Catalysis and regulation of cyclic dinucleotide specific phosphodiesterases

Rao, Feng

2010

Rao, F. (2010). Catalysis and regulation of cyclic dinucleotide specific phosphodiesterases. Doctoral thesis, Nanyang Technological University, Singapore.

https://hdl.handle.net/10356/43545

https://doi.org/10.32657/10356/43545

CATALYSIS AND REGULATION OF CYCLIC DINUCLEOTIDE SPECIFIC PHOSPHODIESTERASES

RAO FENG

SCHOOL OF BIOLOGICAL SCIENCES

NANYANG TECHNOLOGICAL UNIVERSITY

A thesis submitted to the Nanyang Technological University

in fulfillment of the requirements for the degree of

Doctor of Philosophy

2010

ACKNOWLEDGEMENTS

First of all, I would like to express my sincere gratitude to my supervisor, Dr. Liang Zhao-Xun. This thesis would not be completed without his generous support, encouragement, and careful scientific guidance. Special gratitude also goes to my mentor Dr. Susana Geifman, her generous support and guidance is an integral part of this PhD journey.

I wish to extend my sincere appreciation to our collaborators who have helped me with experiments along the way: Drs. Tang Kai and Li Bing for MS analysis; Dr. Konstantin Pervushin and Mr. Edward Tan for NMR structure determination; Drs. Julien Lescar's and Tilman Schirmer's lab for x-ray crystallography of RocR. Drs. Mu Yuguang, Yang Ye, and Lin Xin for computational docking; Dr. Li Jinming for bioinformatic analysis; and Mr Chia Wei Sheng for help to read through the thesis.

I would like to thank all members in our lab: Chuah Lay Cheng Mary, Ho Chun Loong Lawrence, Ho Qin Shi, Ji Qiang, Kong Rong, Liew Chong Wai, Elavazhagan Murugan, Qi Yaning, Swathi Pasunooti, Sun Huihua, and Zhang Dongwei, for always being there to help me. The days we spent together in the lab are memorable. Thanks also go to the FYP and attachment students Tok Yan Ling, Chong Hui Shan, Angeline Liu, Lim Li Shi, Goh Lan Pei, See Rui Yin, Delon Toh, and Ishin Soehano for their efforts and help.

I would also like to thank Dr. K Kobayashi for sharing the bacillus bacterial strains; Dr. K. Fukuyama for sharing protein expression plasmids; Dr. Peter Droge for explaining the process of DNA replication genetics; Drs. Peter Chung, Yoon Ho Sup, Choi Bo Hwa, and Wu Bin for sharing reagents.

Last but not least, I am greatly indebted to my whole family: my grandparents, my parents, my brother, my sister, my wife, and my daughter. The love and support I received from them bring me courage to complete this journey.

Table of Contents

SUMMARY

Over the last several years, cyclic-di-GMP (c-di-GMP) has been firmly established as a major bacterial second messenger, with proteins containing GGDEF, EAL and HD-GYP domains catalyzing its synthesis or degradation. Meanwhile, the existence of the structural analog c-di-AMP in prokaryotes has also come to light.

In this dissertation, I first present data to show that RocR, an EAL domain response regulator from *P. aeruginosa*, catalyzes the hydrolysis of c-di-GMP, but not c-di-AMP. Using a combination of site-directed mutagenesis and computational docking, specific functions were assigned for the conserved residues in the active site, which include five essential residues for Mg^{2+} binding and a sixth essential residue (E^{352}) that may function as a general-base catalyst, assisting the deprotonation of a Mg^{2+}-coordinated water to generate the nucleophilic hydroxide ion. This study provides an explanation at molecular level for the lack of catalytic activity for about 20 percent of the EAL domains encoded by the bacterial genomes.

