription (21), we found that CarD did not stabilize the RNAP-promoter complex against challenge by the competitor heparin (Figure S3.2). This appeared to be a function of the *rrnB* promoter sequence, since the RNAP was capable of forming stable open complexes on other promoters (e.g., RNA-1; Figure 3S.2, (39)). We conclude that it is likely that stabilization by CarD does occur at some level at this and other activated promoters, resulting in increased transcription, but that the stabilization is of a different intermediate complex that is not competitor stable, or is on a time scale not detected in our experiments.

In contrast, the promoter for CarD displayed properties different from the activated promoters. Its basal activity was high, and while it does not contain the thymine at -10 element position -7, it contains another element associated with formation of stable complexes, an extended -10 sequence (40). Mutational analysis indicated that the extended -10 sequence was essential for the strong activity of the promoter in the absence of CarD, and that the sequences just upstream of the extended -10 element also affected the basal activity and the ability of CarD to inhibit the promoter.

Work on other promoters has indicated an important functional role for sequence features in the spacer region just upstream of the extended -10 element. In one study, this spacer region sequence was termed the “Z element” (30). Although the Z element was proposed to be a base-specific recognition element, later structural studies with other RNAP-promoter complexes suggested that the DNA backbone in the Z element region interacts with the RNAP β’ zipper sequence (41). Interestingly, substitutions in the *carD* promoter spacer made these variants insensitive to CarD inhibition (Figure 3.4B and 3.4D). Another study implicated the base at position -18 in promoter function. This DNA backbone at this position interacts with the linker between sigma regions 3 and 2, and mutational analysis showed that a G at this position was correlated with reduced promoter activity (29). We show here that -18G in the *carD* promoter reduced the extent of negative regulation by CarD.

Consistent with the proposal that inhibition by CarD results from stabilization of the open complex, we found that the presence of CarD resulted in higher levels of abortive transcription products from the *carD* promoter. The ratio of abortive to

productive transcripts from the *carD* promoter was increased in the presence of CarD, a property characteristic of promoters that are limited for escape from the promoter. Consistent with this proposal, introduction of a -7T in the *carD* promoter, which could further stabilize the complex, reduced the basal activity from the promoter compared to the wild-type promoter, but the extent of its inhibition was not different from that of the wild type promoter. The presence of a -7T or an extended -10 element may not be the only determinant of a stable open complex and identification of other negatively regulated promoters (if they exist) would require a systematic genome-wide approach.

Potential role of autoregulation in determining levels of CarD during log phase growth

We present here evidence of CarD autoregulation of its own promoter by inhibiting transcription. Autoregulation could maintain homeostasis during log phase growth. The molar concentration of CarD during aerobic growth is in the micromolar range, well above the IC_{50} needed for inhibition (and the EC_{50} needed for activation), a concentration where most or all RNAP-promoter complexes would be occupied (21). DksA also feedback regulates its own expression (in conjunction with ppGpp) in order to maintain a constant level of the transcription factor (12), but it inhibits transcription by a different mechanism.

It is unclear how autoregulation would impact the decrease in CarD protein levels observed when *R. sphaeroides* cells enter stationary phase. The lower CarD levels in stationary would be too low to achieve occupancy of the RNAP-promoter complex and thus there must be another mechanism for reducing CarD protein levels. Previous work

found that the *C. cresentus* CarD homolog is a substrate of ClpX (23). Both *C. cresentus* and *R. sphaeroides* CarD have a C-terminal AA motif (Figure 3.S3), which is important for ClpXP proteolysis. Thus, the decrease in CarD levels in stationary phase could potentially be due to protein turnover mediated by ClpXP.

Precise control of CarD protein levels could be important for regulation of many genes, including many important for ribosome biogenesis (13, 21). We found that the *carD* promoter is unusually active, and this could account for the high intercellular levels in *R. sphaeroides* cells during log growth. CarD regulation of its own expression would prevent cells from producing too little or too much CarD than what is needed for ribosome expression. When cells enter stationary phase, however, CarD protein levels decrease which would lead to global downregulation of transcription of genes involved in ribosome biogenesis. The gradual decrease in CarD protein levels (~8 hours) would be too slow to account for the rapid decrease in transcription that is characteristic of a stringent response (e.g. *R. sphaeroides* rRNA synthesis decreases in ~20 minutes during light downshifts) (42, 43). In chapter 4, I will test whether the *R. sphaeroides* DksA homolog (38) synergizes with ppGpp to inhibit transcription from CarD activated promoters.

SUPPLEMENTAL FIGURES

Figure 3.S1: Identification of lengths of abortive transcripts from the *carD* promoter. (A) The *carD* promoter sequence with the key promoter features (-35 and -10 elements) and the initiating nucleotide and +8 position. (B) To identify the abortive RNAs produced from the *carD* promoter, in vitro transcription was assayed with an NTP mix containing all four ribonucleotide (lanes 1 to 3) or a mix containing only ATP and CTP (lanes 4 and 5), which should produce a prominent +8 RNA (see *Materials and Methods* for further detail). In vitro transcription on the *carD* promoter fragment with RNAP (120nM) without CarD (lanes 1 and 4) or with CarD (4800nM; lanes 2,3 and 5) in transcription buffer (50mM NaCl).

A

-35
-10
+1
+8
GTTGGAATCAGCCGCATTTCGTGGTATGAATGTTTCACAAACAAG

B

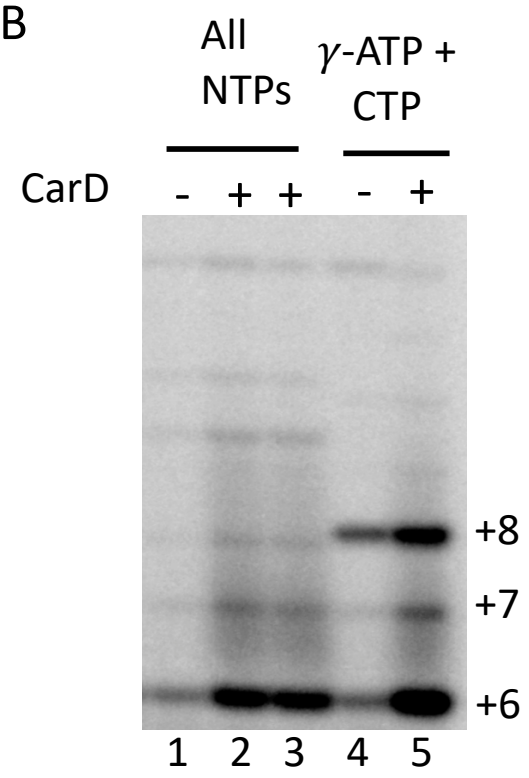

All NTPs
γ-ATP + CTP
CarD
- + + - +
+8
+7
+6
1 2 3 4 5

Figure 3.S2: CarD does not stabilize the *rrnB* promoter against the competitor haparin. in vitro transcription of the $rrnB_{Rsp}$ promoter with $RNAP_{Rsp}$ (20nM) with or without CarD (1280nM) in 170mM NaCl (lanes 1 to 4). To measure complex stability, RNAP-promoter complexes were preformed in the absence of ribonucleoside triphosphates (rNTPs) and presence of CarD (1280nM) at 30℃. After the addition of heparin (50µg/mL final), aliquots were sampled at the indicated times. Samples were added to pre-aliquoted rNTPs, incubated for 10 mins and terminated with stop solution. For time 0, aliquots of the preformed complex were added to a pre-aliquoted mix of rNTPs and heparin, incubated for 10 mins and terminated with stop solution. Transcripts were resolved on 6% acrylamide–7M urea gels. CarD did not stabilize the RNAP-*rrnB* complex against competition with heparin. RNAP formed a heparin stable complex with the plasmid encoded *RNA*-1 promoter.

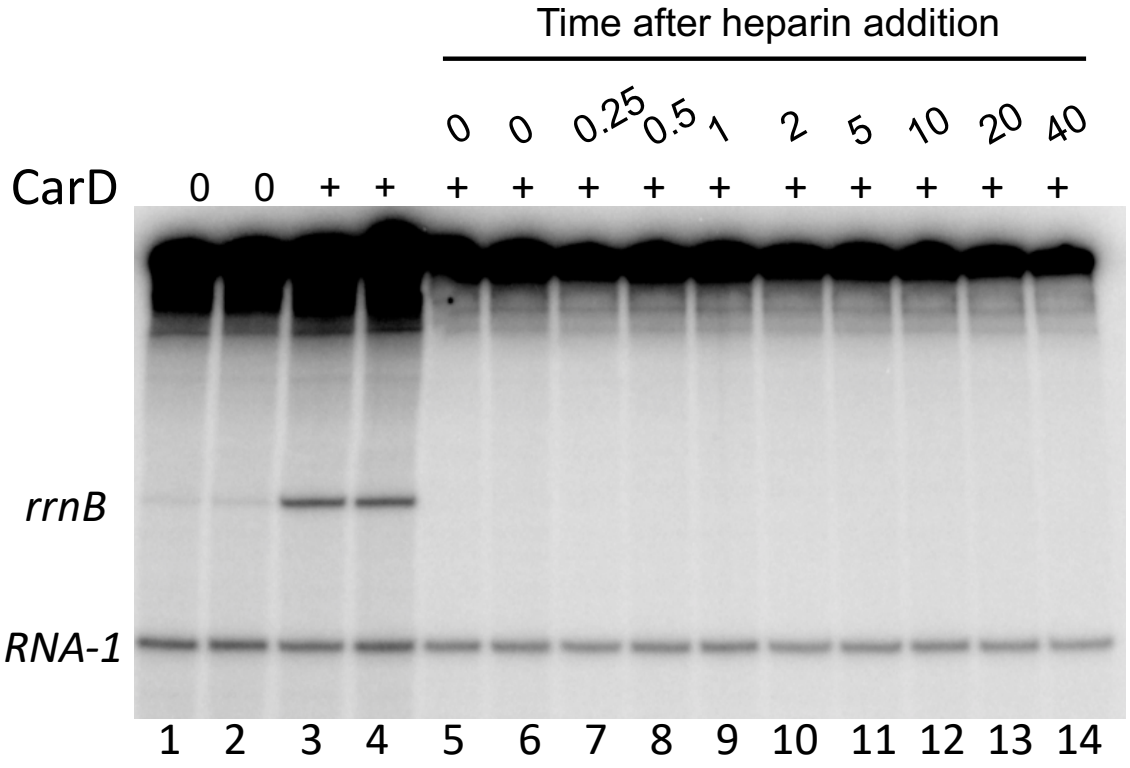

Time after heparin addition
0 0 0.25 0.5 1 2 5 10 20 40
CarD
0 0 + + + + + + + + + + + +
rrnB
RNA-1
1 2 3 4 5 6 7 8 9 10 11 12 13 14

Figure 3.S3: Clustal Omega alignment of CarD from *Rhodobacter sphaeroides* 2.4.1 (R) and *Caulobacter crescentus* NA1000 (C). The C-terminal AA motif (boxed) was previously shown to be important for ClpXP dependent proteolysis of *C. crescentus* CarD (CdnL).

```
R     MTKTKKPEFRPNEFVVYPAHGVGRIISIEEQEIAGIRLELFVISFEKDKMTLRVPTHKAT   60
C     MSKTGL-EFSVGDHVVYPAHGVGSIQAIETQEVAGMSLEVYVITFDHEKMTLRVPTKKAK   59
      *.** . **  .:.********* * :** **:**: **::**:*:::********:**.

R     EVGMRSLSTPDVVTKALDTLKGKARVKRAMWSRRAQEYEQKINSGDLMSIAEVVRDLHRT   120
C     TAGLRPLAEGGTVSQALTTLKGRARVKRTMWSRRAQEYEAKINSGDLISIAEVVRDLHRA   119
       .*:* *:  ..*::** ****:*****:********** *******:***********:

R   DDQREQSYSERQLYEAALERLTREVAAVSGVDEAGAQKAVDAVLVSRAA    169
C   ENQPEQSYSERQLYESALDRMAREVAAIERIDREAAIGILTKSLV-KAA    167
    ::* ***********:**:*::*****:. :*. .*   :   ** .**
```

MATERIALS AND METHODS

Bacterial strains and growth conditions. Bacterial strains are listed in Table 3.1**.** *E. coli* strains were grown in either LB Lennox media or LB agar plates at 37°C supplemented with kanamycin (30 µg/mL) or chloramphenicol (20 µg/mL) when required.

Purification of *R. sphaeroides* RNAP Holoenzyme and SUMO-tagged CarD wild type and variant proteins. *R. sphaeroides* holoenzyme was purified using Ni-agarose affinity chromatography from aerobically grown cells containing the chromosomally encoded His10-β' (RLG14765) as described (21).

The $carD_{Rsp}$ gene was cloned into the pETSUMO vector for overexpression and purification from *E. coli* BL21 using its IPTG-inducible T7 promoter as described (21). Untagged purified CarD was stored in storage buffer (20 mM Tris-HCl [pH 8.0], 100 mM NaCl, 2 mM DTT, 50% glycerol). Protein concentration was assayed by the Bradford assay using BSA as a standard. For in vitro transcription assays, CarD dilutions were prepared in CarD storage buffer.

Construction of in vitro transcription templates. *R. sphaeroides* promoter fragments were PCR amplified from *R. sphaeroides* 2.4.1 chromosomal DNA and inserted into pRLG770 (6) using the oligonucleotides listed in Table 3.2. Each oligonucleotide set created promoter fragments with the described endpoints relative to the predicted transcription start site based on the presence of predicted -10 and -35 elements. The resulting promoter fragments contained sequences for ligation to linearized pRLG770 (amplified using oligonucleotides RLG7057 and RLG7056) by Hi-Fi Assembly (NEB). Plasmids were electroporated into *E. coli* XL1-blue, and those confirmed to have the

correct insert size on gels were sequenced with oligonucleotides RLG4252 and RLG4253. Promoter variants were created by Multi-site Lightning QuikChange (Agilent) as described above using the mutagenic primers listed in Table 3.2. Plasmid DNA was purified from *E. coli* XL1-blue by miniprep (Qiagen), phenol extracted, diluted to 200 ng/µL, and stored at -20°C.

in vitro transcription. Multiple round transcription assays were carried out using supercoiled plasmid templates containing the promoters listed in Table 3.3 using *R. sphaeroides* RNAP (20 nM) in 25µL reactions. Template DNA (5 pmol or 400 ng) was incubated in transcription buffer (20 mM Tris-HCl [pH 7.9], 10 mM MgCl2, 1 mM DTT, 0.1 mg/ml BSA, and either 170 or 100 mM NaCl, as indicated, containing NTPs (200 µM ATP, GTP, CTP, 10µM UTP, and 1.2 µCi α-^{32}P UTP [Perkin Elmer]) with CarD or CarD storage buffer as indicated at room temperature for 10 min.

Transcription was initiated by addition of RNAP and reactions were incubated at 30°C for 20 min. The reactions were terminated by addition of 25 µL water-saturated phenol, aqueous and phenol phases were separated by brief centrifugation, and 16 µL of the aqueous phase containing the transcript was transferred to 16µL 2X loading solution (8M urea, 10 mM EDTA, 2X TBE, 1% SDS and 0.02% bromophenol blue). Transcripts were denatured by heating the samples at 95°C for 1 min, resolved on 6% polyacrylamide 7M urea gels, imaged using a Typhoon phosphorimager, and quantified with ImageQuant 5.2 software.

Abortive transcription. A *carD* promoter fragment for use as a linear transcription template was PCR amplified from pRLG15221 using oligonucleotides RLG8660 and RLG8661 (Table 3.4) resulting in a 122 nucleotide DNA product (Table 3.5). The

fragment was inspected on a 2% agarose gel and then purified using a Qiagen PCR clean up kit and resuspended in water. Abortive transcription reactions were carried out as previously described (44) using 100nM PCR amplified template and 120nM $RNAP_{Rsp}$ in 10μL reactions. Template DNA was incubated in transcription buffer (50mM NaCl, 10mM Tris-HCl [pH8], 10mM $MgCl_2$, 0.1mM DTT) containing NTPs (500 μM GTP, CTP, UTP, 25μM ATP, and 8 μCi γ-^{32}P ATP [Perkin Elmer]) with CarD (4800nM) or CarD storage buffer as indicated. Reactions were initiated with the addition of $RNAP_{Rsp}$ and incubated at 30°C for 30 min. The reactions were terminated with the addition of 10μL stop solution (95% formamide, 10mM EDTA, 0.05% xylene cyanol and 0.05% bromophenol blue). Transcripts were denatured by heating at 90°C for 30 seconds and resolved on a denaturing gel (20% bis-acrylamide [19:1 acrylamide:bis-acrylamide], 7M urea and 0.5XTBE) on a electrophoresis apparatus with a anode buffer containing 0.5XTBE and a cathode buffer containing 0.5XTBE and 0.75M sodium acetate. The gel was wrapped in saran wrap and exposed to a phosphoimager plate overnight. Transcripts and abortive products were imaged on a Typhoon phosphorimager and quantified with ImageQuant 5.2 software. The abortive yield was calculated by dividing the total signal for the abortive RNAs by the total signal for each RNA species in the lane. The productive yield is calculated by dividing the signal for the full-length transcription by the total signal for each RNA species in the lane. The abortive-to-productive ratio was calculated by dividing the abortive yield by the productive yield. Equations used to calculate abortive probability were derived from (44):

$$P_i = \frac{X_i}{100\% - \sum_2^{i-1} X_i}$$

where X_i is the abortive yield, in percent, of the *i*th abortive RNA. The denominator gives the fraction of RNA polymerase-promoter complexes that survives to the *i*th position. The abortive probability for the *i*th position, Pi, in percent, is the probability that the transcript will be aborted from the initial transcription complex at the *i*th position. As the 6-mer was the first RNA in the series, the abortive probability for this position was not calculated.

The size marker ladder was generated from radiolabeling RNA from the Decade^{TM} Markers System (Invitrogen) using the manufacturer's protocol with the exception: the ladder was not heated before loading onto gel.

Determination of abortive RNA length. To determine the length of each abortive RNA from the carD promoter, transcription was performed as above with an NTP mix containing (500 CTP and 25μM ATP, and 8 μCi γ-^{32}P ATP). The *carD* promoter initially transcribed region contains only A and C for the first 8 nucleotide (Figure 3S1B). The NTP mix with only ATP and CTP should produce a prominent RNA for the 8-mer RNA as RNAP will not be able to elongate onto the 9th position as it is a G.

Table 3.1: Strain and Plasmid used in study

Strains used in this study

Strains	Genotype	Source
BL21 (DE3) pLysS	*E. coli* str. B F^- *ompT gal dcm lon hsdS*$_B$($r_B^-m_B^-$) λ(DE3 [*lacI lacUV5-T7p07 ind1 sam7 nin5*]) [*malB*$^+$]$_{K\text{-}12}$(λ^S) pLysS[*T7p20 ori*$_{p15A}$](Cm^R)	Novagen
E. coli XL1-Blue	Δ(*mcrA*)*183* Δ(*mcrCB-hsdSMR-mrr*)*173 endA1 supE44 thi-1 recA1* gyrA96 relA1 lac *[F'* proAB lacI^q*Z*Δ*M15 Tn10 (Tet^r)]*	Stratagene
E. coli XL10-Gold	*endA1 glnV44 recA1 thi-1 gyrA96 relA1 lac Hte Δ(mcrA)183 Δ(mcrCB-hsdSMR-mrr)173 tet^R F'[proAB lacI^{q}ZΔM15 Tn10(Tet^R Amy Cm^R)]*	Stratagene
RLG14765	RLG14946 with a C-terminal thrombin-His_{10} chromosomally tagged β' (*rpoC*)	(21)

Plasmids used in this study

Plasmids	Description	Source
pRLG770	Transcription vector	(6)
pETSUMO	pET derived plasmid that encodes a Ulp1 cleavable His_{10}SUMO tag. ORFs are cloned in frame with the coding sequence of the SUMO tag.	Invitrogen
p15053	CarD W91A substitution	(21)
p15434	CarD W91L substitution	(21)
p15435	CarD Q31A, I33A and R53A substitution	(21)

Table 3.2: Oligonucleotides used for constructing *in vitro* transcription vectors

Primers to linearize RLG770

Oligonucleotide	Sequence
7057	5'-AAGCTTGGGTCCCACCTGACC-3'
7058	5'-GAATTCTTGAAGACGAAAGGGCC-3'

Primers to screen and sequence promoters inserted into RLG770

Oligonucleotide	Sequence
4252	5'-CCCGAAAAGTGCCACCTGACG-3'
4253	5'-CAGTTCCCTACTCTCGCATGG-3'

Primers used to construct *in vitro* transcription plasmids[a]

Plasmid	Oligonucleotide	Sequence	Description
p15221	8006	GGCCCTTTCGTCTTCAAGAATTCC ACGCATTTCTCCGAAAAGCCC	*R. sphaeroides carD* promoter, -161 to +50
	8131	GGTCAGGTGGGACCCAAGCTTACC TCAAGCTGAGAGCTGCG	

[a] For each plasmid constructed, both the forward and reverse primers are given

List of mutagenic primers used to construct plasmids[a]

Plasmid	Oligonucleotide	Sequence	Description
p15487	8375	AGCCGCATTTCGTGGTATG AtTGTTTCACAAACAAGC	*R. sphaeroides carD* promoter A-7T
p15504	8286	GTTGGAATCAGCCGCATTT CGacGTATGAATGTTTCACA AACAA	*R. sphaeroides carD* promoter ext-10 from -15TG-14 to -15AC-14
p15643	8817	GCCGGAGTTGGAATCAGCC GCgTcTgcTGGTATGAATGTT TCACAAAC	*R. sphaeroides carD* promoter spacer -21ATTTCG-16 to -21GTCTGC-16
p15644	8818	CCCGCCGGAGTTGGAATCA GCCGCtaaagcTGGTATGAAT GTTTCACAAACAAGC	*R. sphaeroides carD* promoter spacer -21ATTTCG-16 to -21TAAAGC-16
p15646	8820	CGCCGGAGTTGGAATCAGC CGgATTggaTGGTATGAATG TTTCACAAAC	*R. sphaeroides carD* promoter spacer -22CATTTCG-16 to -22GATTGGA-16
p15681	8831	CCGGAGTTGGAATCAGCCG CATTaCGTGGTATGAAT	*R. sphaeroides carD* promoter spacer T-18A
p15683	8833	CCGGAGTTGGAATCAGCCG CATTgCGTGGTATGAAT	*R. sphaeroides carD* promoter spacer T-18G

[a] Mutations introduced by each oligonucleotide are shown in lowercase

Table 3.3: Plasmids used for *in vitro* transcription

Plasmids	**Promoter**	**Endpoints[a]**	**Source**
p14564	*R. sphaeroides rrnB*	-230 to +50	(21)
p15221	*R. sphaeroides carD*	-161 to +50	This study
P15487	P15221, A-7T		This study
p15504	p15222, ext-10 -15TG-14 to -15AC-14		This study
p15643	p15222, -21GTCTGC-16		This study
p15644	p15222, -21TAAAGC-16		This study

p15646	p15222, -22GATTGGA-16		This study
p15681	p15222, T-18A		This study
p15683	p15222, T-18G		This study

[a] Endpoints are relative to the predicted transcription start site, unless a reference is listed for the source, in which case endpoints are from what was published previously.

