

INTEGRATING TWO SIGNALS: THE RESPONSE REGULATOR, VPSR, REQUIRES
C-DI-GMP AND PHOSPHORYLATION TO DRIVE TRANSCRIPTION INITIATION OF
BIOFILM GENES IN *VIBRIO CHOLERAE*

By

Meng-Lun Hsieh

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Biochemistry and Molecular Biology-Doctor of Philosophy

2018

ABSTRACT

INTEGRATING TWO SIGNALS: THE RESPONSE REGULATOR, VPSR, REQUIRES C-DI-GMP AND PHOSPHORYLATION TO DRIVE TRANSCRIPTION INITIATION OF BIOFILM GENES IN *VIBRIO CHOLERAE*

By

Meng-Lun Hsieh

The mechanistic role of the small bacterial second messenger cyclic di-GMP (c-di-GMP) in transcription initiation has remained unclear. In *Vibrio cholerae*, the causative agent of the disease cholera, VpsR is the master Enhancer Binding Protein (EBP) that binds c-di-GMP to increase biofilm gene expression at the biofilm biogenesis promoter, P_{vpsL} , *in vivo*. Unlike typical EBPs that activate RNA polymerase (RNAP) containing the alternate sigma (σ) factor, σ_{54} , VpsR has several different features: 1) it lacks conserved residues needed to bind to σ_{54} and hydrolyze ATP; 2) it retains a highly conserved aspartic acid (D59) residue, which is typically phosphorylated; and 3) it activates P_{vpsL} in the absence of σ_{54} *in vivo*. These features all suggest a different unknown mechanism of transcription activation.

To address this mechanism, we have established an *in vitro* system and show for the first time that c-di-GMP is sufficient to directly activate transcription with another activator at P_{vpsL} . Unlike other regulators which use c-di-GMP to promote oligomerization and/or increase DNA binding affinity, the presence of c-di-GMP neither affects VpsR oligomerization nor significantly changes the affinity of VpsR for P_{vpsL} DNA. Instead, $KMnO_4$ and DNase I footprinting reveal that the P_{vpsL}/σ_{70} -RNAP/VpsR/c-di-GMP complex forms the open transcription bubble and adopts a different conformation from that formed by P_{vpsL}/σ_{70} -RNAP with or without c-di-GMP or VpsR.

To investigate the role of the D59 residue, we have characterized the phosphodeficient variant D59A and the phosphomimetic D59E. While both variants dimerize and bind DNA with

$K_{d(\text{app})}$ s similar to that of wildtype (WT), D59E activates transcription and forms the open transcription bubble while D59A yields basal transcription. DNase I footprints of the transcription complex made with D59E resemble those made with WT, while footprints with D59A resemble those of RNAP alone. We have also developed a method to denature and renature VpsR (VpsR^{REN}). In the absence of the high-energy phosphate donor, acetyl phosphate, VpsR^{REN} now resembles D59A. Addition of acetyl phosphate results in a VpsR^{REN} that behaves like the previously purified WT VpsR and D59E in *in vitro* transcriptions, electrophoretic mobility shift assays (EMSAs), DNase I and KMnO₄ footprinting.

Lastly, we have also explored whether VpsR has additional regulatory gene expression roles. We have identified three new promoters that are regulated by VpsR and c-di-GMP *in vitro*: *rbmA*, *rbmF*, and *vpsU*. Interestingly, like P_{vpsL}, the regulated genes of each promoter represent the first gene of their gene cluster or operon. Similar to P_{vpsL}, binding of c-di-GMP does not alter protein-DNA contacts at these promoters in the absence of σ 70-RNAP. *In vitro* transcriptions require both c-di-GMP and VpsR, and the positions of the VpsR binding sites at these promoters reveal that VpsR can utilize both Class I and Class II activation to upregulate gene expression.

In conclusion, VpsR represents a novel c-di-GMP dependent transcription regulator. Not only does it use c-di-GMP to promote open complex formation, but our data also suggests that phosphorylation is simultaneously required for that process. As the master regulator of biofilm formation, VpsR directly activates a set of *vps* and *rbm* promoters using different mechanisms of transcription activation. Understanding these mechanisms not only provide a new paradigm in c-di-GMP-dependent transcription activation and elucidate mechanistic processes that regulate biofilm formation, but also provide the foundation needed for the development of novel chemical inhibitors against *V. cholerae* and biofilm-based nosocomial infections

To my parents and grandparents

ACKNOWLEDGEMENTS

I would like to thank my mentors, Dr. Deborah Hinton and Dr. Chris Waters. I am immensely grateful for their guidance and mentorship with which this work would not have been possible. With both their expertise in biochemistry and microbial genetics, I am so fortunate to learn from two experts in two complementary fields. Thank you for the endless opportunities, support, and freedom to explore, making me into a better and more well-rounded scientist. I am especially grateful to Debbie who took a chance and allowed me to join her lab first as a NIH post-baccalaureate fellow and later as a graduate student. Though I initially found it daunting to share a bay together, I can't envision sitting anywhere else! Thank you for all the invaluable skills both inside and outside the lab. You have truly helped me grown as a scientist and a person.

I would also like to thank my committee members: Dr. David Arnosti, Dr. Bill Henry, and Dr. Leslie Kuhn, who provided valuable advice, support, and perspectives. I would also like to acknowledge my preliminary exam chair, Dr. Zach Burton.

In the DO/PhD program, I would like to thank Dr. Justin McCormick, Dr. Brian Schutte, and Mrs. Bethany Heinlin, for their constant support and understanding while navigating both medical school and graduate school.

I would like to acknowledge all past and current members of the Hinton lab and Waters lab. Thank you for always making research stimulating and challenging, yet still fun, lively, and interesting. I can't imagine doing research anywhere else. In particular, I am grateful to Dr. Kyung Moon who has provided invaluable advice, troubleshooting, protocols, and ideas. I am also thankful for Leslie Knipling for suggestions, help with protein purification, and careful manuscript editing. This dissertation would also have not been possible without the pounds of chocolate,

seaweed, and Biscoff cookies as well as liters of coffee that I have consumed. Also thank you little red hen.

I would like to thank my parents, sister, and grandparents for their encouragement, patience, and unconditional support. I am especially thankful to my dad for introducing me to the world of research and science.

Thank you, Andrew Shubin, for your support, compassion, and understanding. You always inspire me as I follow you in your footsteps as a physician-scientist.

Lastly, I would like to acknowledge the funding support that I received at Michigan State University MMG, BMB, College of Natural Science, DO/PhD program as well as NIH intramural program, NIDDK, and GPP program. Thank you NIGMS and NRSA for the F30 award.

TABLE OF CONTENTS

LIST OF TABLES	ix
LIST OF FIGURES	x
KEY TO SYMBOLS AND ABBREVIATIONS	xiv
CHAPTER 1: Introduction	1
1.1 cyclic-di-GMP signaling	2
1.2 <i>Vibrio cholerae</i> as a pathogen	6
1.3 Biofilm formation and regulation in <i>Vibrio cholerae</i>	8
1.4 Transcription initiation and classes of activation	11
1.5 c-di-GMP dependent transcription regulators	15
1.6 Quorum sensing in <i>Vibrio cholerae</i>	20
1.7 Two-component signal transduction pathways	23
1.8 Summary	27
CHAPTER 2: VpsR and c-di-GMP together drive transcription initiation to activate biofilm formation in <i>Vibrio cholerae</i>	29
Preface	30
2.1 Introduction	31
2.2 Materials and Methods	35
2.3 Results	42
2.4 Discussion	59
CHAPTER 3: VpsR directly activates transcription of multiple biofilm genes in <i>Vibrio cholerae</i>	63
Preface	64
3.1 Introduction	65
3.2 Materials and Methods	68
3.3 Results	72
3.4 Discussion	79
CHAPTER 4: VpsR requires c-di-GMP and phosphorylation to drive transcription initiation to activate biofilm formation in <i>Vibrio cholerae</i>	84
Preface	85
4.1 Introduction	86
4.2 Materials and Methods	91
4.3 Results	97
4.4 Discussion	112
CHAPTER 5: Conclusion and future directions	118

APPENDIX	132
REFERENCES	136

LIST OF TABLES

Table 1.1: List of c-di-GMP-dependent transcription activators and repressors	16
Table A.1: Table of plasmids and primers used in this study	133
Table A.2: Strains of <i>Vibrio cholerae</i> and <i>Escherichia coli</i> used in this study	135

LIST OF FIGURES

Figure 1.1:	Synthesis and degradation of c-di-GMP and effector targets	3
Figure 1.2:	c-di-GMP controls a wide variety of phenotypes	5
Figure 1.3:	Regulation of biofilm genes in <i>Vibrio cholerae</i>	11
Figure 1.4:	Different classes of transcription activation at sigma70-dependent promoters and sigma54-dependent promoters	15
Figure 1.5:	Quorum sensing and c-di-GMP signaling in <i>Vibrio cholerae</i>	23
Figure 1.6:	Phosphotransmission and phosphorelay schematic in bacterial two-component systems	24
Figure 1.7:	Synthesis and degradation of acetyl phosphate	26
Figure 2.1:	CLUSTAL Omega alignment of the AAA+ domain of VpsR with other bacterial Enhancer Binding Proteins	34
Figure 2.2:	<i>vpsL-lux</i> gene reporter assays in <i>Vibrio cholerae</i> and <i>Escherichia coli</i> S17- <i>lpir</i>	42
Figure 2.3:	Summary of <i>in vitro</i> primer extension and DNase I and KMnO ₄ footprinting at P _{<i>vpsL</i>}	43
Figure 2.4:	VpsR and c-di-GMP activate transcription at P _{<i>vpsL</i>} by approximately 7-fold <i>in vitro</i>	44
Figure 2.5:	Increasing c-di-GMP alone or VpsR alone has no effect on basal level transcription activity at P _{<i>vpsL</i>} <i>in vitro</i>	45
Figure 2.6:	VpsR only needs the proximal binding site to activate transcription at P _{<i>vpsL</i>}	45
Figure 2.7:	Identification of +1 P _{<i>vpsL</i>} TSS using <i>in vitro</i> and <i>in vivo</i> primer extensions	47
Figure 2.8:	KMnO ₄ footprinting assigns the +1 TSS at the A located 59 bp upstream of the <i>vpsL</i> translation start site	48
Figure 2.9:	KMnO ₄ footprinting of transcription complexes at P _{<i>vpsL</i>} on template DNA	49
Figure 2.10:	VpsR does not hydrolyze ATP	50

Figure 2.11:	VpsR, c-di-GMP, RNAP, and P _{vpsL} DNA rapidly form active and heparin- resistant transcription complexes in single round <i>in vitro</i> transcriptions	51
Figure 2.12:	VpsR forms dimers <i>in vitro</i> with or without c-di-GMP	53
Figure 2.13:	VpsR binds P _{vpsL} DNA with similar affinity with or without c-di-GMP	54
Figure 2.14:	DNase I footprinting of P _{vpsL} complexes on nontemplate DNA and template DNA	55
Figure 2.15:	Discrete P _{vpsL} /RNAP complexes are formed in EMSAs	56
Figure 3.1:	<i>Vibrio cholerae</i> genomic organization of genes involved in <i>Vibrio</i> polysaccharide (VPS) synthesis and ribomatrix protein production	66
Figure 3.2:	VpsR binds P _{rbmA} , P _{rbmF} , and P _{vpsU} with and without c-di-GMP	72
Figure 3.3:	VpsR and c-di-VpsR and c-di-GMP activate transcription at P _{rbmA} , P _{rbmF} , and P _{vpsU} <i>in vitro</i>	73
Figure 3.4:	Determination of +1 transcriptional start site (+1) <i>in vivo</i> and <i>in vitro</i> for P _{rbmA} , P _{rbmF} , and P _{vpsU}	74
Figure 3.5:	DNase I footprinting of VpsR with and without c-di-GMP at template P _{rbmA} P _{rbmF} , and P _{vpsU}	76
Figure 3.6:	DNase I footprinting of transcription complexes containing RNAP alone or RNAP with VpsR and c-di-GMP at template P _{rbmA} , P _{rbmF} , and P _{vpsU}	77
Figure 3.7:	Promoter sequence and summary of DNase I footprinting of P _{rbmA} , P _{rbmF} , P _{vpsU} , and P _{vpsL}	78
Figure 3.8:	Mechanisms of VpsR transcription activation at different promoters	80
Figure 4.1:	CLUSTAL Omega alignment of the N-terminal domain of VpsR with other Enhancer Binding Proteins	97
Figure 4.2:	VpsR D59 residue is required for <i>in vivo</i> and <i>in vitro</i> activities at P _{vpsL}	98
Figure 4.3:	<i>In vitro</i> and <i>in vivo</i> primer extensions at P _{vpsL} require c-di-GMP and WT VpsR or the phosphomimetic D59E variant	99
Figure 4.4:	WT and VpsR D59E form the open complex with RNAP and c-di-GMP at P _{vpsL} in KMnO ₄ footprinting	100

Figure 4.5:	WT VpsR, VpsR D59A, and D59E form dimers <i>in vitro</i> with or without c-di-GMP	101
Figure 4.6:	Trypsin proteolysis of WT VpsR, D59A, and D59E with or without c-di-GMP	102
Figure 4.7:	WT VpsR, VpsR D59A, and D59E bind P _{vpsL} DNA with similar affinities with or without c-di-GMP	103
Figure 4.8:	DNase I footprinting of P _{vpsL} complexes containing WT VpsR, VpsR D59A, and VpsR D59E on nontemplate DNA and template DNA	104
Figure 4.9:	Acetyl phosphate has no effect on transcription activation <i>in vitro</i> when using the WT VpsR isolated without denaturation	106
Figure 4.10:	<i>In vitro</i> transcription activation using VpsR ^{ren} requires acetyl phosphate (Ac~P)	108
Figure 4.11:	Formation of the open complex requires VpsR ^{ren} and Ac~P at P _{vpsL}	108
Figure 4.12:	VpsR ^{ren} and VpsR ^{ren} ~P form dimers <i>in vitro</i> with or without c-di-GMP	109
Figure 4.13:	Trypsin proteolysis of VpsR ^{ren} and VpsR ^{ren} ~P demonstrates no significant changes in protein conformation with or without c-di-GMP	109
Figure 4.14:	VpsR ^{ren} and VpsR ^{ren} ~P bind P _{vpsL} similarly with or without c-di-GMP	110
Figure 4.15:	DNase I footprinting of P _{vpsL} transcription complexes containing WT VpsR, VpsR ^{ren} , and VpsR ^{ren} ~P nontemplate DNA and template DNA	110
Figure 4.16:	<i>vpsL-lux</i> gene reporter assays in <i>E. coli</i> MG1655 and Δ <i>ackA</i> Δ <i>pta</i>	112
Figure 4.17:	Mechanistic model of c-di-GMP-dependent transcription activation at P _{vpsL}	115
Figure 5.1:	Three-plasmid genetic screen to isolate VpsR mutants that are blind to c-di-GMP or constitutively active in the absence of c-di-GMP	122
Figure 5.2:	Characterization of VpsR mutants binding to c-di-GMP by DRaCALA	123
Figure 5.3:	Single round <i>in vitro</i> transcriptions of VpsR mutants	124
Figure 5.4:	Bacterial adenylate cyclase two hybrid assay of VpsR +/- c-di-GMP and α subunit domains	125
Figure 5.5:	Miller assay investigating interactions between VpsR +/- c-di-GMP and α	126

KEY TO SYMBOLS AND ABBREVIATIONS

α	Alpha
AHL	Acyl homoserine lactone signal molecules
AI	Autoinducer
AAA+	ATPases Associated with diverse cellular Activities
AC	Adenylate cyclase
ADP	Adenosine-5'-disphosphate
ATP	Adenosine-5'-triphosphate
β	Beta
BACTH	Bacterial adenylate cyclase two-hybrid
BS ³	Bis(sulfosuccinimidyl)suberate
BSA	Bovine serum albumin
bp	basepair
cAMP	cyclic adenosine monophosphate
c-di-GMP	cyclic dimeric guanosine 3',5'-monophosphate
CT	Cholera enterotoxin
CTD	C-terminal domain
CTP	Cytosine-5'-triphosphate
CV	crystal violet
DGC	diguanylate cyclase enzyme
DNA	deoxyribonucleic acid
DTT	Dithiothreitol
dNTP	deoxyribonucleotide triphosphate

DRaCALA	Differential Radial Capillary Action of Ligand Assay
ds	double stranded
EBP	Enhancer binding protein
EDTA	Ethylenediaminetetraacetic acid
EMSA	Electrophoretic mobility shift assay
FeBABE	Iron(S)-1-(p-bromoacetamidobenzyl) ethylenediaminetetraacetate
GEMM	Genes for the environment, membranes, and motility
GMP	Guanosine monophosphate
GTP	Guanosine-5'-triphosphate
H-T-H	Helix-Turn-Helix
HCD	High cell density
Hpt	His-containing phosphotransfer
HK	Histidine kinase
IPODHR	<i>In vivo</i> Protein Occupancy Display High Resolution
IPTG	Isopropyl β -D1-thiogalactopyranoside
K_d	Dissociation constant
Kglu	Potassium glutamate
KMnO ₄	Potassium permanganate
LB	Lysogeny broth, a rich medium for bacterial growth
LC-MS/MS	Liquid chromatography coupled with tandem mass spectrometry
LCD	Low cell density
MBEC	Minimum Biofilm Eradication Concentration
MSHA	Mannose-sensitive hemagglutinin pili

NTD	N-terminal domain
NTP	Nucleotide triphosphate
OD	Optical density
PAGE	Polyacrylamide gel electrophoresis
Poly(dI-dC)	Poly(deoxyinosinic-deoxycytidylic) acid sodium salt
pGpG	5'-phosphoguanylyl-(3'-5')-guanosine
(p)ppGpp	guanosine pentaphosphate or tetraphosphate
PCR	Polymerase chain reaction
PDE	Phosphodiesterase enzyme
PEI	Polyethylenimine
QS	Quorum sensing
RBS	Ribosomal binding site
REC	Receiver domain
R.L.U	Relative luminescence units
RNA	Ribonucleic acid
RNAP	RNA polymerase
rNTP	ribonucleoside triphosphate
RPc	Closed transcription complex
RPi	Initiating transcription complex
RPo	Open transcription complex
RR	Response regulator
σ	sigma
SDS	Sodium dodecyl sulfate

ss	single stranded
TAE	Tris/Acetic acid/EDTA
TBE	Tris/Borate/EDTA
TCP	Toxin co-regulated pilus
TCS	Two-component signal transduction
Tris	2-amino-2-hydroxymethyl-propane-1,3-diol
TSS	Transcriptional start site
UTP	Uridine-5'-triphosphate
VPS	<i>Vibrio</i> polysaccharide
ω	omega
WT	Wildtype

CHAPTER 1:

Introduction

1.1 cyclic-di-GMP signaling

From relative obscurity to one of the most ubiquitous bacterial signaling molecules, cyclic dimeric (3' → 5') guanosine monophosphate (c-di-GMP) was first discovered by Moshe Benziman and colleagues in 1987 as the long-sought allosteric activator of cellulose biosynthesis in *Acetobacter xylinum* (1). Despite its initial discovery, only ten papers mentioned c-di-GMP from 1987 to 2000, nine of which came from the Benziman laboratory (2). These papers addressed a variety of topics that are still investigated today including c-di-GMP signaling in other bacteria such as *Agrobacterium tumefaciens*, identification of diguanylate cyclases (DGCs) and phosphodiesterases (PDEs), c-di-GMP concentration within the bacterial cell, screens of DGC inhibitors, and interkingdom crosstalk (3-13).

Around the beginning of the century, other investigators began to appreciate that the GGDEF and EAL or HD-GYP domains, named for the amino acids in their active sites, are widespread regulators of c-di-GMP signaling in many organisms. GGDEF domain are diguanylate cyclase enzymes (DGCs) that synthesize c-di-GMP while EAL or HD-GYP domains are phosphodiesterase enzymes (PDEs) that degrade c-di-GMP (14-16). While GGDEFs synthesize c-di-GMP from two molecules of GTP, EALs break down and linearize c-di-GMP into 5'-pGpG, which can be further degraded by the nanoRNase Orn, while HD-GYPs directly degrade c-di-GMP to two GMPs (Fig. 1.1). Synthesis of c-di-GMP is dependent on the active site (A site) of the GGDEF domain. Point mutations (except a D to E mutation) eliminate enzymatic activity (17). Additionally, to help with excess GTP consumption and overproduction, most DGCs contain a secondary site (I site) which is separated from the A site by a linker of five amino acids (18). This I site is characterized by a RXXD motif.

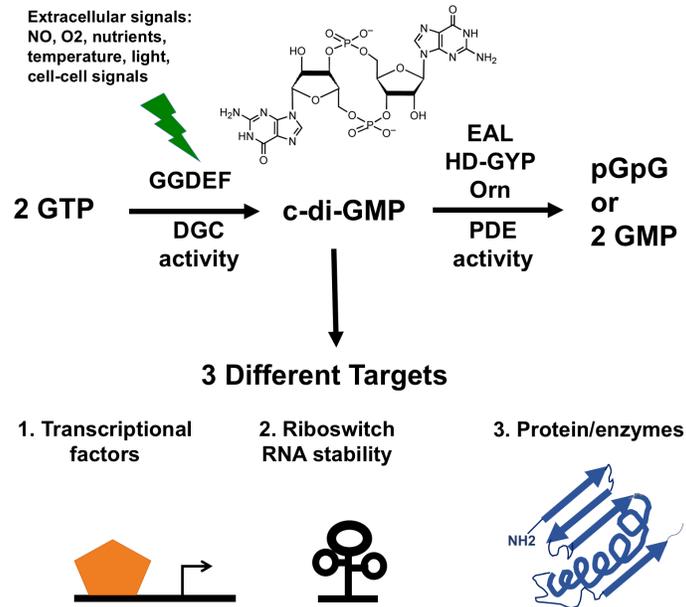


Figure 1.1: Synthesis and degradation of c-di-GMP and effector targets. Synthesis and degradation of c-di-GMP is dependent on diguanylate cyclases (DGCs) containing modular sensory GGDEF domains and phosphodiesterases (PDEs) containing EAL or HD-GYP domains, respectively. As c-di-GMP levels rise, c-di-GMP exerts its effects on three different targets: transcriptional factors, riboswitches, and proteins.

Most enzymes that mediate c-di-GMP signaling contain one of these three domains, but some contain both a synthesis and degradation domain. For example, BphG1 in *Rhodobacter sphaeroides* preferentially activates the DGC domain via allosteric regulation by its amino-terminal bacteriophytochrome domain while SrcC in *Vibrio parahaemolyticus* modulates the GGDEF-EAL domains with accessory proteins, ScrA and ScrB (19,20). C-di-GMP signaling networks are complex as most bacteria encode numerous DGCs and PDEs. For example, *Escherichia coli* K12 contains 12 GGDEFs, 10 EALs, and 7 GGDEF-EAL hybrids and *Vibrio cholerae* contains 31 GGDEFs, 12 EALs, and 10 hybrids (21,22). This speaks to the intricacies involved in regulating c-di-GMP concentrations.

Similar to two-component signal transduction systems which use a sensor histidine kinase and a response regulator (RR), synthesis and degradation of c-di-GMP to control cellular

concentration levels are dependent on signal inputs. Some GGDEF, EAL, and HD-GYP domains contain modular N-terminal sensory input domains with one or more transmembrane helices placed in the periplasm of Gram-negative bacteria while others are located in the cytoplasm (22). Environmental and cellular signals perceived by the bacterium include oxygen, light, starvation, redox conditions, antibiotics, polyamines, or intercellular signaling molecules, such as heme- or flavin-associated PAS domains (2,23-25).

Along with the discovery of c-di-GMP as a regulator of cellulose biosynthesis, we now know today that c-di-GMP is involved in regulating a wide variety of phenotypes. These phenotypes include type two and type six secretion systems, DNA repair, virulence, cell cycle progression, antibiotic production, and regulating the transition between biofilm formation and motility via stimulation of the biosynthesis of exopolysaccharide substances in biofilms and inhibition of various forms of motility (Fig. 1.2) (2,26,27). In order for c-di-GMP to exert its wide variety of different effects, it must first bind to different effectors, allosterically altering its structure and function. These effectors bind c-di-GMP with a dissociation constant (K_d) ranging from 1 nM to 2 μ M (2). Currently, there are three main types of c-di-GMP effectors: proteins, including those of the PilZ family and I site effectors; transcription factors; and the GEMM (genes for the environment, membranes, and motility) encoding riboswitches. The most well-studied are the PilZ family of proteins which are named after the *Pseudomonas aeruginosa* type IV pilus control protein. These c-di-GMP-binding domains are often linked to GGDEF, EAL, or HD-GYP domains on their carboxyl terminus and bind to c-di-GMP to regulate phenotypes via protein-protein interactions. I site effectors have degenerate A sites within the GGDEF domains. These proteins include *V. cholerae* CdgG which binds to c-di-GMP to control rugosity, biofilm formation, and motility as well as the *Caulobacter crescentus* PopA which binds to c-di-GMP to

sequester the replication inhibitor and cell cycle regulator CtrA, targeting it for subsequent degradation (28,29). Transcription factors bind to c-di-GMP to upregulate or downregulate gene expression. For example, *V. cholerae* VpsR binds to c-di-GMP to directly activate transcription with RNA polymerase containing $\sigma 70$, upregulating expression of polysaccharide genes (30). On the other hand, the *V. cholerae* transcription factor FlrA binding to c-di-GMP alters the DNA-binding domain of this protein to inhibit expression of motility genes. Along with proteins, c-di-GMP is also a ligand for riboswitches (31). Riboswitches are RNAs that alter gene expression upon binding to a target ligand. C-di-GMP binds to the 5'-untranslated regions of the highly conserved GEMM RNA domain, which typically contains one of two similar architectures: a specific tetraloop (type 1) or a tetraloop receptor sequence (type 2) (32). These GEMM RNAs are often found residing upstream of DGC and PDE open reading frames or genes controlled by c-di-GMP such as *V. cholerae* *tfoY* and *gbpA* (32). Thus, regulation occurs at many different levels including transcriptional, post-transcriptional, and direct allosteric regulation at the protein level.

High c-di-GMP levels

- Sessility: biofilms, EPS, matrix
- Virulence: T6SS, intracellular replication, phagocytosis resistance
- Transmission
- Environmental persistence
- Cell morphology
- Heterocyst development
- Cell cycle control
- Photosynthesis
- Cell-cell communication
- Detergent resistance

Low c-di-GMP levels

- Motility: swimming, swarming, twitching, gliding
- Virulence: T3SS, invasion, proinflammatory cytokines
- Phage resistance
- Heavy metal resistance
- Hyphae formation
- Antibiotic production



Figure 1.2: c-di-GMP controls a wide variety of phenotypes. High levels and low levels c-di-GMP exert different phenotypic behaviors.

1.2 *Vibrio cholerae* as a pathogen

V. cholerae is a rod-shaped gram-negative bacterium with one single polar flagellum. It causes the well-known and potentially fatal intestinal disease, cholera, which was first described by John Snow in 1854 and identified in 1883 by Robert Koch. Upon ingestion of the bacteria through contaminated water and or food, within three to four hours, a previously healthy individual will start secreting profuse rice-watery diarrhea. This leads to rapid dehydration with an average loss of 20 mL/kg, which can proceed to hypovolemic shock, metabolic acidosis, and ultimately death without proper treatment within twenty-four to forty-eight hours. Thus, cholera is a rapidly progressing disease. During outbreaks, the fatality ratio is up to 50% in vulnerable groups, but can be under 1% with proper treatment (33). Treatment includes immediate fluid resuscitation with antibiotics as adjunct treatment. Adequate sanitation, water, and hygiene infrastructures are the mainstay preventative measures, and oral cholera vaccines have recently been found to be effective and inexpensive in high-risk endemic areas.

Every year, cholera causes an estimated three to five million cases and 100,000 to 120,000 deaths (34). The disease is endemic in 69 countries located in Asia, Africa, and the Americas (35). These countries have poor water sanitation and hygiene facilities and are often plagued by natural disasters, such as famines and/or humanitarian crises, such as wars. Though *V. cholerae* has hundreds of serogroups, only two serogroups, O139 and O1, cause disease. The O1 serogroup can be further divided into two biotypes, classical and El Tor. First discovered in the Ganges Delta, cholera has since spread beyond Asia with over seven recorded pandemics. Six of the seven pandemics belong to the classical biotype and occurred in the 1800s beginning in 1817 while the last and seventh pandemic, caused by the El Tor biotype, started in 1961 and continues today. From there, El Tor is also further classified into four major clonal groups: Australian clone; U.S.

Gulf Coast clone; seventh pandemic clone; and Latin American clone. Virtually all cases worldwide are now derived from El Tor. O139 strains are rarely isolated.

V. cholerae's natural habitat is the marine environment including coastal saline waters and estuaries. Often found associated with shellfish and zooplankton, chitin is the predominant source of carbon and nitrogen for the bacterium (36). In addition, it induces natural competence to facilitate horizontal gene transfer (36,37). In response to nutritional deficiencies, *V. cholerae* enters a viable, but non-culturable state, living in biofilms (38,39). In response to zooplankton blooms, *V. cholerae* then proliferates rapidly (40).

Spreading via the fecal-oral route, humans are the only known hosts of *V. cholerae*. Infection requires ingestion of 10^8 bacteria in healthy volunteers when challenged. It is possible that lower doses are sufficient in impaired individuals, such as those with impaired gastric barriers (41). Upon ingestion, *V. cholerae* then colonizes the small intestines, secreting cholera toxin. Cholera toxin, first described in rabbits in 1959 and later discovered in the 1970s, uses ADP-ribosylation to induce intracellular cyclic AMP signaling, thereby triggering massive fluid secretion (42-44). The protein is 84 kDa in size with one 28 kDa toxic-active A subunit and a ring of five identical 11.6 kDa B-subunits (45). After assembly of the toxin in the periplasm and subsequent secretion into the intestines, the B subunit binds to the epithelial cell surface receptor, GM1, and is then endocytosed and transported to the endoplasmic reticulum (ER) in a retrograde manner (46). Within the ER, the A subunit dissociates from the B pentamer and enters the cytosol in order to catalyze ADP-ribosylation of adenylate cyclase (AC), locking AC in a GTP-bound activated state (47). This leads to a stimulation of the cAMP-responsive protein kinase channel, the cystic fibrosis transmembrane conductance regulator chloride channel (48). Stimulation results

in an increase in chloride and bicarbonate export and a decrease in sodium uptake, resulting in subsequent water excretion into the intestines due to osmotic imbalance.

Bacterial detachment and dissemination are the last stages of the infectious cycle. Infected individuals typically secrete the bacteria for two to three days while symptomatic patients can secrete for two or more weeks (49). To cope with nutrient limitations upon shedding into the environment, the bacteria upregulate many processes including recycling of amino acids, proteins, and cell wall components as well as utilization of alternative carbon sources such as succinate (36,50).

1.3 Biofilm formation and regulation in *Vibrio cholerae*

V. cholerae produces biofilms to aid in environmental dissemination, persistence, and survival from nutrient limitation and predation by bacteriophages and protozoa (51-53). Biofilms are comprised of aggregates of extracellular matrix consisting of nucleic acids, proteins, and sugars. Found on both biotic and abiotic surfaces, *V. cholerae* has a preference for forming biofilms on zooplankton, phytoplankton, and chitin rain (54). Along with the environment, biofilms have also been found within the human hosts and play an important role in host to host transmission. Removal of particles larger than 20 mm lead to a decrease in cholera cases, suggesting that manual biofilm removal hinders bacterial transmission (55,56). Biofilm aggregates have also been found within stools of infected individuals, suggesting a hyper-infectious phenotype during host to host transmission (57,58).

Biofilm formation is a complex multistep process. First, the bacterium scans and attaches to the surface. Initial roaming and attachment are mediated by *V. cholerae*'s single polar flagellum which is powered by a sodium ion motor and the mannose-sensitive hemagglutinin pili (MSHA)

(36). Upon attachment, extracellular matrix is secreted. Important components of the *Vibrio* extracellular matrix include extracellular DNA, matrix proteins such as RbmA, RbmC, and Bap1, and VPS (*Vibrio* polysaccharides). Both matrix proteins and VPS are expressed in a set of gene operons or clusters called VcBMC (*V. cholerae* biofilm matrix cluster). VcBMC is comprised of six ribomatrix proteins (*rbmA*, *rbmC*, and *rbmBDEF*) flanked by 12 genes located within *vps-1* (*vpsU*, *vpsA-K*) and six genes located within *vps-2* (*vpsL-vpsQ*) (59-62). 8.3 kilo base pairs separate these two *vps* operons. Despite the base pair separation distance, *Vibrio* biofilm formation is orchestrated in a sequential manner. Over 50% of *Vibrio* biofilms are comprised of VPS and during biofilm formation, it is synthesized and secreted first after attachment (59-64). While VPS is essential for biofilm three-dimensional structure, matrix proteins are important for biofilm support and cell-cell adhesion. After secretion of *vps*, RbmA matrix proteins are synthesized next and secreted via the type II secretion system (65). Upon accumulation of RbmA on the cell surface, which also helps facilitate cell-cell interaction, Bap1 is then secreted at the attachment interface while RbmC is secreted at discrete sites on the cell surface (66). Thus, maturation of the biofilm results in organized clusters of *vps*, Bap1, RbmA, and RbmC.

Unsurprisingly, biofilm regulation is complex and requires multiple transcriptional regulators, small RNAs, and various signaling molecules including c-di-GMP, cyclic adenosine monophosphate (cAMP), and guanosine 3'-diphosphate 5'-triphosphate and guanosine 3',5'-bis(diphosphate) ((p)ppGpp). Transcriptional activators include VpsR, VpsT, and AphA while transcriptional repressors include H-NS and HapR (Fig. 1.3). Although both VpsR and VpsT bind c-di-GMP directly with a K_d of 1.6 μM and 3.2 μM , respectively, and both activators regulate overlapping biofilm genes such as *vps-1*, *vps-2*, and *vpsT*, VpsR is considered the master regulator of biofilm formation (67-71). Deletion of *vpsR* leads to larger defects in biofilm formation

compared to deletion of *vpsT* (53,72,73). Along with VpsR, AphA, the master regulator of virulence expression, also upregulates VpsT and Vps-1 (69). Contrary to VpsR, HapR is the master biofilm repressor. HapR production is regulated through quorum sensing. Under low cell density conditions, *hapR* mRNA remains untranslated due to the sRNA chaperone Hfq and the sRNAs, Qrr1-4 (70,74). Thus, in the absence of HapR, biofilm genes are unrepressed under low cell density conditions. Furthermore, the histone-like protein, H-NS also act as a major repressor of biofilm gene expression by directing binding to *vps-1*, *vps-2*, and *vpsT* (71,75-77). Lastly, both c-di-GMP and ppGpp upregulate biofilm formation while cAMP downregulates biofilm formation. As mentioned earlier, c-di-GMP directly interacts with both transcriptional activators, VpsR and VpsT. The role of c-di-GMP appears different for both activators. While c-di-GMP enhances VpsT-DNA binding (71), c-di-GMP is required for VpsR to form the active open complex with RNA polymerase containing $\sigma 70$ (30). Similar to c-di-GMP, (p)ppGpp upregulates biofilm genes including *vpsR* and *vpsT*. While *vpsR* transcription requires all three (p)ppGpp synthesis enzymes, RelA, SpoT, and RelV, *vpsT* transcription only requires RelA (78-80). Unlike c-di-GMP and (p)ppGpp, cAMP downregulates biofilm genes by binding to CRP and decreasing expression of multiple genes including *rbmA*, *rbmC*, *bap1*, *vpsR*, *vpsT*, and *vpsL* as well as upregulating the biofilm repressor, HapR (81-83).

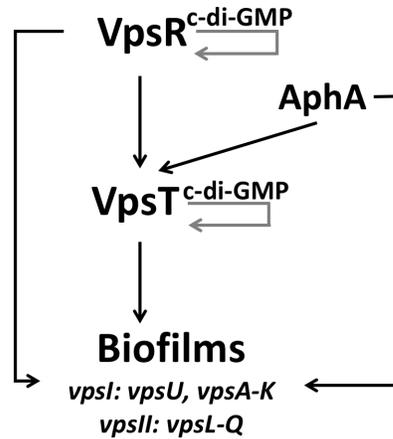


Figure 1.3: Regulation of biofilm genes in *Vibrio cholerae*. Transcription activation of biofilm genes include the activators, VpsR, VpsT, and AphA and the signaling molecule, c-di-GMP. H-NS and HapR act as repressors at both *vpsT* and *vpsI* and *vpsII* (not shown).

1.4 Transcription initiation and classes of activation.

Transcription is a fundamental process that is conserved across all three domains of life (84). It is the first step of gene expression and is thus a highly-regulated process. At the heart of this activity is the catalytic enzyme, RNA polymerase (RNAP) which is comprised of a core of five subunits (β , β' , two α s, and ω) and a specificity factor, sigma (σ) (85). σ factors are important because RNAP core enzyme alone cannot recognize promoter DNA or initiate transcription efficiently.

σ factors can be classified into two main classes: $\sigma 70$ family and $\sigma 54$ family based on their phylogenetic relatedness (86,87). Primary σ factors, such as $\sigma 70$ in *E. coli*, are housekeeping σ s that are primarily responsible for exponential growth (85). The domains of primary σ factors are divided into Regions 1-4, based on structure and function. Alternative σ factors are used during other times of stress or different growth conditions, such as nitrogen utilization. While $\sigma 54$ does not share sequence similarity with $\sigma 70$, other alternate σ factors, such as extracytoplasmic function

(ECF) σ and the stationary phase σ , σ^s , share sequence similarity with Regions 2 and 4 or Regions 2, 3, and 4, respectively (87-90).

Binding of σ to RNAP core generates the RNAP holoenzyme. For transcription initiation using the $\sigma 70$, double stranded DNA (dsDNA) is first recognized, forming an unstable short-lived complex called the closed complex (RPc) (91,92). Without σ factor, RNAP alone cannot efficiently recognize the promoter DNA (87). Recognition of the promoter DNA includes multiple contacts between RNAP and the DNA by σ Regions 2, 3, and/or 4, and/or α CTDs. α CTDs can recognize the minor grooves of the UP elements, an AT-rich region between -40 and -60 relative to the +1 transcriptional start site (TSS) (93). The helix-turn-helix (H-T-H) motif of $\sigma 70$ Region 4.2 can interact with the -35 element, $^{-35}\text{TTGACA}^{-30}$; Region 3 interacts with the extended -10, $^{15}\text{TGn}^{-13}$; and Region 2.4 interacts with the first double stranded (ds) base of the -10 element, $^{12}\text{TATAAT}^{-7}$ (87,91,94). While elements upstream of the -12 are recognized as dsDNA, elements downstream of the -12 are recognized as single stranded DNA (89-92).

After formation of the RPc with RNAP core, $\sigma 70$, and promoter DNA, kinetically favorable isomerizations quickly transition RPc to the heparin-resistant stable open complex, RPo (95,96). Within the RPo, major conformational changes occur within polymerase and the dsDNA bends approximately 90°, forming a transcription bubble from -12/-11 to approximately +5 relative to the +1 TSS (84,96-98). In this bubble σ Region 2.3 interacts with the top (nontemplate) strand of the -10 element from -11 to -7 (ATAAT) while σ Region 1.2 can interact with the ‘discriminator’ located immediately downstream of the -10 and containing a GC-rich motif +1 TSS (84,96-98). This bubble allows the template strand of the DNA to descend into the active site with the TSS at the proper position for RNA synthesis (84). Addition of NTPs entering through the secondary channel allows RPo to proceed to the initially transcribing complex, RPi (96). Short

RNA chains approximately two to thirteen nucleotides, also known as abortives, are first made (96). Because Region 3.2 is still occupying the RNA exit channel, synthesis of RNA chains longer than 13 nucleotides is inhibited (96). However, eventually the longer RNA can displace Region 3.2, leading to promoter clearance and formation of the elongation complex, usually with the release of σ (99). This is also known as promoter escape.

For many promoters, RNAP containing $\sigma 70$ can activate transcription alone. This is known as basal expression. However, other $\sigma 70$ -dependent promoters, which deviate from the ideal consensus sequences, require activators and/or signaling molecules to regulate transcription. There are two main classes of activation: Class I and Class II. In Class I, activators bind to upstream sites of the promoter and contact the α CTDs (Fig. 1.4) (85). In Class II activation, activators bind to a site directly upstream or overlapping the -35 element and contact $\sigma 70$ Region 4 and/or α NTDs (Fig. 1.4) (85). Both classes of activators recruit RNAP, although Class II activation also typically promotes RPo formation. However, there are other types of activation that use different mechanisms. For example, MerR family activators induce a conformational change in the promoter DNA, shortening the suboptimal distance between the -35 and -10 elements that is present in these promoters (85). In another example, called σ appropriation, a co-activator binds to RNAP, altering the H-T-H structure of $\sigma 70$ Region 4, while the activator binds to the promoter DNA and the rearranged Region 4 (100). While most activators do not require binding to small molecules to upregulate gene expression, others utilize signaling molecules such as (p)ppGpp, cAMP, and c-di-GMP to activate transcription. For example, binding of cAMP to *E. coli* CRP causes a coil-to-helix transition, positioning the DNA binding domains in the proper orientation for DNA binding and subsequent transcription activation (101).

Unlike $\sigma 70$, transcription initiation differs dramatically with $\sigma 54$. Though both σ factors bind to core polymerase, they share very little sequence similarities. Unlike $\sigma 70$ -dependent activation, which can transcribe certain promoters in the absence of additional regulators, $\sigma 54$ -dependent transcription absolutely requires an activator. These activators are ATPases and typically bind to sites around 80 to 150 bps upstream of the TSS (102-104). Because this distal binding is reminiscent of eukaryotic enhancer binding proteins, these activators are also commonly known as bacterial enhancer binding proteins (EBPs). Along with utilizing activators, $\sigma 54$ also recognizes different DNA elements and contains very distinct conserved domains. Instead of recognizing -35 and -10 elements, $\sigma 54$ interacts with -12 GC and -24 GG elements (102-104). With its two highly conserved domains (Region 1 and Region 3) linked by a flexible linker domain (Region 2), Region 1 of $\sigma 54$ interacts with core RNAP, the -12 element, and then EBP via the EBP's GAFTGA motif. Region 2 is involved in RNAP isomerization and Region 3 contains a putative H-T-H motif which interacts strongly with both the -12 and the -24 elements (102-104).

Unsurprisingly, with its drastic differences in sequence and binding interactions, the mechanism of $\sigma 54$ -dependent transcription initiation is significantly different (Fig. 1.4). While the $\sigma 70$ -RNAP first forms a typically unstable R_{Pc}, the $\sigma 54$ -RNAP is stable. Transition of $\sigma 54$ -RNAP into R_{Po} not only requires bending of the DNA, typically facilitated by integration host factor or small heteromeric proteins, but most importantly utilizes energy derived from EBP ATP hydrolysis (Fig. 1.4) (102-104). Upon R_{Po} formation, the complex is now competent for transcription initiation and regulation of a set of genes, including those for nitrogen utilization, motility, and fimbriae synthesis.

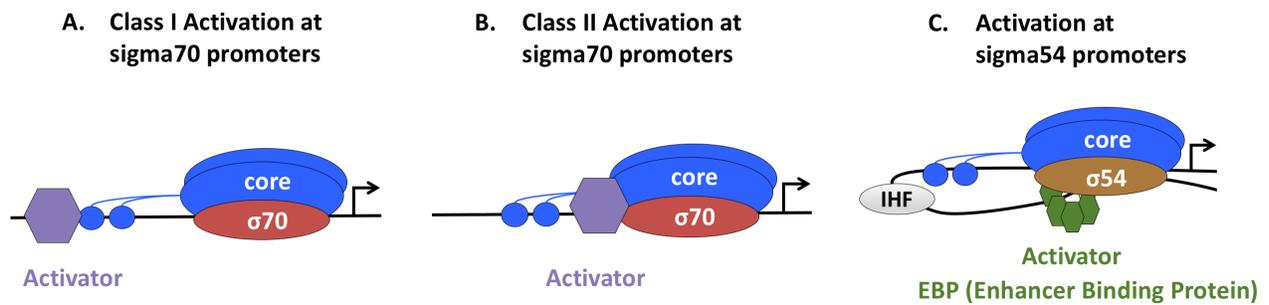


Figure 1.4: Different classes of transcription activation at sigma70-dependent promoters and sigma54-dependent promoters. (A) Activators (purple) in Class I activation of sigma70 ($\sigma70$) promoters bind to upstream sites and contact the α CTDs (blue circles). (B) Activators in Class II activation of $\sigma70$ promoters bind to proximal sites surrounding the -35 element, or overlapping the -35 element. (C) Activators, also known as enhancer binding proteins, bind to -100 to -150 relative to the +1 TSS and interact with the $\sigma54$ -dependent promoter as a hexamer via ATP hydrolysis, utilizing IHF to mediate DNA looping and $\sigma54$ binding.

1.5 c-di-GMP-dependent transcription factors

As mentioned earlier, some regulators function alone with RNAP to alter gene expression while others additionally use signaling molecules such as (p)ppGpp, cAMP, or c-di-GMP. Currently there are 11 transcriptional regulators that bind to c-di-GMP to activate or repress gene expression (Table 1.1). Below, I will briefly describe each regulator's mechanism of action.