The EAL domains belong to a family of $(\beta/\alpha)_8$ barrel fold enzymes that contain a functional active site loop (loop 6) for substrate binding and catalysis. I found that mutations in the loop 6 region of RocR significantly altered the catalytic activity of the EAL domain. The impact of the mutations ranges from apparent substrate inhibition to alteration of oligomeric structure. Meanwhile, mutation of the putative phosphorylation site ($D^{56}\rightarrow N^{56}$) located in the receiver domain lowered the K_m for c-di-GMP by 15-fold, thereby allowing much more efficient catalysis. Hydrogen-deuterium (H/D) exchange by mass spectrometry revealed that the decrease in K_m correlates with a change of solvent accessibility in the loop 6 region. I further examined the *Ax*DGC2 protein from *A. xylinus*, which contains a catalytically incompetent EAL domain with a highly degenerate loop 6. The catalytic activity of the stand-alone EAL domain of DGC2 was recovered by restoring loop 6. In conjunction

with the structural data of two EAL domains, I propose that loop 6 not only mediates the dimerization of EAL domain, but also controls c-di-GMP and Mg^{2+} ion binding. Thus, loop 6 may play an important role in the regulation of the activity of the EAL domain. Importantly, sequence analysis of the 5862 EAL domains in the bacterial genomes revealed that a large number of the EAL domains harbor a degenerate loop 6 and may be catalytically incompetent.

In the second half of the dissertation, I present data demonstrating that YybT, a DHH/DHHA1 domain signaling protein from *B. subtilis*, hydrolyzes c-di-AMP and c-di-GMP to generate the linear dinucleotides 5'-pApA and 5'-pGpG *in vitro*. The data suggest c-di-AMP could be the physiological substrate for YybT, given the physiologically relevant Michaelis-Menten constant (K_m) and the presence of YybT family proteins in bacteria lacking c-di-GMP signaling network. The stringent response alarmone ppGpp was found to be a strong competitive inhibitor of the DHH/DHHA1 domain, suggesting a direct link between ppGpp and c-di-AMP signaling. In addition, the atypical GGDEF domain of YybT exhibits unexpected ATPase activity, distinct from the common diguanylate cyclase activity for GGDEF domains.

Lastly, I unexpectedly found that the Per-Arnt-Sim (PAS) domains of YybT family proteins contain a heme binding site, despite the absence of any obvious endogenous ligand for heme coordination. The ferrous heme iron of YybT proteins forms six-coordinated complex with carbon monoxide (CO), but penta-coordinated complex with nitric oxide (NO). The binding of the heme suppresses the phosphodiesterase activity of the C-terminal DHH/DHHA1 domain by reducing the k_{cat}/K_m by 267-fold, whereas binding of NO to the ferrous heme YybT stimulates the PDE activity by 3-fold. In vivo studies showed that the *B. subtilis* Δ*yybT* mutant exhibits more than 100-fold lower survival rate when challenged with H_2O_2 stress, while NO is known to play an important role in combating H_2O_2 mediated cell

killing. The results reveal a distinct heme-binding PAS domain and suggest that the PAS domain may function as regulatory domain by sensing heme or NO in stress signaling.

