Functional analysis of GGDEF-EAL didomain-containing proteins in cyclic-di-GMP signaling

Qi, Yaning

2011

Qi, Y. N. (2011). Functional analysis of GGDEF-EAL didomain-containing proteins in cyclic-di-GMP signaling. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/44595>

<https://doi.org/10.32657/10356/44595>

**NANYANG
TECHNOLOGICAL
UNIVERSITY**

**FUNCTIONAL ANALYSIS OF GGDEF-EAL
DIDOMAIN-CONTAINING PROTEINS IN
CYCLIC-DI-GMP SIGNALING**

QI YANING

SCHOOL OF BIOLOGICAL SCIENCES

2011

**FUNCTIONAL ANALYSIS OF GGDEF-EAL
DIDOMAIN-CONTAINING PROTEINS IN
CYCLIC-DI-GMP SIGNALING**

QI YANING

SCHOOL OF BIOLOGICAL SCIENCES
NANYANG TECHNOLOGICAL UNIVERSITY

A thesis submitted to the Nanyang Technological University

In partial fulfillment of the requirements for the degree of

Doctor of Philosophy

2011

ACKNOWLEDGEMENTS

I would like to dedicate this thesis to my parents for their endless love in the past twenty-six years. Without their support and encouragement, this thesis would never have been written.

My heartfelt thanks go to my supervisor Dr Liang Zhao-xun for his invaluable guidance and advice throughout the PhD project. It is my great honor to have worked with him. Besides the fancy techniques and scientific ideas, the enthusiasm and perseverance for science that he endowed me would be my greatest treasure during the past four years.

A special appreciation to my dear boyfriend Dr Wang Chunming for his accompany throughout my tough days.

I wish to extent my sincere thanks to our collaborators that have helped me in one way or another: Dr Tang Kai for hydrogen/deuterium exchange mass spectrometry; Dr Julien Lescar for protein crystallography and structural analysis; Dr Li Jinming for the bioinformatic studies.

I also would like to thank all my labmates: Rao Feng, Kong Rong, Sun Huihua, Laurence Ho, Mary Chua, Liew Chong wai, Ela, Swathi, Ho Qinshi, Ji Qiang and Zhang Dong wei for always being there to help me. I will never forget the days we have been together. Thanks also go to the FYP and attachment students Xie Kai ling, Luo Zhen, Goh Tian ying, Lee Gok chuan for their efforts and help.

Last but not least, I would like to thank my dear friends Wang Huijuan, Feng Xia, Jia Yan, Wang Lu, Cao Lei, Luo Dan, Wei Li, Zhuang Tao, Zheng Jia wen, Wu Yu, Meng Wei, Wang Wei and Wang Xi. My life here in Singapore would be dull and unbearable without you.

TABLE OF CONTENTS

ACKNOWLEDGEMENT	I
TABLE OF CONTENTS	II
LIST OF FIGURES	V
LIST OF TABLES	VIII
ABBREVIATION.	IX
SUMMARY.....	XIII
CHAPTER 1 INTRODUCTION	1
1.1 Cyclic-di-GMP as a signaling molecule.....	1
1.2 Molecular mechanism of c-di-GMP metabolism	5
1.2.1 Diguanylate cyclases.....	6
1.2.2 C-di-GMP specific phosphodiesterases	7
1.2.3 Regulation of DGC and PDE-A activity by regulatory domains	9
1.3 GGDEF-EAL didomain proteins.....	10
1.3.1 Didomain proteins with enzymatically active GGDEF domains.....	10
1.3.2 Didomain proteins with enzymatically active EAL domains	11
1.3.3 Didomain proteins with both enzymatically active GGDEF and EAL domains	21
1.3.4 Didomain proteins with enzymatically inactive GGDEF and EAL domains	24
1.4 Objectives and organization of this dissertation.....	25
CHAPTER 2 CATALYTIC MECHANISM OF THE EAL DOMAIN OF PA2567.....	29
2.1 Introduction	29
2.2 Methods and Materials	32
2.2.1 Gene cloning and site-directed mutagenesis	32
2.2.2 Protein expression and purification	32
2.2.3 Steady-state enzymatic activity assay	34

2.2.4 Bioinformatic analysis and structural modeling	35
2.3 Results	36
2.3.1 Purification and characterization of PA2567	36
2.3.2 Site-directed mutagenesis study	37
2.4 Discussion	39
2.4.1 PA2567 supports the proposed EAL domain catalytic mechanism	39
2.4.3 Classification of EAL domains	44
2.5 Summary	48
CHAPTER 3 REGULATION OF C-DI-GMP SYNTHESIS IN AxDGC2	50
3.1 Introduction	50
3.2 Methods and Materials	53
3.2.1 Gene cloning and site-directed mutagenesis	53
3.2.2 Protein expression and purification	54
3.2.3 Characterization of the protein-associated flavin and c-di-GMP.....	55
3.2.4 Enzymatic assay of DGC activity	56
3.2.5 Enzymatic assay of PDE activity	58
3.2.6 Measurement of FAD to protein binding stoichiometry	59
3.2.7 Determination of the redox potential of flavin in AxDGC2.....	60
3.2.8 Sequence alignment and Structural modeling.....	60
3.3 Results	61
3.3.1 Purification and characterization of AxDGC2	61
3.3.2 AxDGC2 binds FAD cofactor in the PAS domain	62
3.3.3 Regulation of the DGC activity by the oxidation state of FAD	64
3.3.4 Effects of mutations in the PAS domain on DGC activity	68
3.3.5 Effects of mutations in the PAS domain on redox potential	74
3.3.6 Effects of mutations in I-site on c-di-GMP binding.....	76

3.3.7 Phosphodiesterase activity of the EAL domain	77
3.4 Discussion	79
3.5 Summary	85
CHAPTER 4 INSIGHT INTO THE FUNCTION OF THE DEGENERATE EAL DOMAIN OF THE C-DI-GMP SIGNALING PROTEIN FIMX.....	87
4.1 Introduction	87
4.2 Methods and Materials	90
4.2.1 Gene cloning and site-directed mutagenesis	90
4.2.2 Protein expression and purification	91
4.2.3 Characterization of the protein-associated c-di-GMP	92
4.2.4 Amide hydrogen-deuterium (H/D) Exchange by mass spectrometry	93
4.2.4.1 HPLC column packing and system configuration	95
4.2.4.2 Hydrogen-Deuterium Exchange	96
4.2.4.3 Data analysis	98
4.2.5 Isothermal titration calorimetry (ITC)	101
4.2.6 Structural modeling	102
4.3 Results	102
4.3.1 Preliminary characterization of full-length FimX and FimX ₄₃₆₋₆₉₁	102
4.3.2 Probe the interaction between FimX and c-di-GMP by amide H/D exchange	103
4.3.3 Effect of the mutation of the EVL motif on c-di-GMP binding	111
4.4 Discussion	115
4.5 Summary	118
CHAPTER 5 CONCLUSIONS AND FUTURE WORKS.....	120
5.1 Conclusions	120
5.2 Future outlooks	123
REFERENCES.....	126
PUBLICATION LIST.....	139

LIST OF FIGURES

Figure 1.1	C-di-GMP acts as a global signaling molecule	4
Figure 1.2	Synthesis and degradation of c-di-GMP	5
Figure 1.3	Structure of GGDEF domain of PleD (PDB code: 2V0N).	7
Figure 1.4	Domain organization of representative GGDEF and EAL proteins	13
Figure 2.1	Domain structure of PA2567 protein.....	31
Figure 2.2	SDS-PAGE and gel filtration characterization of PA2567	37
Figure 2.3	C-di-GMP hydrolysis by PA2567	38
Figure 2.4	Effect of the E464A mutation on the steady-state kinetics of PA2567	40
Figure 2.5	Detailed views of EAL domain catalytic active sites	43
Figure 2.6	Structural model of EAL _{RocR} with the residue of loop 6 and Glu ²⁶⁸ highlighted	44
Figure 2.7	Summary of EAL domains	47
Figure 3.1	Comparison of the domain organization of the three PAS domain-containing proteins in c-di-GMP signaling	53
Figure 3.2	SDS-PAGE and gel filtration analysis of purified AxDGC2 protein.....	61
Figure 3.3	Absorption spectrum of AxDGC2	63
Figure 3.4	The MALDI-MS result of FAD bound by AxDGC2.....	64
Figure 3.5	AxDGC2 reduced by sodium dithionite.....	66
Figure 3.6	DGC activity of oxidized and reduced AxDGC2	67
Figure 3.7	Comparison of the diguanylate cyclase activity of the oxidized and reduced forms of AxDGC2 and the FAD-free N94A mutant.....	68

Figure 3.8	Comparison of PAS _{AxDGC2} domain with other flavin-binding PAS domains.....	70
Figure 3.9	Comparison of AxDGC2 and the N94A mutant	71
Figure 3.10	Comparison of AxDGC2 and its mutants	73
Figure 3.11	Modified FAD in mutant H62A	74
Figure 3.12	Nernst plots of log(ox/red) phenosafranin as a function of log (ox/red) AxDGC2-FAD	75
Figure 3.13	Analysis of I-site	77
Figure 3.14	Catalytic activity of the EAL domain of AxDGC2 towards non-physiological phosphodiester substrates	78
Figure 3.15	Comparison of the phosphodiesterase activity of the oxidized and reduced forms of AxDGC2 with the use of 1 mM sodium dithionite	79
Figure 3.16	A schematic illustration of the relationship between oxygen level, cellular c-di-GMP concentration and cellulose synthesis in <i>A. xylinum</i>	85
Figure 4.1	Domain organization and sequence alignment of <i>Pseudomonas aeruginosa</i> FimX	90
Figure 4.2	The general overview of the experimental setup of Hydrogen/Deuterium Exchange-Mass Spectrometry	94
Figure 4.3	SDS-PAGE and gel filtration characterization of the purified FimX protein and FimX ₄₃₆₋₆₉₁	104
Figure 4.4	Sequence coverage map of FimX with the secondary structure indicated	107
Figure 4.5	H/D exchange profiles for the peptides that do not exhibit significant changes upon c-di-GMP binding.....	108
Figure 4.6	Effect of c-di-GMP binding on deuteration pattern in the EAL domain	112

Figure 4.7	Effect of c-di-GMP binding on the deuteration pattern in the N-terminal REC domain and the linker region.....	113
Figure 4.8	Isothermal titration calorimetry (ITC) analysis of the effect of mutation in the EVL motif on c-di-GMP binding.	114

LIST OF TABLES

Table 2.1	Primers used in PA2567 cloning and site-directed mutagenesis	32
Table 2.2	Steady-state kinetic parameters for PA2567 and its mutants ...	39
Table 3.1	Primers used in AxDGC2 cloning and site-directed mutagenesis.....	54
Table 3.2	Maxima of absorption spectra of riboflavin-containing PAS proteins.....	63
Table 3.3	FAD: protein stoichiometry and steady-state kinetic parameters of AxDGC2 and its mutants	72
Table 3.4	Redox potentials of AxDGC2 and its mutants	75
Table 4.1	Primers used in FimX cloning and site-directed mutagenesis...	91

ABBREVIATIONS

2'3'-cAMP	2'3'-cyclic adenosine monophosphate
3'-AMP	3'-adenosine monophosphate
5'-pGpG	5'-linear diguanylic acid
6'-OH-FAD	6-hydroxyl-flavin adenine dinucleotide
aa	Amino acids
ACN	Acetonitrile
A-site	Activity site
<i>A. vinelandii</i>	<i>Azotobacter vinelandii</i>
<i>A.xylinum</i>	<i>Acetobacter xylinum</i>
bis-pNPP	bis- <i>p</i> -nitrophenyl phosphate
BLUF	Blue-light photoreceptor domain
β-Me	β- mercaptoethanol
BV	Biliverdin
c-di-GMP	Cyclic-di-guanosine monophosphate
cGMP	Cyclic guanosine monophosphate
DGC	Diguanylate cyclase
DSF	Differential scanning fluorimetry
DTT	Dithiothreitol
EAL	Glutamate-alanine-leucine domain harboring c-di-GMP specific PDE activity
<i>E.coli</i>	<i>Escherichia coli</i>
EDTA	Ethylenediaminetetraacetic acid
EPS	Extracellular polysaccharides
ESI-MS	Electrospray ionization-mass spectrometry
FA	Formic acid
FAD	Flavin adenine dinucleotide

FMN	Flavin mononucleotide
FPLC	Fast performance liquid chromatography
FRET	Fluorescence resonance energy transfer
GGDEF	Glycine-glycine-asparatate-glutamate-glutamate domain harboring DGC activity
GFP	Green fluorescent protein
GMP	Guanosine monophosphate
GTP	Guanosine triphosphate
	Histidine kinases, adenylyl cyclases, methyl binding proteins and phosphatases
HAMP	
H/D	Hydrogen/deuterium
	Histidine-aspartic acid-glycine-tyrosine-proline
HD-GYP	domain
HDX-MS	Hydrogen/deuterium exchange mass spectrometry
HPLC	High performance liquid chromatography
HMWO	High molecular weight oligomer
I.D.	Inner diameter
IPTG	Isopropyl- β -D-Thiogalactopyranoside
I-site	Inhibitory site
ITC	Isothermal titration calorimetry
k_{cat}	Turnover number
K_d	Dissociation constant
kDa	KiloDaltons
K_i	Inhibitor constant
K_M	The Michaelis constant
<i>K. pneumonia</i>	<i>Klebsiella pneumonia</i>
LB	Luria-Bertani Media
LC/MS	Liquid chromatography-mass spectrometry
LOV	Light, oxygen or voltage

MS	Mass spectrometry
<i>M. smegmatis</i>	<i>Mycobacterium smegmatis</i>
MS/MS	Tandem mass spectrometry
MW	Molecular Weight
MALDI	Matrix-assisted laser desorption/ionization
m/z	mass to charge
Ni ²⁺ -NTA	Nickel-nitrilotriacetic acid
NMR	Nuclear magnetic resonance
PAS	Per-Arnt-Sim domain
PCR	Polymerase chain reaction
PDB	Protein data bank
PDE	Phosphodiesterase
	Phosphodiesterase A, c-di-GMP specific hydrolysis
PDE-A	activity by EAL domain
PEEK	Polyetheretherketones
PilZ	Pilis domain acting as c-di-GMP receptor
PMSF	Phenylmethanesulphonyl fluoride
pNPP	<i>p</i> -Nitrophenyl Phosphate
<i>P. aeruginosa</i>	<i>Pseudomonas aeruginosa</i>
<i>P. fluorescens</i>	<i>Pseudomonas fluorescens</i>
<i>psi</i>	Pound-force per square inch
REC	Response regulator
<i>rpm</i>	Revolutions per minute
RP-HPLC	Reverse phase-high pressure liquid chromatography
RT	Retention time
<i>S. Typhimurium</i>	<i>Salmonella Typhimurium</i>
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide electrophoresis
TFA	Trifluoro acetic acid
Tfp	Type IV pili

TIC	Total ion current
TM	Transmembrane
UV-Vis	Ultraviolet-visible
<i>V. cholerae</i>	<i>Vibrio cholerae</i>
<i>V. parahaemolyticus</i>	<i>Vibrio parahaemolyticus</i>
WT	Wild type
<i>X. campestris</i>	<i>Xanthomonas campestris</i>
<i>Xcc</i>	<i>Xanthomonas campestris</i> pv. <i>campestris</i>
<i>Y. pestis</i>	<i>Yersinia pestis</i>

SUMMARY

C-di-GMP is a ubiquitous bacterial second messenger that regulates twitching motility, virulence secretion, cell surface adhesion and biofilm formation. The intracellular concentrations of c-di-GMP are controlled by the GGDEF domain proteins with diguanylate cyclase (DGC) activity, and the EAL domain or HD-GYP proteins with c-di-GMP specific phosphodiesterase (PDE) activity. The aim of this study is to investigate the functions of three representative GGDEF-EAL didomain-containing signaling proteins, with the major foci on their catalytic and regulatory mechanisms.

The first protein under investigation is PA2567, a protein from the opportunistic pathogen *Pseudomonas aeruginosa* that contains a GAF domain and the GGDEF-EAL didomain. Enzymatic assays revealed that PA2567 contains a catalytically competent EAL domain but an inactive GGDEF domain. Site-directed mutagenesis in combination with kinetic studies suggested the critical roles of several conserved residues and a functional loop in the putative substrate-binding pocket. The results are supportive of a one-metal-ion or two-metal-ion catalytic mechanism and the existence of a large number of inactive EAL domains.

The second protein is AxDGC2 from *Acetobacter xylinum*, which contains a Per-Arnt-Sim (PAS) domain and the GGDEF-EAL didomain. AxDGC2 exhibits DGC activity but not PDE activity. We found that the PAS domain of

AxDGC2 binds flavin adenosine dinucleotide (FAD) as cofactor, and that the DGC activity is modulated by the redox status of the FAD cofactor, with the oxidized form exhibiting higher catalytic activity and apparent substrate inhibition. The results indicate that AxDGC2 is a signaling protein that regulates cellular c-di-GMP level in response to the change of cellular redox status or oxygen concentration. Mutagenesis studies demonstrated that despite the perturbation of redox potential and the unexpected modification of FAD in one of the mutants, none of the single mutations introduced in the PAS domain to disrupt FAD binding and signal transduction was able to completely block the signal transmission to the GGDEF domain. We propose that the change of FAD redox state can still trigger structural changes in the PAS domain by using substituted hydrogen-bonded water networks. Together with the oxygen-dependent activity of the homologous AxPDEA1, the study provided fresh insight into the relationships between oxygen level, c-di-GMP concentration and cellulose synthesis in *Acetobacter xylinum*.

Lastly, FimX is a GGDEF-EAL didomain-containing protein that is indispensable for the normal twitching motility in *Pseudomonas aeruginosa*. We found that the GGDEF and EAL domains of FimX to be catalytically inactive, and that the inactive EAL domain of FimX binds c-di-GMP with high affinity. We further found that the binding of c-di-GMP in FimX triggers a conformational change in the remote REC domain and the linker region by hydrogen-deuterium (H/D) exchange coupled mass spectrometry. The result

provides a molecular explanation for the change of cellular localization for FimX upon the abolishment of the c-di-GMP binding capability. Given the large number of non-enzymatic EAL domains encoded by bacterial genomes, the molecular mechanisms underlying the c-di-GMP induced conformational changes and signal propagation may be conserved in other non-enzymatic EAL domains.

CHAPTER 1 INTRODUCTION

1.1 Cyclic-di-GMP as a signaling molecule

Bis-(3',5')-cyclic dimeric guanosine monophosphate (cyclic di-GMP or c-di-GMP) was first identified as an activator of the cellulose synthetase complex in *Acetobacter xylinum* 20 years ago (107). In recent years, researchers have discovered an increasingly important role of c-di-GMP in bacterial pathogenesis. The phenotypes mediated by c-di-GMP range from twitching motility, biofilm formation, virulence gene expression, cell-cell signaling to pili and flagellar synthesis (59, 105-106, 127) (**Figure 1.1**). In a short period of several years, c-di-GMP has established itself as a ubiquitous second messenger in the bacterial kingdom.

One striking role of c-di-GMP is to mediate various bacterial motilities that include swimming, swarming, and twitching motility. Swimming in liquid and swarming across surfaces are powered by the flagella, whereas twitching motility is mediated by the repeated extension and retraction of Type IV pili (Tfp) (37, 81). Recent studies have shown that the functions of both flagella and Tfp are under the control of c-di-GMP. Down regulation of flagellar motility by c-di-GMP has been demonstrated in *Salmonella Typhimurium* and *Vibrio cholerae* (7, 118), while the regulation of the transcription of several proteins involved in flagellar synthesis by c-di-GMP is also revealed in several bacterial strains. Deletion of the gene that encodes FimX, a c-di-GMP signaling protein

in *Pseudomonas aeruginosa*, inhibits the Tfp assembly on the cell surface, demonstrating the control of twitching motility by c-di-GMP (55, 61).

C-di-GMP also regulates the production of extracellular polysaccharides (EPS), a key component of bacterial biofilms. Microbial cells adhere to each other or to the surface by embedding themselves in the EPS to form structured biofilm microcolonies. Biofilm forms on living or non-living surfaces, and represents a prevalent mode of microbial life in natural, industrial and hospital settings. Microbial cells growing in a biofilm are physiologically distinct from the planktonic cells or free-swimming cells. C-di-GMP seems to control the switch between the motile planktonic and sedentary biofilm-associated life styles. High level of intracellular c-di-GMP tends to promote biofilm formation and repress twitching motility and flagellar synthesis, and *vice versa*. Such regulatory role has been observed in numerous pathogens such as *V. cholerae*, *S. Typhimurium*, *P. aeruginosa*, and *Yersinia pestis* (36, 51, 64, 118, 136-137).

The third aspect of bacterial pathogenesis under the control of c-di-GMP is virulence gene expression. Increasing studies by genetic screening and transcriptional profiling have linked the regulation of virulence to c-di-GMP signaling in various bacteria. The most prominent example is the CLP protein from *Xanthomonas campestris*. Deletion of *clp* gene in *X. campestris* has changed the transcriptional expression level of 299 genes including a few genes that encode transcription factors (47). This cAMP receptor-like protein was later identified as a c-di-GMP receptor. Structural and biochemical studies have

revealed that the binding between CLP and the target promoter DNA is abrogated by c-di-GMP bound to CLP with micromolar affinity. Together, these findings suggest that the CLP protein regulates the virulence genes expression in *X. campestris* and this regulation could be negatively regulated by cellular c-di-GMP concentration (16).

In the past several years, the roles of c-di-GMP in bacterial physiology and pathogenesis have been examined intensively (**Figure 1.1**). How c-di-GMP is distributed in the bacterial cells is another topic under investigation. In a recent work, genetically encoded fluorescence resonance energy transfer (FRET)-based biosensors have been developed to monitor the concentration of c-di-GMP within bacterial cells by microscopy. It was found that there is a tight spatial and temporal regulation of c-di-GMP in the presence of diverse receptors, including PilZ domain proteins, mRNA riboswitch, and c-di-GMP binding transcription factors (1, 88, 111, 122). Christen and colleagues engineered a set of (FRET)-based biosensors by fusing PilZ proteins, one of the c-di-GMP receptors, with fluorescent tags. It is observed that the fluorescence properties of these biosensors could reflect cellular c-di-GMP levels (21).

Lastly, c-di-GMP has been found as a potent immunostimulant that exhibits promising potentials as a vaccine adjuvant for cancer treatment and prevention of bacterial infection. C-di-GMP inhibits the proliferation of various cancer cell lines by locking majority of cells in the S (DNA synthesis) phase of the cell cycle (82, 90). C-di-GMP can also stimulate protective innate immunity against

various bacterial infections. For example, in pneumococcal diseases, subcutaneous administration of c-di-GMP elevates the expression of chemokines and type 1 cytokines and invokes the accumulation of neutrophils, α/β T cells and activated natural killer (NK) cells. Hence, it is promising to further explore c-di-GMP as a vaccine adjuvant to activate innate host defense against respiratory and systemic infection (90).

Figure 1.1 C-di-GMP acts as a global signaling molecule. The arrows pointing down indicate the physiological functions regulated by c-di-GMP such as cellulose synthesis, twitching motility, biofilm formation and virulence gene expression in bacteria. The arrows pointing up indicate the novel application of c-di-GMP in biotechnology like biosensors and vaccine adjuvant.

1.2 Molecular mechanism of c-di-GMP metabolism

The synthesis of c-di-GMP is catalyzed by the diguanylate cyclases (DGCs) proteins that contain the GGDEF domains. The GGDEF domains catalyze the formation of c-di-GMP from two GTP molecules. The hydrolysis of c-di-GMP requires phosphodiesterases (PDEs) that contains either a conserved EAL domain (PDE-A) to produce linear molecule 5'-pGpG or a HD-GYP domain to generate GMP. The intracellular concentration of c-di-GMP is believed to be regulated by the ubiquitous GGDEF, EAL, and HD-GYP domain proteins in bacterial cells (**Figure 1.2**) (59). Although there are various proteins with single GGDEF or EAL domains, a number of GGDEF and EAL containing proteins contain both domains in tandem.

Figure 1.2 Synthesis and degradation of c-di-GMP. Conversion of two molecules of GTP to c-di-GMP by diguanylate cyclase (DGC) by the GGDEF domain and hydrolysis from c-di-GMP to 5'-pGpG and GMP by the phosphodiesterase-A(PDE-A) protein with an EAL domain or HD-GYP domain.

1.2.1 Diguanylate cyclases

Study of the GGDEF domain proteins from bacteria such as *Pseudomonas fluorescens*, *S. typhimurium*, *Y. pestis*, and *Escherichia coli* confirmed that the DGC activity can be attributed to the GGDEF domain proteins (15, 64, 68, 74). Overexpression of DGC proteins has been demonstrated to increase cellular concentration of c-di-GMP (93, 112, 118).

The GGDEF domain of the *Caulobacter crescentus* protein PleD has been characterized in detail by Hecht and coworkers (48, 94). PleD contains a C-terminal GGDEF domain and two N-terminal receiver domains. The protein controls cell transition from the swarmer (flagellated form) to the stalked form (surface attached form) (35). *In vitro* enzymatic assays provide direct evidence to support that the GGDEF domain exhibits DGC activity when it forms a homodimer in the proper conformation. Structural determination of PleD in complex with the GTP substrate analog reveals a two Mg^{2+} ions assisted catalytic mechanism. Synthesis of c-di-GMP is catalyzed by the cooperative action of the two GGDEF domains that each binds a GTP molecule. Site-directed mutagenesis suggests that the conserved signature motif GGDEF constitutes the catalytic active sites along with other substrate binding residues (92, 145) (**Figure 1.3A**).

Determination of the PleD crystal structure also revealed a product inhibitory site (I-site) that is independent of the active sites. The I-site consists of an RxxD motif that can bind c-di-GMP to cause noncompetitive product

inhibition (13, 19, 145) (**Figure 1.3C**). Sequence alignment reveals that more than 60% of the GGDEF domains from the bacterial genomes contain a putative I-site with a RxxD motif (R can be replaced by a different charged residue) (19, 114).

Figure 1.3 Structure of GGDEF domain of PleD (PDB code: 2V0N) from *Caulobacter crescentus* with the active site (A-site) and product inhibitory site (I-site) highlighted (145). (A) Close up view of the I-site with c-di-GMP bound. (B) Overview of the GGDEF domain of PleD with GTP bound to the A-site and c-di-GMP bound to the I-site. (C) Active sites in complex with the substrate analog GTP α S and two magnesium ions.

1.2.2 C-di-GMP specific phosphodiesterases

The name of the ubiquitous EAL domain originates from the conserved signature EAL (Glu-Ala-Leu) motif. It contains approximately 250-amino-acids and is encoded by genes in most sequenced bacterial genomes (115). Benziman and colleagues first revealed the roles of EAL and GGDEF domains in regulating cellulose synthesis in *A. xylinum* (126). Later, based on the study of *Salmonella enterica* protein YhjH and *V. cholerae* protein VieA, the two EAL

domains have been confirmed to possess PDE activity since over-expression of these proteins reduced the intracellular level of c-di-GMP (33, 129). Mark Gomelsky and colleagues identified a single EAL domain protein YahA and demonstrated its c-di-GMP specific PDE activity *in vitro* (115). Intriguingly, as demonstrated by the results from this study and other labs, some EAL domain proteins are enzymatically incompetent due to the mutation of the signature EAL motif and other key catalytic residues (84, 88, 125, 138).