Table 1.1: List of c-di-GMP-dependent transcription activators and repressors. With only eleven published c-di-GMP-dependent transcription regulators (8 activators and 3 repressors), the field is still novel. Each regulator binds c-di-GMP with a dissociation constant (K_d) ranging from less than 1 micromolar to 20 micromolar. A variety of different c-di-GMP binding pockets are used and a wide variety of different genes are regulated by these c-d-GMP-dependent transcription factors.

<u>Transcription factor</u>	<u>Organism</u>	<u>Family</u>	<u>Activator or repressor upon c-di-GMP binding</u>	<u>Crystal structure</u>	<u>K_d for c-di-GMP binding</u>	<u>c-di-GMP binding pocket or important residues</u>	<u>Functions controlled</u>
VpsT	<i>Vibrio cholerae</i>	LuxR-like	Activator	Yes	3.2 μ M	W[F/L/M][T/S]R	Biofilms and DNA repair
VpsR	<i>Vibrio cholerae</i>	EBP	Activator	No	1.6 μ M	unknown	Biofilms, virulence, and T2SS
FleQ	<i>Pseudomonas aeruginosa</i>	EBP	Activator	Yes	20 μ M	LFR ¹⁴⁴ S motif; R ¹⁸⁵ , N ¹⁸⁶ , ExxxR ³³⁴	Motility and Biofilms
BriR	<i>Pseudomonas aeruginosa</i>	MerR	Activator	Yes	2.2 μ M	R ³¹ , Y ⁴⁰ , R ⁶⁷ , R ⁸⁶ , Y ²⁷⁰	Multidrug transport
MrkH	<i>Klebsiella pneumoniae</i>	PilZ domain	Activator	Yes	0.24 μ M	RxxxR and D/NxSxxG	Fimbriae expression
Bcam1349 (BerA)	<i>Burkholderia cenocepacia</i>	CRP-like	Activator	No	~10 μ M	unknown	Biofilms
LtmA	<i>Mycobacterium smegmatis</i>	TetR	Activator	No	0.83 μ M	unknown	Lipid transport and metabolism
BerB	<i>Burkholderia cenocepacia</i>	EBP	Activator	No	~3 μ M	unknown	Biofilms
FliA	<i>Vibrio cholerae</i>	EBP	Repressor	No	2.4 μ M	R ^{13b} and R ¹⁷⁶	Motility
BldD	<i>Streptomyces venezuelae</i>	---	Repressor	Yes	2.5 μ M	RXD-X ₈ -RXXD	Vegetative growth and sporulation
Clp	<i>Xanthomonas campestris</i>	CRP-like	Repressor	Yes	3.5 μ M	D ¹⁰ , R ¹⁵⁴ , R ¹⁵⁶ , D ¹⁷⁰	Virulence

V. cholerae contains three known c-di-GMP-dependent transcription factors: FlrA, VpsT, and VpsR. FlrA is an EBP that binds c-di-GMP with a K_d of 2.4 μM ; binding to c-di-GMP abrogates DNA binding, thereby inhibiting transcription of *flrBC* and repressing flagellar biosynthesis genes (31). The c-di-GMP binding pocket of FlrA include R135 and R176 located within both the REC and AAA+ domains (31). On the other hand, c-di-GMP interacts with VpsT and VpsR to upregulate overlapping biofilm genes (53,67,72,73). VpsT belongs to the LuxR/FixJ/CsgD family of transcription regulators. Two c-di-GMP molecules bind a VpsT dimer with a K_d of 3.2 μM using the four-residue motif W[F/L/M][T/S]R (68). Binding of c-di-GMP enables VpsT to bind biofilm promoters and upregulate their gene expression (71). In the presence of H-NS, a highly abundant transcriptional silencer and nucleoid organizer that binds to AT-rich sequences, VpsT-c-di-GMP together functions as an anti-H-NS repressor (77). Interestingly, at the *rpoS* promoter, VpsT-c-di-GMP binds to two identified transcription initiation sites, repressing transcription of *rpoS* (105). On the other hand, unlike VpsT, binding of c-di-GMP has no effect on VpsR dimerization ability and DNA binding affinity at the promoter for *vpsL* (P_{vpsL}) (30). Unlike all previously characterized c-di-GMP-dependent transcription regulators, c-di-GMP is required to form the competent open complex with RNAP containing $\sigma 70$ at P_{vpsL} (30). Alignment of VpsR with other c-di-GMP-dependent transcription regulators does not reveal any conserved c-di-GMP binding residues (data not shown) (106). Though the binding pocket is unknown, VpsR binds c-di-GMP with a K_d of 1.6 μM *in vitro* (69).

Pseudomonas aeruginosa contains two known c-di-GMP-dependent transcription factors: FleQ and BrlR. FleQ is the best characterized c-di-GMP dependent transcriptional regulator. FleQ is an EBP that has been suggested to work with both RNAP containing $\sigma 70$ as well as RNAP containing $\sigma 54$ (107-111). Together with the ATPase FleN, FleQ upregulates both flagellar and

exopolysaccharide synthesis in response to low or high c-di-GMP concentrations, respectively (107-111). At the exopolysaccharide *pel* promoter, FleQ binds to two sites and functions as a repressor at low c-di-GMP concentrations, but upon binding to c-di-GMP immediately switches to an activator due to release of DNA-binding at the proximal binding site near the TSS. FleQ binds to c-di-GMP with a K_d of 4.1 μ M using three key motifs: LFR¹⁴⁴S motif (R-switch), R¹⁸⁵ and N¹⁸⁶ in SD1 (post-Walker A), and ExxxR³³⁴ in SD2 (106). FleQ alone forms dimers, trimers, and hexamers in solution, which is unusual for an EBP; however, addition of c-di-GMP stalls this oligomerization and stabilizes the protein in a dimeric conformation both in the absence and presence of ATP (106). Like FleQ, BrlR also activates transcription in the presence of c-di-GMP. Belonging to the MerR family of transcription activators which binds between the -10 and -35 to shorten the elongated promoter configuration, BrlR upregulates at least two multidrug efflux pumps as well as its own promoter, enhancing antibiotic drug tolerance (112,113). BrlR binds c-di-GMP with a K_d of 2.2 μ M; this stimulates DNA binding and increases and stabilizes the dimeric conformation from the monomeric conformation (113). Crystal structures reveal that there are two binding sites, both located on the N-terminal DNA binding domain (114,115). Upon binding to c-di-GMP, the H-T-H and the flexible coiled-coiled linker undergo twists and bends, altering the spacing and orientation of the DNA-binding domains to enhance DNA binding (114). Important residues in the first binding site include R31, Y40, and R270, which interact with the Hoogsteen edge of the guanine base, stack against the edge of the guanine base, and form hydrogen bonds with the phosphorous group of c-di-GMP, respectively (115). The second binding site is located between two arginine residues (R66 and R86), and binding to c-di-GMP is mediated by a hydrophobic pocket formed by V60, P61, A64, and F93 (115).

Klebsiella pneumoniae MrkH binds to c-di-GMP with a K_d of 2.4 μM to stimulate interactions with the Mrk box of *mrkABCDF*, upregulating type three fimbriae synthesis (116-118). Crystal structures reveal that the MrkH monomer binds to an intercalated c-di-GMP dimer using two PilZ motifs, RxxxR and D/NxSxxG, and a new motif (HSDSGK) in the N-terminal domain (119). Binding does not change the monomeric oligomeric state, but does reveal a large 138° interdomain rotation (119).

Both *Burkholderia cenocepacia* Bcam1349 (also known as BerA) and BerB upregulate transcription of biofilm genes in a c-di-GMP-dependent manner. Belonging to the CRP/FNR family of transcriptional regulators, binding of Bcam1349 to c-di-GMP significantly enhances binding to cellulose and fimbriae synthesis genes as well as the Bcam1330-Bcam1341 gene cluster involved in the synthesis of extracellular biofilm matrix components (120,121). Binding of c-di-GMP to Bcam1349 was crudely estimated to have a K_d of 10 μM (121). On the other hand, BerB belongs to the EBP family of transcription factors and binds c-di-GMP with a K_d of 3 μM (122). While binding of Bcam1349 to c-di-GMP significantly enhanced DNA binding affinity, binding of BerB to c-di-GMP did not alter DNA binding affinity (120-122).

Similar to Bcam1349, Clp is also a c-di-GMP-dependent transcription regulator belonging to the CRP/FNR family. Found in *Xanthomonas campestris*, Clp binds to its promoter DNA in the absence of any ligand (123). However, upon binding to c-di-GMP with a K_d of 3.5 μM , Clp no longer binds the DNA, ceasing transcription of virulence genes (123). Important c-di-GMP binding residues include D70, R154, R156, and D170 (123).

One c-di-GMP-dependent transcriptional regulator has been identified in a Gram-positive bacteria. Controlling the expression of at least 167 genes, *Streptomyces venezuelae* BldD sits at the apex of the regulatory cascade of multicellular progression and development, repressing

sporulation genes during vegetative growth (124,125). Interestingly, the crystal structure revealed that binding of tetrameric c-di-GMP stabilized the dimeric conformation of BldD using the bipartite RXD-X₈-RXXD c-di-GMP interaction signature sequence (126). CTD BldD binds c-di-GMP with a K_d of 2.5 μM, thereby increasing dimerization to subsequently enhance DNA binding around the -10 element to repress gene expression (126). As the only transcription factor that uses a tetrameric c-di-GMP, interestingly, the mechanism of c-di-GMP binding to BldD occurs in a sequential manner in which c-di-GMP dimers first bind to motif 2 (RXXD) and then to motif 1 (RXD) (127).

Belonging to neither Gram positive nor Gram negative families, *Mycobacterium smegmatis* also contains a c-di-GMP-responsive transcription regulator from the TetR-type H-T-H domain family. This regulator, LtmA, is c-di-GMP-dependent and broadly activates 37 lipid transport and metabolism genes (128). With unknown, unidentified, and non-conserved binding motifs, LtmA binds c-di-GMP with a K_d of 0.83 μM, stimulating DNA binding affinity (128).

The number of c-di-GMP-dependent transcription factors is quite small as it is difficult to bioinformatically predict these gene expression regulators. Though c-di-GMP binding motifs among each regulator vary to some degree, similar mechanisms are used to activate and/or inactivate gene expression in the presence and/or absence of c-di-GMP. Except for VpsR, these c-di-GMP-dependent transcription factors use c-di-GMP to form the correct oligomeric state and subsequently enhance or inhibit DNA binding.

1.6 Quorum sensing in *Vibrio cholerae*

Along with c-di-GMP, quorum sensing (QS) is also another method that bacteria use to quickly sense and adapt to environmental changes. These two signaling pathways are intricately

intertwined and regulate many functions including biofilm, motility, and virulence (129). However, unlike c-di-GMP signaling, QS controls gene expression using chemical signals called autoinducers (AIs) to sense the local population density (130). As bacterial population increases, such as seen in high cell density (HCD), AIs accumulate in the environment. A decrease in bacterial population, as seen in low cell density (LCD), leads to a reduction in AI levels.

QS in *V. cholerae* predominantly uses two quorum sensing receptors, CqsS and LuxPQ, as well as two AIs, CAI-1 and AI-2, respectively. CAI-1, (S)-3-hydroxytridecan-4-one, is synthesized by CqsA using two substrates: SAM and decanonyl-coenzyme A (131-133). AI-2 is synthesized by LuxS, which converts the SAM cycle intermediate, S-ribosylhomocysteine, to 4,5-di-hydroxy-2,3-pentanedione (DPD) and homocysteine; DPD then spontaneously converts into AI-2 (134-136). Two parallel circuits with two distinct receptors are utilized by *V. cholerae* to sense the two different AIs (Fig. 1.5).

Under LCD, no AIs are synthesized and as a result, CqsS and LuxPQ, which are histidine kinases, become autophosphorylated. The phosphate is then transferred to the EBP, LuxO (137,138). Together with σ_{54} , phosphorylated LuxO activates transcription of four sRNAs, also known as quorum regulatory sRNAs 1-4 (Qrr sRNAs 1-4) (138). These sRNAs are Hfq-dependent and together activate or repress 20 mRNAs (139-141). For example, they activate mRNA translation of the LCD master regulator, AphA, and inhibit mRNA translation of the HCD master regulator, HapR, leading to upregulation of virulence and biofilm genes and downregulation of Hap protease and other HCD genes, respectively (141). Qrr sRNAs basepair to the 5' untranslated region of the *aphA* mRNA, inducing a conformational change in the mRNA secondary structure and revealing the ribosomal binding site (RBS) (142,143). AphA then activates *tcpPH*, TcpPH

activates *toxT*, and finally, ToxT activates expression of virulence factors. On the other hand, Qrr sRNAs basepair to the RBS of *hapR* mRNA, occluding the RBS from translation (141,144).

The opposite is observed under HCD in which AI concentrations are elevated (Fig. 1.5). High levels of AI inhibit autophosphorylation of CqsS and LuxPQ, leading to activation of the phosphatase domains of these receptors (145). As a result, LuxO is dephosphorylated and inactive, eliminating expression of *qrr*. Without expression of the Qrr sRNAs that bind to the RBS of *hapR* mRNA, *hapR* translation now increases under HCD. Mutual repression is also observed in which HapR represses *aphA* and AphA represses *hapR* (142,146,147).

Integration of the two signaling pathways, QS and c-di-GMP, is seen in *V. cholerae* (Fig. 1.5). Under HCD, HapR is expressed and represses expression of 14 GGDEFs and EALs, leading to decreased c-di-GMP levels and biofilm formation (70). Along with directly decreasing c-di-GMP concentrations under HCD, QS and c-di-GMP together also regulate the expression of the LCD master regulator, AphA, and the biofilm transcriptional activator, VpsT, under LCD. Activation of either *aphA* or *vpsT* requires c-di-GMP and VpsR, which commonly occurs at LCD, while HapR represses at HCD (69,70,146). For both promoters, VpsR and HapR bind to overlapping sites, demonstrating that control of gene expression at these promoters is mutually exclusive (147). Thus, at the *aphA* and *vpsT* promoters, QS and c-di-GMP together fine-tune expression of these genes by functioning as a regulatory checkpoint.

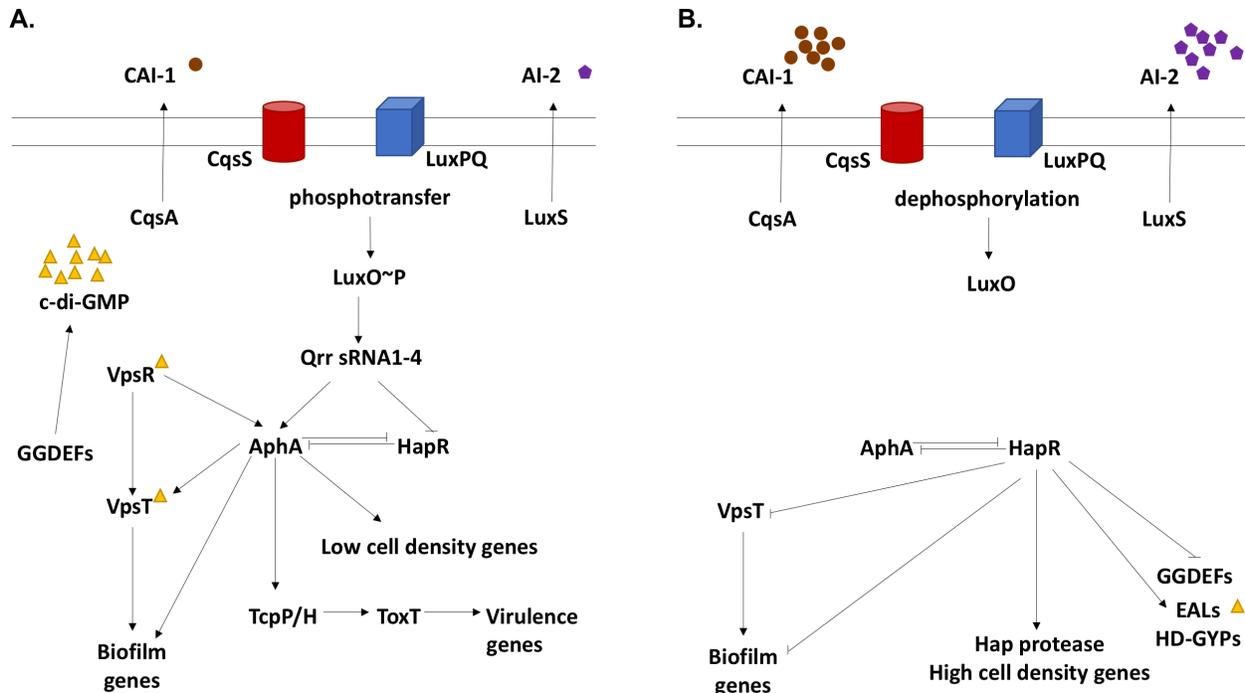


Figure 1.5: Quorum sensing and c-di-GMP signaling in *Vibrio cholerae*. Integration of these two signals in *V. cholerae* occurs at *aphA*. (A) Under low cell density, autoinducers, CAI-1 and AI-2, are not synthesized. This leads to autophosphorylation of CqsS and LuxPQ and their resulting phosphoryl transfer to LuxO, subsequently activating transcription of Qrr sRNA1-4. Together with Hfq, these sRNAs upregulate *aphA* and downregulate *hapR*. AphA then activates expression of biofilm genes and virulence genes and bolsters *hapR* repression. AphA is additionally positively activated by c-di-GMP and the master biofilm regulator, VpsR. (B) At high cell density, CqsA and LuxS synthesize CAI-1 and AI-2, respectively. This leads to activation of the phosphatase activities of CqsS and LuxPQ, leading to the dephosphorylation of LuxO. Without expression of Qrr sRNA1-4, HapR is now expressed, activating Hap protease and high cell density genes as well as inactivating biofilm genes and *aphA*. HapR also represses GGDEF and activates EALs and HD-GYPs, leading to decreased c-di-GMP levels.

1.7 Two-component signal transduction pathways

Finally, the most prevalent and abundant signaling pathway is the two-component system (TCS) transduction pathway that can be found in prokaryotes, archaea, and some eukaryotes. Most bacteria encode for hundreds of TCSs and since they are important in a wide array of phenotypes including cellular survival, development, and virulence, TCSs have been well-studied (148,149).

TCSs are typically comprised of a transmembrane sensor histidine kinase (HK) and its soluble cytoplasmic cognate response regulator (RR). While HKs are comprised of a sensory domain and a transmitter domain containing a dimerization domain with a conserved His residue and an ATPase domain catalyzing phosphorylation, the RRs have a N-terminal receiver domain that accepts phosphorylation linked to a C-terminal effector domain. With over 60 characterized effector domains, these domains can be classified by their enzymatic activities and binding characteristics which include binding to DNA, RNA, ligand, and/or protein. Most RRs are DNA-binding transcription factors regulating gene expression (150). Phosphotransfer or phosphorelay links this separation between the HK and RR (Fig. 1.6).

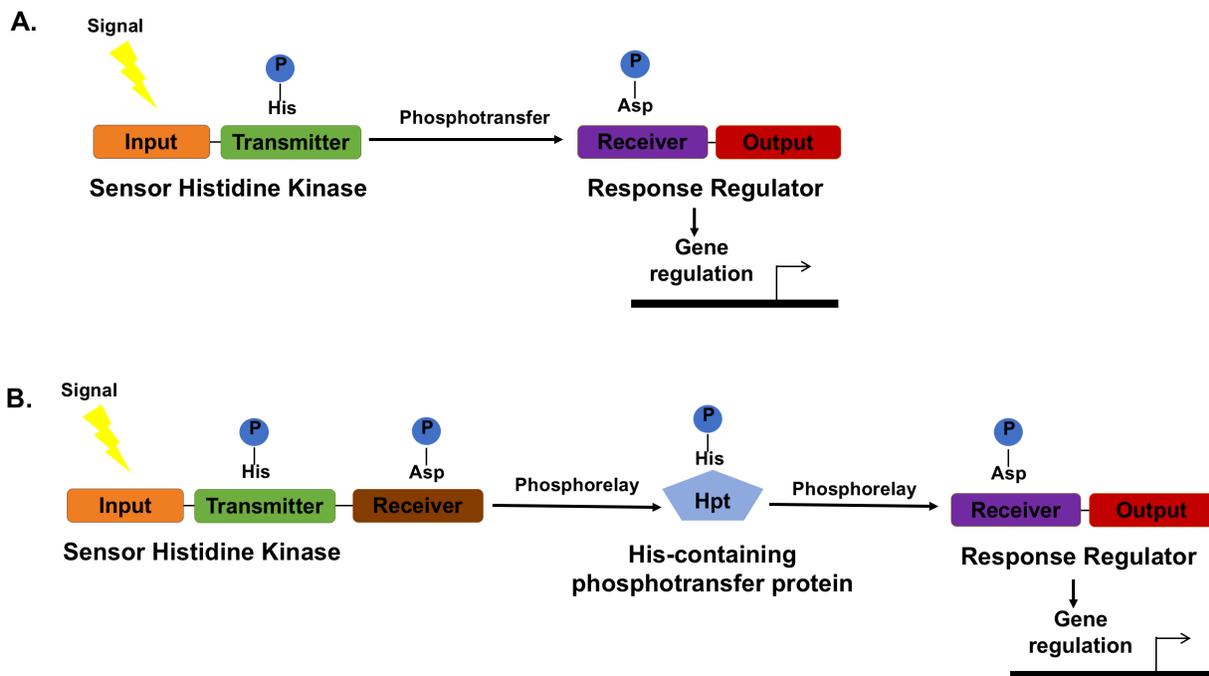


Figure 1.6: Phosphotransmission and phosphorelay schematic in bacterial two-component systems. (A) In phosphotransfer, such as the *E. coli* EnvZ-OmpR two-component system (TCS), the sensor histidine kinase senses a signal leading to autophosphorylation of the histidine residue within the transmitter domain and directly transfers the phosphoryl group to the aspartate on the receiver domain of the response regulator (RR). (B) In phosphorelay, extra phosphoryl transmission steps as well as additional proteins (Hpt: His-containing phosphotransfer proteins) are involved. As characterized by the *B. subtilis* sporulation pathway, the sensor HK senses a

Figure 1.6 (cont'd): signal, autophosphorylates a histidine residue, and proceeds to transfer the phosphoryl group to the aspartate on its own receiver domain. The phosphoryl group is then transferred to the histidine residue of the Hpt protein before finally phosphorylating the aspartate on the receiver domain of the RR.

In response to a specific environmental signal, the SK first autophosphorylates a histidine residue within the transmitter domain. The HK then transfers its phosphoryl group to its cognate RR. This results in a conformational change that alters the activity of the effector domain. *E. coli* EnvZ-OmpR is the most well-characterized TCS utilizing this simple phosphotransfer system. In the phosphorelay scheme, the HK usually contains both the His- and Asp-containing domains. A His-containing phosphotransfer (Hpt) protein is also used to mediate phosphorelay. Thus, phosphorelay requires additional phosphoryl transfer steps as well as proteins. More specifically, the phosphoryl group is transferred from the His to the Asp of the HK, then from the Asp of the HK to the His of the Hpt protein, and then finally from the His of the Hpt protein to the Asp of the RR (151), generating a His-Asp-His-Asp transfer chain. Addition of extra phosphoryl transfer steps and phosphotransfer proteins allow the bacteria to finely sense, amplify, and/or tune signals and modulate gene expression. TCS examples using phosphorelay include *B. subtilis* sporulation genes. Upon phosphorylation of one of four HKs (KinA, KinB, KinC, or KinD), the HK phosphoryl group is first transferred to the RR Spo0F, which then transfers the phosphoryl group to Spo0B (152). Finally, Spo0B transfers that phosphoryl group to Spo0A, regulating various sporulation genes (152).

Alignment of RRs reveals three conserved signature residues involved in phosphorylation: an aspartic acid residue that receives the phosphoryl group from the HK and is located at the end of the third β -strand; and two aspartates or one aspartate and one glutamate, which are involved in Mg^{2+} ion binding and are located within the loop connecting $\beta 1$ and $\alpha 1$ (153). A variety of different RRs can utilize acetyl phosphate as phosphoryl donors, simplifying *in vitro* assays by eliminating

the need to include the cognate histidine kinase (154). Acetyl phosphate is a central metabolite that is the intermediate of the phosphotransacetylase (Pta) and acetate kinase (Acka) pathway. Pta converts acetyl CoA to acetyl phosphate, regenerating coenzyme A and an inorganic phosphate; AckA converts acetyl phosphate to acetate, generating ATP from ADP (Fig. 1.7). Other high energy phosphoryl donors include carbamoyl phosphate and γ -glutamyl phosphate. Many RRs also mimic a nonphosphorylated state upon mutation from an aspartate to an alanine and likewise, mimic a constitutively phosphorylated state upon mutation from an aspartate to a glutamate.

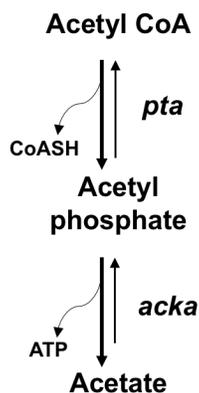


Figure 1.7: Synthesis and degradation of acetyl phosphate. Acetyl phosphate (Ac~P) is formed and broken down in the Acka-Pta pathway. Pta enzyme uses inorganic phosphate and acetyl CoA to form Ac~P and regenerate coenzyme A (CoASH). Acka enzyme then degrades acetyl phosphate into acetate, generating ATP as a byproduct.

Phosphorylation of RRs induces major conformational changes within the RR (153). For many RRs such as those in the OmpR/PhoB family, phosphorylation drives dimerization of the RR, stimulating DNA binding. The receiver domain of *E. coli* PhoB demonstrates the most general form of oligomerization within the OmpR/PhoB family: a 4-5-5 dimer at the α 4- β 5- α 5 surface (155). This oligomerization provides the driving force for DNA binding.

Though it was previously believed that phosphorylation was required for DNA binding and subsequent transcription activation, recent studies reveal that DNA binding does not require phosphorylation at certain promoters. For example, *Bordetella pertussis* BvgA binds to the *fim3* promoter with and without phosphorylation (156). Instead, phosphorylation alters protein-DNA contacts and is required to form the correct competent open transcription complex with RNAP (156). *M. tuberculosis* PhoP also binds to its own promoter in its non-phosphorylated form, but the role of non-phosphorylated vs phosphorylated PhoP has not yet been determined (157,158). In *Salmonella enterica*, the NarL-like regulator, SsrB, requires phosphorylation to activate Salmonella Pathogenicity Island-2 genes via RNAP interactions (159), but does not require phosphorylation to upregulate biofilm genes in both a D56A mutant as well as its HK (SsrA) knockout (160).

Regulation of phosphorylation of RRs must also require dephosphorylation of RRs in order to terminate a specific response. Dephosphorylation rates vary for RRs and typically involve insertion of the amino acid side chain of the phosphatase to initiate nucleophilic attack on the water molecule of the phosphoryl group (153). While some RRs contain dedicated phosphatases, other RRs use their cognate HK's phosphatase activity for dephosphorylation.

1.8 Summary

In conclusion, *V. cholerae* is an important human pathogen that uses multiple signaling pathways, allowing it to adapt to both the aquatic reservoir and the human intestines. It naturally forms biofilms to aid in environmental transmission, survival and persistence. Like many other bacteria, c-di-GMP is the key bacterial second messenger that regulates biofilm formation in *V. cholerae*. Despite decades of c-di-GMP research, the mechanism by which c-di-GMP directly

activates transcription still remains largely unknown. This dissertation investigates this mechanism of regulation by determining how c-di-GMP directly interacts with the *V. cholerae* response regulator, VpsR, to increase biofilm gene expression using a combination of *in vivo* and *in vitro* approaches. A better understanding of the transcriptional mechanism by which c-di-GMP induces biofilm formation is not only important for the development of new antimicrobials against biofilm-based infections via c-di-GMP signaling, but also provide insights for a new paradigm in transcription activation.

CHAPTER 2:

VpsR and c-di-GMP together drive transcription initiation to activate biofilm formation in *Vibrio cholerae*

The findings presented in this chapter have been previously published:

Meng-Lun Hsieh, Deborah M. Hinton, and Christopher Waters (2018). VpsR and c-di-GMP together drive transcription initiation to activate biofilm formation in *Vibrio cholerae*. *Nucleic Acids Research*, 46(17): 8876-8887.

Minor edits have been made to this chapter to conform to dissertation requirements.

Preface

The small molecule cyclic di-GMP (c-di-GMP) is known to affect bacterial gene expression in myriad ways. In *Vibrio cholerae in vivo*, the presence of c-di-GMP together with the response regulator VpsR results in transcription from P_{vpsL} , a promoter of biofilm biosynthesis genes. VpsR shares homology with enhancer binding proteins that activate σ_{54} -RNAP, but it lacks conserved residues needed to bind to σ_{54} -RNAP and to hydrolyze ATP, and P_{vpsL} transcription does not require σ_{54} *in vivo*. Consequently, the mechanism of this activation has not been clear. Using an *in vitro* transcription system, we demonstrate activation of P_{vpsL} in the presence of VpsR, c-di-GMP, and σ_{70} -RNAP. c-di-GMP does not significantly change the affinity of VpsR for P_{vpsL} DNA or the DNase I footprint of VpsR on the DNA, and it is not required for VpsR to dimerize. However, DNase I and $KMnO_4$ footprints reveal that the σ_{70} -RNAP/VpsR/c-di-GMP complex on P_{vpsL} adopts a different conformation from that formed by σ_{70} -RNAP alone, with c-di-GMP, or with VpsR. Our results suggest that c-di-GMP is required for VpsR to generate the specific protein-DNA architecture needed for activated transcription, a previously unrecognized role for c-di-GMP in gene expression.

2.1 Introduction

Biofilm formation and its persistence on catheters, pacemakers, sutures, and other indwelling medical devices account for the vast majority of the two million healthcare-associated annual infections and approximately 100,000 deaths per year in the United States (161). These biofilm-based infections impose an estimated annual \$94 billion in excess medical costs (162). Forming on both biotic and abiotic surfaces, biofilms are aggregates of microbial communities encased by a matrix of extracellular polymeric substances (163). Biofilms, which are formed by almost all bacteria, play a significant role in environmental persistence, dissemination, and transmission as well as protection from environmental stressors such as nutrient limitation, predation, and bacteriophages (56,164-167). However, most concerning of all, biofilms dramatically decrease susceptibility to antimicrobial agents, posing a serious threat to public health.

Because biofilms are recalcitrant to conventional antibiotic therapies and represent a major clinical obstacle, it is essential to understand the molecular mechanisms responsible for biofilm gene expression. A central regulator of biofilm formation is the second messenger cyclic dimeric guanosine monophosphate (c-di-GMP). Present in about 85% of bacteria, c-di-GMP is synthesized by diguanylate cyclases (DGCs), which typically contain a conserved GGDEF motif, and is degraded by phosphodiesterases, which contain a conserved EAL or HD-GYP motif (168). Generally, high levels of c-di-GMP increase biofilm formation and decrease motility, while low levels of c-di-GMP exert the opposite effect (169). Along with biofilm formation and motility, c-di-GMP also regulates a diverse array of phenotypes including quorum sensing, virulence, cell cycle control, secretion, bacterial predation, and stress responses (169). Although c-di-GMP has been extensively studied since its discovery in 1987 (1) and many groups have studied the

mechanisms by which c-di-GMP interacts with effectors (68,106,113,116,119,126-128), mechanism(s) by which c-di-GMP might be needed to directly modulate RNA polymerase (RNAP) in transcription have not been elucidated.

Catalyzing transcription is the multi-subunit enzyme RNAP. Bacterial RNAP is an ~500 kDa enzyme comprised of two large subunits (beta and beta'), two alpha subunits, one omega subunit, and a promoter specificity factor, sigma (σ) (85). Although the primary σ , such as $\sigma 70$ in *Escherichia coli*, is used for the expression of most genes during exponential growth, alternate σ factors, which are either related to $\sigma 70$ or belong to the $\sigma 54$ family, are used under other growth conditions or times of stress (85). The first step in transcription is initiation, a multi-step process that can be controlled by various regulators. During transcription initiation with $\sigma 70$ -RNAP, polymerase first binds to double-stranded DNA elements in the -10 and -35 regions, forming closed complex that is typically unstable (84,91,92). Isomerization to the open complex proceeds rapidly and requires unwinding and bending of the DNA, major conformational changes within RNAP, and formation of the transcription bubble from -11 to ~+3 (84). Upon addition of ribonucleoside triphosphates (rNTPs), the complex transitions to the initiating complex where small abortive RNAs are synthesized and released prior to promoter clearance (84). While RNAP catalyzes transcription efficiently at promoters with optimal -35 and -10 consensus sequences, activators are typically needed to regulate promoters with suboptimal sequences. Some activators additionally use second messenger molecules such as cyclic adenosine monophosphate, guanosine pentaphosphate, or c-di-GMP to modulate gene expression (170).

In *V. cholerae*, an important pathogen that causes the acute diarrhea disease cholera and uses biofilms to aid in environmental transmission, survival, and pathogenesis, VpsR is the master regulator that activates biofilm gene transcription *in vivo* in the presence of high levels of c-di-

GMP and also binds c-di-GMP with a $K_{d(\text{app})}$ of 1.6 μM *in vitro* (53,69-72,171,172). VpsR is known to activate promoters for *vpsL* and *vpsT*, genes within the biofilm biosynthesis operons. Furthermore, VpsR also directly activates expression of other phenotypes in response to c-di-GMP such as acetoin biosynthesis, the transcription factor *tfoY*, and the *eps* operon encoding the type II secretion system (26,69,173), suggesting that this transcription factor is the hub for a central network of c-di-GMP transcriptional control in *V. cholerae*. However, despite the abundance of evidence showing the positive regulatory role of VpsR and c-di-GMP in activating gene expression *in vivo*, previous work has not recapitulated this result *in vitro*.

Based on amino acid sequence homology, VpsR has been classified as an atypical enhancer binding protein (EBP) (69,71). Classic EBPs utilize σ_{54} to activate transcription and are comprised of three conserved domains: an N-terminal receiver (REC) domain, a central AAA+ domain (ATPase Associated with diverse cellular Activities) involved in ATP hydrolysis and binding to σ_{54} , and a C-terminal helix-turn-helix DNA-binding domain (174). Although VpsR has overall homology to EBPs, several residues known to be required for specific EBP functions are not conserved. Not only does VpsR lack the GAFTGA motif involved in binding to σ_{54} , but the highly conserved aspartate (D) and glutamate (E) residues in the Walker B domain involved in ATP hydrolysis are asparagine (N) and aspartate (D) residues in VpsR (Fig. 2.1). Furthermore, microarray analyses demonstrate that transcription from promoters known to be regulated by VpsR does not change in a σ_{54} (*rpoN*-) mutant (72), and sequence analyses indicate that the VpsR-activated promoters do not contain the well-conserved -24 GG and -12 GC consensus sequences utilized by σ_{54} -RNAP. Instead, some of these promoters have reasonable matches to the consensus -10 element of promoters dependent on a primary σ factor, such as σ_{70} .



Figure 2.1: CLUSTAL Omega alignment of the AAA+ domain of VpsR with other bacterial Enhancer Binding Proteins. AAA+ domain of VpsR (VpsRVc) is aligned with other Enhancer Binding Proteins: NtrC from *V. cholerae* (NtrCVc); NtrC from *E. coli* (NtrCEco); HydG from *E. coli* (HydGEco); AlgB from *Pseudomonas putida* (AlgBPpu); and HupR from *Rhodobacter capsulatus* (HupRRca). Asterisks indicate conserved signature residues. The missing GAFTGA motif in VpsR is boxed in red and the nonconserved ND residues within the Walker B domain is boxed in purple. Red indicates hydrophobic residues (A, I, L, M, F, W, V, C), magenta indicates positively charged residues (K, R), blue indicates negatively charged residues (E, D), and green indicates all other residues.

Here we have developed an *in vitro* transcription system demonstrating activated transcription from the VpsR-activated promoter for the *vpsL* gene (P_{vpsL}) in the presence of VpsR, c-di-GMP, and $\sigma 70$ -RNAP. We have used DNase I and $KMnO_4$ footprinting to characterize the protein-DNA complex made by $\sigma 70$ -RNAP alone with P_{vpsL} versus complexes made by $\sigma 70$ -RNAP with VpsR and/or c-di-GMP. Surprisingly, we find that c-di-GMP together with VpsR is needed to generate the correct protein-DNA interactions required for an active transcription complex with $\sigma 70$ -RNAP. Our results provide a new paradigm in c-di-GMP-dependent transcription activation.

2.2 Materials and Methods

DNA.

pMLH06 (*vpsL* short promoter) and pMLH07 (*vpsL* long promoter) contain the *vpsL* promoter from -97 to +213 and from -393 to +213, respectively, cloned into the EcoRI and HindIII restriction enzyme sites of pRLG770 (175). pMLH09 (*vpsL-lux* long promoter) and pMLH10 (*vpsL-lux* short promoter) contain the *vpsL* promoter from -393 to +213 and from -97 to +213, respectively, cloned into the SpeI and BamHI restriction sites of pBBRlux (74). pMLH17 was generated by cloning the wildtype *vpsR* gene into the EcoRI and HindIII restriction sites of pHERD20T (176). Inserts were obtained as PCR products, which had been amplified with primers from *V. cholerae* genomic DNA (BH1514) using Pfu Turbo polymerase (Stratagene). Inserts and vectors were digested with the appropriate restriction enzymes and cloning was performed using standard techniques. Primer sequences are available upon request.

pMLH11 is a pET28b(+) derivative (Novagen) that contains *vpsR* cloned between the NdeI and XhoI restriction sites. PCR was used to amplify the pET28b(+) vector for restrictionless cloning. Gibson Assembly Master Mix (New England Biolabs) was used to assemble the PCR products and vectors according to manufacturer's instructions.

pCMW75 contains an active *V. harveyi* DGC, *qrgB*, and was used for expression of high levels of c-di-GMP in gene reporter assays (70). pCMW98 contains an inactive *V. harveyi* DGC, *qrgB*, and used for expression of low levels of c-di-GMP (70).

Fragments containing P_{vpsL} used for EMSAs and DNase I footprinting were obtained as PCR products using Pfu Turbo polymerase (Stratagene) and upstream and downstream PCR primers, which anneal from positions -97 to +113 relative to the transcription start site (TSS). To radiolabel the DNA, nontemplate or template primer was treated with T4 polynucleotide kinase

(Affymetrix) in the presence of [γ - 32 P] ATP prior to PCR. The radiolabeled PCR products were purified as described (177).

Strains and growth conditions.

E. coli ElectroMAX DH10B (Invitrogen) or *E. coli* DH5 α (New England Biolabs) were used for cloning, and BL21(DE3) or Rosetta2 (DE3) pLysS (New England Biolabs) were used for protein production. The *V. cholerae* strains, $\Delta vpsL$ and $\Delta vpsL \Delta vpsR$, used in this study were derived from the El Tor biotype strain C6707str2 (178). For *lux*-fusion assays, strains containing a mutation in *vpsL* were used to prevent cellular aggregation, allowing us to obtain accurate readings of reporter gene expression at high levels of c-di-GMP by preventing cellular aggregation (70). High c-di-GMP was synthesized by production of the DGC QrgB from the plasmid pCMW75 after addition of 0.5 mM IPTG (70). Strains were grown at 37°C in Luria-Bertani broth (LB) (1% tryptone, 0.5% yeast extract, 1% NaCl at pH 7.5). LB agar medium contained 1.5% (wt/vol) granulated agar (Acumedia). Antibiotics were added at the following concentrations: ampicillin at 100 μ g/mL, chloramphenicol at 100 μ g/mL, and kanamycin at 100 μ g/mL. For *lux* fusion assays shown in Fig. 2.2, *E. coli* S17- λ pir (179) or *V. cholerae* strains containing indicated plasmids were grown overnight in LB medium supplemented with appropriate antibiotics. Cells were then diluted 1:200 in fresh LB supplemented with appropriate antibiotics and grown to an OD₆₀₀ of approximately 0.5. A final concentration of 0.2% arabinose and/or 1 mM IPTG was added to the medium to induce VpsR and/or c-di-GMP synthesis, respectively. Luminescence was measured using an Envision multilabel counter (PerkinElmer) and *lux* expression was reported in relative luminescent units (RLU; counts min⁻¹ mL⁻¹/OD₆₀₀ unit). Assays were repeated with at least two biological replicates and three technical replicates.

Proteins.

E. coli RNAP core was purchased from Epicenter Technologies. *E. coli* $\sigma 70$ was purified as previously described (180). VpsR protein was isolated from Rosetta2 (DE3)/pLysS (Novagen) containing pMLH11, which was grown at 37°C with shaking at 220 rpm in 200 mL of LB containing 50 $\mu\text{g}/\text{mL}$ kanamycin and 25 $\mu\text{g}/\text{mL}$ chloramphenicol to an OD_{600} of approximately 0.5. Cultures were placed on ice, IPTG (final concentration of 0.5 mM) was added, and the cells were then incubated at 16°C with shaking at 150 rpm for 16 hr. After centrifugation at 13,000 x g, cells were harvested and then sonicated in 30 mL of sonication buffer [20 mM sodium phosphate (pH 7.8), 400 mM NaCl, and 7 mM beta-mercaptoethanol]. Centrifugation at 17,500 x g removed insoluble materials, and the soluble fraction was subjected to chromatography on a 1 mL Ni-NTA column (Qiagen). The column was washed with wash buffer [20 mM sodium phosphate (pH 6.0), 400 mM NaCl, and 7 mM beta-mercaptoethanol] containing increasing amounts of imidazole: 0 mM (10 mL), 5 mM (10 mL), 50 mM (5 mL), 100 mM (5 mL), 150 mM (5 mL), 200 mM (5mL), and 250 mM (5 mL). Purified VpsR eluted with fractions containing 150 mM-200 mM imidazole. These fractions were pooled and dialyzed in VpsR buffer [20 mM sodium phosphate (pH 6.0), 150 mM NaCl, 7 mM beta-mercaptoethanol, and 20% glycerol] prior to storage at -80°C. Protein concentrations were determined by comparison with known amounts of RNAP core after SDS-PAGE and gel staining with Colloidal Blue (Invitrogen). To determine if VpsR co-purified with c-di-GMP, 1 μM of VpsR (100 μL) was heated at 95°C for 5 minutes and pelleted by centrifugation. The resulting supernatant was examined using ultrahigh-pressure liquid chromatography (UPLC)–tandem mass spectrometry (MS-MS) as previously described (181).

Thin-layer Chromatography (TLC) ATPase assay

Reactions (2.5 μ l) containing 3 pmol of VpsR, 20 μ M [γ - 32 P] ATP at 2×10^5 dpm/pmol, 50 μ M c-di-GMP (when indicated), and 1 X transcription buffer [40 mM Tris-acetate (pH 7.9), 150 mM potassium glutamate, 4 mM magnesium acetate, 0.1 mM EDTA (pH 7.0), 0.01 mM DTT, and 100 μ g/mL BSA] were incubated for 10 min at 37°C and aliquots (2 μ L) were spotted on polyetherimide (PEI) membranes (Sigma) and allowed to dry. Calf intestinal phosphatase was used as a positive control while buffer alone or BSA were used as negative controls. PEI plates were developed in 0.85M KH_2PO_4 (pH 3.4), autoradiographed, and the images scanned using a Powerlook 2100XL densitometer.

BS³ Crosslinking.

A solution of VpsR buffer containing 1.5 μ M of VpsR, 5 mM of BS³ crosslinker (Thermo Scientific), and as indicated, 50 μ M of c-di-GMP was incubated at room temperature for 30 min. Reactions were quenched with the addition of Tris-Cl (pH 7.5) to a final concentration of 0.1 mM. Proteins were separated by SDS-PAGE on a 10-20% (wt/vol) Tris-tricine gel (Invitrogen) and stained with Colloidal Blue (Invitrogen). To mimic transcription conditions, 5 mM of BS³ was added to a solution containing 0.6 μ M of VpsR and 12.5 μ M of c-di-GMP in transcription buffer. After 30 min of incubation at room temperature, reactions were quenched as described above. Proteins were separated on 10-20% (wt/vol) Tricine gels (Invitrogen) and stained with SilverXpress Silver Stain (Invitrogen) according to manufacturer's instructions.

Electrophoretic Mobility Shift Assays.

Protein-DNA complexes were formed by incubating 5 nM of 32 P-labeled DNA, and as

indicated, VpsR (final concentration from 0.2 μ M to 2 μ M), 0.16 μ M of reconstituted RNAP (σ :core ratio of 2.5:1), and unless indicated otherwise, 50 μ M of c-di-GMP (final volume of 10 to 20 μ l) at 37°C for 10 min in transcription buffer. A 1 μ l solution of 1 mg/mL of poly(dI-dC) or 500 μ g/ml heparin was added to VpsR-DNA complexes or transcription complexes, respectively. Reactions containing VpsR-DNA complexes were loaded onto 5% (wt/vol) nondenaturing polyacrylamide gels already running at 100 V in 1 X Tris/borate/EDTA (TBE) buffer. Samples were electrophoresed for 1.5 h. Transcription complexes were loaded onto 4% (wt/vol) nondenaturing, polyacrylamide gels already running at 100 V in 1 X TBE buffer. After loading, voltage was increased from 100 V to 380 V, and samples were electrophoresed for 3 h. After autoradiography, films were scanned on a Powerlook 2100XL densitometer and analyzed with Quantity One software (Bio-Rad). $K_{d(app)}$ s were calculated as the concentration of VpsR needed to shift 50% of the free DNA.