Table 3.4: Oligonucleotides used for generating *carD* promoter fragment for abortive transcription

Oligonucleotide	Sequence	Description
8660	5' – GATAACGGTTTTGTGGAAACCCGACC – 3'	Generates 120 nucleotide DNA fragment containing Rsp *carD* promoter, -72 to +50
8661	5' – ACCTCAAGCTGAGAGCTGCG – 3'	

Table 3.5: Sequence of *carD* promoter fragment in abortive transcription assay

Sequence (5' to 3')	Description
GATAACGGTTTTGTGGAAACCCGACCCCCCGCCGGA G<u>TTGGAA</u>TCAGCCGCATTTCGTGG**TATGAA**TGTTTC**<u>A</u>** CAAACAAGCACGGAAGCCTTCACCCACCTCGCAGCTC TCAGCTTGAGGT	-35 element is underlined, -10 element is bold and TSS (+1) is bold and underlined

REFERENCES:

1. Haugen SP, Ross W, Gourse RL. 2008. Advances in bacterial promoter recognition and its control by factors that do not bind DNA. Nat Rev Micro 6:507–519.
2. Saecker RM, Record MT, deHaseth PL. 2011. Mechanism of Bacterial Transcription Initiation: RNA Polymerase - Promoter Binding, Isomerization to Initiation-Competent Open Complexes, and Initiation of RNA Synthesis. J Mol Biol 412:754–771.
3. Henderson KL, Evensen CE, Molzahn CM, Felth LC, Dyke S, Liao G, Shkel IA, Record MT. 2019. RNA Polymerase: Step-by-Step Kinetics and Mechanism of Transcription Initiation. Biochemistry 58:2339–2352.
4. Feklístov A, Sharon BD, Darst SA, Gross CA. 2014. Bacterial Sigma Factors: A Historical, Structural, and Genomic Perspective. Annual Review of Microbiology 68:357–376.
5. Appleman JA, Ross W, Salomon J, Gourse RL. 1998. Activation of Escherichia coli rRNA Transcription by FIS during a Growth Cycle. J Bacteriol 180:1525–1532.
6. Ross W, Thompson JF, Newlands JT, Gourse RL. 1990. E.coli Fis protein activates ribosomal RNA transcription in vitro and in vivo. EMBO J 9:3733–3742.
7. Lee DJ, Minchin SD, Busby SJW. 2012. Activating Transcription in Bacteria. Annual Review of Microbiology 66:125–152.
8. Sanchez-Vazquez P, Dewey CN, Kitten N, Ross W, Gourse RL. 2019. Genome-wide effects on Escherichia coli transcription from ppGpp binding to its two sites on RNA polymerase. PNAS 201819682.
9. Paul BJ, Berkmen MB, Gourse RL. 2005. DksA potentiates direct activation of amino acid promoters by ppGpp. PNAS 102:7823–7828.
10. Gopalkrishnan S, Ross W, Chen AY, Gourse RL. 2017. TraR directly regulates transcription initiation by mimicking the combined effects of the global regulators DksA and ppGpp. PNAS 114:E5539–E5548.
11. Gourse RL, Chen AY, Gopalkrishnan S, Sanchez-Vazquez P, Myers A, Ross W. 2018. Transcriptional Responses to ppGpp and DksA. Annual Review of Microbiology 72:163–184.
12. Chandrangsu P, Lemke JJ, Gourse RL. 2011. The dksA promoter is negatively feedback regulated by DksA and ppGpp. Molecular Microbiology 80:1337–1348.
13. Srivastava DB, Leon K, Osmundson J, Garner AL, Weiss LA, Westblade LF, Glickman MS, Landick R, Darst SA, Stallings CL, Campbell EA. 2013. Structure and function of CarD, an essential mycobacterial transcription factor. PNAS 110:12619–12624.
14. Gallego-García A, Mirassou Y, García-Moreno D, Elías-Arnanz M, Jiménez MA, Padmanabhan S. 2014. Structural Insights into RNA Polymerase Recognition and Essential Function of Myxococcus xanthus CdnL. PLOS ONE 9:e108946.

15. Bae B, Chen J, Davis E, Leon K, Darst SA, Campbell EA. 2015. CarD uses a minor groove wedge mechanism to stabilize the RNA polymerase open promoter complex. eLife Sciences e08505.
16. Rammohan J, Ruiz Manzano A, Garner AL, Stallings CL, Galburt EA. 2015. CarD stabilizes mycobacterial open complexes via a two-tiered kinetic mechanism. Nucleic Acids Res 43:3272–3285.
17. Zhu DX, Garner AL, Galburt EA, Stallings CL. 2019. CarD contributes to diverse gene expression outcomes throughout the genome of *Mycobacterium tuberculosis*. Proc Natl Acad Sci USA 116:13573–13581.
18. Stallings CL, Stephanou NC, Chu L, Hochschild A, Nickels BE, Glickman MS. 2009. CarD is an essential regulator of rRNA transcription required for Mycobacterium tuberculosis persistence. Cell 138:146–159.
19. Galburt EA. 2018. The calculation of transcript flux ratios reveals single regulatory mechanisms capable of activation and repression. Proc Natl Acad Sci USA 115:E11604–E11613.
20. Boyaci H, Saecker RM, Campbell EA. 2019. Transcription initiation in mycobacteria: a biophysical perspective. Transcription 0:1–13.
21. Henry KK, Ross W, Myers KS, Lemmer KC, Vera JM, Landick R, Donohue TJ, Gourse RL. 2020. A majority of Rhodobacter sphaeroides promoters lack a crucial RNA polymerase recognition feature, enabling coordinated transcription activation. PNAS 117:29658–29668.
22. Burger BT, Imam S, Scarborough MJ, Noguera DR, Donohue TJ. 2017. Combining Genome-Scale Experimental and Computational Methods To Identify Essential Genes in *Rhodobacter sphaeroides*. mSystems 2:e00015-17.
23. Gallego-García A, Iniesta AA, González D, Collier J, Padmanabhan S, Elías-Arnanz M. 2017. Caulobacter crescentus CdnL is a non-essential RNA polymerase-binding protein whose depletion impairs normal growth and rRNA transcription. Scientific Reports 7:srep43240.
24. Woldemeskel SA, Daitch AK, Alvarez L, Panis G, Zeinert R, Gonzalez D, Smith E, Collier J, Chien P, Cava F, Viollier PH, Goley ED. 2020. The conserved transcriptional regulator CdnL is required for metabolic homeostasis and morphogenesis in Caulobacter. PLOS Genetics 16:e1008591.
25. Myers KS, Vera JM, Lemmer KC, Linz AM, Landick R, Noguera DR, Donohue TJ. 2020. Genome-Wide Identification of Transcription Start Sites in Two Alphaproteobacteria, Rhodobacter sphaeroides 2.4.1 and Novosphingobium aromaticivorans DSM 12444. Microbiol Resour Announc 9.
26. Vera JM, Ghosh IN, Zhang Y, Hebert AS, Coon JJ, Landick R. 2020. Genome-Scale Transcription-Translation Mapping Reveals Features of Zymomonas mobilis Transcription Units and Promoters. mSystems 5.
27. Schlüter J-P, Reinkensmeier J, Barnett MJ, Lang C, Krol E, Giegerich R, Long SR, Becker A. 2013. Global mapping of transcription start sites and promoter motifs in

the symbiotic α-proteobacterium Sinorhizobium meliloti1021. BMC Genomics 14:156.

28. Cortes T, Schubert OT, Rose G, Arnvig KB, Comas I, Aebersold R, Young DB. 2013. Genome-wide Mapping of Transcriptional Start Sites Defines an Extensive Leaderless Transcriptome in Mycobacterium tuberculosis. Cell Reports 5:1121–1131.
29. Singh SS, Typas A, Hengge R, Grainger DC. 2011. Escherichia coli σ 70 senses sequence and conformation of the promoter spacer region. Nucleic Acids Research 39:5109–5118.
30. Yuzenkova Y, Tadigotla VR, Severinov K, Zenkin N. 2011. A new basal promoter element recognized by RNA polymerase core enzyme: Novel class of bacterial promoters. The EMBO Journal 30:3766–3775.
31. Jensen D, Manzano AR, Rammohan J, Stallings CL, Galburt EA. CarD and RbpA modify the kinetics of initial transcription and slow promoter escape of the Mycobacterium tuberculosis RNA polymerase. Nucleic Acids Res https://doi.org/10.1093/nar/gkz449.
32. Carpousis AJ, Gralla JD. 1980. Cycling of ribonucleic acid polymerase to produce oligonucleotides during initiation in vitro at the lac UV5 promoter. Biochemistry 19:3245–3253.
33. Henderson KL, Felth LC, Molzahn CM, Shkel I, Wang S, Chhabra M, Ruff EF, Bieter L, Kraft JE, Jr MTR. 2017. Mechanism of transcription initiation and promoter escape by E. coli RNA polymerase. Proceedings of the National Academy of Sciences 115:E3032–E3040.
34. Durfee T, Hansen A-M, Zhi H, Blattner FR, Jin DJ. 2008. Transcription Profiling of the Stringent Response in Escherichia coli. Journal of Bacteriology 190:1084–1096.
35. Paul BJ, Barker MM, Ross W, Schneider DA, Webb C, Foster JW, Gourse RL. 2004. DksA: A Critical Component of the Transcription Initiation Machinery that Potentiates the Regulation of rRNA Promoters by ppGpp and the Initiating NTP. Cell 118:311–322.
36. Rutherford ST, Villers CL, Lee J-H, Ross W, Gourse RL. 2009. Allosteric control of Escherichia coli rRNA promoter complexes by DksA. Genes Dev 23:236–248.
37. Browning DF, Busby SJW. 2016. Local and global regulation of transcription initiation in bacteria. Nat Rev Micro 14:638–650.
38. Heyduk E, Heyduk T. 2014. Next Generation Sequencing-Based Parallel Analysis of Melting Kinetics of 4096 Variants of a Bacterial Promoter. Biochemistry 53:282–292.
39. Lennon CW, Lemmer KC, Irons JL, Sellman MI, Donohue TJ, Gourse RL, Ross W. 2014. A Rhodobacter sphaeroides Protein Mechanistically Similar to Escherichia coli DksA Regulates Photosynthetic Growth. mBio 5.
40. Voskuil MI, Chambliss GH. 2002. The TRTGn Motif Stabilizes the Transcription Initiation Open Complex. Journal of Molecular Biology 322:521–532.

41. Bae B, Feklistov A, Lass-Napiorkowska A, Landick R, Darst SA. 2015. Structure of a bacterial RNA polymerase holoenzyme open promoter complex. eLife 4:e08504.

42. Eccleston Jr. ED, Gray ED. 1973. Variations in ppGpp levels in Rhodopseudomonas spheroides during adaptation to decreased light intensity. Biochemical and Biophysical Research Communications 54:1370–1376.

43. Cost HB, Gray ED. 1967. Rapidly labeled RNA synthesis during morphogenesis. Biochimica et Biophysica Acta (BBA) - Nucleic Acids and Protein Synthesis 138:601–604.

44. Hsu LM. 1996. [4] Quantitative parameters for promoter clearance, p. 59–71. *In* Methods in Enzymology. Academic Press.

CHAPTER FOUR
ppGpp/DksA inhibits CarD-activated transcription

This chapter is a draft of a manuscript that focuses on ppGpp and DksA regulation of transcription in *R. sphaeroides* using in vivo and in vitro approaches. This project is a continuing collaboration with the Donohue lab (UW-Madison). It will be submitted for publication in 2021 with the following author list: Kemardo K. Henry, Wilma Ross, Bryan D. Lakey, Kevin S. Myers, Timothy J. Donohue and Richard L. Gourse. I performed all the experiments except Figure 4.6 which was done by BDL. BDL and KSM analyzed the data for Figure 4.6, and BDL provided portions of the text relating to experiments performed for Figure 4.6.

ABSTRACT

The model organism *Escherichia coli* increases the levels of the alarmone ppGpp in response to amino acid starvation. ppGpp and its binding partner DksA directly regulate the transcription of over 700 genes, including rRNA synthesis. The photosynthetic α-proteobacterium *Rhodobacter sphaeroides* accumulates ppGpp transiently in response to changes in light intensity rather than amino acid starvation, and the DksA homolog in this bacterium appears to play a role in photosynthetic growth. We found that DksA and ppGpp inhibited CarD-activated *R. sphaeroides* promoters, including those for rRNA and r-protein. Genome-wide expression studies identified significant changes in gene expression when *dksA* was deleted in *R. sphaeroides*, further suggesting that DksA is an important regulator of *R. sphaeroides* gene expression.

INTRODUCTION:

ppGpp and DksA directly regulate transcription of more than 700 genes in *Escherichia coli* (1). For example, upon amino acid starvation, DksA/ppGpp binds directly to RNA polymerase (RNAP), reduces transcription from rRNA, r-protein and other promoters, and stimulates transcription from amino acid biosynthetic as well as other gene promoters (2–4). Structural and biochemical studies showed that ppGpp binds to two sites on RNA polymerase: site one is at the interface of the RNA polymerase ω and β' subunits while site two is formed at the interface of the transcription factor DksA and the β‘ rim-helix (3, 5, 6). Together, ppGpp and DksA work to inhibit transcription by destabilizing open complex formation on some promoters, while they activate some other promoters by shifting the equilibrium towards open complex formation (2, 7).

Studies of *E. coli* rRNA transcription have provided a framework for understanding promoter recognition by RNAP in general. However, negative regulation of rRNA promoters in most other bacteria, for example photosynthetic bacteria, has not been explored in great detail. Like virtually all bacteria studied to date, the photosynthetic bacterium *Rhodobacter sphaeroides* synthesizes ppGpp. However, in *R. sphaeroides*, ppGpp appears to control responses to light rather than to amino acid availability as in *E. coli* (8, 9).

R. sphaeroides is a gram-negative α-proteobacterium capable of aerobic and anaerobic respiration and anoxygenic photosynthesis (10). When grown aerobically, *R. sphaeroides* growth is similar to that of other gram-negative bacteria, utilizing oxygen as the terminal electron acceptor. However, once oxygen levels drop, *R. sphaeroides*

undergoes a morphological change to prepare for photosynthetic growth (11). This switch in growth capabilities is followed by the invagination of the cytoplasmic membrane where red-pigmented light-harvesting complexes are inserted (12, 13).

Recently we showed that a DksA homolog in *R. sphaeroides* (*Rsp*2654) functions similarly to *E.* coli DksA, inhibiting transcription of *E. coli* rRNA promoters by *E. coli* RNAP *in vitro* in conjunction with ppGpp (14). $DksA_{Rsp}$ functions with *E. coli* RNAP to complement an *E coli* strain deleted for *dksA*, and synergizes with ppGpp *in vitro* to significantly reduce the lifetime of complexes formed with either *R. sphaeroides* RNAP or *E. coli* RNAP and either the *lacUV5* promoter or the plasmid encoded *RNA-1* promoter (14). *R. sphaeroides* cells deleted for $dksA_{Rsp}$ have reduced pigmentation and grow poorly under photosynthetic conditions, indicating an as yet undefined role for $DksA_{Rsp}$ in photosynthetic growth. However, direct regulation of *R. sphaeroides* promoters, including rRNA promoters, by DksA/ppGpp had not been demonstrated previously with *R. sphaeroides* RNAP, because the promoters were transcribed very poorly *in vitro*. Previous studies of DksA homologs from other bacterial species were performed in non-cognate systems, using *E. coli* promoters and *E. coli* RNAP (e.g. (15, 16)).

The transcription factor CarD is well characterized in mycobacterial species where it was shown to stabilize open complex formation and activate rRNA transcription (17–19) in vitro. The *R. sphaeroides* CarD homolog was recently shown to be needed for transcription from *R. sphaeroides* ribosomal RNA promoters and some non-ribosomal promoters (20). A large number of *R. sphaeroides* promoters lack the critical thymine at the last position of the -10 element (-7T), but CarD can activate transcription

from these promoters *in vitro* (20, 21). The finding that CarD activated rRNA transcription in *R. sphaeroides* provided us with the opportunity to investigate the role of DksA/ppGpp on transcription from rRNA and other promoters in vitro.

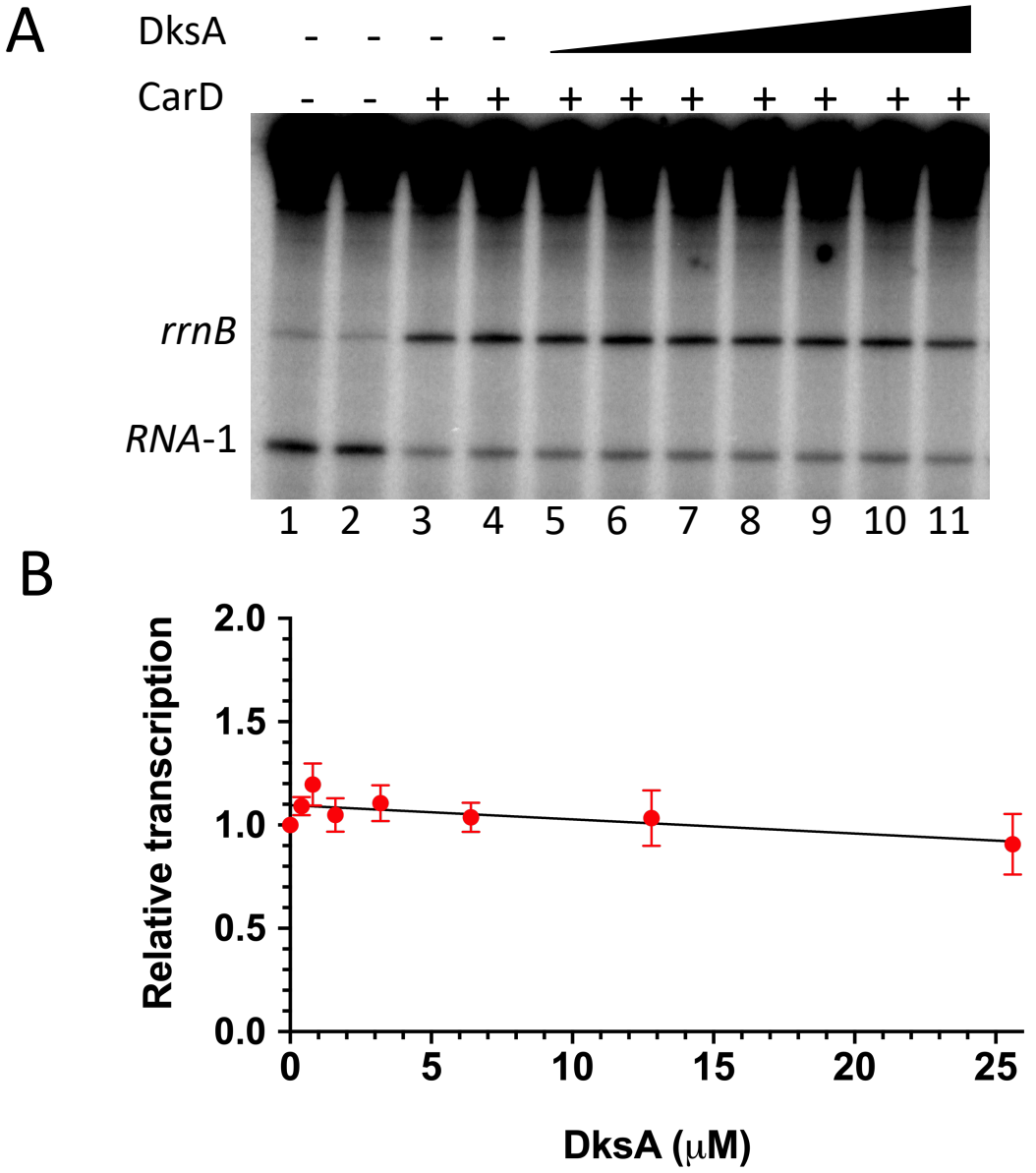
Having established an in vitro system for analysis of *R. sphaeroides* rRNA transcription in vitro using RNAP holozyme and CarD, we demonstrate here that ppGpp and its co-factor $DksA_{Rsp}$ directly negatively regulate $CarD_{Rsp}$-activated promoters, particularly for rRNA and other genes involved in translation, but it does not inhibit promoters that form more stable complexes with *R. sphaeroides* RNAP like the CarD promoter itself. We further show that *R. sphaeroides* strains containing or lacking the *dksA* gene show differential regulation of several hundred genes *in vivo* during both aerobic and photosynthetic growth, indicating a global role for DksA in *R. sphaeroides* gene expression.