HD-GYP domain is a subfamily of the HD family metal-dependent phosphohydrolases. The first HD-GYP domain protein demonstrated to possess the c-di-GMP specific PDE activity is from the plant pathogenic *Xanthomonas campestris* pv. *campestris* (*Xcc*) (109). In *Xcc*, the two-component signaling transduction system *rpfC* and *rpfG* controls the synthesis of extracellular polysaccharide (EPS), expression of virulence factors and biofilm formations. The HD-GYP domain regulatory protein RpfG and the protein RpfC function as the response regulators of the two-component signaling system. Mutation of *rpfG* leads to a coordinated reduction in the synthesis of the EPS expression, which leads to the suppressed biofilm formation. In consistent with its *in vivo* functions, RpfG exhibits c-di-GMP specific phosphodiesterase activity *in vitro*, directly generating GMP (3, 28). However, it should be noted that there are much less frequent occurrence of HD-GYP domains than EAL domains in bacterial genomes. Therefore, it is clear that the EAL domain proteins are the major PDEs for c-di-GMP degradation.

1.2.3 Regulation of DGC and PDE-A activity by regulatory domains

Most GGDEF and EAL or HD-GYP domains are associated with putative regulatory and sensor domains. The most common sensor domains include REC (Response Regulator), PAS (Per-Arnt-Sim), BLUF (Blue-Light Photoreceptor), and GAF domains (41, 57, 103, 124, 132, 148) (**Figure 1.4**). The potential input signals that regulate the enzymatic activities are likely to be diverse. For example, the PAS domains have been identified as sensors for light, redox-potential, and oxygen, often through binding cofactors such as flavin and heme (26, 56, 140). In *A. xylinum*, it has been shown that oxygen level regulates the PDE activity of protein AxPDEA1, through its heme-binding PAS domain (14). The GAF domain has been known for small ligand binding and protein-protein interactions. For example, in *Saccharomyces cerevisiae*, a phosphodiesterase binds the regulatory molecule cGMP with its N-terminal GAF domain (79). It is most likely that the GGDEF and EAL domains are modulated by GAF domains in bacteria given that the widespread occurrence of GAF-EAL and GAF-GGDEF proteins. Moreover, c-di-GMP synthesis or degradation is likely to couple with two-component signaling pathways considering that some GGDEF or EAL domains are conjugated with the CheY-like response receiver (REC) domains. A notable example is the REC-REC-GGDEF protein PleD from *C. crescentus*. Phosphorylation of its first REC domain induces dimerization and leads to the activation of the GGDEF domain (17, 92).

1.3 GGDEF-EAL didomain proteins

Although many proteins contain either the GGDEF or EAL domain, approximately one third of known GGDEF and EAL containing proteins contain both domains. In most cases, the tandem GGDEF-EAL didomains are located at the C-terminus with regulatory domains at the N-terminus (**Figure 1.4**) For example, the *P. aeruginosa* genome encodes 17 GGDEF, 5 EAL and 16 GGDEF-EAL domain proteins (35). Equally striking, there are 31 GGDEF, 22 EAL and 10 GGDEF-EAL domain proteins in *V. cholerae* (35). Theoretically, the GGDEF-EAL domain proteins could exhibit either DGCs or PDE activity, or they could possess both activities. As demonstrated in this thesis, there are also GGDEF-EAL domain proteins exhibit that neither DGC nor PDE activity.

1.3.1 Didomain proteins with enzymatically active GGDEF domains

We first discuss several GGDEF-EAL didomain proteins that only exhibit DGC activity. Three *cdg* operons were found to control the cellular turnover of c-di-GMP in *A. xylinum*. Each of the three operon is organized with a *pdeA* gene upstream of a *dgc* gene (126). A total of six isoenzymes encoded by these gene operons were identified. The isoenzymes AxDGC1-3 and AxPDE-A1-3 share similar domain organization as PAS-GGDEF-EAL (**Figure 1.4**). Genetic disruption analysis demonstrated that the three proteins AxDGC1-3 exhibit DGC activity *in vivo*, whereas AxPDE-A1-3 exhibit PDE-A activity (14, 95).

Their N-terminal PAS domains were first proposed to function as regulatory domains, which was confirmed for AxPDE-A1 later.

MorA, a trans-membrane protein, is found to regulate the timing of flagellar development and affects motility, chemotaxis, and biofilm formation among diverse *Pseudomonas* species (18, 83). Bioinformatic analysis shows that *morA* homologues are ubiquitous among all *Pseudomonas* genomes and that they share a high sequence similarity. The domain organizations of various MorA proteins are well conserved. The MorA family proteins are normally present as single copies in the genome and always contain a trans-membrane domain. In cytoplasm, there is a sensory domain consisting of PAS/PAC motif that is responsible for sensing environment cues such as light, redox potential, or oxygen, as well as a GGDEF-EAL didomain located at the C-terminus (**Figure 1.4**). Genetic evidence showed that disruption of *morA* enhanced swimming motility and chemotaxis in *Pseudomonas. putida*. It also restricts the timing of flagellar development since the deletion of *morA* is induced to form flagellar in all growth phase. In addition, the deletion of *morA* decreases the biofilm formation in *P. aeruginosa*. Taken together, the data suggested that MorA may act as a DGC protein in various *Pseudomonas* species regulating flagellar synthesis and biofilm formation (18).

1.3.2 Didomain proteins with enzymatically active EAL domains

Several GGDEF-EAL didomain proteins have been demonstrated to exhibit

PDE-A activity only. For instance, contrast to AxDGC1-3, one of the isoenzyme PDEA1-3 proteins A_xPDEA1 exhibits PDE-A activity and regulated by reversible binding of oxygen to the heme sensor bound by its PAS domain (**Figure 1.4**). Enzymatic activity assay reveals that the apo-A_xPDEA1 has less than 2% of the PDE-A activity of the holo-A_xPDEA1. Reconstitution of the apo-protein with heme restored the full activity. The results suggest that in A_xPDEA1 the heme-PAS functions as a regulatory domain to regulate the activity of the EAL domain. The oxy-form of the protein exhibits lower catalytic activity than that of the deoxy-protein. In brief, it is implicated that A_xPDEA1 is less active at high level oxygen condition and more active at low level oxygen environment. In addition, it suggests that c-di-GMP and cellulose synthesis in *A. xylinum* would increase under aerobic conditions and decrease during oxygen depletion. The heme-binding properties of A_xPDEA1 are significantly different from those of other oxygen induced heme sensors. For example, the rate of autoxidation of A_xPDEA1 (half-life > 12 h) is the lowest for this type of heme protein so far (14).

Figure 1.4 Domain organization of representative GGDEF and EAL proteins. GGDEF and EAL domains are colored in red and green, respectively. Regulatory domains are shown in blue and cyan. Predicted trans-membrane helices (TM) are indicated. Domains are not drawn to scale.

The second protein that exhibits PDE-A activity only is *EcDos* from *E. coli*, which possesses a PAS-PAS-GGDEF-EAL domain organization (**Figure 1.4**). The N-terminal PAS domain of *EcDos* was proposed to act as a direct O₂ or NO responsive heme-based sensor (39, 42). The GGDEF domain in *EcDos* does not exhibit DGC activity and so far the specific roles of the GGDEF domain in the catalysis and regulation of *EcDos* remain unclear. On the contrary, the EAL domain exhibits PDE-A activity by converting c-di-GMP into the linear dinucleotide 5'-pGpG. Shimizu and colleagues found that both the Fe(III) and Fe(II) heme bound *EcDos* proteins exhibit PDE-A activity toward c-di-GMP (27 and 61 min⁻¹, respectively) and that the activity of the Fe(II) protein form is obviously enhanced by the binding of either O₂ (126 min⁻¹) or CO (143 min⁻¹) (113). Mutagenesis coupled kinetic analysis suggest that a methionine residue, the Met⁹⁵ coordinated to the Fe(II) center plays a critical role in signal transduction. Conformational changes around Met⁹⁵ caused by the binding of external gas signals unlock the catalytic activity by activating the downstream EAL domain (130). Hence, *EcDos* appears to be a new heme-based gas sensor which response to more than one type of gaseous molecule.

The protein CC3396 is a GGDEF-EAL didomain protein from *C. crescentus* that exhibits PDE-A activity (**Figure 1.4**). The EAL domain convert c-di-GMP into the linear dinucleotide 5'-pGpG in the presence of Mg²⁺ ions but is strongly inhibited by the Ca²⁺ ions. The GGDEF domain contains a degenerate GEDEF motif and does not exhibit DGC activity. The degenerate

GGDEF domain specifically binds GTP (K_d of 4 μ M) and allosterically regulates the PDE-A activity by lowering the K_M for substrate c-di-GMP to a significant level. Mutagenesis analysis has shown that the mutant lacking the gene CC3396 reduced the cellular PDE activity more than 80%. In addition, enzymatic assays and UV cross-link experiments with purified full-length protein and single domain fragments reveal that the PDE-A activity is confined within the EAL domain of CC3396 (20). Meanwhile, although the degenerate GGDEF domain lacks catalytic activity, it retains the ability to bind GTP and function as a regulatory domain. It is further postulated that some GGDEF domains could act as regulatory domains for the EAL domain in a similar fashion.

The protein CdpA contains both GGDEF and EAL domains but only possesses PDE-A activity *in vitro*. The PDE-A activity is regulated by the enzymatically inactive GGDEF domain, which is similar to the CC3396 (**Figure 1.4**). The *cdpA* gene was first annotated as VC0130 from *V. cholerae*. The CdpA protein contains a tandem GGDEF-EAL domain, with a GVGEW replacing the GGDEF motif and an ECL instead of EAL motif. Despite a mutation in the EAL motif (ECL), CdpA exhibits PDE-A activity *in vitro*. The cellular c-di-GMP level of $\Delta cdpA$ mutant is four-fold higher than that of the parent strain. Overexpression of CdpA in the mutant led to undetectable levels of c-di-GMP. Together with previous enzymatic assay data, these results confirm CdpA as an active PDE-A protein. To test whether the degenerate

GGDEF domain (GVGEW) of CdpA play a similar regulatory role as in CC3396, mutagenesis and genetic experiments were conducted. Over-expression of a GGDEF domain mutant *cdpA* (AVGAW) enhances motility compared to the vector-alone strain, but less effective than the WT *cdpA*. In addition, the *cdpA* (Δ ECL) mutant had no effect on biofilm formation. Meanwhile, the over-expression of *cdpA* (AVGAW) decreased biofilm formation by four-fold, indicating that partial PDE-A activity remains in the *cdpA* (AVGAW) mutant. These results indicate that the GVGEW domain of CdpA is required for the full PDE-A activity of EAL domain (128).

The gene *bifA* (PA4367) from pathogen *P. aeruginosa* encodes the protein BifA of 688 amino acids with a MW of ~78kDa. The BifA protein is predicted to contain an N-terminal transmembrane domain, which anchors the protein to the inner membrane with the C-terminal GGDEF-EAL didomains resided in the cytoplasm (**Figure 1.4**). Deletion of the PA4367 gene results in a severe defect in swarming motility and a hyperbiofilm phenotype indicating that the gene *bifA* plays a regulatory role in biofilm formation. Biochemical evidence suggests that the protein possesses PDE activity *in vitro*, capable of cleaving both the synthetic substrate bis-*p*NPP and the native substrate c-di-GMP. Nevertheless, there is no detectable DGC activity of BifA presumably due to the altered signature motif in the GGDEF domain (GGDQF). Mutagenesis analysis shows that the Asp residue in the signature EAL motif is critical for the PDE-A activity. Consistent with these *in vitro* observations, *in vivo* studies

show that the $\Delta bifA$ mutant exhibits increased cellular concentration of c-di-GMP and increased synthesis of polysaccharide required for the hyperbiofilm formation than the wild type. However, expression of either the GGDEF motif mutant (GGDQF-AAAAA) or EAL motif mutant (EAL-AAL) proteins in the $\Delta bifA$ mutant strain failed to suppress the hyperbiofilm phenotype of the $\Delta bifA$ mutant. These results indicate that both the EAL domain and the GGDQF domain of BifA are required for proper function *in vivo* (66).

The protein PigX is the first characterized GGDEF-EAL didomain protein in *Serratia*, and also the first example of GGDEF/EAL domain proteins involved in the regulation of antibiotic production. PigX, identified in *Serratia* strain ATCC 39006, is predicted to contain two N-terminal transmembrane helices and a tandem GGDEF-EAL domain (**Figure 1.4**). Fineran and colleagues demonstrated that PigX plays a key role in regulating a conditional biosurfactant production and swarming phenotype in *Serratia* strain. Bioassay, gene fusion, primer extension, and proteomic experiments revealed that PigX controls the biosynthesis of the secondary metabolite: prodigiosin, by repressing the transcription of the pig biosynthetic gene operon (*pigA-pigO*). Bioinformatic analysis indicates that PigX contains a degenerate GGDEF (YHSDF) signature motif, indicating that the protein is unable to catalyze the synthesis of c-di-GMP. Moreover, PigX does not contain the RxxD motif conserved in most active DGCs for product inhibition. Analysis of EAL domain

of PigX demonstrates that the conserved EAL signature motif is replaced by ELI. Complementation experiments show that a triple alanine substitution mutant is not able to complement the *pigX* mutant for the Pig production. The EAL domain alone fully complements Pig biosynthesis in the *pigX* mutant strain back to wild type levels. Such genetic evidence suggests that PigX functions as a PDE to regulate the cellular c-di-GMP turnover and transcription of genes involved in the production of secondary metabolites (32).

The protein CdgC from *V. cholerae* was identified as a regulator of rugose colony development, biofilm formation, and flagellar motility (74). CdgC harbors both GGDEF and EAL domains, with the GGDEF domain lacking the conserved GG[D/E]EF signature motifs. It was predicted that this protein only functions as a phosphodiesterase (**Figure 1.4**). After arabinose induction, the overexpression of CdgC (*pcdgC*) exhibited a significant increase (42%) in motility compared to the control strain. Over-expression of a mutant CdgC with an altered EAL motif (*pcdgC*-AAL) did not cause any significant change in motility, indicating that the EAL domain of CdgC is active and is responsible for increased flagellar motility. It is also revealed that the GGDEF domain of *cdgC* is inactive, since it does not lead to any change in flagellar motility when the EAL domain is mutated. Together with the previous study that the overexpression of CdgC decreased the *vps* gene expression, the result suggests that CdgC acts as a phosphodiesterase. It was found that CdgC regulates *vps* gene expression in *V.cholerae* through a regulatory network composed of the

protein products of *vpsT*, *vpsR*, and *hapR* and other possible regulatory proteins that are required for polysaccharide synthesis, extracellular protein secretion, flagellar biosynthesis, and virulence factor expression. Mutagenesis analysis has shown that CdgC negatively regulates the expression of *vps* gene indirectly by repressing the expression of *vpsR* and *vpsT* and increasing the expression of *hapR*. Because VpsR and VpsT are required for the transcription of the *vps* gene, HapR, which is shown to positively regulate expression of flagellar biosynthesis genes, can negatively regulate *vps* gene expression through the suppression of *vpsR* and *vpsT* expression in the *cdgC* mutant, and results in a reduction in flagellar gene expression (75).

The protein PA2567 was first identified in the screening for genes that can restore the ability of swarming, virulence gene expression and preventing biofilm formation in several swarming-negative phenotype mutants in *P. aeruginosa*. Biochemical study showed that the protein PA2567 exhibited c-di-GMP specific phosphodiesterase-A activity *in vitro* (109). Sequence analysis of PA2567 indicates that the protein is comprised of an N-terminal GAF domain, and a tandem GGDEF-EAL domain at C-terminus. In contrast to the active EAL domain, the GGDEF signature motif of PA2567 is replaced by SPTRF and predicted to be a catalytically incompetent DGC (**Figure 1.4**).

The last characterized protein ScrG with 568 amino acids contains a PAS domain, a GGDEF domain and an EAL domain. It is the second GGDEF-EAL-encoding locus identified to be involved in the regulation of

swarming and sticking in *Vibrio parahaemolyticus*. Bioinformatic analysis predicts that the ScrG is an active PDE with the conserved EAL motif and an inactive GGDEF domain due to the altered GGDEF (HDDDF) signature motif (**Figure 1.4**). To confirm the enzymatic activity of ScrG, genetic disruption and mutation were used to observe the phenotypic differences. Over-expression of ScrG induced flagellar *laf:lux* gene expression in liquid and suppressed biofilm formation. Expression of a mutant form of the protein with the EAL domain truncated (ScrG_{ΔEAL}) abolished the activity of ScrG. An alanine substitution for the glutamate (E350A) in the EAL signature motif also completely disrupted the ability of ScrG to activate *laf* gene expression. It was further shown that over-expression of the full length ScrG protein decreased cellular c-di-GMP level, whereas over-expression of the truncated protein ScrG_{ΔEAL} raised the c-di-GMP level. Besides, the defected mutants of ScrG exhibited altered swarming, flagellar synthesis and colony morphology (63). Three alanine substitution mutations to the HDDDF motif (HDADF, HDDAF, HDDAA) were introduced into ScrG together with the truncated ScrG_{ΔEAL}. None of the mutants seemed to be critical for the activity associated with the GGDEF domain in the absence of the EAL domain. Sequence analysis suggested that ScrG possesses an enzymatically inactive diguanylate cyclase domain. Taken together, these data support the prediction that ScrG acts as phosphodiesterase in *V. parahaemolyticus*.

1.3.3 Didomain proteins with both enzymatically active GGDEF and EAL domains

Intriguingly, recent studies have revealed several bifunctional GGDEF-EAL didomain proteins with both DGC and PDE-A activities. BphG1, the first bifunctional GGDEF-EAL didomain protein identified in c-di-GMP signaling, is a photochromic protein from *Rhodobacter sphaeroides*. The 92 kDa protein harbors a PAS-GAF-PHY-GGDEF-EAL domain organization and the PAS domain binds biliverdin (BV) covalently as a light sensor (**Figure 1.4**). Sequence analysis suggests that both the GGDEF and EAL signature motif are highly conserved. However, the full-length BphG1 exhibits c-di-GMP specific PDE activity only *in vitro* with the PDE-A activity independent of light intensity. An approximately 31 kDa protein was found to be co-purified with BphG1 through the affinity chromatography. It turns out to be the entire EAL domain of BphG1, and this EAL domain fragment retains the PDE-A activity. A truncated BphG1 protein without the EAL domain was cloned and purified in the same condition as the full-length protein. Surprisingly, the *holo* form of the truncated BphG1 displays light-dependent DGC activity. These results suggest that BphG1 is a potentially bifunctional GGDEF-EAL domain protein with both c-di-GMP synthesis and hydrolysis activities. The catalytic activity can be switched from PDE to DGC by cleaving the C-terminal EAL domain (131).

ScrC is the first GGDEF-EAL didomain protein identified in *V. parahaemolyticus* to influence the cellular c-di-GMP concentration. The protein

regulates swarming motility, polysaccharide production, and flagellar synthesis gene expression. The *scrC* gene is part of the three-gene *scrABC* operon that encodes the ScrC protein. ScrC harbors two transmembrane domains at N-terminus (~300 amino acids in length) and a tandem GGDEF-EAL didomain at the cytoplasmic C-terminus (11)(**Figure 1.4**). Both the GGDEF and EAL domains of ScrC are highly conserved and were predicted to possess both the DGC and PDE activities. Similar to ScrG, deletion of the EAL domain of the ScrC, produced the similar phenotype as the deletion of the entire gene or operon, strongly suggesting that the dominant role of ScrC is to degrade c-di-GMP *in vivo*. In the absence of *scrA* and *scrB*, over-expression of *scrC* suppresses swarming gene expression and enhances polysaccharide gene *cps* transcription. However, overproduction of *scrC* in the presence of *scrA* and *scrB* induces swarming and prevents *cps* transcription. Thus, ScrC appears to be a diguanylate cyclase in the absence of ScrA and ScrB but a phosphodiesterase in the presence of these two partners. Genetic evidence further verified the catalytic activities for the GGDEF and EAL domains, respectively. The mutation of the EAL motif to AAL suppressed the ability of ScrABC to regulate the expression of the *laf* and *cps* gene. Interestingly, instead of inactivating the diguanylate cyclase activity, alanine substitution of Asp₄₂₈ in the GGDEF motif affects the putative phosphodiesterase activity, as ScrABC_{D428A} is not able to activate *laf* expression (31). Compared to the bifunctional GGDEF-EAL didomain protein BphG1 from *R. sphaeroides*, ScrC does not need to be

physically cleaved to exhibit DGC activity. The activity of ScrC appears to be regulated by other two proteins encoded by the same operon. The periplasmic protein ScrB resembles an extracellular binding protein. It may bind input signals and interact with ScrC. The role of the cytoplasmic protein ScrA in this signal transduction is still unclear. It is hypothesized that ScrC is associated with ScrA and ScrB and exhibits phosphodiesterase activity. Environmental cues could regulate the protein-protein interaction and switch the PDE activity to DGC activity.

The protein MSDGC-1 from *Mycobacterium smegmatis*, is a three-domain protein with the GGDEF-EAL didomains at the C-terminus and a GAF domain at the N-terminus (**Figure 1.4**). Kumar and colleagues cloned, expressed and purified this protein to demonstrate that it exhibits both DGC and PDE-A activity. The GGDEF domain converts the substrate GTP into c-di-GMP firstly, and then c-di-GMP would be degraded by the EAL domain into 5'-pGpG. Besides biochemical evidence, gene deletion of MSDGC-1 affects the long-term survival of *M. smegmatis* under the conditions of nutritional starvation. To understand the regulation of MSDGC-1, the individual GGDEF and EAL domain was cloned, expressed and purified separately. Intriguingly, both the stand-alone GGDEF and EAL domain are enzymatically inactive. Furthermore, addition of one domain to the other for prolonged incubation did not enhance any activity. The result indicates that both GGDEF and EAL domains of MSDGC-1 are required for the DGC and PDE-A activities as the

deletion of the single domain would abolish the catalytic activity (134).

1.3.4 Didomain proteins with enzymatically inactive GGDEF and EAL domains

In contrast to the bifunctional GGDEF-EAL domain proteins, several studies reveal that some didomain proteins exhibit neither DGC nor PDE activities. The *E.coli* protein YhdA, a regulator for the *csrB* and *csrC* gene expression belongs to a two-component signaling system. YhdA was predicted to be a membrane bound protein that contains a GGDEF-EAL didomain with highly degenerate GGDEF and EAL motifs (**Figure 1.4**) (125). Genetic evidence suggests that the YhdA affects the expression of a *csrC-lacZ* reporter. However, there is no direct evidence suggesting that YhdA binds to the *csrB* or *csrC* promoters. The membrane bound protein YhdA lacks the EAL motif and probably lacks c-di-GMP specific phosphodiesterase activity. The GGDEF signature motif is also very poorly conserved in YhdA (HRSDF). Deletion of the *Salmonella yhdA* gene (STM3375) had no apparent effect on the formation of extracellular polysaccharides and motility (120). This suggests that YhdA exhibits neither PDE-A nor DGC activity and may function as non-enzymatic domains. The presence of a putative 100 amino acid trans-membrane region and a putative coiled-coil linker domain suggests that the protein is likely to be regulated by external signals or membrane-bound factors.

FimX is a cytoplasmic protein (76 kDa) required for normal twitching

motility and biofilm formation in the opportunistic pathogen *P. aeruginosa* (55, 61, 67). FimX comprises four domains that include a CheY-like REC domain, a PAS domain, a GGDEF and an EAL domain (**Figure 1.4**). The functions of the REC and PAS domains remain unknown despite the overall sequence homology shared with other REC and PAS domains. The REC domain is unlikely to function as an archetypal phosphoreceiver domain considering it lacks the essential Asp residue required for phosphorylation, whereas the PAS domain could act as a sensor domain for an unidentified signal (55). FimX was first identified as a phosphodiesterase and is regulated by the degenerate GGDEF domain. However, recent solution and crystallographic studies from Sondermann and co-workers have demonstrated that the EAL domain of *P. aeruginosa* FimX is inactive and capable of binding c-di-GMP with high affinity (below 1 nM) (86). Farah and co-workers further reported that the EAL domain of the FimX homolog in *Xcc* not only binds c-di-GMP, but also interacts directly with the PilZ protein, a key protein in type IV pili synthesis (45).

1.4 Objectives and organization of this dissertation

The overall objective of this Ph.D. project is to elucidate the catalytic and regulatory mechanisms of GGDEF-EAL didomain-containing proteins in c-di-GMP signaling. As mentioned above, c-di-GMP signaling pathway plays vital and dynamical roles in dictating many aspects of bacterial behavior, and

the regulation is executed by the GGDEF/EAL domain-containing proteins in various cases. About one third of the GGDEF or EAL proteins contain both domains in tandem at the C-terminus and often present in conjunction with the sensor domains such as PAS, GAF, and REC. It is of great interest to understand how these didomain proteins are regulated by the sensor domains for controlling c-di-GMP concentration. We set out by selecting three representative GGDEF-EAL proteins and utilizing biochemical and biophysical approaches to probe the function and mechanism of the proteins. The three GGDEF-EAL didomain-containing proteins include PA2567, FimX from *P. aeruginosa* and AxDGC2 from *A. xylinum*.

At the beginning of the research project in 2006, the structural and catalytic mechanism of the GGDEF domain proteins were relatively well understood through the studies on PleD and WspR. In contrast, the lack of information on the structure and catalytic mechanism for the EAL domain presented an initial challenge for us. Together with the study of the EAL domain protein RocR in our lab, the didomain protein PA2567 was examined. Site-directed mutagenesis and kinetic measurements were performed to probe the catalytic mechanism. Study of PA2567 assisted the classification of catalytically active and inactive EAL domains. The elucidation of catalytic and regulatory mechanisms of EAL domains has contributed to the effort of deciphering the c-di-GMP signaling pathway.

The second protein AxDGC2 is one of the GGDEF-EAL proteins that

regulate cellulose synthesis in the obligate aerobe *A. xylinum*. It was predicted to contain a putative active GGDEF and inactive EAL domain. Previous sequence analysis suggested that the N-terminal sensory PAS domain could regulate the catalytic activity. Thus, it serves as a model system for the study of the regulation of GGDEF-EAL didomain protein activity by the sensory PAS domain.

The third protein FimX is a multi-domain signaling protein required for normal twitching motility, type IV pili synthesis and biofilm formation in the opportunistic pathogen *P. aeruginosa*. FimX localizes at one single pole in the bacterial cell and contains both highly degenerate GGDEF and EAL domains. The N-terminal REC and PAS domains were predicted to be sensor domains for perceiving input signal or interaction domains for interacting with other proteins. FimX provides us with a platform to study the function of a GGDEF-EAL didomain protein with putative non-enzymatic GGDEF and EAL domains.

In brief, we aim to examine the three GGDEF-EAL didomain-containing proteins with functionally diverse GGDEF and EAL domains. The detailed elucidation of catalytic and regulatory mechanism of the three didomain proteins would yield insights into the molecular mechanism of the c-di-GMP signaling network. Towards this goal, this dissertation is divided into the following five chapters.