***In Vitro* Transcriptions.**

Multiple and single round *in vitro* transcriptions were performed in 5 μ L reactions containing 0.02 pmol of supercoiled template, 0 to 3.0 pmol of VpsR, 0 to 50 μ M of c-di-GMP, reconstituted RNAP (0.2 pmol of σ 70 plus 0.05 pmol of core), and transcription buffer. Unless otherwise indicated, samples were incubated at 37°C for 10 min prior to the addition of a solution (1 μ l) containing rNTPs (2.86 mM ATP, GTP CTP, and 71 μ M [α -³²P] UTP at 5 x 10⁴ dpm/pmol) with and without 500 ng heparin. After incubation for 10 min at 37°C, reactions were collected on dry ice, formamide load solution (15 μ L) was added, and aliquots were electrophoresed on 4% (wt/vol) polyacrylamide, 7 M urea denaturing gels for 2500 volt-hrs in 0.5 X TBE buffer. After electrophoresis, gels were exposed to X-ray films, films were scanned, and radioactivity was

quantified as described above.

Primer Extensions.

Primer extension analyses of RNA generated *in vitro* were performed according to manufacturer's instructions (Promega) by using AMV Reverse Transcriptase. A sample (5 μ L) of the *in vitro* transcription reaction was added to 6 μ L of primer mixture containing 2 X AMV Primer Extension buffer and 2 pmol of 32 P-labeled primer, which annealed +103 bp downstream of the +1 transcriptional start site. Aliquots were electrophoresed on 8% (wt/vol) polyacrylamide, 7 M urea denaturing gels for 4000 volt-hours in $\frac{1}{2}$ X TBE. Densitometry and quantification were performed as described above.

In vivo RNA was obtained from WN310 containing pMLH17 and pCMW75 or pCMW98 grown in LB with 100 μ g/mL ampicillin and 100 μ g/mL kanamycin at 37° C with shaking at 220 rpm to an OD₆₀₀ of approximately 0.5. Cells were harvested by centrifugation and RNA was extracted using the RNeasy kit (Qiagen), and an on-column DNase I digestion (Qiagen) was performed according to manufacturer's instructions. After elution, 5 μ g of total RNA in 5 μ L was added to the primer mixture and subsequent steps for primer extension reactions were performed as described above.

DNase I footprinting.

Solutions were assembled as described for EMSAs using 0.04 μ M DNA, and as indicated, 1.4 μ M VpsR, 50 μ M c-di-GMP, and/or 0.16 μ M of reconstituted RNAP (σ :core ratio of 2.5:1). After incubation with poly(dI-dC) (complexes lacking RNAP) or heparin (complexes with RNAP) for 15 s, 0.3 U of DNase I in 1.5 μ L was added. Solutions were incubated for an additional 45 s at 37° C and then immediately loaded onto 4% (wt/vol) nondenaturing, polyacrylamide gels already

running at 100 V in 1 X TBE buffer. Upon loading samples, voltage was increased from 100 V to 380 V, and samples were electrophoresed for 3 h. After autoradiography, the protein/DNA complexes were excised and extracted DNA was electrophoresed on denaturing gels as described (156).

Potassium Permanganate Footprinting.

For potassium permanganate (KMnO_4) footprinting, solutions were assembled as described for DNase I footprinting. After addition of 500 ng of heparin, KMnO_4 was added to a final concentration of 2.5 mM, solutions were incubated at 37°C for 2.5 min, quenched with 5 μL of 14 M 2-mercaptoethanol, and further processed as described (156).

2.3 Results

In the presence of c-di-GMP, VpsR activates σ 70/RNAP at the *vpsL* promoter (P_{vpsL}) *in vitro*.

It has previously been demonstrated in *V. cholerae* that deletion of *vpsR* eliminates expression of a P_{vpsL} -driven *lux* (71,182) and that high levels of c-di-GMP yield greater levels of P_{vpsL} -*lux* transcription (70). Thus, we sought to analyze the effects of both VpsR and c-di-GMP at P_{vpsL} using gene reporter fusion assays. We find that the presence of VpsR and c-di-GMP activates P_{vpsL} -*lux* expression in either *V. cholerae* (Fig. 2.2A) (69,70,182) or *E. coli* (Fig. 2.2B).

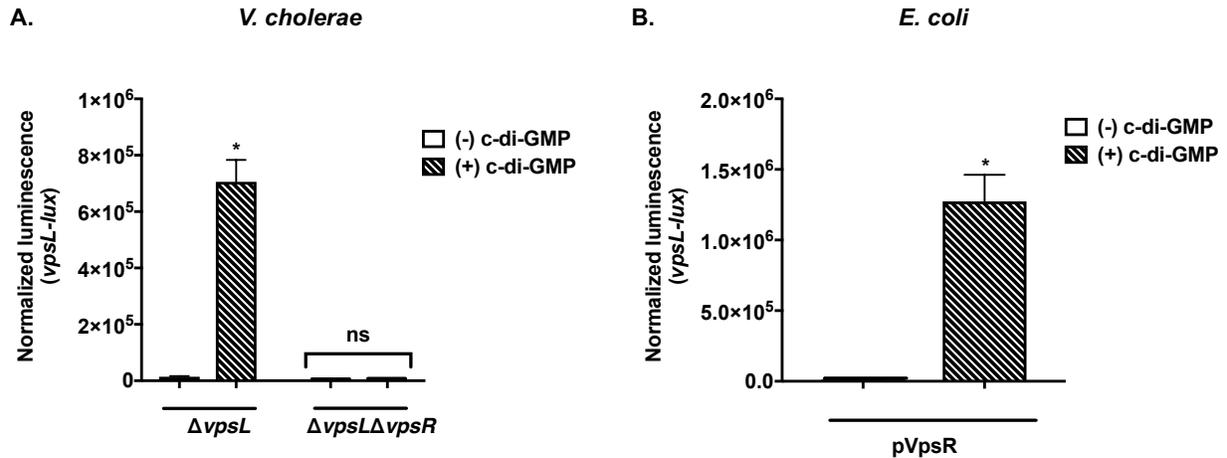


Figure 2.2: *vpsL-lux* gene reporter assays in *Vibrio cholerae* and *Escherichia coli* S17-lpir. Both (A) *V. cholerae* and (B) *E. coli* strains contained pMLH10 (*vpsL-lux* short) and either pCMW98 (for low levels of c-di-GMP) or pCMW75 (for high levels of c-di-GMP), while the *E. coli* strain contained an additional plasmid, pMLH17, for VpsR overexpression. Error bars indicate standard deviation of at least three independent cultures analyzed by one-way ANOVA with Tukey's HSD *posthoc* analysis (* $p < 0.05$; ns, not significant).

Thus, no specific *Vibrio* factors other than VpsR are required for activation. Consequently, we established an *in vitro* transcription system and performed various footprinting assays using *E. coli* RNAP. These analyses are detailed below and summarized in Fig. 2.3. It should be noted that the transcription start site (TSS), which we determined by primer extension analyses of both *in vivo*

RNA isolated from *V. cholerae* and *in vitro* RNA described below, differs from previously reported locations (69,71).

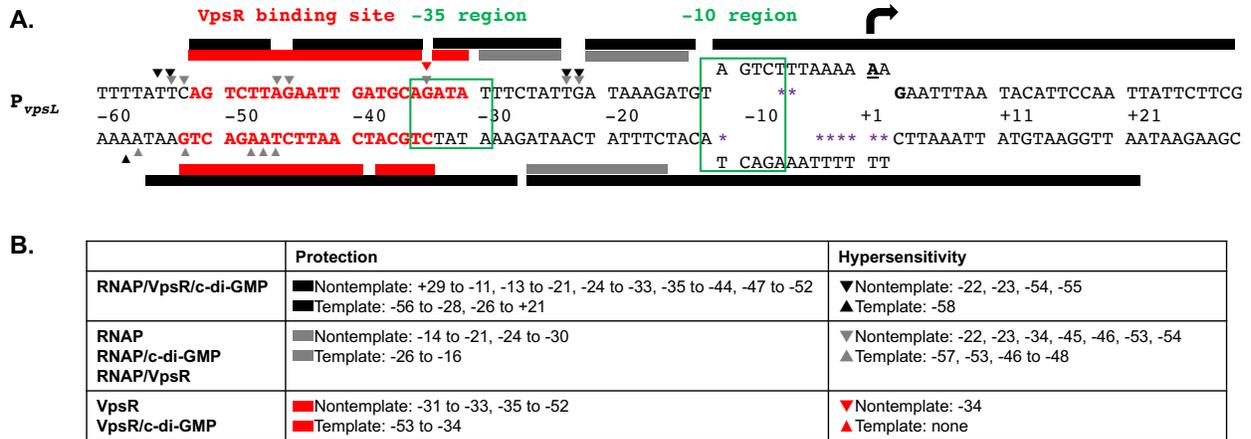


Figure 2.3: Summary of *in vitro* primer extension and DNase I and KMnO₄ footprinting at P_{vpsL}. (A) Sequence of P_{vpsL} from -60 to +30. Bold and underlined A with black arrow at +1 and bold G (+3) represent the transcription start sites determined by primer extensions; the -10 element and the -35 region are labeled and boxed in green; sequences in bold and red denote the VpsR binding site. Protection sites from DNase I footprinting and hypersensitivity sites are depicted as rectangular boxes and triangles, respectively, either above (nontemplate) or below (template) the sequences: grey, RNAP with or without c-di-GMP or VpsR; black, RNAP with VpsR and c-di-GMP; red, VpsR with or without c-di-GMP. The open transcription bubble detected using KMnO₄ footprinting is shown as separated ss DNA from position -11 to +2 with sites of KMnO₄ cleavage indicated as purple asterisks. (B) Summary of positions of protection and hypersensitivity sites on nontemplate and template strand DNA.

As a transcription template, we constructed pMLH06, which contains *V. cholerae* DNA from -97 to +213 relative to the *vpsL* TSS inserted upstream of the *rrnBT1* terminator. Previous EMSAs and DNase I footprinting analyses indicated that VpsR binds to the promoter for *vpsL* (P_{vpsL}) at both a promoter distal site (-297 to -336) and a promoter proximal site (-31 to -52). However, promoter-fusion expression studies have demonstrated that the downstream site is sufficient for activation (69,71).

Using *E. coli* core RNAP reconstituted with $\sigma 70$, we found that the presence of VpsR and c-di-GMP activates transcription at P_{vpsL} by ~ 7-fold (Fig. 2.4). Both VpsR and c-di-GMP are

required for this activation. Addition of varying amounts of either c-di-GMP alone or VpsR alone to RNAP does not alter the basal level observed with RNAP alone, while addition of both c-di-GMP and VpsR to RNAP results in a dose-dependent activation (Figs. 2.4 and 2.5). We also determined that only the downstream VpsR binding site is required for this activation *in vitro* (Fig. 2.6), consistent with the results obtained with *lux*-fusion assays (Fig. 2.2) as well as other studies (69,71).

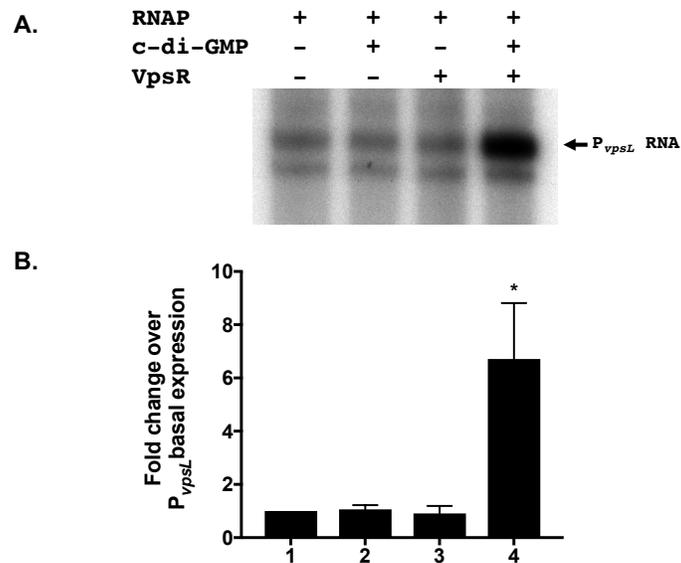


Figure 2.4: VpsR and c-di-GMP activate transcription at P_{vpsL} by approximately 7-fold *in vitro*. (A) Representative gel showing P_{vpsL} RNA obtained after multiple round *in vitro* transcription reactions using plasmid template P_{vpsL} with RNAP alone (lane 1), RNAP and c-di-GMP (lane 2), RNAP and VpsR (lane 3), and RNAP, VpsR, and c-di-GMP (lane 4). (B) Graph showing the level of P_{vpsL} transcription relative to that with RNAP alone (basal) obtained from at least three independent experiments (one-way ANOVA with Tukey's HSD (honest significant difference) *posthoc* analysis, *p < 0.05).

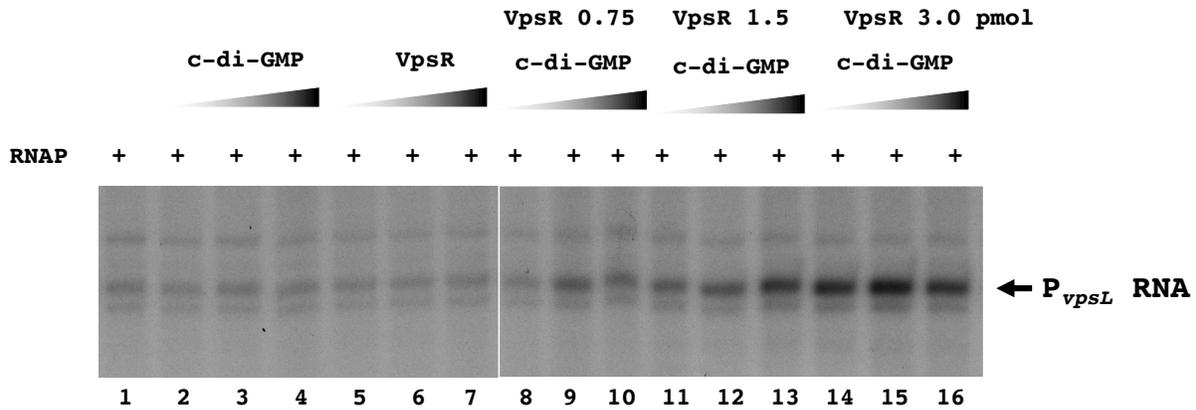
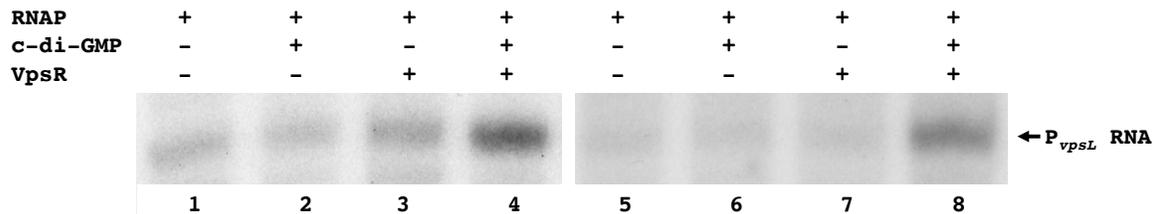


Figure 2.5: Increasing c-di-GMP alone or VpsR alone has no effect on basal level transcription activity *in vitro*. 0, 62.5, 125, or 250 pmol of c-di-GMP or 0, 0.75, 1.5, 3 pmol of VpsR were used. A representative gel of P_{vpsL} RNA obtained after single round transcription reactions with RNAP and the indicated components is shown.

A.



B.

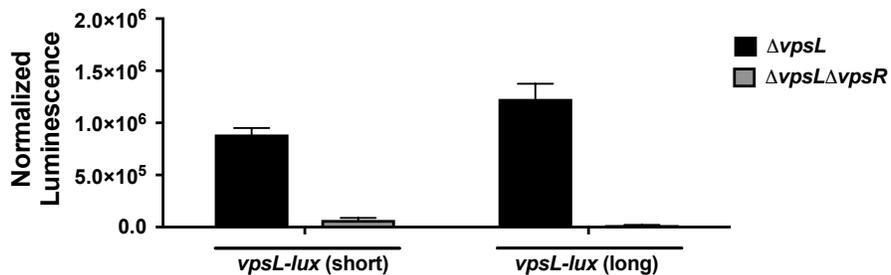


Figure 2.6: VpsR only needs the proximal binding site to activate transcription at P_{vpsL} . (A) Single round *in vitro* transcription reactions were performed with pMLH06 (*vpsL* short promoter containing the proximal VpsR binding site) (lanes 1-4) or with pMLH07 (*vpsL* long promoter containing both the distal and proximal VpsR binding sites) (lanes 5-8). (B) *vpsL-lux* gene reporter assays in *V. cholerae*. Strains contain pCMW75 to induce high levels of c-di-GMP and either pMLH10 (*vpsL-lux* short, P_{vpsL} with the proximal VpsR binding site) or pMLH09 (*vpsL-lux* long, P_{vpsL} with both the proximal and distal VpsR binding sites).

Primer extension analyses and KMnO₄ footprinting identifies the TSS of P_{vpsL}.

Previous primer extension analyses of *V. cholerae* RNA isolated from cells have identified multiple 5' ends for the RNA occurring upstream of the *vpsL* coding sequence. These included an A, a T (most abundant), a G, and an A nucleotide, located 37, 39, 57, and 59 bases upstream of the assigned GUG translation start site, respectively (71,76). However, these positions were determined using exponentially growing *V. cholerae*, which should have low levels of c-di-GMP. To determine the TSS *in vitro*, we performed primer extensions using RNA synthesized from our *in vitro* transcription reactions. This analysis identified two 5' ends whose presence is stimulated when reactions contain RNAP together with both VpsR and c-di-GMP (Fig 2.7A, lane 9): the 'A' located 59 bases upstream of the *vpsL* GUG, which was one of the ends observed previously and is indicated as the +1 in Fig. 2.3, and the 'G' located 57 bases upstream of the GUG. To assign the TSS in the presence of high levels of c-di-GMP *in vivo*, we isolated RNA from *V. cholerae* grown with high intracellular concentrations of c-di-GMP. We again observed these two sites as well as other downstream 5'-ends (Fig. 2.7B, lane 4). Given that the farthest upstream end(s) seen *in vivo* align with start sites seen *in vitro* and the TSS determined by KMnO₄ footprinting (below), we propose that the farthest upstream sites are in fact the start of the *vpsL* RNA and the other ends observed *in vivo* arise from RNA processing or degradation.

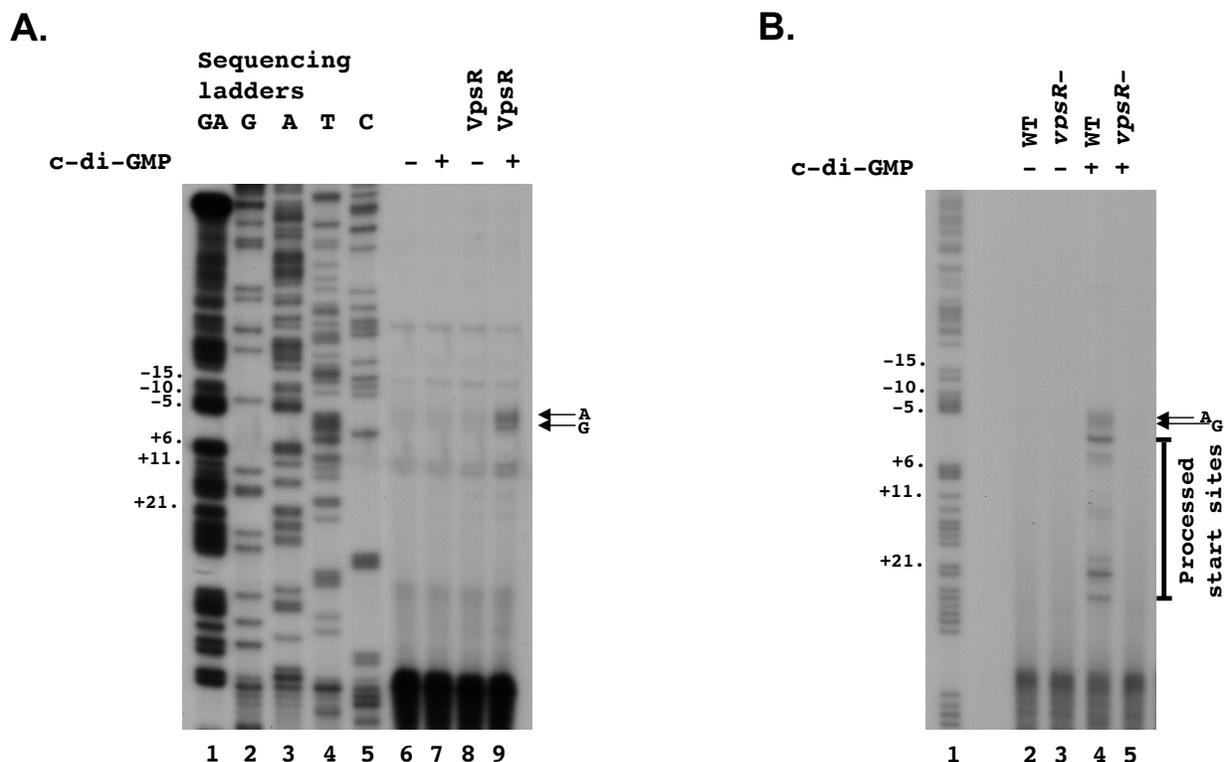


Figure 2.7: Identification of $P_{vpsL} +1$ TSS using *in vitro* and *in vivo* primer extensions. RNA was isolated from *in vitro* transcriptions (A) or *V. cholerae* (B). Two major primer extension products, which are observed only in the presence of both VpsR and c-di-GMP are indicated with arrows.

In a transcription open complex, the single-stranded (ss) transcription bubble typically occurs from -11 to $\sim +3$ (91,92). KMnO_4 footprinting, which selectively oxidizes thymines in ss DNA, is considered the ‘gold’ standard for observing the position of this transcription bubble and by extension the position of the +1 TSS (95). In this analysis, the ‘T’ at position -11 on the template strand marks the start of the ss region of the DNA within the open complex and is thus the farthest upstream reactive T in the analysis. Reactive thymines can extend to the end of the bubble (position $\sim +3$). In our assay, we challenged complexes with the addition of heparin, which typically destabilizes closed complexes but does not impact open complexes. Thus, we conclude that any oxidized thymines we are observing arise from the stable open complex.

When we incubated P_{vpsL} with RNAP/VpsR/c-di-GMP, we observed reactive thymines on the template strand at positions -11 and -4 to +2 (Fig. 2.8A, lane 2) and reactive thymines on the nontemplate at positions -6 and -7 (Fig. 2.8B, lane 2) relative to the 'A' that is 59 bases upstream of the GUG. These reactive bases identify the open complex and are consistent with our identification of the transcription start site at position -59 relative to start of the gene. Furthermore, this analysis defines the σ^{70} -10 recognition element of P_{vpsL} as $^{-12}\text{TAGTCT}^{-7}$.

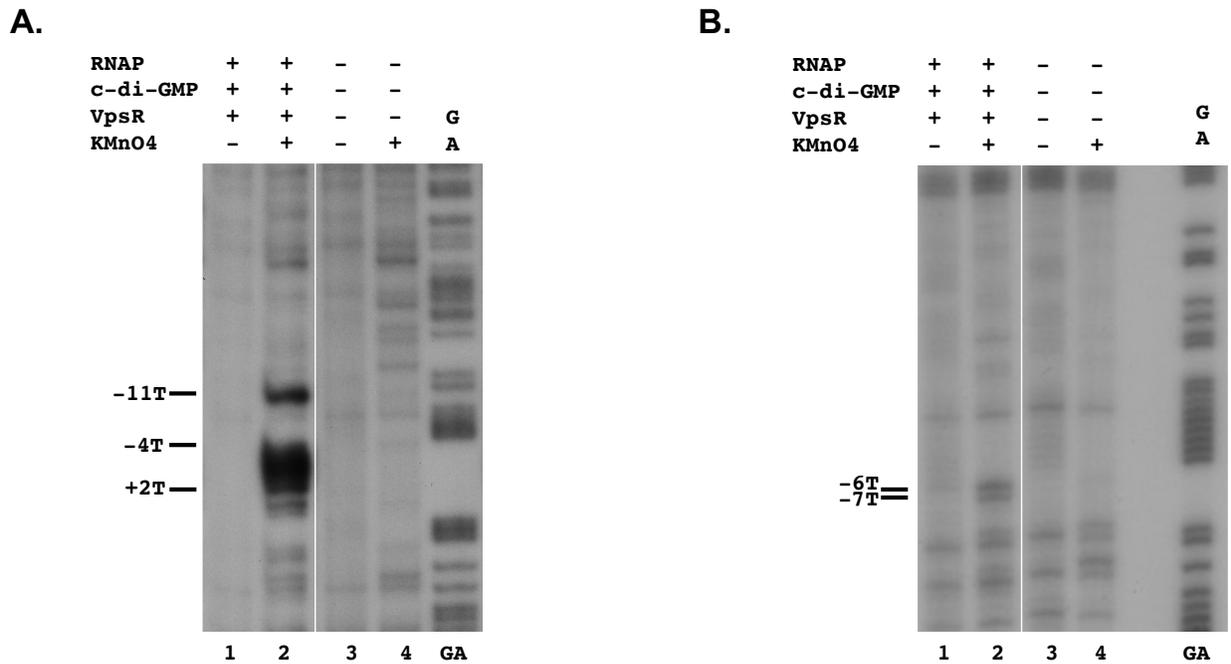


Figure 2.8: KMnO₄ footprinting assigns the +1 TSS at the A located 59 bp upstream of the *vpsL* translation start site. Reactive thymines within the transcription bubble are observed at positions -11, -4, -3, -2, -1, +1, and +2 on template DNA (A). Reactive thymines within the transcription bubble are also observed at positions -6 and -7 on nontemplate DNA (B). GA corresponds to G+A ladder.

We also used KMnO₄ footprinting to investigate open complex formation when only some of the components are present. As expected from the basal transcription that we observed with RNAP alone (Fig. 2.4), we observed a low level of reactive thymines at the same positions in the complexes made by RNAP in the absence of c-di-GMP and VpsR (Fig. 2.9, lane 2). Addition of

either VpsR or c-di-GMP to RNAP did not stimulate this basal level of transcription (Fig. 2.4) or the amount of reactive thymines (Fig. 2.9, lanes 4 and 6). Thus, KMnO₄ analyses indicate that RNAP together with both VpsR and c-di-GMP is needed to form the maximum level of open complex a P_{vpsL}, consistent with our *in vitro* transcription results (Fig. 2.4). Interestingly, these analyses also indicated the presence of reactive thymines at other positions within the basal complexes (Fig. 2.9, lanes 2, 4, and 6), which were relatively much less abundant in the activated complex (Fig. 2.9, lane 8). It is possible then that in the absence of both VpsR and c-di-GMP, RNAP is promiscuous in promoter choice utilizing different start sites such as the ones previously reported (71,76).

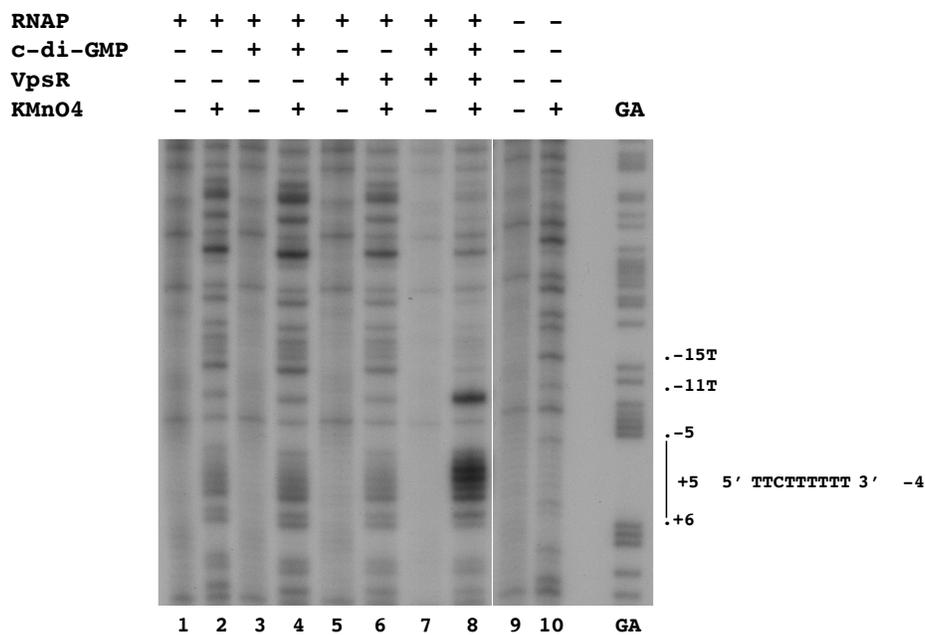


Figure 2.9: KMnO₄ footprinting of transcription complexes at P_{vpsL} on template DNA. Reactions were assembled with the indicated components and P_{vpsL} DNA. Cleavages are seen at positions -11, -4, -3, -2, -1, +1, and +2 on template DNA, indicative of an open transcription bubble. The sequence from +5 to -4 is indicated. GA corresponds to G+A ladder.

Nevertheless, it is important to note that RNAP, VpsR, and c-di-GMP form an open complex in the absence of ATP. This is unlike classic enhancer binding proteins which use ATPase activity to

drive open complex formation. To investigate whether VpsR has ATPase activity, we assayed ATP hydrolysis in the presence and absence of c-di-GMP (Fig. 2.10, lanes 3 and 4). No ATPase activity was detected.

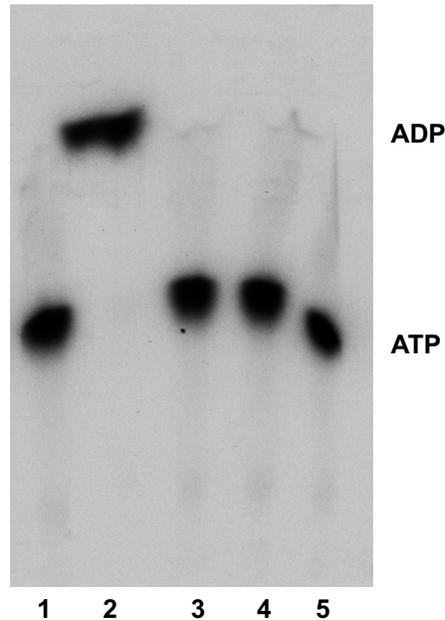


Figure 2.10: VpsR does not hydrolyze ATP. Reactions containing buffer alone (lane 1), 3 pmol of calf intestinal phosphatase as a positive control (lane 2), 3 pmol of VpsR without and with c-di-GMP (lanes 3 and 4, respectively), or 3 pmol BSA as a negative control (lane 5) in 1 X transcription buffer were incubated with $[\gamma\text{-}^{32}\text{P}]\text{ATP}$ for 10 min at 37°C, and products were separated via thin-layer chromatography on PEI membranes. The positions of ATP and ADP are indicated.

In the presence of VpsR and c-di-GMP, RNAP rapidly forms a heparin-resistant, stable complex at P_{vpsL} .

To determine the rate of P_{vpsL} open complex formation in the presence of RNAP, VpsR, and c-di-GMP, we incubated P_{vpsL} DNA with these components for various times before adding heparin together with rNTPs. As the addition of heparin will destabilize any unstable complexes, the subsequent level of single round transcription reflects the level of stable complex present at that time point. As seen from our KMnO_4 analyses (Figs. 2.8 and 2.9), this represents the open complex.

As seen in Fig. 2.11A, no additional activation was observed after the first time point of 1 min, indicating that the open complex forms rapidly in the presence of VpsR and c-di-GMP.

To determine the stability of the open complex, we incubated proteins and DNA for 10 min and then added heparin for varying time periods prior to the addition of rNTPs. Either a one min or 15 min heparin incubation yielded similar levels of transcription (Fig. 2.11B), indicating that the open complex is stable for at least several min. The stability of open complex to heparin inhibition was not dependent upon the addition of c-di-GMP or VpsR as basal expression exhibited equivalent stability to heparin addition at one versus 15 min although total transcript levels were reduced. Thus, we conclude that RNAP rapidly forms a stable open complex at the *vpsL* promoter and the amount of open complex formation is stimulated by VpsR and c-di-GMP.

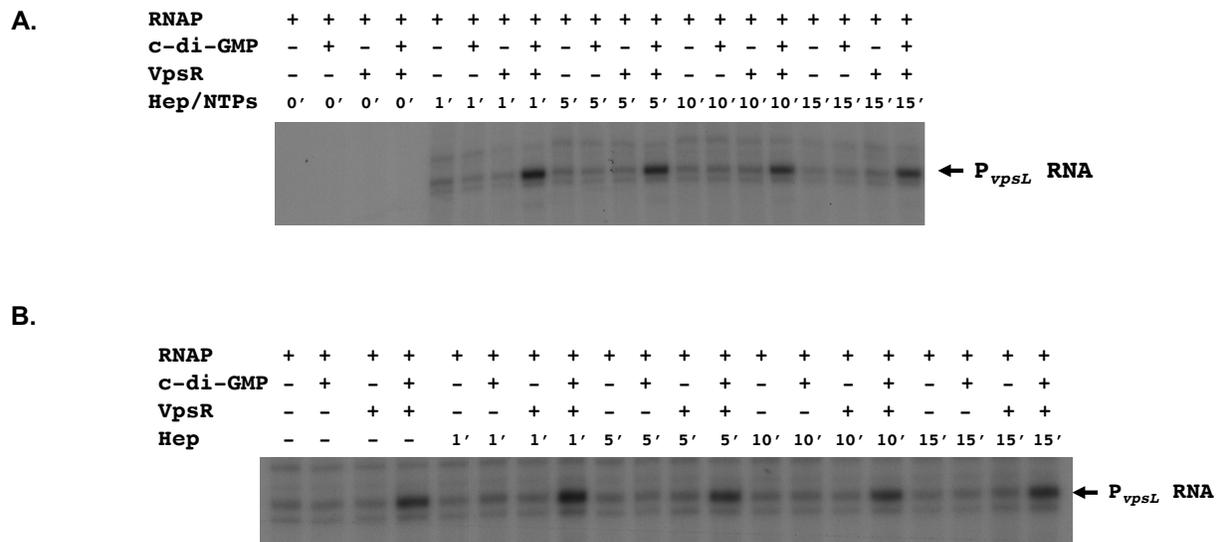


Figure 2.11: VpsR, c-di-GMP, RNAP, and *P_{vpsL}* DNA rapidly form active and heparin-resistant transcription complexes in single round *in vitro* transcriptions. Transcription complexes were formed for 0, 1, 5, 10, or 15 minutes prior to simultaneous addition of heparin and rNTPs (A). Open complexes were challenged with heparin for 0, 1, 5, 10, or 15 min prior to the addition of rNTPs (B). Each assay was performed at least three times. A representative gel of *P_{vpsL}* RNA is shown.

Dimerization and DNA binding by WT VpsR with and without c-di-GMP are similar.

Multiple studies have found that many transcription factors require c-di-GMP binding to facilitate dimerization or the formation of higher order structures (68,106,119,126). Consequently, we used BS³ crosslinking, which generates nonspecific amine to amine covalent bonds, to investigate whether VpsR forms oligomers and if so, whether this formation is affected by c-di-GMP. As seen in Fig. 2.12A, WT VpsR dimers are observed in the presence of BS³ crosslinker, and the amount of this crosslinked species is similar in the presence or absence of c-di-GMP. To make sure that our transcription conditions did not affect these results, we also tested the BS³ crosslinking using the same protein concentration and buffer conditions that were used for transcription. In this case, the presence of Tris buffer, which quenches the crosslinking reaction, reduces the overall amount of crosslinking, but again there is no significant difference in the presence or absence of c-di-GMP (Fig. 2.12B). On these silver stained gels, three crosslinked bands are observed with and without c-di-GMP. We infer that these bands represent different crosslinked VpsR dimer conformations using different amines since BS³ is a nonspecific amine to amine crosslinker. These results indicate that unlike other characterized c-di-GMP binding proteins (68,106,119,126), VpsR does not require c-di-GMP to dimerize. However, it is possible that the presence of c-di-GMP could change the conformation of the formed dimers.

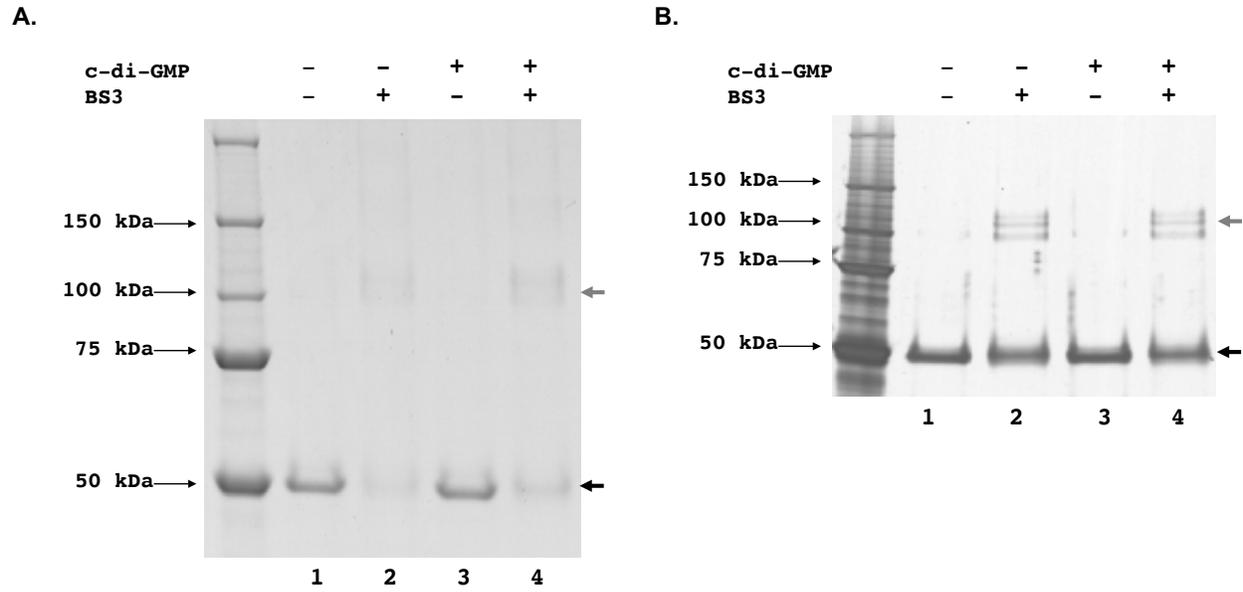


Figure 2.12: VpsR forms dimers *in vitro* with or without c-di-GMP. (A) Samples containing 1.5 μM of VpsR with and without 50 μM c-di-GMP were treated with the chemical cross-linker BS³ as indicated and separated on a 10-20% (wt/vol) Tricine gel that was stained with Colloidal Blue. (B) Samples containing 0.6 μM of VpsR with and without 12.5 μM c-di-GMP in transcription buffer were treated with the chemical cross-linker BS³ as indicated and separated on a 10-20% (wt/vol) Tricine gel that was silver-stained. Far left lane of each panel contains marker proteins, whose molecular weights are indicated. Black arrows indicate position of VpsR monomer (~50kDa) and grey arrows indicate position of VpsR dimer (~100kDa). Each sample was repeated independently three times, and a representative gel image is shown.

Along with oligomerization, c-di-GMP also plays an important role in helping transcription factors bind the DNA (68,106,113,116,119,126-128). Thus, we asked whether the ability of c-di-GMP to stimulate open complex formation could arise by promoting the interaction of VpsR with the DNA. We tested this possibility by determining the apparent dissociation constant ($K_{d(\text{app})}$) for VpsR binding to P_{vpsL} in the presence or absence of c-di-GMP. $K_{d(\text{app})}$ s were calculated by determining the concentration of VpsR needed to shift 50% of the free DNA. We found that the presence of c-di-GMP did not enhance VpsR binding to the DNA [2.22 μM (+/- 0.64 μM) without c-di-GMP, 1.66 μM (+/- 1.00 μM) with c-di-GMP (Fig. 2.13)], indicating that c-di-GMP does not alter the affinity of VpsR for the DNA. Our results are consistent with a previous study that did

not observe differences in VpsR binding to *vpsL* in the presence or absence of c-di-GMP using EMSAs, though these experiments were only done at one concentration of VpsR and dissociation constants were not measured (71).

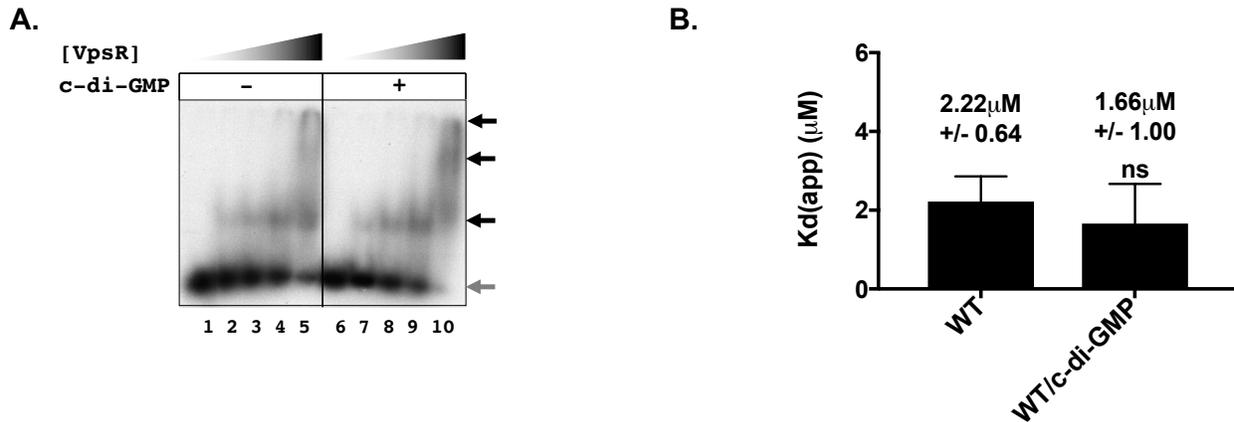
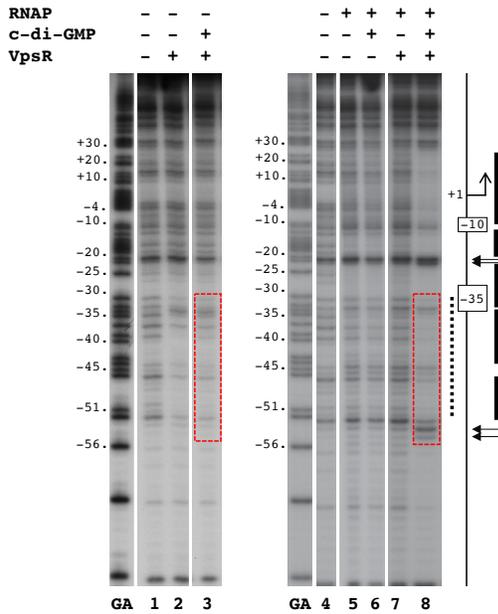


Figure 2.13: VpsR binds P_{vpsL} DNA with similar affinity with or without c-di-GMP. Representative gels showing the retardation of ^{32}P -labeled DNA harboring -97 to +103 of P_{vpsL} with increasing VpsR concentrations from 0 μM to 2 μM either in the absence (lanes 1-5) or presence (lanes 6-10) of 50 μM c-di-GMP. Black arrows indicate retarded complexes while grey arrow indicates free DNA. (B) Quantitation of EMSAs. Apparent DNA-binding dissociation constants ($K_{d(\text{app})}$) were calculated as the concentration of VpsR needed to retard 50% of the free DNA. Values from at least three EMSAs were analyzed using one-way ANOVA with Tukey's HSD *posthoc* analysis (ns, not significant).

We also performed DNase I footprinting to determine whether there are different VpsR-DNA contacts in the presence or absence of c-di-GMP. To make sure that we were only observing the footprint of the stable protein complex of interest, we challenged the complexes with poly(dI-dC), treated them with DNase I, and then isolated the complexes from EMSA gels before isolating the DNA. Similar to a previous study, which identified the proximal VpsR binding site from -31 to -52 using nontemplate P_{vpsL} in the absence of c-di-GMP (71), we found that VpsR with or without c-di-GMP protected the DNA from -31 to -52 on nontemplate P_{vpsL} and from -34 to -53 on template P_{vpsL} (Fig. 2.14A and B, lanes 2 and 3). Thus, we did not observe any significant differences between the VpsR-DNA contacts whether c-di-GMP was present or absent.

A. Nontemplate



B. Template

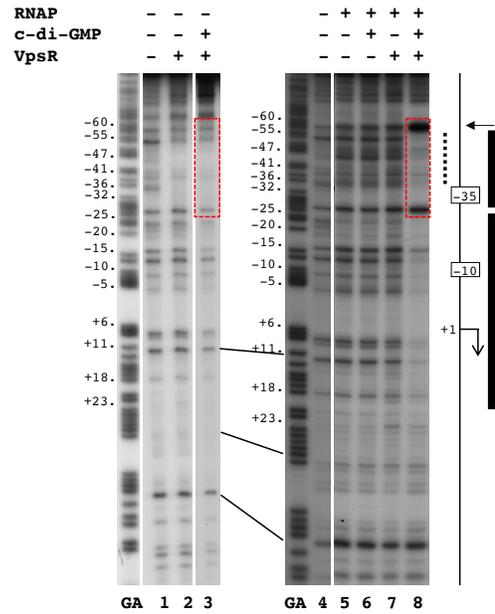


Figure 2.14: DNase I footprinting of P_{vpsL} complexes on nontemplate DNA and template DNA. GA corresponds to G+A ladder. VpsR, c-di-GMP, and/or RNAP are present as indicated. To the right of each gel image, a schematic indicates the -10 and -35 regions and the +1. The VpsR binding site is indicated as a dashed black line. DNase I protection regions and hypersensitivity sites seen with the activated complex of RNAP, VpsR, c-di-GMP, and DNA are depicted as black rectangles and horizontal arrows. The dashed red boxes indicate the regions of DNA where the protection/enhancement within and immediately adjacent to the VpsR binding site changes when comparing complexes containing RNAP with or without VpsR or c-di-GMP to the activated complex.