We show by Western blotting of aerobically grown cultures that the DksA and CarD proteins are maintained at a constant concentration during logarithmic growth, and both decline steeply on the same time scale in stationary phase. In contrast, the relative amounts of CarD and DksA differ in photosynthetically grown anaerobic cultures, where CarD declines much earlier than DksA during the approach to stationary phase. These results suggest that the differential control of the CarD and DksA concentrations plays a role in the photosynthetic and possibly other lifestyles of this α-proteobacterium.

RESULTS

***R. sphaeroides* DksA does not inhibit transcription of the CarD activated *R. sphaeroides rrnB* promoter in the absence of ppGpp.** Having identified CarD as a required transcription factor for *R. sphaeroides* rRNA transcription by *R. sphaeroides* RNAP, we designed a assay to test the regulatory effects of ppGpp and DksA on gene transcription in vitro. To visualize transcription from the *rrnB* promoter, we used a saturating concentration of CarD to maximize transcription, as reported previously (Figure 4.1A; compare lanes 1 and 2 with lanes 3 and 4; (20)). Addition of a range of concentrations of *R. sphaeroides* DksA to the reaction resulted in little or no inhibition of transcription (Figure 4.1A, lanes 5 to 11). This finding was particularly striking because *E. coli* DksA inhibits *E. coli rrnB* transcription *by E. coli* RNAP at moderate concentrations (IC_{50} of ~1.3 μM; (2, 7, 22)) in the absence of ppGpp, whereas *R. sphaeroides* DksA inhibited *rrnB* transcription only minimally, even at the highest concentration tested (~25 μM; Figure 4.1B). This observation is consistent with previously observed effects of *R. sphaeroides* DksA in which it was shown to have a lesser inhibitory effect than *E. coli* DksA on transcription from the *E. coli rrnB* P1 promoter using *E. coli RNAP* in the absence of ppGpp (14) and is not due to the purified DksA being inactive.

Figure 4.1: DksA$_{Rsp}$ alone (without ppGpp) does not inhibit transcription from the *R. sphaeroides rrnB* promoter. (A) in vitro transcription of the *rrnB*$_{Rsp}$ promoter with RNAP$_{Rsp}$ (20nM), with (+ = 1280nM) or without CarD, and with DksA where indicated (wedge indicates DksA concentration range of 0.4-25.6μM; lanes 5-11) in 170mM NaCl. CarD activates transcription from the *rrnB* promoter ~13-fold (lanes 3,4 vs 1,2) while DksA$_{Rsp}$ has little or no inhibitory effect on the CarD activated *rrnB* promoter (lanes 5-11 vs lanes 3,4). (B) Relative transcription of the *R. sphaeroides rrnB* promoter in the presence of increasing concentration of DksA. Transcription was normalized to the CarD-activated *rrnB* signal (without DksA). Error bars indicate standard deviation from 3 separate assays.


A
DksA - - - -
CarD - - + + + + + + + + +
rrnB
RNA-1
1 2 3 4 5 6 7 8 9 10 11
B
Relative transcription
2.0
1.5
1.0
0.5
0.0
0 5 10 15 20 25
DksA (μM)

ppGpp inhibits transcription from *R. sphaeroides rrnB* in the absence of DksA. Next, we tested whether ppGpp alone, in the absence of DksA, inhibited transcription of the CarD-activated *rrnB* promoter (Figure 4.2A, lanes 5 to 10). We found that ppGpp inhibited *rrnB* transcription ~ 3-fold with an IC50 of ~7μM (Figure 4.2B). The magnitude of transcription inhibition is similar to the fold inhibition observed for the *E. coli rrnB* P1 promoter in transcription by *E. coli* RNAP (3). (We note that the *R. sphaeroides* and *E. coli* assays differ in that an additional transcription factor was not required for detection of transcription of *E. coli rrnB* P1).

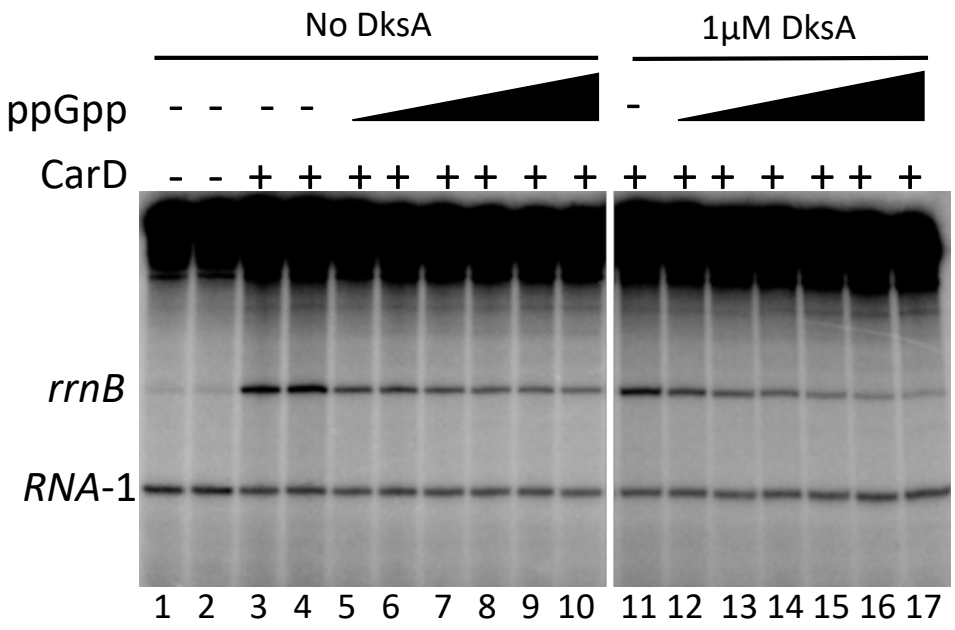
In *E. coli*, the effect of ppGpp in the absence of DksA is mediated by direct binding of ppGpp to RNAP at the interface of the ω and β‘ subunits (referred to as ppGpp Binding Site 1) (3, 5, 6, 23). The effect of ppGpp on transcription by *R. sphaeroides* RNAP (Figure 4.2) is consistent with conservation between *E. coli* and *R. sphaeroides* of almost all the residues involved in ppGpp binding to Site 1 (Figure 4.S1) and with the binding of radio-labeled ppGpp (^{32}P-ppGpp) to purified *R. sphaeroides* RNAP in a differential radial capillary action of ligand assay (DRaCALA) (Figure 4.S2).

We did not attempt to construct an RNAP$_{Rsp}$ strain deleted for the ω subunit (required for formation of ppGpp binding to Site 1), because the ω subunit is essential in *R. sphaeroides* (25). Therefore, we could not purify *R. sphaeroides* RNAP lacking ω to directly test the effect of binding of ppGpp to Site 1 versus to an RNAP without Site 1.

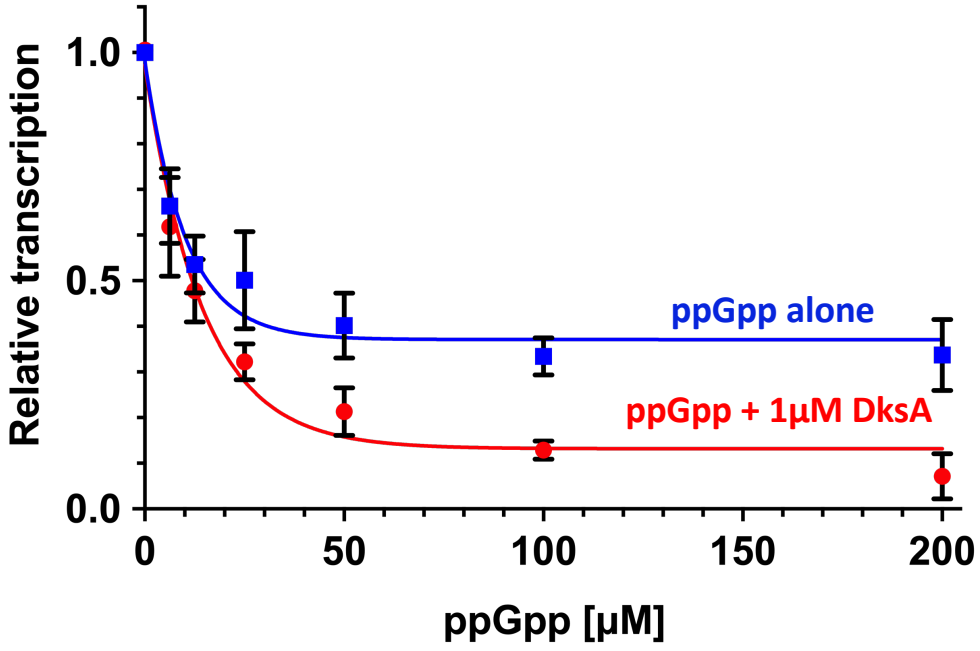
DksA$_{Rsp}$ and ppGpp synergistically inhibit *rrnB* transcription *in vitro* by *R. sphaeroides* RNAP. The effect of DksA$_{Eco}$ on *E. coli* RNAP is increased by direct

Figure 4.2: ppGpp and DksA synergistically inhibit *rrnB* transcription. (A) in vitro transcription of the $rrnB_{Rsp}$ promoter with $RNAP_{Rsp}$ (20nM) with (+=1280nM) or without CarD, with ppGpp where indicated (wedge indicates ppGpp concentration range of 0.625 to 200μM), and with (lanes 11-17) or without (lanes 1-10) 1μM DksA, in 170mM NaCl. ppGpp inhibits the CarD activated *rrnB* promoter ~3-4 fold (left panel) and ppGpp along with DksA synergistically inhibits rrnB transcription ~20-fold. The middle portion of the gel was removed for clarity. (B) Relative transcription of the *R. sphaeroides rrnB* promoter as a function of ppGpp concentration, with or with 1μM DksA, from experiments like that in (A). Transcription was normalized to the CarD-activated *rrnB* signal (without DksA or ppGpp). The half-max of inhibition is 7μM with ppGpp alone and 10μM in the presence of DksA. Error bars indicate standard deviation from 3-5 separate assays.

A


No DksA
1µM DksA
ppGpp
CarD
rrnB
RNA-1
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17

B


Relative transcription
1.0
0.5
0.0
0
50
100
150
200
ppGpp [µM]
ppGpp alone
ppGpp + 1µM DksA

binding of ppGpp to RNAP at the interface of DksA and the β‘ subunit rim helix (referred to as ppGpp Binding Site 2) (3, 5, 6). Although $DksA_{Rsp}$ alone did not inhibit transcription from the $rrnB_{Rsp}$ promoter (Figure 4.1A), ppGpp and $DksA_{Rsp,}$ together inhibited transcription at a DksA concentration that had little or no effect on transcription alone (1μM). ppGpp/DksA inhibited *rrnB* transcription ~10 to 20-fold, with a ppGpp IC_{50} of 10μM (Figure 4.2A, lanes 11 to 17; Figure 4.2B). Thus, ppGpp and ppGpp/DksA are able to inhibit gene transcription even in the presence of CarD. The fold inhibition of $rrnB_{Rsp}$ gene transcription by $DksA_{Rsp}$ + ppGpp together was similar to that previously reported for inhibition of rRNA transcription by *E. coli* RNAP with ppGpp/$DksA_{Eco}$ (3). The results in Figure 4.2 are consistent with the model that $DksA_{Rsp}$ forms a ppGpp binding site (Site 2) at its interface with the β‘ rim-helix and with the observation that almost all the residues that interact with ppGpp in *E. coli* β' and in *E. coli* DksA are conserved in *R. sphaeroides* (Figure 4.S1).

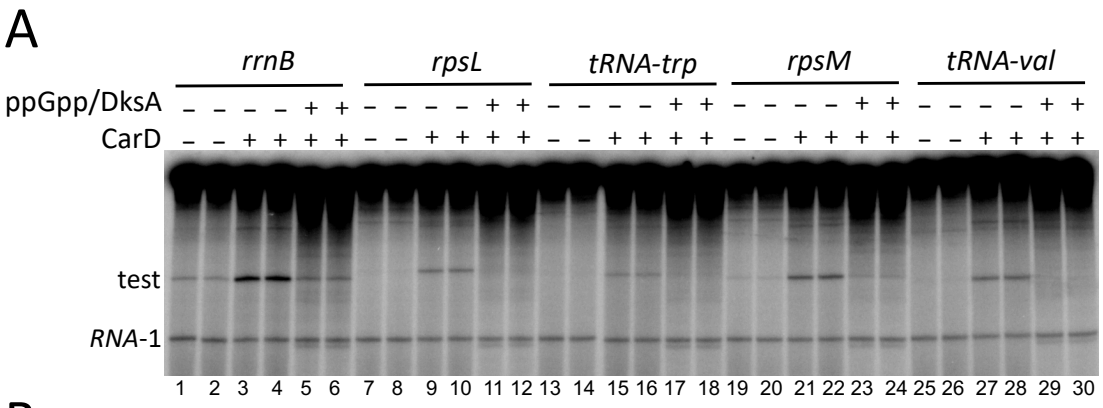
Our findings on the synergism between the effects of $DksA_{Rsp}$ and ppGpp are also consistent with previous results with $DksA_{Rsp}$ and ppGpp using an assay that reported their effects on the half-life of the *lacUV5* promoter complex formed with *R. sphaeroides* RNAP (14). $DksA_{Rsp}$ and ppGpp reduced the half-time of the $RNAP_{Rsp}$-*lacUV* promoter complex ~50-fold, whereas ppGpp reduced the lifetime of the promoter complex only 4-fold (14).

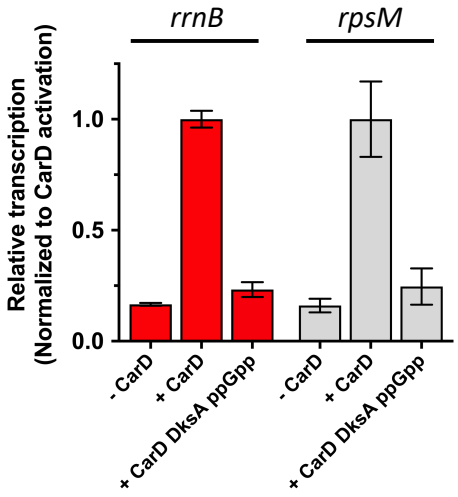
$DksA_{Rsp}$ and ppGpp inhibit transcription from other CarD-activated promoters. In *E. coli*, in addition to regulating rRNA transcription, ppGpp and DksA directly regulate transcription of hundreds of promoters, many of which are promoters related to components of the translation apparatus, including ribosomal proteins (1, 4). To

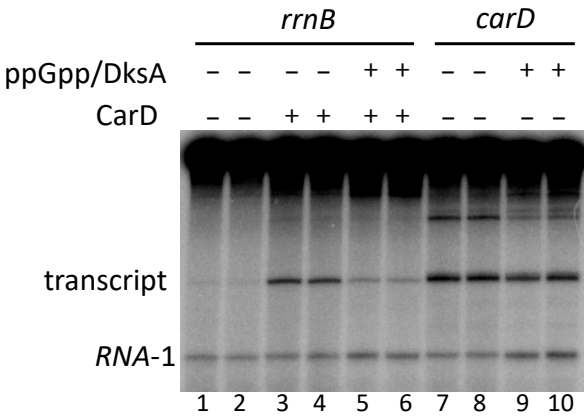
investigate whether DksA$_{Rsp}$ and ppGpp play a similar role in *R. sphaeroides*, we studied 4 promoters that were previously shown to be activated by CarD (20). Similar to the transcription assay using the *R. sphaeroides rrnB* promoter, *in vitro* transcription was performed on reactions containing CarD (for activation), 100μM ppGpp, and 1μM DksA. At these concentrations of ppGpp and DksA, *rrnB* transcription was inihibited ~10-fold) (Figure 4.2).

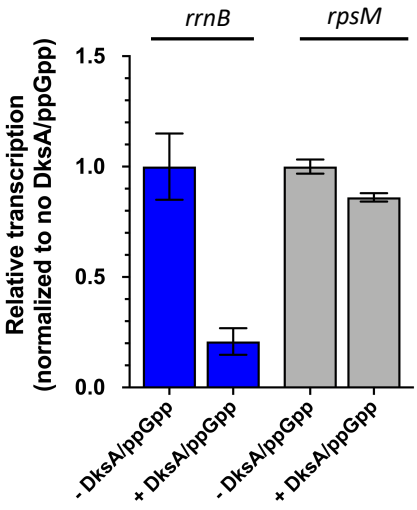
ppGpp and DksA inhibited transcription from all 4 CarD-activated promoters tested, as well as from the *rrnB* promoter (Figure 4.3, compare lanes 3-4, 9-10, 15-16, 21-22 and 27-28 with lanes 5-6, 11-12, 17-18, 23-24 and 29-30 respectively). The inhibited promoters tested included the promoters for *rpsM* (which transcribes a polycistronic mRNA encoding the r-proteins S13 and S11), for tRNA-val (valine tRNA), for *rpsL* (the mRNA for r-proteins S12, S7, for the fusA1 gene, which codes for the translation elongation factor EF-G), and for tRNA-trp (tryptophan tRNA). The concentration of ppGpp/DksA used in these assays inhibited transcription from the CarD-activated *rpsM* promoter ~4-fold and from the CarD-activated *rrnB* promoter ~4-fold (Figure 4.3B). The lower fold-inhibition observed for the *rrnB* promoter compared to what was observed in Figure 4.2 could be because the solution conditions used to test CarD activation of the non-rRNA promoters included only 100mM NaCl, rather than the 170mM NaCl used in Figure 4.2). The fold inhibition of the tRNA-val, tRNA-trp, and *rpsL* promoters could not be calculated since transcription in the presence of CarD and DksA/ppGpp was too low to quantify. However, for each of the promoters tested here, DksA and ppGpp together strongly inhibited transcription even though the promoter was activated by CarD.

Figure 4.3: ppGpp and DksA synergistically inhibit transcription from some *R. sphaeroides* promoters, but do not inhibit transcription from the *carD* promoter. (A) In vitro transcription with *R. sphaeroides* RNAP (20nM) of *R. sphaeroides* promoters in 100mM NaCl. Duplicate lanes without CarD, with 1280 nM CarD, or with 1280nM CarD plus 100μM ppGpp and 1μM DksA are shown for each promoter. (B) Relative transcription of the *rrnB* and *rpsM* promoters without CarD, with CarD, or with CarD plus DksA and ppGpp, from experiments like that in (A). Values are normalized to the +CarD samples. Error bars are generated from 2 independent assays. Transcript levels for basal and ppGpp/DksA inhibited transcription from *rpsL*, *tRNA-trp* and *tRNA-val* were too low to quantify. (C) In vitro transcription of either the *rrnB* or *carD* promoter with RNAP (20nM) in 170mM NaCl. For the *rrnB* promoter, duplicate lanes without CarD, with CarD (1280nM), or with CarD and ppGpp (100μM) and DksA (1μM) (as in (A), but at 170mM NaCl). For the *carD* promoter duplicate lanes with no CarD or ppGpp/DksA or with CarD (1280nM) and ppGpp (100μM) and DksA (1μM). (D) Relative transcription of the *rrnB* promoter (+ CarD/ + CarD and DksA and ppGpp) or the *carD* promoter with or without DksA and ppGpp. Error bars are generated from 2 independent assays.