List of Figures

List of Tables

Abbreviations

2'3'-cAMP	2'3'-cyclic adenosine monophosphate
3'-AMP	3'-adenosine monophosphate
5'-pGpG	5'-linear diguanylic acid
aa	Amino acids
ACN	Acetonitrile
AMPPNP	adenylyl imidodiphosphate, a nonydrolyzable ATP analog
A-site	Activity site
bis-pNPP	bis-*p*-nitrophenyl phosphate
BLUF	Blue-light photoreceptor domain
β-ME	β- mercaptoethanol
ca.	Circa: approximately
c-di-GMP	Cyclic-di-guanosine monophosphate
c-di-AMP	Cyclic-di-adenosine monophosphate
cGMP	Cyclic guanosine monophosphate
CH	Casein hydrolysate
CORM2	Tricarbonyldichlororuthenium (II) dimer
DAC	Di-adenylate cyclase
DGC	Di-guanylate cyclase
DTH	Sodium Dithionite
DTT	Dithiothreitol
dsRNA	double-stranded RNA
EAL	"Glu-ala-leu" motif containing domain harboring c-di-GMP PDE activity
EDTA	Ethylenediaminetetraacetic acid
ESI-MS	Electrospray ionization-mass spectrometry
FA	Formic acid
FAD	Flavin adenine dinucleotide
FMN	Flavin mononucleotide
FPLC	Fast performance liquid chromatography
GGDEF	"Gly-gly-asp-glu-phe" motif containing domain harboring DGC activity
GMP	Guanosine monophosphate
GTP	Guanosine triphosphate
HAMP	Histidine kinases, adenylyl cyclases, methyl binding proteins and phosphatases
H/D	Hydrogen/deuterium
HD-GYP	Histidine-aspartic acid-glycine-tyrosine-proline domain
HDX-MS	Hydrogen/deuterium exchange mass spectrometry
HPLC	High performance liquid chromatography
HMWO	High molecular weight oligomer
Hr	Hour
IC_{50}	The half maximal inhibitory concentration
IPTG	Isopropyl-β -D-Thiogalactopyranoside
I-site	Inhibitory site
ITC	Isothermal titration calorimetry

k_{cat}	Turnover number
K_d	Dissociation constant
K_i	Inhibition constant
K_m	The Michaelis constant
KCN	Potassium cyanide
kDa	KiloDaltons
L-NAME	N_ω-Nitro-L-arginine methyl ester hydrochloride
LB	Luria-Bertani Media
LC/MS	Liquid chromatography-mass spectrometry
LOV	Light, oxygen or voltage
MS	Mass spectrometry
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
NONOate	A family of nitric oxide (NO) donors
NOS	Nitric oxide synthase
pAp	3'-phosphoadenosine 5'-phosphate
PAS	Per-Arnt-Sim domain
PCR	Polymerase chain reaction
PDB	Protein data bank
PDE	Phosphodiesterase
PDE-A	Phosphodiesterase A, c-di-GMP specific hydrolysis activity by EAL domain
PEP	Phosphenolpyruvate
PMSF	Phenylmethylsulphonyl fluoride
pNPP	*p*-Nitrophenyl Phosphate
ppGpp	Guanosine tetraphosphate
REC	The phosphoryl receiver domain of response regulator
rpm	Revolutions per minute
RP-HPLC	Reverse phase-high pressure liquid chromatography
RT	Room temperature
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide electrophoresis
SNP	Sodium nitroprusside
TEAB	Triethylammonium bicarbonate
TBAB	Tetrabutyl ammonium bromide
TM	Transmembrane
UV-Vis	Ultraviolet-visible
WT	Wild type

Publications

1. **Rao, F.,** Yang, Y., Qi, Y., Liang, Z-X. Catalytic mechanism of C-di-GMP specific phosphodiesterase: a study of the EAL domain containing protein RocR from Psudomonas aeruginosa. ***J. Bacteriol.*** 2008, 190:3622-31*(The work in this publication constitutes Chapter 2, Yang Y performed computational docking study).*

2. **Rao, F.,** Qi, Y., Chong, H.S., Kotaka, M., Li, B., Lescar, J., Tang, K., Liang, Z-X. The functional role of a conserved loop in EAL domain-based c-di-GMP specific phosphodiesterase. ***J. Bacteriol.*** 2009, 191:4722-31 (Commentary: J. Bacteriol. 2009, 191:4697-700). *(The work in this publication constitutes Chapter 3, Chong HS, Li B, Tang K performed the MS analysis, Qi Y provided data for the PA2567 protein).*

3. **Rao, F.,** See, RY., Zhang, D., Toh, D.C., Liang, Z-X. YybT is a signaling protein that contains a cyclic-di-nucleotide phosphodiesterase domain and a GGDEF domain with ATPase activity. ***J. Biol. Chem.*** 2010, 285:473-82. *(The work in this publication constitutes Chapter 4, See RY, Zhang D, and Toh DC provided experimental assistance)*