It is important to note that in the previous study, DNase I footprinting was performed in the absence of any competitor (71), while here we used poly(dI-dC) and isolated footprinting complexes from EMSA gels. We observed no protection when complexes made with VpsR +/- c-di-GMP were challenged with heparin (data not shown), even though the VpsR/RNAP/c-di-GMP/ P_{vpsL} transcription complex is stable to heparin challenge (Fig. 2.11B). We conclude that the presence of RNAP with c-di-GMP stabilizes VpsR binding to the DNA, forming a heparin-resistant complex.

DNase I footprinting analyses suggest that c-di-GMP is needed to form the active transcriptional protein/DNA architecture at P_{vpsL} .

While RNAP alone, RNAP/c-di-GMP, and RNAP/VpsR all yield basal transcription from P_{vpsL} , activated transcription requires RNAP, VpsR, and c-di-GMP. To investigate whether the protein-DNA interactions differed between the basal and activated transcription complexes, we performed DNase I footprinting. Again, we challenged the complexes with heparin and extracted the stable complexes from EMSA gels before isolating the DNA (Fig. 2.15) to ensure that we were observing contacts made within the stable open complex. The observed protection patterns and hypersensitive sites are summarized in Fig. 2.2.

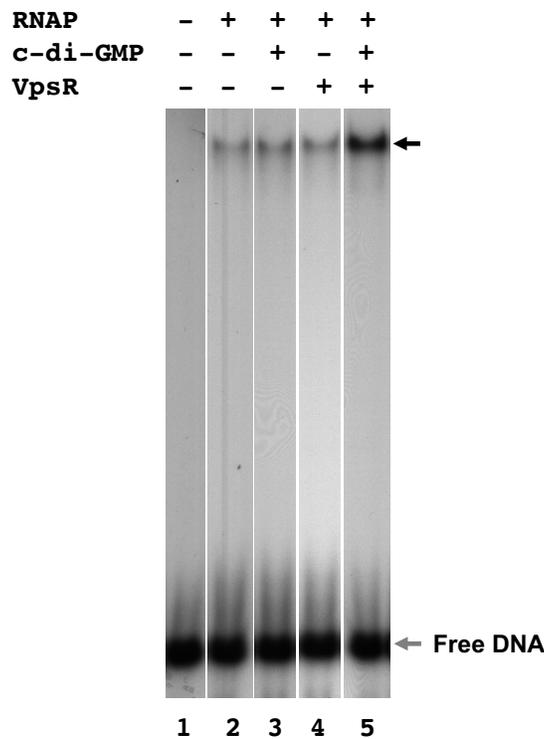


Figure 2.15: Discrete P_{vpsL} /RNAP complexes are formed in EMSAs. Native polyacrylamide gel showing the complexes isolated for DNase I footprinting by incubation of ^{32}P -end labeled nontemplate P_{vpsL} fragment from -97 to +103 with buffer alone (lane 1), RNAP (lane 2), RNAP and c-di-GMP (lane 3), RNAP and VpsR (lane 4), or RNAP and VpsR and c-di-GMP (lane 5) in 1 X transcription buffer. Grey arrow indicates free DNA while black arrow indicates discrete complexes with RNAP.

DNase I footprints of basal complexes formed with RNAP alone, RNAP/c-di-GMP, or RNAP/VpsR at P_{vpsL} were similar. On the nontemplate strand, protection was observed from -14 to -21 and -24 to -30 with hypersensitive sites at -22, -23, -34, -45, -46, -53, -54 (Fig. 2.14A, lane 5-7). On the template strand, protection was present from -26 to -16 with hypersensitive sites at -57, -53, and -46 to -48 (Fig. 2.14B, lane 5-7). Because the $KMnO_4$ footprinting (detailed above) indicated the presence of an open bubble in these heparin resistant complexes, we conclude that these are the contacts present within open complexes for basal transcription at P_{vpsL} . The presence of neither VpsR nor c-di-GMP alone to RNAP significantly affects these contacts.

In contrast, the activated complex of RNAP/VpsR/c-di-GMP at P_{vpsL} generated distinct footprints. On the nontemplate strand, strong protection was observed from +29 to -11, -13 to -21, -24 to -33, -35 to -44, and -47 to -52 with hypersensitivity sites at -22, -23, -54, and -55 (Fig. 2.14A, lane 8). On the template strand, strong protection was seen from -56 to -28 and -26 to +21 with hypersensitivity sites at -58 (Fig. 2.14B, lane 8). This pattern is consistent with the formation of a typical stable open complex of RNAP and an activator or in our case, RNAP and VpsR/c-di-GMP at P_{vpsL} .

Interestingly, a comparison of the DNase I footprints obtained with VpsR/RNAP/c-di-GMP versus VpsR/c-di-GMP reveal differences in the protection/cleavage patterns within the VpsR/c-di-GMP binding site of -31 to -53 as well as within the immediate upstream and downstream regions (Compare patterns in Fig. 2.2 and regions within the red dashed boxes in Fig. 2.14). For example, on the nontemplate strand, the addition of RNAP to VpsR/c-di-GMP yielded enhanced protection downstream and within the downstream portion of the binding site (-25 to -32 and -35 to -40) and enhanced cleavage in the upstream portion (-45, -54, and -55). On the template strand, addition of RNAP to VpsR/c-di-GMP yielded enhanced cleavage (-46 to -48)

within the VpsR/c-di-GMP-binding site and more protection (-53 to -55) and a hypersensitivity site (-58) upstream of the binding site. Because footprints between VpsR alone versus VpsR/c-di-GMP are identical, both RNAP and c-di-GMP are required to facilitate these protein-DNA contact changes within the VpsR-binding site in the activated complex. Thus, these results suggest that the binding of VpsR to its DNA site is altered by the presence of both RNAP and c-di-GMP and/or that contacts between RNAP and the DNA are altered by the presence of both VpsR/c-di-GMP.

Taken together, the footprints suggest that the transcription complex formed by VpsR, c-di-GMP, and RNAP at P_{vpsL} is competent because it achieves a different architecture. The presence of RNAP alone or with either c-di-GMP or VpsR does not generate this particular protein-DNA conformation.

2.4 Discussion

Biofilm formation by bacteria imposes an enormous medical cost, both in suffering and in the price of treatment. Consequently, understanding the regulation of biofilm formation is crucial to the prevention and treatment of bacterial disease. A central player in biofilm formation is the second messenger c-di-GMP, which has previously been shown to be required for the activity of several transcriptional activators including VpsR, the master regulator of biofilm formation in *V. cholerae*. By developing the first *in vitro* transcription assay with c-di-GMP, we have demonstrated that c-di-GMP works with VpsR in a novel way to stimulate transcription by RNAP at P_{vpsL} , a promoter for biofilm biogenesis genes. Surprisingly, unlike other characterized regulators that use c-di-GMP, such as *Klebsiella pneumoniae* MrkH, *Mycobacterium smegmaticum* LtmA, *Streptomyces coelicolor* BldD, *V. cholerae* VpsT, and *Pseudomonas aeruginosa* FleQ and BrlR (68,106,113,116,119,126-128), VpsR does not require c-di-GMP to oligomerize or bind to the DNA. VpsR dimers form with or without c-di-GMP, and the presence of the second messenger does not substantially affect the affinity of VpsR for the DNA or the protein-DNA contacts made by VpsR alone at P_{vpsL} . Instead, c-di-GMP is needed to observe distinct protein-DNA contacts within the activated transcription complex of $\sigma 70$ -RNAP/VpsR/c-di-GMP. How the presence of c-di-GMP results in these contacts is not clear. However, it could be needed to generate a particular VpsR conformation that is active for transcription and/or by promoting needed contacts between VpsR and $\sigma 70$ -RNAP. In fact, the position of the VpsR binding site immediately upstream of the -35 region suggests that VpsR should function as a Class II activator that can interact with $\sigma 70$ region 4 and/or alpha CTDs.

Besides the novelty of activation, VpsR is also unusual as an atypical EBPs. Classic EBPs interact with $\sigma 54$ -RNAP at a promoter, utilizing ATPase to form homomeric hexamers to generate

the energy needed to form a stable open complex (102,183). However, VpsR, like other atypical EBPs, lacks the GAFTGA motif needed for interaction with σ 54-RNAP and has nonconserved amino acids in the Walker B motif involved in ATP hydrolysis. Atypical EBPs that utilize σ 70 rather than σ 54 may represent an evolutionary link between these two very different σ class families. To date, five atypical EBPs have been characterized: *E. coli* TyrR, *Rhodobacter capsulatus* HupR, *Myxococcus xanthus* HsfA, *Pseudomonas putida* PhhR, and *Brucella abortus* NtrX (184-189). While all five atypical EBPs contain variations in the GAFTGA motif responsible for binding to σ 54, some contain nonconsensus Walker A or Walker B motifs involved in ATP binding and hydrolysis (190). Recently, the crystal structure of *B. abortus* NtrX was solved, representing the first full-length crystal structure of a NtrC-like response regulator as well as the first full-length crystal structure of an atypical EBP. However, unlike VpsR, NtrX functions as a repressor at the pYX promoter and does not bind c-di-GMP (189). Thus, it appears that atypical EBPs may function by varied mechanisms. Nevertheless, the remaining four atypical EBPs work with σ 70-RNAP in the absence of c-di-GMP to activate transcription. How the activity of these non-canonical EBPs is regulated remains to be determined for most of these transcription factors, but we show here that VpsR represents the first EBP and first atypical EBP that is dependent on a second messenger to directly activate transcription with RNAP.

In addition to VpsR, two other EBPs, FlrA in *V. cholerae* and FleQ from *Pseudomonas aeruginosa*, are also directly controlled by c-di-GMP. However, unlike VpsR, these regulators are typical EBPs and contain conserved elements needed for σ 54-dependent transcription. While binding of c-di-GMP to FlrA inhibits its ability to bind to the *flrBC* promoter to promote transcription activation (31), binding of c-di-GMP to FleQ has more complex effects. FleQ can regulate transcription at promoters containing σ 54 or σ 70 elements in *P. aeruginosa*, but it is

unclear whether FleQ directly activates transcription with these σ factors *in vitro*. Like FlrA, binding of c-di-GMP to FleQ represses flagellar genes in *P. aeruginosa*, but also derepresses and activates the *pel* biofilm extracellular polysaccharide gene cluster *in vivo* (106-108). Thus, VpsR, FlrA, and FleQ appear to function on a continuum with each transcription factor having different dependencies on σ_{54} or σ_{70} as well as different responses to c-di-GMP. While FlrA and FleQ bind c-di-GMP via conserved arginine residues that flank a central cavity between the N-terminal receiver domain and central AAA+ domain (106), VpsR lacks these arginines, instead having a methionine and glutamate at those positions. The mechanism by which VpsR binds to c-di-GMP is therefore unknown.

Along with the proximal VpsR binding site from -31 to -53 at P_{vpsL} , interestingly, a second VpsR binding site lies far upstream of P_{vpsL} at -297 to -336. These binding sites differ in both sequence, length, and protection intensities. Using DNase I footprinting with VpsR alone in the absence of c-di-GMP, the protection pattern was stronger at the distal site versus the proximal site (71). While VpsR protected the sequence TTTCTCAAAAATAATTCAATGTAAATCCAAAATGTAATAC at the distal site, VpsR protected the sequence AGTCTTAGAATTGATGCAGATA at the proximal site (71). Although this distal site has no effect in our *in vitro* transcription assays with purified proteins as well as no effect in transcriptional fusion studies when truncated, it appears that VpsR binding here is needed to relieve H-NS repression *in vivo* (71). The downstream portion of the distal VpsR binding site overlaps the first distal H-NS binding site (76). Thus, we speculate that at P_{vpsL} , VpsR acts as an anti-H-NS repressor, blocking H-NS binding at the distal promoter site. In between the proximal and distal VpsR binding sites, a VpsT binding site is present from -238 to -192. Previous work demonstrates that VpsT acts solely as an antirepressor of H-NS at P_{vpsL} and *in vitro* transcription

studies in our laboratory show that VpsT does not directly activate transcription at P_{vpsL} (data not shown). This allows for additional H-NS regulation in which both VpsR-binding to the distal promoter site and VpsT-binding downstream together help prevent H-NS from first binding the site overlapping the distal VpsR binding site. Upon freeing up the DNA from H-NS binding, VpsR/c-di-GMP may then bind to the proximal binding site to directly activate transcription with RNAP. Other VpsR sites appear at various locations relative to the TSS of various genes. VpsR binds and regulates *vpsT* with a site at -149 to -119, *aphA* with a site at -88 to -70, and *epsC* with a site from -50 to -33 (69,147,173), and *in silico* analyses have identified conserved VpsR boxes present at other locations, including promoters for *rbmA*, *rbmB*, *rbmC*, *rbmE*, *vpsU*, *vpsR*, *cdgC*, and *bap1* (71). H-NS sites have also been identified at some of these promoters (*vpsL*, *vpsT*, *rbmA*, *rbmB*, and *rbmC* (75-77)). Thus, we speculate that in general, promoter distal VpsR binding sites may correlate with a role in relieving H-NS repression, while promoter proximal sites may correlate with VpsR/c-di-GMP activation with RNAP. Such a mechanism may be similar to that used by *S. typhirium* SsrB. During biofilm formation, SsrB binds the DNA and displaces H-NS to relieve H-NS silencing and enable transcription activation of *csgD*, the master regulator of biofilms (191). However, at promoters of Salmonella Pathogenicity Island-2 SPI-2 genes, SsrB interacts with RNAP to activate transcription (160). The role of VpsR's diverse and numerous binding sites remain unclear and future studies in determining the differing roles of VpsR in transcriptional activation versus relieving H-NS repression are in progress.

CHAPTER 3:

VpsR directly activates transcription of multiple biofilm genes in *Vibrio cholerae*

Preface

Vibrio cholerae biofilm formation is important for its survival, dissemination, and persistence. Regulation of biofilm formation is complex and requires several transcriptional activators, repressors, and the ubiquitous signaling molecule, cyclic di-GMP (c-di-GMP). VpsR, the master positive regulator of biofilm formation, was previously shown to be a Class II transcription activator at the *vpsL* promoter with RNA polymerase containing $\sigma 70$ and c-di-GMP. In this study, we have identified additional biofilm promoters that are directly activated by VpsR in the presence of $\sigma 70$ containing RNAP and c-di-GMP. These promoters include P_{rbmA} , P_{rbmF} , and P_{vpsU} and are all located upstream of the first gene of each gene operon or cluster involved in biofilm formation. While VpsR binds to -82 to -59 at the *rbmA* promoter relative to the +1 transcriptional start site, VpsR binds to -99 to -57 at the *rbmF* promoter and -52 to -31 at the *vpsU* promoter relative to the TSS. Not only do these binding sites demonstrate that VpsR can function as both a Class I and Class II transcriptional activator, but the binding site at *rbmF* reveals that VpsR can bind to two sites in one intergenic region, simultaneously utilizing two classes of transcription activation to upregulate biofilm gene expression. Biofilm formation is complex and the master regulator, VpsR, directly binds and activates the first gene of multiple biofilm operons or gene clusters.

3.1 Introduction

Vibrio cholerae, the causative agent of the enteric disease cholera, is responsible for 3-5 million infections as well as 100,000-120,000 deaths per year (192,193). Symptoms of cholera include profuse rice water diarrhea, which subsequently leads to hypotonic shock and death within 12 hours in the absence of proper treatment (194). Though cholera is not often encountered in developed countries, it is still a significant public health problem in developing countries due to the lack of proper water sanitation facilities.

During its life cycle, *V. cholerae* lives in both the planktonic and biofilm state. In the human host as well as in its aquatic reservoir, *V. cholerae* forms biofilms to aid in survival, transmission, and persistence (36,57,58,164,167). Not only do biofilms shield the bacteria from harsh environmental conditions and stresses, they also protect the bacteria from protozoa and bacteriophage predation as well as nutrient limitation (56,164-167). Biofilms are comprised of matrix proteins, extracellular DNA, and *Vibrio* polysaccharides (VPS). Over 50% of the *Vibrio* biofilm matrix mass is comprised of VPS (59-64). Essential for the three-dimensional structure observed in biofilms, *V. cholerae* secretes VPS after initial attachment to either biotic or abiotic surfaces (66). Both the biofilm structure modulator A protein RbmA as well as the biofilm matrix protein Bap1 are also secreted from the cell surface, promoting cell-cell adhesion and cell-VPS adherence (61,62,66). As the biofilm grows and develops, other matrix proteins, outer-membrane vesicles, and extracellular DNA help encase the biofilm (63,66,195).

Genes involved in VPS synthesis are located on the *V. cholerae*'s larger chromosome and organized into two operons, *vps-1* and *vps-2* (59,60). *Vps-1* contains 12 genes, *vpsU*, *vpsA-K* while *vps-2* contains six genes, *vpsL-vpsQ* (59,60). 8.3 kilo base pairs separate these two operons and contain six ribomatrix genes: *rbmA*, *rbmC*, and *rbmBDEF* (Fig. 3.1) (59,61,62).

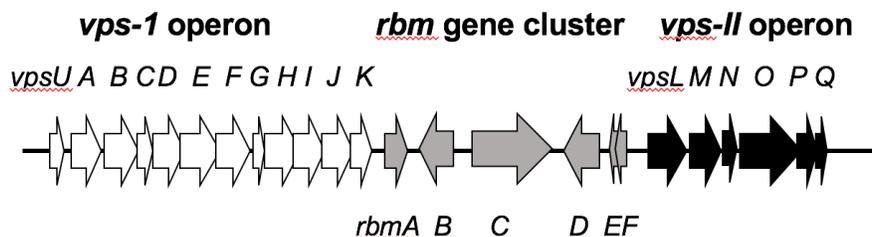


Figure 3.1: *Vibrio cholerae* genomic organization of genes involved in *Vibrio* polysaccharide (VPS) synthesis and ribomatrix protein production. VPS synthesis is comprised of two *vps* operons: *vps*-1 (white arrows) and *vps*-2 (black arrows). Ribomatrix protein production is comprised of the *rbm* gene cluster (grey arrows) and *bap1* (not shown). Arrow direction depicts transcription direction. Image is approximately to scale.

Since biofilm formation requires an assembly of multiple gene products, it is not surprising that this process is highly regulated with multiple signaling pathways, transcriptional repressors and activators, and regulatory sRNA. Among these effectors is the ubiquitous signaling molecule, cyclic di-GMP (c-di-GMP), a keystone regulator of biofilm formation in many bacterial species. c-di-GMP is formed from two GTP molecules by diguanylate cyclase enzymes that contain a GGDEF domain and then degraded into pGpG or two GMP molecules by phosphodiesterase enzymes containing either PDE or HD-GYP domains. Generally, high levels of c-di-GMP increase biofilm formation while low levels of c-di-GMP promote motility (168). Along with regulating the switch between biofilm formation and motility, c-di-GMP also controls a vast array of different phenotypes including DNA repair, type III secretion system, predation, development, fimbriae synthesis, and stress response (26,27,169).

The transcriptional activators, VpsR and VpsT are required to activate biofilm formation in *V. cholerae* (30,53,68-72,147). c-di-GMP binds both VpsR and VpsT with a dissociation constant (K_d) of 1.6 μ M and 3.2 μ M, respectively (68,69). However, at the biofilm biogenesis promoter, P_{vpsL} , c-di-GMP binds VpsT to enhance DNA binding while the presence or absence c-di-GMP does not promote VpsR-DNA binding (30,71). Instead, c-di-GMP is required to initiate

and activate transcription with both RNA polymerase containing the primary sigma (σ) factor, $\sigma 70$, and VpsR (30). Based on amino acid homology, VpsR is classified as an enhancer binding protein (EBP). Regulators in the EBP family use $\sigma 54$ and ATP hydrolysis to activate transcription. However, VpsR is an atypical EBP. Not only does VpsR lack the GAFTGA motif responsible for interacting with $\sigma 54$, but VpsR is also missing conserved residues within the Walker B domain involved in ATP hydrolysis. *In vitro* transcriptions demonstrate that VpsR activation of transcription with $\sigma 70$ -RNAP and c-di-GMP does not require ATP hydrolysis (30).

Transcription activators commonly function by a Class I or a Class II mechanism or a combination of both classes. In Class I transcription activation, activators bind to upstream regions typically around -60 and in Class II transcription activation, activators interact directly upstream or overlapping the -35 element (85). In the case of P_{vpsL} , VpsR functions as a class II activator. It interacts with a region overlapping the -35 element to subsequently initiate transcription with c-di-GMP and RNAP via the formation of the open complex (30).

In this study, we report our findings that similar to its induction of P_{vpsL} , VpsR requires RNAP containing $\sigma 70$ and c-di-GMP to activate transcription at P_{rbmA} , P_{rbmF} , and P_{vpsU} . Analogous to P_{vpsL} , VpsR binding at P_{vpsU} overlaps the -35 element, while VpsR binding sites at P_{rbmA} and P_{rbmF} are located farther upstream and centered around the -60. Though VpsR's DNA binding affinities vary in the presence and absence of c-di-GMP at each promoter, specific DNA-protein interactions do not change in the presence of the signaling molecule. Our findings demonstrate that VpsR is an adaptable activator, functioning as both a Class I and Class II transcription activator to induce additional promoters of genes essential for biofilm formation in *V. cholerae*.

3.2 Materials and Methods

DNA.

In vitro transcription templates, pMLH40 containing the *rbmA* promoter from -222 to +120 relative to the +1 transcription start site (TSS), pMLH41 containing the *rbmF* promoter from -209 to +167 relative to the TSS, and pMLH42 containing the *vpsU* promoter from -130 to +155 relative to the TSS were cloned into the EcoRI and HindIII restriction enzyme sites of pRLG770 (175) (74). Inserts were first obtained as PCR products, which had been amplified with primers (Table 1) from *V. cholerae* genomic DNA (BH1514) using Pfu Turbo polymerase (Stratagene). Vectors were digested with the appropriate restriction enzymes, ligations were performed using standard Gibson techniques (196), and resulting products were transformed into DH5 α .

pMLH11 is a pET28b(+) derivative (Novagen) that contains *vpsR* cloned within the NdeI and XhoI restriction sites constructed as previously described (NAR citation).

Fragments used for EMSAs and DNase I footprinting were obtained as PCR products using Pfu Turbo polymerase (Stratagene) and upstream and downstream PCR primers, which anneal from positions -222 to +120 relative to the TSS of *rbmA*, -209 to +167 relative to TSS of *rbmF*, and -130 to +155 relative to TSS of *vpsU*. To radiolabel the DNA, template primer was treated with T4 polynucleotide kinase (Affymetrix) in the presence of [γ -³²P] ATP prior to PCR. Radiolabeled PCR products were purified as described (177).

Proteins.

E. coli RNAP core was purchased from NEB. *E. coli* σ 70 was purified as previously described (180). VpsR protein was isolated from Rosetta2 (DE3)/pLysS (Novagen) containing pMLH11 and grown and purified as previously described (NAR citation). Protein concentrations

were determined by comparison of known amounts of core RNAP after SDS-PAGE and gel staining with Colloidal Coomassie Blue (Invitrogen).

***In Vitro* Transcriptions.**

Single round *in vitro* transcriptions were performed in 5 μ L reactions containing 0.02 pmol of supercoiled template, 3.0 pmol of VpsR, 0 or 12.5 μ M of c-di-GMP, reconstituted RNAP (0.2 pmol of σ 70 plus 0.05 pmol of core), and transcription buffer [40 mM Tris-acetate (pH 7.9), 150 mM potassium glutamate, 4 mM magnesium acetate, 0.1 mM EDTA (pH 7.0), 0.01 mM DTT, and 100 μ g/mL BSA]. Samples were first incubated at 37°C for 10 min prior to the addition of rNTPs (final concentrations of 220 μ M ATP, 220 μ M GTP, 220 μ M CTP, 11 μ M [α -³²P] UTP at 6.5 x 10⁵ dpm/pmol) and 500 ng heparin. After incubation for 10 min at 37°C, reactions were collected on dry ice, formamide load solution (15 μ L) was added, and aliquots were electrophoresed on 4% (wt/vol) polyacrylamide, 7 M urea denaturing gels for 2500 volt-hrs in 0.5 X TBE buffer. After electrophoresis, gels were exposed to X-ray films, films were scanned using a Powerlook 2100XL (check) densitometer and analyzed with Quantity One software (Bio-Rad).

Primer Extensions.

Primer extension products generated from RNA isolated from *in vitro* transcriptions were obtained described above and processed according to manufacturer's instructions (Promega) and as described previously (30). Briefly, 5 μ L of the *in vitro* transcription reaction was added to 6 μ L of a primer mixture containing 2 X AMV Primer Extension buffer, reverse transcriptase, and 2 pmol of ³²P-labeled primer (*rbmA* primer, which anneals 69 bp downstream of the TSS of *rbmA*; *rbmF* primer, which anneals 113 bp downstream of the TSS of *rbmF*; or *vpsU* primer, which

anneals 103 bp downstream of the TSS of *vpsU*. Samples were electrophoresed on 8% (wt/vol) polyacrylamide, 7 M urea denaturing gels for 5000 volt-hours in ½ X TBE. Imaging, densitometry, and quantification were performed as described above.

In vivo V. cholerae RNA was obtained from WN310 containing pMLH17 and pCMW75 or pCMW98. Cells were grown and harvested and RNA was extracted as previously described (30). Primer extension products were performed according to manufacturer's instructions (Promega) and as described previously (30).

Electrophoretic Mobility Shift Assays.

VpsR-DNA complexes and transcription complexes with RNAP were formed using 0.05 pmol of DNA, as previously described in transcription buffer (30). To ensure observation of specific complexes, a 1 µl solution of 1 mg/mL of poly(dI-dC) or 500 µg/ml heparin was added to VpsR-DNA complexes (20 µL) or transcription complexes (10 µL), respectively. After addition of competitor, VpsR-DNA complexes were loaded onto a 5% (wt/vol) nondenaturing polyacrylamide gel already running at 100 V/hr in Tris/borate/EDTA (TBE) buffer and subsequently electrophoresed for 1.5 h at 100 V/hr. Transcription complexes were loaded onto a 4% (wt/vol) nondenaturing, polyacrylamide gel already running at 100 V in 1 X TBE buffer and electrophoresed for 3 h at 380 V/hr. After autoradiography, films were scanned as described above. $K_{d(app)}$'s were calculated as the concentration of VpsR needed to shift 50% of the free DNA.

DNase I footprinting.

Solutions in transcription buffer were assembled as described above for EMSAs using 0.04 µM DNA, and as indicated, 1.4 µM VpsR, 50 µM c-di-GMP, and/or 0.16 µM of reconstituted

RNAP. After incubation at 37° C for 10 min, poly(dI-dC) or heparin was added to VpsR-DNA complexes and transcription complexes with RNAP, respectively, for 15 s at 37° C. To initiate DNase I cleavage reaction, 0.3 U of DNase I in 1.5 µL was added. Solutions were incubated for an additional 30 s at 37° C, immediately loaded onto a 4% (wt/vol) nondenaturing, polyacrylamide gel already running at 100 V in 1 X TBE buffer, and electrophoresed for 3 h at 380 V/hr. After autoradiography, the protein/DNA complexes were excised and extracted DNA was electrophoresed on denaturing gels as described (30,156).

3.3 Results

VpsR binds the promoter regions of *rbmA*, *rbmF* and *vpsU* with or without c-di-GMP.

Previous in silico sequence analysis using MEME (multiple EM for motif elicitation) revealed VpsR motifs upstream of the following genes: *vpsR*, *vpsL*, *aphA*, *VCA0075*, *rbmA*, *vpsA*, *vpsT*, *vpsU*, *bap1*, *cdgC*, and *rbmC* (71). Of these promoters, VpsR binding sites have been identified at *aphA*, *vpsL*, and *vpsT* (30,69,71,147). To assess whether VpsR bound to the promoters of these other genes predicted by MEME, we performed gel retardation assays. Similar to what we observed with P_{vpsL} , VpsR bound to P_{rbmA} , P_{rbmF} , and P_{vpsU} and the presence of c-di-GMP was not required for this binding (Fig. 3.2). However, for P_{rbmA} and P_{rbmF} , VpsR binding affinity improved by 50% and ~5-fold, respectively, in the presence of c-di-GMP (Fig. 3.2A and B). In contrast, no significant difference in VpsR affinity for P_{vpsU} (Fig. 3.2C) or P_{vpsL} (30) was observed with or without c-di-GMP.

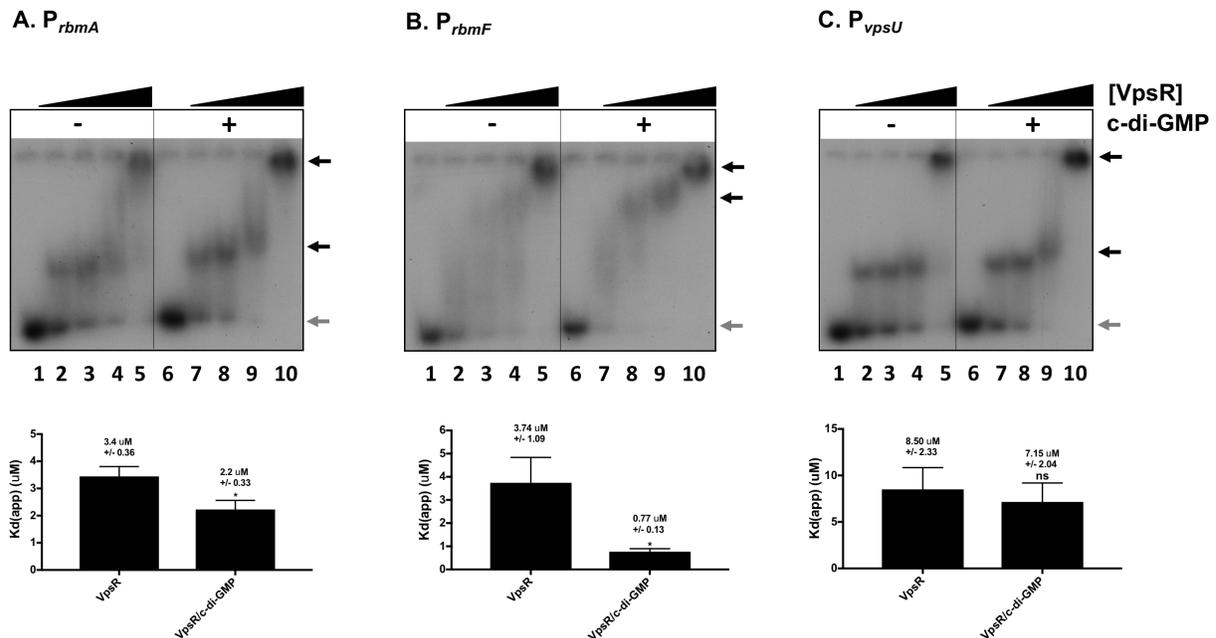


Figure 3.2: VpsR binds P_{rbmA} , P_{rbmF} , and P_{vpsU} with and without c-di-GMP. Representative gel retardation assays of 32 P-labeled DNA harboring -222 to +120 of P_{rbmA} , -209 to +167 of P_{rbmF} , and -130 to +155 of P_{vpsU} relative to the +1 TSS with increasing VpsR concentrations from 0 to 2.5 μ M either in the absence of c-di-GMP (lanes 1-5) or presence of 50 μ M c-di-GMP (lanes 6-10).

Samples were incubated at room temperature for 10 minutes prior to electrophoresis on 5% TBE polyacrylamide gel. (D) Apparent DNA-binding dissociation constants $K_{d(app)}$ has been calculated as the concentration of VpsR needed to retard 50% of the free DNA. Values from at least three gel retardation assays were analyzed using one-way ANOVA with Tukey's HSD *posthoc* analysis (ns, not significant; * $P < 0.05$).

VpsR-dependent activation of the *rbmA*, *rbmF* and *vpsU* promoters requires both VpsR and c-di-GMP *in vitro*.

To determine if VpsR and c-di-GMP directly activate transcription at P_{rbmA} , P_{rbmF} , and P_{vpsU} , we constructed *in vitro* transcription templates containing P_{rbmA} (-222 to +120; pMLH40), P_{rbmF} (-209 to +167; pMLH41), and P_{vpsU} (-130 to +155; pMLH42) Each insert was located upstream of the *rrnBPI* terminator within the vector pRLG770 (175). Using *E. coli* RNAP (New England Biolabs) with $\sigma 70$ and supercoiled templates, we found that both VpsR and c-di-GMP were needed to directly activate transcription from these promoters (Fig. 3.3A-C, lanes 4). RNAP, VpsR, and c-di-GMP activated transcription at P_{rbmA} , P_{rbmF} , and P_{vpsU} by approximately 3-fold, 4-fold, and 17-fold over basal level, respectively. RNAP alone or RNAP with either c-di-GMP or VpsR only yielded a basal level of transcription (Fig. 3.3A-C, lanes 1-3).

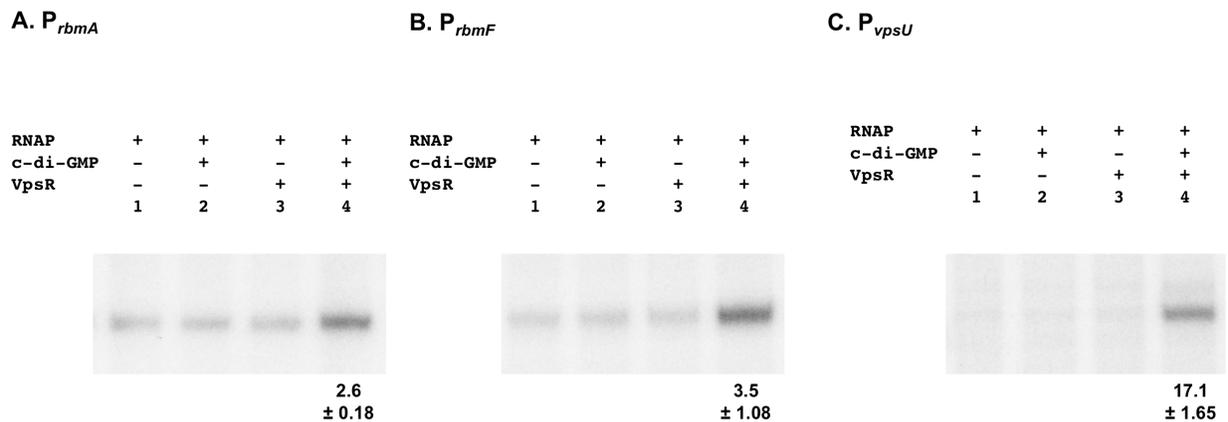


Figure 3.3: VpsR and c-di-VpsR and c-di-GMP activate transcription at P_{rbmA} , P_{rbmF} , and P_{vpsU} *in vitro*. Representative images of single round *in vitro* transcription from supercoiled plasmid templates at (A) P_{rbmA} , (B) P_{rbmF} , and (C) P_{vpsU} with RNAP containing $\sigma 70$ (lane 1), RNAP and c-di-GMP (lane 2), RNAP and VpsR (lane 3), and RNAP and VpsR and c-di-GMP (lane 4).

Figure 3.3 (cont'd): Activation fold change for RNAP/VpsR/c-di-GMP relative to basal level expression of RNAP alone is below the image. All reactions were repeated in triplicates.

***In vivo* and *in vitro* primer extension analyses define the +1 transcriptional start site for *rbmA*, *rbmF*, and *vpsU*.**

To determine the +1 TSS of the RNA generated from P_{rbmA} , P_{rbmF} , and P_{vpsU} after *in vitro* transcription reactions, we performed primer extension analyses. These analyses revealed TSSs at the C nucleotide located 19 bps upstream of the TTG start codon of *rbmA* (Fig. 3.4A, lane 8), at the A and T nucleotides located 44 and 46 bps upstream of the ATG start codon of *rbmF* (Fig. 3.4B, lane 8), and at the T nucleotide located 53 bps upstream of the ATG start codon of *vpsU* (Fig. 3.4C, lane 8).

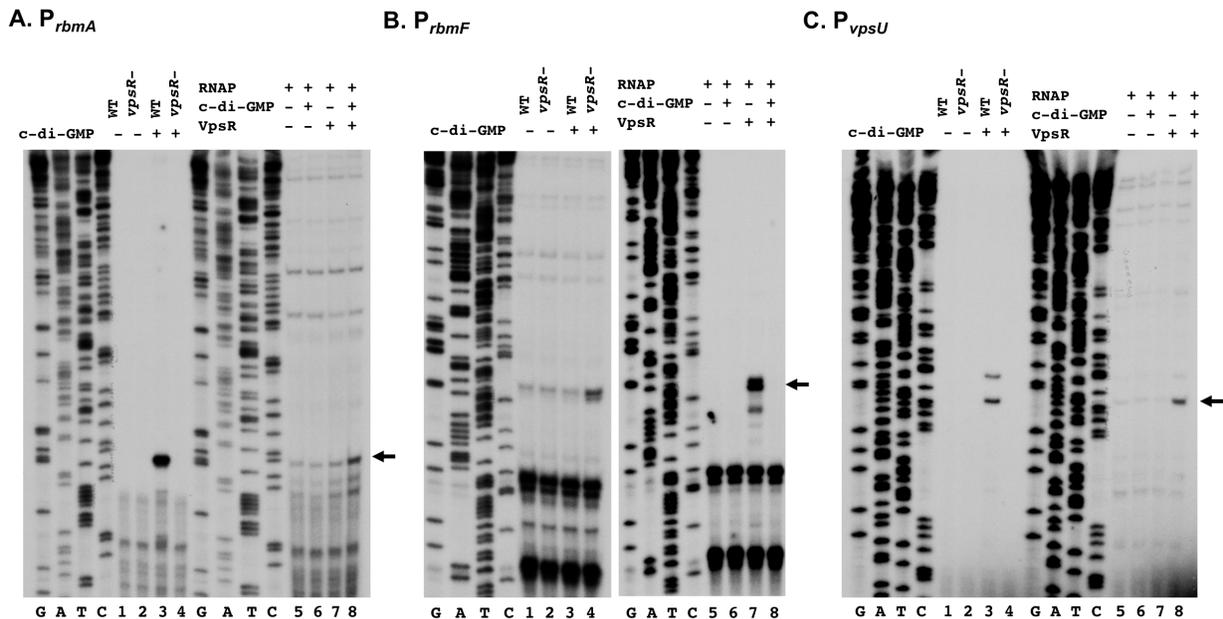


Figure 3.4: Determination of +1 transcriptional start site (+1) *in vivo* and *in vitro* for P_{rbmA} , P_{rbmF} , and P_{vpsU} . Primer extension analyses were performed with ^{32}P -labeled primer that hybridized 69, 113, and 103 basepairs downstream of the *rbmA*, *rbmF*, and *vpsU* translation start site, respectively. RNA was isolated from *V. cholerae* (lanes 1-4) or *in vitro* transcriptions (lanes 5-8). (A) For P_{rbmA} , the most prominent *in vivo* and *in vitro* transcripts corresponding to the C located 19 bp upstream of the *rbmA* translation start site. (B) For P_{rbmF} , the most prominent *in vivo* and *in vitro* transcripts corresponding to the T located 46 bp upstream of the *rbmF* translation start

Figure 3.4 (cont'd): site. (C) For P_{vpsU} , the most prominent *in vivo* and *in vitro* transcripts corresponding to the T located 53 bp upstream of the *vpsU* translation start site.

Primer extension analyses of RNA isolated from *V. cholerae* grown in the presence and absence of c-di-GMP as well as from a $\Delta vpsR$ strain were confirmatory. As seen using *in vitro* RNA, we again observe primary primer extension products corresponding to the C TSS 19 bps upstream of the TTG start codon of *rbmA* (Fig. 3.4A, lane 3), the T 46 bp upstream of the ATG start codon of *rbmF* (Fig. 3.4B, lane 3), and the T located 53 bps upstream of the +1 ATG (Fig. 3.4C, lane 3) with RNA isolated from *V. cholerae*. These +1 positions are not seen in a $\Delta vpsR$ mutant or under low c-di-GMP levels (Fig. 3.4 A-C, lanes 1, 2, and 4), indicating that as seen *in vitro*, this activation is dependent on both VpsR and c-di-GMP.

VpsR complexes with RNAP and c-di-GMP bind DNA in typical Class I and Class II transcription complex fashion.

Since the presence of c-di-GMP altered the DNA binding affinity of VpsR to P_{rbmA} and P_{rbmF} , but not P_{vpsU} in gel retardation assays (Fig. 3.2), we were curious to see if there were any changes in specific DNA bp contacts. Using DNase I footprinting, we discovered that VpsR bound *rbmA* from -89 to -52 (Fig. 3.5A), VpsR bound *rbmF* from -99 to -57 (Fig. 3.5B), and VpsR bound *vpsU* from -52 to -31 (Fig. 3.5C) relative to the +1 TSS of each promoter. This binding pattern, which did not change in the presence and absence of c-di-GMP, is summarized in Fig. 3.7.

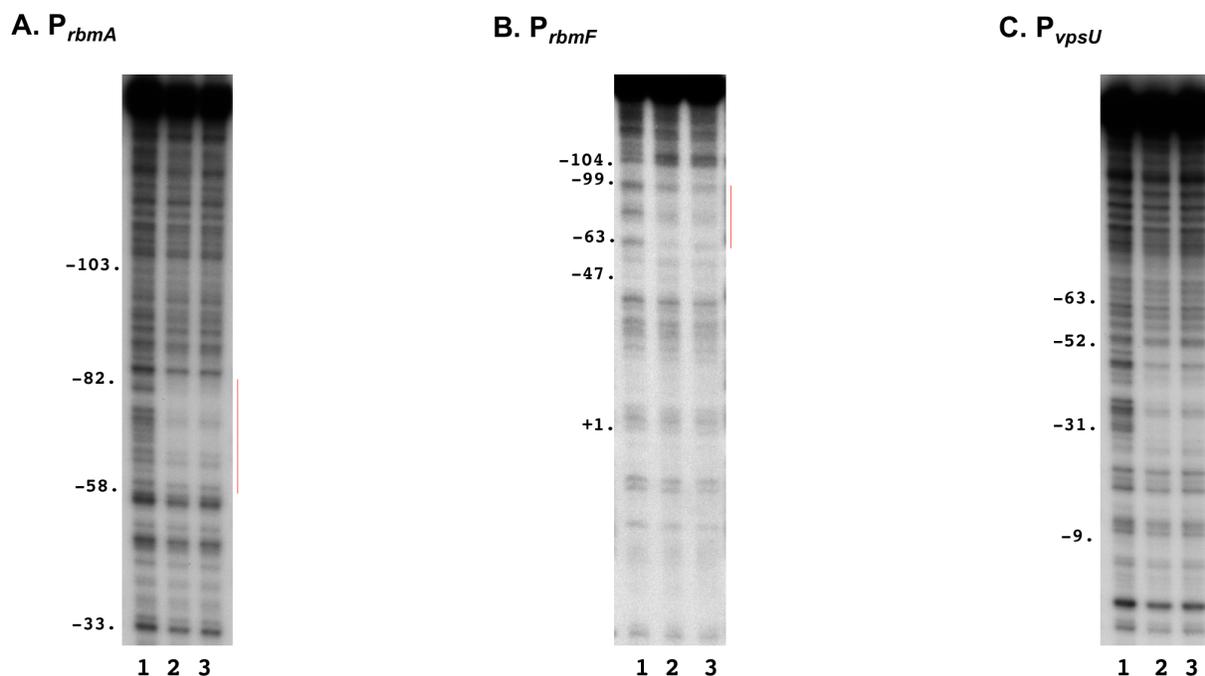


Figure 3.5: DNase I footprinting of VpsR with and without c-di-GMP at template P_{rbmA} , P_{rbmF} , and P_{vpsU} . VpsR with and without c-di-GMP complexes were formed and treated with poly(dI-dC) while transcription complexes were formed and treated with heparin at template (A) P_{rbmA} , (B) P_{rbmF} , and (C) P_{vpsU} . All complexes were treated with DNase I prior to isolation from gel retardation assays. VpsR binding site is indicated to the right of each image with a solid red line.

Next, we wanted to determine the bp contacts of the activated transcription complex of RNAP/VpsR/c-di-GMP at all three promoters. To ensure that we are observing footprints of the activated transcription complex, we performed DNase I footprinting on the template strands in combination with gel retardation assays. This enables us to also remove all uncut background DNA. Both *rbmF* and *vpsU* promoters demonstrated typical footprints for an activated transcription complex. At P_{rbmF} , protection extended from -96 to +23 with hypersensitivity sites at -97, -35, and -26, consistent with VpsR binding at its site centered at -78 and RNAP binding at the core promoter sequence (Fig. 3.6B). At P_{vpsU} , protection extended from -57 to +22 with a hypersensitivity site at -58, consistent with the VpsR binding site centered at -41, and RNAP located immediately downstream (Fig. 3.6C). Unlike P_{rbmF} , although RNAP protection was similar

in the absence or presence of VpsR/c-di-GMP, changes in hypersensitivity sites were observed only in the presence of VpsR/c-di-GMP at P_{vpsU} . Unfortunately, we were unable to obtain clear DNase I footprints for VpsR and RNAP at P_{rbmA} despite multiple attempts (Fig. 3.6A). This result suggests that the activated transcription complex at P_{rbmA} is unusually unstable.