A
rrnB
rpsL
tRNA-trp
rpsM
tRNA-val
ppGpp/DksA
CarD
test
RNA-1
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30


B
rrnB
rpsM
Relative transcription
(Normalized to CarD activation)
1.0
0.5
0.0
- CarD
+ CarD
+ CarD DksA ppGpp
- CarD
+ CarD
+ CarD DksA ppGpp

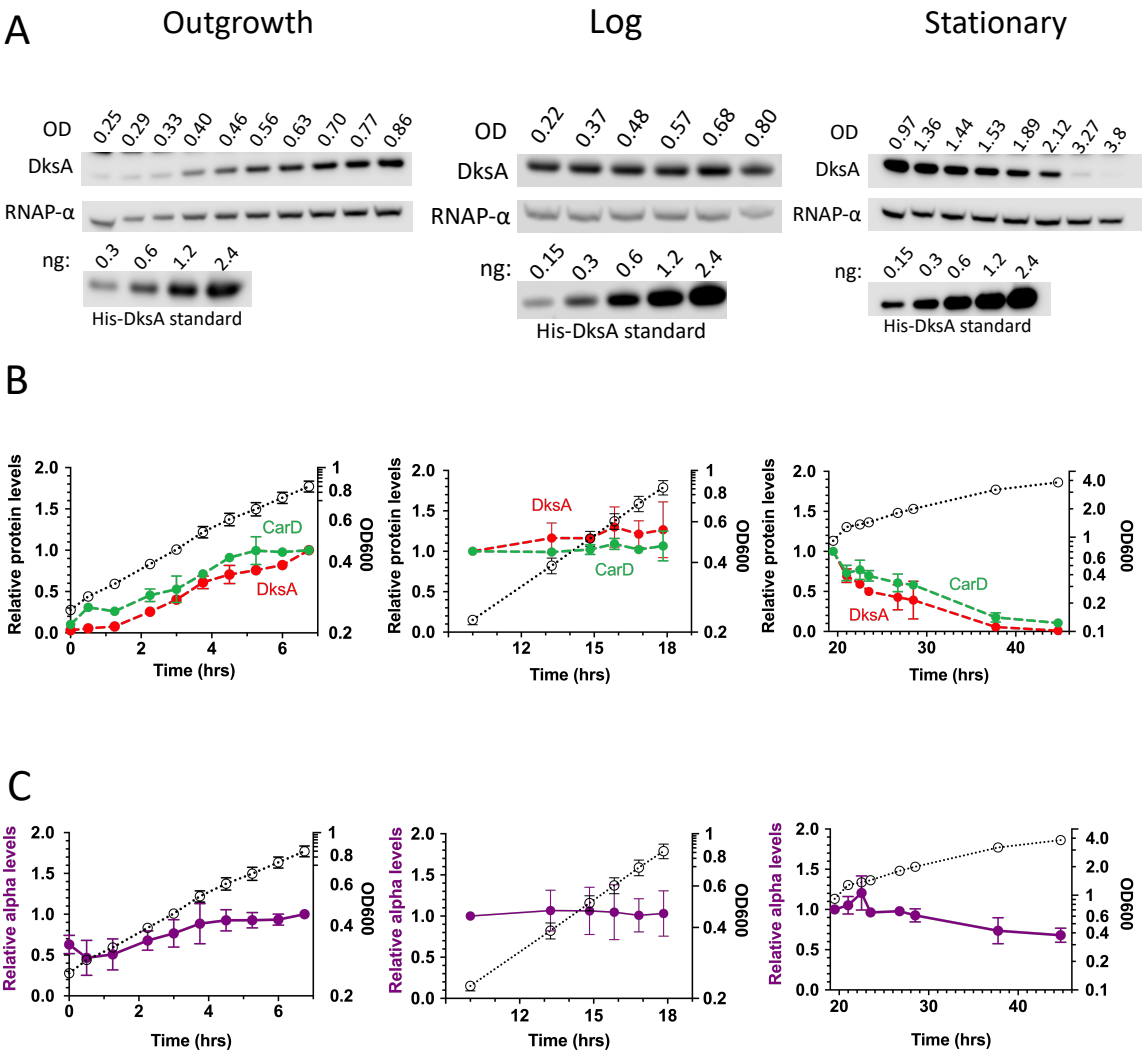

C
rrnB
carD
ppGpp/DksA
CarD
transcript
RNA-1
1 2 3 4 5 6 7 8 9 10


D
rrnB
rpsM
Relative transcription
(normalized to no DksA/ppGpp)
1.5
1.0
0.5
0.0
- DksA/ppGpp
+ DksA/ppGpp
- DksA/ppGpp
+ DksA/ppGpp

DksA$_{Rsp}$ with ppGpp does not inhibit transcription from the *carD* promoter. CarD activates transcription from a large number of *R. sphaeroides* promoters and has no effect on other promoters (20). In Chapter 3, I presented evidence that CarD inhibits transcription from its own promoter by limiting the rate of promoter escape. Moreover, CarD protein levels are constant during log growth and decrease upon entry into stationary phase, a time when ribosome synthesis decreases and ppGpp levels increase (20, 26). To address whether the decrease in CarD protein levels during stationary phase is due to inhibition of its transcription by DksA and ppGpp, we tested their effects on transcription from the *R. sphaeroides carD* promoter *in vitro*. Under the conditions tested, ppGpp and DksA failed to significantly inhibit transcription from the *carD* promoter, while the *rrnB* promoter (positive control) was inhibited ~10-fold (Figure 4.3C, compare lanes 3-4 and 7-8 with 5-6 and 9-10 respectively; Figure 4.3D). The findings of Chapter 3 suggested that the *carD* promoter forms stable open complexes. Stable open complexes are not inhibited by DksA and ppGpp (2, 27).

DksA and CarD levels change similarly in vivo during aerobic growth. Our data indicate that ppGpp/DksA and CarD can each affect transcription by *R. sphaeroides* RNAP in vitro, and DksA/ppGpp can inhibit promoters activated by saturating concentrations of CarD (Figures 4.1, 4.2; (20)). Thus, we conclude that the cellular concentrations of each factor determine when it will regulate transcription of different promoters in vivo. CarD levels were measured during outgrowth from stationary phase, during logarithmic growth, and in stationary phase in aerobically grown cultures by

Figure 4.4: DksA and CarD levels at outgrowth, log and stationary phase in aerobically grown cells. Western blots were performed from *R. sphaeroides* cells grown aerobically at 30 ℃ and sampled at the indicated times. Saturated culture were diluted in fresh minimal media at time 0, and samples were removed at the indicated OD_{600} values for Western blots. $DksA_{Rsp}$ levels were determined with a polyclonal antibody raised against $DksA_{Rsp}$ (9) and $RNAP_{Rsp}$ was determined with an antibody for RNAP-α subunit. Purified DksA standards were analyzed in parallel to determine cellular concentration of DksA. DksA and RNAP-α levels were determined in the same samples. CarD protein levels (green) were previously determined from the same samples (see *Materials and Methods*). (A) DksA and RNAP-α levels in outgrowth, log and stationary phase cells. Relative CarD and DksA (B) and RNAP-α purple (C) levels normalized to overlapping OD_{600}. Error ranges were determined from two experiments for each growth phase.


A
Outgrowth
Log
Stationary
OD
DksA
RNAP-α
ng:
His-DksA standard
B
Relative protein levels
OD600
Time (hrs)
CarD
DksA
C
Relative alpha levels
OD600
Time (hrs)

Western blotting using an antibody to $CarD_{Rsp}$ (Chapter 2, Figure 2; (20)). Western blots were performed on the same protein lysates as were assayed with antibodies to CarD but with antibodies to $DksA_{Rsp}$. We found previously that CarD levels increased rapidly after dilution of stationary phase aerobically grown cells into fresh medium and they plateaued during log phase at a cellular concentration of ~1-2μM. CarD then decreased ~10-fold to ~100nM when the cells entered stationary phase (Figure 4.4B). (The CarD data shown here in Figure 4.4B were reported in reference (20) but are repeated here for comparison with DksA levels).

Like CarD, DksA levels also increased rapidly during outgrowth although they appeared to reach a plateau somewhat later than CarD (Figure 4.4B, left and middle panels). The levels of CarD, DksA and the α subunit of RNAP (used as a control) remained constant from 10 to 18 hours after inoculation (Figures 4.4B and C, center panels). As the cells approached stationary phase, however, CarD and DksA levels both decreased at least 10-fold, while the cellular abundance of the RNAP alpha subunit remained relatively constant. Viable cell measurements indicated that the reductions in CarD and DksA protein levels were not from a loss of cell viability (Figure 4.S4). Using a cell volume of 0.87μm for aerobically grown cells (28), along with purified DksA samples as standards, we estimated that DksA concentrations in the cell were ~0.3-0.6μM during log phase and <60nM during stationary phase.

DksA and CarD levels change differentially in vivo during photosynthetic growth. Next, we determined CarD and DksA levels in photosynthetically grown cells using

Western blotting. Two minimal media cultures were grown, one culture to monitor outgrowth from stationary phase, and the second culture to monitor logarithmic growth, entry into stationary phase and stationary phase growth. Samples were taken at time points indicated and Western blots probing levels of CarD, DksA and RNAP-α were carried out as described in Methods (Figure 4.5).

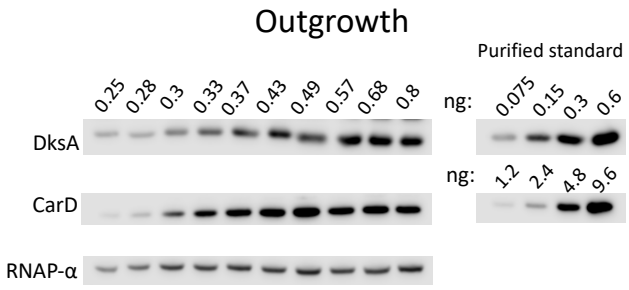
During outgrowth, DksA and CarD concentrations increased about 10- and 30-fold respectively, while the level of RNAP-α subunit increased only ~2-fold (Figure 4.5A, 4.5C and 4.5D). CarD levels increased at a faster rate than DksA levels, reached a peak at ~4 hours after inoculation, and declined to its log phase plateau level by ~7 hr (at an OD_{600} of ~0.9; Figure 4.5C).

As the cells proceeded through log phase (right panels in Figure 4.5C and 4.5D), CarD and DksA levels remained relatively constant from ~11 to 18 hours after inoculation (from OD_{600} 0.15 to ~0.5). However, the CarD concentrations began to decrease at ~18 hr after inoculation, well before the DksA levels decreased, and continued to decline steeply out to 40 hours (Figure 4.5C). The drop in CarD levels during this time were not due to a drop in cell viability (Figure 4.S3B). CarD levels dropped ~20 fold to about ~90nM. (Interestingly, the OD_{600} at which CarD levels decreased in the outgrowth and log phase cultures was ~ the same, at OD_{600} about 0.6).

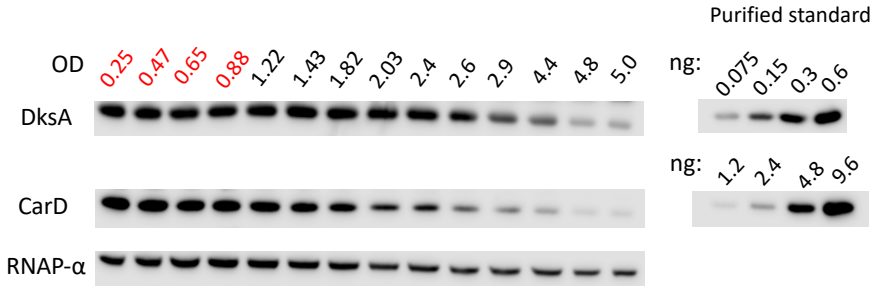
In contrast to the decline in CarD levels beginning at ~18 hours, DksA levels remained constant for an additional ~8 hr (to 24 hours of growth, corresponding to an OD of 2.0; Figure 4.5B and 4.5C), and then declined after ~28 hr. Using a cell volume of 1.49μm^3 for photosynthetically grown cells (28) and the purified standards for CarD and DksA (Figure 4.5B), we determined that the CarD and DksA cellular concentrations

Figure 4.5: DksA and CarD levels at outgrowth and log to stationary phase in photosynthetically growth cells. Western blots were performed from *R. sphaeroides* cells grown photosynthetically at 30 ℃ and sampled at the indicated times. Saturated cultures were diluted in fresh minimal media at time 0, and samples were removed at the indicated OD_{600} values for Western blots. The outgrowth experiment (A) and the log-stationary (B) are from separate growth experiments. $DksA_{Rsp}$ and CarD levels were determined with polyclonal antibodies raised against $DksA_{Rsp}$ or $CarD_{Rsp}$ and $RNAP_{Rsp}$ was determined with an antibody for RNAP-α subunit. Purified DksA or CarD standards were analyzed in parallel to determine cellular concentration of each factor. CarD or DksA and RNAP-α levels were determined in the same samples. (A) Western blot of DksA, CarD, and RNAP-α in outgrowth cells. (B) DksA, CarD and RNAP-α levels cells grown from log (red OD_{600} values) into stationary phase (black OD_{600} values). The log phase of the experiment (OD_{600} value 0.25 to 0.88) was blotted separately with the purified standards to determine the cellular concentration of either CarD or DksA. (C) Relative levels of CarD (green), DksA (red) and OD_{600} values (black) plotted as a function of time of growth after inoculation of the culture in either the outgrowth culture or the log to stationary phase culture. (D) Relative RNAP-α levels (purple) levels and OD600 values from the same cultures and samples as in (C), plotted as in normalized to overlapping OD_{600} values. Error ranges were determined from two experiments for each growth phase experiment.

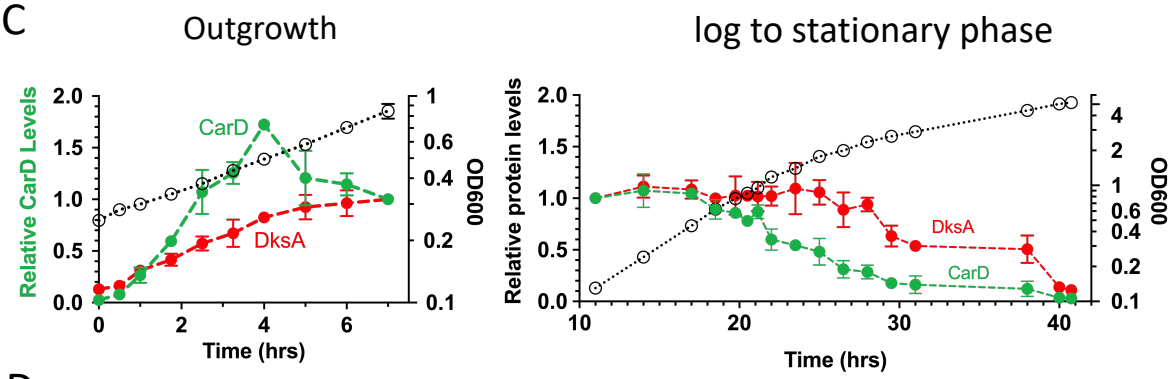
A



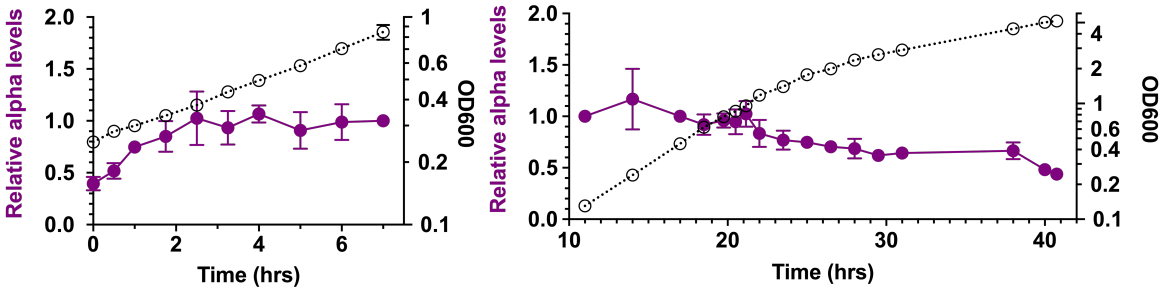
B log to stationary phase



C



D



were ~1.8μM and ~0.2μM respectively at ~10 hours after inoculation (Figure 4.5C, right panel). The levels of RNAP, as indicated by the levels of RNAP-α, changed about only about 2-fold as cells entered stationary phase (Figure 4.5C), suggesting that the differences in CarD and DksA protein levels are not due to global changes in protein expression but rather were specific to DksA and CarD.

Deletion of the *dksA* gene results in widespread changes in gene expression. ppGpp regulates more than 700 genes in *E. coli* within 5 min of induction (1). Deletion of the *dksA* gene or the *relA* gene (the major ppGpp synthetase) have dramatic effects on cell physiology and gene expression profiles (1–3). *R. sphaeroides* cells lacking *dksA* grow poorly under photosynthetic conditions (14). We found here that the DksA protein concentration remains relatively constant during log and early stationary phase under photosynthetic conditions (Figure 4.5C). In addition, DksA and ppGpp regulated rRNA and r-protein transcription with *R. sphaeroides* RNAP, suggesting that together they have a dramatic effect on gene expression in *R. sphaeroides*.

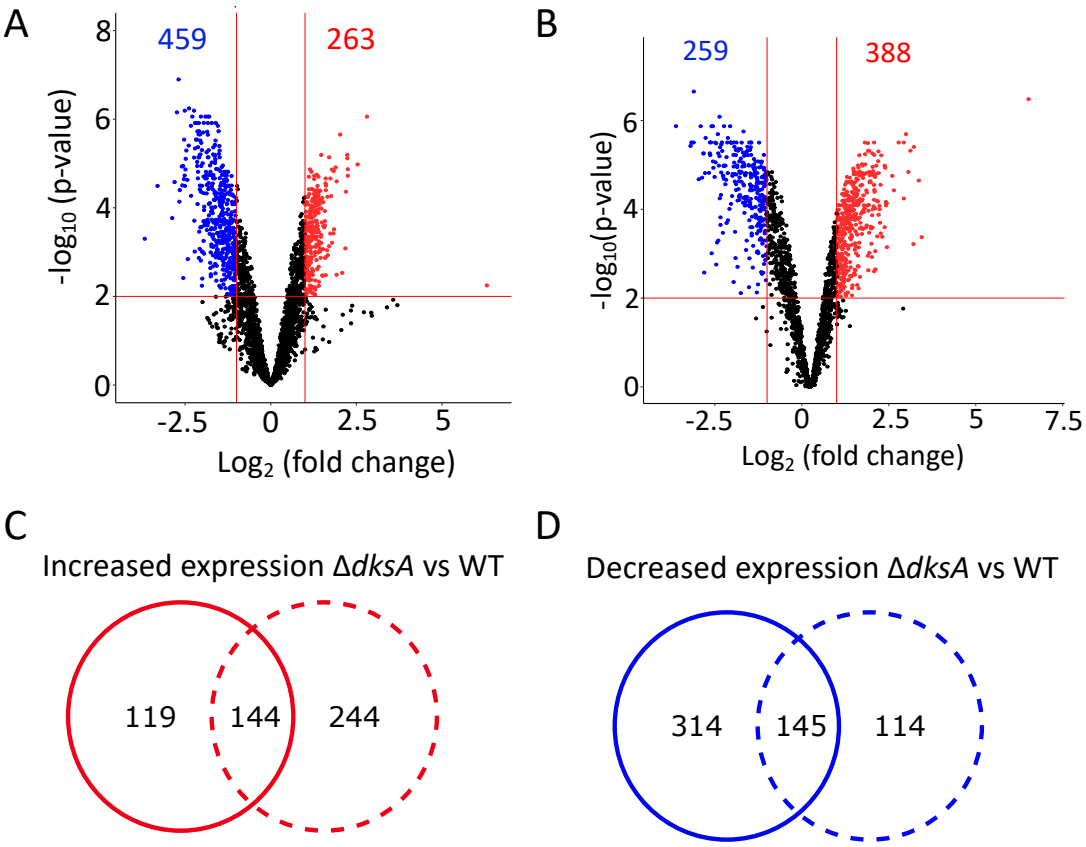
To further examine the genome-wide impact of *dksA* has on gene expression in *R. sphaeroides*, we compared the gene expression pattern in a wild-type *R. sphaeroides* strain to that of a Δ*dksA* mutant strain. Total RNA was extracted from cells grown to mid-log (to an OD_{600} of ~0.5 from a starting OD_{600} of 0.01) in minimal medium under either photosynthetic or aerobic conditions and RNA was sequenced by RNA-seq. Three biological replicates were collected for each strain under each growth condition.

To compare the RNA transcript abundance for genes in each of the two strains, the RNA-seq results are displayed as volcano plots for aerobic and photosynthetic conditions (Figure 4.6A and 4.6B). Genes whose relative RNA transcript abundance changed twofold or more when comparing the Δ*dksA* mutant and wild-type strain (log2 >1 or -1) with a p value < 0.05) were considered statistically significant (see *Materials and Methods*). We identified 722 genes whose transcripts changed in aerobically grown cells (459 decreased and 263 increased) and 647 genes whose transcripts changed in photosynthetically grown cells (259 decreased and 388 increased) when comparing the Δ*dksA* mutant and wild-type strains (Figure 4.6A and 4.5B). Expression of 144 genes increased and 153 genes decreased in both aerobic and photosynthetic growth. A *dksA* deletion in *E. coli* was also found to have a similarly large effect on gene expression in microarray experiments (29).