4. **Rao, F.,** Soehano, I., Ji, Q., Laing, Z-X. Unusual Heme-Binding PAS Domain from YybT Family Proteins *J. Bacteriol.* 2011, 193:1543-1551. (*The work in this submitted manuscript constitutes Chapter 5, Soehano I and Ji Q provided experimental assistance)*

Related publications not included in this thesis

5. **Rao, F.,** Pasunooti S., Ng Y., Zhuo W., Lim L., Liu A.W., Liang Z-X. Enzymatic synthesis of c-di-GMP using a thermophilic diguanylate cyclase. ***Anal. Biochem.*** 2009, 389:138-42.
6. **Rao, F.,** Qi, Y., Murugan, E., Pasunooti, S., Ji, Q. 2',3'-cAMP hydrolysis by metal-dependent phosphodiesterases containing DHH, EAL, and HD domains is non-specific: implications for PDE screening. ***Biochem. Biophys. Res. Commun.*** 2010, 398:500-505.
7. Qi, Y., **Rao, F.,** Luo, Z., Liang, Z_X. A flavin cofactor-binding PAS domain regulates C-di-GMP synthesis in AxDGC2 from *Acetobacter xylinum*. ***Biochemistry*** 2009, 48:10275-85.
8. Kotaka, M., Dutta, S., Lee, H.C., Lim, M., Wong, Y., **Rao, F.,** Mitchell, E.P., Liang, Z-X, Lescar, J. Expression, purification and preliminary crystallographic analysis of *Pseudomonas aeruginosa* RocR protein. ***Acta Crystallogr. F*** 2009, 65: 1035-1038.
9. Liang, Z-X., **Rao, F.** Engineering a thermophilic diguanylate cyclase for large scale production of Cyclic-di-GMP. 2009, **US Patent** (provisional No: 61/156,826).

Chapter One: Overall introduction

Since the first discovery of adenosine 3', 5'-monophosphate (cAMP) in the 1950s, it has become increasingly clear that small molecules are often employed as second second messengers due to their flexibility to integrate a variety of signals for spatiotemporal control and signal amplification. The role of cAMP and guanosine 3', 5'-monophosphate (cGMP) as second messengers has been well established in eukaryotes. Two small nucleotides, cAMP and guanosine-3', 5'-bis (pyrophosphate) (ppGpp), have also been well-established as messengers in prokaryotes (81, 122). Here, my PhD research focuses on the recently established second messenger cyclic-di-guanosine monophosphate (C-di-GMP or cyclic-di-GMP) and on cyclic-di-adenosine monophosphate (C-di-AMP), which is potentially a fourth second messenger **(Fig. 1.1)**.