Chapter 1 reviews the roles of c-di-GMP signaling in bacteria and the

regulation of c-di-GMP by the GGDEF and EAL domain proteins. Some of the characterized GGDEF-EAL didomain proteins are briefly introduced with their domain organization, enzymatic activity, and associated phenotypes.

Chapter 2 presents the study of the first didomain-containing protein PA2567. The results help to establish the catalytic mechanism of the EAL domains. Elucidation of the catalytic mechanism enabled us to classify EAL domains according to enzymatic activity.

Chapter 3 discusses AxDGC2. The full-length protein exhibits only DGC activity and is regulated by the redox status of the FAD-bound PAS domain. Biochemical and mutagenesis evidence provide interesting insights into the regulatory mechanism as well as the regulation of c-di-GMP level and cellulose synthesis in the obligate aerobe *A. xylinum*.

Chapter 4 is devoted to FimX. The EAL domain of FimX was found to function as a c-di-GMP binding receptor rather than c-di-GMP degradation domain. Hydrogen-deuterium (H/D) exchange study by mass spectrometry reveals that the binding of c-di-GMP to the EAL domain triggers a conformational change in the REC domain and adjacent linker.

Chapter 5 concludes these studies and suggests future work.

CHAPTER 2 CATALYTIC MECHANISM OF THE EAL DOMAIN OF PA2567

2.1 Introduction

EAL domain proteins are the major phosphodiesterases for maintaining the cellular concentration of c-di-GMP in bacteria. Given the pivotal roles of EAL domains in the regulation of virulence secretion and biofilm formation in pathogenic bacteria (53, 129), the elucidation of the catalytic and regulatory mechanisms would assist the development of antibacterial agents and contribute to the effort of deciphering the cyclic-di-GMP signaling pathway.

Previous studies have indicated the existence of at least two classes of EAL domains, one class with apparent c-di-GMP specific PDE-A activity and the other class without apparent catalytic activity. The genomes of bacteria that contain the c-di-GMP signaling network encode multiple copies of EAL domain proteins, with many of them associated with an adjacent GGDEF domain. The utilization of the GGDEF domain could be a major strategy for regulating the catalytic activity of the EAL domains. For example, the PDE-A activity of CC3396 from *C. crescentus* can be stimulated by the binding of GTP to its degenerated GGDEF domain (20). Meanwhile, the putative inactive EAL domains occurred in the GGDEF-EAL proteins that with active GGDEF domains raise the possibility that some EAL domains may function as regulatory rather than catalytic domains (88, 138).

Although the detailed catalytic mechanism of the EAL domains was

unknown at the start of the project, the crystal structures of two EAL domain proteins TdEAL and YkuI had been determined (PDB codes: 2R6O, 2BAS) (84, 133) by two structural genomic groups. EAL domains adopt a $(\beta/\alpha)_8$ barrel fold that contains two extended strands, including an anti-parallel strand. From the previous biochemical studies of EAL domain proteins, it was known that Mg^{2+} or Mn^{2+} ion are required for hydrolyzing cyclic-di-GMP (115, 129). It was also found that Zn^{2+} and Ca^{2+} ion strongly inhibited the enzymatic activity, presumably by dislodging the Mg^{2+} ion. The Glu in the EAL (or EXL) signature motif seems to be essential for the enzymatic activity, considering that the Glu to Ala mutants in two EAL domain proteins abolished their phosphodiesterase activities (20, 129). In addition, Schmidt and coworkers suggested that other conserved motifs, including a DDFGTG motif, might be essential for the catalytic activity (115).

The single EAL domain-containing protein RocR was used as a model system for studying the catalytic mechanism by our group. In total, the functions of 14 conserved polar residues in RocR were assigned and a general base-catalyzed mechanism with the assistance of the Mg^{2+} ion was proposed (99). Additionally, the conserved DFG(T/A)GYSS motif forming loop 6 in RocR, was proposed to not only mediate the dimerization of EAL domain to stabilize the protein quaternary structures but also control the binding of c-di-GMP and Mg^{2+} ion (97).

To further examine whether the proposed catalytic mechanism also applies

to the EAL domain in the GGDEF-EAL context, my study focused on the didomain-containing protein PA2567. The *PA2567* gene was first identified in the screening for genes that can restore the ability of swarming, virulence gene expression and prevent biofilm formation in several swarming-negative mutants in *P. aeruginosa* (109). Biochemical study showed that the protein PA2567 exhibited c-di-GMP specific phosphodiesterase-A activity *in vitro*. Sequence analysis of PA2567 indicates that the protein is comprised of an N-terminal GAF domain, and a tandem GGDEF-EAL domain at the C-terminus. In contrast to the active EAL domain, the GGDEF signature motif of PA2567 is replaced by SPTRF and predicted to be catalytically incompetent in synthesizing c-di-GMP (**Figure 2.1**).

One Glu (Glu³⁵²) and a functional loop 6 in RocR have been proposed to play a critical role in catalysis. To test whether the Glu residue and loop 6 are only limited to RocR and whether there is a difference in catalytic mechanism between the single EAL domain-containing proteins and didomain-containing proteins, site-directed mutagenesis and kinetic measurement were performed with PA2567.

Figure 2.1 Domain organization of PA2567 protein.

2.2 Methods and Materials

2.2.1 Gene cloning and site-directed mutagenesis

The genomic DNA of *P. aeruginosa* PAO-1 (ATCC) was isolated following standard procedures. The gene encoding PA2567 from *P. aeruginosa* was amplified by PCR using the Expand High-Fidelity Kit (Roche). The amplified DNA fragment was cloned into the pET-46 Ek/LIC vector (Novagen). Site-directed mutagenesis was carried out using the Quick Change mutagenesis kit (Stratagene) following manufacturer's instructions. The correct sequences of the cloned and mutated genes were confirmed by the complete nucleotide sequencing (1st Base). All the primers used for PCR and mutagenesis are listed in Table 2.1.

Table 2.1 Primers used in PA2567 cloning and site-directed mutagenesis

Primer	Relevant sequence
PA2567Forward	5'-GACGACGACAAGATGGCAACCCACCC-3'
PA2567Reverse	5'-GAGGAGAAGCCCGGTAAACGCCGCAGCCAG-3'
PA2567 _{E464A} Forward	5'-GAGCTGGAATTCACCGCAAGCGTACTCATCCGCGAC-3'
PA2567 _{E464A} Reverse	5'-GTCGCGGATGAGTACGCTTGCGGTGAATTCCAGCTC-3'
PA2567 _{E548Q} Forward	5'-TATCGGGTGGTCGCGCAAGGCATAGAGACCCACGACAC-3'
PA2567 _{E548Q} Reverse	5'-GTGTCGTGGGTCTCTATGCCTTGCGCGACCAACCCGATA-3'
PA2567 _{S493A} Forward	5'-ATCGCGGTGGACGACGCCGGCACCGGCTACAGCAA-3'
PA2567 _{S493A} Reverse	5'-TTGCTGTAGCCGGTGCCGGCGTCGTCCACCGCGAT-3'
PA2567 _{F498A} Forward	5'-CTTCGGCACCGGCTACGCCAACTGGACCTACCTG-3'
PA2567 _{F498A} Reverse	5'-CAGGTAGGTCCAGTTGGCGTAGCCGGTGCCGAAG-3'

2.2.2 Protein expression and purification

The plasmids harboring the genes were transformed into *E. coli* strain

BL21(DE3) (Novagen). The cells were grown in Luria-Bertani (LB) media at 37 °C with vigorous shaking (220 *rpm*) until the OD_{600nm} reached 0.6-0.8. Isopropyl-β-D-thiogalactopyranoside (IPTG, 0.4 mM) was added to induce protein expression and the culture was grown for additional 16 hours at 16 °C. Cells were harvested by centrifugation for 10 min at 5000 *rpm* at 4 °C. The cell pellet was frozen and thawed before the cells were lysed by sonication in 40 ml of lysis buffer (50 mM NaH₂PO₄, pH 7.0, 5 mM β-ME, 20 mM imidazole, 0.2 mM PMSF). The cell extract was centrifuged at 18,000 *rpm* for 30 min. All the purification steps described below were performed at 4 °C. The supernatant was filtered and loaded onto 1 ml of Ni²⁺-NTA resin (GE healthcare) that had been pre-packed into a column. The flow-through was collected and passed through the column again. The column was washed with 50 ml of washing buffer (lysis buffer supplemented with 50 mM imidazole). The proteins were eluted by using the elution buffer (lysis buffer supplemented with 300 mM imidazole). After SDS-PAGE gel analysis, fractions with purity higher than 95% were pooled together and desalted by using either a PD-10 desalting column (GE-Healthcare) or gel filtration column Superdex 200 (GE-Healthcare) with the AKTA FPLC system. The desalting buffer for PD-10 and gel filtration is 50 mM Tris-HCl (pH 8.0), 100 mM NaCl, 5% glycerol, and 1 mM DTT. The molecular weight of the proteins was estimated based on the standard curve generated by using the standard proteins that included ferritin (440 kDa), aldolase (158 kDa), conalbumin (75 kDa), ovalbumin (44 kDa) and blue dextran (for void volume

determination). The bright-yellow looking proteins were concentrated using Amicon concentrator (Millipore) and were stored at -80 °C after the measurement of protein concentration by Bradford assay.

2.2.3 Steady-state enzymatic activity assay

The substrate c-di-GMP used for kinetic measurement was produced enzymatically by using an engineered thermophilic DGC protein (96). The product 5'-pGpG was monitored by using the Agilent LC1200 HPLC system (Mobile phase: 20 mM Triethylammonium bicarbonate (pH 7.0), 9% Methanol, 1 ml/min) with an XDB-C18 column (4.6×150 mm) and confirmed by MALDI mass spectrometry with a molecular weight of 709.108 (calculated M.W. 709.11). The enzymatic reaction was performed by incubating the enzyme and c-di-GMP at 23 °C in 100 mM Tris-Cl (pH 8.0), 50 mM KCl and 10 mM MgCl₂. The reaction was stopped by adding 1/20 reaction volume of 1 M CaCl₂ and then boiled at 95 °C for 5 min. Following centrifugation at 14,000 *rpm* for 5 min to remove protein aggregate, the supernatant was filtered and loaded onto the HPLC. Initial velocity at a certain substrate concentration was obtained from a series of reactions with varying incubation time. The total turnover was kept to ensure the accurate measurement of initial velocities within the linear range. Initial velocity was measured at 6-12 substrate concentrations. The turnover number (k_{cat}) and the Michaelis-Menten constant (K_M) were obtained by fitting the initial velocities at various substrate concentrations to the

Michaelis-Menten equation with the exception of the protein that exhibits substrate inhibition using the software Prism (Graphpad). For wild type PA2567, kinetic data were fit using Michaelis-Menten equation. However, for the mutant that exhibits substrate inhibition, the kinetic data were fit using a model that assumes the substrate binds to the enzyme at a productive and a non-productive (or inhibitory) binding site (**Scheme 2.1**).

The curves were obtained by fitting the kinetic data to Eq. 2.1 derived from the above model with steady-state and rapid equilibrium assumptions. In the equation, K_i is the inhibition constant, α and β are the factors by which the K_M , K_i and V_{max} change when the second substrate is bound at the non-productive site.

$$v = \frac{V_{max} (1/K_m + \beta[S]/\alpha K_i K_m)}{1/[S] + 1/K_m + 1/K_i + [S]/\alpha K_m K_i} \quad \text{Eq. 2.1}$$

2.2.4 Bioinformatic analysis and structural modeling

The sequence alignments were conducted by the ClustalW and Esprict Server. (<http://align.genome.jp/>; <http://esprict.ibcp.fr/ESPrict/ESPrict/>). The structural model of the EAL domain was constructed with the coordinates of the

TdEAL structure (PDB code: 2R6O) as the template in Swiss-Model Server (116).

2.3 Results

2.3.1 Purification and characterization of PA2567

The full-length PA2567 with an N-terminal (His)₆-tag was purified by Ni²⁺-NTA-sepharose chromatography from the soluble extract of *E.coli* BL21(DE3) harboring pET-46-PA2567 plasmid. Immediately after elution, the protein was applied to a gel-filtration column to remove imidazole. The protein was eluted with a predicted mass around 140 kDa indicating a dimeric protein (**Figure. 2.2B**). The typical yield of the purified PA2567 was ~10 mg per liter cell culture. SDS-PAGE showed that the protein is ~95% homogeneous and the molecular mass is in agreement with that predict from the sequence (66kDa including the (His)₆-tag at **Figure. 2.2A**). The purified PA2567 was stored at -80 °C and found to be catalytically active for several months.

The enzymatic activity was assayed by incubating the enzyme with c-di-GMP and Mg²⁺. Product analysis was carried out by the Reverse-phased HPLC system equipped with an XDB-C18 column. The retention time of reaction product on the HPLC column was identical to that of 5'-pGpG, generated from the hydrolysis of c-di-GMP by RocR (99) (**Figure 2.3A**). The identity of 5'-pGpG was further confirmed by MALDI mass spectrometry with a molecular weight of 709.108 (calculated M.W. 709.11). (**Figure 2.3B**)

Figure 2.2 SDS-PAGE and gel filtration characterization of PA2567. (A) SDS-PAGE analysis of the purified PA2567. Lane 1: protein ladder. Lane 2: PA2567. (B) Gel filtration chromatogram of PA2567. Proteins were loaded and separated using 50 mM Tris-HCl (pH 8.0), 100 mM NaCl, 5% glycerol and 1 mM DTT.

2.3.2 Site-directed mutagenesis study

The EAL domain of PA2567 was found to be catalytically active, with a k_{cat} of $0.39 \pm 0.03\text{s}^{-1}$ (**Table 2.2**). Mutation of the residue Glu⁵⁴⁸ (corresponding to Glu³⁵² in RocR) to Gln reduced k_{cat} to an undetectable level, consistent with the observation for RocR and supporting the critical role of Glu⁵⁴⁸ in catalysis (94).

The conserved loop 6 that contains a DFG(T/A)GYSS motif was found in most characterized EAL domains with enzymatic activity (108). Previously, it was shown that the mutation in the highly conserved residue Glu²⁶⁸ in RocR that stabilizes loop 6 significantly alter the catalytic properties of RocR. Mutations of two residues on loop 6, Phe²⁹⁷ and Ser³⁰², affects the oligomeric state of the protein, indicating that loop 6 is also crucial for maintaining the quaternary structure of RocR (92).

Figure 2.3 C-di-GMP hydrolysis by PA2567. (A) HPLC trace showing the c-di-GMP hydrolysis by PA2567 (solid line) and RocR (dashed line). (B) The MALDI-MS result of product 5'-pGpG.

Surprisingly, the single mutations of Phe⁴⁹³ and Ser⁴⁹⁸ in PA2567, which are equivalents of Phe²⁹⁷ and Ser³⁰² in RocR, caused the protein to form inclusion bodies during protein expression. Meanwhile, the mutation of Glu⁴⁶⁴, the equivalent of Glu²⁶⁸ in RocR, caused strong substrate inhibition at high substrate concentration, with an inhibition constant of $30.4 \pm 12.0 \mu\text{M}$ (**Figure 2.4 and Table 2.2**). The mutation E464A did not perturb the oligomeric state of

the dimeric PA2567 according to size exclusion chromatography (**Figure 2.2B**). The results of PA2567 provided further support for the conclusion from RocR study that the loop 6 plays critical roles in stabilizing the overall protein structure and participating in catalysis, probably through mediating substrate binding.

Table 2.2 Steady-state kinetic parameters for PA2567 and its mutants ^a

PA2567 enzyme	k_{cat} (s ⁻¹)	K_M (μM)	k_{cat}/K_M (s ⁻¹ μM ⁻¹)	K_i (μM)
WT	0.39 ± 0.03	5.2 ± 1.3	$(7.5 ± 2.0) × 10^{-2}$	
E548Q mutant	ND ^b	ND	ND	
E464A mutant	1.1 ± 0.5	13.0 ± 5.0	$(8.4 ± 5.0) × 10^{-2}$	30.4 ± 12.0
S493A mutant	ND	ND	ND	
F498A mutant	ND	ND	ND	

^a Conditions for PA2567 and its mutant were 100 mM Tris-Cl (pH8.0) (23 °C), 50 mM KCl, and 10 mM MgCl₂.

^b ND, not determined due to inactivity caused by protein instability or extremely low activity (<10⁵-fold less active than wild-type PA2567).

2.4 Discussion

2.4.1 PA2567 supports the proposed EAL domain catalytic mechanism

In previous study of RocR, site-directed mutagenesis was conducted on 14 highly conserved polar residues and 7 of them led to a >10⁵-fold decrease in the turnover number. 4 of these 7 residues were found to be indispensable for Mg²⁺

coordination (**Figure 2.5A**). A one-metal-ion catalytic mechanism was proposed in RocR by our lab. The Glu³⁵² was hypothesized to be a general base catalyst (99). However, recent observation of two Mg²⁺ ions in the crystal structure of TdEAL and BlrP1 raised the possibility of a two-metal-ion catalytic mechanism with the Glu³⁵² coordinated by a second Mg²⁺ ion (5) (**Figure 2.5B**). In RocR, TdEAL and BlrP1, mutation of this Glu rendered all of the three proteins completely inactive, confirming the critical role of the residue in catalysis. Since this Glu residue plays essential roles in both the one metal and two metal mechanisms, the release of the EAL domain crystal structure of RocR and PA2567 in future might help us solve the controversial problem.

Figure 2.4 Effect of the E464A mutation on the steady-state kinetics of PA2567. The curves were generated by fitting the kinetic data of the wild-type and mutant enzyme to the Michaelis-Menten equation and equation 2.1, respectively.

Study of various $(\beta/\alpha)_8$ barrel fold enzymes suggested that a functional loop (loop 6) seems to be preserved during evolution in $(\beta/\alpha)_8$ barrel fold proteins (121). The results of previous studies on RocR protein suggest an equally important role for loop 6 in the hydrolysis of c-di-GMP for the $(\beta/\alpha)_8$ barrel fold EAL domains. The single mutations in the loop 6 region exerted profound impact on the structure and enzymatic activity of RocR. The mutation of the loop residue Phe²⁹⁷ and the loop-stabilizing Glu²⁶⁸ caused the formation of high molecular weight oligomer (HMWO) and significantly changed both k_{cat} and K_M . The change in oligomeric state is likely to be a result of disruption of the dimer interface, as inferred from the observation that loop 6 is located at the dimer interface in the crystal structure of TdEAL and EAL_{YkuL}. The observed antiparallel loop-loop interaction between the residues from loop 6 may stabilize the dimeric structure (84). It is conceivable that the mutation of Glu²⁶⁸ or the residues on the loop may disrupt the loop-loop interaction and result in the dissociation of the dimer and the formation of HMWO. On the other hand, the mutation of Asp²⁹⁶ and Ser³⁰², two residues from loop 6 in RocR, led to significant changes in catalytic parameters without altering the oligomeric structure (97) (**Figure 2.6**). These data indicate that the loop 6 plays a critical structural role in maintaining the quaternary structure of the protein.

Similar results were obtained with the GGDEF-EAL didomain-containing protein PA2567 in our study. Although the mutation E464A (E268A in RocR) did not totally abolish the PDE activity of PA2567 and the k_{cat} is only two-fold

less than that of the wild type PA2567, the mutant exhibits substrate inhibition at high c-di-GMP level and an increased K_M in contrast to the wild type PA2567. The observed substrate inhibition and greater K_M indicated that the perturbation of the loop could lead to the alteration of the binding affinity for c-di-GMP. The single mutations of Phe⁴⁹³ and Ser⁴⁹⁸ residues located on loop 6 in PA2567, which are equivalents of Phe²⁹⁷ and Ser³⁰² in RocR caused the protein to form inclusion bodies during protein expression. The formation of inclusion body indicates that protein stability has been altered. Together, the mutations of the residues on loop 6 once again confirm the catalytic and structural roles played by the loop. We proposed that loop 6 is crucial not only for stabilizing the dimer interface to stabilize the quaternary structure of RocR and PA2567 but also for the binding of Mg²⁺ ions and c-di-GMP. The conformational change in loop 6 caused by mutation or regulatory signal will impact catalysis in one of the following ways. First, the conformational change in the loop can affect the dimerization of the protein, which would result in a change of oligomeric structure and even protein stability. Second, binding of c-di-GMP can be affected as a result of the alternation of the interaction between c-di-GMP and the loop. Mutations that affect the interactions between loop 6 and substrate could impede the ability of loop 6 to undergo conformational change for effective substrate binding would affect catalysis. Third, change in the loop conformation can severely hamper catalysis by dislocating the essential Mg²⁺ ion-binding residue. Given the sensitivity of the catalytic activity to the loop

conformation and the position of the Mg^{2+} -binding residue, we postulate that some EAL domains could be regulated through controlling the conformation of loop 6.

Figure 2.5 Detailed views of EAL domain catalytic active sites, the residue numbers for the protein RocR and BlrP1 are shown, with the corresponding residue numbers for EAL_{PA2567} in parentheses. (A) The one-metal-ion catalytic mechanism proposed in RocR. The four residues that coordinate Mg^{2+} are colored in blue. The two residues that we also mutate in PA2567 are colored in red, including the general base catalyst. (B) Two-metal-ion catalytic mechanism proposed in BlrP1 (5). The manganese ions (violet) are coordinated by several carboxylic acid residues including Glu³⁵⁹.

Unlike the EAL domain protein RocR, which is adjacent to the regulatory REC domain, there is a degenerate GGDEF domain between the putative regulatory domain and EAL domain in PA2567. Due to the lack of residues for GTP binding, this GGDEF domain is unlikely to function as the one in CC3396 for allosteric control of PDE activity (20). We have experimentally confirmed that the GGDEF domain does not bind GTP (data not shown). The function of the degenerate GGDEF domain in PA2567 and many other GGDEF-EAL proteins remain mysterious.

Figure 2.6 Structural model of EAL_{RocR} with the residue of loop 6 and Glu²⁶⁸ highlighted, along with the corresponding residue numbers for EAL_{PA2567} in parentheses. The hydrogen bonds formed between Glu²⁶⁸ and the loop residues Gly³⁰⁰ and Ser³⁰² are presented by the broken lines. The Mg²⁺ ion is shown as the ball.

2.4.3 Classification of EAL domains

Based on the hypothesized catalytic mechanism for RocR and PA2567, we attempted to rationalize the lack of enzymatic activity for some of the characterized EAL domain proteins. We first examined the sequences of 24 EAL domains with catalytic activity characterized by either *in vivo* or *in vitro* assays. These domains are grouped into catalytically active and inactive EAL domains based on the conservation of critical catalytic residues and loop 6 (**Figure 2.7A**). The lack of the enzymatic activity for 12 inactive EAL domains can be easily explained by the absence of one or more catalytic residues. For example, FimX is catalytically inactive because it lacks three residues for coordinating the Mg²⁺ ion.

By examining the protein sequences of the 5,862 EAL domains encoded by the bacterial genomes, we further categorize the EAL domains in the database into three classes. The class 1 EAL domains possess conserved catalytic residues and a conserved loop 6 and function as active PDE-As (**Figure 2.7A, B**). It is revealed that 4,809 of the 5,862 EAL domains contain the catalytic residues that include the general base catalyst and four residues for Mg^{2+} coordination. Among the 4,809 EAL domains, 2,895 of them contain a conserved loop 6 [DFG(A/S/T)(G/A)(Y/F)(S/A/T)(S/A/G/V/T)] and 1,914 of them contain a degenerate loop 6 (**Figure 2.7B**). The 2,895 EAL domains are likely to be catalytically active PDE domains that belong to Class 1, whereas the 1,914 EAL domains with a degenerate loop 6 are categorized as class 2 EAL domains, as exemplified by the YkuI and AxDGC2. The EAL domains of YkuI and AxDGC2 were found to be catalytically inactive toward c-di-GMP by *in vitro* enzymatic assay (84, 95). However, it remains to be seen whether the EAL domain can be "activated" by the putative regulatory domains in YkuI and AxDGC2. Thus, Class 2 EAL domains contain conserved catalytic residues and a degenerated loop 6. In addition to the 4,809 EAL domains, 1,053 EAL domains were found to lack one or more of the essential catalytic residues and are categorized as Class 3 (**Figure 2.7B**). Among the 1,053 EAL domains belonged to Class 3, the majority of them (990) contain a degenerate loop 6 (**Figure 2.7A, B**). The lack of enzymatic activity for Class 3 proteins is supported by previous characterization of inactive EAL domains as well as

three recently reported ones. For instance, the EAL domain of the protein LapD from *P. fluorescens* Pf0-1 contains a highly degenerate loop 6 and lacks an essential residue for Mg^{2+} binding. It was proposed that LapD functions as a c-di-GMP-binding receptor rather than a PDE domain (88). Similar to LapD, the GGDEF-EAL didomain protein FimX from *P. aeruginosa* is also found to bind c-di-GMP via its EAL domain instead of hydrolyzing it due to the altered loop 6 and the absence of several essential residues including two for Mg^{2+} binding and the general base catalyst (86). Moreover, the *E. coli* protein YcgF contains an EAL domain that was initially assumed to be a c-di-GMP-hydrolyzing domain. However, the EAL domain contains a Met residue at the position of the general base catalyst and a highly degenerate loop 6. Recent studies showed that the EAL domain does not exhibit PDE activity but participates in protein-protein interactions instead (138).

47

(essential residues) asterisks. The loop 6 is underlined. The EAL domains shown are from the following proteins: PA2567, BifA, and FimX from *Pseudomonas aeruginosa* (indicated by the Pa suffix after the hyphen and protein) (55, 66, 109); TdEAL from *Thiobacillus denitrificans* (TdEAL-Td); VieA from *Vibrio cholera* (VieA-Vc) (129); CC3396 from *Crescentus caulobacter* (CC3396-Cc) (20); YcgF, YahA, Dos, YciR, CsrD, and YegE from *Escherichia coli* (indicated by the Ec suffix after the hyphen and protein) (40, 115, 125, 138, 146); BphG from *Rhodobacter sphaeroides* (BphG-Rs) (131); PdeA1, DGC1, DGC2, and DGC3 from *Acetobacter xylinum* (indicated by the Ax suffix after the hyphen and protein) (14, 95, 126); HmsP from *Yersinia pestis* (Hmsp-Yp) (9); GcpC from *Salmonella enterica* (GcpC-Se) (36); STM1344 and STM3375 from *Salmonella Typhimurium* (indicated by the St suffix after the hyphen and protein) (117); and LapD from *Pseudomonas fluorescens Pf0-1* (LapD-Pf) (88). The sequences were aligned using MultAlin (<http://bioinfo.genopole-toulouse.prd.fr/multalin/multalin.html>), and the figure was generated using ESPript 2.2. (B) Pie chart summary of the classification of the 5,862 EAL domains from bacterial genomes according to the conservation of catalytic residues and loop 6.