Many of the genes involved in expression of the light harvesting complex had either reduced expression or increased expression in the Δ*dksA* strain compared to the wild-type strain under photosynthetic growth conditions (Figure 4.6C and 4.D). There were significant numbers of genes whose expression changed at least twofold in both the Δ*dksA* mutant and wild-type strain regardless of growth condition (Figure 4.6C and 4.6D). Genes that encode components of the translation apparatus, such as the r-protein genes for L11 (*rplP*) and S15 (*rpsO*) and tRNA synthetase genes were increased in the Δ*dksA* strain during aerobic and photosynthetic growth (Figure 4.S4), while genes that encode proteins involved in amino acid transport were down in both conditions. In addition, the expression of *rpsM* and *rpsL*, two genes that are directly inhibited by DksA and ppGpp (Figure 4.3A) were upregulated in the Δ*dksA* mutant,

further supporting the evidence (Figure 4.3) that DksA directly inhibits expression of these genes. There was no change in the expression of CarD, which confirms our observation *in vitro* that DksA and ppGpp do not regulate this gene. Our results suggest that DksA plays a significant role in global gene expression, either directly or indirectly, in *R. sphaeroides*.

Figure 4.6: Global gene expression differences between Δ*dksA* and wild-type (WT) *R. sphaeroides* strains grown aerobically (A) or photosynthetically (B). Volcano plots show genes whose expression are statistically significant as a function to the magnitude of change (log_2 fold change). Genes with statistically significant differential expression cluster to the upper right and left quadrants. Additionally, genes that decreased in expression in Δ*dksA* vs WT are blue, while genes that increased in expression are red. Venn diagram showing the number of genes that increase (C) or decrease (D) at least twofold between Δ*dksA* and WT grown either aerobically (solid lines) or photosynthetically (broken lines).


A
459
263
-log10 (p-value)
Log2 (fold change)
B
259
388
-log10(p-value)
Log2 (fold change)
C
Increased expression ΔdksA vs WT
119
144
244
D
Decreased expression ΔdksA vs WT
314
145
114

DISCUSSION

Regulation of rRNA and ribosome synthesis in *E. coli* is controlled by variations in the levels of ppGpp. ppGpp is induced by various kinds of starvation, and together with its cofactor DksA, ppGpp inhibits rRNA synthesis (29). In contrast, *R. sphaeroides* produces ppGpp in response to light downshifts and likely other kinds of nutritional downshifts (9, 30). Studying the regulation of *R. sphaeroides* rRNA transcription has provided us with insights into the transcriptional network that utilizes CarD, ppGpp, and DksA. The bacterium's requirement for rRNA and ribosome synthesis may increase under conditions of photosynthetic growth when vast amounts of photosynthetic membrane proteins must be synthesized. Consistent with these metabolic differences, the signals for induction of ppGpp in α-proteobacterial differ from those in *E. coli* (9, 31). In contrast to *E. coli*, starvation for single amino acids in the alpha proteobacterium Caulobacter did not induce ppGpp (31), and in *R. sphaeroides* a downshift in light intensity induced ppGpp and decreased rRNA synthesis (9).

We have previously shown that *R. sphaeroides* DksA and ppGpp together regulate transcription of *E. coli* promoters by *R. sphaeroides* RNAP in vitro (14). However, we were not previously able to ask whether DksA/ppGpp act directly on *R. sphaeroides* rRNA promoters due to our inability to transcribe these promoters in a purified in vitro transcription system. The finding that CarD is a required cofactor for transcription from rRNA promoters and that it is needed for many other promoters as well in *R. sphaeroides* (20) has now enabled such measurements. Unexpectedly, we found that DksA$_{Rsp}$ alone has little inhibitory effect, but that ppGpp and DksA$_{Rsp}$ together are strong inhibitors of CarD-dependent *R. sphaeroides* rRNA, r-protein, and tRNA

transcription. Our data suggest that ppGpp functions together with $DksA_{Rsp}$ to inhibit expression from promoters involved in synthesis of the translation apparatus.

Our in vivo data indicate a role for DksA in the overall pattern of gene expression under both aerobic and photosynthetic growth conditions. Many of the genes affected differ under the two growth conditions, whereas some are affected under both conditions. These data suggest candidates for genes directly affected by DksA and its cofactor ppGpp. However, further experiments will be required to determine the direct impacts of DksA with its cofactor ppGpp during growth transitions when ppGpp levels would be expected to change.

Regulatory features of *R. sphaeroides* rRNA promoters

While many bacterial phyla appear to encode only DksA (beta and gamma proteobacteria) or only CarD (Actinomycetes, Firmicutes or Thermus/Deinococcus), both factors are present in most alpha and delta proteobacteria. Interestingly, both CarD and DksA/ppGpp function by binding directly to RNAP to alter its interactions with promoter DNA, but they bind to distinct regions of the RNAP and affect its function in different ways. CarD binds to the RNAP beta subunit near the region of sigma that is responsible for -10 recognition, and it also interacts directly with DNA at the upstream edge of the -10 hexamer, at or near the site of nucleation of strand separation, to facilitate promoter melting and to help stabilize open complex formation (17, 19, 32). In contrast, DksA/ppGpp binds to the RNAP secondary channel, allosterically altering the conformation of mobile regions of RNAP involved in isomerization to the open complex.

DksA/ppGpp ultimately inhibits transcription by reducing open complex occupancy at promoters that make intrinsically unstable open complexes (27). Several conformational changes in RNAP induced by binding of the related transcription factor TraR (6) and models for the mechanism of transcription inhibition have recently been reported (33, 34).

A majority of R. sphaeroides promoters lack a -7T at the last position of the -10 element. Having a non-consensus base at this position can significantly affect the kinetics of promoters, particularly, the rate of isomerization. We previously found that CarD activates promoters that lack this critical base. Coincidently, DksA and ppGpp were able to negatively regulate these promoters, suggesting that although CarD was able to stabilize the RNAP-promoter to allow transcription, the inherent unstable nature of having a -7T, makes it so that these promoters can be destabilize by DksA and ppGpp. Here, we propose that these promoters were evolved to lack a -7T both for coordinate activation by CarD and inhibition by DksA/ppGpp. Rather than acting like classical transcription factors, CarD, DksA, and ppGpp regulate promoters by interacting with RNAP and taking advantage of the inherent instability of the RNAP-promoter complex rather utilizing a DNA binding site to confer promoter specific of the effects of the factor.

Our results indicate that under conditions when both CarD and DksA/ppGpp are present in *R. sphaeroides*, DksA/ppGpp can still inhibit the promoter even when it is activated by CarD in order to decrease rRNA transcription. We suggest that the proposed stabilization of the complex by $CarD_{Rsp}$ occurs at an early step in the transcription initiation pathway and that DksA/ppGpp decreases occupancy of a later

step in RPo complex formation in order to prevent transcription initiation. The binding sites of CarD and DksA on RNAP are sufficiently far removed from each other so are unlikely to sterically occlude each other's binding (6, 19, 34), although subtle effects of binding of one factor on binding of the other cannot be ruled out.

We note that *R. sphaeroides* $RNAP_{Rsp}$-$rrnB_{Rsp}$ promoter complexes are extremely sensitive to competitors like heparin, and that unlike the situation with $RNAP_{Mtb}$ and $CarD_{Mtb}$ (35), $CarD_{Rsp}$ is not able to stabilize these complexes to competitor challenge. As a result, we suggest that *R. sphaeroides* rRNA promoters have evolved to be exquisitely balanced: $CarD_{Rsp}$ stabilizes the promoter complexes sufficiently to facilitate propagation of strand separation and RPo formation but not so much as to make the complexes so stable that they are insensitive to inhibition by DksA/ppGpp. Future studies will be needed to test this hypothesis.

SUPPLEMENTAL FIGURES

Figure 4.S1: Comparison of the amino acid sequence of portions of the *E. coli* (*Eco*) and *R. sphaeroides* (*Rsp*) RNAP subunits (A) or of full length *E. coli* or *R. sphaeroides* DksA (B) containing residues involved in ppGpp binding. Residues that make up the ppGpp binding site 1 are red and the residues for site 2 are blue. Alignment of the amino acids residues for the RNAP ω subunit (A), the β' subunit (B) and DksA (C). Alignments were made using Clustal Omega. Residue numbers are indicated at the end of each line. Conserved or substituted residues for the ppGpp site 1 residues in ω and β' are shown in red and boxed (identified genetically and biochemically (5)), red and underlined (identified structurally (23)) and red (identified structurally (6)). Conserved or substituted residues for the ppGpp site 2 in β' and DksA are shown in blue and boxed (identified genetically and biochemically (3)) and blue and underlined (identified structurally (6)). The ppGpp binding site 2 *E. coli* β' Y626 (green) is not conserved in *R. sphaeroides* β'.

A Omega

```
Eco  MARVTVQDAVEKIGNRFDLVLVAARRARQMQVGGKDPLVPEENDKTTVIALREIEEGLIN  60
Rsp  MARVTVEDCVDKVPNRFELVMLAAHRAREIASGSSL-TIDRDNDKNPVVALREIAEETQS  59
     *****:*.*:*: ***:**::**:***::  *..   : .:***. *:***** *   .
```

B β’

```
Eco  GPYLRLHQCGLPKKMALELFKPFIYGKLELRGLATTIKAAKKMVEREEAVVWDILDEVIR  417
Rsp  GPELKLHQCGLPKKMALELFKPFIYSRLEAKGLSSTVKQAKKLVEKERPEVWDILDEVIR  420
     ** *:*********************.:** :**::*:* ***:**:*.  *********

Eco  LGKKAISKMLNTCYRILGLKPTVIFADQIMYTGFAYAARSGASVGIDDMVIPEKKHEIIS  655
Rsp  LRKKDVQNVIDTVYRYCGQKESVIFCDQIMGMGFREAFKAGISFGKDDMLIPDTKWPIVN  656
     * ** :.:::* **  * * :***.****  **  * ::* *.* ***:**:.*  *:.

Eco  EAEAEVAEIQEQFQSGLVTAGERYNKVIDIWAAANDRVSKAMMDNLQTETVINRDGQEEK  715
Rsp  EVRDQVKEFEQQYMDGLITQGEKYNKVVDAWSKCSDKVAGEMMAEISAVRYD----DAGA  712
     *.. :* *:::*: .**:* **:*****:* *: ..*:*:  ** ::.:        :
```

C DksA

```
Eco  -------MQEGQNRKTSSLSILAIAGVEPYQEKPGEEYMNEAQLAHFRRILEAWRNQLRD  53
Rsp  MTVNHISEPAGLQQRAAA-MKAEIFLPEDYRPAENEPFMNERQLEYFRRKLLNWKQELLD  59
               * ::::::     *   * *:   .* :*** ** :*** *  *:::* *

Eco  EVDRTVTHMQDEAANFPDPVDRAAQEEEFSLELRNRDRERKLIKKIEKTLKKVEDEDFGY  113
Rsp  QSAETIEGLQESGRNVPDIADRASEETDRALELRTRDRQRKLVAKIDAALRRIEAGEYGY  119
     :  .*:  :*:.. *.** .***::* : :****.***:***: **: :*:::*  ::**

Eco  CESCGVEIGIRRLEARPTADLCIDCKTLAEIREKQMAG-    151
Rsp  CEVTGEPISLKRLDARPIATMTLEAQERHERRERVHRDE   158
     **  *  *.::**:*** * : ::.:   * **:   .
```

Figure 4.S2: *R. sphaeroides* RNAP binds radiolabeled ppGpp in DRaCALA assay. ^{32}P-ppGpp was incubated with either 2μM *E. coli* or *R. sphaeroides* RNAP or in buffer only. The reaction was spotted onto nitrocellulose filters. Duplicate filters are shown for each RNAP and the no RNAP control.

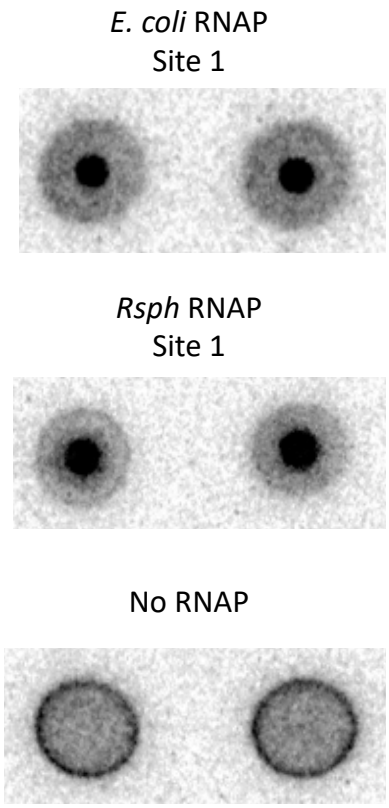


Figure 4.S3: Relative levels of DksA$_{Rsp}$ and CarD and CFU/mL in aerobically (A) or photosynthetically (B) grown *R. sphaeroides* cultures. CFU/mL and protein levels were determined in samples from stationary phase cultures and plotted as a function of time. Error bars are from 2 experiments.

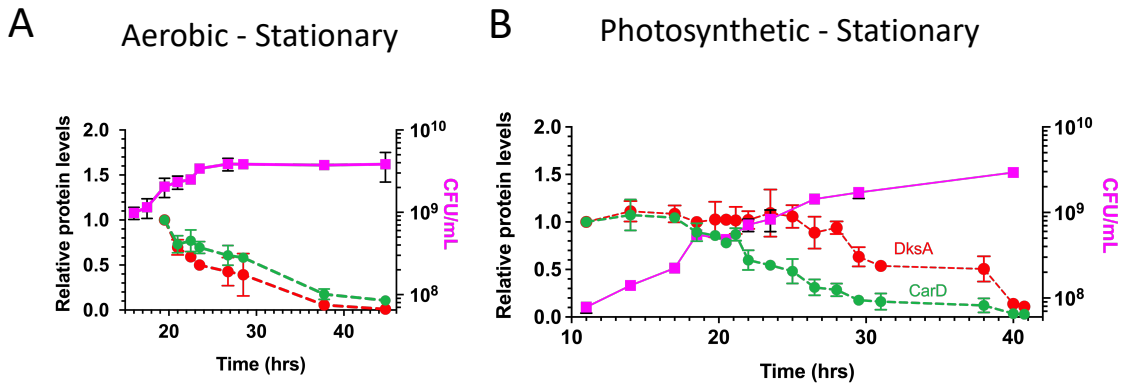
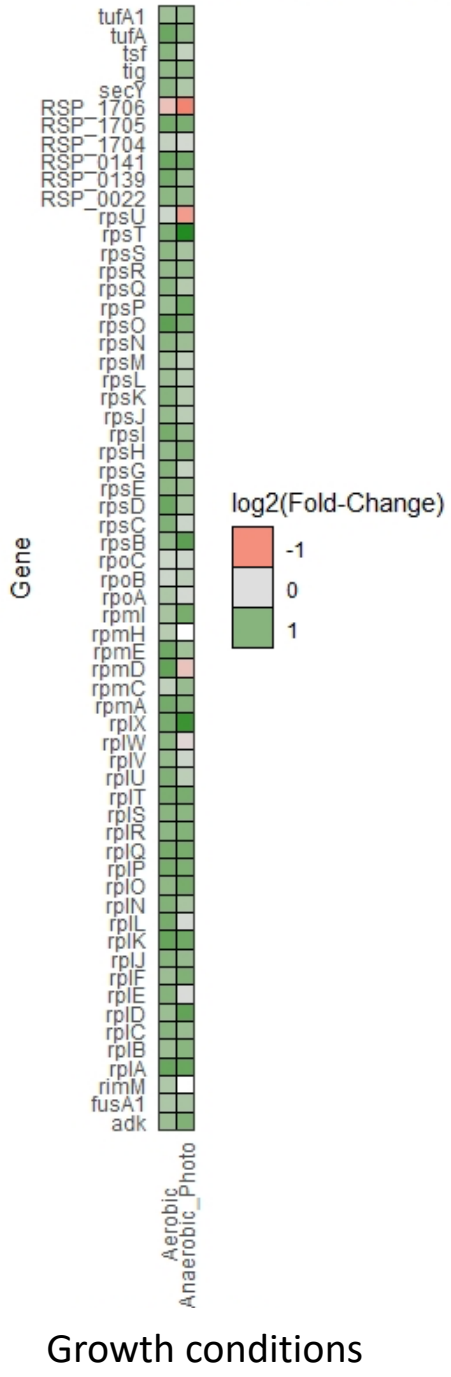


Figure 4.S4: Heat-map showing expression changes for gene involved in translation that were upregulated in the *R. sphaeroides* Δ*dksA* vs wild-type strains. Genes that were up-regulated are green (twofold), no change (grey) and down-regulated (red).



MATERIALS AND METHODS

Bacterial strains and growth conditions. Bacterial strains are listed in Table 4.S1**.** *E. coli* strains were grown in either LB Lennox media or LB agar plates at 37°C supplemented with kanamycin (30 µg/mL) or chloramphenicol (20 µg/mL) when required. *R. sphaeroides* 2.4.1 strains were grown in succinate-based minimal medium (SIS) (36) at 30°C with kanamycin 25 µg/mL or sucrose (10%) when appropriate. For liquid growth conditions, *R. sphaeroides* 2.4.1 was grown in 500 mL cultures with constant sparging with 69% N_2, 30% O_2 and 1% CO_2 at 30°C for aerobic growth and 98.5% N_2, 0.5% O_2 and 1% CO_2 at 30°C in front of an incandescent light with a light intensity of 10 W/m^2 for photosynthetic growth. For aerobic and photosynthetic growth curves, sample cultures were removed by positive pressure from the main culture and optical density was determined at OD_{600} and viable cell counts were determined by plating on solid minimal media.

Purification of *R. sphaeroides* RNAP Holoenzyme and SUMO-tagged CarD wild type and variant proteins. *R. sphaeroides* holoenzyme was purified using Ni-agarose affinity chromatography from aerobically grown cells containing the chromosomally encoded His10-β' (RLG14765) as previously described (20) and stored in buffer containing 5 mM Tris-Cl [pH 8.0], 50 mM KCl, 5 mM $MgCl_2$, 0.1 mM EDTA and 50% glycerol at -20°C.

The $carD_{Rsp}$ gene was cloned into the pETSUMO vector for overexpression and purification from *E. coli* BL21 (RLG15456) using its IPTG-inducible T7 promoter as described (20). Untagged purified CarD was stored in storage buffer (20 mM Tris-HCl [pH 8.0], 100 mM NaCl, 2 mM DTT, 50% glycerol). Protein concentration was assayed

by the Bradford assay using BSA as a standard. For in vitro transcription assays, CarD was diluted into storage buffer.

Purification of DksA$_{Rsp}$. DksA$_{Rsp}$ was purified as previously described (14). In short: BL21 (DE3) cells transformed with pRLG10348 (expresses N-terminal His$_6$-thrombin-tagged DksA$_{Rsp}$ under the expression of an IPTG inducible T7 promoter) were grown at 37°C to an OD$_{600}$ of ~0.5-0.6 in LB media supplemented with kanamycin. Cells were induced with the addition of 1mM IPTG for 3 hrs and harvested by centrifugation. Cells were resuspended into 20mL of lysis buffer (10mM Tris-Cl pH 8, 300mM NaCl, 5% glycerol (vol/vol), 2mM β-mercaptoethanol, 5mM imidazole) supplemented with phenylmethylsulfonyl fluoride (23 µg/ml), lysozyme (0.26 µg/ml) and sodium deoxycholate (0.1%) and lysed by sonication. Next, the cells were passed through a nickel resin (Ni-NTA) column pre-equilibrated with lysis buffer + 5mM imidazole, washed with lysis buffer + 10mM imidazole and His-tagged DksA was eluted into lysis buffer + 0.3M imidazole. The eluted His$_6$-thrombin tagged DksA was dialyzed overnight at 4°C into thrombin cleavage buffer (20mM Tris-HCl [pH8], 150mM NaCl and 5mM $CaCl_2$), after which the protein concentration was determined by the Bradford assay using bovine serum albumin as a standard. The His$_6$-thrombin tag was removed from DksA$_{Rsp}$ by incubating with biotinylated-thrombin protease (Millipore-Sigma) at a ratio of 0.75units/mg (thrombin/tagged DksA) at room temperature for 3.5 hrs. The liberated His-tag and biotinylated thrombin were removed by passage through a combined Ni-NTA and streptavidin agarose column (~0.3mL Ni-NTA/0.12mL streptavidin packed) pre-equilibrated with thrombin cleavage buffer. The flow-through containing DksA was dialyzed overnight against storage buffer (20 mM Tris-HCl [pH 8.0], 100 mM NaCl,

0.1mM DTT, 50% glycerol). Protein concentration was assayed by the Bradford assay. Proteins were stored at -20°C. For in vitro transcription assays, DksA was diluted into storage buffer to the desired concentration.

In vitro transcription. Multiple round transcription assays were carried out using supercoiled plasmid templates containing the promoters listed in Table 4.2 using *R. sphaeroides* RNAP (20 nM) in 25μL reactions. Template DNA (5 pmol or 400 ng) was incubated in transcription buffer (20 mM Tris-HCl [pH 7.9], 10 mM MgCl2, 1 mM DTT, 0.1 mg/ml BSA, and either 170 or 100 mM NaCl, as indicated, containing NTPs (200 μM ATP, GTP, CTP, 10μM UTP, and 1.2 μCi α-^{32}P UTP [Perkin Elmer]) with CarD or storage buffer as indicated and the reactions were incubated at room temperature for 10 min. For reactions with DksA$_{Rsp}$ and ppGpp, DksA and/or ppGpp were added at the indicated concentration after the addition of CarD and reactions were incubated at room temperature for 10 min. Reactions without DksA, storage buffer was added, and reactions without ppGpp, the reactions were supplied water.