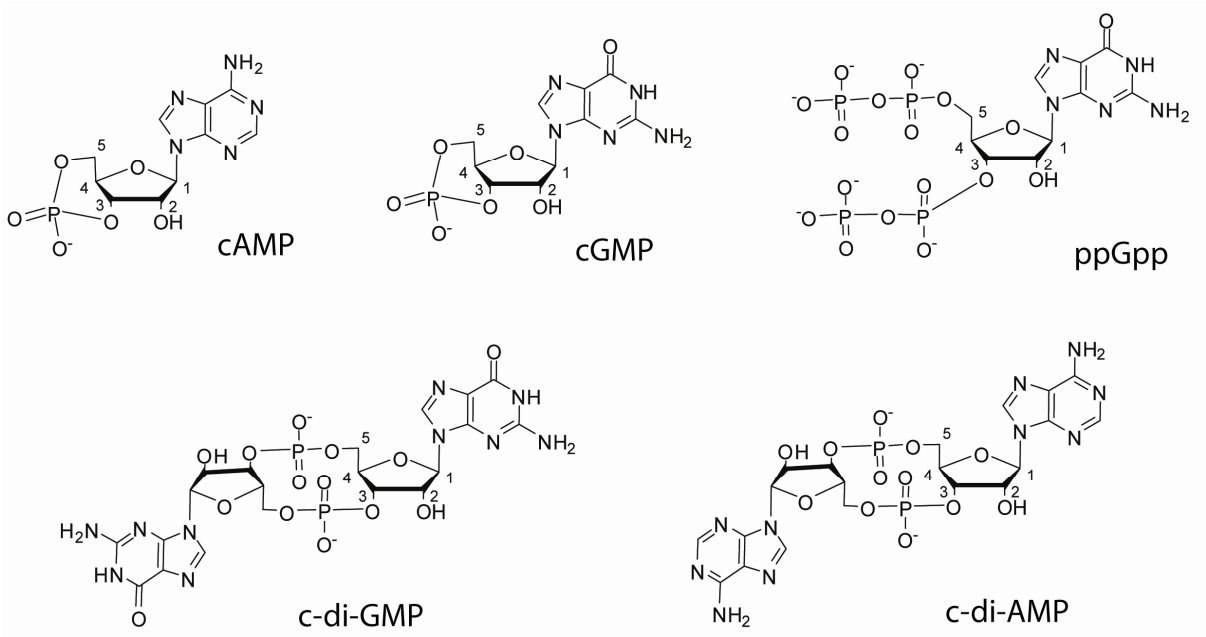


Figure 1.1 Structures of the nucleotide signaling messengers.

1.1 C-di-GMP

C-di-GMP was first discovered in *Gluconacetobacter xylinus* 20 years ago as an allosteric regulator of cellulose synthesis by Benziman and co-workers (139). However, interest in c-di-GMP signaling systems increased dramatically only recently, where genome sequencing revealed that the GGDEF and EAL domains responsible for controlling c-di-GMP level are present in a large number (from 20 to 60 proteins) in many species (43). Today, c-di-GMP has emerged as a global bacterial second messenger, with a large body of evidence suggesting that c-di-GMP plays a variety of roles in regulating cellular functions that include cell-cell communication, biofilm formation, twitching motility, differentiation, antibiotic production and virulence gene expression (59, 69-70, 137-138, 164).

Initial genetic studies in *G. xylinus* suggested that proteins containing GGDEF and EAL domains (named after the conserved amino acid motifs "GGDEF" and "EAL" found in the respective domains) regulate cellular c-di-GMP level (163). Later, a family of HD-GYP domain proteins was also proposed to be c-di-GMP phosdiesterases (PDEs), but they are not as widely present as the EAL domain proteins (43). The proposed roles of GGDEF and EAL were supported by further biochemical, genetic and bioinformatic studies (5, 23, 116, 151, 172). These studies have established the GGDEF domain as a diguanylate cyclase (DGC) domain that synthesizes c-di-GMP from two GTP molecules, and the EAL domain as a c-di-GMP specific PDE that degrades c-di-GMP to 5'-pGpG **(Fig. 1.2A)**. The functions of the GGDEF and EAL domains were confirmed when the enzymatic activities of purified GGDEF (115, 144) and EAL (28, 146, 165) domain proteins were characterized *in vitro* by the groups of Gomelsky and others. Meanwhile, diverse physiological functions of GGDEF and EAL proteins also came to light. In *Caulobacter crecentus*, the pole localization and c-di-GMP producing activity of the GGDEF domain containing response regulator PleD is controlled in a cell-cycle dependant manner, suggesting a role for c-di-GMP in cell cycle progression (115). More prominently, many GGDEF and EAL proteins are implicated in controlling the switch