2.5 Summary

EAL domain-based cyclic di-GMP (c-di-GMP)-specific phosphodiesterases play important roles in bacteria by regulating the cellular concentration of the dinucleotide messenger c-di-GMP. To elucidate the catalytic mechanism of EAL domain proteins, the *P. aeruginosa* protein PA2567 with an N-terminal GAF domain and the GGDEF-EAL didomain was cloned, expressed and purified to homogeneity. Enzymatic assays revealed that PA2567 contains a catalytically competent EAL domain but an inactive GGDEF domain. Site-directed mutagenesis in combination with kinetic studies suggested that the EAL domains hydrolyze the phosphodiester bond of c-di-GMP using a one metal-ion assisted mechanism. In addition to the five essential residues

involved in Mg^{2+} binding, the essential residue Glu⁵⁴⁸ (in PA2567) or Glu³⁵² (in RocR) could function as a general base catalyst assisting the deprotonation of Mg^{2+} -coordinated water to generate the nucleophilic hydroxide ion in the one-metal mechanism, or coordinate a second Mg^{2+} according to the two-metal mechanism proposed by Barends TR and colleagues. We also proposed that the conserved loop 6 in EAL domain not only mediates the dimerization of the EAL domain to stabilize the quaternary structure of RocR and PA2567 but also controls c-di-GMP and Mg^{2+} ion binding. Together with the previous kinetic study of RocR, these results are supportive of the Mg^{2+} or Mn^{2+} ion dependent catalytic mechanism proposed for EAL domains. Importantly, based on the sequence analysis of the 5,862 EAL domains in the bacterial genomes, EAL domains can be categorized into three classes. We hope that the functional classification of EAL domains described here can provide some assistance in elucidating the evolution and biological functions of the versatile EAL domains.

CHAPTER 3 REGULATION OF C-DI-GMP SYNTHESIS IN AXDGC2

3.1 Introduction

The cyclic dinucleotide c-di-GMP has emerged as a major signaling messenger in bacteria for regulating a variety of cellular functions and behaviors (50, 105). The cellular concentration of c-di-GMP is maintained by the proteins that contain GGDEF, EAL or HD-GYP domain (114). The GGDEF domains named after the conserved GGDEF motif function as diguanylate cyclases (DGC) for c-di-GMP synthesis, whereas the EAL domains and HD-GYP domains characterized by the signature EAL or HD/GYP motifs function as phosphodiesterases (PDE) for c-di-GMP hydrolysis (50, 106, 118). Many GGDEF, EAL and HD-GYP domains are associated with putative sensor domains for perceiving various environmental signals (34-35). The characterization of the sensor domains and associated signals represents one of the major challenges towards understanding c-di-GMP signaling. Considering that c-di-GMP controls such phenotypes as virulence expression and biofilm formation, the identification of the signals that regulate the cellular c-di-GMP level may also be crucial for understanding host-pathogen interaction and bacterial pathogenicity.

A large number of the putative sensor domains associated with GGDEF, EAL and HD-GYP domains are the ubiquitous Per-Arnt-Sim (PAS) domains found in all kingdoms of life (49, 132). PAS domains are best known as sensor

domains for perceiving changes in oxygen concentration, light intensity, voltage and redox potential (38, 132). Several heme-containing PAS domains sense gaseous ligands including O₂, CO and NO have been characterized in details (38, 132, 139, 144). For example, the heme-containing PAS domains of AxPDEA1 and EcDOS regulate the activity of the EAL domain by reversible binding to O₂ (14, 130). The structural change in the PAS domain induced by O₂ binding can be transmitted to modulate the catalytic activity of the output domain.

In addition to heme-binding PAS domains, two classes of flavin-binding PAS domains have been under intensive investigation in recent years. The first class, as represented by the BLUF and LOV photoreceptors, uses the flavin cofactor as chromophore for sensing blue light (2, 5, 22, 41, 85). The isoalloxazine moiety of the flavin responds to photon activation by triggering structural change in the protein scaffold. The molecular mechanisms for the light-driven structural change varies among the PAS domains, ranging from reversible formation of covalent adduct between the flavin and a cysteine in LOV domains to reorganization of the hydrogen-bond network in BLUF domains (22, 41). The second class of flavin-binding PAS domains perceives the change of oxygen tension or redox status in the surroundings, as exemplified by the PAS domains of *Azotobacter vinelandii* and *Klebsiella pneumonia* NifL, *E. coli* Aer and putatively *M. capsulatus* MmoS (6, 52, 142). Biochemical and structural studies on Aer and NifL have suggested that the

change of the oxidation state of FAD would induce structural changes in the PAS domain and output domain (52, 100, 119, 141). It has also been demonstrated that the FAD cofactor of *KpNifL* can exchange electron with menaquinone in *K. pneumonia*, suggesting that the menaquinone pool could be the physiological signal perceived by the PAS domain of *KpNifL* (135). Since the reduced FADH₂ state can be rapidly oxidized by O₂, the PAS_{Aer} and PAS_{NifL} domains were also suggested to function as indirect oxygen sensors. The observation of a putative oxygen channel in the crystal structure of PAS_{NifL} has provided further support for oxygen sensing (62).

AxDGC2 is one of the proteins that regulate cellulose synthesis in the obligate aerobe *A. xylinum* (126). Although AxDGC2 shares similar domain organization with two previously characterized heme-binding proteins AxPDEA1 and EcDOS (14, 42) (**Figure 3.1**), we found that the PAS domain of AxDGC2 binds FAD as cofactor instead, and that the redox state of the flavin regulates the catalytic activity of the adjacent GGDEF domain for synthesizing c-di-GMP. Mutagenesis studies and kinetic measurement were carried out to probe the roles of several residues in cofactor binding and signal transduction. The results provided direct evidence for the role of the flavin-containing PAS domain as redox/oxygen sensor. Together with the observed regulation of AxPDEA1 by O₂ (14), the results also underscore the regulation of c-di-GMP concentration and cellulose synthesis by using both heme and flavin-containing PAS domains for O₂ response in the obligate aerobe *A. xylinum*.

Figure 3.1 Comparison of the domain organization of the three PAS domain-containing proteins in c-di-GMP signaling (AxDGC2, AxPDEA1 from *Acetobacter xylinum*; EcDOS from *Escherichia coli*).

3.2 Methods and Materials

3.2.1 Gene cloning and site-directed mutagenesis

The gene encoding AxDGC2 from *A. xylinum* was obtained from Genscript (NJ, USA) and was ligated into the pET-26(b+) vector (Novagen) between the NdeI and XhoI restriction sites. The resulted plasmid was used as a template for PCR amplification of the isolated EAL domain (AxDGC2₃₀₆₋₅₇₄). The Expanded High-Fidelity Kit (Roche) was used for PCR and the amplified DNA fragments were also cloned into the pET-26(b+) vector via the NdeI and XhoI restriction sites as well. Site-directed mutagenesis was carried out using the Quick Change mutagenesis kit (Stratagene) following manufacturer's instructions. The correct sequences of the cloned and mutated genes were determined by the complete nucleotide sequencing (1st Base). All the primers used for PCR and mutagenesis are listed in **Table 3.1**.

Table 3.1 Primers used in AxDGC2 cloning and site-directed mutagenesis

Primer	Relevant sequence
AxDGC2₃₀₆₋₅₇₄Forward	5'-AAACATATGCGCAATCTGCGTGAAC'
AxDGC2₃₀₆₋₅₇₄Reverse	5'-AAACTCGAGCAGGGTAACGC-3'
N94A Forward	5'- CGCGGCAACATCTGTGCCCCGTGCGAAAGATGGTAG-3'
N94A Reverse	5'- CTACCATCTTTTCGCACGGGCACAGATGTTGCCGCG-3'
D217A Forward	5'-CTGACCCATCCGGATGCGCCGGTGAGCCGTCTG-3
D217A Reverse	5'-CAGACGGCTCACCGGCGCATCCGGATGGGTCTAG-3'
H62A Forward	5'-CTGGTTGGCAGCACCGCCCGCATTGTGAATAGCGG-3'
H62A Reverse	5'- CCGCTATTCAATGCGGGCGGTGCTGCCAACCAG-3'
N66A Forward	5'- GCACCCACCGCATTGTGGCTAGCGGCTATCATGATG-3'
N66A Reverse	5'- CATCATGATAGCCGCTAGCCACAATGCGGTGGGTGC-3'
R125A Forward	5'- CGGCTATGTGGCGAGCGCCTTTGAAATTACCGAACT-3'
R125A Reverse	5'- GTTCGGTAATTTCAAAGGCGCTCGCCACATAGCCGG-3'

3.2.2 Protein expression and purification

The plasmids harboring the genes were transformed into *E. coli* strain BL21(DE3) (Novagen). The cells were grown in Luria-Bertani (LB) media at 37 °C with vigorous shaking (220 *rpm*) until the OD_{600nm} reached 0.6-0.8. Isopropyl-β-D-thiogalactopyranoside (IPTG, 0.5 mM) was added to induce protein expression and the culture was grown for additional 16 hours at 16 °C. Cells were harvested by centrifugation for 10 min at 5000 *rpm* at 4 °C. The cell pellet was frozen and thawed before the cells were lysed by sonication in 40 ml of lysis buffer (50 mM NaH₂PO₄, pH 7.0 (pH 8.5 for AxDGC2₃₀₆₋₅₇₄), 300 mM NaCl (200 mM for AxDGC2₃₀₆₋₅₇₄), 5 mM β-ME, 20 mM imidazole). The cell extract was centrifuged at 18,000 *rpm* for 30 min. All the purification steps described below were performed at 4 °C. The supernatant was filtered and loaded onto 1 ml of Ni²⁺-NTA resin (GE healthcare) that had been pre-packed

into a column. The flow-through was collected and passed through the column again. The column was washed with 50 ml of washing buffer (lysis buffer supplemented with 50 mM imidazole). The proteins were eluted by using the elution buffer (lysis buffer supplemented with 500 mM imidazole). After SDS-PAGE gel analysis, fractions with purity higher than 95% were pooled together and desalted by using either a PD-10 desalting column (GE-Healthcare) or gel filtration column Superdex 200 (GE-Healthcare) with the AKTA FPLC system. The desalting buffer for PD-10 and gel filtration is 50 mM NaPi (pH 8.0), 50 mM KCl, 1 mM DTT. The molecular weight of the protein was estimated based on the standard curve generated by using the standard proteins that included ferritin (440 kDa), aldolase (158 kDa), conalbumin (75 kDa), ovalbumin (44 kDa) and blue dextran (for void volume determination). The bright-yellow looking proteins were concentrated using Amicon concentrator (Millipore) and were stored at -80 °C after the measurement of protein concentration by Bradford assay.

3.2.3 Characterization of the protein-associated flavin and c-di-GMP

UV-Vis spectrum was taken using the UV-1650PC spectrophotometer (Shimadzu) equipped with thermostat and quartz cuvette. For HPLC analysis, AxDGC2 and the isolated EAL domain AxDGC2₃₀₆₋₅₇₄ was denatured by 1% TFA and the protein precipitate was removed by centrifugation at 14,000 *rpm* for 20 min. The supernatant was loaded onto the Agilent LC1200 HPLC system

equipped with an XDB-C18 column (4.6 × 150 mm). The mobile phase for analysis of cofactor bound by AxDGC2 was a gradient from 100% H₂O + 0.045% TFA to 100% acetonitrile + 0.045% TFA in 100 min with a flow rate of 1 ml/min. 100 μM FMN and FAD standards were also applied to the column in the same condition. C-di-GMP bound by both the AxDGC2 and its truncated construct were analyzed by mobile phase (20 mM Triethylammonium bicarbonate (pH 7.0), 9% Methanol), with an isocratic flow rate of 1 ml/min in the same fashion by HPLC.

3.2.4 Enzymatic assay of DGC activity

The progress of c-di-GMP synthesis catalyzed by the diguanylate cyclase domain of AxDGC2 was monitored by using the Agilent LC1200 HPLC system (Mobile phase: 20 mM Triethylammonium bicarbonate (pH 7.0), 9% Methanol, 1 ml/min) with an XDB-C18 column (4.6×150 mm). The enzymatic reaction was performed by incubating the enzyme and GTP at 23 °C in 100 mM Tris-Cl (pH 8.0), 50 mM KCl and 10 mM MgCl₂. The reaction was stopped by adding 1/20 reaction volume of 0.5 M EDTA and then boiled at 95 °C for 5 min. Following centrifugation at 14,000 *rpm* for 5 min to remove protein aggregate, the supernatant was filtered and loaded onto the HPLC. Anaerobiosis was established in a Coy anaerobic chamber with the anaerobic buffer flushed with nitrogen for 0.5~1 hr. The reduced form AxDGC2 was generated via adding 2 mM sodium dithionite to the protein solution under anaerobic condition. The

redox reaction was monitored spectroscopically and the fully reduced sample was mixed with an equal volume of oxygen saturated buffer and exposed to air to re-oxidize the protein. The xanthine oxidase system was also used to reduce the AxDGC2 for comparison. The oxidized and reduced AxDGC2 were incubated with various concentrations of GTP and the initial velocities were obtained from a series of reactions with different incubation time. The total turnover was controlled to ensure the accurate measurement of initial velocities within the linear range. The enzymatic assay for the AxDGC2₁₋₃₀₅ was the same as the full-length protein. The turnover number (k_{cat}) and the Michaelis-Menten constant (K_M) were obtained by fitting the initial velocities at various substrate concentrations to the Michaelis-Menten equation with the exception of the proteins that exhibit substrate inhibition using the software Prism (Graphpad). The kinetic data for these proteins were fit using a model that assumes the substrate binds to the enzyme at a productive and a non-productive (or inhibitory) binding site (**Scheme 3.1**). For Reduced form AxDGC2, kinetic data were fit using Michaelis-Menten equation instead of substrate inhibition equation.

The curves were obtained by fitting the kinetic data to **Eq. 3.1** derived from the above model with steady-state and rapid equilibrium assumptions. In the

equation, K_i is the inhibition constant, α and β are the factors by which the K_M , K_i and V_{max} change when the second substrate is bound at the non-productive site.

$$v = \frac{V_{max} (1/K_m + \beta[S]/\alpha K_i K_m)}{1/[S] + 1/K_m + 1/K_i + [S]/\alpha K_m K_i} \quad \text{Eq. 3.1}$$

3.2.5 Enzymatic assay of PDE activity

C-di-GMP was synthesized using the thermophilic DGC enzyme described previously (96). The reaction condition for phosphodiesterase assay was 100 mM Tris-Cl (pH 8.0), 25 mM KCl and 10 mM MgCl₂ or MnCl₂. The buffer used for the five non-physiological substrates: Thymidine-5-monophosphate-*p*-nitrophenyl ester sodium salt (thymidine-*p*-NPP), O-(4-Nitrophenylphosphoryl) choline, Bis-(*p*-nitrophenyl) phosphate sodium salt (Bis-*p*-NPP), 4-Nitrophenyl phosphate disodium salt hexahydrate and phospho (enol) pyruvic acid monopotassium salt) was 50 mM Tris-Cl (pH 8.0) with 25 mM MgCl₂. The reaction was monitored spectroscopically and the OD reading at 410 nm was recorded. For both thymidine-*p*-NPP and bis-*p*-NPP, the measurement of the steady-state kinetic parameters was carried out by monitoring the OD reading at 410 nm by a UV-spectroscope. The kinetic parameters were obtained by fitting the initial velocities at various substrate concentrations to Michaelis-Menten equation using the software Prism (GraphPad).

2',3'-cAMP was used as another non-physiological substrate for the PDE activity. The production of 3'-AMP was monitored by using the Agilent LC1200 HPLC system (Mobile phase: 20 mM Triethylammonium bicarbonate (pH 7.0), 9% Methanol, 1 ml/min) with an XDB-C18 column (4.6×150 mm). The enzymatic reaction was performed by incubating the enzyme and 2'3'-cAMP at 23 °C in 100 mM Tris-Cl (pH 8.0), 50 mM KCl and 10 mM MgCl₂. The reaction was stopped by adding 1/20 reaction volume of 0.5 M EDTA and then boiled at 95 °C for 5 min. Following centrifugation at 14,000 *rpm* for 5 min to remove protein aggregate, the supernatant was filtered and loaded onto the HPLC. The kinetic parameters were obtained by fitting the initial velocities at various substrate concentrations to Michaelis-Menten equation using the software Prism (GraphPad).

3.2.6 Measurement of FAD to protein binding stoichiometry

The protein concentration was measured by Bradford method, whereas the content of the FAD was measured following the release of the cofactor from the protein by acid denaturation and the concentration determined by using a molar extinction coefficient of free FAD at 450 nm of 11.3 mM⁻¹cm⁻¹ and AxDGC2 at 280 nm of 44.1 mM⁻¹cm⁻¹. The concentration of the 6-OH-FAD was determined by using an extinction coefficient of 22.6 mM⁻¹cm⁻¹ at 427 nm (87). The FAD: protein stoichiometry was obtained from the FAD to protein ratio.

3.2.7 Determination of the redox potential of flavin in AxDGC2

The redox potentials of AxDGC2 and its mutants were determined with the established method described by Uchechi et al recently (141). Briefly, a 200 μ l solution of 5-10 μ M AxDGC2 in 50 mM Tris-Cl pH 8.0 was incubated anaerobically with 5 μ M phenosafranin as the redox indicator at room temperature. A stock solution of 30 mM sodium dithionite was added in 1 μ l aliquots inside the anaerobic chamber. The concentrations of oxidized and reduced AxDGC2-bound FAD and phenosafranin were measured after each dithionite addition by monitoring the absorbance of FAD at 450 nm (H62A at 425 nm) and phenosafranin at 522 nm using the UV-Vis spectrophotometer. The redox potential was determined by plotting $\log (\text{Ox/Red})$ phenosafranin versus $\log (\text{Ox/Red})$ AxDGC2.

3.2.8 Sequence alignment and Structural modeling

The sequence alignments were conducted by the ClustalW and Esript Server. (<http://align.genome.jp/>; <http://esript.ibcp.fr/ESPript/ESPript/>). The structural model of the PAS_{AxDGC2} domain was constructed using the Swiss-Model Server (4). The crystal structure of PAS-A domain of MmoS (PDB code: 3EWK) was used as a template (142).

3.3 Results

3.3.1 Purification and characterization of AxDGC2

About 20-30% of the recombinant AxDGC2 was expressed as soluble protein and purified by metal affinity and size-exclusion chromatography. The typical yield of purified AxDGC2 was ~1 mg/L cell culture. SDS-PAGE revealed the protein to be ~95% homogeneous with a molecular mass in agreement with that predicted from the sequence (65 kDa including the (His)₆-tag at **Figure. 3.2A**). The protein was found in two major fractions after eluted from the gel-filtration column, with a yellow dimeric fraction and a colorless high-molecular-weight oligomeric (HMWO) fraction in the void volume (**Figure. 3.2B**). Purified AxDGC2 was stored at -80°C and was stable for several months.

Figure 3.2 SDS-PAGE and gel filtration analysis of purified AxDGC2 protein. (A) The SDS-PAGE of AxDGC2. Lane1: protein ladder. Lane 2: AxDGC2. (B) The gel filtration of AxDGC2 on Superdex 200. Proteins were loaded and separated using 50 mM Tris-Cl (pH 8.0), 50 mM NaCl and 1 mM DTT.

3.3.2 AxDGC2 binds FAD cofactor in the PAS domain

The absorption spectrum of the dimeric fraction of AxDGC2 suggested that the protein binds an oxidized flavin (FAD or FMN) cofactor, with the characteristic absorbance of 375, 450 and 470 nm for flavin-binding proteins (**Figure 3.3, Table 3.2**). The flavin cofactor was dissociated from the protein after heat or trifluoroacetic acid (TFA) treatment, suggesting that the cofactor is non-covalently bound by the protein. The cofactor was confirmed to be FAD by comparison with the standard by HPLC (**Figure 3.3, Inset**) as well as MALDI-MS analysis (M.W. 787.1) (**Figure 3.4**). The colorless HMWO fraction did not exhibit enzymatic activity, and thus, only the dimeric protein was further investigated. The FAD to protein stoichiometry for the dimeric form was measured to be 0.41 ± 0.1 , indicating relatively weak affinity of the PAS domain for the cofactor. A few residues for flavin binding were predicted by sequence comparison with other flavin-binding PAS domains. We mutated a highly conserved residue (Asn⁹⁴) that putatively interacts with the isoalloxazine ring of FAD. The mutant N94A was purified as a FAD-free protein with a colorless appearance, confirming that the FAD is bound by the PAS domain.

Table 3.2 Maxima of absorption spectra of riboflavin-containing PAS proteins.

<i>Protein</i>	<i>Cofactor</i>	Wavelength (nm)
AxDGC2	FAD	375, 451, 480
AvNifL	FAD	358, 372, 446, 470
Aer	FAD	375, 450, 470
MmoS _(PAS-A)	FAD	376, 444, 470
BLUF_AppA _(lit state)	FAD	367, 384, 457, 483
BLUF_AppA _(dark state)	FAD	363, 378, 443, 471
LOV _(NPH1)	FMN	380, 430, 450, 470

Note: AxDGC2 from *A. xylium*; AvNifL from *Azotobacter vinelandii* (119); Aer from *E. coli*(8); MmoS from *Methylococcus capsulatus*(141); BLUF_Appa from *Synechocystis* sp. PCC6803 (80); LOV-NPH1 from *Arabidopsis thaliana*(22).

Figure 3.3 Absorption spectrum of AxDGC2. Inset: HPLC analysis of the cofactor bound by AxDGC2. (Denatured AxDGC2 sample (black line), FAD standard (red dash line), FMN standard (green dash line))

Figure 3.4 The MALDI-MS result of FAD bound by AxDGC2.

3.3.3 Regulation of the DGC activity by the oxidation state of FAD

Enzymatic assays revealed that AxDGC2 exhibited DGC activity by converting GTP to c-di-GMP when incubated with GTP and Mg^{2+} . In contrast, AxDGC2 did not degrade c-di-GMP even after prolonged incubation of the protein with c-di-GMP and Mg^{2+} or Mn^{2+} . Because some flavin-containing PAS domains function as photoreceptors for perceiving blue light and regulating the activity of output domains, enzymatic assays were performed to test whether the PAS domain of AxDGC2 functions as a photoreceptor, the protein was illuminated for 10-15 min using a continuous light source and the absorption spectrum was immediately taken. Unlike the BLUF and LOV domain-containing proteins, the absorption spectrum of AxDGC2 did not exhibit any noticeable change upon illumination. Moreover, enzymatic assays

performed under the light and dark conditions showed that the rate of c-di-GMP synthesis was not altered, and that the EAL domain remained inactive under the light conditions. The light-independent catalytic activity of AxDGC2 suggested that the PAS domain is not involved in light sensing.

Further enzymatic assays were carried out to test whether the redox status of the FAD cofactor modulates the catalytic activity of GGDEF or EAL domain. The FAD cofactor in AxDGC2 was readily reduced by sodium dithionite under anaerobic conditions. The complete conversion of the oxidized FAD to the reduced form was evident from the changes in absorption spectrum, with the disappearance of the characteristic 451 and 480 nm bands. The reduced AxDGC2 underwent rapid oxidation upon exposure to the air as evidenced by the recovery of the absorption spectrum for the oxidized FAD (**Figure 3.5**). Such reversible oxidation-reduction has been documented for the PAS domains of Aer, NifL and MmoS (6, 52, 141).

To examine the catalytic activity of the oxidized and reduced AxDGC2, the reduced protein was prepared in an anaerobic chamber by incubating the protein with sodium dithionite; whereas the oxidized AxDGC2 was obtained by exposing the reduced protein to air to ensure that the concentrations of the reduced and oxidized enzyme remain identical. Neither the oxidized nor the reduced AxDGC2 could degrade c-di-GMP, suggesting the EAL domain remains inactive regardless of the redox status of the flavin. In contrast, the incubation of GTP, Mg^{2+} with the enzyme revealed that the oxidized AxDGC2

exhibited higher efficiency than the reduced protein in synthesizing c-di-GMP (**Figure 3.6A**). Steady-state kinetic measurements were subsequently carried out to characterize the activity of the oxidized and reduced AxDGC2. The results showed that the oxidized form not only exhibited higher catalytic rate, but also displayed noticeable substrate inhibition in the 0-500 μM [GTP] range (**Figure 3.6B**).

Figure 3.5 AxDGC2 reduced by sodium dithionite. The spectrum of reduced AxDGC2 (green line) was recorded after incubation of the oxidized AxDGC2 (red solid line) with sodium dithionite. The spectrum of the re-oxidized AxDGC2 (red dash line) was recorded after mixing the reduced AxDGC2 with an equal volume of oxygenated buffer and exposing to air for 30 min.

Figure 3.6 DGC activity of oxidized and reduced AxDGC2. (A) HPLC analysis of product formation for oxidized and reduced AxDGC2. Reaction conditions are: 100 mM Tris-Cl (pH 8.0), 50 mM KCl, 10 mM MgCl₂, 10 μM GTP. (B) Initial rate measurement for the oxidized and reduced AxDGC2. Reaction buffer was 100 mM Tris-Cl (pH8.0), 50 mM KCl, 10 mM MgCl₂.

Comparison of the kinetic parameters revealed a 7.8-fold greater k_{cat} and an inhibition constant of K_i of 52 ± 27 μM for the oxidized form (**Table 3.3**). Measurement of the kinetic parameters for the oxidized AxDGC2 in oxygenated or anaerobic buffer did not seem to affect the activity of the oxidized protein, indicating that the difference in catalytic activity is not due to the presence of O₂. To further confirm that the low catalytic activity for the reduced form was not due to the inhibition of sodium dithionite, the xanthine oxidase/xanthine/methyl viologen system was used to reduce AxDGC2 instead (77). The FAD of AxDGC2 could be readily reduced by the enzyme system and the oxidized AxDGC2 still exhibited higher catalytic activity (**Figure 3.7A**). Moreover, the presence of sodium dithionite did not affect the catalytic activity of the FAD-free mutant N94A (**Figure 3.7B**), suggesting that the observed differences in enzymatic activity is indeed caused by the redox change in the

FAD cofactor.

Figure 3.7 (A) Comparison of the diguanylate cyclase activity of the oxidized and reduced forms of AxDGC2 using the xanthine oxidase system. 15 μM AxDGC2, 400 μM xanthine and 1.5 μM Benzyl Viologen were mixed in 100 mM Tris-Cl (pH8.0), 50 mM KCl buffer. The anaerobiosis had been established by flushing the Coy chamber with N_2 , a catalytic amount of xanthine oxidase was added in. The reduction of AxDGC2 was monitored by UV spectrometry. Reaction buffer of DGC activity assay was 100 mM Tris-Cl (pH8.0), 50 mM KCl, 10 mM MgCl_2 . Final substrate concentration was 100 μM and the enzyme concentration was 2 μM . (B) Comparison of the catalytic efficiency of the FAD-free N94A mutant in the presence or absence of sodium dithionite under anaerobic conditions. The reaction conditions are similar to the ones described in Material and Methods.

3.3.4 Effects of mutations in the PAS domain on DGC activity

The $\text{PAS}_{\text{AxDGC2}}$ domain shares sequence identity and similarity with the PAS domains of Aer, *KpNifL*, *AvNifL* and MmoS (**Figure 3.8A**). Particularly, $\text{PAS}_{\text{AxDGC2}}$ shares significant sequence identity (45%) and similarity (76%) with the PAS_{MmoS} domain. The high sequence homology between $\text{PAS}_{\text{AxDGC2}}$ and PAS_{MmoS} (PDB code: 3EWK) allowed us to build a reasonably reliable structural model for the PAS domain. Some of the key residues in *AvNifL* and Aer for the propagation of structural changes have been identified based on

structural and mutagenesis studies (62, 101). Moreover, the crystal structure of PAS_{AvNiFL} determined by Moffat and coworkers revealed a hydrogen bond network formed by the residues Glu⁷⁰, Ser³⁹, His¹³³ and two water molecules in the proximity of the N5 atom of the FAD isoalloxazine ring. Interestingly, although the main residue (Asn⁹⁴) crucial for FAD binding is conserved in PAS_{AxDGC2}, the residues near the N5 atom that could play important roles in signal propagation seem to be different from PAS_{AvNiFL} (62).