Transcription was initiated by addition of RNAP and reactions were incubated at 30°C for 20 min. The reactions were terminated by addition of 25 μL water-saturated phenol, aqueous and phenol phases were separated by brief centrifugation, and 16 μL of the aqueous phase containing the transcript was transferred to 16μL 2X loading solution (8M urea, 10 mM EDTA, 2X TBE, 1% SDS and 0.02% bromophenol blue). Transcripts were denatured by heating the samples at 95°C for 1 min, resolved on 6% polyacrylamide 7M urea gels, imaged using a Typhoon phosphorimager, and quantified with ImageQuant 5.2 software.

Visualizing Binding of ^{32}P-ppGpp by differential radial capillary action of ligand assay (DRaCALA). Binding of ^{32}P-ppGpp to *E. coli* or *R. sphaeroides* holo-RNAP (in buffer: 10 mM Tris-Cl [pH 8.0], 100 mM KCl, 10 mM MgCl2, 0.1 mM EDTA) was determined by DRaCALA (37) with a modified protocol for ^{32}P-ppGpp purification (3). ^{32}P-ppGpp (a gift from Angela Myers, Gourse lab) was synthesized using a crude ribosome extract containing functional RelA to catalyze the transfer of the γ-phosphate from γ-^{32}P-ATP (PerkinElmer) onto GDP (Amresco). ^{32}P-ppGpp was then purified by PEI cellulose thin layer chromatography in 1.5 M KH_2PO4 [pH 3.4], eluted from the TLC plate in 4 M LiCl, precipitated, and then stored in aliquots at -80°C as. To visualize binding of ^{32}P-ppGpp to RNAP, 15 µL DRaCALA binding reactions containing ~5 nM ^{32}P-ppGpp and 2.0 µM RNAP in DRaCALA buffer (15mM Tris-Cl [pH8], 30mM KCl, 10mM MgCl2 and 1mM DTT). The binding assays were incubated at room temperature for 12 min and duplicate 4.0 µL aliquots of each sample were spotted onto dry Protran BA85 nitrocellulose membrane filters. Samples on the membranes were dried, covered with saran wrap and visualized on a phosphoimager.

Western blot analysis. *R. sphaeroides* cells were harvested from aerobically and photosynthetically grown cultures at the OD_{600} indicated. Western blot was performed on (20). For the aerobic growth experiment: separate cultures were grown for each of the growth phases analyzed: outgrowth, log and stationary phase. For outgrowth, saturated culture (OD600 ~3.5) was diluted (1:12.5) into fresh media to a starting OD600 of 0.25. For the log phase culture, an aliquot of a stationary phase culture was diluted 1:50 to a starting OD of 0.02 into fresh medium at time 0, and the culture was allowed to grow for 10 hours to an OD of ~0.2 before sampling, and samples were then

collected at intervals from 10 to 18 hours, during which time the culture grew exponentially. For the stationary culture experiment, the culture was inoculated from a stationary phase culture to a starting OD of ~0.0125 and was sampled at intervals from 15 to 45 hours, during which it entered and continued in stationary phase.

For photosynthetically growth cells, for monitoring outgrowth a saturated culture (OD_{600} ~5) was diluted ~1:20 into fresh media to a starting OD_{600} of 0.25 at time 0, and samples were removed at the indicates times. For analysis of cells during log phase and stationary phase the same culture was used. An aliquot of a stationary phase culture was diluted 1:100 to a starting OD of 0.015 into fresh medium at time 0, and cultures were allowed to grow for 11 hours to an OD of ~0.12 before sampling. Samples were collected at intervals from 11 to 41 hours, during which the culture grew exponentially and into stationary phase.

For each sample collected (at the indicated OD_{600}) 2OD equivalent of cells were pelleted, resuspended in 200 µL of 1X Laemmli loading buffer (75 mM Tris-HCl [pH 8], 2% SDS, 0.5 M β-mercaptoethanol, 10% glycerol, 0.2% Triton X100, 0.02% bromophenol blue) and heated at 95°C for 10 min. Aliquots of these whole cell lysates (10 µL/lane or 0.1 OD_{600} equivalent) and purified standards were separated on 12% NuPAGE Bis-Tris gels (Invitrogen) in MES buffer, and proteins were transferred onto a nitrocellulose membrane (GE Healthcare). The membrane was incubated for 2 hrs with the primary antibody against CarD (20) or for 2hrs with the primary antibody for $DksA_{Rsp}$ (14) at a 1:10,000 dilution at room temperature in blocking buffer (phosphate buffered saline [pH 7.4], 0.1% Tween20 (PBST) and 5% milk). The membrane was subsequently washed 3 times with PBST and then incubated with secondary antibody (goat anti-rabbit

IgG horseradish peroxidase conjugate [Promega]; added at 1:10,000) for 1 hr in blocking buffer. The membrane was then washed 3 times with PBST, incubated in SuperSignal West Dura Extended Duration Substrate (Thermo Scientific), and bands were detected with the LI-COR Odyssey Fc Imaging System. The intensity of each band was quantified using Image Studio Lite software (LI-COR). Polyclonal antibody to the α subunit of *R. sphaeroides* RNAP (WI515) was a gift from Karen Wassarman (University of Wisconsin-Madison). CarD and DksA protein concentrations in the whole cell lysate were interpolated from the standard curve.

RNA Isolation and Sequencing. *R. sphaeroides strains for RNA isolation* were grown either photosynthetically or aerobically in 500 mL cultures. At an OD_{600} of ~0.5, 44 mL of cells were harvested and combined with 6 mL of ice cold stop solution (95% EtOH and 5% acid phenol:chloroform (5:1, pH 4.5)) and pelleted by centrifugation at 4°C for 10 minutes at 6,000 x *g*. Cell pellets were resuspended in 2 mL of lysis solution (2% sodium dodecyl sulfate (SDS), 16mM EDTA) and heated at 65°C for 5 minutes. Nucleic acids were extracted using 2 mL acid phenol:chloroform:IAA (pH 4.5, ThermoFisher Scientific) heated to 65°C, mixed by inversion, incubated at 65°C for 5 minutes, and centrifuged at room temperature for 7 minutes at 20,000 x *g*. The aqueous phase was removed, and extraction was performed twice more. Next, 2 mL of chloroform was added, mixed, and centrifuged once more at room temperature at for 7 minutes at 20,000 x *g*. The aqueous layer was taken and mixed with 1/10 volume of 3M sodium acetate and equal volume of isopropanol and incubated at -20°C for >1 hour to precipitate nucleic acids, after which, RNA was pelleted by centrifugation at 4°C for 30 minutes at 16,000 x *g*. The pellet was washed twice with 1 mL of 75% EtOH (prepared

with RNase free water). Once all residual ethanol had been removed by evaporation, the pellet was resuspended in 85 µL of RNase free water. Samples were treated with RNase-free DNase (10 µL 10x DNase buffer, 2 µL DNase, 3 µL RNasin (Promega)) and then purified with a RNeasy kit (Qiagen).

RNA-seq library preparation and sequencing was performed at the University of Wisconsin-Madison Biotechnology Center. Libraries for sequencing were created using the Illumina TruSeq Stranded Total RNA kit (Illumina) following the standard protocol. RNA-seq libraries were sequenced on an Illumina NextSeq in 2x151 reads using the standard protocol. All FASTQ files were processed through the same pipeline. Reads were trimmed using Trimmomatic version **0.3** (38) with the default settings except for a HEADCROP of 5, LEADING of 3, TRAILING of 3, SLIDINGWINDOW of 3:30, and MINLEN of 36. After trimming the reads were aligned to the *Rhodobacter sphaeroides 2.4.1* genome sequence (GenBank accession ASM1290v2) using Bowtie2 version 2.2.2 (39) with default settings except the number of mismatches was set to 1. Aligned reads were mapped to gene locations using HTSeq version 0.6.0 (40) using default settings except that the “reverse” strandedness argument was used. To remove low expressed genes, genes that lacked read counts of at least 20 cpm in at least 3 samples were removed from further analysis. The R package edgeR (version 3.30.3) (41) was used to identify significantly differentially expressed genes from pairwise analyses, using false discovery rate (FDR) (42) less than 0.05 as a significance threshold. Genes with a fold change of at least 2 between the two pairwise conditions were considered of highest biological importance. Raw sequencing reads were normalized using the reads per kilobase per million mapped reads (RPKM).

Table 4.1: Strain and Plasmid used in study

Strains used in this study

Strains	Genotype	Source
BL21 (DE3) pLysS	*E. coli* str. B F^- *ompT gal dcm lon hsdS*$_B$($r_B^-m_B^-$) λ(DE3 [*lacI lacUV5-T7p07 ind1 sam7 nin5*]) [*malB*$^+$]$_{K\text{-}12}$(λ^S) pLysS[*T7p20 ori*$_{p15A}$](Cm^R)	Novagen
BL21 (DE3)	*E. coli str. B F$^-$ ompT gal dcm lon hsdS$_B$($r_B^-m_B^-$) λ(DE3 [lacI lacUV5-T7p07 ind1 sam7 nin5]) [malB$^+$]$_{K\text{-}12}$(λ^S)*	Novagen
E. coli XL1-Blue	Δ(*mcrA*)*183* Δ(*mcrCB-hsdSMR-mrr*)*173 endA1 supE44 thi-1 recA1* gyrA96 relA1 lac *[F'* proAB lacI^{q}ZΔM15 *Tn*10 *(Tet^r)]*	Stratagene
RLG14946	*Rhodobacter sphaeroides* 2.4.1 (Wild type lab strain)	
RLG14947	RLG14946 with a scarless chromosomal deletion of *dksA* (*RSP_*2654)	(14)
RLG14765	RLG14946 with a C-terminal thrombin-His$_{10}$ chromosomally tagged β' (*rpoC*)	(20)
RLG15456	BL21 (DE3) pLysS transformed with pETSUMO His10-CarD$_{Rsp}$	(20)

Plasmids used in this study

Plasmids	Description	Source
pRLG770	Transcription vector	(43)
pETSUMO	pET derived plasmid that encodes a Ulp1 cleavable His$_{10}$SUMO tag. ORFs are cloned in frame with the coding sequence of the SUMO tag.	Invitrogen
p15433	R. sphaeroides *carD* cloned into pETSUMO	(20)
p10348	pET33 His6-HMK DksA$_{Rsp}$ (RSP_2654)	(14)
p10349	BL21DE3 transformed with p10348	(14)

Table 4.2: Plasmids used for *in vitro* transcription

Plasmids	Promoter[a]	Endpoints	Source
p14564	*R. sphaeroides rrnB*	-230 to +50	(20)
p15221	*R. sphaeroides carD*	-161 to +50	This study
p15224	*R. sphaeroides rpsL*	-132 to +65	(20)
p15228	*R. sphaeroides rpsM*	-156 to +50	(20)
p15227	*R. sphaeroides tRNA-trp* (*RSP_*4298)	-106 to +56	(20)
p15472	*R. sphaeroides tRNA-val* (*RSP_*4309)	-79 to +47	(20)

[a] Promoters are named with the first gene in the operon

REFERENCES:

1. Sanchez-Vazquez P, Dewey CN, Kitten N, Ross W, Gourse RL. 2019. Genome-wide effects on Escherichia coli transcription from ppGpp binding to its two sites on RNA polymerase. PNAS 201819682.
2. Paul BJ, Barker MM, Ross W, Schneider DA, Webb C, Foster JW, Gourse RL. 2004. DksA: A Critical Component of the Transcription Initiation Machinery that Potentiates the Regulation of rRNA Promoters by ppGpp and the Initiating NTP. Cell 118:311–322.
3. Ross W, Sanchez-Vazquez P, Chen AY, Lee J-H, Burgos HL, Gourse RL. 2016. ppGpp Binding to a Site at the RNAP-DksA Interface Accounts for Its Dramatic Effects on Transcription Initiation during the Stringent Response. Molecular Cell 62:811–823.
4. Lemke JJ, Sanchez-Vazquez P, Burgos HL, Hedberg G, Ross W, Gourse RL. 2011. Direct regulation of Escherichia coli ribosomal protein promoters by the transcription factors ppGpp and DksA. PNAS 108:5712–5717.
5. Ross W, Vrentas CE, Sanchez-Vazquez P, Gaal T, Gourse RL. 2013. The Magic Spot: A ppGpp Binding Site on E. coli RNA Polymerase Responsible for Regulation of Transcription Initiation. Molecular Cell 50:420–429.
6. Molodtsov V, Sineva E, Zhang L, Huang X, Cashel M, Ades SE, Murakami KS. 2018. Allosteric Effector ppGpp Potentiates the Inhibition of Transcript Initiation by DksA. Molecular Cell 69:828-839.e5.
7. Rutherford ST, Lemke JJ, Vrentas CE, Gaal T, Ross W, Gourse RL. 2007. Effects of DksA, GreA, and GreB on Transcription Initiation: Insights into the Mechanisms of Factors that Bind in the Secondary Channel of RNA Polymerase. Journal of Molecular Biology 366:1243–1257.
8. Acosta R, Lueking DR. 1987. Stringency in the absence of ppGpp accumulation in Rhodobacter sphaeroides. J Bacteriol 169:908–912.
9. Eccleston Jr. ED, Gray ED. 1973. Variations in ppGpp levels in Rhodopseudomonas spheroides during adaptation to decreased light intensity. Biochemical and Biophysical Research Communications 54:1370–1376.
10. Mackenzie C, Eraso JM, Choudhary M, Roh JH, Zeng X, Bruscella P, Puskás Á, Kaplan S. 2007. Postgenomic Adventures with Rhodobacter sphaeroides. Annual Review of Microbiology 61:283–307.
11. Kaplan S. 2002. Photosynthesis genes and their expression in Rhodobacter sphaeroides 2.4.1: a tribute to my students and associates. Photosynthesis Research 73:95–108.
12. Kiley PJ, Kaplan S. 1988. Molecular genetics of photosynthetic membrane biosynthesis in Rhodobacter sphaeroides. Microbiol Rev 52:50–69.
13. Pemberton JM, Horne IM, McEwan AG. 1998. Regulation of photosynthetic gene expression in purple bacteria. Microbiology 144:267–278.

14. Lennon CW, Lemmer KC, Irons JL, Sellman MI, Donohue TJ, Gourse RL, Ross W. 2014. A Rhodobacter sphaeroides Protein Mechanistically Similar to Escherichia coli DksA Regulates Photosynthetic Growth. mBio 5.
15. Furman R, Biswas T, Danhart EM, Foster MP, Tsodikov OV, Artsimovitch I. 2013. DksA2, a zinc-independent structural analog of the transcription factor DksA. FEBS Letters 587:614–619.
16. Blaby-Haas CE, Furman R, Rodionov DA, Artsimovitch I, de Crécy-Lagard V. 2011. Role of a Zn-independent DksA in Zn homeostasis and stringent response. Molecular Microbiology 79:700–715.
17. Srivastava DB, Leon K, Osmundson J, Garner AL, Weiss LA, Westblade LF, Glickman MS, Landick R, Darst SA, Stallings CL, Campbell EA. 2013. Structure and function of CarD, an essential mycobacterial transcription factor. PNAS 110:12619–12624.
18. Rammohan J, Ruiz Manzano A, Garner AL, Stallings CL, Galburt EA. 2015. CarD stabilizes mycobacterial open complexes via a two-tiered kinetic mechanism. Nucleic Acids Res 43:3272–3285.
19. Bae B, Chen J, Davis E, Leon K, Darst SA, Campbell EA. 2015. CarD uses a minor groove wedge mechanism to stabilize the RNA polymerase open promoter complex. eLife Sciences e08505.
20. Henry KK, Ross W, Myers KS, Lemmer KC, Vera JM, Landick R, Donohue TJ, Gourse RL. 2020. A majority of Rhodobacter sphaeroides promoters lack a crucial RNA polymerase recognition feature, enabling coordinated transcription activation. PNAS 117:29658–29668.
21. Myers KS, Vera JM, Lemmer KC, Linz AM, Landick R, Noguera DR, Donohue TJ. 2020. Genome-Wide Identification of Transcription Start Sites in Two Alphaproteobacteria, Rhodobacter sphaeroides 2.4.1 and Novosphingobium aromaticivorans DSM 12444. Microbiol Resour Announc 9.
22. Gopalkrishnan S, Ross W, Chen AY, Gourse RL. 2017. TraR directly regulates transcription initiation by mimicking the combined effects of the global regulators DksA and ppGpp. PNAS 114:E5539–E5548.
23. Zuo Y, Wang Y, Steitz TA. 2013. The Mechanism of E. coli RNA Polymerase Regulation by ppGpp Is Suggested by the Structure of their Complex. Molecular Cell 50:430–436.
24. Burger BT, Imam S, Scarborough MJ, Noguera DR, Donohue TJ. 2017. Combining Genome-Scale Experimental and Computational Methods To Identify Essential Genes in *Rhodobacter sphaeroides*. mSystems 2:e00015-17.
25. Murray HD, Schneider DA, Gourse RL. 2003. Control of rRNA Expression by Small Molecules Is Dynamic and Nonredundant. Molecular Cell 12:125–134.
26. Rutherford ST, Villers CL, Lee J-H, Ross W, Gourse RL. 2009. Allosteric control of Escherichia coli rRNA promoter complexes by DksA. Genes Dev 23:236–248.

27. Lemmer KC, Alberge F, Myers KS, Dohnalkova AC, Schaub RE, Lenz JD, Imam S, Dillard JP, Noguera DR, Donohue TJ. 2020. The NtrYX Two-Component System Regulates the Bacterial Cell Envelope. mBio 11.
28. Gourse RL, Chen AY, Gopalkrishnan S, Sanchez-Vazquez P, Myers A, Ross W. 2018. Transcriptional Responses to ppGpp and DksA. Annual Review of Microbiology 72:163–184.
29. Campbell TB, Lueking DR. 1983. Light-mediated regulation of phospholipid synthesis in Rhodopseudomonas sphaeroides. J Bacteriol 155:806–816.
30. Boutte CC, Crosson S. 2011. The complex logic of stringent response regulation in Caulobacter crescentus: starvation signalling in an oligotrophic environment. Molecular Microbiology 80:695–714.
31. Hubin EA, Fay A, Xu C, Bean JM, Saecker RM, Glickman MS, Darst SA, Campbell EA. 2017. Structure and function of the mycobacterial transcription initiation complex with the essential regulator RbpA. eLife 6:e22520.
32. Chen J, Chiu C, Gopalkrishnan S, Chen AY, Olinares PDB, Saecker RM, Winkelman JT, Maloney MF, Chait BT, Ross W, Gourse RL, Campbell EA, Darst SA. 2020. Stepwise Promoter Melting by Bacterial RNA Polymerase. Molecular Cell https://doi.org/10.1016/j.molcel.2020.02.017.
33. Chen J, Gopalkrishnan S, Chiu C, Chen AY, Campbell EA, Gourse RL, Ross W, Darst SA. 2019. E. coli TraR allosterically regulates transcription initiation by altering RNA polymerase conformation. eLife 8:e49375.
34. Davis E, Chen J, Leon K, Darst SA, Campbell EA. 2015. Mycobacterial RNA polymerase forms unstable open promoter complexes that are stabilized by CarD. Nucleic Acids Res 43:433–445.
35. Sistrom WR. 1960. A Requirement for Sodium in the Growth of Rhodopseudomonas spheroides. Microbiology 22:778–785.
36. Roelofs KG, Wang J, Sintim HO, Lee VT. 2011. Differential radial capillary action of ligand assay for high-throughput detection of protein-metabolite interactions. PNAS 108:15528–15533.
37. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120.
38. Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. 4. Nature Methods 9:357–359.
39. Anders S, Pyl PT, Huber W. 2015. HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics 31:166–169.
40. Robinson MD, McCarthy DJ, Smyth GK. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26:139–140.

41. Benjamini Y, Hochberg Y. 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society Series B (Methodological) 57:289–300.
42. Ross W, Thompson JF, Newlands JT, Gourse RL. 1990. E.coli Fis protein activates ribosomal RNA transcription in vitro and in vivo. EMBO J 9:3733–3742.

CHAPTER FIVE

Conclusions and Future Direction

In this chapter, I will briefly summarize the major findings of this thesis and suggest some subject for further investigation.