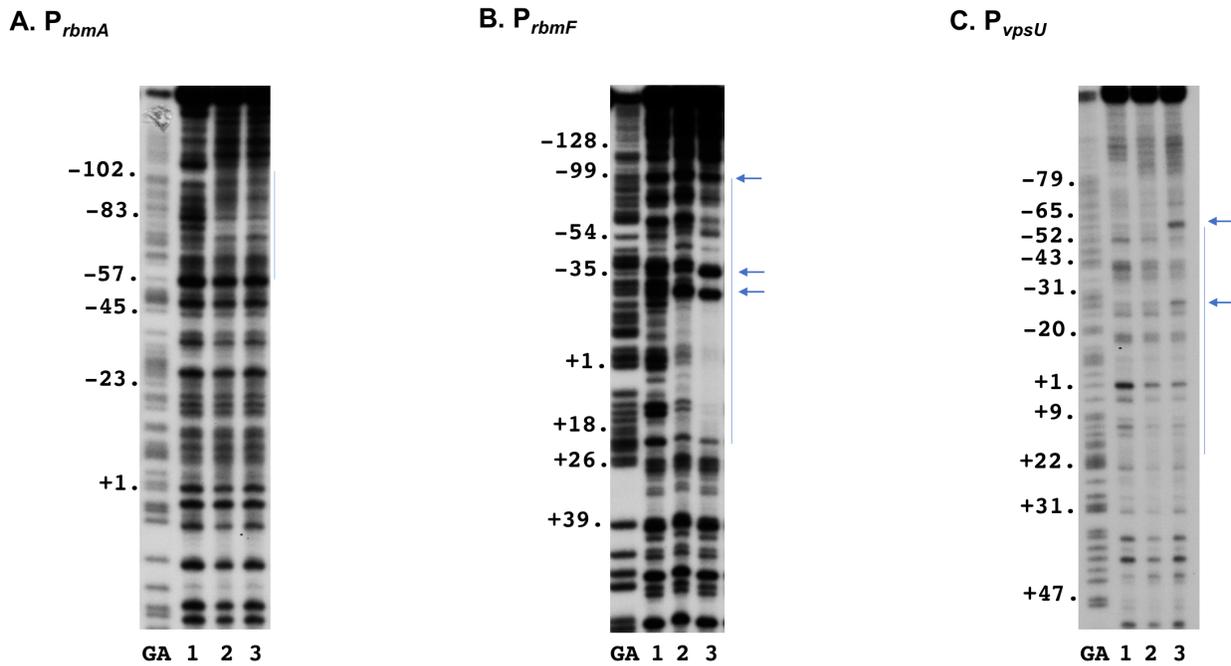


Figure 3.6: DNase I footprinting of transcription complexes containing RNAP alone or RNAP with VpsR and c-di-GMP at template P_{rbmA} , P_{rbmF} , and P_{vpsU} . Transcription complexes were formed, challenged with heparin, and treated with DNase I prior to isolation from gel retardation assays at template (A) P_{rbmA} , (B) P_{rbmF} , and (C) P_{vpsU} . GA represents G+A ladder. Lane 1 contains DNA only, lane 2 contains RNAP, and lane 3 contains RNAP/VpsR/c-di-GMP. Protection sites of transcription complexes are indicated by a blue line and hypersensitivity sites are represented by blue arrow, both to the right of each image.

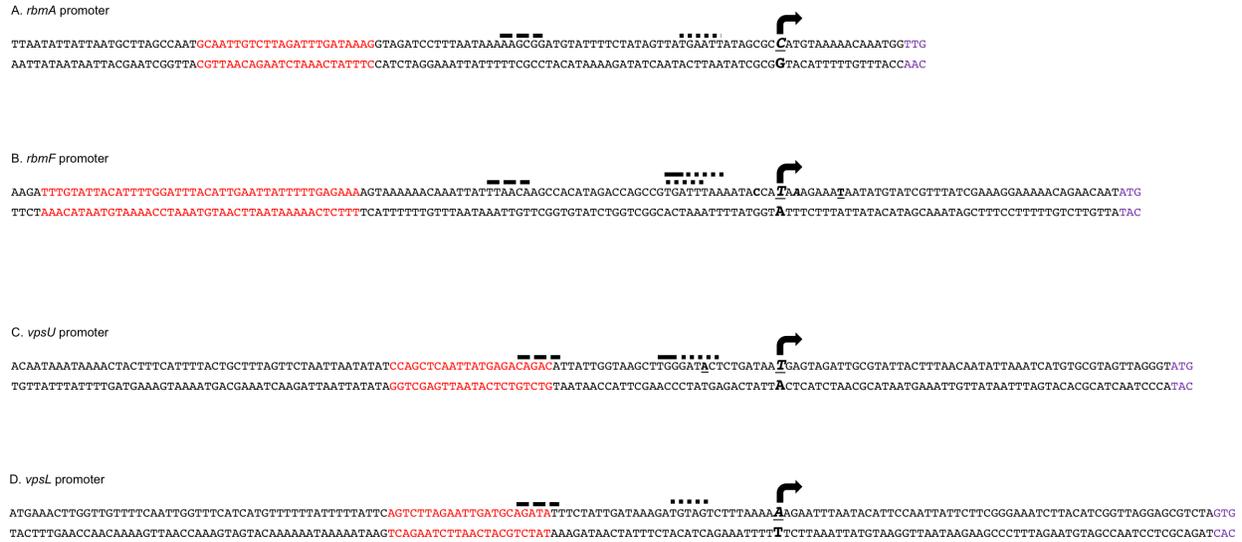


Figure 3.7: Promoter sequence and summary of DNase I footprinting of P_{rbmA} , P_{rbmF} , P_{vpsU} , and P_{vpsL} . (A) P_{rbmA} contain sequences from -103 to +22. (B) P_{rbmF} contain sequences from -103 to +49. (C) P_{vpsU} contain sequences from -103 to +56. (D) P_{vpsL} contain sequences from -103 to +62. VpsR binding sites are in red, +1 transcriptional start site is in purple, predominant +1 transcriptional start site from both *V. cholerae* RNA and *in vitro* transcription RNA are bold, italicized, and underlined with a black arrow above, alternative +1 TSS from *V. cholerae* RNA are bold and underlined, alternative +1 TSS from RNA isolated from *in vitro* transcriptions are bold and italicized, potential -10 consensus elements contain a dotted black line above the sequence, potential extended -10 sites contain a solid black line above the sequence, and the -35 region contain a dashed black line above the sequence.

3.4 Discussion

A variety of different mechanisms are used to direct RNAP to specific promoters to activate transcription in response to environmental cues and growth signals. Factors can either first bind to the promoter to subsequently guide RNAP to initiate transcription or factors can first interact with RNAP to then change its promoter preference (85). For promoters that use transcriptional activators, upon directing RNAP to the DNA, these activators typically either stabilize the initial RNAP-DNA complex and/or accelerate the transition from the closed unstable complex, R_{Pc}, into the stable open complex, R_{Po} (85).

Activators can be generally categorized into four different classes based on promoter binding location and RNAP contacts. In Class I activation, the activator binds upstream of RNAP (typically around -60 or upstream) and interacts with one or both α CTDs (85). In Class II activation, the activator binds adjacent to or overlapping the -35 element, directly contacting σ 70 region 4 and/or the α NTD (85). While most transcriptional activators can be classified as either a Class I or Class II activator, *Bordetella pertussis* uses a combination of both Class I and Class II at P_{flaB}. At this promoter, BvgA requires contacts with σ 70 region 4 as well as interaction with α CTDs (197). Another set of regulators alter DNA conformation to shorten the suboptimal distance between the -35 and -10 element, allowing RNAP to bind. The most well-understood activator of this class is MerR. Binding of MerR to the promoter causes a twist in the -35 and -10 spacer, triggering transcription of genes needed for efflux pumps (198). In our last mechanism of activation, σ appropriation, σ 70 region 4 is remodeled by a small protein to redirect σ DNA binding and alter σ protein binding by small proteins encoded by bacteriophage T4 (199). More specifically, T4 uses the co-activator AsiA to remodel the helix-turn-helix of *E. coli* σ 70 Region 4 (200), thereby allowing σ to bind the T4 activator MotA, which interacts with a motif within T4

middle promoters. This hijacks *E. coli* RNAP and results in recognition of the T4 middle promoters rather than host *E. coli* promoters (201).

VpsR, the master regulator of *V. cholerae* biofilm genes, uses a combination of Class I and Class II mechanism of activation at biofilm promoters as demonstrated in this chapter (Fig. 3.8). Based on VpsR binding site locations at these promoters, Class I is seen at *rbmA* and *rbmF* promoter while Class II is seen at *vpsL* and *vpsU* promoters.

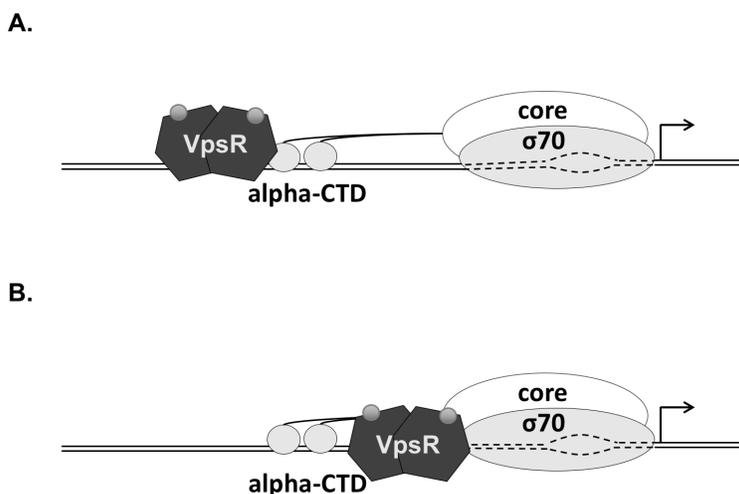


Figure 3.8: Mechanisms of VpsR transcription activation at different promoters. Diagrams of the RNAP/VpsR/c-di-GMP transcription complex in (A) Class I activation at P_{vpsL} and P_{vpsU} and (B) Class II activation at P_{rbmA} and P_{rbmF} .

This ability to use two different classes of activation is also seen in *E. coli* cAMP receptor protein (CRP; also known as CAP, catabolite acceptor protein). At the Class I *lac* promoter, CRP binds upstream of RNAP at a site centered around -61.5, and transcription activation is mediated by protein-protein interactions between the surfaced-exposed β -turn of CAP and one of the α CTDs (202). This increases the affinity of RNAP for promoter DNA, activating transcription. On the other hand, at the Class II *galP1* promoter, CRP binds to a site centered around -42, overlapping the -35 element, and transcription activation is mediated by three protein-protein interactions

between the α CTD, α NTD and σ 70 region 4 (202). Unlike class I activation, CRP class II activation not only recruits RNAP, but also facilitates formation of RPc to RPo (202). Interestingly, we find that the presence of c-di-GMP modestly or more significantly improves the DNA binding affinity of VpsR for the Class I *rbmA* and *rbmF* promoters, while its presence does not affect VpsR's DNA binding affinity at the Class II *vpsU* and *vpsL* promoters. This is consistent with the idea that VpsR can use different activation mechanisms at different promoters. Our gel shift assays demonstrate an increase in transcription complex formation at Class II promoters, but not at Class I promoters during DNase I footprinting assays (data not shown). In addition, significantly increased protection was observed with the activated transcription complex of RNAP/VpsR/c-di-GMP at both P_{*vpsL*} (30) and P_{*rbmF*} (Fig. 3.6). However, the RNAP protection sites were similar in the absence or presence of VpsR/c-di-GMP at P_{*vpsU*} (Fig. 3.6). These additional hypersensitivity sites suggest that perhaps at the *vpsU* promoter, instead of recruiting RNAP and enhancing the DNA binding affinity of the transcription complex, VpsR/c-di-GMP instead play a role in altering the DNA architecture required for transcription activation. While it was previously demonstrated that c-di-GMP directly drives open complex formation and is required to rapidly form the stable heparin-resistant open complex at the Class II *vpsL* promoter (30), whether c-di-GMP plays the same and/or differing roles at the P_{*rbmA*}, P_{*rbmF*}, and P_{*vpsU*} promoters remains to be determined.

The intergenic region between the divergently transcribed *rbmF* and *vpsL* genes contains two VpsR binding sites. It has previously been suggested that the distal VpsR binding site (relative to the +1 TSS) functions as a H-NS anti-repressor site (30). Our *in vitro* transcriptions and primer extension analyses in this chapter demonstrate that this distal VpsR binding site is instead directly involved in activating the divergent *rbmF* promoter. These two VpsR binding sites are approximately 237 bp apart. Though VpsR forms a dimer in the presence and absence of c-di-

GMP in solution (30), perhaps in the presence of DNA containing two binding sites, VpsR can then form higher order structures. This would then allow DNA looping, further preventing H-NS binding and repression while simultaneously bolstering transcription. It would be interesting to see if there is temporal control in transcription activation of these two divergently transcribed promoters mediated by the VpsR's DNA binding affinity and/or RNAP recruitment and isomerization.

Along with VpsR, the repressor H-NS and the anti-repressor VpsT also directly regulate all four promoter regions. Gel retardation assays show that H-NS and VpsT/c-di-GMP bind *rbmA* and *vpsU* promoters as well as the intergenic region between *vpsL* and *rbmF* (71,76). Chromatin immunoprecipitation (ChIP) assay demonstrated that deletion of *vpsT* resulted in higher H-NS promoter occupancy while overexpression of c-di-GMP leads to a decrease in H-NS promoter occupancy at *rbmA*, *rbmF*, *vpsL*, and *vpsU* (76). Nevertheless, because the expression from these promoters was still present in the absence of both H-NS and VpsT (76), we also propose that at these promoters, VpsT solely functions as an antirepressor and as demonstrated in our *in vitro* transcriptions, VpsR is the main transcriptional activator with RNAP.

Taken together, our data supports a complex model of *vps* and *rbm* regulation dependent on H-NS, VpsT, VpsR, and c-di-GMP. While VpsT/c-di-GMP are important in releasing the DNA from negative regulation by H-NS, VpsR/c-di-GMP are important in positive regulation, directly activating transcription with RNAP containing $\sigma 70$. Because these promoters vary in VpsR-DNA binding affinity in the absence and presence of c-di-GMP, we speculate that VpsR's mechanism of activation is dependent on both the concentration of VpsR and the level of c-di-GMP to increase transcription. This two-tiered type of regulation allows fine-tuning of gene expression. It is important to note that VpsR also contains a highly conserved aspartic residue at position 59. As an

orphan response regulator, there is no cognate histidine kinase directly upstream of *vpsR* and currently, the kinase involved in phosphorylating VpsR remains undiscovered. If VpsR is indeed phosphorylated, the combination of phosphorylation, c-di-GMP levels, and VpsR concentration truly allows exquisite fine-tuning and diversity in transcription activation mechanisms at VpsR regulated-promoters.

CHAPTER 4:

VpsR requires c-di-GMP and phosphorylation to drive transcription initiation to activate biofilm formation in *Vibrio cholerae*

Preface

Two-component signal transduction systems (TCSs), typically comprised of a sensor histidine kinase (HK) and a response regulator (RR), are used by bacteria to respond to environmental stimuli and extracellular signals. With 43 predicted HKs and 52 putative RRs, *Vibrio cholerae* is well-adapted to environmental changes as it transitions between the aquatic reservoir and the human host. Upon a signal input, the HK is phosphorylated and then transfers its phosphate to the RR. Because few studies have researched the role of phosphorylation in transcription initiation, we have investigated the role of the highly conserved D59 residue of the orphan response regulator VpsR *in vivo* and *in vitro* at the biofilm biogenesis promoter P_{vpsL} . Using the typically phosphodeficient variant D59A and the phosphomimetic D59E, we determined that both variants dimerize and bind DNA with $K_{d(app)}$ s similar to that of wildtype (WT) while VpsR D59E activates transcription and D59A does not. DNase I and $KMnO_4$ footprints of the transcription complex made with D59E resemble those made with WT, while footprints with D59A resemble those of RNAP alone. Lastly, we have also developed a new method to purify VpsR (VpsR^{ren}) using denaturation and renaturation. VpsR^{ren} resemble the D59A variant in dimerization, DNA-binding affinity, and DNase I and $KMnO_4$ footprints while addition of acetyl phosphate to VpsR^{ren} (VpsR^{ren}~P) yields similar results to WT and D59E variant. Our data suggests that both c-di-GMP and phosphorylation of VpsR are required to initiate transcription and generate the specific protein-DNA architecture needed for activated transcription.

4.1 Introduction

Rapid detection and adaptation to fluctuating environmental changes are essential for survival and proliferation. Not only must the microorganism sense a multitude of extracellular signals, but it must then transmit those signals to alter and control gene expression. These signals tend to be physical and chemical parameters such as temperature, pH, osmolarity, autoinducers, ions, and host cell contact (203). In bacteria, the prevalent signaling pathways are the two-component signal transduction systems (TCSs). TCSs are comprised of a membrane-bound histidine kinase (HK) and a soluble cytoplasmic response regulator (RR). Though separated, these two proteins are linked by a phosphotransmission or phosphorelay pathway and are generally located adjacent to each other within genomes (204). After a signal is perceived by the sensory domain of the HK, the His residue of the transmitter domain of the HK is phosphorylated by the ATPase domain. In phosphotransmission, the HK then transfers its phosphoryl group to the Asp of the receiver domain of the cognate RR while in phosphorelay, the phosphoryl group is transferred from the His to the Asp of the HK, then to the His of the His-containing phosphotransfer (Hpt) protein, and then finally to the Asp of the RR. Phosphorylation of the RR often induces a conformational change, altering DNA-binding properties and controlling gene expression. As the phosphorylation of an aspartic acid residue is typically unstable and labile (205,206), the phosphate group can be lost passively or a specific phosphatase can remove the phosphorylation, leading to an inactive RR, which can then be phosphorylated again. Currently, over 5000 bacterial genomes have been sequenced with over 70,000 predicted HKs and over 500,000 predicted RRs (207).

Vibrio cholerae is the Gram-negative motile bacterial pathogen that causes the gastrointestinal disease, cholera. In areas where cholera is endemic, disease occurrence parallels seasonal pattern and climate changes (164). In regions where cholera is nonendemic, upon

introduction of the pathogenic bacteria, poor sanitation leads to rapid bacterial dissemination via the fecal-oral route (208). Despite causing illness in human hosts, *V. cholerae* is a natural aquatic inhabitant and survives many environmental niches including the tropics such as the Bay of Bengal and temperate waters world-wide, including Australia, Italy, and Sweden (209-211). With over 43 predicted HKs and 52 putative RRs (182), *V. cholerae* employs TCSs as a major tool to sense and respond to these environmental changes, helping it survive stressors such as changes in temperature and salinity, and to resist predators such as bacteriophages and protists.

Several TCSs have been characterized in *V. cholerae*. These TCSs regulate behaviors such as quorum sensing, virulence, nutrient availability, and biofilm formation. Biofilm formation is crucial to the bacterium since it facilitates survival in the aquatic environment and transmission to the human host. Of the 52 putative RRs, 12 RRs have been shown to regulate biofilm formation positively and negatively. For example, VpsR, VpsT, LuxO, and VxrB activate biofilm formation while PhoB, VarA, VieA, VarR, VC1348, and VCA0210 repress biofilm formation (30,67,72,74,182,212-217). VpsR is the master regulator of biofilm formation (30,53,60,71,72). We have recently shown that in the presence of small molecule, cyclic-di-guanosine monophosphate (c-di-GMP), VpsR directly activates transcription *in vitro* with RNA polymerase (RNAP) containing the primary sigma (σ) factor, σ 70 (30). Based on amino acid homology, VpsR has been classified as an enhancer binding protein (EBP). EBPs have three distinct domains: a receiver (REC) domain, an ATPases Associated with diverse cellular Activities (AAA+) domain, and a helix-turn-helix (H-T-H) domain. For many EBPs, the AAA+ domain interacts with the alternate σ factor, σ 54. However, VpsR is an atypical EBP since both its REC and AAA+ domains contain substitutions at highly conserved residues needed for activity. Although VpsR is considered an orphan response regulator due to the lack of cognate histidine kinase directly

upstream of the *vpsR* gene, a highly conserved aspartic acid residue (D59) is present within the REC domain, suggesting that it might be phosphorylated (Fig. 4.1). This idea is bolstered by a previous study, which has demonstrated that cells complemented with the phosphodeficient D59A variant cannot form biofilms, while cells complemented with the phosphomimetic D59E variant produce biofilms at levels similar to wildtype (WT) (171). However, we have shown previously that purified VpsR, heterologously produced in *E. coli*, is active for *in vitro* transcription without needing treatment with a HK or acetyl phosphate, an unexpected result given the usual instability of a phosphorylated aspartic acid residue (205,206). Thus, the phosphorylation status of VpsR has remained unknown, a kinase responsible for phosphorylating VpsR has remained unidentified, and the role of phosphorylation in transcription activation has remained to be determined.

Along with TCSs, the highly ubiquitous second messenger, c-di-GMP, also regulates biofilm formation. c-di-GMP is generated from two GTP molecules by diguanylate cyclases (DGCs) containing GGDEF domains. c-di-GMP is then degraded into GMP or pGpG by phosphodiesterases (PDEs) containing conserved EAL or HD-GYP domains (218). The *V. cholerae* genome contains 31 GGDEF, 22 EAL, 9 HD-GYP, and 10 combined GGDEF-EAL domain proteins (219). Levels of c-di-GMP controls the transition between the motile planktonic state and the sedentary biofilm lifestyle negatively and positively, respectively (169). Along with biofilms and motility, other phenotypes regulated by c-di-GMP include quorum sensing, DNA repair, predation, stress response, virulence, and cell cycle progression (26,27,169).

As the master regulator of biofilm formation, VpsR activates several biofilm promoters including P_{rbmA} , P_{rbmF} , P_{vpsL} , P_{vpsT} , P_{vpsU} (Chapter 3). Transcription initiation is a multistep process mediated by RNAP. RNAP core is comprised of five subunits: 2 α s, β , β' , and ω . During the first step of this process, core binds the σ factor to form the RNAP holoenzyme. Primary σ factors

facilitate recognition of -35 and -10 consensus DNA elements. Upon binding to the DNA, RNAP forms the unstable closed complex (R_{Pc}). The R_{Pc} contains double-stranded (ds) DNA and is typically very short-lived, quickly isomerizing to the stable open complex (R_{Po}). Major conformational changes are seen in the R_{Po}/DNA complex including DNA bending and unwinding of the DNA from -11 to ~ +5 relative to the +1 transcriptional start site (TSS) (84,91,92). While some promoters are active with RNAP alone, other promoters with suboptimal sequences use activators, including RRs, to activate transcription at various steps of initiation, such as recruitment of RNAP, formation of the open complex, and/or promoter clearance (85). Class I and Class II are the major classes of activation (95). During Class I activation, the activator binds to a site centered upstream of the core promoter sequence typically around -60 and interacts with the α CTDs, while during Class II activation, the activator binds to sites overlapping or immediately adjacent to the -35 element and interacts with σ 70 region 4 and/or α NTDs (85). As seen at P_{rbmA} , P_{rbmF} , P_{vpsL} , and P_{vpsU} , VpsR can function as either a Class I and Class II activator with RNAP and c-di-GMP, allowing VpsR to regulate a large collection of promoters (Chapter 3).

Here we have explored the role of VpsR phosphorylation at P_{vpsL} using the phosphomimetic variant D59E, the phosphodeficient variant D59A, and a specially purified VpsR (VpsR^{ren}) that requires the presence of acetyl phosphate for function. For most RRs, phosphorylation has been found to promote dimerization and/or DNA binding by the RR, leading to an active RR or RR/DNA that can then recruit RNAP. However, we find that phosphorylation of VpsR is not needed for these properties. Instead phosphorylation is needed to generate the correct architecture of the activated transcription complex containing VpsR, c-di-GMP, RNAP, and P_{vpsL} . Our results indicate that together with c-di-GMP, the activity of VpsR can be regulated through

phosphorylation, suggesting that regulation of biofilm biogenesis in *V. cholerae* is dependent on the phosphorylation state of this master regulator.

4.2 Materials and Methods

DNA.

pMLH06 contains the *vpsL* promoter from -97 to +213 cloned into the EcoRI and HindIII restriction enzyme sites of pRLG770 (175). pMLH10 contains the *vpsL* promoter from -97 to +213 cloned into the SpeI and BamHI restriction sites of pBBRlux (74). pMLH17 containing *vpsR* wildtype was constructed as previously described (30). pMLH18 and pMLH19 contain the *vpsR* D59A and D59E, respectively, and were constructed from pMLH17 using QuikChange (Agilent). Primer sequences are available upon request.

pMLH11 is a pET28b(+) derivative (Novagen) that contains *vpsR* cloned as previously described (30). PCR was used to amplify the pET28b(+) vector for restrictionless cloning. Gibson Assembly Master Mix (New England Biolabs) was used to assemble the PCR products and vectors according to manufacturer's instructions. pMLH14 and pMLH15 contain the *vpsR* variant genes encoding VpsR D59A and VpsR D59E, respectively, which were constructed from pMLH11 using QuikChange (Agilent).

pCMW75 contains an active *Vibrio harveyi* DGC, *qrgB*, and was used for generating high levels of c-di-GMP in gene reporter assays (70). pCMW98 contains an inactive *V. harveyi* DGC, *qrgB*, and was used for generating low levels of c-di-GMP (70).

Fragments containing P_{vpsL} used for EMSAs and DNase I footprinting were obtained as PCR products as previously described (30), using primers that annealed from positions -97 to +113 relative to the transcription start site (TSS) of *vpsL*.

Strains and growth conditions.

Cloning strains include *E. coli* ElectroMAX DH10B (Invitrogen) and *E. coli* DH5 α (New England Biolabs), while protein synthesis strains include BL21(DE3) and Rosetta2 (DE3)/pLysS (New England Biolabs). The *V. cholerae* strains, CW2034 ($\Delta vpsL$), WN310 ($\Delta vpsL \Delta vpsR$), and BP52 ($\Delta vpsR$) were derived from the El Tor biotype strain C6707str2 (178). Strains carrying mutations in *vpsR* were constructed using the pKAS32 suicide vector as described (220). For *lux*-fusion assays in *V. cholerae*, strains containing a deletion in *vpsL* were used to prevent cellular aggregation. This allowed us to obtain accurate readings of reporter gene expression at high levels of c-di-GMP by preventing cellular aggregation (70). Overnight strains were diluted 1:100 and 0.5mM IPTG was added to induce the DGC, QrgB, from the plasmid pCMW75, leading to the synthesis of high concentrations of c-di-GMP (70). Strains were grown and *lux* fusion assays with at least two biological replicates and three technical replicates were performed as previously described (30). For *lux*-fusion assays in *E. coli*, MG1655 and $\Delta ackA \Delta pta$ strains were used. Overnight strains were diluted 1:100, immediately induced with 0.02% arabinose to synthesize VpsR and grown to an OD600 of approximately 0.3 prior to 0.5mM IPTG induction of DGC. *Lux* fusion assays were performed as previously described (30).

Proteins.

E. coli RNAP core was purchased from Epicenter Technologies and New England Biolabs (NEB). *E. coli* $\sigma 70$ was purified as previously described (180). WT VpsR, D59A, and D59E protein variants were expressed and purified from Rosetta2 (DE3)/pLysS (Novagen) containing pMLH11, pMLH14, or pMLH15, respectively, as previously described (30). Nonphosphorylated VpsR^{ren} was expressed and purified as previously described with the modifications below (30).

After centrifugation of the induced culture at 13,000 x g for 10 minutes, the cell pellet was resuspended in 20 mL Binding Buffer [20 mM Tris-HCl (pH 7.5), 500 mM NaCl, 5 mM imidazole, and 1 mM PMSF], sonicated, centrifuged at 17,500 x g for 20 minutes, and resuspended again in Binding Buffer. Upon centrifugation at 17,500 x g for 15 min, the pellet was resuspended in 8 ml Binding Buffer containing 6 M Urea (Binding Buffer/6M urea), rocked gently for an hour, and centrifuged at 17,500 x g for 15 minutes. The supernatant was carefully removed to a new centrifuge bottle, centrifuged as before, and filtered through a 0.8 micron syringe filter. After pre-equilibration of the Ni-NTA His-binding resin (QIAGEN) with Binding Buffer/6M urea, 2 ml of the supernatant was mixed with 2 ml of the resin slurry and gently rocked for at least one hour, then loaded into a 10 ml disposable column (Bio-Rad) and washed first with 12 ml Binding Buffer/6 M Urea and then with 12 ml Wash Buffer [20 mM Tris-HCl (pH 7.5), 500 mM NaCl, 10 mM imidazole, 6 M urea]. VpsR was eluted twice with 3 ml Elute Buffer #1 (20 mM Tris-HCl [pH 7.5], 500 mM NaCl, 10 mM imidazole, 6 M Urea, 60 mM imidazole) and then twice with 3 ml Elute Buffer #2 (20 mM Tris-HCl [pH 7.5], 500 mM NaCl, 10 mM imidazole, 6 M Urea, 200 mM imidazole). The eluted protein was renatured by sequential dialysis in the following buffers: 1 X Reconstitution Buffer (RB) (50 mM Tris-HCl [pH 7.5], 1 mM EDTA, 0.1% Triton X-100, 20% glycerol, 1 mM DTT) containing 6M urea, three times; ½ X RB containing 6M Urea; ½ X RB containing 3M urea; 1 X RB containing 3M urea; ½ X RB containing 3M urea; ½ X RB; 1 X RB, three times; and VpsR Storage Buffer (20 mM Tris-HCl [pH 7.5], 1 mM EDTA, 10 mM NaCl, 20% glycerol, 0.1 mM DTT), four times. Dialyzed protein was stored at -20°C. Protein concentrations were determined by comparison with known amounts of RNAP core after SDS-PAGE and gel staining with Colloidal Blue (Invitrogen).

***In vitro* phosphorylation of VpsR with acetyl phosphate**

0.0039g of lithium potassium acetyl phosphate (Sigma) was dissolved and resuspended in 20mM Tris, pH 7.5 to make a final concentration of 200mM acetyl phosphate. A final concentration of 20mM acetyl phosphate was added to VpsR^{ren} and incubated at room temperature for 30 min, forming VpsR^{ren}-P.

BS³ Crosslinking.

Crosslinking reactions containing 1.5 μ M of VpsR, 5 mM of BS³ crosslinker (Thermo Scientific), and 50 μ M of c-di-GMP, as indicated, were incubated at room temperature for 30 min. After quenching with a final concentration of 0.1 mM Tris-Cl (pH 7.5), proteins were separated by SDS-PAGE on a 10-20% (wt/vol) Tris-tricine gel (Invitrogen) and stained with Colloidal Blue (Invitrogen) and/or SilverXpress (Invitrogen).

Electrophoretic Mobility Shift Assays.

Complexes were formed as described previously (30). 5 nM of ³²P-labeled DNA was incubated as indicated with the following: VpsR (final concentration from 0.2 μ M to 2 μ M); 0.16 μ M of reconstituted RNAP (σ :core ratio of 2.5:1); and 50 μ M of c-di-GMP at 37°C for 10 min in transcription buffer. 1 μ g of poly(dI-dC) or 500 μ g of heparin was added to VpsR-DNA complexes or transcription complexes, respectively. VpsR-DNA complexes were loaded onto 5% (wt/vol), nondenaturing polyacrylamide gels already running at 100 V in 1 X Tris/borate/EDTA (TBE) buffer and electrophoresed for 1.5 h. Transcription complexes were loaded onto 4% (wt/vol) nondenaturing, polyacrylamide gels already running at 100 V in 1 X TBE buffer. Samples were electrophoresed for 3 h at 380 V/h. After autoradiography, films were scanned on a Powerlook

2100XL densitometer. Using the Quantity One software (Bio-Rad), $K_{d(app)}$ s were calculated by determining the VpsR concentration needed to shift 50% of the free DNA.

***In Vitro* Transcriptions.**

Single round *in vitro* transcriptions contained 0.02 pmol of supercoiled template, 0 to 3.0 pmol of VpsR, 0 to 50 μ M of c-di-GMP, reconstituted RNAP (0.2 pmol of σ 70 plus 0.05 pmol of core), 1 X transcription buffer [40 mM Tris-acetate (pH 7.9), 150 mM potassium glutamate, 4 mM magnesium acetate, 0.1 mM EDTA (pH 7.0), 0.01 mM DTT, and 100 μ g/mL BSA] in a total volume of 5 μ l. After incubation at 37°C for 10 min, a solution (1 μ l) containing rNTPs (2.86 mM ATP, GTP CTP, and 71 μ M [α -³²P] UTP at 5 x 10⁴ dpm/pmol) and 500 ng heparin were added to each reaction. Reactions were incubated for another 10 min at 37°C and subsequently collected on dry ice with the addition of formamide load solution (15 μ L). 10 μ L aliquots were electrophoresed on 4% (wt/vol) polyacrylamide, 7 M urea denaturing gels for 2500 volt-hrs in 0.5 X TBE buffer. After electrophoresis, gels were exposed to X-ray films, films were scanned, and radioactivity was quantified as described above.

DNase I footprinting.

Solutions were assembled as described above for EMSAs. DNA (0.04 μ M) was incubated with the following as indicated: 1.4 μ M VpsR; 50 μ M c-di-GMP; and/or 0.16 μ M of reconstituted RNAP (σ :core ratio of 2.5:1). After an incubation of 10 min at 37°C, heparin was added to complexes with RNAP for 15 s to ensure binding specificity prior to addition of a 1.5 μ L solution containing 0.3 U of DNase I. No additional competitor was added to complexes containing just VpsR and DNA + - c-di-GMP since these reactions already contained 1 μ g of poly(dI-dC). After

an additional 30 s incubation at 37° C to initiate DNase I cleavage reactions, samples were immediately loaded onto 4% (wt/vol) nondenaturing, polyacrylamide gels already running at 100 V in 1 X TBE buffer. Samples were electrophoresed for 3 h at 380 V/h. After autoradiography, the protein/DNA complexes were excised and extracted DNA was isolated electrophoresed on denaturing gels as described (156).

Potassium Permanganate Footprinting.

For potassium permanganate (KMnO₄) footprinting, reactions were assembled as described above for DNase I footprinting and as described previously (30). Briefly, after addition of 500 ng of heparin to ensure transcription complex stability, a final concentration of 2.5 mM KMnO₄ was added prior to incubation at 37°C for 2.5 min. Reactions were quenched with 5 µL of 14 M 2-mercaptoethanol, and further processed as described (156).

Trypsin proteolysis.

VpsR (1.8 µg) in 1 X Transcription Buffer with and without 50 µM c-di-GMP, as indicated, was incubated for 10 min at 37°C. Trypsin (36 ng, Thermo Scientific) was added to the reaction and aliquots were removed 2, 5, 10, 30, and 60 min after incubation at 37°C. Samples were quenched with 4 X SDS-PAGE loading buffer (200 mM Tris, 6.8, 8% SDS, 0.4% bromophenol blue, 40% glycerol), heated at 95°C, loaded onto a 10-20% (wt/vol) Tris-tricine gel (Invitrogen) and stained with SilverXpress (Invitrogen).

impairs activation while VpsR D59E behaves similarly to WT *in vivo* (Fig. 4.2A) and *in vitro* (Fig. 4.2B, 4.2C). Our primer extension analyses confirmed these results (Fig. 4.3).

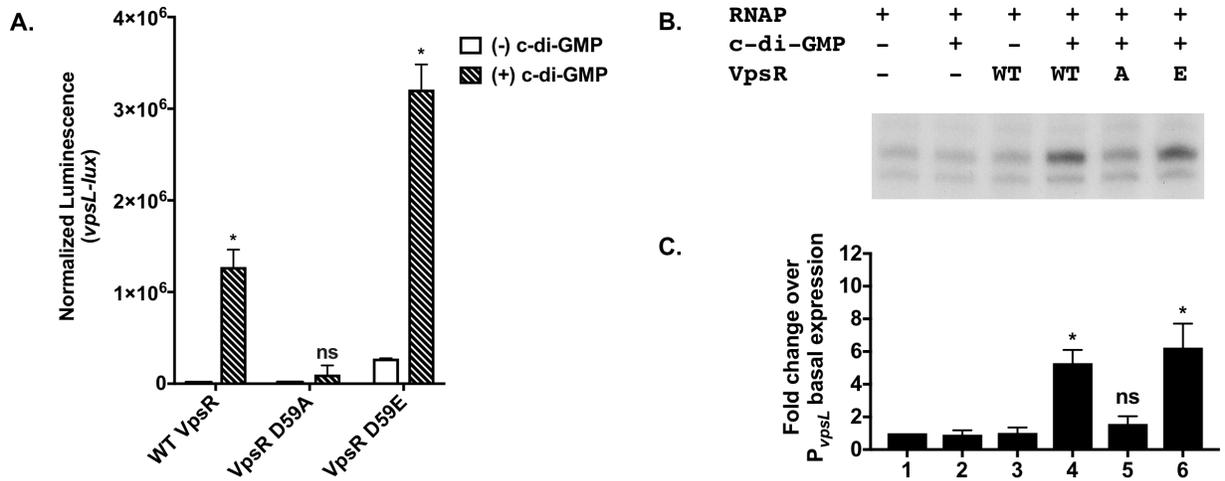
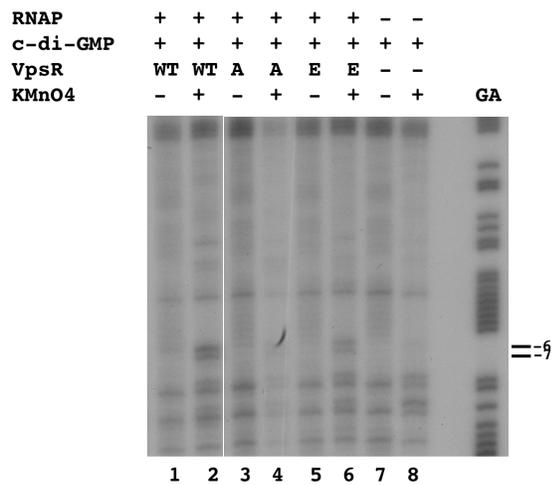


Figure 4.2: VpsR D59 residue is required for *in vivo* and *in vitro* activities at P_{vpsL} . (A) Expression of *vpsL-lux* was determined in WT VpsR, VpsR D59A, or VpsR D59E mutations under low or high c-di-GMP conditions in *E. coli*. Cells were grown to an OD_{500} of approximately 0.5 and induced with a final concentration of 0.2% arabinose to induce VpsR synthesis and/or 1 mM IPTG to induce c-di-GMP synthesis. Error bars indicate standard deviation of three independent cultures. (B) Single round *in vitro* transcription from plasmid template P_{vpsL} with *E. coli* RNAP containing $\sigma 70$ and either WT VpsR, VpsR D59A, or VpsR D59E with or without c-di-GMP. (C) Histogram of at least three independent experiments of multiple round *in vitro* transcriptions with standard deviations.

Similar to our previous results with strains containing WT VpsR (30), we identified two predominant transcriptional start sites (TSS) using RNA isolated from *in vitro* transcriptions as well as from *V. cholerae*: an A nucleotide and a G nucleotide located 59 bp and 57 bp upstream of the *vpsL* AUG, respectively (Fig. 4.3). These 5' ends are observed using RNA made with the phosphomimetic variant D59E (Fig. 4.3A, lane 6; Fig. 4.3B, lane 7), and significantly decreased with the RNA made with the phosphodeficient variant D59A (Fig. 4.3A, lane 5; Fig. 4.3B, lane 6).

A. Nontemplate



B. Template

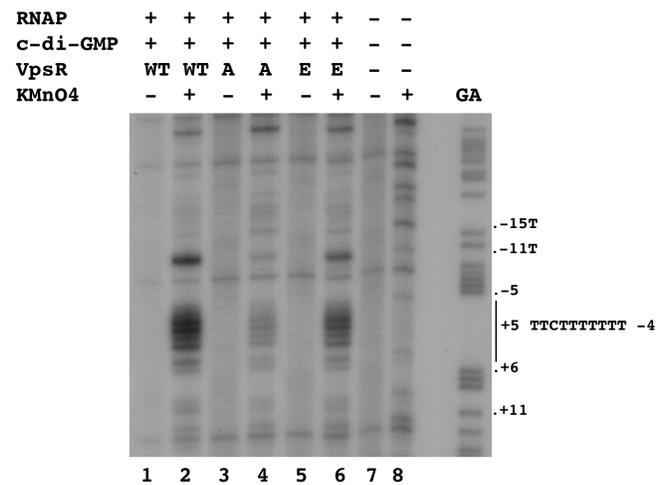


Figure 4.4: WT and VpsR D59E form the open complex with RNAP and c-di-GMP at P_{vpsL} in $KMnO_4$ footprinting. (A) Reactive thymines within the transcription bubble are observed at positions -6 and -7 on nontemplate DNA. (B) Reactive thymines within the transcription bubble are observed at positions -11, -4, -3, -2, -1, +1, and +2 on template DNA. GA corresponds to G+A ladder.

Dimerization and DNA binding by WT VpsR, VpsR D59A, and VpsR D59E are similar.

Multiple studies have found that for many RRs, phosphorylation promotes protein dimerization, which then enhances DNA binding (223-227). We previously demonstrated using BS^3 crosslinking that VpsR forms dimers in the absence and presence of c-di-GMP (30). To determine if VpsR phosphorylation affected this oligomer formation, we repeated the BS^3 crosslinking using the D59A and D59E variants. As seen in Fig. 4.5, the level of crosslinked dimer is similar using WT VpsR, VpsR D59A, or VpsR D59E in the presence or absence of c-di-GMP. These results indicate that unlike many other phosphorylated RRs (223-227), VpsR does not require phosphorylation to dimerize.

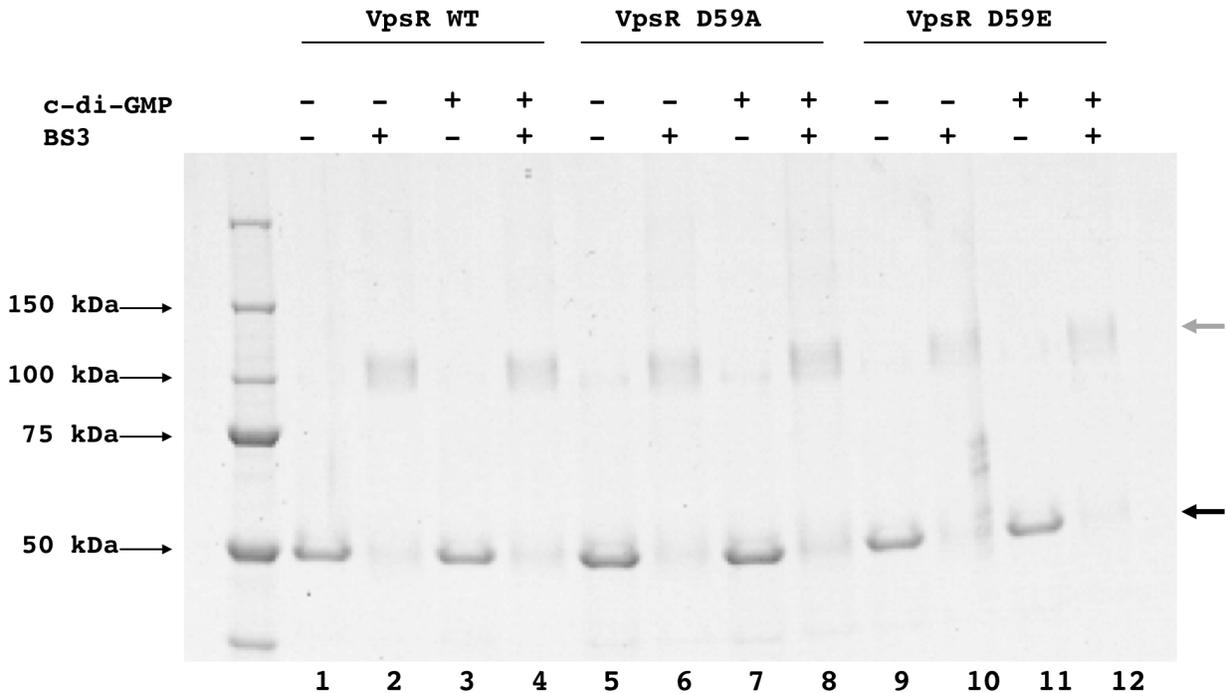


Figure 4.5: WT VpsR, VpsR D59A, and VpsR D59E form dimers *in vitro* with or without c-di-GMP. Samples containing 1.5 mM of VpsR with and without 50 mM c-di-GMP were treated with the chemical cross-linker BS³ as indicated and separated on a 10-20% (wt/vol) Tricine gel that was stained with Colloidal Blue. Black arrows indicate position of VpsR monomer (~50kDa) and grey arrows indicate position of VpsR dimer (~100kDa). Each sample was repeated independently three times, and a representative gel image is shown.