According to the structural model of PAS_{AxDGC2}, the residue Asn⁹⁴ directly interacts with the isoalloxazine moiety of FAD via hydrogen binding (**Figure 3.8C**). This residue is highly conserved in flavin-binding PAS domains and it has been shown that the mutation of the equivalent residue in Aer resulted in altered phenotype in *E. coli* (6). We found that the mutation of Asn⁹⁴ abolished FAD binding to AxDGC2, as evidenced by the colorless appearance of the protein and the absence of the FAD spectrum (**Figure 3.9A**). Enzymatic assay showed that the N94A mutant converted GTP to c-di-GMP with a k_{cat} that is 3.8-fold smaller than that of AxDGC2. (**Figure 3.9B, Table 3.3**).

In addition, the structural model suggested that the residue Asn⁶⁶ is likely to form hydrogen bonds with 2'-OH group or the ribityl chain and the O2 atom of the isoalloxazine ring. Surprisingly, the mutant N66A exhibited a higher FAD to protein stoichiometry (0.68 ± 0.1) than the wild type (0.41) (**Table 3.3**), suggesting Asn⁶⁶ is not as important as Asn⁹⁴ in binding the cofactor. Meanwhile, the oxidized and reduced forms of the N66A mutant still exhibited

a difference of 2.8-fold in k_{cat} (**Figure 3.10C**).

Figure 3.8 Comparison of PAS_{AxDGC2} domain with other flavin-binding PAS domains. (A). Sequence alignment of flavin-binding PAS domains. The four residues examined in this study are indicated by the arrows. (B). Phylogenetic relationship analysis of the PAS domains. (AxDGC2 from *Acetobacter xylium*; NifL from *Azotobacter vinelandii* (119); Aer from *Escherichia coli*(8); MmoS from *Methylococcus capsulatus*(141); BLUF_Appa from *Synechocystis sp. PCC6803* (80); LOV-NPH1 from *Arabidopsis thaliana*(22).) (C). Structural model of the PAS_{AxDGC2} domain with the isoalloxazine ring and the four residues examined shown as sticks. The putative hydrogen bonds between the isoalloxazine ring and the residues are represented by the dash lines.

It has been proposed that the change of flavin redox state induces structural changes in PAS_{AvNifL} via reorganizing the hydrogen-bond network in the flavin-binding pocket, and that the reorganization of the hydrogen bond network is initiated by the protonation and deprotonation of the N5 atom of the

isoalloxazine ring (60, 62). The structural model of PAS_{AxDGC2} suggests that His⁶² and Arg¹²⁵ are located near the N5 atom and are the most likely participants in forming hydrogen bonds with N5 through one or more water molecules (**Figure 3.8C**). To assess their roles in signal transduction, the two residues were mutated individually for spectroscopic and kinetic characterization. The FAD to protein stoichiometries for the H62A and R125A mutants were measured to be 0.16 ± 0.06 and 0.3 ± 0.1 respectively. The R125A mutation did not seem to significantly alter the binding environment of the FAD cofactor as evidenced by the absorption spectrum; whereas the absorption spectrum of the H62A mutant has been changed drastically, suggesting either an altered binding environment or modified FAD (**Figure 3.10A**).

Figure 3.9 Comparison of AxDGC2 and the N94A mutant. (A). Absorption spectra of AxDGC2 and the N94A mutant. (B). Comparison of the catalytic activity of AxDGC2 and the N94A mutant. Reaction buffer was 100 mM Tris-Cl (pH8.0), 50 mM KCl, 10 mM MgCl₂.

Table 3.3 FAD: protein stoichiometry and steady-state kinetic parameters of AxDGC2 and its mutants.

Enzyme	FAD:protein stoichiometry	Redox state	k_{cat} (min^{-1})	K_{M} (μM)	$k_{\text{cat}}/K_{\text{M}}$ ($\text{min}^{-1} \mu\text{M}^{-1}$)	K_{i} (μM)
WT	0.4	Ox	5.6 ± 2.4	176 ± 89	$(3.2 \pm 2.1) \times 10^{-2}$	52 ± 27
		Red	0.71 ± 0.1	112 ± 18	$(0.6 \pm 0.1) \times 10^{-2}$	-
H62A	0.1	Ox	7.2 ± 4.2	138 ± 106	$(5.2 \pm 5.0) \times 10^{-2}$	157 ± 135
		Red	1.7 ± 0.4	196 ± 97	$(0.9 \pm 0.5) \times 10^{-2}$	-
N66A	0.68	Ox	2.8 ± 0.8	60 ± 29	$(4.7 \pm 2.7) \times 10^{-2}$	371 ± 233
		Red	1.0 ± 0.5	189 ± 112	$(0.5 \pm 0.4) \times 10^{-2}$	50 ± 31
N94A	-	-	1.5 ± 1.3	239 ± 237	$(0.6 \pm 0.8) \times 10^{-2}$	38 ± 37
R125A	0.3	Ox	8.0 ± 3.0	117 ± 59	$(6.9 \pm 4.4) \times 10^{-2}$	138 ± 77
		Red	0.62 ± 0.1	107 ± 20	$(0.5 \pm 0.1) \times 10^{-2}$	-

To test whether the FAD has been modified, we denatured the mutant protein and found that the free cofactor indeed exhibited altered retention time and absorption spectrum (**Figure 3.11**). A survey of the literature suggested that the modified FAD is most likely to be the 6-hydroxyl FAD (6-OH-FAD) in the neutral pH. 6-OH-FAD is known to occur naturally in several enzymes with a signature peak at ~430 nm and a broad peak (~600 nm, for the anionic form at high pH) (58, 78, 87), which is consistent with the observation that a broad peak at 600 nm would be lost when we decreased the pH of H62A mutant to an acidic condition by TFA 1% (**Figure 3.11B**). Despite the modification, the H62A mutant could still be readily reduced by sodium dithionate or the xanthine oxidase system. Enzymatic assays revealed that the oxidized forms of the R125A and H62A mutants exhibited similar substrate inhibition as the wild

type protein. They exhibited 12.9 and 4.2-fold greater k_{cat} than their reduced forms respectively, relative to the 7.8 fold difference for AxDGC2 (**Figure 3.10B, D**). The difference in catalytic rate between the proteins should be taken with caution since the comparison is complicated by the different FAD to protein stoichiometry.

Figure 3.10 Comparison of AxDGC2 and its mutants. (A). absorption spectra of AxDGC2 and the three mutants. (B-D). Comparison of the catalytic activity of the oxidized and reduced forms of the three AxDGC2 mutants. Reaction buffer was 100 mM Tris-Cl (pH8.0), 50 mM KCl, 10 mM MgCl_2 .

Figure 3.11 Modified FAD in mutant H62A. (A) Comparison of the HPLC chromatograms of FAD and the modified FAD from the mutant H62A. (B) UV-Vis spectra of FAD and the modified FAD from the mutant H62A.

3.3.5 Effects of mutations in the PAS domain on redox potential

In *A. xylinum* cells, the FAD cofactor of AxDGC2 is presumably oxidized by O₂ and reduced by an unknown partner *via* electron transfer. The rate of electron transfer, which could be crucial for the physiological function of the cofactor, is dependent on the redox potential and reorganization energy of the protein-embedded cofactor (24, 73, 89). To investigate the roles of the residues in modulating redox potential, we measured the oxidation potential of AxDGC2 and its mutants (**Table 3.4, Figure 3.12**). The wild type AxDGC2 exhibits a redox potential of -280 ± 3.0 mV, whereas the mutants N66A, R125A exhibit slightly higher redox potentials of -275 ± 0.8 mV, -272 ± 2.2 mV respectively. The H62A mutant that contains the modified FAD exhibited a higher potential of -265 ± 1.3 mV, which matches the potential of -265 mV measured for a

6-OH-FAD reconstituted ferredoxin-NADP⁺ reductase (150). The redox potentials are lower compared to those of the free flavin (-220 mV) and *AvNifL* (-227 mV), but comparable to the potentials of -290 mV for *MmoS* and -277 mV for *KpNifL*.

Table 3.4 Redox potentials of *AxDGC2* and its mutants

Enzyme	WT	H62A	N66A	R125A	<i>AvNifL</i> ^a	<i>KpNifL</i> ^b	<i>MmoS</i> ^c
Redox potential (mV)	-280 ± 3.0	-265 ± 1.3	-275 ± 0.8	-272 ± 2.3	-226 (pH 8.0)	-277 ± 5 (pH 8.0)	-291.2 (pH 8.0)

a. Data from ref. (77) , b. Data from ref. (65), c. Data from ref. (141).

Figure 3.12 Nernst plots of log(ox/red) phenosafranin as a function of log (ox/red) *AxDGC2*-FAD. The redox potentials of the FAD in *AxDGC2* and the mutants were obtained from the y intercept, where log (ox/red) phenosafranin equals zero.

3.3.6 Effects of mutations in I-site on c-di-GMP binding

AxDGC2 contains two putative catalytic domains (GGDEF and EAL) for c-di-GMP synthesis and degradation respectively. HPLC analysis of the extract from the denatured protein solution revealed another small ligand bound by the protein in addition to FAD. The small ligand was identified to be c-di-GMP by comparing with the standard using HPLC (**Figure 3.13A**). C-di-GMP is most likely to be bound by the GGDEF domain because the binding of c-di-GMP by an inhibitory site (I-site) is well known for GGDEF domain (25, 96, 145). Sequence alignment showed that a few key residues, including one of the Arg residues for c-di-GMP binding in the I-site, are not conserved in *AxDGC2* (**Figure 3.13B**). To test whether the c-di-GMP is indeed bound at the I-site, we mutated a conserved residue (Asp²¹⁷) in the putative I-site. HPLC analysis of the denatured protein solution of mutant D217A showed that the protein was no longer associated with c-di-GMP (**Figure 3.13A**), suggesting that the GGDEF domain of *AxDGC2* contains an intact I-site even the residues are not strictly conserved.

Figure 3.13 Analysis of I-site. (A) HPLC analysis of the binding of c-di-GMP by the GGDEF domain at I-site showed that D217A mutation abolished c-di-GMP binding. (B) Sequence alignment of the GGDEF domains of AxDGC1, 2, 3 and two orthodox GGDEF domains of WspR and PleD. The conserved GGDEF motif and the residues in the I-site are highlighted by the underlining bar and arrows, respectively.

3.3.7 Phosphodiesterase activity of the EAL domain

The EAL domain of AxDGC2 lacks a conserved loop 6 (DFG(T/A)GYSS) found in catalytically active EAL domains (97). Instead, it contains a highly degenerate loop 6 consisting of N⁴⁷³FGKGITVL⁴⁸¹. We recently found that the restoration of the loop of the isolated EAL domain by mutating three (N473D, K476T and I478Y) residues is sufficient to recover the enzymatic activity of the domain in hydrolyzing c-di-GMP (97). Although the oxidized and reduced forms of AxDGC2 did not hydrolyze c-di-GMP, the wild type AxDGC2 could hydrolyze the non-physiological phosphodiesterase substrates thymidine *p*NPP and bis *p*NPP, but not the phosphatase substrates 4-nitrophenyl phosphate, O-(4-nitrophenylphosphoryl) choline and phosphor(enol) pyruvic acid. Steady state kinetic measurements showed that $V_{\max}$ was reached at a millimolar substrate concentration for thymidine *p*NPP, while no saturation was observed

for the structurally simpler bis-*p*NPP at comparable substrate concentration level (**Figure 3.14A**). AxDGC2 was further found to be able to hydrolyze 2', 3'-cyclic AMP (cAMP), a non-natural cyclic nucleotide (**Figure 3.14B**). The isolated EAL domain also exhibited similar activity toward the non-physiological substrates, but not c-di-GMP. In addition, we confirmed that the catalytic activity of the EAL domain is independent of the FAD redox state (**Figure 3.15**).

Figure 3.14 Catalytic activity of the EAL domain of AxDGC2 towards non-physiological phosphodiester substrates. (A). Hydrolysis of Bis-*p*-NPP and thymidine-*p*-NPP by AxDGC2. (B). Hydrolysis of 2', 3'-cAMP by AxDGC2. Reaction conditions: 100 mM Tris-Cl (pH8.0), 50 mM KCl, and 10 mM MgCl₂.

Figure 3.15 Comparison of the phosphodiesterase activity of the oxidized and reduced forms of AxDGC2 with the use of 1 mM sodium dithionite. Reaction buffer was 50 mM Tris-Cl (pH8.0), 25 mM MgCl₂.

3.4 Discussion

The cellulose-producing bacterium *A. xylinum* has been a model system for studying cellulose biosynthesis in bacteria. Cellulose synthesis is regulated by the reversible binding of the messenger c-di-GMP to a subunit of the cellulose synthesizing machinery (107-108). Six homologous proteins (AxPDEA1, 2, 3 and AxDGC1, 2, 3) were found to control the cellular concentration of c-di-GMP in *A. xylinum* (126). Although the six proteins share similar domain organization with a PAS-GGDEF-EAL arrangement, *in vivo* analysis suggested that AxPDEA1, 2, 3 are responsible for c-di-GMP degrading; whereas AxDGC1, 2, 3 for c-di-GMP synthesis (126). It was also found that the binding of O₂ by the heme cofactor of the PAS domain of AxPDEA1 suppresses the activity of the EAL domain, implying that high O₂ tension would retard the hydrolysis of c-di-GMP and lead to higher c-di-GMP concentration (14). Despite the similar

domain composition shared by *AxDGC2* and *AxPDEA1*, the results from the current study revealed several major differences between the two proteins. First, the PAS_{*AxDGC2*} domain binds FAD, in contrast to the heme in *AxPDEA1*. This observation is consistent with the highest sequence similarity shared by PAS_{*AxDGC2*} and the PAS domains of *MmoS*, *Aer* and *NifL*. Second, *AxDGC2* contains a catalytically inactive EAL domain; whereas *AxPDEA1* contains a catalytically active EAL domain, consistent with the observation from *in vivo* studies and the mutations in a functional loop (97, 126). Third, while high O₂ concentration stimulates the activity of the GGDEF domain of *AxDGC2* for c-di-GMP synthesis, it represses the EAL domain of *AxPDEA1* for c-di-GMP degradation.

The mutagenesis studies not only confirmed that Asn⁹⁴, but not Asn⁶⁶, is indispensable for FAD binding, but also probed the roles of residues His⁶² and Arg¹²⁵ in propagating the redox signal from FAD to the surrounding protein scaffold. Recent studies have suggested that the reversible protonation of the N5 atom of the isoalloxazine ring initiates the structural change in the FAD-containing *AvNifL* and BLUF through hydrogen-bonded networks (60, 62). Arg¹²⁵ is located near the N5 atom of the isoalloxazine ring and may affect the protonation/deprotonation of the N5 atom; whereas His⁶² stacks on the isoalloxazine ring and may form hydrogen bond with N5 through water molecules. Importantly, the site resided by His⁶² is occupied by an essential cysteine residue that forms covalent adduct with the FAD upon light irradiation

in the light-sensing LOV domain (23), and is occupied by a glutamic acid residue that plays a critical role initiating the reorganization of hydrogen-bond network in AvNifL(62). Our results showed that although the two single mutations caused some minor changes in redox potential and catalytic activity, they did not completely disrupt the transmission of the structural change to the GGDEF domain, which would be manifested by the redox-independent DGC activity. The observation for the H62A mutation is particularly revealing given that the signal transmission has not been blocked despite the modification of the FAD cofactor at the 6-position and perturbation of local structure caused by the deletion of the side chain. The hydroxylation of the FAD is likely to result from the attack of the molecular oxygen residing near position 6, similar to the mechanism proposed for trimethylamine dehydrogenase based on an ^{18}O incorporation experiment (76). It is important to point out that the low FAD:protein stoichiometry observed for the proteins indicates that the protein solutions contain a large portion of apoprotein whose activity does not change with oxidation or reduction. Hence, the observed differences in catalytic rate between the reduced and oxidized proteins are likely to be even greater when the proteins are fully populated by the FAD. On the basis of these observations, we propose that substituted hydrogen-bonded water networks may form in the mutants, and that the protonation/deprotonation of the N5 atom can induce structural change in the protein scaffold through the newly formed hydrogen-bonded water networks. As observed in the flavoprotein mandelate

dehydrogenase, change of flavin redox state can be transmitted through a hydrogen-bonded water network to induce conformational change (123). Therefore, the data do not completely rule out the possibility that the hydrogen-bonded network is still intact in the mutants. Further structural study of the PAS_{AvNiFL} domain will reveal whether there are other residues for organization of the hydrogen-bonded network near the N5 atom and disclose the effect of the mutation on the local protein structure and the hydrogen-bonded network.

It should be stressed that although the mutations did not totally disrupt the communication between the PAS and the GGDEF domains, the three residues could still be crucial for the physiological function of the protein by affecting O₂ binding and electron transfer, and thus, the mutations may still be able to cause phenotypic changes. The binding of O₂ could be affected considering that O₂ must first bind near the N5 atom prior to the oxidation of the FAD (62). Rapid electron transfer between the PAS_{AxDGC2} domain and its redox partners that include the oxidant O₂ and the unknown reductant is also critical for fast response to the change of cellular redox status. Regardless of the identity of the redox partner, the rates of reduction and oxidation of AxDGC2 will be dictated by the redox potential of the protein-embedded FAD (24, 72). In addition to the altered redox potential for the mutants, the reorganization energy that represents the kinetic barrier for electron transfer is likely to be affected as well given that the environment in the FAD-binding pocket has been perturbed.

Intriguingly, the PDE activity assay further revealed that although the EAL domain of AxDGC2 did not degrade c-di-GMP under the experimental conditions, it is a competent PDE domain that can hydrolyze some phosphodiester bond-containing substrates. A one-metal-ion catalytic mechanism for EAL domain proteins was proposed based on the biochemical study on RocR (99); whereas the recent crystallographic study of Brbp1 suggested a two-metal-ion mechanism (5). Considering the lack of an Asp residue for binding the second metal ion in AxDGC2, the observed PDE activity seems to suggest that the EAL domain uses a single-metal-ion mechanism for hydrolyzing the nitrophenol substrates and 2', 3'-cAMP. Furthermore, the observation that the catalytic activity towards c-di-GMP can be recovered by restoring loop 6 also raised the possibility that the activity of the EAL domain towards c-di-GMP can be activated by an unknown signal. A few bifunctional GGDEF-EAL domain proteins with catalytically active GGDEF and EAL domains have been documented (69, 131). Interestingly, a recent study of ScrG, a GGDEF-EAL didomain protein in *V. parahaemolyticus*, showed that the protein exhibits DGC or PDE activity, depending on the presence of another protein partner (63). Hence, it remains to be seen whether AxDGC2 is regulated by another signal in addition to the redox signal.

Finally, the present study demonstrated that the change of the redox status of the flavin-cofactor modulates the enzymatic activity of the GGDEF domain in AxDGC2, and thus, provided support for the notion that the FAD-containing

PAS domain of AxDGC2 functions as redox/oxygen sensor. The PAS domains of AxDGC1 and AxDGC3 are likely to function as redox/oxygen sensors as well given the high sequence similarity (>90%) shared by the PAS domains of AxDGC1, 2 and 3. Considering that the estimated free cellular [GTP] is in the range of 100-400 μ M (12, 46), the different rates observed for oxidized and reduced states could mirror the different efficiency in degrading c-di-GMP for AxDGC1, 2 and 3 in the cells. It is also important to point out that the low FAD to protein stoichiometry observed for AxDGC2 indicates that the protein solutions contained a large portion of *apo*-protein whose activity does not change with oxidation/reduction. Hence, the observed difference in catalytic rate between the reduced and oxidized AxDGC2 is likely to be even greater when the protein is fully populated by FAD in the cells. Together with an early study on AxPDEA1 (14, 43, 108), the data would indicate that high cellular O₂ concentration will simultaneously stimulate the DGC activity of AxDGC1-3 and suppress the PDE activity of AxPDE1-3, on the assumption that all the six proteins are constitutively expressed. Consequently, the cellular c-di-GMP level will rise with oxygen concentration to enable the binding of c-di-GMP to the membrane-embedded cellulose synthase for activating cellulose synthesis (Figure 3.16).

Figure 3.16 A schematic illustration of the relationship between oxygen level, cellular c-di-GMP concentration and cellulose synthesis in *Acetobacter xylinum*.

This model is consistent with the observation that cellulose was produced by better aerated cells at the air-water interface in the static culture of *A. xylinum* (14, 108). The utilization of flavin and heme-containing PAS domains by the synchronized regulation of DGC and PDE activities may be crucial for achieving fast O₂ response in the obligate aerobe *A. xylinum*.

3.5 Summary

The cytoplasmic protein AxDGC2 regulates cellulose synthesis in the obligate aerobe *A. xylinum* by controlling the cellular concentration of the cyclic dinucleotide messenger c-di-GMP. AxDGC2 contains a Per-ARNT-SIM (PAS) domain and two putative catalytic domains (GGDEF and EAL) for c-di-GMP metabolism. We found that the PAS domain of AxDGC2 binds a

flavin adenine dinucleotide cofactor non-covalently. The redox status of the FAD cofactor modulates the catalytic activity of the GGDEF domain for c-di-GMP synthesis, with the oxidized form exhibiting higher catalytic activity and apparent substrate inhibition. The results suggest that AxDGC2 is a signaling protein that regulates cellular c-di-GMP level in response to the change of cellular redox status or oxygen concentration. Moreover, several residues predicted to be involved in FAD binding and signal transductions were mutated to examine the impact on redox potential and catalytic activity. Despite the perturbation of redox potential and the unexpected modification of FAD in one of the mutants, none of the single mutations was able to completely disrupt the signal transmission to the GGDEF domain, indicating that the change of FAD redox state can still trigger structural changes in the PAS domain probably by using substituted hydrogen-bonded water networks. Meanwhile, although the EAL domain of AxDGC2 was found to be catalytically inactive towards c-di-GMP, it was capable of hydrolyzing some phosphodiester bond-containing substrates. The observed catalytic activity raised the possibility that AxDGC2 is under the control of an additional signal and the EAL domain can be activated for c-di-GMP degradation. Together with the oxygen-dependent activity of the homologous AxPDEA1, the results provided new insight into the relationships between oxygen level, c-di-GMP concentration and cellulose synthesis in *A. xylinum*.

CHAPTER 4 INSIGHT INTO THE FUNCTION OF THE DEGENERATE EAL DOMAIN OF THE C-DI-GMP SIGNALING PROTEIN FIMX

4.1 Introduction

The emergence of the cyclic dinucleotide c-di-GMP as a major signalling molecule in bacteria in recent years has unveiled a previously hidden signalling network that controls a variety of bacterial behaviours (50, 59, 106, 114). Some of the behaviours, such as twitching motility, surface adhesion, biofilm formation and virulence expression, may contribute to the pathogenicity of the pathogenic bacteria during infection. The c-di-GMP network is comprised of a large number of signalling proteins as well as riboswitches responsible for c-di-GMP synthesis, degradation and recognition (114). EAL and GGDEF domain-containing proteins are the most prevalent c-di-GMP signalling proteins. The canonical GGDEF and EAL domains function as diguanylate cyclases (DGC) for c-di-GMP synthesis and phosphodiesterases (PDE) for c-di-GMP degradation respectively. Recent studies have, however, unearthed a significant number of degenerate or non-catalytic GGDEF and EAL domains (19, 29, 88, 98, 138). The discovery of the non-catalytic GGDEF and EAL domains adds another layer of complexity to the c-di-GMP signalling network that has yet to be fully understood.

Upon the elucidation of the catalytic mechanism of EAL domains, it was clear that many EAL domains are catalytically incompetent due the lack of the

critical catalytic residues. The EAL domain of FimX was among some of the inactive EAL domains identified. FimX is a cytoplasmic protein (76 kDa) required for normal twitching motility and biofilm formation in the opportunistic pathogen *P. aeruginosa* (55, 61, 67). Although the EAL domain of FimX was first suggested to function as a phosphodiesterase domain, it was later established to be a non-catalytic domain (45, 61, 97, 99). Recent studies from Sondermann and co-workers have demonstrated that the EAL domain of *P. aeruginosa* FimX is capable of binding c-di-GMP with high affinity (86). The crystal structure of the stand-alone EAL domain in complex with c-di-GMP determined by the same group reveals the binding of c-di-GMP without the assistance of metal ion. Farah and co-workers further reported that the EAL domain of the FimX homolog in *Xcc* not only binds c-di-GMP, but also interacts directly with the PilZ protein, a key protein in type IV pili biogenesis (45). In addition to the EAL domain, FimX contains three other protein domains that include a CheY or REC (phosphoreceiver-like) domain, a PAS (Per-Sim-ARNT) domain, and a GGDEF domain (**Figure. 4.1A**). The functions of the three domains remain unknown despite the overall sequence similarity shared with other homologous domains. Notably, the defective REC domain is unlikely to function as an archetypal phosphoreceiver domain because it lacks the essential Asp residue required for phosphorylation (55) (**Figure 4.1B**).

Mattick and co-workers first reported that Green fluorescent protein (GFP)-tagged FimX adopts a unipolar localization in *P. aeruginosa* (55). The

cellular localization of FimX was shown to be critical for the function of FimX in the assembly of unipolar type IV pili (55, 61). *In vivo* studies demonstrated that the N-terminal REC domain and the adjacent linker region are indispensable for retaining the unipolar localization of FimX, indicating that the localization sequence resides in the N-terminal region (55, 61). It is puzzlingly that deletion of the EAL domain or mutation of the signature EVL motif of the EAL domain has a dramatic effect on protein localization by greatly reducing the probability of the protein to adopt the unipolar distribution (61). A molecular mechanism that can rationalize the *in vivo* observation is still lacking. In this study, we examined FimX using the method of amide hydrogen/deuterium (H/D) exchange-coupled mass spectrometry. The highly sensitive H/D exchange method revealed that the binding of c-di-GMP to the EAL domain triggers a distinctive conformational change in the REC domain and the adjacent linker region. Binding assays demonstrated that mutation of the EVL motif completely abolishes the c-di-GMP binding capability of FimX. Together, the results provide a novel molecular explanation for the mediation of protein function and cellular location by c-di-GMP through an allosteric regulation mechanism.