CONCLUSIONS

I undertook this thesis project to determine how rRNA promoters are regulated in the photosynthetic bacterium, *Rhodobacter sphaeroides*. In *E. coli*, rRNA synthesis is regulated by the transcription factor DksA and the alarmone ppGpp (1, 2). Our lab, in collaboration with the Donohue lab, previously characterized an *R. sphaeroides* DksA homolog that was functionally similar to the *E. coli* DksA homolog (3). *R. sphaeroides* cells harboring a chromosomal deletion of the *dksA* gene grew poorly under photosynthetic conditions. Previous work demonstrated that *R. sphaeroides* produces ppGpp in response to light downshifts rather than in response to amino acid starvation as in *E. coli* (4, 5). Together, these observations suggested that rRNA regulation is tied to photosynthetic growth in *R. sphaeroides*. To test whether DksA regulates transcription, Chris Lennon, a former graduate in our lab, tried to measure transcription from *R. sphaeroides* rRNA promoters using *R. sphaeroides* RNA polymerase ($RNAP_{Rsp}$), but he detected little or no transcription *in vitro.*

In Chapter 2, I addressed why transcription from *R. sphaeroides* rRNA promoters was so weak using *R. sphaeroides* RNAP. I designed a reliable method for purifying *R. sphaeroides* RNAP that involved creating a *R. sphaeroides* strain carrying a C-terminally his-tagged β' subunit on the chromosome and purifying the resulting RNAP by nickel affinity. Although the purified $RNAP_{Rsp}$ was catalytically active on some test promoters from *E. coli* (*RNA 1*, *lacUV5*, and *rrnB* P1), it transcribed *Rsph* rRNA promoters very inefficiently, whereas in most moderate to fast growing bacterial species, rRNA promoters are among the most active promoters in the cell. This was puzzling.

We got the first hint of an explanation when we examined the sequences of the three *R. sphaeroides* rRNA promoters and found that the last position (-7) of the -10 element was an adenine rather than a thymine. In other bacteria, it has been shown using biochemical, structural, and genetic studies that this position is crucial for promoter recognition by RNAP. I found that $RNAP_{Rsp}$ could transcribe *R. sphaeroides* rRNA promoters *in vitro* when a suboptimal position in the -10 element was mutated to the *E. coli* consensus sequence. I also found that many R. sphaeroides promoters, such as the ones for tRNA and r-proteins, lacked a -7T. Genome-wide transcription start site mapping studies followed by bioinformatic analysis showed that a majority of *R. sphaeroides* promoters lack -7T (6, 7), whereas in most non-alpha proteobacteria, this position is a T in more than 90% of promoters. These results suggested to me that transcription from *R. sphaeroides* promoters might require an additional factor that compensates for the absence of the non-consensus base in the -10 element.

I found that the *R. sphaeroides* CarD homolog, a transcription factor that activated rRNA transcription in *Mycobacterium tuberculosis*, activated transcription of *R. sphaeroides* promoters with $RNAP_{Rsp}$ (8). CarD activated transcription from rRNA promoters along with some other promoters involved in translation, transcription, and general cell physiology (Chapter 2). Amino acid substitutions in CarD that are essential for CarD function in other species are required for $CarD_{Rsph}$ to activate *R. sphaeroides* rRNA promoters. We also found that promoters that are dependent on CarD for activation also lacked -7T, suggesting a strong sequence correlation among promoters requiring CarD for activity. CarD protein levels are constant during log growth when *R. sphaeroides* is grown under aerobic conditions and its levels decrease when cells enter

stationary phase. From these observations, I hypothesized that *R. sphaeroides* may have evolved with promoters under coordinate control of CarD, and thus the expression of these genes might be coordinated with the levels of CarD.

In Chapter 3, I found that CarD can also function as a negative regulator of transcription, inhibiting transcription from its own promoter. Basal activity from the *carD* promoter was strong, even though the promoter lacked a -7T. I correlated this activity with the presence of an extended -10 sequence and not with the identity of the base at -7 as I had observed in the other *R. sphaeroides* promoters I had studied. I found that CarD inhibited transcription from the *carD* promoter ~4-fold, and the extent of inhibition was reduced by substitutions in amino acid residues important for CarD's ability to activate transcription. Since key residues that play a role in the ability of CarD to activate also contribute to its ability to inhibit transcription, I hypothesized CarD made similar contacts to both kinds of promoters, but the kinetic properties of the two classes of promoters differed. Because the *carD* promoter produced significantly more abortive RNAs when inhibited by CarD, I surmised that its mechanism of inhibition might be by preventing RNAP from escaping the promoter. Although CarD had previously been proposed to be capable of inhibition (9, 10), my findings demonstrated that there is in fact direct inhibition of a promoter by CarD.

Creation of a reliable *in vitro* transcription system and identification of CarD as an essential component of the *R. sphaeroides* transcription apparatus allowed me to test whether DksA and ppGpp negatively regulated *R. sphaeroides* rRNA transcription (Chapter 4). Interestingly, I found that the *R. sphaeroides* DksA homolog did not have an effect on rRNA transcription by itself *in vitro*, in contrast to the *E. coli* homolog (11,

12). In contrast, ppGpp by itself inhibited transcription from the *R. sphaeroides rrnB* promoter to a similar extent as described with *E. coli* RNAP (13, 14). Even though $DksA_{Rsp}$ by itself did not inhibit transcription under the conditions I used *in vitro*, it worked synergistically with ppGpp to inhibition transcription strongly from the *R. sphaeroides rrnB* promoter. I also demonstrated that DksA/ppGpp inhibited transcription from promoters for r-protein and tRNA genes, indicating that ribosome biogenesis is part of the stringent response in *R. sphaeroides*.

In Chapters 2 and 4, I found that CarD and DksA protein levels behaved similarly in aerobically grown cells, in that their levels were constant during log growth and decreased during stationary phase. However, Bryan Lakey from the Donohue lab and I found that protein levels of the two factors were different under photosynthetic conditions the – DksA levels were constant during log growth and into early stationary phase, while CarD levels quickly increased during outgrowth and then decreased during mid-to-late log growth. The different patterns displayed by CarD and DksA levels during photosynthetic growth could provide insights about when these factors are important for cellular function. In Chapter 4, I showed that DksA and ppGpp inhibited CarD-activated transcription *in vitro*. Knowing when ppGpp levels are induced during *R. sphaeroides* growth will illuminate when DksA likely inhibits transcription in cells (See Appendix A).

Studies of *E. coli* rRNA transcription have provided us with a framework for understanding promoter recognition by RNAP in general. However, as the study of transcription has extended into bacteria from different phylogenetic niches, there is a growing appreciation for the complexity of transcription regulation in different organisms. For example, transcription of rRNA is regulated by the essential transcription factors

RbpA and CarD in *M. tuberculosis* (15–17), while rRNA transcription is regulated by initiating NTPs in *B. subtilis* (18). When I first started this project, the expectation was that transcription in *R. sphaeroides* would be similar to *E. coli*, since both species are proteobacteria. In some ways, this expectation has proven correct, but as the data presented in this thesis reveal, *R. sphaeroides* appears to have evolved a system of transcription regulation that is a hybrid of those found in *E. coli* and *M. tuberculosis* while in some ways transcription in *R. sphaeroides* has its own unique properties. My hope is that the work presented in this thesis will provide a greater appreciation for the complexity of transcription regulation in *R. sphaeroides*.

FUTURE DIRECTIONS

Here I present some further subjects for future studies.

Why did *R. sphaeroides* (and other alpha-proteobacteria) evolve to have a so many promoters without a thymine at the last position of the -10 element?

We proposed in Chapter 2 that these promoters may have evolved -10 elements so that global gene expression could be coordinated by a single transcription factor. If so, how did the CarD-dependent α-proteobacterial regulatory circuits evolve, i.e., how did such a large fraction of α-proteobacterial promoters evolve to lack -7T even though RNAP retained the residues responsible for recognizing this key base in the -10 hexamer? If protein-coding changes occurred first, followed by changes in the cis-acting binding sites for those regulatory proteins, during the evolution of transcription circuits as has been proposed recently (19), how could so many -10 elements in α-proteobacterial

promoters have lost -7T in a concerted fashion? In fact, three participants (RNAP, CarD, and the promoter sequence) and perhaps other regulatory factors acting on these promoters all would have had to co-evolve to accommodate CarD-dependent regulatory circuits in the α proteobacteria. Exactly how this could have happened is an interesting evolutionary problem.

What regulates CarD protein levels during stationary phase?

The findings of Chapter 3 indicate that CarD autoregulates transcription from its own promoter. Autoregulation would keep CarD protein levels constant; however, we observed that CarD protein levels decline as cells enter stationary phase under both aerobic and photosynthetic conditions. Since CarD levels dropped in stationary phase, if CarD were the only factor that regulates its levels, one might expect an increase in transcription from the *carD* promoter and a subsequent increase in CarD levels in stationary phase, neither of which was observed. Therefore, the decline in CarD levels might result from post-transcriptional mechanisms. We found that DksA and ppGpp do not affect CarD transcription. In *Caulobacter crecentus*, the CarD homolog was shown to be a substrate for ClpXP proteolysis. Perhaps this mechanism could account for the turnover of CarD (20).

Together, the data suggest that there must be a mechanism for turning off transcription from the *carD* promoter during stationary phase. This mechanism could be dynamic and tied to nutritional signals that could increase CarD protein levels rapidly during outgrowth under both aerobic and photosynthetic conditions (Chapter 2 and 4).

When is ppGpp induced in *R. sphaeroides* cells?

When photosynthetically grown *R. sphaeroides* cells experience a light downshift, the growth response is characteristic of stringent control in that there is sharp increase in ppGpp levels followed by a decrease in rRNA synthesis and a corresponding change in ribosome synthesis to adjust cells to the new growth rate (4, 21). In addition to being induced in response in an abrupt change in nutrient availability, ppGpp is also induced when *E. coli* cells enter stationary phase (22). When *E. coli* cells enter stationary phase, there is an induction in ppGpp levels and $DksA_{Eco}$ helps downregulate transcription of genes involved in ribosome biogenesis (2, 23). Similar to the situation in E. coli, in Chapter 4, we found that $DksA_{Rsp}$ and ppGpp work synergistically to inhibit transcription from the promoters of rRNA and r-proteins. The points at which ppGpp levels are induced will be crucial in understanding when $DksA_{Rsp}$ is potentiated to affect gene expression (see Appendix A).

Why does *R. sphaeroides* DksA by itself does not inhibit rRNA transcription?

In *E. coli*, DksA by itself is capable of inhibiting rRNA transcription *in vitro*. In chapter 4, we found that the *R. sphaeroides* DksA homolog does not inhibit rRNA transcription by itself. The amino acid residues that are important for DksA functions are conserved in the DksA homologs from both species (Chapter 4; (3)). There could be subtle structural differences in DksA that could have arisen from the differences in amino acid sequence that may affect whether DksA alone can bind to RNAP and inhibit transcription. $DksA_{Rsp}$ was shown to bind to the secondary channel of $RNAP_{Eco}$ and inhibited transcription from

the *E. coli* rRNA promoter *rrnB* P1. To test whether the lack of inhibition is due to poor binding, we could use a binding assay such as an iron cleavage assay (3, 24). Another hypothesis is that DksA alone is not capable of antagonizing the RNAP-promoter complex once it has been stabilized by CarD and thus requiring ppGpp to be a more potent inhibitor.

What is the role of ppGpp and DksA in photosynthetic growth?

A previous study from our lab found that cells harboring a chromosomal deletion of *dksA* grew poorly under photosynthetic conditions, and assembly of the photosynthetic pigment-protein complex was impaired in the mutant strain (3). With the knowledge that ppGpp is produced when *R. sphaeroides* cells are downshifted in light (4, 5), we surmised that DksA and ppGpp must play a regulatory role in gene expression during photosynthetic growth. To better understand the global gene expression consequence of a DksA deletion, Bryan Lakey, a graduate student in the Donohue lab, performed RNA-seq on RNA extracted from *R. sphaeroides* wild-type and the Δ*dksA* strain. Preliminary findings of the analysis of the RNA-seq data is described in Chapter 4. Further analysis of the data should narrow down some candidates for genes involved in the photosynthetic growth defect caused by the absence of *dksA*.

What genes are directly regulated by ppGpp and DksA in vivo?

In Chapter 4, we showed that there are significant numbers of genes whose expression was changed by the absence of *dksA*. However, we were unable to distinguish which of these changes in gene expression were direct or indirect. One approach of identifying

genes that are directly regulated by ppGpp and DksA would be to conditionally induce ppGpp under non-starvation conditions by overexpression of the ppGpp synthetase domain of *E. coli* RelA (RelA') (25). This approach was successfully used by our lab to determine which were directly regulated by $DksA_{Eco}$ and ppGpp. We could employ a similar approach, by expression of RelA' from a replicable plasmid in *R. sphaeroides* (26) and determine global gene expression changes by RNA sequencing. Next, with the in vitro transcription system developed in Chapter 2, we can test which of these changes are direct.

Do DksA/ppGpp and CarD bind *R. sphaeroides* RNAP at the same time?

DksA/ppGpp and CarD appear to affect transcription in opposite directions – DksA/ppGpp destabilize open complex formation on promoters that make relatively unstable RNAP-promoter complexes to begin with while CarD stabilizes open complex formation on the same promoters. Structural studies showed that CarD binds to the beta1 lobe (15, 27) while DksA/ppGpp binds to the secondary channel and ppGpp binds to the omega-beta prime interface (11, 28). These regions on RNAP are non-overlapping and are at least ~59Å apart on RNAP. Structural data indicate that ppGpp site 1 and site 2 can be occupied at the same time (28), and thus it is conceivable that CarD could be bound to RNAP while these sites are occupied. There were no major conformation changes that were observed in the beta1 lobe when RNAP site 2 is occupied by DksA/ppGpp (28); however, these structural studies were done with RNAP from *E. coli*, which has no CarD homolog. This raises an interesting question, in addition

to destabilizing open complex formation, do DksA and ppGpp binding lead to a conformational change in *R. sphaeroides* RNAP that occludes CarD binding?

REFERENCES:

1. Gourse RL, Chen AY, Gopalkrishnan S, Sanchez-Vazquez P, Myers A, Ross W. 2018. Transcriptional Responses to ppGpp and DksA. Annual Review of Microbiology 72:163–184.
2. Paul BJ, Ross W, Gaal T, Gourse RL. 2004. rRNA Transcription in Escherichia coli. Annual Review of Genetics 38:749–770.
3. Lennon CW, Lemmer KC, Irons JL, Sellman MI, Donohue TJ, Gourse RL, Ross W. 2014. A Rhodobacter sphaeroides Protein Mechanistically Similar to Escherichia coli DksA Regulates Photosynthetic Growth. mBio 5.
4. Eccleston Jr. ED, Gray ED. 1973. Variations in ppGpp levels in Rhodopseudomonas spheroides during adaptation to decreased light intensity. Biochemical and Biophysical Research Communications 54:1370–1376.
5. Acosta R, Lueking DR. 1987. Stringency in the absence of ppGpp accumulation in Rhodobacter sphaeroides. J Bacteriol 169:908–912.
6. Myers KS, Vera JM, Lemmer KC, Linz AM, Landick R, Noguera DR, Donohue TJ. 2020. Genome-Wide Identification of Transcription Start Sites in Two Alphaproteobacteria, Rhodobacter sphaeroides 2.4.1 and Novosphingobium aromaticivorans DSM 12444. Microbiol Resour Announc 9.
7. Vera JM, Ghosh IN, Zhang Y, Hebert AS, Coon JJ, Landick R. 2020. Genome-Scale Transcription-Translation Mapping Reveals Features of Zymomonas mobilis Transcription Units and Promoters. mSystems 5.
8. Henry KK, Ross W, Myers KS, Lemmer KC, Vera JM, Landick R, Donohue TJ, Gourse RL. 2020. A majority of Rhodobacter sphaeroides promoters lack a crucial RNA polymerase recognition feature, enabling coordinated transcription activation. PNAS https://doi.org/10.1073/pnas.2010087117.
9. Galburt EA. 2018. The calculation of transcript flux ratios reveals single regulatory mechanisms capable of activation and repression. Proc Natl Acad Sci USA 115:E11604–E11613.
10. Jensen D, Manzano AR, Rammohan J, Stallings CL, Galburt EA. CarD and RbpA modify the kinetics of initial transcription and slow promoter escape of the Mycobacterium tuberculosis RNA polymerase. Nucleic Acids Res https://doi.org/10.1093/nar/gkz449.
11. Ross W, Sanchez-Vazquez P, Chen AY, Lee J-H, Burgos HL, Gourse RL. 2016. ppGpp Binding to a Site at the RNAP-DksA Interface Accounts for Its Dramatic

Effects on Transcription Initiation during the Stringent Response. Molecular Cell 62:811–823.

12. Gopalkrishnan S, Ross W, Chen AY, Gourse RL. 2017. TraR directly regulates transcription initiation by mimicking the combined effects of the global regulators DksA and ppGpp. PNAS 114:E5539–E5548.
13. Ross W, Vrentas CE, Sanchez-Vazquez P, Gaal T, Gourse RL. 2013. The Magic Spot: A ppGpp Binding Site on E. coli RNA Polymerase Responsible for Regulation of Transcription Initiation. Molecular Cell 50:420–429.
14. Barker MM, Gaal T, Josaitis CA, Gourse RL. 2001. Mechanism of regulation of transcription initiation by ppGpp. I. Effects of ppGpp on transcription initiation in vivo and in vitro1. Journal of Molecular Biology 305:673–688.
15. Srivastava DB, Leon K, Osmundson J, Garner AL, Weiss LA, Westblade LF, Glickman MS, Landick R, Darst SA, Stallings CL, Campbell EA. 2013. Structure and function of CarD, an essential mycobacterial transcription factor. PNAS 110:12619–12624.
16. Rammohan J, Ruiz Manzano A, Garner AL, Prusa J, Stallings CL, Galburt EA. 2016. Cooperative stabilization of Mycobacterium tuberculosis rrnAP3 promoter open complexes by RbpA and CarD. Nucl Acids Res gkw577.
17. Hubin EA, Tabib-Salazar A, Humphrey LJ, Flack JE, Olinares PDB, Darst SA, Campbell EA, Paget MS. 2015. Structural, functional, and genetic analyses of the actinobacterial transcription factor RbpA. PNAS 112:7171–7176.
18. Krásný L, Gourse RL. 2004. An alternative strategy for bacterial ribosome synthesis: Bacillus subtilis rRNA transcription regulation. EMBO J 23:4473–4483.
19. Britton CS, Sorrells TR, Johnson AD. 2020. Protein-coding changes preceded cis-regulatory gains in a newly evolved transcription circuit. Science 367:96–100.
20. Gallego-García A, Iniesta AA, González D, Collier J, Padmanabhan S, Elías-Arnanz M. 2017. Caulobacter crescentus CdnL is a non-essential RNA polymerase-binding protein whose depletion impairs normal growth and rRNA transcription. Scientific Reports 7:srep43240.
21. Campbell TB, Lueking DR. 1983. Light-mediated regulation of phospholipid synthesis in Rhodopseudomonas sphaeroides. J Bacteriol 155:806–816.
22. Murray HD, Schneider DA, Gourse RL. 2003. Control of rRNA Expression by Small Molecules Is Dynamic and Nonredundant. Molecular Cell 12:125–134.
23. Paul BJ, Barker MM, Ross W, Schneider DA, Webb C, Foster JW, Gourse RL. 2004. DksA: A Critical Component of the Transcription Initiation Machinery that Potentiates the Regulation of rRNA Promoters by ppGpp and the Initiating NTP. Cell 118:311–322.
24. Lennon CW, Gaal T, Ross W, Gourse RL. 2009. Escherichia coli DksA Binds to Free RNA Polymerase with Higher Affinity than to RNA Polymerase in an Open Complex. J Bacteriol 191:5854–5858.

25. Sanchez-Vazquez P, Dewey CN, Kitten N, Ross W, Gourse RL. 2019. Genome-wide effects on Escherichia coli transcription from ppGpp binding to its two sites on RNA polymerase. PNAS 201819682.

26. Ind AC, Porter SL, Brown MT, Byles ED, Beyer JA de, Godfrey SA, Armitage JP. 2009. Inducible-Expression Plasmid for Rhodobacter sphaeroides and Paracoccus denitrificans. Appl Environ Microbiol 75:6613–6615.

27. Bae B, Chen J, Davis E, Leon K, Darst SA, Campbell EA. 2015. CarD uses a minor groove wedge mechanism to stabilize the RNA polymerase open promoter complex. eLife Sciences e08505.

28. Molodtsov V, Sineva E, Zhang L, Huang X, Cashel M, Ades SE, Murakami KS. 2018. Allosteric Effector ppGpp Potentiates the Inhibition of Transcript Initiation by DksA. Molecular Cell 0.

APPENDIX A

Measuring ppGpp in *Rhodobacter sphaeroides* cells by LC-MS

The goal of this work is to determine ppGpp levels in *R. sphaeroides* cells by LC-MS. The data will be included in a manuscript that is currently in preparation that investigates rRNA regulation by DksA and ppGpp (Chapter 4). The experiments described below were performed by me and are preliminary. This work is a collaboration with the Amador-Noguez, Wang and Donohue labs. This work will be continued by Bryan Lakey, a graduate student in the Donohue lab.