However, the dimers formed by the variants might have different protein conformations. To investigate this, we performed limited trypsin proteolysis in the presence and absence of c-di-GMP (Fig. 4.6). Although there were subtle differences in these patterns among the proteins, there were no differences that suggested a different conformation for the active WT and D59E proteins vs. the impaired D59A protein. These results are again consistent with the idea that phosphorylation is not needed to generate a particular dimer conformation. Presence of c-di-GMP (Fig. 4.6, lanes 6-10) vs. absence of c-di-GMP (Fig. 4.6, lanes 1-5) also do not reveal significant differences in digestion patterns. Thus, like phosphorylation, c-di-GMP do not impart large conformational changes within the protein.

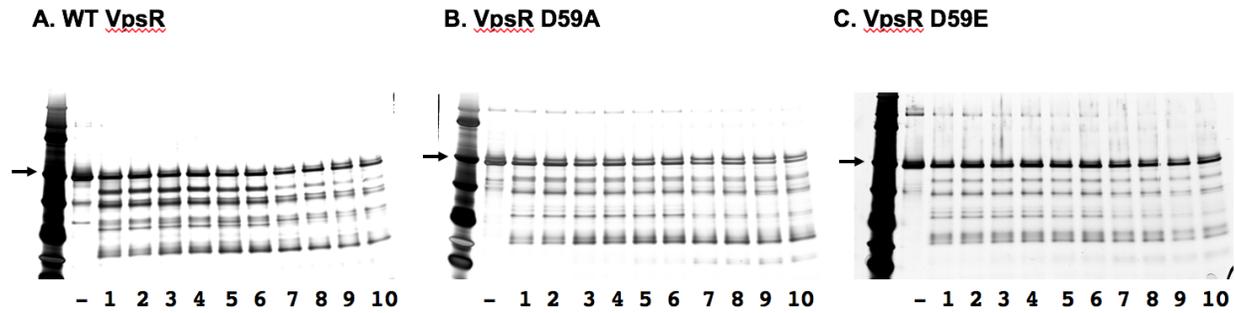


Figure 4.6: Trypsin proteolysis of WT VpsR, VpsR D59A, and VpsR D59E with or without c-di-GMP. Samples containing 1.6 μg of (A) WT VpsR, (B) VpsR D59A, or (C) VpsR D59E were incubated in the absence of c-di-GMP (lanes 1-5) or presence of 50 μM c-di-GMP (lanes 6-10) for 10 min. After digestion with trypsin for 2 min (lanes 1 and 6), 5 min (lanes 2 and 7), 10 min (lanes 3 and 8), 30 min (lanes 4 and 9), and 60 min (lanes 5 and 10), samples were separated on a 10-20% (wt/vol) Tricine gel and stained with SilverXpress.

Next, we used gel retardation assays to investigate whether the phosphodeficient and phosphomimetic D59 substitutions affect the affinity of VpsR for the DNA (Fig. 4.7). We previously reported an apparent dissociation constant ($K_{d(\text{app})}$) for WT VpsR binding to P_{vpsL} by determining the amount of VpsR needed to shift 50% of the DNA. We found that the presence of c-di-GMP did not enhance VpsR binding to the DNA [2.22 μM (+/- 0.64 μM) without c-di-GMP, 1.66 μM (+/- 1.00 μM) with c-di-GMP] (30). VpsR D59A and VpsR D59E with and without c-di-GMP bound the DNA in similar ranges, from 1.04 μM to 1.82 μM (Fig. 4.7D).

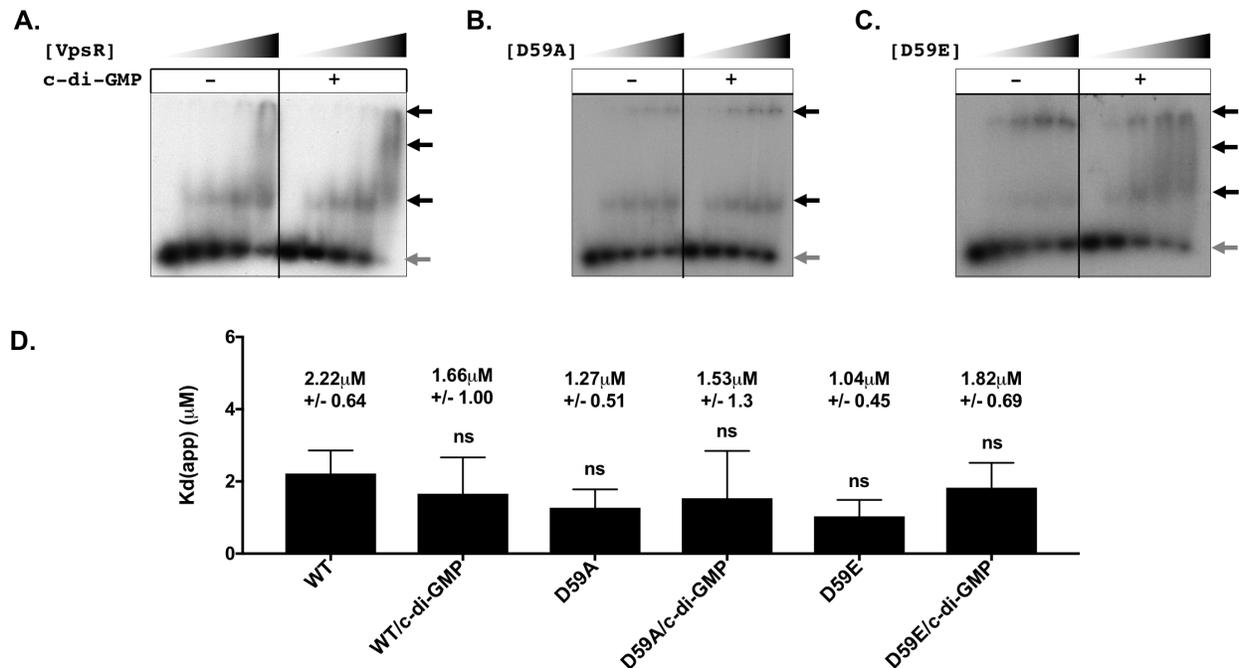


Figure 4.7: WT VpsR, VpsR D59A, and D59E bind P_{vpsL} DNA with similar affinities with or without c-di-GMP. (A-C) Representative gels showing the retardation of the ^{32}P -labeled DNA harboring -97 to +103 of P_{vpsL} with increasing VpsR concentrations either in the absence (lanes 1-5) or presence (lanes 6-10) of c-di-GMP. Black arrows indicate retarded complexes while grey arrow indicates free DNA. (D) Quantitation of gel retardation assays. Apparent DNA-binding dissociation constants ($K_{d(\text{app})}$) were calculated as the concentration of VpsR needed to retard 50% of the free DNA. Values from at least three EMSAs were analyzed using one-way ANOVA with Tukey's HSD *posthoc* analysis (ns, not significant).

To investigate whether the substitutions affect the protein-DNA contacts within the VpsR/DNA complex, we performed DNase I footprinting in the presence or absence of c-di-GMP. To ensure that we were observing the footprint of the stable and specific protein complex of interest, we challenged the complexes with poly(dI-dC), treated them with DNase I, and then isolated the complexes from gel retardation assays prior to isolating the DNase I-cleaved DNA. Like previous studies, the DNase I footprint patterns for WT VpsR were similar in the absence and presence of c-di-GMP, and binding site for WT VpsR extended from -31 to -52 on nontemplate P_{vpsL} and from -34 to -53 on template P_{vpsL} (Fig. 4.8A and 4.8B, lanes 2 and 5) (30,71). The D59A and D59E variants protected the DNA similarly (Fig. 4.8A and 7B, lanes 3, 4, 6, and 7), although

protection was somewhat weaker overall. Thus, we conclude that there are no significant differences in the protein-DNA interactions among the variants in the presence or absence of c-di-GMP.

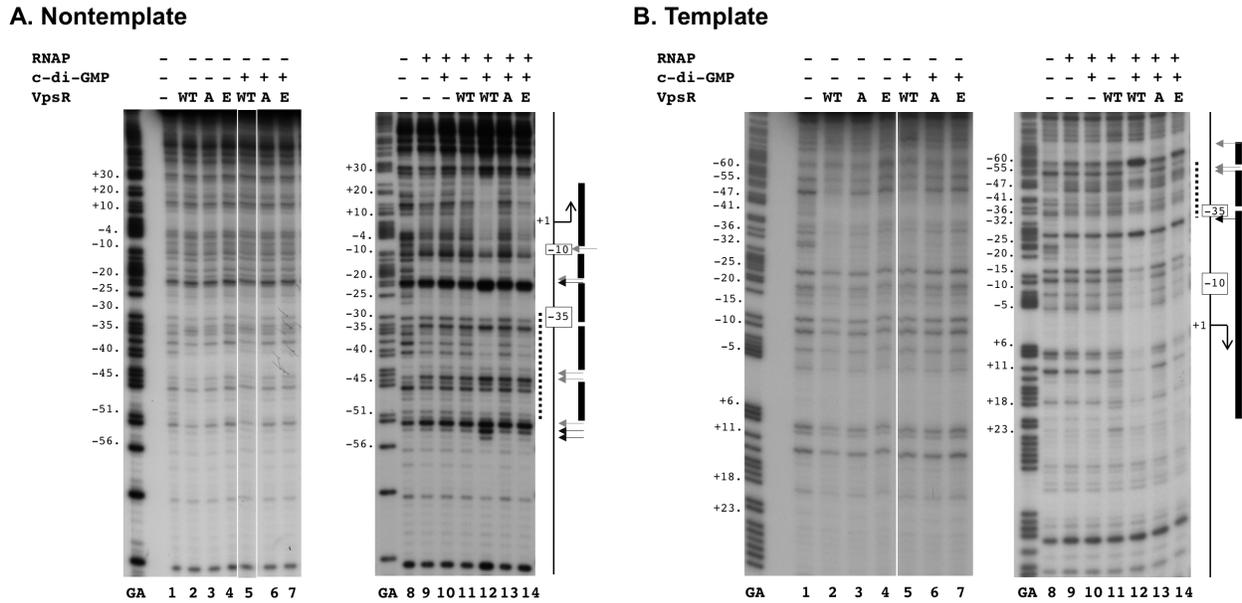


Figure 4.8: DNase I footprinting of P_{vpsL} containing WT VpsR, VpsR D59A, and VpsR D59E complexes on nontemplate DNA and template DNA. GA corresponds to G+A ladder. VpsR, c-di-GMP, and/or RNAP are present as indicated. To the right of each gel image, a schematic indicates the -10 and -35 regions and the +1. The VpsR binding site is indicated as a dashed red line. DNase I protection regions and hypersensitivity sites seen with the activated complex of RNAP, VpsR, c-di-GMP, and DNA are depicted as black rectangles and horizontal arrows.

DNase I footprinting analyses suggest that phosphorylation is needed to form the active transcriptional protein/DNA architecture at P_{vpsL} .

Although either RNAP alone or RNAP/c-di-GMP/D59A yielded a basal level of transcription from P_{vpsL} , both RNAP/c-di-GMP/VpsR and RNAP/c-di-GMP/D59E resulted in activated transcription. Given that the activities of WT VpsR and the D59A and D59E variants were similar in the dimerization and DNA binding assays, these results suggested that the architecture of the transcription complexes made in the presence of c-di-GMP by WT VpsR or VpsR D59E might be the significant difference between the active proteins and the inactive D59A

variant. To investigate whether protein-DNA interactions differed among complexes formed with RNAP and the variants, we performed DNase I footprinting. To minimize nonspecific background DNA cleavage sites and to ensure that the observed footprints originated specifically from the transcription complexes, we treated the complexes with DNase I and then isolated the stable complexes from gel retardation assays prior to isolating the DNA.

We previously reported the DNase I footprints of RNAP alone, RNAP/c-di-GMP, RNAP/VpsR, and RNAP/c-di-GMP/VpsR at P_{vpsL} (30). We repeated these here to compare the complexes made with WT VpsR vs. the phosphovariants. As observed before, the DNase I footprints obtained with the basal transcription complexes formed with RNAP alone, RNAP/c-di-GMP, or RNAP/VpsR at P_{vpsL} were similar and consistent with the formation of a weak transcription complex, while the activated complex of RNAP/VpsR/c-di-GMP at P_{vpsL} generated distinct footprints, consistent with an activated open complex in which the DNA is protected from -58 to +30 (30). In particular, enhanced protection in the downstream portion of the VpsR binding site and enhanced cleavage in the upstream portion of the VpsR binding site were only observed in the presence of RNAP/c-di-GMP (Fig. 4.8A and 4.8B, lane 11), suggesting that the presence of RNAP/c-di-GMP alters the VpsR DNA binding pattern.

Though both the D59A and D59E variants bound the DNA similarly in the presence and absence of c-di-GMP, the presence of the phosphodeficient VpsR D59A, which is inactive for transcription activation (Fig. 4.2), did not significantly alter the basal footprint made by RNAP alone (Fig. 4.8A and 4.8B, lanes 9-11, and 13)]. However, the footprints obtained with VpsR D59E (Fig. 4.8A and 4.8B, lane 14), although weaker, were similar to those of the activated complex (Fig. 4.8A and 4.8B, lane 12). Furthermore, the hypersensitivity sites at -23 on the nontemplate

strand and -59 on the template strand were observed with either WT VpsR and the D59E variant (Fig. 4.8A and 4.8B, lanes 12 and 14).

Taken together, the footprints again demonstrate that VpsR D59 is a residue that is required for activation and suggest that in addition to c-di-GMP, phosphorylation of VpsR or the acidic nature of the D59 residue are required to achieve the distinct, activated transcription complex.

An inactive, nonphosphorylated VpsR, purified after denaturation/renaturation, becomes active after addition of acetyl phosphate.

In our previous study and the experiments described to this point, we used WT VpsR that was heterologously produced in *E. coli* from a plasmid construct (30). This purified protein was active even though it was not subsequently treated with a kinase or with a chemical reagent, such as acetyl phosphate, to ensure phosphorylation. Furthermore, addition of acetyl phosphate to the purified protein did not increase the level of transcription (Fig. 4.9).

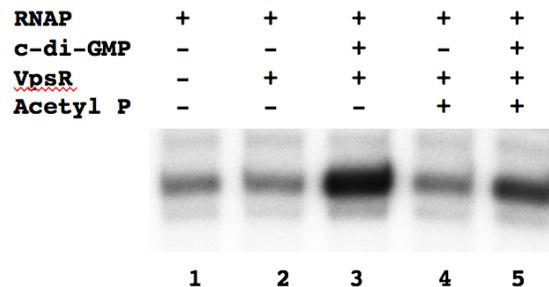


Figure 4.9: Acetyl phosphate has no effect on transcription activation *in vitro* when using the WT VpsR isolated without denaturation. Multiple round *in vitro* transcription from plasmid template P_{vpsL} with *E. coli* RNAP containing σ 70 and VpsR and as indicated, c-di-GMP and/or acetyl phosphate (~P).

Given the results with the phospho-variants, we speculated that perhaps our purified WT VpsR was phosphorylated *in vivo* and retained this phosphorylation throughout the purification process.

However, phosphorylation of an aspartic acid tends to be quite labile (205,206), and we would not expect that it could be retained. Consequently, we considered the following possibilities: 1) D59 is a crucial residue that can be replaced by a glutamic acid, but it is not actually phosphorylated; 2) phosphorylation of D59, which occurs *in vivo* either by acetyl phosphate or through an unknown non-cognate *E. coli* sensor kinase, is atypically stable; or 3) phosphorylation of D59 *in vivo* converts VpsR to a different conformation that then is maintained even if phosphorylation is lost. We hypothesized that if direct phosphorylation of D59 of VpsR was needed, then a purification protocol that involved a denaturation/renaturation step might result in a protein that either lost the phosphorylation or lost the conformation triggered by phosphorylation. It would then need to be phosphorylated again. Previous VpsR purification methods, including ours, have involved purifying the protein from the soluble lysate even though most of the VpsR protein remains insoluble within the pellet. Consequently, we purified VpsR from the insoluble pellet by denaturation in buffer containing 6M urea and then slowly renatured the protein. We refer to this protein as VpsR^{ren}.

Acetyl phosphate has been used with other RRs as a way to phosphorylate an aspartic acid residue in the absence of the HR (154). Consequently, we tested VpsR^{ren} in *in vitro* transcription reactions (Fig. 4.10), KMnO₄ footprinting analyses (Fig. 4.11), BS³ crosslinking studies (Fig. 4.12), trypsin digestions (Fig. 4.13), gel retardation assays (Fig. 4.14), and DNase I footprinting analyses (Fig. 4.15) in the presence or absence of acetyl phosphate.

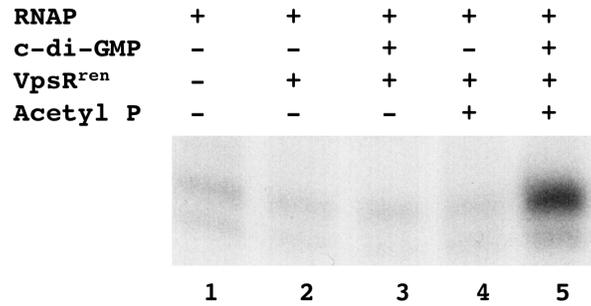


Figure 4.10: *In vitro* transcription activation using VpsR^{ren} requires acetyl phosphate (Ac~P). Single round *in vitro* transcription from plasmid template P_{vpsL} with *E. coli* RNAP containing $\sigma 70$ and either VpsR^{ren} or VpsR^{ren}~P generated by the addition of Ac~P

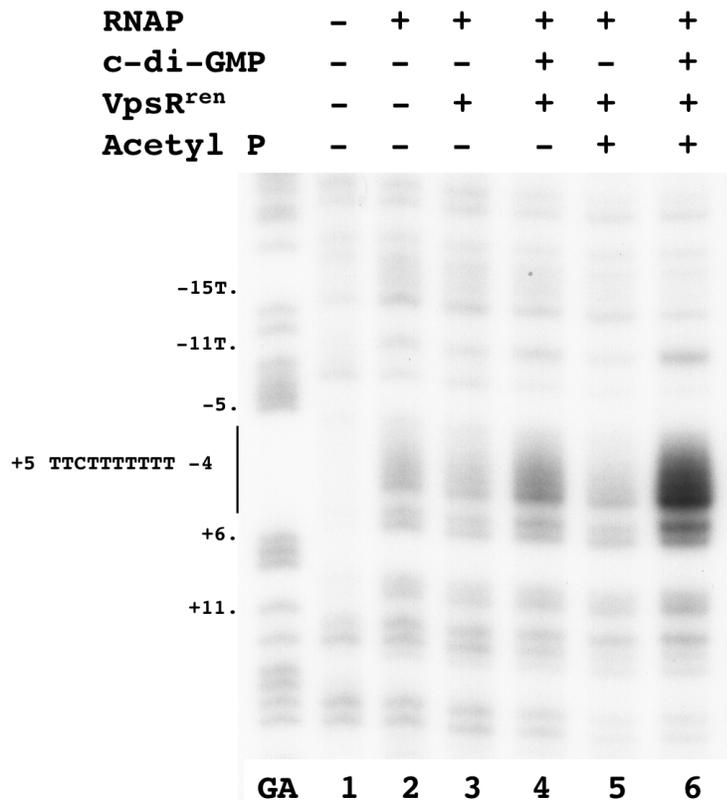


Figure 4.11: Formation of the open complex requires VpsR^{ren} and Ac~P at P_{vpsL}. Denaturing gels show the products obtained after KMnO₄ footprinting. Reactive thymines within the transcription bubble are observed at positions -11, -4, -3, -2, -1, +1, and +2 on template DNA. GA corresponds to G+A ladder.

	VpsR ^{ren}		VpsR ^{ren} ~P		VpsR ^{ren} c-di-GMP		VpsR ^{ren} ~P c-di-GMP	
Acetyl P	-	-	+	+	-	-	+	+
c-di-GMP	-	-	-	-	+	+	+	+
BS3	-	+	-	+	-	+	-	+

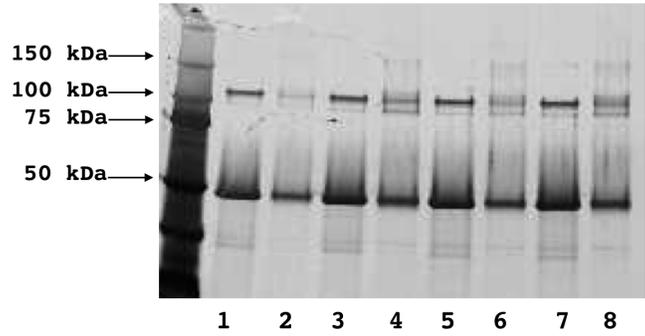


Figure 4.12: VpsR^{ren} and VpsR^{ren}~P form dimers *in vitro* with or without c-di-GMP. Samples containing 1.5 mM of VpsR^{ren} or VpsR^{ren}~P with and without 50 mM c-di-GMP were treated with the chemical cross-linker BS³ as indicated and separated on a 10-20% (wt/vol) Tricine gel that was stained with Colloidal Coomassie Blue. Black arrows indicate position of VpsR monomer (~50kDa) and grey arrows indicate position of VpsR dimer (~100kDa).

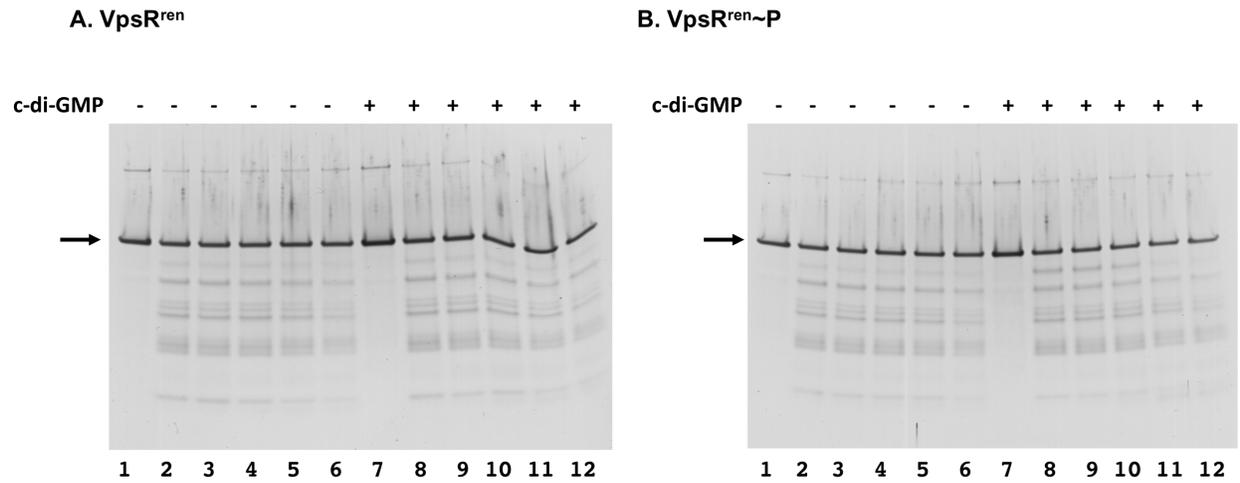
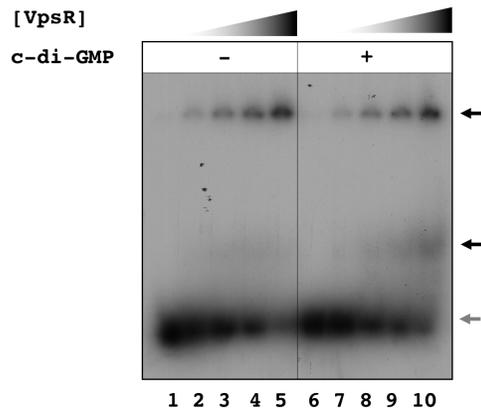


Figure 4.13: Trypsin proteolysis of VpsR^{ren} and VpsR^{ren}~P demonstrates no significant changes in protein conformation with or without c-di-GMP. Samples containing 1.6 μg of protein was incubated in the absence of c-di-GMP (lanes 1-5) or presence of 50 μM c-di-GMP (lanes 6-10) for 10 min. Samples were then digested with trypsin for 2 min (lanes 2 and 8), 5 min (lanes 3 and 9), 10 min (lanes 4 and 10), 30 min (lanes 5 and 11), and 60 min (lanes 6 and 12) prior to quenching with 4 X SDS loading dye. A 10-20% (wt/vol) Tricine gel used to separate cleaved products and stained with SilverXpress, is shown.

A. VpsR^{ren}



B. VpsR^{ren~P}

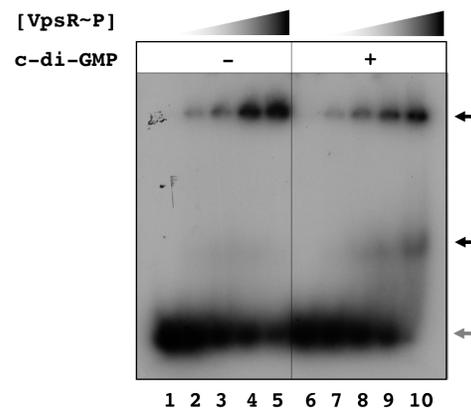
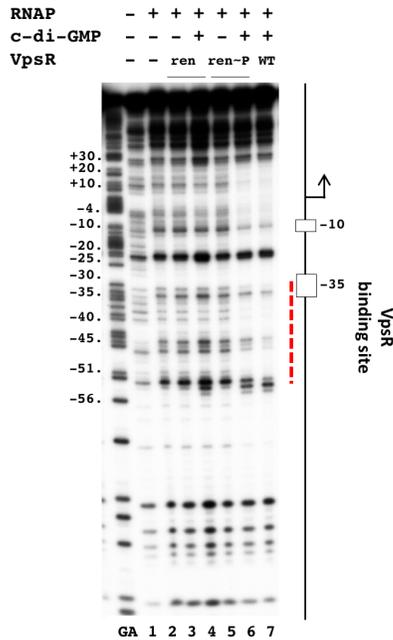


Figure 4.14: VpsR^{ren} and VpsR^{ren~P} bind P_{vpsL} similarly with or without c-di-GMP. Representative gels showing the retardation of the ³²P-labeled DNA harboring -97 to +103 of P_{vpsL} with increasing (A) VpsR^{ren} or (B) VpsR^{ren~P} concentrations either in the absence (lanes 1-5) or presence (lanes 6-10) of c-di-GMP. Black arrows indicate retarded complexes while grey arrow indicates free DNA.

A. Nontemplate



B. Template

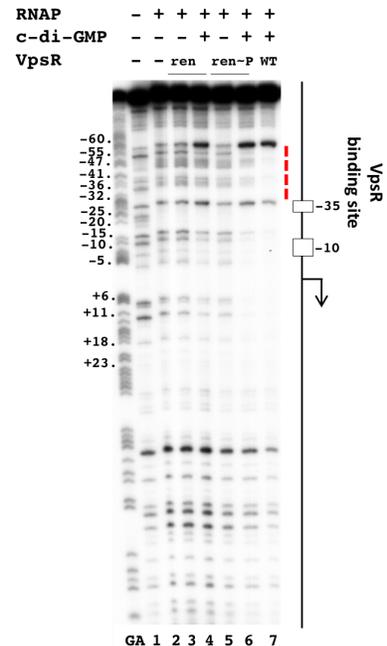


Figure 4.15: DNase I footprinting of P_{vpsL} transcription complexes containing WT VpsR, VpsR^{ren}, and VpsR^{ren~P} on nontemplate DNA and template DNA. GA corresponds to G+A ladder. VpsR WT, VpsR^{ren}, VpsR^{ren~P}, c-di-GMP and/or RNAP are present as indicated. To the right of each gel image, a schematic indicates the -10 and -35 regions and the +1. The VpsR binding site is indicated as a dashed red line.

In all of these assays, VpsR^{ren} in the absence of acetyl phosphate resembled the phosphodeficient variant D59A. However, in the presence of acetyl phosphate, VpsR^{ren} resembled our previously purified WT VpsR and the phosphomimetic D59E variant. In particular, activation of transcription *in vitro* (Fig. 4.10, lane 5), KMnO₄ footprinting, which measures the level of open complex (Fig. 4.11, lane 6), and the distinct DNase I footprints observed with the activated transcription complex (Fig. 4.15, lane 6) were only seen with VpsR^{ren} in the presence of acetyl phosphate. The addition of acetyl phosphate had little to no effect on VpsR dimerization (Fig. 4.12) and protein conformation (Fig. 4.13) as well as DNA binding (Fig. 4.14) in the absence and presence of c-di-GMP. This was expected since these properties were similar whether using WT VpsR, the D59E variant, or the D59A variant.

These results indicate that treatment of VpsR^{ren} with acetyl phosphate is required to generate a protein that is active in transcription, but is not needed for dimerization or DNA binding. Taken with the highly conserved D59 residue, the ability of the D59E variant to replace WT VpsR, and the ability of acetyl phosphate to generate an active VpsR^{ren}, we conclude that phosphorylation of VpsR at D59 is needed for VpsR's transcriptional activity. However, we cannot eliminate the formal possibility that acetyl phosphate acetylates the protein rather than working as a phospho-donor. Our mass spectrometry analyses of previously purified VpsR did not detect phosphorylation or acetylation (data not shown). Given the lability of a phosphorylated aspartic acid residue, it is highly unlikely that we would detect phosphorylation. However, acetylation is usually a stable modification. Consequently, we conclude that along with c-di-GMP, phosphorylation is indeed needed to drive open complex formation to initiate transcription at P_{vpsL}.

Transcription activation by VpsR *in vivo* requires *ackA pta*

Acetyl phosphate is the intermediate of the phosphotransacetylase (Pta)-acetate kinase (AckA) pathway. Pta converts acetyl-CoA and inorganic phosphate to acetyl phosphate and CoA while AckA converts acetyl phosphate to acetate, generating ATP from ADP in the process. Thus, in the absence of Pta, acetyl phosphate is not generated. To investigate whether the level of acetyl phosphate *in vivo* affects transcription from P_{vpsL} , we performed transcriptional *lux* fusion assays in *E. coli* WT MG1655 and MG1655 $\Delta ackA \Delta pta$ strains containing plasmids that generate high levels of c-di-GMP and elevated VpsR concentrations (Fig. 4.16). We find that the level of *lux* activity is ~6-fold lower in the $\Delta ackA \Delta pta$ strain compared to WT MG1655, implicating acetyl phosphate in the generation of active VpsR *in vivo*.

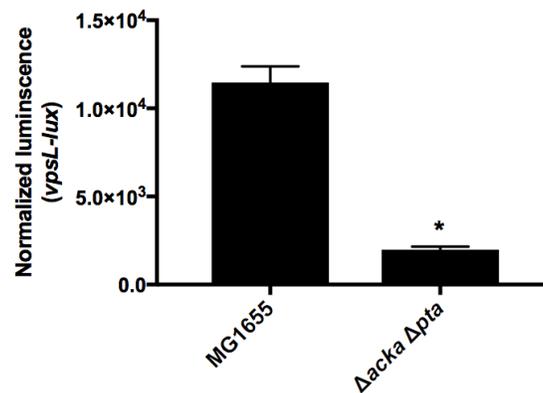


Figure 4.16: *vpsL-lux* gene reporter assays in *E. coli* MG1655 and $\Delta ackA \Delta pta$. Both strains, (A) MG1655 and (B) $\Delta ackA \Delta pta$ contained pMLH10 (*vpsL-lux*), pCMW75 (for high levels of c-di-GMP), and pMLH17 (for VpsR overexpression). Error bars indicate standard deviation of at least three independent cultures analyzed by one-way ANOVA with Tukey's HSD *posthoc* analysis (* $p < 0.05$; ns, not significant).

4.4 Discussion

Regulation of TCSs allows bacteria to quickly adapt to a wide variety of different environments. Despite the extensive biochemical characterizations and identifications of hundreds

of TCSs, research uncovering the role of RR phosphorylation in activating transcription is still emerging. The best characterized phosphorylated RR is the OmpR/PhoB family. Phosphorylation of proteins within this family promotes dimerization, which then significantly increases DNA binding affinity (228). On the other hand, phosphorylation of *Bordetella pertussis* BvgA by its cognate HR BvgS results in a more complicated situation. At the promoter P_{flaB} , phosphorylation of BvgA is needed to improve DNA binding (229). However, at another *B. pertussis* virulence promoter P_{fim3} , phosphorylation of BvgA alters specific protein-DNA contacts in the presence of RNAP (156). Without this phosphorylation, BvgA acts as a repressor by forming an incompetent stable transcription complex with RNAP at P_{fim3} (156). In the case of VpsR, how or even whether phosphorylation affects function has been unclear. There is no cognate kinase upstream of *vpsR*, and VpsR purified after synthesis in *E. coli* from a plasmid-borne *vpsR* is active *in vitro* without the addition of a kinase or acetyl phosphate (30). However, a highly conserved aspartic acid, D59, is present within VpsR, and biofilm formation by *V. cholerae*, which is activated by VpsR, is impaired by the phosphodeficient substitution D59A, but is similar to WT with the phosphomimetic substitution D59E (171). These results strongly suggest that phosphorylation might play a role.

We have investigated whether VpsR is actually affected by phosphorylation by testing the functions of the phosphodeficient D59A variant, the phosphomimetic D59E variant, and VpsR^{ren}, a WT VpsR, which was purified by a protocol involving denaturation/renaturation. Our previous purification procedure involved using the minor fraction of VpsR available in the soluble fraction. We speculated this protein may be phosphorylated by acetyl phosphate or through an unknown non-cognate *E. coli* sensor kinase *in vivo*. It is then active either because it retains this phosphorylation or because phosphorylation generates an active conformation that remains after

phosphorylation is lost. We find that the distinct footprint of the activated complex, open complex formation, and transcription activation all require c-di-GMP and either VpsR^{ren} that has been treated with acetyl phosphate or the D59E variant. VpsR^{ren} that has not been treated with acetyl phosphate behaves like the D59A variant and does not activate transcription above the basal level seen with RNAP alone. Our results are consistent with the idea that phosphorylation of VpsR at residue D59 is indeed needed to generate a protein that can activate transcription. However, it should be noted that we have been unsuccessful in our attempts to detect phosphorylation of VpsR either by Phos-tag gel or by mass spectrometry (data not shown).

Interestingly, formation of the active transcription complex at P_{vpsL} requires both c-di-GMP and phosphorylation of VpsR (Figs. 4.2 and 4.10). Unlike the majority of c-di-GMP-dependent transcription factors and RRs requiring phosphorylation (68,106,113,116,119,126-128,228), neither c-di-GMP nor phosphorylation significantly affect VpsR's ability to bind P_{vpsL} (Fig. 4.7) and form dimers (Figs. 4.5 and 4.12) (30). This then leads to the question, is VpsR phosphorylation dependent on c-di-GMP or vice versa? Our previous mass spectrometry analyses indicate that our previously purified WT VpsR lacks c-di-GMP (30). Since VpsR^{ren} is also devoid of c-di-GMP and behaves similarly to this previously purified VpsR and the phosphomimetic D59E variant, VpsR may have a preference for phosphorylation over c-di-GMP binding. Thus, we hypothesize that VpsR would be phosphorylated first and then bind c-di-GMP second.

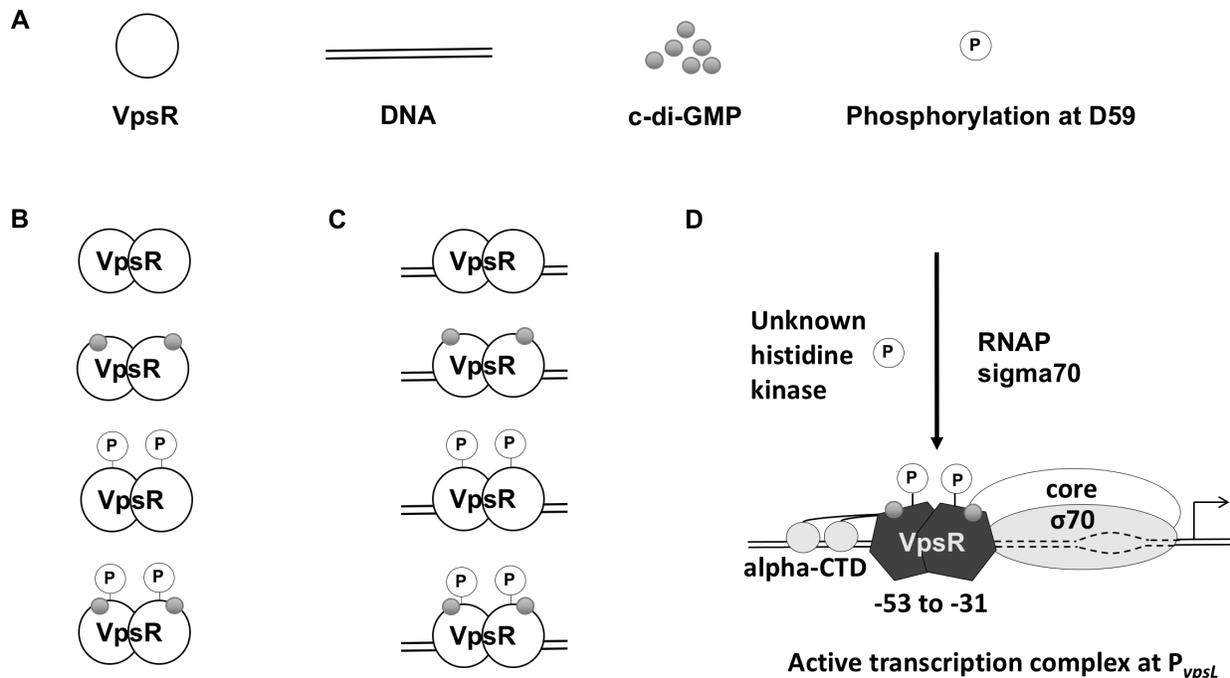


Figure 4.17: Mechanistic model of c-di-GMP-dependent transcription activation at P_{vpsL} . (A) Major players for transcription initiation at P_{vpsL} include VpsR, DNA, c-di-GMP, and possible phosphorylation of VpsR at the D59 residue. (B) VpsR forms a dimer in the presence and absence of phosphorylation at D59 and/or c-di-GMP. (C) VpsR binds the DNA in the presence and absence of phosphorylation and/or c-di-GMP. (D) Upon phosphorylation of VpsR at the D59 residue by an unknown kinase and the presence of c-di-GMP, VpsR adopts a new conformation (now represented as dark hexagons), enabling it to activate transcription and to form the open complex with RNAP.

Our results are consistent with the idea that our heterologously produced VpsR is phosphorylated as it is produced in *E. coli* either by crosstalk from a *E. coli* HK or by acetyl phosphate. We used the Keio collection to screen for an *E. coli* kinase that might be responsible for VpsR phosphorylation. However, we were unsuccessful in these attempts (data not shown), suggesting that when VpsR is produced in *E. coli*, acetyl phosphate is the phosphodonor. Whether there is a specific HK in *V. cholerae* responsible for VpsR phosphorylation and a phosphatase that dephosphorylates is not known. Given the fact that the ability of VpsR to activate P_{vpsL} in a reporter assay is diminished 10-fold in *V. cholerae* lacking *pta*, it appears that acetyl phosphate contributes to VpsR phosphorylation *in vivo* either directly or indirectly. Despite gene expression evidence

supporting RR phosphorylation by acetyl phosphate, the mechanism remains controversial. Several mechanisms have been proposed to explain this link: 1) phosphorylation provides an increase in ATP energy; 2) acetyl phosphate binds an ATPase, resulting in increased activity; or 3) acetyl phosphate directly donates a phosphate independent of a HK (154). The direct phosphodonor model is the most attractive given that about 100 genes respond to the status of the acetyl phosphate pool *in vivo* and that acetyl phosphate can donate to a large subset of RRs *in vitro* including OmpR, NRI, RcsB, and BvgA (154). For these RRs that respond to acetyl phosphate *in vivo*, they appear to possess one of three qualities: their cognate HK contains a phospho-aspartate phosphatase activity (HKP) that functions mainly as a phosphatase; they exist in greater excess over their cognate HK or HKP; or they lack a cognate HK. As mentioned earlier, VpsR lacks a cognate HR, further supporting the possibility that acetyl phosphate is the phosphodonor for VpsR.

Although both c-di-GMP and phosphorylation play a role in transcription initiation by forming the active stable and competent open complex with RNAP, it is possible that phosphorylation alone may also play a distinct role in upregulating gene expression. Because the phosphodeficient D59A variant can still bind to the promoter DNA, we hypothesize that at certain promoters, VpsR could act independently of RNAP and upregulate transcription via anti-repression. For example, while VpsR binds to the *epsC* promoter, *vpsA* promoter, and *vpsT* promoter in gel retardation assays (26,69,71,230), *in vitro* transcriptions at these promoters demonstrate no differences in activation levels relative to basal level expression (data not shown). All three promoters are also directly repressed by H-NS (75-77). Whether VpsR functions as an anti-H-NS repressor remains to be determined.

In conclusion, we have discovered that VpsR is a unique and unusual RR. To our knowledge, VpsR is the first RR that combines two signals, c-di-GMP and phosphorylation, to

directly activate transcription with RNAP to upregulate biofilm genes. A transcriptome comparing WT *E. coli* cells to those of cells that either accumulated acetyl phosphate or cells that were unable to synthesize acetyl phosphate demonstrated that increased levels of acetyl phosphate correlated with decreased expression in flagella biosynthesis as well as increased expression of biofilm genes, type 1 pilus assembly, and stress response (231). This is reminiscent of c-di-GMP-dependent gene regulation in *V. cholerae*. Thus, perhaps VpsR has evolved to use both c-di-GMP and phosphorylation to positively activate biofilm biogenesis genes. As the master regulator of biofilm formation as well as other processes such as T2SS and virulence, VpsR upregulates a wide variety of different promoters with binding sites having various lengths and DNA sequences. Incorporating two signals instead of simply one would allow VpsR to vary its mechanism of transcription activation, allowing fine-tuning of gene expression. Future studies determining the structural conformational changes that occur upon binding to c-di-GMP and/or phosphorylation are in progress.

CHAPTER 5:
Conclusions and Future Directions

Biofilm formation and persistence on implantable devices are problematic in numerous settings including hospitals, industries, and developing countries. Not only do biofilms exhibit significantly decreased antimicrobial susceptibility, but they also prevent proper healing of chronic wounds. Thus, it is thus crucial to understand how biofilm formation is regulated in bacteria to develop new strategies to prevent their formation. The second messenger cyclic dimeric guanosine monophosphate (c-di-GMP) is a central signaling molecule that controls biofilm formation as well as motility, quorum sensing, virulence, cell cycle regulation, and other processes in the vast majority of bacteria (169). However, the molecular mechanisms by which c-di-GMP directly controls transcription remains largely uncharacterized. Previous studies in our laboratory and other laboratories revealed that transcription of the *Vibrio cholerae* biofilm genes is regulated by both c-di-GMP and the atypical enhancer binding protein (EBP) VpsR (30,53,60,71-73,147), although the mechanism by which this occurs remained unknown until my recent work detailed in this dissertation (Chapters 2-4).

The overarching aim of this dissertation is to determine how c-di-GMP signaling regulates biofilm formation to develop new strategies to prevent and treat biofilm-based bacterial infections. To elucidate the mechanism by which c-di-GMP binds and modulates regulators to directly control transcription, I have utilized both *in vivo* and *in vitro* approaches to study how VpsR interacts with c-di-GMP and regulates target genes using *V. cholerae* as a model organism. In chapter 2, I determined that c-di-GMP is required and sufficient to form the activated transcription complex with RNAP containing $\sigma 70$ and VpsR. With an *in vitro* transcription activation of approximately 7-fold over basal expression, c-di-GMP plays an important role in forming the transcription bubble (30). Unlike other c-di-GMP-dependent transcription regulators, c-di-GMP does not play a significant role in VpsR's DNA binding affinity, DNA basepair (bp) contacts in the absence of

RNAP, and dimerization ability (30). In chapter 3, I uncovered new Class I and Class II biofilm promoters directly regulated by both VpsR and c-di-GMP. Lastly, in chapter 4, I further investigated the role of VpsR phosphorylation at the highly conserved D59 residue within the REC domain. My data suggests that along with c-di-GMP, phosphorylation is also required to form the transcription bubble. Like c-di-GMP, phosphorylation does not play a significant role in DNA binding affinity, DNA bp contacts, and oligomerization. Whether phosphorylation is required to bind to c-di-GMP remains to be determined.

Though I have made progress in understanding VpsR's mechanism of transcription activation, my dissertation has led to fascinating unanswered questions, opening many new future avenues of research. A major breakthrough was the development of a new method to purify high quantities of VpsR. These new purification methods should enable us to obtain crystal structures of VpsR in collaboration with Dr. Matthew Neiditch's lab at the University of Rutgers. We plan to solve VpsR structures in four different forms: VpsR, VpsR + c-di-GMP, VpsR~P, and VpsR~P + c-di-GMP. Not only will these structures reveal the conformational changes upon phosphorylation with acetyl phosphate, but they will also give us insights into the dimerization pocket and c-di-GMP binding pocket. Though my trypsin proteolysis experiments demonstrate no striking differences in protein digestion patterns/protein conformation in the presence and absence of phosphorylation and/or c-di-GMP, the conformational changes may be subtle, preventing the visualization of these changes in a trypsin proteolysis assay or the conformational changes may occur only in the presence of RNAP containing $\sigma 70$. Solving the structure of VpsR will enable us to model VpsR with available structures of RNAP. Alternately, we may be able to obtain cryo-EM structures of the activated transcription complex. Understanding how VpsR interacts with the DNA with and without c-di-GMP and/or phosphorylation within the activated transcription complex

with RNAP will give us significant insights into the mechanism of VpsR's unusual c-di-GMP-dependent transcription initiation process.