Figure 4.1 Domain organization and sequence alignment of *Pseudomonas aeruginosa* FimX. (A) Domain organization of FimX with the sequence of the linker between the REC and PAS domain highlighted. (B) Comparison of the CheY-like REC domain of FimX with CheY domain of other proteins. The Asp residue for phosphorylation is indicated red asterisk. The sources of the CheY sequences are as follows: CheY_PleD_Pa, CheY of PleD_ *Pseudomonas aeruginosa* (AAG04845); CheY_Ec, CheY of *Escherichia coli* (NP_416396); CheY_Ba, CheY of *Bacillus amyloliquefaciens* (CP_000560.1)

4.2 Methods and Materials

4.2.1 Gene cloning and site-directed mutagenesis

The gene encoding the full-length FimX was first amplified by PCR from the genomic DNA of *P. aeruginosa* PAO-1 using the Expand High fidelity kit (Roche). The gene fragment was cloned into the expression vector pET 26b(+) (Novagen) between the NdeI and NotI restriction sites. The DNA fragment that encodes the EAL domain (FimX₄₃₆₋₆₉₁) was cloned into the expression vector pET 28a(+) (Novagen) between the NdeI and XhoI restriction sites. All the

plasmids harboring the gene constructs and the (His)₆-Tag encoding sequence were transformed into *E. coli* strain BL21(DE3) for protein expression. Site-directed mutagenesis was carried out using the Quick Change mutagenesis kit (Stratagene) following manufacturer's instructions. The correct sequences of the cloned and mutated genes were determined by the complete nucleotide sequencing (1st Base). All the primers used for PCR and mutagenesis are listed in **Table 4.1**.

Table 4.1 Primers used in FimX cloning and site-directed mutagenesis

Primer	Relevant sequence
FimX Forward	5'-CATATGATGGCCATCGAAAAGAAAACC-3'
FimX Reverse	5'-GCGGCCGCTTCGTCTCCCGAGGAGAA-3'
FimX₄₃₆₋₆₉₁ Forward	5'-TTTCATATGGCCGCCGCCAGCGC -3'
FimX₄₃₆₋₆₉₁ Reverse	5'-GCGGCCGCTTCGTCTCCCGAGGAGAA-3'
FimX_{EVL-AAA} Forward	5'-CCACGAGAACTACGCGGCGGCCCTGCGCCTGCTC-3'
FimX_{EVL-AAA} Reverse	5'-GAGCAGGCGCAGGGCCGCCGCGTAGTTCTCGTGG -3'

4.2.2 Protein expression and purification

The plasmids harboring the genes were transformed into *E.coli* strain BL21(DE3) (Novagen). 2 ml of the culture inoculated by the cell stock was added to one liter of LB medium. Bacterial culture was grown at 37 °C up to OD of 0.8 before being induced with 0.5 mM IPTG at 16 °C for 16 hours. After harvest by centrifugation, the cell pellets were lysed with the 40 ml lysis buffer containing 50 mM NaH₂PO₄ (pH 7.0), 300 mM NaCl, 5 mM β-mercaptoethanol and 20 mM imidazole. After centrifugation at 20,000 *rpm* for 30 min, the

supernatants were filtered and then incubated with 2 ml of Ni^{2+} -NTA resin (Qiagen) for 30 min. The resin was washed with 50 ml of wash buffer (lysis buffer with 50 mM imidazole). The proteins were eluted with the elution buffer (lysis buffer with 300 mM imidazole). After SDS-PAGE gel analysis, fractions with purity higher than 95% were pooled together. Size exclusion chromatograph were carried out at 4 °C using the AKTA FPLC system equipped with a Superdex 200 HR 16/60 column (GE Healthcare) for the full-length FimX and a Superdex 75 HR 16/60 column for FimX₄₃₆₋₆₉₁. C-di-GMP was found to be co-purified with the full-length FimX and the EAL domain. To produce FimX without c-di-GMP binding, the purified proteins were treated with the c-di-GMP specific phosphodiesterase RocR and Mg^{2+} (99). The ligand-free proteins used in this study were separated from RocR by size-exclusion chromatography. All the purified proteins were flash frozen in liquid N_2 and stored in -80 °C freezer after the measurement of concentration by Bradford method. The final storage buffer was 20 mM Tris-Cl (pH8.0), 100 mM NaCl, 5% glycerol, and 2 mM DTT. All the mutant proteins were expressed and purified following the same procedure for the wild type proteins.

4.2.3 Characterization of the protein-associated c-di-GMP

UV-Vis spectrum was taken using the UV-1650PC spectrophotometer (Shimadzu) equipped with thermostat and quartz cuvette. For HPLC analysis, full-length wild type FimX, mutants and the isolated EAL domain FimX₄₃₆₋₆₉₁

were denatured by 1% TFA and the protein precipitate was removed by centrifugation at 14,000 *rpm* for 20 min. The supernatant was loaded onto the Agilent LC1200 HPLC system equipped with an XDB-C18 column (4.6 × 150 mm). C-di-GMP bound by the full-length wild type FimX, the mutants and its truncated construct and mutants were analyzed by mobile phase (20 mM Triethylammonium bicarbonate (pH 7.0), 9% Methanol), with a flow rate of 1 ml/min in the same fashion by HPLC.

4.2.4 Amide hydrogen-deuterium (H/D) Exchange by mass spectrometry

Hydrogen/deuterium (H/D) exchange is a commonly used approach to investigate protein dynamics. The exchange rate between the hydrogen of backbone amide and isotopically-labeled water is used as evaluation of the extent of local and global protein conformational changes. This method is conventionally used in other biophysical instruments such as NMR spectrometry, but there are certain limitations. Therefore, mass spectrometry (MS) is now widely applied in probing protein dynamics as it provides high sensitivity in detecting low concentration of protein including large proteins that are not suitable for NMR studies.

The ionization method used in H/D exchange with mass spectrometry (HDX-MS) can be electrospray ionization (ESI) or matrix-assisted laser desorption ionization (MALDI). To increase the resolution of the results generated by mass spectrometry, protein samples are subjected to digestion to

peptides. In-exchange reaction of protein is first quenched by low pH and low temperature, which is then followed by pepsin digestion (54). HPLC is interfaced with mass spectrometry to separate the digested peptides and mass spectrometry analysis is then performed (**Figure 4.2**). Our HDX system was set up follow the protocol from Lee T and colleagues (70)

Figure 4.2 The general overview of the experimental setup of hydrogen/deuterium exchange-mass spectrometry. Protein with amide hydrogen is incubated with D_2O and deuteration occurs at the backbone amide. Deuterium incorporation is quenched by low pH and low temperature. Protein is then subjected to pepsin digestion and further processed by HPLC and mass spectrometry analysis (The figure is adopted from reference (54))

4.2.4.1 HPLC column packing and system configuration

Capillary HPLC columns were assembled by packing reversed-phase resin into fused-silica tubings (Polymicro Technologies, Phoenix, United States). The outlet tubing (320 μm I.D.) was 5-10 cm in length and the inlet tubing (530 μm I.D.) was 15-20 cm in length. They were connected with 1 inch of outlet tubing being inserted into the inlet tubing and the connection point was covered by Epoxy glue (Epoxy Technology, Billerica, United States). The inlet end of glued capillary was filled with Teflon patch, the C18 extraction disk (3M EmporeTM, Saint Paul, United States). Reversed-phase POROS 20 R1 resin (Applied Biosystems, Foster City, United States) was mixed with 100% ethanol to become resin slurry. It passed through the capillary by 200 *psi* of helium gas generated by pressure injection cell (Next Advance, Averill Park, United States). Teflon patch was pushed upwards and trapped at the connection point where the outlet diameter is relatively smaller and it served as a seal so that the resin filled the inlet from top to the bottom.

The HPLC sample injector valve, model 9010 (Rheodyne, Rohnert Park, United States) was coupled with a MicroPro[®] high pressure (up to 10,000 psi) HPLC pump (Eldex, Napa, United States). Polyetheretherketones (PEEK) sample loops (Agilent Technologies, Foster City, United States) were used for sample loading (500 μL) and solvent pre-cooling (2 mL). Syringes for sample injection (250 μL) and gradient buffer loading (50 μL) were wrapped with parafilm (Alcan Packaging, Neenah, United States) to prevent heat transmission

(syringe with 23 gauge blunt needle is preferable). The packed capillary HPLC column was connected to mass spectrometry by 10FEP tubing sleeves of 0.018 inch I.D. (Upchurch Scientific Inc, Oak Harbor, United States). A large ice bucket was used for submerging the injector valve, sample loop, solvent pre-cooling loop and HPLC reversed-phase column. Low temperature (0 °C) was maintained to minimize back-exchange of amide hydrogen. A stainless steel needle port was used for injection by syringe (Rheodyne, Rohnert Park, United States)

4.2.4.2 Hydrogen-Deuterium Exchange

Purified FimX was exchanged into 20 mM sodium phosphate buffer (pH 7.4) that contains 100 mM sodium chloride and 2 mM DTT. 8 µL of c-di-GMP free FimX (5 mg/ml) was pre-incubated with 2 µL of c-di-GMP (250 µM) at room temperature for 1 min. C-di-GMP was synthesized enzymatically using a thermophilic DGC protein as described previously (96). For the control group, 2 µL HPLC grade water was added instead of c-di-GMP. After pre-incubation, protein sample was transferred to ice (0 °C). Meanwhile 90 µL D₂O (Sigma Aldrich, Wilwaukee, United States) was pre-chilled in 10 °C water bath and then added into the prepared protein sample. Immediately, the mixture was transferred into the 10 °C water bath to perform the H/D exchange for several lengths of time: 10 s, 1 min, 5 min, 20 min, 40 min, 60 min and 90 min. After D₂O incubation, protein sample was transferred to -10 °C water bath for 5 s to

cool it rapidly. To quench the in-exchange reaction, 90 μL of 0 $^{\circ}\text{C}$ 25 mM sodium citrate (Sigma-Aldrich, St. Louis, United States) / 25mM sodium succinate (Sigma-Aldrich, St. Louis, United States) at pH 2.4 in LC/MS grade water was added immediately. This acidified protein sample was incubated further in -10 $^{\circ}\text{C}$ for 20 s and then being transferred to ice where 10 μL of pepsin (Sigma- Aldrich, St. Louis, United States) was added immediately. The ratio of FimX and pepsin is 1 to 1.

Protein/pepsin sample (200 μL) at pH 2.4 was injected immediately using parafilm- wrapped syringe into 500 μL PEEK sample loop which was rinsed with 0.1% formic acid (FA) in HPLC grade water. One minute after protease addition, the injector valve was switched to inject position to inject all digested sample into the column which was equilibrated with 0.1% FA previously. Flow rate was controlled by the HPLC pump at 40 $\mu\text{L}/\text{min}$. After 5 min of injection onto the column, injector valve was switched back to load position for the washing of column. Meanwhile, sample loop was rinsed with 2 mL of HPLC grade water to remove D_2O crystal leftover. Step gradients which were kept on ice (0 $^{\circ}\text{C}$) before was then injected into the sample loop in the order of 40 μL of 80% acetonitrile (ACN) in 0.1% FA followed by 17.5 μL each of 40, 30, 25, 22.5, 20, 17.5, 15, 12.5, 10, 7.5% ACN in 0.1% FA. After 6 min of column washing, the flow rate of HPLC pump was reduced to 20 $\mu\text{L}/\text{min}$ and the column was then connected to the electrospray source of mass spectrometry. Injector valve was switched to inject position to inject the gradient buffer onto

the column and elute the peptides for mass spectrometry analysis.

A non-deuterated sample was used as control by adding 90 μL of HPLC grade water instead of 90 μL of D_2O . This result was collected for the peptide identification based on the MS/MS tandem spectra. Another zero-time control was performed to measure the low-extent artifactual in-exchange at backbone protons after quenching of deuteration and pepsin digestion. This was done by adding 90 μL of quenching buffer (pH 2.4) to 10 μL of protein sample (0 $^{\circ}\text{C}$) followed by putting another tube containing 90 μL of D_2O into 10 $^{\circ}\text{C}$ water bath. Acidified protein mixture was transferred to -10 $^{\circ}\text{C}$ water bath prior to addition of D_2O (10 $^{\circ}\text{C}$). Subsequent steps were accomplished as described above.

4.2.4.3 Data analysis

Xcalibur software was used for spectrum analysis and data extraction. The H/D exchange data were processed by using the program HX-Express. Zero time-point control or the “artifactual in-exchange” control was performed by adding the quenching buffer to the protein solution before exposure to D_2O . Measured peptide masses were corrected for artifactual in-exchange at $t=0$, normalized to 100% D_2O , and corrected for back exchange following the empirical method described by Hoofnagle *et al* (54).

The MS and MS/MS spectra from raw files could be extracted by using the Raw Extracter software (The Scripps Research Insititute). A relatively small

database was built including FimX and several other proteins from *P. aeruginosa*. These MS/MS files were then searched against the database for peptide identity by using the ProLuCID search engine (The Scripps Research Insititute). A quantity of digested peptides then could be identified and the sequence coverage of FimX protein could be calculated.

The raw data of each HDX run could be opened as a total ion chromatograph by using Xcalibur. Different peptides appeared as different peaks in the chromatograph. To find the Retention Time for each peptide, the “Displayed Options” dialog box was opened firstly to label the charge status of all the peaks. Secondly, the “Ranges” dialog box was opened and the mass range of a particular peptide was entered in the “Plot properties” dialog box to select the corresponding peak with correct charge status from all the isotopic peaks falling to this mass range. Thirdly, the m/z and signal intensity of a particular peptide each was saved in a text file and opened in a spreadsheet using the Microsoft Excel. The m/z and signal intensity were placed in the first and second column respectively. HX express software was used to calculate the weighted average mass for each peptide at time t ($M_{t,avg}$) by the **Eq. 4.1**, where $\sum (m/z \times \text{intensity})$ is the sum of multiplication of m/z ratio of each isotopic form and its corresponding intensity, $\sum (\text{intensity})$ is the sum of intensity of all isotopic forms.

$$M_{t,avg} = [\sum (m/z \times \text{intensity}) / \sum (\text{intensity})] \times z - z \quad (\text{Eq. 4.1})$$

Low extent of artifactual in-exchange occurs after quenching buffer was added to stop the in-exchange. The corrected peptide mass is calculated based on the zero-time control experiment data by **Eq. 4.2**:

$$M_{t,corr(IE)} = (M_{t,avg} - LM_{inf,90}) / (1-L); \quad L = (M_0 - M_{calc}) / (M_{inf,90} - M_{calc}) \quad (\text{Eq. 4.2})$$

$M_{t,corr(IE)}$ is the artifactual in-exchange corrected peptide mass at time t , $M_{t,avg}$ is the observed weighted average mass at time t , $M_{inf,90}$ is the theoretical mass of the peptide at infinite exchange where 90% of nonproline backbone amide hydrogens are incorporated with deuterium, M_0 is the observed average mass in zero-time control, M_{calc} is the theoretical average mass of the peptide, L is the fraction of artifactual in-exchange.

Following digestion and before mass determination, deuteriums at backbone amide would slowly back-exchange to hydrogen during HPLC separation in water. There is an empirical formula (**Eq. 4.3**) derived to calculate the fractional back-exchange (102)

$$BE = L \times (\%H_2O/\%D_2O) + [(peptide\ elution\ time\ from\ HPLC\ in\ min + 6min) \times 1\%] (\text{Eq. 4.3})$$

In **Eq. 4.3**, $(\%H_2O/\%D_2O)$ is the scaling factor that reflects the H_2O solvent in the back-exchange normalized to D_2O in the artifactual in-exchange.

This estimation accounts for approximately 10% of observed back-exchange of the peptide on the column prior to elution in each minute. Together with estimation of artifactual in-exchange and back-exchange, the peptide mass can be corrected by the **Eq. 4.4**:

$$M_{t,corr}(BE) = M_{t,corr}(IE) + [BE \times (M_{t,corr}(IE) - M_{calc})] \quad (\text{Eq. 4.4})$$

The final corrected time-dependent H/D exchange data were fit by nonlinear least squares fitting as described previously by using **Eq. 4.5** (97). N is the total number of deuterons incorporated over the observed course for each peptide, and A , B and C correspond to the number of amides exchanging with the rate constants k_1 , k_2 and k_3 , respectively (97, 147). The number of non-exchanging amides is calculated from the total number of the backbone amides (N_H) in the peptide, excluding proline residues.

$$Y = N - (Ae^{-k_1 t} + Be^{-k_2 t} + Ce^{-k_3 t}) \quad (\text{Eq. 4.5})$$

4.2.5 Isothermal titration calorimetry (ITC)

The dissociation constants (K_d) and stoichiometric ratio for the interaction between FimX and c-di-GMP were measured by isothermal titration calorimetry (ITC) using a ITC₂₀₀ calorimeter (Microcal). Calorimetric titration of c-di-GMP (1 mM in the syringe; 1 μ L injections) and full-length FimX,

FimX₄₃₆₋₆₉₁ or FimX mutants (40 μ M in the cell) was performed at 25 °C in the assay buffer that contains 10 mM Tris-Cl (pH 8.0), 100 mM NaCl. A time spacing of 240 s was set between injections. ITC data were analyzed by integrating the heat effects after being normalized to the amount of injected protein. Data fitting was based on a single-site binding model using the embedded software package (Microcal) to obtain the dissociation constants.

4.2.6 Structural modeling

The structural models for the REC and PAS domains of FimX were built using the structures of the REC domain of PleD (PDB code: 1W25) and PAS domain of FixJ (PDB code: 1DRM) as templates. The final model of the full-length FimX was built based on the FimX model from Sondermann and coworkers, with the structural models of the REC and PAS domain incorporated into the original model.

4.3 Results

4.3.1 Preliminary characterization of full-length FimX and FimX₄₃₆₋₆₉₁

The full-length FimX expressed in *E. coli* was purified by metal-affinity and size-exclusion chromatography. The typical yield of the purified FimX was ~30 mg/L cell culture for full-length and ~20 mg/L cell culture for the FimX₄₃₆₋₆₉₁. SDS-PAGE showed that the protein is ~95% homogeneous and the

molecular mass is in agreement with that predicted from the sequence with a calculated mass of 76 kDa for full-length FimX and 28 kDa for FimX₄₃₆₋₆₉₁ (**Figure 4.3**). The gel-filtration chromatography suggests that the oligomerization state of the full-length FimX is dimeric and that the FimX₄₃₆₋₆₉₁ is monomeric (**Figure 4.3**). A significant amount of c-di-GMP was found to bind the purified FimX as evidenced by the absorption spectrum of the protein and the HPLC analysis of the supernatant following heat treatment of the protein. To remove c-di-GMP from the protein, the purified protein was treated with the c-di-GMP specific phosphodiesterase RocR and Mg²⁺ (99). The ligand-free FimX used in this study was separated from RocR by size-exclusion chromatography. Purified full-length FimX and FimX₄₃₆₋₆₉₁ were stored at -80 °C and seemed to be stable for several months.

4.3.2 Probe the interaction between FimX and c-di-GMP by amide H/D exchange

The binding of c-di-GMP by FimX does not induce a change in global structure or oligomeric state because the free and c-di-GMP bound FimX proteins appear to be almost identical in solution scattering and analytic ultracentrifugation analysis (86). We postulated that the effect of c-di-GMP binding might be exerted through a local conformational change that was undetectable by the scattering and sedimentation methods. Amide H/D exchange by mass spectrometry is a highly sensitive technique for detecting

ligand-binding site and conformational changes in proteins by monitoring the deuteration level of the peptide segments (54). We reasoned that Amide H/D exchange by mass spectrometry would be an ideal tool for detecting small local conformational changes in FimX.

Figure 4.3 SDS-PAGE and gel filtration characterization of FimX and FimX₄₃₆₋₆₉₁. (A) SDS-PAGE analysis of the purified FimX. Lane 1: protein ladder. Lane 2: FimX. (B) Gel filtration chromatogram of FimX on Superdex 200. (C) SDS-PAGE analysis of the purified FimX₄₃₆₋₆₉₁. Lane 1: protein ladder. Lane 2: FimX₄₃₆₋₆₉₁. (D) Gel filtration chromatogram of FimX₄₃₆₋₆₉₁ on Superdex 75.

H/D exchange by mass spectrometry was performed following an established protocol (71, 97, 102). The peptides generated by pepsin digestion were identified by a MS/MS experiment. H/D exchange of FimX was

conducted in the absence and presence of c-di-GMP. A total of 38 peptides that encompass 86 % of the protein sequence were chosen for data analysis. The collation of 38 peptides starting from the N-terminus of FimX is shown in **Figure 4.4**. The peptides include peptides 1 to 8 from the REC domain and the adjacent linker, peptides 9 to 14 from the PAS domain, peptides 15 to 26 from the GGDEF domain and peptides 27 to 38 from the EAL domain. The deuteration patterns of every peptide in the absence and presence of c-di-GMP was compared by examining the time-dependent deuteration profiles. Overall, 30 of the 38 peptides exhibited almost identical H/D exchange patterns (**Figure 4.5**). Meanwhile, statistically significant differences in H/D exchange rates and deuteration levels are observed for eight peptides. The eight peptides are found exclusively in the EAL domain, the REC domain, and the adjacent linker region as discussed below.

In the EAL domain, six peptides exhibited reduced levels of deuteration, indicating a decreased solvent accessibility for the six peptides upon c-di-GMP binding (**Figure 4.6A**). The six peptides were mapped onto the crystal structure of the EAL domain of FimX (PDB code: 3HV8) (**Figure 4.6B and C**). Peptides 29 and 38 exhibited the greatest suppression in deuteration. The crystal structure shows that both peptides contain residues that are in direct contact with the guanine moiety of c-di-GMP. In addition, Arg⁴⁷⁹ of peptide 29 is involved in charge-charge interaction with the phosphate group. Notably, peptide 29 also harbors the signature motif EVL and a short helix that was seen

to undergo a rigid body shift upon c-di-GMP binding (86). Minor suppression of deuteration was observed for four peptides (peptides 28, 32, 33 and 36) in the EAL domain. Peptide 28 consists of the first β -strands of the central barrel, whereas peptides 33 and 36 contain β -strands and loops from the c-di-GMP binding pocket. A small reduction in deuteration was observed for peptide 32, which is the only peptide that does not make direct contact with c-di-GMP. In the crystal structure of the stand-alone EAL domain, the N-terminal helix ($\alpha 0$) becomes unstructured upon c-di-GMP binding (86). No difference in deuteration is observed between the free and c-di-GMP bound forms of the helix $\alpha 0$ -containing peptide 27 (**Figure 4.5**), indicating that the secondary structure of the helix does not change in the full-length FimX.

Figure 4.4 Sequence coverage map of FimX with the secondary structure indicated. The peptides used for the analysis of deuterium pattern are represented by the underlying bars. The red and yellow bars represent the peptides exhibiting altered deuterium incorporation upon the binding of c-di-GMP, whereas the grey bars represent the peptides with unchanged deuterium incorporation.

Figure 4.5 continued.

Figure 4.5 H/D exchange profiles for the peptides that do not exhibit significant

changes upon c-di-GMP binding. Time courses of deuteration are shown for the peptides in the absence (●) and presence (○) of 10 μ M c-di-GMP. Weighted average masses of peptides were corrected for artifactual in-exchange and back-exchange.

In contrast to the EAL domain, none of the peptides in the adjacent GGDEF and PAS domains showed significant change in deuteration. For instance, the exchange profile for the peptide (peptide 22) that contains the GDSIF motif in the GGDEF domain revealed that only two amide hydrogens were substituted with deuterons during H/D exchange regardless of the presence of c-di-GMP (**Figure 4.7A**). The lack of changes indicates that c-di-GMP binding has little effect on the conformation or flexibility of the GGDEF and PAS domains. In contrast, two peptides from the N-terminal REC domain and the adjacent linker exhibited reduction in deuteration upon c-di-GMP binding (**Figure 4.7A**). One peptide (peptide 6) resides in the REC domain, and the other (peptide 8) resides in the adjacent linker region (**Figure 4.7B**). Peptide 6 contains the α 5 helix, β 4 strand and part of the α 4 helix and constitutes most of the α 4- β 4- α 5 face that is known as the interaction surface for CheY-like proteins. Meanwhile, peptide 8 consists of part of the predicted long helix linker between the REC and PAS domain. The orientation of the REC domain and the position of peptide 8 are arbitrary in the structural model shown in **Figure 4.7B** because the full-length FimX or REC domain structure is unavailable at this moment. The suppressed deuteration levels upon c-di-GMP binding for peptides 6 and 8 indicate a decrease of solvent accessibility in the

region containing the two peptides, likely due to a conformational change as discussed below.

4.3.3 Effect of the mutation of the EVL motif on c-di-GMP binding

In vivo studies showed that the N-terminal REC domain and the adjacent linker region harbors the protein sequence that is indispensable for the unipolar localization of FimX (55, 61). Intriguingly, Kazmierczak *et al* made the observation that the mutation of the E⁴⁷⁵V⁴⁷⁶L⁴⁷⁷ motif to AAA in the EAL domain dramatically alters the unipolar localization of FimX, with the triple mutant protein more likely adopting a bipolar or non-polar distribution. Over-expression of the mutant protein in the $\Delta fimX$ strain also failed to restore twitching motility (61). In the crystal structure of the EAL_{FimX} complexed with c-di-GMP, c-di-GMP is bound in the pocket through the interaction with several polar and non-polar residues including Tyr⁶⁷³, Glu⁶⁵⁴, Asp⁵⁰⁷, Arg⁴⁷⁹ and Leu⁴⁷⁷ (86). The EVL motif may contribute to c-di-GMP binding since Leu⁴⁷⁷ makes direct contact with the ribose and guanine moiety and Glu⁴⁷⁵ interacts with c-di-GMP indirectly through a water-mediated hydrogen-bond network.

Figure 4.6 Effect of c-di-GMP binding on deuteration pattern in the EAL domain. (A). Time courses of deuteration are shown for the peptides from the EAL_{FimX} domain that exhibit significant different deuteration patterns in the absence (●) and presence (○) of 10 μ M c-di-GMP. Weighted average masses of peptides were corrected for artifactual in-exchange and back-exchange and fit by nonlinear least squares to a sum of three exponential terms as described in the Experimental Procedures. (B). Crystal structure of the EAL_{FimX} domain in complex with c-di-GMP (PDB code: 3HV8). The peptides exhibiting decreased deuteration upon the binding of c-di-GMP (stick model) are colored red and yellow. (C). Surface representation of the EAL_{FimX} domain with the region exhibiting solvent protection upon c-di-GMP binding colored red and yellow. The structure shares the same orientation as the one in panel B.

Figure 4.7 Effect of c-di-GMP binding on the deuteration pattern in the N-terminal REC domain and the linker region. (A) Time courses of deuteration are shown for the peptides from the EAL_{FimX} domain that exhibit significant different deuteration patterns in the absence (●) and presence (○) of 10 μ M c-di-GMP. Weighted average masses of peptides were corrected for artifactual in-exchange and back-exchange and fit by nonlinear least squares to a sum of three exponential terms as described in the Experimental Procedures. (B) Structural model of the full-length FimX protein with the regions exhibited altered deuteration colored red and yellow. The peptide that contains the GDSIF motif in the GGDEF domain is colored green.

Given the structural information and the conformational change induced by c-di-GMP as suggested by the H/D exchange results, we further examined whether the effect of the EVL-AAA mutation is due to the perturbation of c-di-GMP binding. A binding assay by isothermal titration calorimetry (ITC) was conducted to compare the wild type FimX and the AAA triple mutant. Full length FimX bound c-di-GMP with a sub-micromolar K_d of $0.14 \pm 0.02 \mu$ M (**Figure 4.8A**), which is close to the reported K_d of 125 nM (86). Remarkably,

the triple mutation reduces the binding affinity for c-di-GMP drastically, with no significant binding observed for the mutant under the experimental conditions where the final concentration of c-di-GMP reaches 100 μ M (**Figure 4.8B**). The disastrous effect of the mutation on the binding is surprising considering that the key residues (Arg⁴⁷⁹ and Glu⁶⁵⁴) for c-di-GMP binding remain in place. At the same time, the lack of c-di-GMP binding for the mutant also confirms that FimX only possesses a single c-di-GMP binding site in the EAL domain.