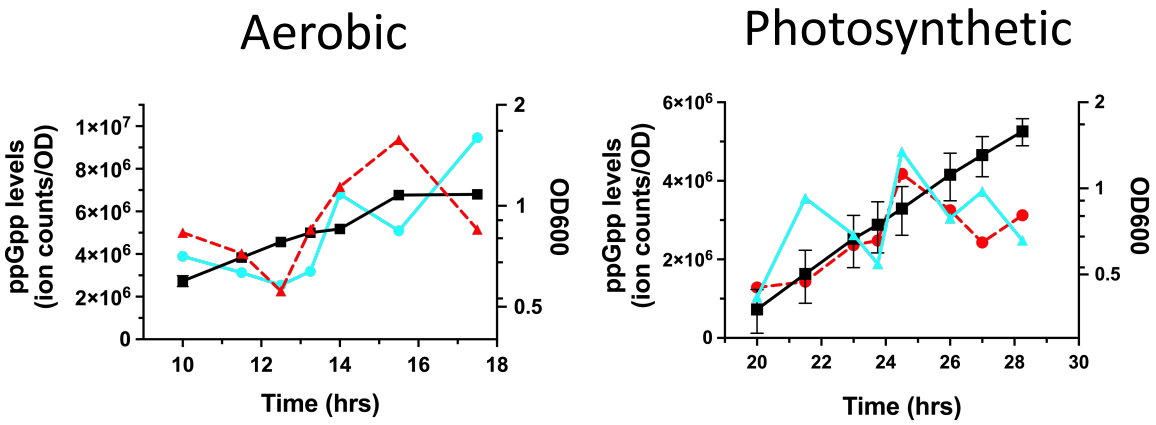
SUMMARY

We found that under photosynthetic conditions, CarD proteins levels decline much earlier than DksA levels, which were constant into stationary phase (Chapter 4). We hypothesize that the decline in CarD levels could be correlated by an induction of ppGpp, and thus DksA will be potentiated to affect gene expression in *R. sphaeroides*. To determine whether ppGpp is induced when *R. sphaeroides* cells enter stationary phase, we will utilize an LC-MS-based method of detecting ppGpp from extracted metabolites. This approach was successfully used to quantitatively detect ppGpp in *B. subtilis* and *E. coli* cells (1, 2).

RESULTS

Metabolites were extracted from *R. sphaeroides* cells grown either under aerobic or photosynthetic conditions. ppGpp levels was monitored throughout log phase and into stationary phase. The method was successful in detecting ppGpp levels in *R. sphaeroides* cells growth either aerobically or photosynthetically; however, the ppGpp levels were inconsistent between samples taken at the same time points (Figure AA). The differences in ppGpp level could be accounted for growth variability between cultures grown at the same time. Future effort is current being dedicated to properly account for these variabilities and minimize sources of error.

Figure AA: Levels of ppGpp levels in aerobically or photosynthetically grown *R. sphaeroides* cells. ppGpp was measured by LC-MS. Normalized ion count is ion count per OD_{600} per unit volume of the culture. Two independent cultures were grown for each growth condition (cyan or red). Error bar on OD_{600} (black plot) represents the average of OD600 from two growth experiments.



MATERIALS AND METHODS

Bacterial strains and growth conditions. *R. sphaeroides* 2.4.1 wild-type (RLG14946) were grown in succinate-based minimal medium (SIS) (3) at 30°C. For liquid growth conditions, *R. sphaeroides* was grown in 500-800 mL cultures with constant sparging with 69% N_2, 30% O_2 and 1% CO_2 at 30°C for aerobic growth and 98.5% N_2, 0.5% O_2 and 1% CO_2 at 30°C in front of an incandescent light with a light intensity of 10 W/m^2 for photosynthetic growth. For aerobic and photosynthetic growth curves, sample cultures were removed by positive pressure from the main culture and optical density was determined at OD_{600}.

Metabolite extraction. For the either the aerobic and photosynthetic growth experiment: separate cultures were grown, and optical density (OD_{600}) was monitored from log to stationary phase. At the indicated OD_{600}, 8OD equivalent of cells were filtered through PTFE membrane (Sartorius). Membranes with cell pellet were submerged into 50mL conical tube containing 3 mL extraction solvent (50:50 (v/v) chloroform/water) to lyse cells, quench metabolism and extract metabolites containing ppGpp. The mixture was incubated on ice for 15minutes, with intermittent vortexing to lyse cells. Afterwards, the mixture was transferred to a 15mL conical, spun at 6000xg for 10min to separate the organic and liquid phase. The liquid phase (1.1mL) was transfer to a 1.5mL Eppendorf tube, centrifuged at 16100xg for 10minutes to remove cell debris. The metabolite profile in the extracted samples were analyzed using an HPLC-tandem MS (HPLC-MS/MS) system as previously described (1, 2). Samples were frozen at -80°C if immediately analysis was not possible.

REFERENCES:

1. Fung DK, Yang J, Stevenson DM, Amador-Noguez D, Wang JD. 2020. Small Alarmone Synthetase SasA Expression Leads to Concomitant Accumulation of pGpp, ppApp, and AppppA in Bacillus subtilis. Front Microbiol 11.
2. Yang J, Anderson BW, Turdiev A, Turdiev H, Stevenson DM, Amador-Noguez D, Lee VT, Wang JD. 2020. The nucleotide pGpp acts as a third alarmone in Bacillus , with functions distinct from those of (p) ppGpp. 1. Nature Communications 11:5388.
3. Sistrom WR. 1960. A Requirement for Sodium in the Growth of Rhodopseudomonas spheroides. Microbiology 22:778–785.

APPENDIX B

Transcriptional studies of CarD from *Caulobacter crescentus* on cognate RNA polymerase and *carD* promoter

The goal of this work is to test whether C. *crescentus* CarD inhibited transcription of the $carD_{Ccr}$ promoter when added to cognate RNA polymerase. This work was discontinued due to time constraints.

SUMMARY

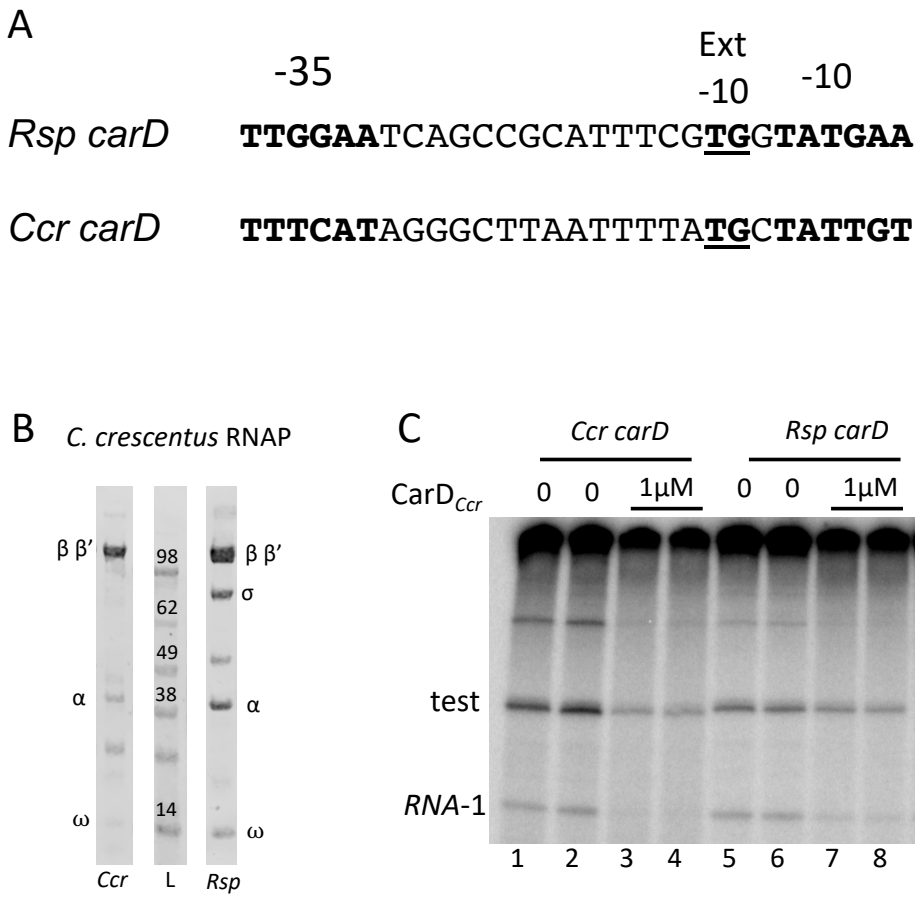
Caulobacter crescentus is an aquatic α-proteobacterium with a markedly asymmetric division pattern. During each cell cycle *C. crescentus* divides producing a warmer and stalked progeny with different cell fates and distinct morphologies (1). This bacterium is highly studied as a model organism for prokaryotic cycle control and cell differentiation. The CarD homolog of *C. crescentus* (CdnL) was recently shown to interact with the RNA polymerase (RNAP) β subunit and localizes at the *rrnA* promoter *in vivo*. While it was found to be nonessential, its deletion causes significant reduction in growth and aberrant cell physiology such as filamentation and more curved cells (2, 3). Gene analysis of Δ*carD* cells found transcriptional downregulation of genes involved in metabolism (2). The $carD_{Ccr}$ is predicted to be σ^{70} dependent and promoter-reporter assay found that activity from the promoter was high and increased during exponential growth and leveled off at stationary phase (3); however, in vitro transcriptional studies of the $carD_{Ccr}$ promoter has not been reported. In Chapter 3, I reported that basal activity from the $carD_{Rsp}$ promoter was high and that $CarD_{Rsp}$ inhibited transcription when added to $RNAP_{Rsp}$. The extended -10 feature is conserved in the *carD* promoter for *R. sphaeroides* and *C. crescentus* (Figure AB.1A). With these observation in mind, I tested whether $CarD_{Ccr}$ had an effect on transcription from the $carD_{Ccr}$ using purified $RNAP_{Ccr}$.

RESULTS

I purified *C. crescentus* RNAP using the purification method I developed for purifying *R. sphaeroides* RNAP (Figure AB.1B). The α, β and β' subunits of $RNAP_{Ccr}$ were in stoichiometric amounts; however, the prep was depleted for σ^{70} and thus

holoenzyme was reconstituted with purified σ^{70} (Karen Wassarman). The purified $RNAP_{Ccr}$ transcribed the *carD* promoters from both *R. sphaeroides* and *C. crescentus*, as well as the plasmid encoded RNA-1 promoter. The intensity of signal of each transcript (*carD*, *RNA*-1 and the large transcripts at the top of the lane) within the lane was reduced when $CarD_{Ccr}$ was added (Figure AB.1C). These results are inconclusive as to whether $CarD_{Ccr}$ had an effect on transcription from the $carD_{Ccr}$ promoter. It is likely that the purified $CarD_{Ccr}$ introduced RNase into the reaction. Future work would need to test whether the $CarD_{Ccr}$ has RNase activity.

Figure AB: Activity of the *R. sphaeroides* or *C. crescentus carD* promoter with *C. crescentus* RNAP. (A) Sequence comparison of the of the *R. sphaeroides (Rsp)* or *C. crescentus (Ccr) carD* promoter, from −35 to -7. The -35 and -10 elments are bold and the extended -10 are bold and underlined. (B) 2.0µg of purified C. crescentus RNAP (Ccr) or *R. sphaeroid*es RNAP (Rsp) were run on a 4-12% Invitrogen NuPAGEBis-Tris gel with MES buffer. The SeaBlue2 size marker (Invitrogen, L) with the indicated MW (kDa). Intervening lanes were removed for clarity. (C) in vitro transcription of the *R. sphaeroides* or *C. crescentus carD* promoter with *C. crescentus* RNAP in the presence or absence of *C. crescentus* CarD (1µM).



MATERIALS AND METHOD

Bacterial strains and growth conditions. Bacterial strains are listed in Table AB-1**.** *E. coli* strains were grown in either LB Lennox media or LB agar plates at 37°C supplemented with kanamycin (30 µg/mL) or chloramphenicol (20 µg/mL) when required. *Caulobacter crescentus* NA1000 strains were grown in peptone yeast extract medium (PYE) at 30°C with kanamycin (30 µg/mL) where required.

Construction of in vitro transcription templates. The $carD_{Ccr}$ promoter was PCR isolated from *C. crescentus* NA1000 genomic DNA and inserted into the pRLG770 (4) using the oligonucleotides listed in Table AB-2. The oligonucleotide set created a promoter fragment with the described endpoints relative to the predicted transcription start site based on the presence of predicted -10 and -35 elements. The resulting promoter fragment contained sequences for ligation to linearized pRLG770 (amplified using oligonucleotides RLG7057 and RLG7056) by Hi-Fi Assembly (NEB). The resulting plasmid was electroporated into *E. coli* XL1-blue, and those confirmed to have the correct insert size on gels were sequenced with oligonucleotides RLG4252 and RLG4253.

Purification of *C. crescentus* RNAP and SUMO-tagged CarD wild type protein. *C. crescentus* RNA polymerase ($RNAP_{Ccr}$ was purified using Ni-agarose affinity chromatography from 2L of cells containing the chromosomally encoded His10-β' (RLG15637: *C. crenscentus* NA1000 with a rpoC::rpoC-HIS_{10} (5)) grown in PYE supplemented with kanamycin (30 µg/mL).Cells were harvested at OD600 ~1 (late log phase), and cell pellets were resuspended in Bugbuster (Novagen), 5 mL of Bugbuster per gram of wet cell, supplemented with 5 µL per gram Lysonase (Novagen), 1X HALT

protease inhibitor (ThermoFisher Scientific), and 23 µg/mL phenylmethylsulfonyl fluoride (PMSF). After incubation at room temperature with gentle rocking for 1 hour, cell debris was removed by centrifugation at 14,000 RPM for 20 min at 4°C, and the cleared lysate was then added to a 1 mL Ni-NTA resin column (Qiagen) pre-equilibrated with RNAP wash buffer (50 mM NaH2PO4 [pH 8], 300 mM NaCl, 20 mM imidazole, 5% glycerol, 0.1% Tween 20, and 5 mM β-mercaptoethanol). After washing the column with RNAP wash buffer, RNAP was subsequently washed with buffer#2 containing (25 mM NaH2PO4 [pH 8], 20 mM Tris-Cl [pH 8.0], 100 mM KCl, 10 mM $MgCl_2$, 150 mM NaCl, 20 mM imidazole, 2.5% glycerol, 0.05% Tween 20, and 2.5 mM β-mercaptoethanol, 0.1 mM EDTA, 1 mM DTT) followed by wash buffer#3 containing (20 mM Tris-Cl pH 8.0, 200 mM KCl, 20 mM $MgCl_2$, 0.2 mM EDTA and 2 mM DTT). Next, RNAP was eluted with wash buffer#3 containing 300mM imidazole, dialyzed twice into 2X RNAP storage buffer (10 mM Tris-Cl [pH 8.0], 100mM KCl, 10 mM MgCl2, 0.1 mM EDTA, 1 mM DTT, 20 µM ZnCl2) at 4°C, and concentrated using a 5 mL spin concentrator (Sartorius) with a 5 kDa molecular weight cutoff. An equal volume of 100% glycerol was added, and the protein concentration was assayed using the Bradford reagent (Bio-Rad) using BSA as a standard. Proteins were stored at -20°C. *R. sphaeroides* holo-enzyme was purified in parallel.

The *carD* $_{Ccr}$ gene was PCR isolated from *C. crescentus* CB15 genomic DNA and cloned into the pETSUMO vector for overexpression and purification from *E. coli* BL21 using its IPTG-inducible T7 promoter as described (6). Untagged purified CarD was stored in storage buffer (20 mM Tris-HCl [pH 8.0], 100 mM NaCl, 2 mM DTT, 50% glycerol). Protein concentration was assayed by the Bradford assay using BSA as a

standard. For in vitro transcription assays, CarD was diluted into CarD storage buffer to the desired concentration.

in vitro transcription. Multiple round transcription assays were carried out using supercoiled plasmid templates containing the promoters listed in Table AB-1 using *R. sphaeroides* RNAP (20 nM) or *C. crescentus* RNAP (10nM) in 25μL reactions. *C. crescentus* RNAP was found to be depleted for σ^{70} thus holoenzyme was constituted by adding purified σ^{70} (gift from Karen Wassarman, UW-Madison) to RNAP at a 1:2 ratio in transcription buffer. Template DNA (5 pmol or 400 ng) was incubated in transcription buffer (20 mM Tris-HCl [pH 7.9], 10 mM MgCl2, 1 mM DTT, 0.1 mg/ml BSA, and either 170 or 100 mM NaCl, as indicated, containing NTPs (200 μM ATP, GTP, CTP, 10μM UTP, and 1.2 μCi α-^{32}P UTP [Perkin Elmer]) with CarD or CarD storage buffer as indicated at room temperature for 10 min.

Transcription was initiated by addition of RNAP and reactions were incubated at 30°C for 20 min. The reactions were terminated by addition of 25 μL water-saturated phenol, aqueous and phenol phases were separated by brief centrifugation, and 16 μL of the aqueous phase containing the transcript was transferred to 16μL 2X loading solution (8M urea, 10 mM EDTA, 2X TBE, 1% SDS and 0.02% bromophenol blue). Transcripts were denatured by heating the samples at 95°C for 1 min, resolved on 6% polyacrylamide 7M urea gels, imaged using a Typhoon phosphorimager.

Table AB.1: Strain and Plasmid used in study

Strains used in this study

Strains	Genotype	Source
BL21 (DE3) pLysS	*E. coli* str. B F^- *ompT gal dcm lon* $hsdS_B(r_B^-m_B^-)$ λ(DE3 [*lacI lacUV5-T7p07 ind1 sam7 nin5*]) [$malB^+$]$_{K-12}$(λ^S) pLysS[*T7p20* ori_{p15A}](Cm^R)	Novagen
E. coli XL1-Blue	Δ(*mcrA*)*183* Δ(*mcrCB-hsdSMR-mrr*)*173 endA1 supE44 thi-1 recA1* gyrA96 relA1 lac *[F'* proAB $lacI^qZ\Delta M15$ *Tn*10 *(Tet')]*	Stratagene
RLG14765	*R. sphaeroides* 2.4.1 with a C-terminal thrombin-His_{10} chromosomally tagged β' (*rpoC*)	(6)
RLG15640	*Caulobacter crescentus* NA1000 (wild type)	
RLG15637	Caulobacter *crescentus* NA1000 (RLG15640) with a C-terminal thrombin-His_{10} chromosomally tagged β' (*rpoC*), Kan^R	(5)

Plasmids used in this study

Plasmids	Description	Source
pRLG770	Transcription vector	(4)
pETSUMO	pET derived plasmid that encodes a Ulp1 cleavable His_{10}SUMO tag. ORFs are cloned in frame with the coding sequence of the SUMO tag.	Invitrogen
15656	C. crescentus CarD gene in pETSUMO	This study
15662	C. crescentus carD promoter in p770	This study
15686	15656 cloned into BL21 (DE3) pLysS	This study

Table AB.2: Oligonucleotides used for constructing *in vitro* transcription vectors

Primers to linearize RLG770

Oligonucleotide	Sequence
7057	5'-AAGCTTGGGTCCCACCTGACC-3'
7058	5'-GAATTCTTGAAGACGAAAGGGCC-3'

Primers to screen and sequence promoters inserted into RLG770

Oligonucleotide	Sequence
4252	5'-CCCGAAAAGTGCCACCTGACG-3'
4253	5'-CAGTTCCCTACTCTCGCATGG-3'

Primers used to construct *in vitro* transcription plasmids[a]

Plasmid	Oligonucleotide	Sequence	Description
p15662	8777	GGCCCTTTCGTCTTCAAGAATTC GATTTCTCAAGCCCAGAGCGTG	*C. cresentus carD* promoter, -160 to +45
	8778	GGTCAGGTGGGACCCAAGCTTG AAACGGCTTCTTCGACAAGACG	

[a] For each plasmid constructed, both the forward and reverse primers are given

Plasmids used for *in vitro* transcription

Plasmids	**Promoter[a]**	**Endpoints[b]**	**Source**
p15221	*R. sphaeroides carD*	-161 to +50	Chapter 3
p15662	*C. cresentus carD*	-160 to +45	This study

REFERENCES:

1. Ryan KR, Shapiro L. 2003. Temporal and Spatial Regulation in Prokaryotic Cell Cycle Progression and Development. Annu Rev Biochem 72:367–394.
2. Woldemeskel SA, Daitch AK, Alvarez L, Panis G, Zeinert R, Gonzalez D, Smith E, Collier J, Chien P, Cava F, Viollier PH, Goley ED. 2020. The conserved transcriptional regulator CdnL is required for metabolic homeostasis and morphogenesis in Caulobacter. PLOS Genetics 16:e1008591.
3. Gallego-García A, Iniesta AA, González D, Collier J, Padmanabhan S, Elías-Arnanz M. 2017. Caulobacter crescentus CdnL is a non-essential RNA polymerase-binding protein whose depletion impairs normal growth and rRNA transcription. Scientific Reports 7:srep43240.
4. Ross W, Thompson JF, Newlands JT, Gourse RL. 1990. E.coli Fis protein activates ribosomal RNA transcription in vitro and in vivo. EMBO J 9:3733–3742.
5. Haakonsen DL, Yuan AH, Laub MT. 2015. The bacterial cell cycle regulator GcrA is a σ70 cofactor that drives gene expression from a subset of methylated promoters. Genes Dev 29:2272–2286.
6. Henry KK, Ross W, Myers KS, Lemmer KC, Vera JM, Landick R, Donohue TJ, Gourse RL. 2020. A majority of Rhodobacter sphaeroides promoters lack a crucial RNA polymerase recognition feature, enabling coordinated transcription activation. PNAS https://doi.org/10.1073/pnas.2010087117.