We can also use our high throughput gene reporter assay to complement VpsR structural studies. Using error-prone PCR with manganese chloride and targeted arginine site-directed mutagenesis, respectively, I have scanned the VpsR protein for mutations that affect binding to c-di-GMP using both a nonbiased and biased approach. Upon expression of *vpsR* mutants under the arabinose promoter in a *V. cholerae* $\Delta vpsR$ background containing a reporter plasmid with *vpsL-lux* and another plasmid with the GGDEF enzyme QrgB, I have isolated *vpsR* mutants belonging to one of four classes: 1) wildtype-like, 2) inactive, 3) blind to c-di-GMP with low transcriptional activity, or 4) constitutively active in the presence or absence of c-di-GMP at a level similar to WT with c-di-GMP (Fig. 5.1, WT green bar). This is a powerful screen since any mutation that renders the protein inactive will be immediately eliminated from subsequent analyses. So far, I have screened approximately 2000 mutants and identified several VpsR residues that appear to be involved in binding to c-di-GMP *in vivo*. These VpsR mutants demonstrate only small differences between VpsR induction (Fig. 5.1, red bars) vs. VpsR/c-di-GMP induction (Fig. 5.1, green bars). I have followed up with one constitutively active mutant (I170T) and four c-di-GMP blind mutants (R193H, R274C, I275C, L325S) for further studies.

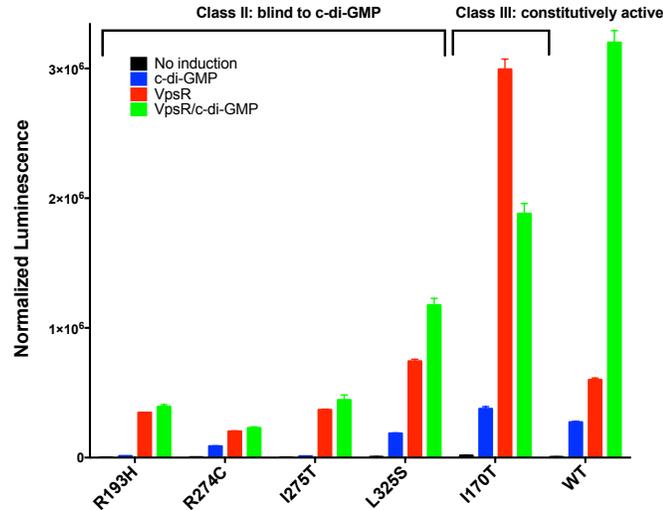


Figure 5.1: Three-plasmid genetic screen to isolate VpsR mutants that are blind to c-di-GMP or constitutively active in the absence of c-di-GMP. VpsR mutants were isolated using error-prone PCR and tested for *vpsL-lux* expression with cells harboring three total plasmids as indicated in the text. Cells were grown to OD₆₀₀ of 0.5 with final concentration of 0.2% arabinose and/or 1 mM IPTG for induction. Blue bar indicates induction of GGDEF to produce high levels of c-di-GMP. Red bar indicates induction of VpsR with arabinose. Green bar indicates induction of both VpsR and c-di-GMP with arabinose and IPTG, respectively. Error bars represent standard deviation with three independent cultures.

To test whether these VpsR mutants can bind c-di-GMP *in vitro*, I have employed Differential Radial Capillary Action of Ligand Assay (DRaCALA). This powerful technique allows the use of induced whole cell lysates to probe protein-ligand interactions, eliminating the need for protein purification (232). After cloning individual VpsR mutations into the pET system and confirming protein expression with SDS-PAGE, I tested each mutant's ability to bind to ³²P-c-di-GMP. My preliminary data demonstrate that the VpsR mutants from Fig. 5.1 all exhibit decreased binding to ³²P-c-di-GMP relative to wildtype VpsR *in vitro* (Fig. 5.2A and 2B).

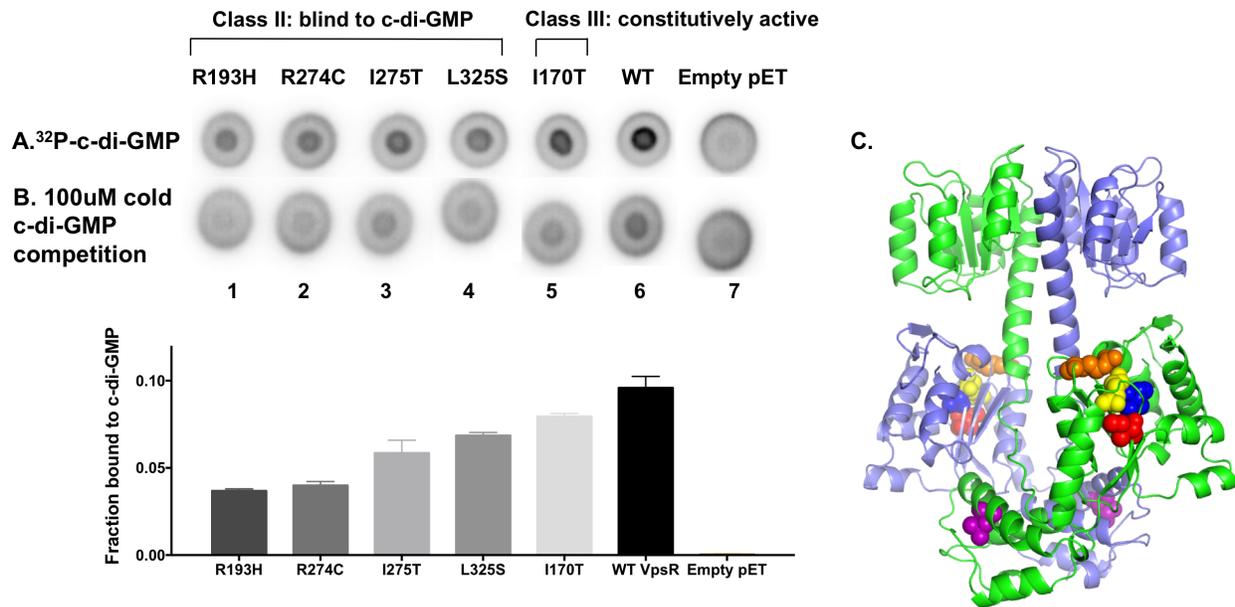


Figure 5.2: Characterization of VpsR mutants binding to c-di-GMP by DRaCALA. (A) 6 μ M of 32 P-c-di-GMP were incubated with VpsR protein lysates and later competed with 100 μ M cold c-di-GMP and dotted on nitrocellulose (B) Three technical replicates were performed and analyzed using Graphpad Prism, 7.0. (C) Mapping mutant amino acid residues on VpsR homology model. Mutants include constitutively active mutant, I170T (red), and c-di-GMP blind mutants, R193 (orange), R274 (yellow), I275 (blue), and L325 (purple). A comparative model of VpsR was generated using NtrC1 crystal structure (PDB: 4I4U). All VpsR mutants map to the AAA+ domain.

Thus far, all mutants are located in the AAA+ domain of VpsR. Of the 5 identified amino acid mutants, 4 mutations are clustered in one location on our VpsR homology model based on NtrC1 structure (PDB: 4I4U) (Fig. 5.2C). I have also purified and characterized two of these mutants, the c-di-GMP blind VpsR R193H and the constitutively active VpsR I170T, in EMSAs and *in vitro* transcriptions to determine their effects on DNA-binding and transcription activation. Though both variants bind DNA and dimerize (data not shown), neither behaves *in vitro* as predicted from their *in vivo* characteristics. R193H is defective in transcription activation, while I170T still requires c-di-GMP for transcription activation (Fig. 5.3). Biofilm phenotypes for each mutant could also be tested using crystal violet staining. In this assay, overnight cultures would be transferred to a 96-well polystyrene plate with a minimum biofilm eradication concentration (MBEC) lid. Biofilms

would then be washed with ethanol and stained with crystal violet prior to quantification at an absorbance of 595 nm (233,234).

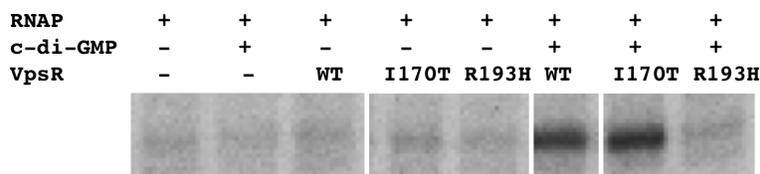


Figure 5.3: Single round *in vitro* transcriptions of VpsR mutants. Reactions were performed with plasmid template P_{vpsL} with *E. coli* RNAP containing σ 70, with or without c-di-GMP, and either WT VpsR, c-di-GMP blind mutant VpsR R193H, or constitutively active mutant VpsR I170T.

In addition, bacterial adenylate cyclase two-hybrid (BACTH) assays can be used to complement these structural studies. These analyses can identify possible protein-protein interactions within the activated transcription complex between VpsR and RNAP or within the VpsR dimerization pocket between VpsR and VpsR in the absence and presence of c-di-GMP. In this system, which has been previously used to study c-di-GMP-dependent protein-protein interactions (235), two fragments of pertussis adenylate cyclase, T25 and T18, need to physically close to synthesize cyclic-AMP, a small molecule activator of the lactose and maltose operons, in a *E. coli cya-* strain (236). A protein or protein domain is fused to T25 and its putative partner is fused to T18. If these proteins interact, T18 and T25 are in close proximity and adenylate cyclase functions, synthesizing cAMP. I constructed a bait plasmid containing VpsR fused to the T18 domain with and without QrgB cloned downstream of the fused VpsR protein. As prey plasmids, I constructed the following fused to the T25 domain: RNAP subunits and VpsR with and without a downstream QrgB. After transformation into BTH101, *E. coli* that lacks the native *cyaA* gene, overnight cultures were spotted onto LB agar supplemented with appropriate antibiotics, IPTG, and X-galactose. Blue colonies represent positive interactions while white colonies denote no

interaction. Dot reporter plates demonstrate strong interactions between VpsR and VpsR with and without c-di-GMP, supporting my BS³ crosslinking studies. I also observed a signal consistent with a strong interaction between VpsR and either full-length RpoA or RpoA^{NTD} (Fig. 5.4). Although no interaction was observed between VpsR and RpoA^{CTD}, this assay is known to have false negatives and thus this result does not exclude a potential interaction.

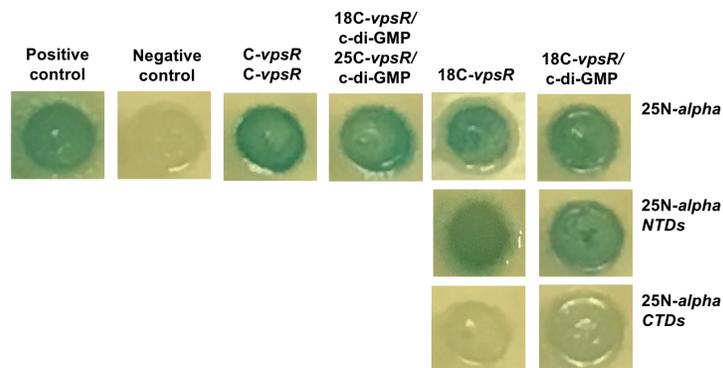
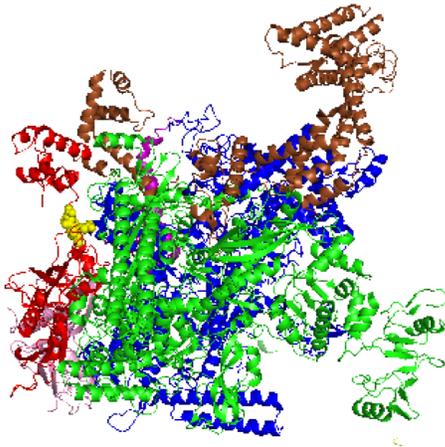


Figure 5.4: Bacterial adenylate cyclase two hybrid (BACTH) assay of VpsR +/- c-di-GMP and α subunit domains. Reconstitution of CyaA, indicative of protein-protein interaction, was detected by β -galactosidase activity on LB plates containing ampicillin (100 μ g/mL), kanamycin (50 μ g/mL), IPTG (0.5mM), and X-gal (40 μ g/mL). Plates were incubated at 30°C for 48 hrs.

I performed Miller assays to quantify the activity of betagalactosidase and further determine strength of protein-protein interactions. Since *E. coli* CRP is the most well-characterized regulator with genetic and biochemical analyses defining the interaction pocket between CRP and α NTD, I used the interacting domain of the α NTD with CRP to select *V. cholerae* *rpoA* residues that could potentially interact with VpsR. I then mutated the corresponding *V. cholerae* *rpoA* residues, E163, E164, D165, and E166, that could potentially be interacting with VpsR (237). My Miller assays demonstrated a modest, but reproducible decrease in binding interaction for these mutants (Fig. 5.5).

A.



B.

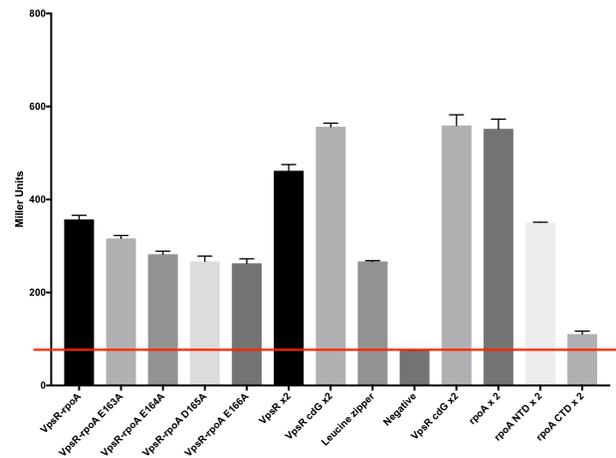


Figure 5.5: Miller assay investigating interactions between VpsR +/- c-di-GMP and α . (A) Mutant α residues (yellow spheres) are mapped to the crystal structure *E. coli* RNAP (PDB: 4YG2). α dimers are represented in red and pink, β represented in green, β' represented in blue, ω represented in purple, and $\sigma 70$ represented in brown. (B) The level of β -galactosidase (in Miller units) observed in the 2-hybrid assays using the indicated protein-protein pairs. Red line indicates negative control. Values above the red line represent positive interactions while values around or below the red line represent negative interactions. Empty vector plasmids were used as negative controls.

In the future, additional *rpoA* as well as *vpsR* residues can be screened either through direct selection or utilizing a random mutagenesis approach with error-prone PCR as described above for the high throughput reporter assay screen. Similar to the high throughput reporter assay, VpsR mutants can be overexpressed and purified and tested in EMSAs, *in vitro* transcriptions, BS³ crosslinking, DRaCALA, and crystal violet assays to examine DNA binding affinity, transcription activation, dimerization, binding to c-di-GMP, and biofilm formation, respectively. Additionally, RpoA mutants can also be tested *in vivo* in a reporter assay, which has been used previously by another group looking at mutant α^{NTD} and CRP interactions (238). In this assay, *rpoA* mutants are overexpressed on one plasmid, the CRP is overexpressed on the second plasmid, and the promoter is fused to *lacZ* on the chromosome (238). Furthermore, because we also routinely purify *E. coli*

RNAP in the Hinton laboratory, we can also express, purify, and test these α mutants in *in vitro* transcriptions. To further validate protein-protein interactions, pull-down assays can be performed with purified His-tagged VpsR bound to c-di-GMP and RNAP subunits. The bait protein, His-tagged VpsR, would be bound to a Ni-NTA affinity column prior to addition of one of the RNAP subunits. Positive protein-protein interactions can be identified on denaturing polyacrylamide gels and stained with Coomassie Blue.

Along with protein-protein and protein-signal interactions, protein-DNA interactions within the activated transcription complex can be further investigated using FeBABE assays. FeBABE is a powerful biochemical cleavage tool that conjugates to cysteine residues (239). Addition of appropriate reagents induces the Fenton reaction and generates a chemical bomb due to production of hydroxyl radicals, thereby cleaving DNA and protein near the chelated site (239). With our VpsR homology model, we can select residues within and/or surrounding the DNA binding domain to mutate into cysteines. We can thereby obtain a molecular map of VpsR/c-di-GMP +/- RNAP transcription complex on the DNA.

Furthermore, with the ability to purify nonphosphorylated VpsR, we can now also look at the role of nonphosphorylation vs. phosphorylation at other VpsR-regulated promoters using acetyl phosphate as the phosphate donor. Recent studies reveal that nonphosphorylated RRs are not simply just inactivated versions of RRs. Instead, they also play a role in transcription activation. For example, *Salmonella enterica* SsrB binds to the *csgD* promoter in the absence of phosphorylation and activates transcription by relieving H-NS silencing (160,191). However, upregulation of Salmonella Pathogenicity Island-2 genes requires phosphorylation of SsrB due to direct interaction with RNAP at this promoter (159). Interestingly, in gene reporter assays at *vpsL* and *vpsT* promoters, VpsR/c-di-GMP and H-NS up-regulate and down-regulate these genes,

respectively (30,69,71,76,77). Gel shift assays and DNase I footprinting demonstrate that both regulators bind to both pieces of DNA (data not shown) (30,69,71,76,77). However, in *in vitro* transcriptions, WT VpsR (presumably phosphorylated) and c-di-GMP together activate *vpsL*, while *vpsT* does not (data not shown) (30), suggesting that perhaps VpsR also uses two modes of transcription activation, similar to SsrB. Further studies using competition assays between VpsR +/- c-di-GMP vs. H-NS can be investigated. *In vitro* transcriptions can also be used to determine if VpsR can function as an anti-H-NS repressor, relieving repression and upregulating gene expression. Since VpsR/c-di-GMP also do not directly activate *epsC*, *tfoY*, and *vpsA* in *in vitro* transcriptions (data not shown), this mechanism may apply at these promoters. Single molecule atomic force microscopy could also be employed to look at protein oligomerization as well as conformational changes associated with DNA binding of each regulator.

Along with looking for downstream effects associated with phosphorylation, we can also investigate *vpsR* regulation. My studies in chapter 4 demonstrate that the *E. coli* $\Delta ackA\Delta pta$ mutant is unable to activate P_{vpsL} in gene reporter assays. These mutants can also be constructed in *V. cholerae*, which encodes one *pta* gene and two *ack* genes: *ack1* and *ack2*. Mass spectrometry can be used to confirm that these mutants contain low levels of acetate and acetyl phosphate. We would expect that activation would be low in a $\Delta ack1\Delta ack2\Delta pta$ mutant but high in a Δpta strain. Though these studies would not rule out the possibility that a kinase may still exist, we can additionally perform a random transposon mutagenesis screen to identify the kinase and/or regulators involved in *vpsR* promoter regulation.

Though I have identified three new promoters directly activated *in vitro* by VpsR, c-di-GMP, and RNAP, it has become increasingly evident that VpsR regulates additional promoters beyond biofilm gene expression. For example, VpsR directly binds to the promoters, *epsC*, *tfoY*,

and *aphA*, regulating T2SS, motility, and virulence, respectively (26,69,147,173). Though I did not observe transcription activation *in vitro* at P_{espC} and P_{tfoY} , it is possible that I did not have the proper conditions for activation. In collaboration with Dr. Peter Freddolino at the University of Michigan, we are searching for additional promoters that are transcriptionally regulated by VpsR and c-di-GMP using a combination of RNA-seq and *In vivo* Protein Occupancy Display High Resolution (IPODHR) (240). We have grown, induced, crosslinked, and harvested WT *V. cholerae* and $\Delta vpsR$ strains containing IPTG-inducible QrgB or QrgB*, allowing us to achieve high and low c-di-GMP levels, respectively. After DNA fragmentation and isolation, we used deep sequencing and looked for genes that have decreased promoter occupancy and decreased gene expression in $\Delta vpsR$. Preliminary data appears promising. Positive hits include many of the promoters that are directly regulated by VpsR and c-di-GMP: *vpsT*, *vpsL*, *aphA*. Full analysis of the data will allow us to identify additional promoters that are bound by and are transcriptionally regulated by both VpsR and c-di-GMP. Upon identification of these promoters, we can then test these promoters in a wide variety of assays that I have already developed: gene reporter assays, *in vitro* transcriptions, EMSAs, and DNase I footprinting.

Lastly, *vps* regulation is a complex and multistep process requiring multiple activators, repressors, and signaling molecules. At the *vpsL* promoter, together with c-di-GMP, VpsR and VpsT upregulate transcription while H-NS represses transcription under LCD *in vivo*. I have purified VpsT and H-NS and have attempted to construct this entire regulation *in vitro*. My data show that VpsT alone with RNAP and c-di-GMP does not activate transcription, but H-NS alone inhibits transcription (Fig. 5.5A and 5.5B). Addition of both activators (VpsR and VpsT) and c-di-GMP as well as the repressor H-NS reveals that activation still occurs (Fig. 5.5E). Mechanistic experiments incubating H-NS first show that VpsR and VpsT can relieve H-NS repression to

activate transcription (Fig. 5.6). Likewise, prior incubation with VpsR and VpsT prevents H-NS from repressing transcription (Fig. 5.6). Optimization of this *in vitro* system by varying concentrations of each regulator will further allow us to understand the mechanism of this complex regulation. To further mimic the system, we can use purified *V. cholerae* RNAP instead of *E. coli* RNAP. With the help of Genscript, we have constructed a plasmid expressing *V. cholerae* RNAP with a His-tag at the N-terminus of *rpoB*. Expression and purification has been thus far unsuccessful, but we plan to try different expression conditions and/or redesign the plasmid.

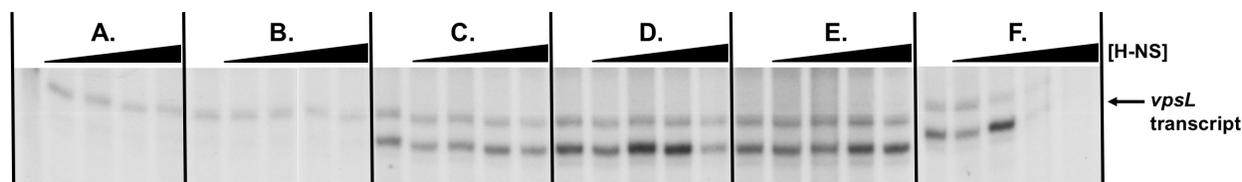


Figure 5.6: VpsR, VpsT, c-di-GMP, and H-NS regulation at P_{vpsL} . Single round *in vitro* transcription reactions were performed to look at the mechanism of transcription regulation at P_{vpsL} . (A) Reactions were performed with RNAP/VpsT/c-di-GMP preincubation prior to H-NS incubation from 0 to 7.5 μ M. (B) Reactions were performed with RNAP preincubation prior to H-NS incubation from 0 to 7.5 μ M. (C) Reactions were performed with RNAP/VpsR/c-di-GMP preincubation prior to H-NS incubation from 0 to 5 μ M. (D) Reactions were performed with RNAP/VpsR/VpsT/c-di-GMP preincubation prior to H-NS incubation from 0 to 5 μ M. (E) All components were incubated on ice simultaneously prior to initiation transcription. (F) Reactions were performed with H-NS incubation prior to incubation with RNAP/VpsR/VpsT/c-di-GMP.

In conclusion, I have made significant progress in understanding the mechanism by which c-di-GMP interacts with transcriptional regulators to directly modulate gene expression with RNAP. I was the first to determine the transcriptional mechanism of c-di-GMP-dependent activation of *V. cholerae* biofilm genes, but I have also investigated the role of VpsR phosphorylation in this mechanism and I have identified additional VpsR-regulated promoters. Based on this work, future studies should be able to construct a molecular map of both protein-DNA and protein-protein interactions detailing how VpsR and c-di-GMP function together and interact with RNAP and the DNA. By exploring this transcriptional mechanism, my research has

had three important major impacts: I have established a new paradigm in c-di-GMP-dependent transcription activation, I have elucidated mechanistic processes that regulate biofilm formation, and most importantly, I have provided the foundation needed for the development of new inhibitors that target *V. cholerae*, biofilm-based infections, and EBP/c-di-GMP regulatory pathways. As atypical EBPs, such as VpsR, and c-di-GMP are widespread in bacteria and responsible for fundamental processes including virulence, biofilm formation, and motility, understanding how this key transcription factor family is regulated by c-di-GMP will allow us to predict novel c-di-GMP-dependent signal transduction pathways in other bacteria and identify new targets to directly modulate this gene expression.

APPENDIX

Table A.1: Table of plasmids and primers used in this study.

Plasmids	Description	Source
pMLH06	<i>vpsL</i> promoter -97 to +213 cloned into pRLG770	(30) Chapter 2
pMLH10	<i>vpsL</i> promoter -97 to +213 cloned into pBBRlux	(30) Chapter 2
pCMW75	IPTG-inducible <i>V. harveyi</i> diguanylate cyclase <i>qrgB</i>	(70)
pCMW98	IPTG-inducible <i>V. harveyi</i> diguanylate cyclase <i>qrgB</i> mutant (GG→AA)	(70)
pMLH17	<i>vpsR</i> cloned into pHERD20T	(30) Chapter 2
pMLH18	<i>vpsR D59A</i> cloned into pHERD20T	Chapter 4
pMLH19	<i>vpsR D59E</i> cloned into pHERD20T	Chapter 4
pMLH11	<i>vpsR</i> cloned into pET28b(+)	(30) Chapter 2
pMLH14	<i>vpsR D59A</i> cloned into pET28b(+)	Chapter 4
pMLH15	<i>vpsR D59E</i> cloned into pET28b(+)	Chapter 4
pMLH40	<i>rbmA</i> promoter -222 to +120 cloned into pRLG770	Chapter 3
pMLH41	<i>rbmF</i> promoter -209 to +167 cloned into pRLG770	Chapter 3
pMLH42	<i>vpsU</i> promoter -130 to +155 cloned into pRLG770	Chapter 3
pMLH48	C18- <i>vpsR</i> for BACTH	Chapter 5
pMLH49	C18- <i>vpsR-qrgB</i>	Chapter 5
pMLH50	C25- <i>vpsR</i> for BACTH	Chapter 5
pMLH54	C18- <i>vpsR-qrgB</i> for BACTH	Chapter 5
pMLH81	N18- <i>rpoA</i> for BACTH	Chapter 5
pMLH82	N18- <i>rpoA</i> NTD for BACTH	Chapter 5

Table A.1 (cont'd):

Plasmids	Description	Source
pMLH83	N18- <i>rpoA</i> CTD for BACTH	Chapter 5
pMLH91	N18- <i>rpoA</i> for BACTH	Chapter 5
pMLH92	N18- <i>rpoA</i> NTD for BACTH	Chapter 5
pMLH93	N18- <i>rpoA</i> CTD for BACTH	Chapter 5
Radiolabeled DNA for EMSA and DNase I footprinting	<i>vpsL</i> promoter -97 to +103	Chapter 2
Radiolabeled DNA for EMSA and DNase I footprinting	<i>rbmA</i> promoter -222 to +120	Chapter 3
Radiolabeled DNA for EMSA and DNase I footprinting	<i>rbmF</i> promoter -209 to +167	Chapter 3
Radiolabeled DNA for EMSA and DNase I footprinting	<i>vpsU</i> promoter -130 to +155	Chapter 3
pKAS VpsR D59A	allelic exchange vector, <i>vpsR</i> D59A	Chapter 4
pKAS VpsR D59E	allelic exchange vector, <i>vpsR</i> D59E	Chapter 4

Table A.2: Strains of *Vibrio cholerae* and *Escherichia coli* used in this study.

Strain name	Description	Reference/Source
ElectroMAX DH10B	<i>E. coli</i> strain used for cloning	Invitrogen
DH5 α	<i>E. coli</i> strain used for cloning	NEB
BL21 (DE3)	<i>E. coli</i> strain used for protein purification	NEB
Rosetta2 (DE3) pLysS	<i>E. coli</i> strain used for protein purification	NEB
S17- λ pir	<i>E. coli</i> strain used for conjugation	(179)
2781	<i>E. coli</i> MG1655	Alan Wolfe
2785	<i>E. coli</i> <i>ackA</i> -, <i>pta</i> -	Alan Wolfe
C6707str2 (WT)	<i>V. cholerae</i> El Tor strain	(178)
CW2034	<i>vpsL</i> -	(70)
WN310	<i>vpsL</i> - <i>vpsR</i> -	(69)
BP52	<i>vpsR</i> -	Chapter 4
BP72	<i>vpsL</i> -, <i>vpsR</i> D59A	Chapter 4
WN314	<i>vpsL</i> -, <i>vpsR</i> D59E	Chapter 4

REFERENCES

REFERENCES

1. Ross, P., Weinhouse, H., Aloni, Y., Michaeli, D., Weinberger-Ohana, P., Mayer, R., Braun, S., de Vroom, E., van der Marel, G.A., van Boom, J.H. *et al.* (1987) Regulation of cellulose synthesis in *Acetobacter xylinum* by cyclic diguanylic acid. *Nature*, **325**, 279-281.
2. Hengge, R. (2009) Principles of c-di-GMP signalling in bacteria. *Nature reviews. Microbiology*, **7**, 263-273.
3. Amikam, D. and Benziman, M. (1989) Cyclic diguanylic acid and cellulose synthesis in *Agrobacterium tumefaciens*. *J Bacteriol*, **171**, 6649-6655.
4. Morgan, J.L., McNamara, J.T. and Zimmer, J. (2014) Mechanism of activation of bacterial cellulose synthase by cyclic di-GMP. *Nat Struct Mol Biol*, **21**, 489-496.
5. Tal, R., Wong, H.C., Calhoon, R., Gelfand, D., Fear, A.L., Volman, G., Mayer, R., Ross, P., Amikam, D., Weinhouse, H. *et al.* (1998) Three *cdg* operons control cellular turnover of cyclic di-GMP in *Acetobacter xylinum*: genetic organization and occurrence of conserved domains in isoenzymes. *J Bacteriol*, **180**, 4416-4425.
6. Schmidt, A.J., Ryjenkov, D.A. and Gomelsky, M. (2005) The ubiquitous protein domain EAL is a cyclic diguanylate-specific phosphodiesterase: enzymatically active and inactive EAL domains. *J Bacteriol*, **187**, 4774-4781.
7. Schirmer, T. (2016) C-di-GMP Synthesis: Structural Aspects of Evolution, Catalysis and Regulation. *Journal of molecular biology*, **428**, 3683-3701.
8. Weinhouse, H., Sapir, S., Amikam, D., Shilo, Y., Volman, G., Ohana, P. and Benziman, M. (1997) c-di-GMP-binding protein, a new factor regulating cellulose synthesis in *Acetobacter xylinum*. *FEBS Lett*, **416**, 207-211.
9. Simm, R., Morr, M., Remminghorst, U., Andersson, M. and Romling, U. (2009) Quantitative determination of cyclic diguanosine monophosphate concentrations in nucleotide extracts of bacteria by matrix-assisted laser desorption/ionization-time-of-flight mass spectrometry. *Anal Biochem*, **386**, 53-58.
10. Amikam, D., Steinberger, O., Shkolnik, T. and Ben-Ishai, Z. (1995) The novel cyclic dinucleotide 3'-5' cyclic diguanylic acid binds to p21ras and enhances DNA synthesis but not cell replication in the Molt 4 cell line. *Biochem J*, **311 (Pt 3)**, 921-927.
11. Ohana, P., Delmer, D.P., Carlson, R.W., Glushka, J., Azadi, P., Bacic, T. and Benziman, M. (1998) Identification of a novel triterpenoid saponin from *Pisum sativum* as a specific inhibitor of the diguanylate cyclase of *Acetobacter xylinum*. *Plant Cell Physiol*, **39**, 144-152.

12. Lieberman, O.J., Orr, M.W., Wang, Y. and Lee, V.T. (2014) High-throughput screening using the differential radial capillary action of ligand assay identifies ebselen as an inhibitor of diguanylate cyclases. *ACS Chem Biol*, **9**, 183-192.
13. Karaolis, D.K., Means, T.K., Yang, D., Takahashi, M., Yoshimura, T., Muraille, E., Philpott, D., Schroeder, J.T., Hyodo, M., Hayakawa, Y. *et al.* (2007) Bacterial c-di-GMP is an immunostimulatory molecule. *J Immunol*, **178**, 2171-2181.
14. Ausmees, N., Mayer, R., Weinhouse, H., Volman, G., Amikam, D., Benziman, M. and Lindberg, M. (2001) Genetic data indicate that proteins containing the GGDEF domain possess diguanylate cyclase activity. *FEMS Microbiol Lett*, **204**, 163-167.
15. Barends, T.R., Hartmann, E., Griese, J.J., Beitlich, T., Kirienko, N.V., Ryjenkov, D.A., Reinstein, J., Shoeman, R.L., Gomelsky, M. and Schlichting, I. (2009) Structure and mechanism of a bacterial light-regulated cyclic nucleotide phosphodiesterase. *Nature*, **459**, 1015-1018.
16. Chan, C., Paul, R., Samoray, D., Amiot, N.C., Giese, B., Jenal, U. and Schirmer, T. (2004) Structural basis of activity and allosteric control of diguanylate cyclase. *Proc Natl Acad Sci U S A*, **101**, 17084-17089.
17. Malone, J.G., Williams, R., Christen, M., Jenal, U., Spiers, A.J. and Rainey, P.B. (2007) The structure-function relationship of WspR, a *Pseudomonas fluorescens* response regulator with a GGDEF output domain. *Microbiology*, **153**, 980-994.
18. Christen, B., Christen, M., Paul, R., Schmid, F., Folcher, M., Jenoe, P., Meuwly, M. and Jenal, U. (2006) Allosteric control of cyclic di-GMP signaling. *The Journal of biological chemistry*, **281**, 32015-32024.
19. Tarutina, M., Ryjenkov, D.A. and Gomelsky, M. (2006) An unorthodox bacteriophytochrome from *Rhodobacter sphaeroides* involved in turnover of the second messenger c-di-GMP. *The Journal of biological chemistry*, **281**, 34751-34758.
20. Ferreira, R.B., Antunes, L.C., Greenberg, E.P. and McCarter, L.L. (2008) *Vibrio parahaemolyticus* ScrC modulates cyclic dimeric GMP regulation of gene expression relevant to growth on surfaces. *J Bacteriol*, **190**, 851-860.
21. Galperin, M.Y. (2004) Bacterial signal transduction network in a genomic perspective. *Environ Microbiol*, **6**, 552-567.
22. Galperin, M.Y., Nikolskaya, A.N. and Koonin, E.V. (2001) Novel domains of the prokaryotic two-component signal transduction systems. *FEMS Microbiol Lett*, **203**, 11-21.
23. Chang, A.L., Tuckerman, J.R., Gonzalez, G., Mayer, R., Weinhouse, H., Volman, G., Amikam, D., Benziman, M. and Gilles-Gonzalez, M.A. (2001) Phosphodiesterase A1, a regulator of cellulose synthesis in *Acetobacter xylinum*, is a heme-based sensor. *Biochemistry*, **40**, 3420-3426.

24. Sasakura, Y., Hirata, S., Sugiyama, S., Suzuki, S., Taguchi, S., Watanabe, M., Matsui, T., Sagami, I. and Shimizu, T. (2002) Characterization of a direct oxygen sensor heme protein from *Escherichia coli*. Effects of the heme redox states and mutations at the heme-binding site on catalysis and structure. *The Journal of biological chemistry*, **277**, 23821-23827.
25. Jenal, U. and Malone, J. (2006) Mechanisms of cyclic-di-GMP signaling in bacteria. *Annu Rev Genet*, **40**, 385-407.
26. Pursley, B.R., Maiden, M.M., Hsieh, M.L., Fernandez, N.L., Severin, G.B. and Waters, C.M. (2018) Cyclic di-GMP Regulates TfoY in *Vibrio cholerae* To Control Motility by both Transcriptional and Posttranscriptional Mechanisms. *Journal of bacteriology*, **200**, 578-617.
27. Fernandez, N.L., Srivastava, D., Ngouajio, A.L. and Waters, C.M. (2018) Cyclic di-GMP Positively Regulates DNA Repair in *Vibrio cholerae*. *Journal of bacteriology*, **200**.
28. Beyhan, S., Odell, L.S. and Yildiz, F.H. (2008) Identification and characterization of cyclic diguanylate signaling systems controlling rugosity in *Vibrio cholerae*. *J Bacteriol*, **190**, 7392-7405.
29. Duerig, A., Abel, S., Folcher, M., Nicollier, M., Schwede, T., Amiot, N., Giese, B. and Jenal, U. (2009) Second messenger-mediated spatiotemporal control of protein degradation regulates bacterial cell cycle progression. *Genes Dev*, **23**, 93-104.
30. Hsieh, M.L., Hinton, D.M. and Waters, C.M. (2018) VpsR and cyclic di-GMP together drive transcription initiation to activate biofilm formation in *Vibrio cholerae*. *Nucleic acids research*, **46**, 8876-8887.
31. Srivastava, D., Hsieh, M.L., Khataokar, A., Neiditch, M.B. and Waters, C.M. (2013) Cyclic di-GMP inhibits *Vibrio cholerae* motility by repressing induction of transcription and inducing extracellular polysaccharide production. *Molecular microbiology*, **90**, 1262-1276.
32. Sudarsan, N., Lee, E.R., Weinberg, Z., Moy, R.H., Kim, J.N., Link, K.H. and Breaker, R.R. (2008) Riboswitches in eubacteria sense the second messenger cyclic di-GMP. *Science*, **321**, 411-413.
33. Dick, M.H., Guillermin, M., Moussy, F. and Chaignat, C.L. (2012) Review of two decades of cholera diagnostics--how far have we really come? *PLoS Negl Trop Dis*, **6**, e1845.
34. Ali, M., Lopez, A.L., You, Y.A., Kim, Y.E., Sah, B., Maskery, B. and Clemens, J. (2012) The global burden of cholera. *Bull World Health Organ*, **90**, 209-218A.
35. Finkelstein, R.A. (1996) In th and Baron, S. (eds.), *Medical Microbiology*, Galveston (TX).

36. Conner, J.G., Teschler, J.K., Jones, C.J. and Yildiz, F.H. (2016) Staying Alive: *Vibrio cholerae*'s Cycle of Environmental Survival, Transmission, and Dissemination. *Microbiol Spectr*, **4**.
37. Meibom, K.L., Blokesch, M., Dolganov, N.A., Wu, C.Y. and Schoolnik, G.K. (2005) Chitin induces natural competence in *Vibrio cholerae*. *Science*, **310**, 1824-1827.
38. Kamruzzaman, M., Udden, S.M., Cameron, D.E., Calderwood, S.B., Nair, G.B., Mekalanos, J.J. and Faruque, S.M. (2010) Quorum-regulated biofilms enhance the development of conditionally viable, environmental *Vibrio cholerae*. *Proceedings of the National Academy of Sciences of the United States of America*, **107**, 1588-1593.
39. Faruque, S.M., Islam, M.J., Ahmad, Q.S., Faruque, A.S., Sack, D.A., Nair, G.B. and Mekalanos, J.J. (2005) Self-limiting nature of seasonal cholera epidemics: Role of host-mediated amplification of phage. *Proceedings of the National Academy of Sciences of the United States of America*, **102**, 6119-6124.
40. Huq, A., Sack, R.B., Nizam, A., Longini, I.M., Nair, G.B., Ali, A., Morris, J.G., Jr., Khan, M.N., Siddique, A.K., Yunus, M. *et al.* (2005) Critical factors influencing the occurrence of *Vibrio cholerae* in the environment of Bangladesh. *Applied and environmental microbiology*, **71**, 4645-4654.
41. Hornick, R.B., Music, S.I., Wenzel, R., Cash, R., Libonati, J.P., Snyder, M.J. and Woodward, T.E. (1971) The Broad Street pump revisited: response of volunteers to ingested cholera vibrios. *Bull N Y Acad Med*, **47**, 1181-1191.
42. De, S.N. (1959) Enterotoxicity of bacteria-free culture-filtrate of *Vibrio cholerae*. *Nature*, **183**, 1533-1534.
43. Field, M., Fromm, D., al-Awqati, Q. and Greenough, W.B., 3rd. (1972) Effect of cholera enterotoxin on ion transport across isolated ileal mucosa. *J Clin Invest*, **51**, 796-804.
44. Gill, D.M. and Meren, R. (1978) ADP-ribosylation of membrane proteins catalyzed by cholera toxin: basis of the activation of adenylate cyclase. *Proc Natl Acad Sci U S A*, **75**, 3050-3054.
45. Lonngren, I. and Holmgren, J. (1973) Subunit structure of cholera toxin. *J Gen Microbiol*, **76**, 417-427.
46. Tsai, B. and Rapoport, T.A. (2002) Unfolded cholera toxin is transferred to the ER membrane and released from protein disulfide isomerase upon oxidation by Ero1. *J Cell Biol*, **159**, 207-216.
47. Wernick, N.L., Chinnapen, D.J., Cho, J.A. and Lencer, W.I. (2010) Cholera toxin: an intracellular journey into the cytosol by way of the endoplasmic reticulum. *Toxins (Basel)*, **2**, 310-325.

48. Gabriel, S.E., Brigman, K.N., Koller, B.H., Boucher, R.C. and Stutts, M.J. (1994) Cystic fibrosis heterozygote resistance to cholera toxin in the cystic fibrosis mouse model. *Science*, **266**, 107-109.
49. Weil, A.A., Begum, Y., Chowdhury, F., Khan, A.I., Leung, D.T., LaRocque, R.C., Charles, R.C., Ryan, E.T., Calderwood, S.B., Qadri, F. *et al.* (2014) Bacterial shedding in household contacts of cholera patients in Dhaka, Bangladesh. *Am J Trop Med Hyg*, **91**, 738-742.
50. Kamp, H.D., Patimalla-Dipali, B., Lazinski, D.W., Wallace-Gadsden, F. and Camilli, A. (2013) Gene fitness landscapes of *Vibrio cholerae* at important stages of its life cycle. *PLoS Pathog*, **9**, e1003800.
51. Wai, S.N., Mizunoe, Y., Takade, A., Kawabata, S.I. and Yoshida, S.I. (1998) *Vibrio cholerae* O1 strain TSI-4 produces the exopolysaccharide materials that determine colony morphology, stress resistance, and biofilm formation. *Appl Environ Microbiol*, **64**, 3648-3655.
52. Matz, C., McDougald, D., Moreno, A.M., Yung, P.Y., Yildiz, F.H. and Kjelleberg, S. (2005) Biofilm formation and phenotypic variation enhance predation-driven persistence of *Vibrio cholerae*. *Proc Natl Acad Sci U S A*, **102**, 16819-16824.
53. Beyhan, S., Bilecen, K., Salama, S.R., Casper-Lindley, C. and Yildiz, F.H. (2007) Regulation of rugosity and biofilm formation in *Vibrio cholerae*: comparison of VpsT and VpsR regulons and epistasis analysis of vpsT, vpsR, and hapR. *Journal of bacteriology*, **189**, 388-402.
54. Tamplin, M.L., Gauzens, A.L., Huq, A., Sack, D.A. and Colwell, R.R. (1990) Attachment of *Vibrio cholerae* serogroup O1 to zooplankton and phytoplankton of Bangladesh waters. *Appl Environ Microbiol*, **56**, 1977-1980.
55. Huq, A., Xu, B., Chowdhury, M.A., Islam, M.S., Montilla, R. and Colwell, R.R. (1996) A simple filtration method to remove plankton-associated *Vibrio cholerae* in raw water supplies in developing countries. *Appl Environ Microbiol*, **62**, 2508-2512.
56. Colwell, R.R., Huq, A., Islam, M.S., Aziz, K.M., Yunus, M., Khan, N.H., Mahmud, A., Sack, R.B., Nair, G.B., Chakraborty, J. *et al.* (2003) Reduction of cholera in Bangladeshi villages by simple filtration. *Proceedings of the National Academy of Sciences of the United States of America*, **100**, 1051-1055.
57. Faruque, S.M., Biswas, K., Udden, S.M., Ahmad, Q.S., Sack, D.A., Nair, G.B. and Mekalanos, J.J. (2006) Transmissibility of cholera: in vivo-formed biofilms and their relationship to infectivity and persistence in the environment. *Proceedings of the National Academy of Sciences of the United States of America*, **103**, 6350-6355.
58. Nielsen, A.T., Dolganov, N.A., Rasmussen, T., Otto, G., Miller, M.C., Felt, S.A., Torreilles, S. and Schoolnik, G.K. (2010) A bistable switch and anatomical site control *Vibrio cholerae* virulence gene expression in the intestine. *PLoS Pathog*, **6**, e1001102.