Figure 4.8 Isothermal titration calorimetry (ITC) analysis of the effect of mutation in the EVL motif on c-di-GMP binding. (A). Binding of c-di-GMP by the full length FimX. (B). Binding of c-di-GMP by the triple mutant FimX (EVL-AAA).

4.4 Discussion

Comparison of the structures of the free and c-di-GMP bound EAL_{FimX} domain suggested that the EAL domain does not undergo large conformational changes upon ligand binding (86). The decreases in solvent accessibility in the EAL domain upon c-di-GMP binding are thus likely caused by direct protein-ligand interaction. The largest decreases are seen in the peptides that form hydrogen-bonds with c-di-GMP, consistent with the idea that formation of hydrogen-bonds between the protein and ligand can impede H/D exchange. The abolishment of the binding for c-di-GMP by the mutation of the EVL motif to AAA reveals the critical roles of the three residues and local structure for c-di-GMP recognition. Considering the highly similar c-di-GMP binding mode shared by EAL domains, such critical role of the EAL or EVL motif in c-di-GMP binding may be common for EAL domains.

The most striking observation of this study is the change of solvent accessibility in the REC domain and the adjacent linker region upon the binding of c-di-GMP. ITC experiments with the full-length FimX and the AAA triple mutant confirm that FimX only harbors a single c-di-GMP binding site in the EAL domain. Thus, the altered solvent accessibility cannot be attributed to the direct binding of c-di-GMP in this region, but most likely due to a long-range conformational change triggered by c-di-GMP binding in the EAL domain. The observation is remarkable because the location of the conformational change is *ca.* 70 Å away from the c-di-GMP binding site, and that no significant changes

are seen in the intervening PAS and GGDEF domains. How the binding signal is transmitted over such a long distance in FimX remains an intriguing question for future exploration. It also should be noted that the data can also be accommodated by a classic allosteric model that assumes the pre-existence of two discrete conformations for the ligand-free FimX. The binding of the ligand c-di-GMP may just shift the equilibrium towards one of the conformations.

Regardless of the nature of signal propagation, the allosteric effect of c-di-GMP binding hints a mechanistic link between the *in vitro* and *in vivo* observations from previous studies and provides an explanation for the mediation of the cellular localization of FimX by c-di-GMP. We propose that the binding of c-di-GMP induces a local conformational change in the region that encompasses most of the $\alpha 4$ - $\beta 5$ - $\alpha 5$ face and the helix linker. Such conformational change could prevent the binding of FimX to its partner, a protein that is putatively localized to the single pole as suggested previously (55, 61). Instead of acting as a phospho-receiving domain, the defective REC domain of FimX would function as an effector domain for the interaction with its partner through the $\alpha 4$ - $\beta 5$ - $\alpha 5$ face and the adjacent helix. Or the $\alpha 4$ - $\beta 5$ - $\alpha 5$ face may be part of the dimerization interface of FimX as proposed by Sondermann et al (86) and the binding to the partner could be solely mediated by the helix linker. In this case, the change in deuteration may also result from a change of the subunit-subunit contact within the dimeric protein, possibly through a rotation of the REC domain. The postulated role of the $\alpha 4$ - $\beta 5$ - $\alpha 5$ face

in hetero- or homo- protein-protein interaction is not unprecedented, considering the common role of the $\alpha 4$ - $\beta 5$ - $\alpha 5$ surface in mediating intra- or inter-protein interaction for CheY-like domain proteins (27, 44, 104, 149). Note that although we did not observe the change of oligomeric structure upon c-di-GMP binding under the experimental conditions, the *in vitro* observations cannot totally rule out the possibility that c-di-GMP binding may affect higher-order oligomer formation in the cells where other protein partners are present and high local FimX concentration may be attained at the pole of the bacterial cell. With the postulated effect of c-di-GMP binding on the interaction between FimX and its protein partners or protein oligomerization, it can be rationalized why the deletion of the EAL domain, or the mutation of the EVL motif, negatively affects the localization of FimX to the single pole for the correct assembly of type IV pili.

Mediation of bacterial motility through the interaction between flagellar proteins and the c-di-GMP binding proteins YcgR and DgrA in *E. coli*, *S. Typhimurium* and *C. crescentus* have been established recently (10, 30, 91, 110). The binding of YcgR and DgrA to their partners is highly dependent on the concentration of the local c-di-GMP pool. The results presented here suggest that c-di-GMP may also act in a similar fashion as an allosteric regulator controlling the binding of FimX to its partner in the mediation of the assembly of type IV pili and twitching motility. In contrast to YcgR and DgrA that contain a PilZ domain for c-di-GMP binding, FimX proteins employ the

non-catalytic EAL domain as the c-di-GMP binding domain. The role of the non-catalytic EAL domain as c-di-GMP binding domain is also reminiscent of the role of the EAL domain of LapD, a *P. fluorescens* protein that senses cellular c-di-GMP level with its cytoplasmic EAL domain and modulates the function of the output domain in the periplasmic space (88). LapD and FimX thus represent an increasing number of EAL domain proteins that have lost the c-di-GMP phosphodiesterase activity and evolved to function as regulatory domains for binding the allosteric regulator c-di-GMP. Given that about 20-25% of the EAL domains encoded by the bacterial genomes are estimated to be catalytically inactive, the use of non-catalytic EAL domain as c-di-GMP binding domain for the regulation of catalytic or binding function of the output domain through an allosteric mechanism may be a rather common phenomenon.

4.5 Summary

FimX is a multidomain signaling protein required for type IV pili biogenesis and twitching motility in the opportunistic pathogen *P. aeruginosa*. FimX is localized to the single pole of the bacterial cell and the unipolar localization is crucial for the correct assembly of type IV pili. FimX contains a non-catalytic EAL domain that lacks the c-di-GMP specific phosphodiesterase activity. It was shown that the deletion of the EAL domain or mutation of the signature EVL motif affects the unipolar localization of FimX. However, it was

not understood how the C-terminal EAL domain could influence protein localization considering that the localization sequence resides in the remote N-terminal region of the protein. By using hydrogen-deuterium (H/D) exchange coupled mass spectrometry, we found that the binding of c-di-GMP to the EAL domain triggers a long-range (*ca.* 70 Å) conformational change in the N-terminal REC domain and adjacent linker. In conjunction with the observation that the mutation of the EVL motif of the EAL domain abolishes the binding for c-di-GMP, the H/D exchange results provide a molecular explanation for the mediation of protein localization and type IV pili biogenesis by the regulator c-di-GMP through a remarkable allosteric regulation mechanism.

CHAPTER 5 CONCLUSIONS AND FUTURE WORKS

5.1 Conclusions

We have selected three GGDEF-EAL didomain-containing proteins and utilized both biochemical and biophysical approaches to probe the functions and mechanisms of these proteins. The results presented in this dissertation have provided insights into the catalytic and regulatory mechanism of the three proteins and shed light on the molecular mechanism of c-di-GMP signaling.

Firstly, the didomain protein PA2567 from *P. aeruginosa* was used as a model to elucidate the catalytic mechanism of EAL domain proteins. Enzymatic assays revealed that PA2567 contains a catalytically active EAL domain but an inactive GGDEF domain. Site-directed mutagenesis in combination with kinetic studies suggests that the essential residue Glu⁵⁴⁸ in PA2567 could assist the deprotonation of the water according to the one-metal mechanism, or coordinate a second Mg²⁺ according to the two-metal mechanism. A conserved loop 6 in EAL domain has been proposed to not only mediate the dimerization of EAL domain to stabilize its quaternary structure but also control c-di-GMP and Mg²⁺ ion binding. Together with the study of RocR, my results are supportive of the Mg²⁺ or Mn²⁺ ion-dependent catalytic mechanism proposed for EAL domains. Importantly, based on the sequence analysis of the 5,862 EAL domains in the bacterial genomes, we categorized EAL domains into three classes to understand the evolution and biological functions of the versatile

EAL domains.

Secondly, we investigated how the cytoplasmic didomain protein AxDGC2 regulates cellulose synthesis in the obligate aerobe *A. xylinum* by controlling the cellular concentration of c-di-GMP. We discovered that the PAS domain of AxDGC2 binds a FAD cofactor non-covalently and the redox status of the FAD cofactor modulates the catalytic activity of the GGDEF domain for c-di-GMP synthesis, with the oxidized form exhibiting higher catalytic activity and apparent substrate inhibition. The results suggest that AxDGC2 is a signaling protein that regulates cellular c-di-GMP level in response to the change of cellular redox status or oxygen concentration. Moreover, several residues predicated to be involved in FAD binding and signal transduction were mutated to examine the impact on redox potential and catalytic activity. However, none of the single mutations was able to completely disrupt the signal transmission to the GGDEF domain, indicating that the change of FAD redox state can still trigger structural changes in the PAS domain probably by using substituted hydrogen-bonded water networks. Meanwhile, although the EAL domain of AxDGC2 was found to be catalytically inactive towards c-di-GMP, it was capable of hydrolyzing some phosphodiester bond-containing substrates. Together with the oxygen-dependent activity of the homologous AxDPEA1, the results provided new insights into the relationships between oxygen level, c-di-GMP concentration and cellulose synthesis in *A. xylinum*.

Thirdly, FimX is a didomain protein from *P. aeruginosa* with both

catalytically inactive GGDEF and EAL domains. We have found that the enzymatically incompetent EAL domain binds c-di-GMP with a dissociate constant of 0.14 μ M. Our hydrogen-deuterium (H/D) exchange study by mass spectrometry on the full-length FimX reveals that the binding of c-di-GMP to the EAL domain triggers an unexpected conformational change in the REC domain and adjacent linker. Alteration of the EVL motif in the EAL domain abolishes the c-di-GMP binding capability, and the loss of c-di-GMP binding capability correlates with the altered cellular localization of FimX. Given the large number of non-enzymatic EAL domains encoded by bacterial genomes, the molecular mechanisms underlying the c-di-GMP induced conformational change and signal propagation may be conserved in other non-enzymatic EAL domains.

Taken together, the three GGDEF-EAL didomain proteins that play vital roles in c-di-GMP signaling serve as model systems for the investigation of the structure-function relationship in these proteins. Study of PA2567 supports our previous proposed catalytic mechanism of EAL domains and establishes the classification scheme of EAL domains encoded by bacterial genomes. Investigation of AxDGC2 reveals the first GGDEF-EAL didomain-containing proteins with the DGC activity regulated by the redox status of the FAD-bound PAS domain. Study of FimX indicates the important roles of the inactive EAL domain in the regulation of protein cellular localization or protein-ligand/protein interaction. Biochemical and biophysical study of the

three GGDEF-EAL didomain proteins described in this dissertation yielded valuable knowledge about the catalytic and regulatory mechanisms of these c-di-GMP signaling proteins. The studies also laid down the foundation for further investigation of the *in vivo* function of these proteins in c-di-GMP signaling.

5.2 Future outlooks

The crystal structures (PDB code: 3HVW) for the GGDEF domain of PA2567 was determined by the Northeast Structural Genomics Consortium. However, the full-length PA2567 structure is still not available. Since there has been no full-length crystal structure of any GGDEF-EAL didomain protein, we would like to determine the structure of the full-length PA2567 protein by collaborating with Dr Julien Lescar's lab. A sitting-drop approach will be used for protein crystallization and the X-ray diffraction will be used to solve the structure. The function of the putative sensor domain GAF of PA2567 remains to be uncovered. A differential scanning fluorimetry (DSF) system has been set up in our lab to screen ligand-protein interactions using a conventional real-time PCR instrument (143). We will investigate the signal sensed by the GAF domain of PA2567 via this method. A compound library is being constructed with a large number of small molecules including nucleotides, amino acids, metabolites and metals ions for ligand screening.

AxDGC2 protein is encoded by one of the three *cdg* operons in the obligate

aerobe *A. xylinum*. Each operon encodes a DGC and a PDE-A protein in tandem. A recent discovery found that two proteins: *EcDosP* and *EcDosC* encoded from the same operon exhibit DGC and PDE-A activity respectively in *E. coli*. These two proteins were found to associate into a functional complex and control the turnover of cellular c-di-GMP in cooperation (139). We will clone and co-express the *AxPDE-A2* protein with *AxDGC2* to investigate the potential interaction between these two proteins. Besides the study of the proteins encoded by *cdg2* operon, we are also interested in the *cdg1* operon. There are two extra genes that encode *Cdg1A* and *Cdg1D* proteins besides the *AxDGC1* and *AxPDEA-1*. Bioinformatic study shows that *Cdg1A* shares high sequence similarity with a transcriptional factor CLP that is found to be a novel c-di-GMP receptor (16). We will clone, express and purify these two proteins and explore their putative regulatory roles.

To further explore the conformational change upon c-di-GMP binding in FimX, we will utilize a thiol-reactive dye to site-specifically label the full-length FimX. Fluorescence spectrometry will be performed to detect any fluorescence intensity change of FimX upon c-di-GMP titration. Our preliminary study also demonstrated that the homolog of FimX (XAC2398) in *Xcc* binds to the PilZ protein (XAC1028) through the EAL domain (45), whereas the binding between *P. aeruginosa* FimX and cognate PilZ seems to be very weak. Hydrogen-deuterium (H/D) exchange coupled mass spectrometry will be performed to investigate the conformational change of XAC2398_{EAL}.

upon binding of XAC1028. In conjugation with site-directed mutagenesis, binding assay will allow us to map out the binding motif in the EAL domain for PilZ interaction, and revealed the molecular basis for the lack of binding affinity between FimX and PilZ in *P. aeruginosa* FimX. Lastly, by comparison of the sequences and structural models, we hope to provide a possible explanation for the divergence of function for FimX in PilZ binding.

REFERENCES

1. Amikam, D., and M. Y. Galperin. 2006. PilZ domain is part of the bacterial c-di-GMP binding protein. *Bioinformatics* 22:3-6.
2. Anderson, S., V. Dragnea, S. Masuda, J. Ybe, K. Moffat, and C. Bauer. 2005. Structure of a Novel Photoreceptor, the BLUF Domain of AppA from *Rhodobacter sphaeroides*. *Biochemistry* 44:7998-8005.
3. Andrade, M. O. 2006. The HD-GYP domain of RpfG mediates a direct linkage between the Rpf quorum-sensing pathway and a subset of diguanylate cyclase proteins in the phytopathogen *Xanthomonas axonopodis pv citri*. *Mol. Microbiol.* 62:537-551.
4. Arnold, K., L. Bordoli, J. Kopp, and T. Schwede. 2006. The SWISS-MODEL Workspace: A web-based environment for protein structure homology modelling. *Bioinformatics* 22:195-201.
5. Barends, T. R. M., E. Hartmann, J. J. Griese, T. Beitlich, N. V. Kirienko, D. A. Ryjenkov, D. A. Reinstein, R. I. Shoeman, M. Gomelsky, and I. Schlichting. 2009. Structure and mechanism of a bacterial light-regulated cyclic nucleotide phosphodiesterase. *Nature* 459:1015-1018.
6. Barry, L. T. 2007. Aer on the inside looking out: paradigm for a PAS-HAMP role in sensing oxygen, redox and energy. *Mol. Microbiol.* 65:1415-1424.
7. Beyhan, S., A. D. Tischler, A. Camilli, and F. H. Yildiz. 2006. Transcriptome and Phenotypic Responses of *Vibrio cholerae* to Increased Cyclic di-GMP Level. *J. Bacteriol.* 188:3600-3613.
8. Bibikov, S. I., L. A. Barnes, Y. Gitin, and J. S. Parkinson. 2000. Domain organization and flavin adenine dinucleotide-binding determinants in the aerotaxis signal transducer Aer of *Escherichia coli*. *Proc. Natl Acad. Sci. USA* 97:5830-5835.
9. Bobrov, A. G., V. Kirillina, and R. D. Perry. 2007. Regulation of biofilm formation in *Yersinia Pestis*. *Adv. Exp. Med. Biol.* 603:201-210.
10. Boehm, A., M. Kaiser, H. Li, C. Spangler, C. A. Kasper, M. Ackermann, V. Kaefer, V. Sourjik, V. Roth, and U. Jenal. Second Messenger-Mediated Adjustment of Bacterial Swimming Velocity. *Cell* 141:107-116.
11. Boles, B. R., and L. L. McCarter. 2002. *Vibrio parahaemolyticus* scrABC, a Novel Operon Affecting Swarming and Capsular Polysaccharide Regulation. *J. Bacteriol.* 184:5946-5954.
12. Buchholz, A., R. Takors, and C. Wandrey. 2001. Quantification of Intracellular metabolites in *Escherichia coli* K12 using liquid chromatographic-electrospray ionization tandem mass spectrometric techniques. *Analytical Biochemistry* 295:129-137.
13. Chan, C., R. Paul, D. Samoray, N. C. Amiot, B. Giese, U. Jenal, and T.

- Schirmer. 2004. Structural basis of activity and allosteric control of diguanylate cyclase. *Proc. Natl. Acad. Sci. USA* 101:17084-9.
14. Chang, A. L., J. R. Tuckerman, G. Gonzalez, R. Mayer, H. Weihouse, G. Volman, D. Amikam, M. Benziman, and M. A. Gilles-Gonzalez. 2001. Phosphodiesterase A1, a regulator of cellulose synthesis in *Acetobacter xylinum*, is a heme-based sensor *Biochemistry* 40:3420-3426.
 15. Chen, J. M., G. J. German, D. C. Alexander, H. Ren, T. Tan, and J. Liu. 2006. Roles of Lsr2 in Colony Morphology and Biofilm Formation of *Mycobacterium smegmatis*. *J. Bacteriol.* 188:633-641.
 16. Chin, K.-H., Y.-C. Lee, Z.-L. Tu, C.-H. Chen, Y.-H. Tseng, J.-M. Yang, R. P. Ryan, Y. McCarthy, J. M. Dow, A. H. J. Wang, and S.-H. Chou. 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.
 17. Cho, H. S., S. Y. Lee, D. L. Yan, X. Y. Pan, J. S. Parkinson, S. Kustu, D. E. Wemmer, and J. G. Pelton. 2000. NMR structure of activated CheY. *J. Mol. Biol.* 297:543-551.
 18. Choy, W.-K., L. Zhou, C. K.-C. Syn, L.-H. Zhang, and S. Swarup. 2004. MorA Defines a New Class of Regulators Affecting Flagellar Development and Biofilm Formation in Diverse *Pseudomonas* Species. *J. Bacteriol.* 186:7221-7228.
 19. Christen, B., M. Christen, R. Paul, F. Schmid, M. Folcher, P. Jenoe, M. Meuwly, and U. Jenal. 2006. Allosteric Control of Cyclic di-GMP Signaling. *J. Biol. Chem.* 281:32015-32024.
 20. Christen, M., B. Christen, M. Folcher, A. Schauerte, and U. Jenal. 2005. Identification and characterization of a cyclic di-GMP-specific phosphodiesterase and its allosteric control by GTP. *J. Biol. Chem.* 280:30829-30837.
 21. Christen, M., H. D. Kulasekara, B. Christen, B. R. Kulasekara, L. R. Hoffman, and S. I. Miller. 2010. Asymmetrical Distribution of the Second Messenger c-di-GMP upon Bacterial Cell Division. *Science* 328:1295-1297.
 22. Christie, J. M., M. Salomon, K. Nozue, M. Wada, and W. R. Briggs. 1999. LOV (light, oxygen, or voltage) domains of the blue-light photoreceptor phototropin (nph1): Binding sites for the chromophore flavin mononucleotide. *Proc. Natl. Acad. Sci. USA* 96:8779-8783.
 23. Crosson, S., S. Rajagopal, and K. Moffat. 2003. The LOV Domain Family: Photoresponsive Signaling Modules Coupled to Diverse Output Domains. *Biochemistry* 42:2-10.
 24. Davidson, V. L. 2000. What Controls the Rates of Interprotein Electron Transfer Reactions. *Acc. Chem. Res.* 33:87-93.
 25. De, N., M. Pirruccello, P. V. Krasteva, N. Bae, R. V. Raghavan, and H. Sondermann. 2008. Phosphorylation-independent regulation of the diguanylate cyclase WspR. *PLoS Biology* 6:e67.

26. Delgado-Nixon, V. M., G. Gonzalez, and M. A. Gilles-Gonzalez. 2000. Dos, a heme-binding PAS protein from *Escherichia coli*, is a direct oxygen sensor. *Biochemistry* 39:2685-2691.
27. Djordjevic, S., P. N. Goudreau, Q. Xu, A. M. Stock, and A. H. West. 1998. Structural basis for methylesterase CheB regulation by a phosphorylation-activated domain. *Proc. Natl. Acad. Sci. USA* 95:1381-1386.
28. Dow, J. M., Y. Fouhy, J. F. Lucey, and R. P. Ryan. 2006. The HD-GYP domain, cyclic di-GMP signaling, and bacterial virulence to plants. *Mol. Plant Microbe Interact.* 19:1378-1384.
29. Duerig, A., S. r. Abel, M. Folcher, M. Nicollier, T. Schwede, N. Amiot, B. Giese, and U. Jenal. 2009. Second messenger-mediated spatiotemporal control of protein degradation regulates bacterial cell cycle progression. *Genes & Dev.* 23:93-104.
30. Fang, X., and M. Gomelsky. A post-translational, c-di-GMP-dependent mechanism regulating flagellar motility. *Mol. Microbiol.* 76:1295-1305.
31. Ferreira, R. B. R., L. C. M. Antunes, E. P. Greenberg, and L. L. McCarter. 2008. *Vibrio parahaemolyticus* ScrC modulates cyclic dimeric GMP regulation of gene expression relevant to growth on surfaces. *J. Bacteriol.* 190:851-860.
32. Fineran, P. C., N. R. Williamson, K. S. Lilley, and G. P. Salmond. 2007. Virulence and prodigiosin antibiotic biosynthesis in *Serratia* are regulated pleiotropically by the GGDEF/EAL domain protein, PigX. *J. Bacteriol.* 189:7653-7662.
33. Frye, J. 2006. Identification of new flagellar genes of *Salmonella enterica* serovar typhimurium. *J. Bacteriol.* 188:2233-2243.
34. Galperin, M. Y. 2006. Structural Classification of Bacterial Response Regulators: Diversity of Output Domains and Domain Combinations. *J. Bacteriol.* 188:4169-4182.
35. Galperin, M. Y., A. N. Nikolskaya, and E. V. Koonin. 2001. Novel domains of the prokaryotic two-component signal transduction systems. *FEMS Microbiol. Lett.* 203:11-21.
36. Garcia, B. 2004. Role of the GGDEF protein family in *Salmonella* cellulose biosynthesis and biofilm formation. *Mol. Microbiol.* 54:264-277.
37. George A. O'Toole, R. K. 1998. Flagellar and twitching motility are necessary for *Pseudomonas aeruginosa* biofilm development. *Molecular Microbiology* 30:295-304.
38. Gilles-Gonzalez, M.-A., and G. gonzalez. 2004. Singal transduction by heme-containing PAS-domain proteins. *J. Appl. Physiol.* 96:774-783.
39. Gilles-Gonzalez, M. A., and G. Gonzalez. 2005. Heme-based sensors: defining characteristics, recent developments, and regulatory hypotheses. *Journal of Inorganic Biochemistry* 99:1-22.
40. Girgrig, H. S., Y. Liu, W. S. Ryu, and S. Tavazoie. 2007. A

- comprehensive genetic characterization of bacterial motility. *PLOS Genet.* 3:1644-1660.
41. Gomelsky, M., and G. Klug. 2002. BLUF: a novel FAD-binding domain involved in sensory transduction in microorganisms. *Trends in Biochemical Sciences* 27:497-500.
 42. Gonzalez, G., E. M. Dioum, C. M. Bertolucci, T. Tomita, M. Ikeda-Saito, M. R. Cheesman, N. J. Watmough, and M.-A. Gilles-Gonzalez. 2002. Nature of the displaceable heme-axial residue in the *EcDos* protein, a heme-based sensor from *Escherichia coli*. *Biochemistry* 41:8414-8421.
 43. Green, J., and M. Pagar. 2004. Bacterial redox sensors. *Nat. Rev. Micro.* 2:954-966.
 44. Guhaniyogi, J., T. Wu, S. S. Patel, and A. M. Stock. 2008. Interaction of CheY with the C-terminal peptide of CheZ. *J. Bacteriol.* 190:1419-1428.
 45. Guzzo, C. R., R. K. Salinas, M. O. Andrade, and C. S. Farah. 2009. PILZ Protein Structure and Interactions with PILB and the FIMX EAL Domain: Implications for Control of Type IV Pilus Biogenesis. *J Mol Biol* 393:848-866.
 46. Hatakeyama, K., T. Harada, and H. Kagamiyama. 1992. IMP dehydrogenase inhibitors reduce intracellular tetrahydrobiopterin levels through reduction of intracellular GTP levels. Indications of the regulation of GTP cyclohydrolase I activity by restriction of GTP availability in the cells. *J. Biol. Chem.* 267:20734-20739.
 47. He, Y.-W., A. Y.-J. Ng, M. Xu, K. Lin, L.-H. Wang, Y.-H. Dong, and L.-H. Zhang. 2007. *Xanthomonas campestris* cell-cell communication involves a putative nucleotide receptor protein Clp and a hierarchical signalling network. *Molecular Microbiology* 64:281-292.
 48. Hecht, G., and A. Newton. 1995. Identification of a novel response regulator required for the swarmer- to-stalked-cell transition in *Caulobacter crescentus*. *J. Bacteriol.* 177:6223-6229.
 49. Hefti, M. H., K. J. Francoijs, S. C. de Vries, R. Dixon, and J. Vervoort. 2004. The PAS fold - A redefinition of the PAS domain based upon structural prediction. *Eur. J. Biochem.* 271:1198-1208.
 50. Hengge, R. 2009. Principles of c-di-GMP signalling in bacteria. *Nat. Rev. Micro.* 7:263-273.
 51. Hickman, J. W., D. F. Tifrea, and C. S. Harwood. 2005. A chemosensory system that regulates biofilm formation through modulation of cyclic diguanylate levels. *Proc. Natl. Acad. Sci. USA* 102:14422-14427.
 52. Hill, S., S. Austin, T. Eydmann, T. Jones, and R. Dixon. 1996. *Azotobacter vinelandii* NIFL is a flavoprotein that modulates transcriptional activation of nitrogen-fixation genes via a redox-sensitive switch. *Proc. Natl Acad. Sci. USA* 93:2143-2148.
 53. Hisert, K. B., M. MacCoss, M. U. Shiloh, K. H. Darwin, S. Singh, R. A. Jones, S. Ehrt, Z. Y. Zhang, B. L. Gaffney, S. Gandotra, D. W. Holden, D. Murray, and C. Nathan. 2005. A glutamate-alanine-leucine (EAL)