between a motile planktonic and sessile biofilm-associated lifestyle of bacteria. Romling and co-workers showed that, in pathogens like *Salmonella enterica* and *Pseudomonas aeruginosa* or the commensal species *E. coli,* the EAL domain activity is associated with reduced c-di-GMP level and a transition from sessility to motility for bacterial species, whereas the GGDEF domain activity is associated with increased c-di-GMP level and a transition from motile to sessile/biofilm lifestyle (150). This notion has been subsequently supported by various studies on GGDEF and EAL proteins in many bacterial species (reviewed in 16). Another comprehensive analysis of all GGDEF and EAL domain proteins in *P. aeruginosa* further implies that these proteins, and hence the messenger c-di-GMP, plays important roles in virulence (86). Thus, the GGDEF and EAL domains are the central players in the c-di-GMP signaling network and their function and regulation deserve careful examination.

Structural studies on the GGDEF domain protein PleD and WspR yielded insights into the catalytic mechanism of c-di-GMP synthesis as well as the regulatory mechanisms through the phosphoreceiver domain and a product inhibition site (I-site) (22, 32). The existence of an inherent I-site in GGDEF domains suggest that c-di-GMP level is tightly controlled. Other structural and biochemical studies on the c-di-GMP binding PilZ domain proteins also demonstrated how ligand binding triggers conformation changes to extert allosteric effect (59, and references therein).

In comparison, how the EAL domain catalyzes the hydrolysis of c-di-GMP was largely unknown. When I first started my research in 2006, it was only known that EAL domain is responsible for c-di-GMP specific hydrolysis (146, 163) and that the conserved EAL motif is indispensible *in vitro* (165). The EAL domain has since been the topic of intensive research in the c-di-GMP field in recent years. From the structural, biochemical and genetic studies conducted by several laboratories, including ours, a considerable amount of information has been gathered about the EAL domain. The most noticeable progress includes the elucidation of the catalytic mechanism and the discovery of catalytically incompetent EAL domains. For

example, we initiated an extensive mutagenesis study on EAL domain protein RocR, which shed light on the essential catalytic residues that were subsequently verified by structural study from other groups (105, 7). The essential role of a conserved loop in the EAL domain identified (128) also echoed the discovery of inactive EAL domain proteins that function as protein-protein interaction module (YcgF) or c-di-GMP receptor (LapD) (174, 112). Notably, the structure of the BLUF-EAL di-domain protein BlrP1 solved by Barends et al (7) illustrated how the change in quanterny structure regulates the EAL domain activity. In this dissertation, I will present results to answer the question of what is the catalytic and regulatory mechanism of EAL domain.

Apart from the diverse regulatory modules controlling c-di-GMP levels, the effector modules that bind and respond to c-di-GMP are also remarkably diverse. The first example is the PilZ-like regulatory domain in the cellulose synthase BcsA, from which c-di-GMP was isolated and discovered (139). Subsequent bioinformatic and biochemical analysis found that the PilZ domain is conserved in many proteins that function as c-di-GMP receptors (1), these proteins can be stand-alone PilZ domain or fused to enzymatic output domains. It is interesting that certain GGDEF and EAL proteins can be highly degenerate, catalytically incompetent, and function as c-di-GMP receptors. The GGDEF protein PopA from *C. crecentus* is inactive as a diguanylate cyclase but contains an intact I-site that serves as a c-di-GMP receptor. Binding of c-di-GMP sequesters PopA and its binding partner CtrA, which is an inhibitor of replication initiation, targets CtrA for preteolytic degradation, allowing cell cycle progression through G1 to S phase (37). Similarly, instead of hydrolyzing c-di-GMP, the degenerate EAL domain protein LapD from *P fluorescens* binds intracellular c-di-GMP and communicate this signal to its periplasmic domain for regulating cell attachment (112). Given that many GGDEF and EAL domains do not contain the catalytically essential residues, more of them may function as c-di-GMP receptors.