59. Fong, J.C., Syed, K.A., Klose, K.E. and Yildiz, F.H. (2010) Role of *Vibrio* polysaccharide (*vps*) genes in VPS production, biofilm formation and *Vibrio cholerae* pathogenesis. *Microbiology*, **156**, 2757-2769.
60. Yildiz, F.H. and Schoolnik, G.K. (1999) *Vibrio cholerae* O1 El Tor: identification of a gene cluster required for the rugose colony type, exopolysaccharide production, chlorine resistance, and biofilm formation. *Proceedings of the National Academy of Sciences of the United States of America*, **96**, 4028-4033.
61. Fong, J.C., Karplus, K., Schoolnik, G.K. and Yildiz, F.H. (2006) Identification and characterization of RbmA, a novel protein required for the development of rugose colony morphology and biofilm structure in *Vibrio cholerae*. *Journal of bacteriology*, **188**, 1049-1059.
62. Fong, J.C. and Yildiz, F.H. (2007) The *rbmBCDEF* gene cluster modulates development of rugose colony morphology and biofilm formation in *Vibrio cholerae*. *Journal of bacteriology*, **189**, 2319-2330.
63. Absalon, C., Van Dellen, K. and Watnick, P.I. (2011) A communal bacterial adhesin anchors biofilm and bystander cells to surfaces. *PLoS Pathog*, **7**, e1002210.
64. Moorthy, S. and Watnick, P.I. (2005) Identification of novel stage-specific genetic requirements through whole genome transcription profiling of *Vibrio cholerae* biofilm development. *Molecular microbiology*, **57**, 1623-1635.
65. Johnson, T.L., Fong, J.C., Rule, C., Rogers, A., Yildiz, F.H. and Sandkvist, M. (2014) The Type II secretion system delivers matrix proteins for biofilm formation by *Vibrio cholerae*. *J Bacteriol*, **196**, 4245-4252.
66. Berk, V., Fong, J.C., Dempsey, G.T., Develioglu, O.N., Zhuang, X., Liphardt, J., Yildiz, F.H. and Chu, S. (2012) Molecular architecture and assembly principles of *Vibrio cholerae* biofilms. *Science*, **337**, 236-239.
67. Casper-Lindley, C. and Yildiz, F.H. (2004) *VpsT* is a transcriptional regulator required for expression of *vps* biosynthesis genes and the development of rugose colonial morphology in *Vibrio cholerae* O1 El Tor. *Journal of bacteriology*, **186**, 1574-1578.
68. Krasteva, P.V., Fong, J.C., Shikuma, N.J., Beyhan, S., Navarro, M.V., Yildiz, F.H. and Sondermann, H. (2010) *Vibrio cholerae* *VpsT* regulates matrix production and motility by directly sensing cyclic di-GMP. *Science*, **327**, 866-868.
69. Srivastava, D., Harris, R.C. and Waters, C.M. (2011) Integration of cyclic di-GMP and quorum sensing in the control of *vpsT* and *aphA* in *Vibrio cholerae*. *Journal of bacteriology*, **193**, 6331-6341.
70. Waters, C.M., Lu, W., Rabinowitz, J.D. and Bassler, B.L. (2008) Quorum sensing controls biofilm formation in *Vibrio cholerae* through modulation of cyclic di-GMP levels and repression of *vpsT*. *Journal of bacteriology*, **190**, 2527-2536.

71. Zamorano-Sanchez, D., Fong, J.C., Kilic, S., Erill, I. and Yildiz, F.H. (2015) Identification and characterization of VpsR and VpsT binding sites in *Vibrio cholerae*. *Journal of bacteriology*, **197**, 1221-1235.
72. Yildiz, F.H., Dolganov, N.A. and Schoolnik, G.K. (2001) VpsR, a Member of the Response Regulators of the Two-Component Regulatory Systems, Is Required for Expression of vps Biosynthesis Genes and EPS(ETr)-Associated Phenotypes in *Vibrio cholerae* O1 El Tor. *Journal of bacteriology*, **183**, 1716-1726.
73. Yildiz, F.H., Liu, X.S., Heydorn, A. and Schoolnik, G.K. (2004) Molecular analysis of rugosity in a *Vibrio cholerae* O1 El Tor phase variant. *Molecular microbiology*, **53**, 497-515.
74. Hammer, B.K. and Bassler, B.L. (2007) Regulatory small RNAs circumvent the conventional quorum sensing pathway in pandemic *Vibrio cholerae*. *Proceedings of the National Academy of Sciences of the United States of America*, **104**, 11145-11149.
75. Ayala, J.C., Wang, H., Benitez, J.A. and Silva, A.J. (2015) RNA-Seq analysis and whole genome DNA-binding profile of the *Vibrio cholerae* histone-like nucleoid structuring protein (H-NS). *Genomics data*, **5**, 147-150.
76. Ayala, J.C., Wang, H., Silva, A.J. and Benitez, J.A. (2015) Repression by H-NS of genes required for the biosynthesis of the *Vibrio cholerae* biofilm matrix is modulated by the second messenger cyclic diguanylic acid. *Molecular microbiology*, **97**, 630-645.
77. Wang, H., Ayala, J.C., Silva, A.J. and Benitez, J.A. (2012) The histone-like nucleoid structuring protein (H-NS) is a repressor of *Vibrio cholerae* exopolysaccharide biosynthesis (vps) genes. *Applied and environmental microbiology*, **78**, 2482-2488.
78. Das, B., Pal, R.R., Bag, S. and Bhadra, R.K. (2009) Stringent response in *Vibrio cholerae*: genetic analysis of spoT gene function and identification of a novel (p)ppGpp synthetase gene. *Mol Microbiol*, **72**, 380-398.
79. Raskin, D.M., Judson, N. and Mekalanos, J.J. (2007) Regulation of the stringent response is the essential function of the conserved bacterial G protein CgtA in *Vibrio cholerae*. *Proc Natl Acad Sci U S A*, **104**, 4636-4641.
80. He, H., Cooper, J.N., Mishra, A. and Raskin, D.M. (2012) Stringent response regulation of biofilm formation in *Vibrio cholerae*. *J Bacteriol*, **194**, 2962-2972.
81. Liang, W., Pascual-Montano, A., Silva, A.J. and Benitez, J.A. (2007) The cyclic AMP receptor protein modulates quorum sensing, motility and multiple genes that affect intestinal colonization in *Vibrio cholerae*. *Microbiology*, **153**, 2964-2975.
82. Liang, W., Silva, A.J. and Benitez, J.A. (2007) The cyclic AMP receptor protein modulates colonial morphology in *Vibrio cholerae*. *Appl Environ Microbiol*, **73**, 7482-7487.

83. Fong, J.C. and Yildiz, F.H. (2008) Interplay between cyclic AMP-cyclic AMP receptor protein and cyclic di-GMP signaling in *Vibrio cholerae* biofilm formation. *J Bacteriol*, **190**, 6646-6659.
84. Decker, K.B. and Hinton, D.M. (2013) Transcription regulation at the core: similarities among bacterial, archaeal, and eukaryotic RNA polymerases. *Annual review of microbiology*, **67**, 113-139.
85. Lee, D.J., Minchin, S.D. and Busby, S.J. (2012) Activating transcription in bacteria. *Annual review of microbiology*, **66**, 125-152.
86. Paget, M.S. and Helmann, J.D. (2003) The sigma70 family of sigma factors. *Genome Biol*, **4**, 203.
87. Feklistov, A., Sharon, B.D., Darst, S.A. and Gross, C.A. (2014) Bacterial sigma factors: a historical, structural, and genomic perspective. *Annual review of microbiology*, **68**, 357-376.
88. Gruber, T.M. and Gross, C.A. (2003) Multiple sigma subunits and the partitioning of bacterial transcription space. *Annual review of microbiology*, **57**, 441-466.
89. Zuo, Y. and Steitz, T.A. (2015) Crystal structures of the *E. coli* transcription initiation complexes with a complete bubble. *Mol Cell*, **58**, 534-540.
90. Bae, B., Feklistov, A., Lass-Napiorkowska, A., Landick, R. and Darst, S.A. (2015) Structure of a bacterial RNA polymerase holoenzyme open promoter complex. *Elife*, **4**.
91. Feklistov, A. and Darst, S.A. (2011) Structural basis for promoter-10 element recognition by the bacterial RNA polymerase sigma subunit. *Cell*, **147**, 1257-1269.
92. Zhang, Y., Feng, Y., Chatterjee, S., Tuske, S., Ho, M.X., Arnold, E. and Ebright, R.H. (2012) Structural basis of transcription initiation. *Science*, **338**, 1076-1080.
93. Ross, W., Ernst, A. and Gourse, R.L. (2001) Fine structure of *E. coli* RNA polymerase-promoter interactions: alpha subunit binding to the UP element minor groove. *Genes Dev*, **15**, 491-506.
94. Murakami, K.S., Masuda, S. and Darst, S.A. (2003) Crystallographic analysis of *Thermus aquaticus* RNA polymerase holoenzyme and a holoenzyme/promoter DNA complex. *Methods Enzymol*, **370**, 42-53.
95. Hook-Barnard, I.G. and Hinton, D.M. (2007) Transcription Initiation by Mix and Match Elements: Flexibility for Polymerase Binding to Bacterial Promoters. *Gene regulation and systems biology*, **1**, 275-293.
96. Saecker, R.M., Record, M.T., Jr. and Dehaseth, P.L. (2011) Mechanism of bacterial transcription initiation: RNA polymerase - promoter binding, isomerization to initiation-

- competent open complexes, and initiation of RNA synthesis. *Journal of molecular biology*, **412**, 754-771.
97. Chakraborty, A., Wang, D., Ebright, Y.W., Korlann, Y., Kortkhonjia, E., Kim, T., Chowdhury, S., Wigneshweraraj, S., Irschik, H., Jansen, R. *et al.* (2012) Opening and closing of the bacterial RNA polymerase clamp. *Science*, **337**, 591-595.
 98. Mekler, V., Kortkhonjia, E., Mukhopadhyay, J., Knight, J., Revyakin, A., Kapanidis, A.N., Niu, W., Ebright, Y.W., Levy, R. and Ebright, R.H. (2002) Structural organization of bacterial RNA polymerase holoenzyme and the RNA polymerase-promoter open complex. *Cell*, **108**, 599-614.
 99. Hsu, L.M. (2002) Promoter clearance and escape in prokaryotes. *Biochimica et biophysica acta*, **1577**, 191-207.
 100. Hinton, D.M., Pande, S., Wais, N., Johnson, X.B., Vuthoori, M., Makela, A. and Hook-Barnard, I. (2005) Transcriptional takeover by sigma appropriation: remodelling of the sigma70 subunit of Escherichia coli RNA polymerase by the bacteriophage T4 activator MotA and co-activator AsiA. *Microbiology*, **151**, 1729-1740.
 101. Popovych, N., Tzeng, S.R., Tonelli, M., Ebright, R.H. and Kalodimos, C.G. (2009) Structural basis for cAMP-mediated allosteric control of the catabolite activator protein. *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 6927-6932.
 102. Bush, M. and Dixon, R. (2012) The role of bacterial enhancer binding proteins as specialized activators of sigma54-dependent transcription. *Microbiology and molecular biology reviews : MMBR*, **76**, 497-529.
 103. Zhang, N. and Buck, M. (2015) A perspective on the enhancer dependent bacterial RNA polymerase. *Biomolecules*, **5**, 1012-1019.
 104. Zhang, N., Darbari, V.C., Glyde, R., Zhang, X. and Buck, M. (2016) The bacterial enhancer-dependent RNA polymerase. *Biochem J*, **473**, 3741-3753.
 105. Wang, H., Ayala, J.C., Benitez, J.A. and Silva, A.J. (2014) The LuxR-type regulator VpsT negatively controls the transcription of rpoS, encoding the general stress response regulator, in *Vibrio cholerae* biofilms. *Journal of bacteriology*, **196**, 1020-1030.
 106. Matsuyama, B.Y., Krasteva, P.V., Baraquet, C., Harwood, C.S., Sondermann, H. and Navarro, M.V. (2016) Mechanistic insights into c-di-GMP-dependent control of the biofilm regulator FleQ from *Pseudomonas aeruginosa*. *Proceedings of the National Academy of Sciences of the United States of America*, **113**, E209-218.
 107. Baraquet, C. and Harwood, C.S. (2015) FleQ DNA Binding Consensus Sequence Revealed by Studies of FleQ-Dependent Regulation of Biofilm Gene Expression in *Pseudomonas aeruginosa*. *Journal of bacteriology*, **198**, 178-186.

108. Baraquet, C. and Harwood, C.S. (2013) Cyclic diguanosine monophosphate represses bacterial flagella synthesis by interacting with the Walker A motif of the enhancer-binding protein FleQ. *Proceedings of the National Academy of Sciences of the United States of America*, **110**, 18478-18483.
109. Baraquet, C., Murakami, K., Parsek, M.R. and Harwood, C.S. (2012) The FleQ protein from *Pseudomonas aeruginosa* functions as both a repressor and an activator to control gene expression from the *pel* operon promoter in response to c-di-GMP. *Nucleic acids research*, **40**, 7207-7218.
110. Jyot, J., Dasgupta, N. and Ramphal, R. (2002) FleQ, the major flagellar gene regulator in *Pseudomonas aeruginosa*, binds to enhancer sites located either upstream or atypically downstream of the RpoN binding site. *Journal of bacteriology*, **184**, 5251-5260.
111. Arora, S.K., Ritchings, B.W., Almira, E.C., Lory, S. and Ramphal, R. (1997) A transcriptional activator, FleQ, regulates mucin adhesion and flagellar gene expression in *Pseudomonas aeruginosa* in a cascade manner. *Journal of bacteriology*, **179**, 5574-5581.
112. Liao, J., Schurr, M.J. and Sauer, K. (2013) The MerR-like regulator BrIR confers biofilm tolerance by activating multidrug efflux pumps in *Pseudomonas aeruginosa* biofilms. *Journal of bacteriology*, **195**, 3352-3363.
113. Chambers, J.R., Liao, J., Schurr, M.J. and Sauer, K. (2014) BrIR from *Pseudomonas aeruginosa* is a c-di-GMP-responsive transcription factor. *Molecular microbiology*, **92**, 471-487.
114. Wang, F., He, Q., Yin, J., Xu, S., Hu, W. and Gu, L. (2018) BrIR from *Pseudomonas aeruginosa* is a receptor for both cyclic di-GMP and pyocyanin. *Nat Commun*, **9**, 2563.
115. Raju, H. and Sharma, R. (2017) Crystal structure of BrIR with c-di-GMP. *Biochemical and biophysical research communications*, **490**, 260-264.
116. Wilksch, J.J., Yang, J., Clements, A., Gabbe, J.L., Short, K.R., Cao, H., Cavaliere, R., James, C.E., Whitchurch, C.B., Schembri, M.A. *et al.* (2011) MrkH, a novel c-di-GMP-dependent transcriptional activator, controls *Klebsiella pneumoniae* biofilm formation by regulating type 3 fimbriae expression. *PLoS Pathog*, **7**, e1002204.
117. Tan, J.W., Wilksch, J.J., Hocking, D.M., Wang, N., Srikhanta, Y.N., Tauschek, M., Lithgow, T., Robins-Browne, R.M., Yang, J. and Strugnell, R.A. (2015) Positive autoregulation of mrkHI by the cyclic di-GMP-dependent MrkH protein in the biofilm regulatory circuit of *Klebsiella pneumoniae*. *Journal of bacteriology*, **197**, 1659-1667.
118. Yang, J., Wilksch, J.J., Tan, J.W., Hocking, D.M., Webb, C.T., Lithgow, T., Robins-Browne, R.M. and Strugnell, R.A. (2013) Transcriptional activation of the *mrkA* promoter of the *Klebsiella pneumoniae* type 3 fimbrial operon by the c-di-GMP-dependent MrkH protein. *PLoS One*, **8**, e79038.

119. Schumacher, M.A. and Zeng, W. (2016) Structures of the activator of *K. pneumoniae* biofilm formation, MrkH, indicates PilZ domains involved in c-di-GMP and DNA binding. *Proceedings of the National Academy of Sciences of the United States of America*, **113**, 10067-10072.
120. Fazli, M., McCarthy, Y., Givskov, M., Ryan, R.P. and Tolker-Nielsen, T. (2013) The exopolysaccharide gene cluster Bcam1330-Bcam1341 is involved in *Burkholderia cenocepacia* biofilm formation, and its expression is regulated by c-di-GMP and Bcam1349. *Microbiologyopen*, **2**, 105-122.
121. Fazli, M., O'Connell, A., Nilsson, M., Niehaus, K., Dow, J.M., Givskov, M., Ryan, R.P. and Tolker-Nielsen, T. (2011) The CRP/FNR family protein Bcam1349 is a c-di-GMP effector that regulates biofilm formation in the respiratory pathogen *Burkholderia cenocepacia*. *Molecular microbiology*, **82**, 327-341.
122. Fazli, M., Rybtke, M., Steiner, E., Weidel, E., Berthelsen, J., Groizeleau, J., Bin, W., Zhi, B.Z., Yaming, Z., Kaefer, V. *et al.* (2017) Regulation of *Burkholderia cenocepacia* biofilm formation by RpoN and the c-di-GMP effector BerB. *Microbiologyopen*, **6**.
123. Chin, K.H., Lee, Y.C., Tu, Z.L., Chen, C.H., Tseng, Y.H., Yang, J.M., Ryan, R.P., McCarthy, Y., Dow, J.M., Wang, A.H. *et al.* (2010) The cAMP receptor-like protein CLP is a novel c-di-GMP receptor linking cell-cell signaling to virulence gene expression in *Xanthomonas campestris*. *Journal of molecular biology*, **396**, 646-662.
124. den Hengst, C.D., Tran, N.T., Bibb, M.J., Chandra, G., Leskiw, B.K. and Buttner, M.J. (2010) Genes essential for morphological development and antibiotic production in *Streptomyces coelicolor* are targets of BldD during vegetative growth. *Molecular microbiology*, **78**, 361-379.
125. Elliot, M.A., Bibb, M.J., Buttner, M.J. and Leskiw, B.K. (2001) BldD is a direct regulator of key developmental genes in *Streptomyces coelicolor* A3(2). *Molecular microbiology*, **40**, 257-269.
126. Tschowri, N., Schumacher, M.A., Schlimpert, S., Chinnam, N.B., Findlay, K.C., Brennan, R.G. and Buttner, M.J. (2014) Tetrameric c-di-GMP mediates effective transcription factor dimerization to control *Streptomyces* development. *Cell*, **158**, 1136-1147.
127. Schumacher, M.A., Zeng, W., Findlay, K.C., Buttner, M.J., Brennan, R.G. and Tschowri, N. (2017) The *Streptomyces* master regulator BldD binds c-di-GMP sequentially to create a functional BldD2-(c-di-GMP)₄ complex. *Nucleic acids research*, **45**, 6923-6933.
128. Li, W. and He, Z.G. (2012) LtmA, a novel cyclic di-GMP-responsive activator, broadly regulates the expression of lipid transport and metabolism genes in *Mycobacterium smegmatis*. *Nucleic acids research*, **40**, 11292-11307.
129. Srivastava, D. and Waters, C.M. (2012) A tangled web: regulatory connections between quorum sensing and cyclic Di-GMP. *Journal of bacteriology*, **194**, 4485-4493.

130. Ng, W.L. and Bassler, B.L. (2009) Bacterial quorum-sensing network architectures. *Annu Rev Genet*, **43**, 197-222.
131. Higgins, D.A., Pomianek, M.E., Kraml, C.M., Taylor, R.K., Semmelhack, M.F. and Bassler, B.L. (2007) The major *Vibrio cholerae* autoinducer and its role in virulence factor production. *Nature*, **450**, 883-886.
132. Ng, W.L., Perez, L.J., Wei, Y., Kraml, C., Semmelhack, M.F. and Bassler, B.L. (2011) Signal production and detection specificity in *Vibrio* CqsA/CqsS quorum-sensing systems. *Molecular microbiology*, **79**, 1407-1417.
133. Wei, Y., Perez, L.J., Ng, W.L., Semmelhack, M.F. and Bassler, B.L. (2011) Mechanism of *Vibrio cholerae* autoinducer-1 biosynthesis. *ACS Chem Biol*, **6**, 356-365.
134. Surette, M.G., Miller, M.B. and Bassler, B.L. (1999) Quorum sensing in *Escherichia coli*, *Salmonella typhimurium*, and *Vibrio harveyi*: a new family of genes responsible for autoinducer production. *Proceedings of the National Academy of Sciences of the United States of America*, **96**, 1639-1644.
135. Schauder, S., Shokat, K., Surette, M.G. and Bassler, B.L. (2001) The LuxS family of bacterial autoinducers: biosynthesis of a novel quorum-sensing signal molecule. *Molecular microbiology*, **41**, 463-476.
136. Chen, X., Schauder, S., Potier, N., Van Dorsselaer, A., Pelczer, I., Bassler, B.L. and Hughson, F.M. (2002) Structural identification of a bacterial quorum-sensing signal containing boron. *Nature*, **415**, 545-549.
137. Freeman, J.A. and Bassler, B.L. (1999) A genetic analysis of the function of LuxO, a two-component response regulator involved in quorum sensing in *Vibrio harveyi*. *Molecular microbiology*, **31**, 665-677.
138. Lilley, B.N. and Bassler, B.L. (2000) Regulation of quorum sensing in *Vibrio harveyi* by LuxO and sigma-54. *Molecular microbiology*, **36**, 940-954.
139. Pappenfort, K. and Vogel, J. (2010) Regulatory RNA in bacterial pathogens. *Cell Host Microbe*, **8**, 116-127.
140. Vogel, J. and Luisi, B.F. (2011) Hfq and its constellation of RNA. *Nature reviews. Microbiology*, **9**, 578-589.
141. Lenz, D.H., Mok, K.C., Lilley, B.N., Kulkarni, R.V., Wingreen, N.S. and Bassler, B.L. (2004) The small RNA chaperone Hfq and multiple small RNAs control quorum sensing in *Vibrio harveyi* and *Vibrio cholerae*. *Cell*, **118**, 69-82.
142. Rutherford, S.T., van Kessel, J.C., Shao, Y. and Bassler, B.L. (2011) AphA and LuxR/HapR reciprocally control quorum sensing in vibrios. *Genes Dev*, **25**, 397-408.

143. Shao, Y. and Bassler, B.L. (2012) Quorum-sensing non-coding small RNAs use unique pairing regions to differentially control mRNA targets. *Molecular microbiology*, **83**, 599-611.
144. Tu, K.C. and Bassler, B.L. (2007) Multiple small RNAs act additively to integrate sensory information and control quorum sensing in *Vibrio harveyi*. *Genes Dev*, **21**, 221-233.
145. Wei, Y., Ng, W.L., Cong, J. and Bassler, B.L. (2012) Ligand and antagonist driven regulation of the *Vibrio cholerae* quorum-sensing receptor CqsS. *Molecular microbiology*, **83**, 1095-1108.
146. Kovacikova, G. and Skorupski, K. (2002) Regulation of virulence gene expression in *Vibrio cholerae* by quorum sensing: HapR functions at the *aphA* promoter. *Molecular microbiology*, **46**, 1135-1147.
147. Lin, W., Kovacikova, G. and Skorupski, K. (2007) The quorum sensing regulator HapR downregulates the expression of the virulence gene transcription factor AphA in *Vibrio cholerae* by antagonizing Lrp- and VpsR-mediated activation. *Molecular microbiology*, **64**, 953-967.
148. Capra, E.J. and Laub, M.T. (2012) Evolution of two-component signal transduction systems. *Annual review of microbiology*, **66**, 325-347.
149. Galperin, M.Y. (2005) A census of membrane-bound and intracellular signal transduction proteins in bacteria: bacterial IQ, extroverts and introverts. *BMC Microbiol*, **5**, 35.
150. Galperin, M.Y. (2010) Diversity of structure and function of response regulator output domains. *Curr Opin Microbiol*, **13**, 150-159.
151. Appleby, J.L., Parkinson, J.S. and Bourret, R.B. (1996) Signal transduction via the multi-step phosphorelay: not necessarily a road less traveled. *Cell*, **86**, 845-848.
152. Jiang, M., Shao, W., Perego, M. and Hoch, J.A. (2000) Multiple histidine kinases regulate entry into stationary phase and sporulation in *Bacillus subtilis*. *Molecular microbiology*, **38**, 535-542.
153. Zschiedrich, C.P., Keidel, V. and Szurmant, H. (2016) Molecular Mechanisms of Two-Component Signal Transduction. *Journal of molecular biology*, **428**, 3752-3775.
154. Wolfe, A.J. (2005) The acetate switch. *Microbiology and molecular biology reviews* : *MMBR*, **69**, 12-50.
155. Bachhawat, P., Swapna, G.V., Montelione, G.T. and Stock, A.M. (2005) Mechanism of activation for transcription factor PhoB suggested by different modes of dimerization in the inactive and active states. *Structure*, **13**, 1353-1363.

156. Boulanger, A., Moon, K., Decker, K.B., Chen, Q., Knipling, L., Stibitz, S. and Hinton, D.M. (2015) Bordetella pertussis fim3 gene regulation by BvgA: phosphorylation controls the formation of inactive vs. active transcription complexes. *Proceedings of the National Academy of Sciences of the United States of America*, **112**, E526-535.
157. Gupta, S., Sinha, A. and Sarkar, D. (2006) Transcriptional autoregulation by Mycobacterium tuberculosis PhoP involves recognition of novel direct repeat sequences in the regulatory region of the promoter. *FEBS Lett*, **580**, 5328-5338.
158. Sinha, A., Gupta, S., Bhutani, S., Pathak, A. and Sarkar, D. (2008) PhoP-PhoP interaction at adjacent PhoP binding sites is influenced by protein phosphorylation. *Journal of bacteriology*, **190**, 1317-1328.
159. Walthers, D., Carroll, R.K., Navarre, W.W., Libby, S.J., Fang, F.C. and Kenney, L.J. (2007) The response regulator SsrB activates expression of diverse Salmonella pathogenicity island 2 promoters and counters silencing by the nucleoid-associated protein H-NS. *Molecular microbiology*, **65**, 477-493.
160. Walthers, D., Li, Y., Liu, Y., Anand, G., Yan, J. and Kenney, L.J. (2011) Salmonella enterica response regulator SsrB relieves H-NS silencing by displacing H-NS bound in polymerization mode and directly activates transcription. *The Journal of biological chemistry*, **286**, 1895-1902.
161. Bryers, J.D. (2008) Medical biofilms. *Biotechnology and bioengineering*, **100**, 1-18.
162. Wolcott, R.D., Rhoads, D.D., Bennett, M.E., Wolcott, B.M., Gogokhia, L., Costerton, J.W. and Dowd, S.E. (2010) Chronic wounds and the medical biofilm paradigm. *Journal of wound care*, **19**, 45-46, 48-50, 52-43.
163. Teschler, J.K., Zamorano-Sanchez, D., Utada, A.S., Warner, C.J., Wong, G.C., Linington, R.G. and Yildiz, F.H. (2015) Living in the matrix: assembly and control of Vibrio cholerae biofilms. *Nature reviews. Microbiology*, **13**, 255-268.
164. Alam, M., Islam, A., Bhuiyan, N.A., Rahim, N., Hossain, A., Khan, G.Y., Ahmed, D., Watanabe, H., Izumiya, H., Faruque, A.S. *et al.* (2011) Clonal transmission, dual peak, and off-season cholera in Bangladesh. *Infection ecology & epidemiology*, **1**, 1-13.
165. Alam, M., Sultana, M., Nair, G.B., Siddique, A.K., Hasan, N.A., Sack, R.B., Sack, D.A., Ahmed, K.U., Sadique, A., Watanabe, H. *et al.* (2007) Viable but nonculturable Vibrio cholerae O1 in biofilms in the aquatic environment and their role in cholera transmission. *Proc Natl Acad Sci U S A*, **104**, 17801-17806.
166. Huq, A., Small, E.B., West, P.A., Huq, M.I., Rahman, R. and Colwell, R.R. (1983) Ecological relationships between Vibrio cholerae and planktonic crustacean copepods. *Applied and environmental microbiology*, **45**, 275-283.
167. Islam, M.S., Jahid, M.I., Rahman, M.M., Rahman, M.Z., Islam, M.S., Kabir, M.S., Sack, D.A. and Schoolnik, G.K. (2007) Biofilm acts as a microenvironment for plankton-

- associated *Vibrio cholerae* in the aquatic environment of Bangladesh. *Microbiology and immunology*, **51**, 369-379.
168. Ryjenkov, D.A., Tarutina, M., Moskvina, O.V. and Gomelsky, M. (2005) Cyclic diguanylate is a ubiquitous signaling molecule in bacteria: insights into biochemistry of the GGDEF protein domain. *Journal of bacteriology*, **187**, 1792-1798.
 169. Romling, U., Galperin, M.Y. and Gomelsky, M. (2013) Cyclic di-GMP: the first 25 years of a universal bacterial second messenger. *Microbiology and molecular biology reviews : MMBR*, **77**, 1-52.
 170. Kalia, D., Merey, G., Nakayama, S., Zheng, Y., Zhou, J., Luo, Y., Guo, M., Roembke, B.T. and Sintim, H.O. (2013) Nucleotide, c-di-GMP, c-di-AMP, cGMP, cAMP, (p)ppGpp signaling in bacteria and implications in pathogenesis. *Chemical Society reviews*, **42**, 305-341.
 171. Lauriano, C.M., Ghosh, C., Correa, N.E. and Klose, K.E. (2004) The sodium-driven flagellar motor controls exopolysaccharide expression in *Vibrio cholerae*. *Journal of bacteriology*, **186**, 4864-4874.
 172. Rashid, M.H., Rajanna, C., Zhang, D., Pasquale, V., Magder, L.S., Ali, A., Dumontet, S. and Karaolis, D.K. (2004) Role of exopolysaccharide, the rugose phenotype and VpsR in the pathogenesis of epidemic *Vibrio cholerae*. *FEMS microbiology letters*, **230**, 105-113.
 173. Sloup, R.E., Konal, A.E., Severin, G.B., Korir, M.L., Bagdasarian, M.M., Bagdasarian, M. and Waters, C.M. (2017) Cyclic Di-GMP and VpsR Induce the Expression of Type II Secretion in *Vibrio cholerae*. *Journal of bacteriology*, **199**, e00106-00117.
 174. Rappas, M., Bose, D. and Zhang, X. (2007) Bacterial enhancer-binding proteins: unlocking sigma54-dependent gene transcription. *Current opinion in structural biology*, **17**, 110-116.
 175. Ross, W., Thompson, J.F., Newlands, J.T. and Gourse, R.L. (1990) E.coli Fis protein activates ribosomal RNA transcription in vitro and in vivo. *The EMBO journal*, **9**, 3733-3742.
 176. Qiu, D., Damron, F.H., Mima, T., Schweizer, H.P. and Yu, H.D. (2008) PBAD-based shuttle vectors for functional analysis of toxic and highly regulated genes in *Pseudomonas* and *Burkholderia* spp. and other bacteria. *Applied and environmental microbiology*, **74**, 7422-7426.
 177. March-Amegadzie, R. and Hinton, D.M. (1995) The bacteriophage T4 middle promoter PuvxX: analysis of regions important for binding of the T4 transcriptional activator MotA and for activation of transcription. *Mol Microbiol*, **15**, 649-660.
 178. Thelin, K.H. and Taylor, R.K. (1996) Toxin-coregulated pilus, but not mannose-sensitive hemagglutinin, is required for colonization by *Vibrio cholerae* O1 El Tor biotype and O139 strains. *Infection and immunity*, **64**, 2853-2856.

179. de Lorenzo, V. and Timmis, K.N. (1994) Analysis and construction of stable phenotypes in gram-negative bacteria with Tn5- and Tn10-derived minitransposons. *Methods Enzymol*, **235**, 386-405.
180. Hsieh, M.L., James, T.D., Knipling, L., Waddell, M.B., White, S. and Hinton, D.M. (2013) Architecture of the bacteriophage T4 activator MotA/promoter DNA interaction during sigma appropriation. *The Journal of biological chemistry*, **288**, 27607-27618.
181. Massie, J.P., Reynolds, E.L., Koestler, B.J., Cong, J.P., Agostoni, M. and Waters, C.M. (2012) Quantification of high-specificity cyclic diguanylate signaling. *Proceedings of the National Academy of Sciences of the United States of America*, **109**, 12746-12751.
182. Teschler, J.K., Cheng, A.T. and Yildiz, F.H. (2017) The Two-Component Signal Transduction System VxrAB Positively Regulates *Vibrio cholerae* Biofilm Formation. *Journal of bacteriology*, **199**.
183. Joly, N., Zhang, N., Buck, M. and Zhang, X. (2012) Coupling AAA protein function to regulated gene expression. *Biochimica et biophysica acta*, **1823**, 108-116.
184. Pittard, J., Camakaris, H. and Yang, J. (2005) The TyrR regulon. *Molecular microbiology*, **55**, 16-26.
185. Studholme, D.J. and Dixon, R. (2003) Domain architectures of sigma54-dependent transcriptional activators. *Journal of bacteriology*, **185**, 1757-1767.
186. Herrera, M.C. and Ramos, J.L. (2007) Catabolism of phenylalanine by *Pseudomonas putida*: the NtrC-family PhhR regulator binds to two sites upstream from the phhA gene and stimulates transcription with sigma70. *Journal of molecular biology*, **366**, 1374-1386.
187. Dischert, W., Vignais, P.M. and Colbeau, A. (1999) The synthesis of *Rhodobacter capsulatus* HupSL hydrogenase is regulated by the two-component HupT/HupR system. *Molecular microbiology*, **34**, 995-1006.
188. Ueki, T. and Inouye, S. (2002) Transcriptional activation of a heat-shock gene, lonD, of *Myxococcus xanthus* by a two component histidine-aspartate phosphorelay system. *The Journal of biological chemistry*, **277**, 6170-6177.
189. Fernandez, I., Cornaciu, I., Carrica, M.D., Uchikawa, E., Hoffmann, G., Sieira, R., Marquez, J.A. and Goldbaum, F.A. (2017) Three-Dimensional Structure of Full-Length NtrX, an Unusual Member of the NtrC Family of Response Regulators. *Journal of molecular biology*, **429**, 1192-1212.
190. Ghosh, T., Bose, D. and Zhang, X. (2010) Mechanisms for activating bacterial RNA polymerase. *FEMS Microbiol Rev*, **34**, 611-627.
191. Desai, S.K., Winardhi, R.S., Periasamy, S., Dykas, M.M., Jie, Y. and Kenney, L.J. (2016) The horizontally-acquired response regulator SsrB drives a *Salmonella* lifestyle switch by relieving biofilm silencing. *Elife*, **5**, e10747.

192. Kaper, J.B., Morris, J.G., Jr. and Levine, M.M. (1995) Cholera. *Clin Microbiol Rev*, **8**, 48-86.
193. Faruque, S.M., Albert, M.J. and Mekalanos, J.J. (1998) Epidemiology, genetics, and ecology of toxigenic *Vibrio cholerae*. *Microbiology and molecular biology reviews* : *MMBR*, **62**, 1301-1314.
194. Charles, R.C. and Ryan, E.T. (2011) Cholera in the 21st century. *Curr Opin Infect Dis*, **24**, 472-477.
195. Altindis, E., Fu, Y. and Mekalanos, J.J. (2014) Proteomic analysis of *Vibrio cholerae* outer membrane vesicles. *Proceedings of the National Academy of Sciences of the United States of America*, **111**, E1548-1556.
196. Gibson, D.G., Young, L., Chuang, R.Y., Venter, J.C., Hutchison, C.A., 3rd and Smith, H.O. (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods*, **6**, 343-345.
197. Decker, K.B., Chen, Q., Hsieh, M.L., Boucher, P., Stibitz, S. and Hinton, D.M. (2011) Different requirements for sigma Region 4 in BvgA activation of the *Bordetella pertussis* promoters P(fim3) and P(fhaB). *Journal of molecular biology*, **409**, 692-709.
198. Hobman, J.L., Wilkie, J. and Brown, N.L. (2005) A design for life: prokaryotic metal-binding MerR family regulators. *Biometals*, **18**, 429-436.
199. Baxter, K., Lee, J., Minakhin, L., Severinov, K. and Hinton, D.M. (2006) Mutational analysis of sigma70 region 4 needed for appropriation by the bacteriophage T4 transcription factors AsiA and MotA. *Journal of molecular biology*, **363**, 931-944.
200. Lambert, L.J., Wei, Y., Schirf, V., Demeler, B. and Werner, M.H. (2004) T4 AsiA blocks DNA recognition by remodeling sigma70 region 4. *The EMBO journal*, **23**, 2952-2962.
201. Hinton, D.M. (2010) Transcriptional control in the prereplicative phase of T4 development. *Virol J*, **7**, 289.
202. Lawson, C.L., Swigon, D., Murakami, K.S., Darst, S.A., Berman, H.M. and Ebright, R.H. (2004) Catabolite activator protein: DNA binding and transcription activation. *Current opinion in structural biology*, **14**, 10-20.
203. Beier, D. and Gross, R. (2006) Regulation of bacterial virulence by two-component systems. *Curr Opin Microbiol*, **9**, 143-152.
204. Ulrich, L.E. and Zhulin, I.B. (2010) The MiST2 database: a comprehensive genomics resource on microbial signal transduction. *Nucleic acids research*, **38**, D401-407.
205. Hess, J.F., Oosawa, K., Kaplan, N. and Simon, M.I. (1988) Phosphorylation of three proteins in the signaling pathway of bacterial chemotaxis. *Cell*, **53**, 79-87.

206. Zapf, J., Madhusudan, M., Grimshaw, C.E., Hoch, J.A., Varughese, K.I. and Whiteley, J.M. (1998) A source of response regulator autophosphatase activity: the critical role of a residue adjacent to the Spo0F autophosphorylation active site. *Biochemistry*, **37**, 7725-7732.
207. Nguyen, M.P., Yoon, J.M., Cho, M.H. and Lee, S.W. (2015) Prokaryotic 2-component systems and the OmpR/PhoB superfamily. *Can J Microbiol*, **61**, 799-810.
208. Morris, J.G., Jr. (2011) Cholera--modern pandemic disease of ancient lineage. *Emerg Infect Dis*, **17**, 2099-2104.
209. de Magny, G.C., Mozumder, P.K., Grim, C.J., Hasan, N.A., Naser, M.N., Alam, M., Sack, R.B., Huq, A. and Colwell, R.R. (2011) Role of zooplankton diversity in *Vibrio cholerae* population dynamics and in the incidence of cholera in the Bangladesh Sundarbans. *Applied and environmental microbiology*, **77**, 6125-6132.
210. Vezzulli, L., Pezzati, E., Moreno, M., Fabiano, M., Pane, L., Pruzzo, C. and VibrioSea, C. (2009) Benthic ecology of *Vibrio* spp. and pathogenic *Vibrio* species in a coastal Mediterranean environment (La Spezia Gulf, Italy). *Microb Ecol*, **58**, 808-818.
211. Collin, B. and Rehnstam-Holm, A.S. (2011) Occurrence and potential pathogenesis of *Vibrio cholerae*, *Vibrio parahaemolyticus* and *Vibrio vulnificus* on the South Coast of Sweden. *FEMS Microbiol Ecol*, **78**, 306-313.
212. Pratt, J.T., McDonough, E. and Camilli, A. (2009) PhoB regulates motility, biofilms, and cyclic di-GMP in *Vibrio cholerae*. *Journal of bacteriology*, **191**, 6632-6642.
213. Sultan, S.Z., Silva, A.J. and Benitez, J.A. (2010) The PhoB regulatory system modulates biofilm formation and stress response in El Tor biotype *Vibrio cholerae*. *FEMS microbiology letters*, **302**, 22-31.
214. Tischler, A.D. and Camilli, A. (2004) Cyclic diguanylate (c-di-GMP) regulates *Vibrio cholerae* biofilm formation. *Molecular microbiology*, **53**, 857-869.
215. Bilecen, K., Fong, J.C., Cheng, A., Jones, C.J., Zamorano-Sanchez, D. and Yildiz, F.H. (2015) Polymyxin B resistance and biofilm formation in *Vibrio cholerae* are controlled by the response regulator CarR. *Infection and immunity*, **83**, 1199-1209.
216. Lenz, D.H., Miller, M.B., Zhu, J., Kulkarni, R.V. and Bassler, B.L. (2005) CsrA and three redundant small RNAs regulate quorum sensing in *Vibrio cholerae*. *Molecular microbiology*, **58**, 1186-1202.
217. McKee, R.W., Kariisa, A., Mudrak, B., Whitaker, C. and Tamayo, R. (2014) A systematic analysis of the in vitro and in vivo functions of the HD-GYP domain proteins of *Vibrio cholerae*. *BMC Microbiol*, **14**, 272.
218. Krasteva, P.V., Giglio, K.M. and Sondermann, H. (2012) Sensing the messenger: the diverse ways that bacteria signal through c-di-GMP. *Protein Sci*, **21**, 929-948.

219. Tamayo, R., Pratt, J.T. and Camilli, A. (2007) Roles of cyclic diguanylate in the regulation of bacterial pathogenesis. *Annual review of microbiology*, **61**, 131-148.
220. Skorupski, K. and Taylor, R.K. (1996) Positive selection vectors for allelic exchange. *Gene*, **169**, 47-52.
221. Correa, N.E., Lauriano, C.M., McGee, R. and Klose, K.E. (2000) Phosphorylation of the flagellar regulatory protein FlrC is necessary for *Vibrio cholerae* motility and enhanced colonization. *Molecular microbiology*, **35**, 743-755.
222. Klose, K.E., Weiss, D.S. and Kustu, S. (1993) Glutamate at the site of phosphorylation of nitrogen-regulatory protein NTRC mimics aspartyl-phosphate and activates the protein. *Journal of molecular biology*, **232**, 67-78.
223. Fiedler, U. and Weiss, V. (1995) A common switch in activation of the response regulators NtrC and PhoB: phosphorylation induces dimerization of the receiver modules. *The EMBO journal*, **14**, 3696-3705.
224. Laguri, C., Stenzel, R.A., Donohue, T.J., Phillips-Jones, M.K. and Williamson, M.P. (2006) Activation of the global gene regulator PrrA (RegA) from *Rhodobacter sphaeroides*. *Biochemistry*, **45**, 7872-7881.
225. Leoni, L., Ascenzi, P., Bocedi, A., Rampioni, G., Castellini, L. and Zennaro, E. (2003) Styrene-catabolism regulation in *Pseudomonas fluorescens* ST: phosphorylation of StyR induces dimerization and cooperative DNA-binding. *Biochemical and biophysical research communications*, **303**, 926-931.
226. Menon, S. and Wang, S. (2011) Structure of the response regulator PhoP from *Mycobacterium tuberculosis* reveals a dimer through the receiver domain. *Biochemistry*, **50**, 5948-5957.
227. Nakashima, K., Kanamaru, K., Aiba, H. and Mizuno, T. (1991) Signal transduction and osmoregulation in *Escherichia coli*. A novel type of mutation in the phosphorylation domain of the activator protein, OmpR, results in a defect in its phosphorylation-dependent DNA binding. *The Journal of biological chemistry*, **266**, 10775-10780.
228. Gao, R. and Stock, A.M. (2009) Biological insights from structures of two-component proteins. *Annual review of microbiology*, **63**, 133-154.
229. Boucher, P.E., Murakami, K., Ishihama, A. and Stibitz, S. (1997) Nature of DNA binding and RNA polymerase interaction of the *Bordetella pertussis* BvgA transcriptional activator at the *fha* promoter. *Journal of bacteriology*, **179**, 1755-1763.
230. Sloup, R.E., Konal, A.E., Severin, G.B., Korir, M.L., Bagdasarian, M.M., Bagdasarian, M. and Waters, C.M. (2017) Cyclic Di-GMP and VpsR Induce the Expression of Type II Secretion in *Vibrio cholerae*. *J Bacteriol*, **199**.

231. Wolfe, A.J., Chang, D.E., Walker, J.D., Seitz-Partridge, J.E., Vidaurri, M.D., Lange, C.F., Pruss, B.M., Henk, M.C., Larkin, J.C. and Conway, T. (2003) Evidence that acetyl phosphate functions as a global signal during biofilm development. *Mol Microbiol*, **48**, 977-988.
232. Roelofs, K.G., Wang, J., Sintim, H.O. and Lee, V.T. (2011) Differential radial capillary action of ligand assay for high-throughput detection of protein-metabolite interactions. *Proceedings of the National Academy of Sciences of the United States of America*, **108**, 15528-15533.
233. Sambanthamoorthy, K., Schwartz, A., Nagarajan, V. and Elasri, M.O. (2008) The Role of msa in Staphylococcus aureus Biofilm Formation. *BMC microbiology*, **8**, 221.
234. Harrison, J.J., Stremick, C.A., Turner, R.J., Allan, N.D., Olson, M.E. and Ceri, H. (2010) Microtiter susceptibility testing of microbes growing on peg lids: a miniaturized biofilm model for high-throughput screening. *Nature protocols*, **5**, 1236-1254.
235. Steiner, S., Lori, C., Boehm, A. and Jenal, U. (2013) Allosteric activation of exopolysaccharide synthesis through cyclic di-GMP-stimulated protein-protein interaction. *The EMBO journal*, **32**, 354-368.
236. Battesti, A. and Bouveret, E. (2012) The bacterial two-hybrid system based on adenylate cyclase reconstitution in Escherichia coli. *Methods*, **58**, 325-334.
237. Murakami, K.S. (2013) X-ray crystal structure of Escherichia coli RNA polymerase sigma70 holoenzyme. *The Journal of biological chemistry*, **288**, 9126-9134.
238. Niu, W., Kim, Y., Tau, G., Heyduk, T. and Ebright, R.H. (1996) Transcription activation at class II CAP-dependent promoters: two interactions between CAP and RNA polymerase. *Cell*, **87**, 1123-1134.
239. Greiner, D.P., Miyake, R., Moran, J.K., Jones, A.D., Negishi, T., Ishihama, A. and Meares, C.F. (1997) Synthesis of the protein cutting reagent iron (S)-1-(p-bromoacetamidobenzyl)ethylenediaminetetraacetate and conjugation to cysteine side chains. *Bioconjugate chemistry*, **8**, 44-48.
240. Freddolino, P.L., Amini, S. and Tavazoie, S. (2012) Newly identified genetic variations in common Escherichia coli MG1655 stock cultures. *Journal of bacteriology*, **194**, 303-306.