- domain protein of *Salmonella* controls bacterial survival in mice, antioxidant defence and killing of macrophages: role of cyclic diGMP. *Mol. Microbiol.* 56:1234-1245.
54. Hoofnagle, A. N., K. A. Resing, and N. G. Ahn. 2003. Protein analysis by hydrogen exchange mass spectrometry. *Annu. Rev. Biophys. Biomol. Struct.* 32:1 - 25.
 55. Huang, B., C. B. Whitchurch, and J. S. Mattick. 2003. FimX, a Multidomain Protein Connecting Environmental Signals to Twitching Motility in *Pseudomonas aeruginosa*. *J. Bacteriol.* 185:7068-7076.
 56. Hunt, S. M., M. Elvin, S. K. Crosthwaite, and C. Heintzen. 2007. The PAS/LOV protein VIVID controls temperature compensation of circadian clock phase and development in *Neurospora crassa*. *Genes Dev.* 21:1964-1974.
 57. Hurley, J. H. 2003. GAF domains: cyclic nucleotides come full circle. *Sci. STKE* 164:pe1.
 58. Igarashi, K., M. F. J. M. Verhagen, M. Samejima, M. Schulein, K.-E. L. Eriksson, and T. Nishino. 1999. Cellobiose dehydrogenase from the fungi *Phanerochaete chrysosporium* and *Humicola insolens*. A flavochemoprotein from *Humicola insolens* contains 6-hydroxy-FAD as the dominant active cofactor. *J. Biol. Chem.* 274:3338-3344.
 59. Jenal, U., and J. Malone. 2006. Mechanisms of cyclic-di-GMP signaling in bacteria. *Annu. Rev. Genet.* 40:385-407.
 60. Jung, A., T. Domratcheva, M. Tarutina, Q. Wu, W. H. Ko, R. L. Shoeman, M. Gomelsky, K. H. Gardner, and L. Schlichting. 2005. Structure of a bacterial BLUF photoreceptor: Insights into blue light-mediated signal transduction. *Proc. Natl. Acad. Sci. USA* 102:12350-12355.
 61. Kazmierczak, B. I., M. B. Lebron, and T. S. Murray. 2006. Analysis of fimX, a phosphodiesterase that governs twitching motility in *Pseudomonas aeruginosa*. *Mol. Microbiol.* 60:1026-1043.
 62. Key, J., M. Hefti, E. B. Purcell, and K. Moffat. 2007. Structure of the redox sensor domain of *Azotobacter vinelandii* NifL at atomic resolution: signaling, dimerization, and mechanism. *Biochemistry* 46:3614-3623.
 63. Kim, Y.-K., and L. L. McCarter. 2007. ScrG, a GGDEF-EAL Protein, Participates in Regulating Swarming and Sticking in *Vibrio parahaemolyticus*. *J. Bacteriol.* 189:4094-4107.
 64. Kirillina, O., J. D. Fetherstone, A. G. Bobrov, J. Abney, and R. D. Perry. 2004. HmsP, a putative phosphodiesterase, and HmsT, a putative diguanylate cyclase, control Hms-dependent biofilm formation in *Yersinia pestis*. *Mol. Microbiol.* 54:75-88.
 65. Klopprogge, K., and R. A. Schmitz. 1999. NifL of *Klebsiella pneumoniae*: Redox characterization in relation to the nitrogen source. *Biochim. Biophys. Acta.* 1431:462-470.
 66. Kuchma, S. L. 2007. BifA, a cyclic-di-GMP phosphodiesterase,

- inversely regulates biofilm formation and swarming motility by *Pseudomonas aeruginosa* PA14. J. Bacteriol. 189:8165-8178.
67. Kulasakara, H. 2006. Analysis of *Pseudomonas aeruginosa* diguanylate cyclases and phosphodiesterase reveals a role for bis-(3'-5')-cyclic-GMP in virulence. Proc. Natl Acad. Sci. USA 103:2839-2844.
 68. Kulasekara, H. D., I. Ventre, B. R. Kulasekara, A. Lazdunski, A. Filloux, and S. Lory. 2005. A novel two-component system controls the expression of *Pseudomonas aeruginosa* fimbrial cup genes. Mol. Microbiol. 55:368-380.
 69. Kumar, M., and D. Chatterji. 2008. Cyclic-di-GMP: a second messenger required for long-term survival, but not for biofilm formation, in *Mycobacterium smegmatis*. Microbiology 154:2942-2955.
 70. Lee, T., A. N. Hoofnagle, K. A. Resing, and N. G. Ahn. 2006. Protein Hydrogen Exchange Measured by Electrospray Ionization Mass Spectrometry, p. 443-449. In E. C. Julio (ed.), Cell Biology (Third Edition). Academic Press, Burlington.
 71. Liang, Z. X., T. Lee, K. A. Resing, N. G. Ahn, and J. P. Klinman. 2004. Thermal-activated protein mobility and its correlation with catalysis in thermophilic alcohol dehydrogenase. Proc Natl Acad Sci U S A 101:9556-9561.
 72. Liang, Z. X., J. M. Nocek, K. Huang, R. T. Hayes, I. V. Kurnikov, D. N. Beratan, and B. M. Hoffman. 2002. Dynamic docking and electron transfer between Zn-myoglobin and cytochrome b(5). J. Am. Chem. Soc. 124:6849-6859.
 73. Liang, Z. X., J. M. Nocek, I. V. Kurnikov, D. N. Beratan, and B. M. Hoffman. 2000. Electrostatic control of electron transfer between myoglobin and cytochrome b(5): Effect of methylating the heme propionates of Zn-myoglobin. J. Am. Chem. Soc. 122:3552-3553.
 74. Lim, B., S. Beyhan, J. Meir, and F. H. Yildiz. 2006. Cyclic-diGMP signal transduction systems in *Vibrio cholerae*: modulation of rugosity and biofilm formation. Mol. Microbiol. 60:331-348.
 75. Lim, B., S. Beyhan, and F. H. Yildiz. 2007. Regulation of *Vibrio polysaccharide* synthesis and virulence factor production by CdgC, a GGDEF-EAL domain protein, in *Vibrio cholerae*. Journal of Bacteriology 189:717-729.
 76. Lu, X., D. Nikolic, D. J. Mitchell, R. B. van Breemen, J. A. Mersfelder, R. Hille, and R. B. Silverman. 2003. A mechanism for substrate-Induced formation of 6-Hydroxyflavin mononucleotide catalyzed by C30A trimethylamine dehydrogenase. Bioorg. Med. Chem. Lett. 13:4129-4132.
 77. Macheroux, P., S. Hill, S. Austin, T. Eydmann, T. Jones, S. O. Kim, R. Poole, and R. Dixon. 1998. Electron donation to the flavoprotein NifL, a redox-sensing transcriptional regulator. Biochem. J. 332:413-419.
 78. Marshall, K. R., M. Gong, L. Wodke, J. H. Lamb, D. J. L. Jones, P. B.

- Farmer, N. S. Scrutton, and A. W. Munro. 2005. The human apoptosis-inducing protein AMID Is an oxidoreductase with a modified flavin cofactor and DNA binding activity. *J. Biol. Chem.* 280:30735-30740.
79. Martinez, S. E., A. Y. Wu, N. A. Glavas, X.-B. Tang, S. Turley, W. G. J. Hol, and J. A. Beavo. 2002. The two GAF domains in phosphodiesterase 2A have distinct roles in dimerization and in cGMP binding. *Proc. Natl. Acad. Sci. USA* 99:13260-13265.
80. Masuda, S., K. Hasegawa, A. Ishii, and T. a. Ono. 2004. Light-Induced Structural Changes in a Putative Blue-Light Receptor with a Novel FAD Binding Fold Sensor of Blue-Light Using FAD (BLUF); Slr1694 of *Synechocystis* sp. PCC6803. *Biochemistry* 43:5304-5313.
81. Mattick, J. S. 2002. TYPE IV PILI AND TWITCHING MOTILITY. *Annual review of Microbiology* 56:289-314.
82. McWhirter, S. M., R. Barbalat, K. M. Monroe, M. F. Fontana, M. Hyodo, N. T. Joncker, K. J. Ishii, S. Akira, M. Colonna, Z. J. Chen, K. A. Fitzgerald, Y. Hayakawa, and R. E. Vance. 2009. A host type I interferon response is induced by cytosolic sensing of the bacterial second messenger cyclic-di-GMP. *Journal of Experimental Medicine* 206:1899-1911.
83. Meissner, A., V. Wild, R. Simm, M. Rohde, C. Erck, F. Bredenbruch, M. Morr, U. Romling, and S. Haussler. 2007. *Pseudomonas aeruginosa* cupA-encoded fimbriae expression is regulated by a GGDEF and EAL domain-dependent modulation of the interacellular level of cyclic diguanylate. *Environ. Microb.* 9:2475-2485.
84. Minasov, G., S. Padavattan, L. Shuvalova, J. S. Brunzelle, D. J. Miller, A. Basle, C. Massa, F. R. Collart, T. Schirmer, and W. F. Anderson. 2009. Crystal structures of YkuI and its complex with second messenger c-di-GMP suggests catalytic mechanism of phosphodiester bond cleavage by EAL domains. *J. Biol. Chem.*:M808221200.
85. Moglich, A., and K. Moffat. 2007. Structural basis for light-dependent signaling in the dimeric LOV domain of the photosensor YtvA. *J. Mol. Biol.* 373:112-126.
86. Navarro, M. V. A. S., D. N., N. Bae, Q. Wang, and H. Sondermann. 2009. Structural Analysis of the GGDEF-EAL Domain-Containing c-di-GMP Receptor FimX. *Structure* 17:1104-1116.
87. Negri, A., V. Massey, and C. H. Williams, Jr. 1987. D-aspartate oxidase from beef kidney. Purification and properties. *J. Biol. Chem.* 262:10026-10034.
88. Newell, P. D., R. D. Monds, and G. A. O'Toole. 2009. LapD is a bis-(3',5')-cyclic dimeric GMP-binding protein that regulates surface attachment by *Pseudomonas fluorescens* Pf0-1. *Proc. Natl. Acad. Sci. USA*:ASAP online article.
89. Nocek, J. M., J. S. Zhou, S. D. Forest, S. Priyadarshy, D. N. Beratan, J.

- N. Onuchic, and B. M. Hoffman. 1996. Theory and Practice of Electron Transfer within Protein-Protein Complexes: Application to the Multidomain Binding of Cytochrome c by Cytochrome c Peroxidase. *Chem. Rev.* 96:2459-2489.
90. Ogunniyi, A. D., J. C. Paton, A. C. Kirby, J. A. McCullers, J. Cook, M. Hyodo, Y. Hayakawa, and D. K. R. Karaolis. 2008. c-di-GMP is an effective immunomodulator and vaccine adjuvant against pneumococcal infection. *Vaccine* 26:4676-4685.
 91. Paul, K., V. Nieto, W. C. Carlquist, D. F. Blair, and R. M. Harshey. The c-di-GMP Binding Protein YcgR Controls Flagellar Motor Direction and Speed to Affect Chemotaxis by a "Backstop Brake" Mechanism. *Mol. Cell.* 38:128-139.
 92. Paul, R. 2007. Activation of the diguanylate cyclase PleD by phosphorylation-mediated dimerization. *J. Biol. Chem.* 282:29170-29177.
 93. Paul, R. 2004. Cell cycle-dependent dynamic localization of a bacterial response regulator with a novel di-guanylate cyclase output domain. *Genes Dev.* 18:715-727.
 94. Paul, R., S. Weiser, N. C. Amiot, C. Chan, T. Schirmer, B. Giese, and U. Jenal. 2004. Cell cycle-dependent dynamic localization of a bacterial response regulator with a novel di-guanylate cyclase output domain. *Genes & Development* 18:715-727.
 95. Qi, Y., F. Rao, and Z.-X. Liang. 2009. A flavin-binding PAS domain regulates c-di-GMP synthesis in AxDGC2 from *Acetobacter xylinum*. *Biochemistry* 48:10275-10285.
 96. Rao, F., S. Pasunooti, Y. Ng, W. Zhuo, L. Lim, W. Liu, and Z.-X. Liang. 2009. Enzymatic synthesis of c-di-GMP using a thermophilic diguanylate cyclase. *Anal. Biochem.* 389:138-142.
 97. Rao, F., Y. Qi, H. S. Chong, M. Kotada, B. Li, J. Lescar, K. Tang, and Z.-X. Liang. 2009. The functional role of a conserved loop in EAL domain-based c-di-GMP specific phosphodiesterase. *J. Bacteriol.* 191:4722-4731.
 98. Rao, F., R. Y. See, D. Zhang, D. C. Toh, Q. Ji, and Z.-X. Liang. 2010. YybT Is a Signaling Protein That Contains a Cyclic Dinucleotide Phosphodiesterase Domain and a GGDEF Domain with ATPase Activity. *Journal of Biological Chemistry* 285:473-482.
 99. Rao, F., Y. Yang, Y. Qi, and Z. X. Liang. 2008. Catalytic mechanism of c-di-GMP specific phosphodiesterase: a study of the EAL domain-containing RocR from *Pseudomonas aeruginosa*. *J. Bacteriol.* 190:3622-3631.
 100. Rebbapragada, A., M. S. Johnson, G. P. Harding, A. J. Zuccarelli, H. M. Fletcher, I. B. Zhulin, and B. L. Taylor. 1997. The Aer protein and the serine chemoreceptor Tsr independently sense intracellular energy levels and transduce oxygen, redox, and energy signals for *Escherichia coli*

- behavior. *Proc. Natl Acad. Sci. USA* 94:10541-10546.
101. Repik, A., A. Rebbapragada, M. S. Johnson, J. Q. Haznedar, I. B. Zhulin, and B. L. Taylor. 2000. PAS domain residues involved in signal transduction by the Aer redox sensor of *Escherichia coli*. *Mol. Microbiol.* 36:806-816.
 102. Resing, K. A., A. N. Hoofnagle, and N. G. Ahn. 1999. Modeling deuterium exchange behavior of ERK2 using pepsin mapping to probe secondary structure. *Am. Soc. Mass. Spec.* 10:685-702.
 103. Robinson, V. L., D. R. Buckler, and A. M. Stock. 2000. A tale of two components: a novel kinase and a regulatory switch. *Nature Structural Biology* 7:626-633.
 104. Robinson, V. L., T. Wu, and A. M. Stock. 2003. Structural Analysis of the Domain Interface in DrrB, a Response Regulator of the OmpR/PhoB Subfamily. *J. Bacteriol.* 185:4186-4194.
 105. Romling, U., and D. Amikam. 2006. Cyclic di-GMP as a second messenger. *Curr. Opin. Microbiol.* 9:218-228.
 106. Romling, U., M. Gomelsky, and M. Y. Galperin. 2005. C-di-GMP: the dawning of a novel bacterial signalling system. *Mol. Microbiol.* 57:629-639.
 107. Ross, P. 1987. Regulation of cellulose synthesis in *Acetobacter xylinum* by cyclic diguanylate. *Nature* 325:279-281.
 108. Ross, P., R. Mayer, H. Weinhouse, D. Amikam, Y. Huggirat, M. Benziman, E. de Vroom, A. Fiddler, P. de Paus, and L. A. Sliedregt. 1990. The cyclic diguanylic acid regulatory system of cellulose synthesis in *Acetobacter xylinum*. Chemical synthesis and biological activity of cyclic nucleotide dimer, trimer, and phosphothioate derivatives. *J. Biol. Chem.* 265:18933-18943.
 109. Ryan, R. P. 2006. Cell-cell signaling in *Xanthomonas campestris* involves an HD-GYP domain protein that functions in cyclic di-GMP turnover. *Proc. Natl Acad. Sci. USA* 103:6712-6717.
 110. Ryjenkov, D. A., R. Simm, U. Romling, and M. Gomelsky. 2006. The PilZ domain is a receptor for the second messenger c-di-GMP - The PilZ domain protein YcgR controls motility in enterobacteria. *J. Biol. Chem.* 281:30310-30314.
 111. Ryjenkov, D. A., R. Simm, U. Romling, and M. Gomelsky. 2006. The PilZ domain is a receptor for the second messenger c-di-GMP: the PilZ protein YcgR controls motility in enterobacteria. *J. Biol. Chem.* 281:30310-30314.
 112. Ryjenkov, D. A., M. Tarutina, O. V. Moskvina, and M. Gomelsky. 2005. Cyclic diguanylate is a ubiquitous signaling molecule in bacteria: Insights into biochemistry of the GGDEF protein domain. *J. Bacteriol.* 187:1792-1798.
 113. Sasakura, Y., T. Yoshimura-Suzuki, H. Kurokawa, and T. Shimizu. 2006. Structure and Function Relationships of EcDOS, a Heme-Regulated

- Phosphodiesterase from *Escherichia coli*. Accounts of Chemical Research 39:37-43.
114. Schirmer, T., and U. Jenal. 2009. Structural and mechanistic determinants of c-di-GMP signalling. Nat Rev Micro 7:724-735.
 115. Schmidt, A. J., D. A. Ryjenkov, and M. Gomelsky. 2005. The ubiquitous protein domain EAL is a cyclic diguanylate-specific phosphodiesterase: Enzymatically active and inactive EAL domains. J. Bacteriol. 187:4774-4781.
 116. Schwede, T., J. Kopp, N. Guex, and M. C. Peitsch. 2003. SWISS-MODEL: an automated protein homology-modeling server. Nucl. Acids. Res. 31:3381-3385.
 117. Simm, R., A. Lusch, A. Kader, M. Andersson, and U. Romling. 2007. Role of EAL-Containing Proteins in Multicellular Behavior of *Salmonella enterica* Serovar Typhimurium. J. Bacteriol. 189:3613-3623.
 118. Simm, R., M. Morr, A. Kader, M. Nimtz, and U. Romling. 2004. GGDEF and EAL domains inversely regulate cyclic-di-GMP levels and transition from sessility to motility. Mol. Microbiol. 53:1123-1134.
 119. Söderbäck, E., F. Reyes-Ramirez, T. Eydmann, S. Austin, S. Hill, and R. Dixon. 1998. The redox- and fixed nitrogen-responsive regulatory protein NIFL from *Azotobacter vinelandii* comprises discrete flavin and nucleotide-binding domains. Mol. Microbiol. 28:179-192.
 120. Sommerfeldt, N. 2009. Gene expression patterns and differential input into curli fimbriae regulation of all GGDEF/EAL domain proteins in *Escherichia coli*. Microbiology 155:1318-1331.
 121. Sterner, R., and B. Hocker. 2005. Catalytic versatility, stability and evolution of the (b/a)₈-barrel enzyme fold. Chem. Rev. 105:4038-4055.
 122. Sudarsan, N., E. R. Lee, Z. Weinberg, R. H. Moy, J. N. Kim, K. H. Link, and R. R. Breaker. 2008. Riboswitches in eubacteria sense the second messenger cyclic di-GMP. Science 321:411-413.
 123. Sukumar, N., A. R. Dewanti, B. Mitra, and F. S. Mathews. 2004. High resolution structures of an oxidized and reduced flavoprotein: THE WATER SWITCH IN A SOLUBLE FORM OF (S)-MANDELATE DEHYDROGENASE. J. Biol. Chem. 279:3749-3757.
 124. Suquet, C., M. Savenkova, and J. D. Satterlee. 2005. Recombinant PAS-heme domains of oxygen sensing proteins: High level production and physical characterization. Protein Expression and Purification 42:182-193.
 125. Suzuki, K., P. Babitzke, S. R. Kushner, and T. Romeo. 2006. Identification of a novel regulatory protein (CsrD) that targets the global regulatory RNAs CsrB and CsrC for degradation by RNase E. Genes Dev. 20:2605-2617.
 126. Tal, R., H. C. Wong, R. Calhoon, D. Gelfand, A. L. Fear, G. Volman, R. Mayer, P. Ross, D. Amikam, H. Weinhouse, A. Cohen, S. Sapir, P. Ohana, and M. Benziman. 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.
127. Tamayo, R., J. T. Pratt, and A. Camilli. 2007. Roles of cyclic diguanylate in the regulation of bacterial pathogenesis. Annu. Rev. Microbiol. 61:131-148.
 128. Tamayo, R., S. Schild, J. T. Pratt, and A. Camilli. 2008. Role of Cyclic Di-GMP during El Tor Biotype *Vibrio cholerae* Infection: Characterization of the In Vivo-Induced Cyclic Di-GMP Phosphodiesterase CdpA. Infect. Immun. 76:1617-1627.
 129. Tamayo, R., A. D. Tischler, and A. Camilli. 2005. The EAL domain protein VieA is a cyclic diguanylate phosphodiesterase. J. Biol. Chem. 280:33324-33330.
 130. Tanaka, A., H. Takahashi, and T. Shimizu. 2007. Critical role of the heme axial ligand, Met95, in locking catalysis of the phosphodiesterase from *Escherichia coli* (Ec DOS) toward cyclic diGMP. J. Biol. Chem. 282:21301-21307.
 131. Tarutina, M., D. A. Ryjenkov, and M. Gomelsky. 2006. An Unorthodox Bacteriophytochrome from *Rhodobacter sphaeroides* Involved in Turnover of the Second Messenger c-di-GMP. J. Biol. Chem. 281:34751-34758.
 132. Taylor, B. L., and I. B. Zhulin. 1999. PAS domains: Internal sensors of oxygen, redox potential, and light. Microbiol. Mol. Biol. Rev. 63:479.
 133. Tchigvintsev, A., X. Xu, A. Singer, C. Chang, G. Brown, M. Proudfoot, H. Cui, R. Flick, W. F. Anderson, A. Joachimiak, M. Y. Galperin, A. Savchenko, and A. F. Yakunin. Structural Insight into the Mechanism of c-di-GMP Hydrolysis by EAL Domain Phosphodiesterases. Journal of Molecular Biology In Press, Uncorrected Proof.
 134. Thormann, K. M. 2006. Control of formation and cellular detachment from *Shewanella oneidensis* MR-1 biofilms by cyclic di-GMP. J. Bacteriol. 188:2681-2691.
 135. Thummer, R., O. Klimmek, and R. A. Schmitz. 2007. Biochemical Studies of *Klebsiella pneumoniae* NifL Reduction Using Reconstituted Partial Anaerobic Respiratory Chains of *Wolinella succinogenes*. J. Biol. Chem. 282:12517-12526.
 136. Tischler, A. D., and A. Camilli. 2004. Cyclic diguanylate (c-di-GMP) regulates *Vibrio cholerae* biofilm formation. Mol. Microbiol. 53:857-869.
 137. Tischler, A. D., and A. Camilli. 2005. Cyclic diguanylate regulates *Vibrio cholerae* virulence gene expression. Infection and Immunity 73:5873-5882.
 138. Tschowri, N., S. Busse, and R. Hengge. 2009. The BLUF-EAL protein YcgF acts as a direct anti-repressor in a blue-light response of *Escherichia coli*. Genes & Dev. 23:522-534.

139. Tuckerman, J. R., G. Gonzalez, E. H. S. Sousa, X. Wan, S. J. A., M. Alam, and M. A. Gilles-Gonzalez. 2009. An oxygen-sensing diguanylate cyclase and phosphodiesterase couple for c-di-GMP control. *Biochemistry* DOI: 10.1021/bi901409g
140. Uchida, T., E. Sato, A. Sato, I. Sagami, T. Shimizu, and T. Kitagawa. 2005. CO-dependent activity-controlling mechanism of heme-containing CO-sensor protein, neuronal PAS domain protein. *Journal of Biological Chemistry* 280:21358-21368.
141. Ukaegbu, U. E., S. Henery, and A. C. Rosenzweig. 2006. Biochemical Characterization of MmoS, a Sensor Protein Involved in Copper-Dependent Regulation of Soluble Methane Monooxygenase. *Biochemistry* 45:10191-10198.
142. Ukaegbu, U. E., and A. c. Rosenzweig. 2009. Structure of the redox sensor domain of *Methylococcus capsulatus* (Bath) MmoS. *Biochemistry* 48:2207-2215.
143. Vedadi, M., F. H. Niesen, A. Allali-Hassani, O. Y. Fedorov, P. J. Finerty, G. A. Wasney, R. Yeung, C. Arrowsmith, L. J. Ball, H. Berglund, R. Hui, B. D. Marsden, P. r. Nordlund, M. Sundstrom, J. Weigelt, and A. M. Edwards. 2006. Chemical screening methods to identify ligands that promote protein stability, protein crystallization, and structure determination. *Proc. Natl. Acad. Sci. USA* 103:15835-15840.
144. Wan, X., J. R. Tuckerman, J. A. Saito, T. A. K. Freitas, J. S. Newhouse, J. R. Denery, M. Y. Galperin, G. Gonzalez, M.-A. Gilles-Gonzalez, and M. Alam. 2009. Globins synthesize the second messenger bis-(3'-5')-Cyclic diguanosine monophosphate in bacteria. *J. Mol. Biol.* 388:262-270.
145. Wassmann, P. 2007. Structure of BeF₃-modified response regulator PleD: implications of diguanylate cyclase activation, catalysis, and feedback inhibition. *Structure (Camb.)* 15:915-927.
146. Weber, H., C. Persavento, G. Possling, G. Tischendorf, and R. Hengge. 2006. Cyclic-di-GMP-mediated signalling within the sigma(S) network of *Escherichia coli*. *Mol. Microbiol.* 62:1014-1034.
147. Weis, D. D., I. J. Kass, and J. R. Engen. 2006. Semi-automated analysis of hydrogen exchange mass spectra using HX-Express. *J. Am. Soc. Mass. Spec.* 17:1700-1703.
148. West, A. H., and A. M. Stock. 2001. Histidine kinases and response regulator proteins in two-component signaling systems. *Trends in Biochemical Sciences* 26:369-376.
149. Ye, S., I. Vakonakis, T. R. Ioerger, A. C. LiWang, and J. C. Sacchettini. 2004. Crystal Structure of Circadian Clock Protein KaiA from *Synechococcus elongatus*. *Journal of Biological Chemistry* 279:20511-20518.
150. Zanetti, G., V. Massey, and B. Curti. 1983. FAD analogues as mechanistic and binding-domain probes of spinach ferredoxin-NADP⁺

reductase. Eur. J. Biochem. 132:201-205.

PUBLICATION LIST

- 1 **YN Qi**, ML Chuah, XM Dong, KL Xie, Z Luo, K Tang, ZX Liang. 'Binding of c-di-GMP in the non-catalytic EAL domain of FimX induces a long-range conformational change.' *Journal of Biological Chemistry* 2011 Jan 28; 286(4):2910-7.
- 2 **YN Qi**, F Rao, Z Luo, ZX Liang. 'A flavin cofactor-binding PAS domain regulates c-di-GMP synthesis in AxDGC2 from *Acetobacter xylinum*'. *Biochemistry*. 2009 Nov 3; 48(43): 10275-85.
- 3 F Rao, **YN Qi**, E Murugan, Pasunooti S, Q Ji. '2', 3'-cAMP hydrolysis by metal-dependent phosphodiesterases containing DHH, EAL, and HD domains is non-specific: Implication for PDE screening.' *Biochem Biophys Res Commun*. 2010 Jul 30; 398(3): 500-5.
- 4 F Rao, **YN Qi**, HS Chong, M Kotaka, B Li, J Li, J Lescar, K Tang, ZX Liang. 'The functional role of a conserved loop in EAL domain-based cyclic di-GMP-specific phosphodiesterase'. *Journal of Bacteriology*. 2009 Aug; 191(15): 4722-31.
- 5 F Rao, Y Yang, **YN Qi**, ZX Liang. 'Catalytic mechanism of cyclic di-GMP-specific phosphodiesterase: a study of the EAL domain-containing RocR from *Pseudomonas aeruginosa*'. *Journal of Bacteriology*. 2008 May; 190(10): 3622-31.