Recent discoveries demonstrate that c-di-GMP can directly modulate protein level by binding to transcription factors. The transcription factor VpsT binds c-di-GMP, through its response regulator receiver (REC) domain, to control extracellular matrix production and motility in *Vibrio cholerae* (82), whereas the transcription factor FleQ from *P. aeruginosa*, which also controls extracellular matrix synthesis, seems to sense c-di-GMP directly through its helix-turn-helix transcription factor domain (60). By binding to another type of transcription factor domain belonging to the CRP/FNR family (168), c-di-GMP exerts control on the quorum-sensing dependent secretion of virulence factors in the plant pathogen *Xanthomonas campestris*. Notably, the same domain fold was previously known to be the receptor for cAMP in other bacterial species (81), suggesting a common theme of direct transcriptional control by nucleotide second messengers in bacteria.

Besides regulating gene transcription, c-di-GMP controls protein expression at the mRNA level. Breaker and coworkers discovered a class of c-di-GMP responsive riboswitch class in many mRNAs, which revealed a variety of c-di-GMP regulon associated with virulence gene expression, pilus formation, and flagellum biosynthesis (161). What's more, the folding changes induced by c-di-GMP binding can result in alternative splicing to modulate alternative mRNA processing. In other words, c-di-GMP can function as an allosteric factor for regulating ribozyme catalysis (92). Taken together, the known effector modules through which c-di-GMP exert its action is already extraordinarily versatile, which underscores the importance of spatiotemporal control of c-di-GMP level. Detailed characterization of the mechanism of degradation by EAL domain, as conducted in this thesis work, should aid the dissection of c-di-GMP signaling network and lay the foundation for inhibitor design for the EAL domain.

1.2 C-di-AMP

The c-di-AMP synthesizing di-adenylate cyclases (DAC) were first discovered in *Bacillus* and *Thermotoga* in 2008 (183), where crystallization of the DNA damage scanning protein DisA serendipitously discovered an enzymatic active site that retained the product c-di-AMP in the N-terminal domain (DUF147, renamed DAC). As exemplified by the N-terminal domain of DisA (9, 183), the DAC proteins, which are structurally unrelated to the DGC proteins, condense two molecules of ATP to form c-di-AMP. The DAC activity of DisA was stimulated when DNA fragments representing branched nucleic acids such as Holliday junctions binds to the c-terminal nucleic acid binding HhH domain (183), suggesting that the level of c-di-AMP is linked to DNA damage repair. Subsequent phylogenetic analysis revealed a widespread distribution of DAC domain proteins in bacteria and archea. While some of them are fused to typical sensor modules such as PAS domain, most of them are associated with domains of unknown function, suggesting that signaling pathways that regulate and respond to c-di-AMP are mainly unexplored (136, 183). In light of the roles of c-di-GMP, it was further postulated that c-di-AMP may be another previously unknown second messenger (136). To fulfil such function, the cellular level of c-di-AMP must be tightly regulated and dynamically adjusted, presumably by a specific phosphodiesterase or efflux pump, or both. When I first inititated this project, there was no report of c-di-AMP degrading or exporting proteins for controlling cellular c-di-AMP level, although c-di-AMP was recently found to be secreted by *Listeria monocytogene* (184). Bioinformatic mining led me to discover a family of YybT proteins that co-occurs with the DAC domain proteins. YybT proteins contain a DHH domain that is well-known as metal-dependant phosphoesterase (3) and its co-occurance with DAC domain proteins lead us to hypothesize that it may function as c-di-AMP specific PDE **(Fig. 1.2B)**.

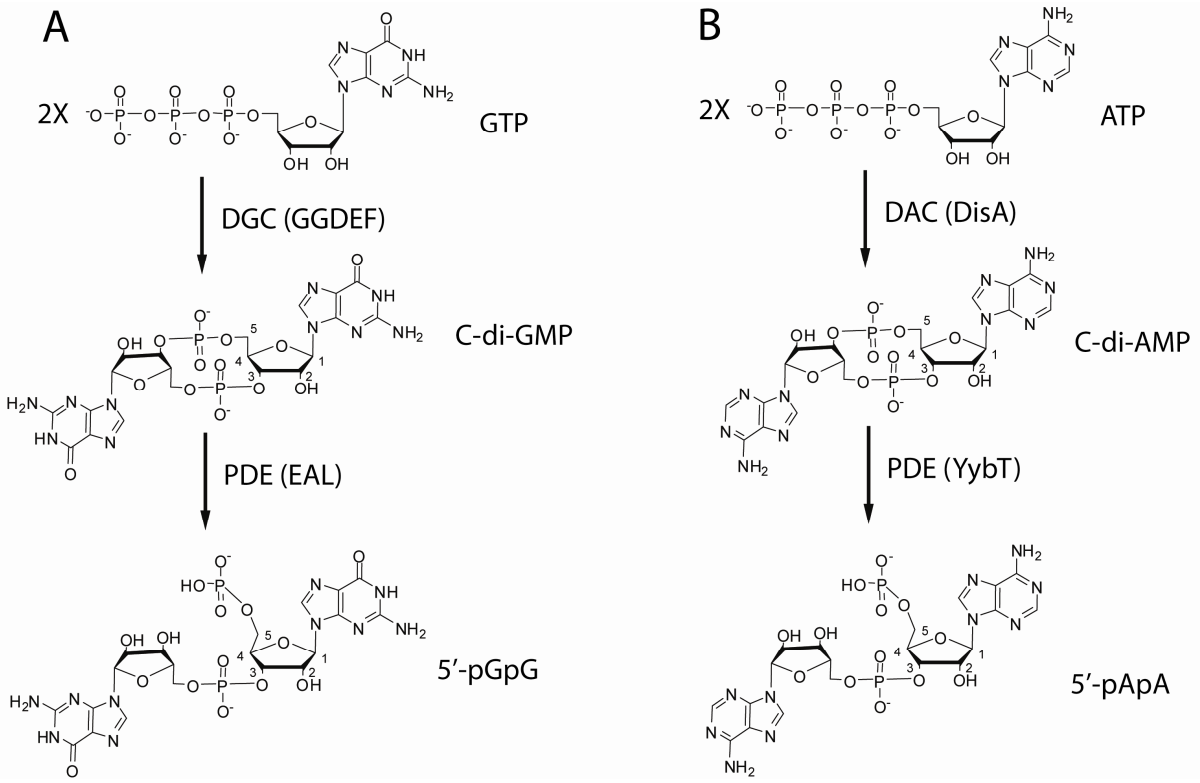


Figure 1.2 The synthesis and degradation of c-di-GMP (A) and c-di-AMP (B).

1.3 Objectives

The overall objective of this PhD project is to elucidate the catalytic and regulatory mechanism of c-di-GMP PDEs and identify c-di-AMP specific PDEs. First, the EAL domain-based c-di-GMP specific PDE will be examined in detail. Three EAL domain proteins (RocR, PA2567, and *Ax*DGC2) that contain catalytically active or inactive EAL domain are examined by biochemical and biophysical studies to probe the catalytic and regulatory mechanisms. Second, the DHH/DHHA1 domain-containing YybT will be studied as the potential c-di-AMP specific PDE. The substrate specificity of the DHH/DHHA1 domain will be examined. The regulation of the DHH/DHHA1 domain by a heme-binding PAS domain will also be investigated.

1.4 Thesis Organization

In Chapter two, I describe the mutagenesis study on RocR and PA2567 conducted to establish the catalytic mechanism of the EAL domain (132). In Chapter three, biochemical and biophysical evidences are presented to demonstrate the important role of a functional loop in the catalysis and regulation of the EAL domain (128). In Chapter four, I present data to argue that the bacillus protein YybT may contain a c-di-AMP specific phosphodiesterase domain (131). The allosteric regulation of PDE activity by an atypical PAS domain of YybT will be discussed in Chapter five. A brief concluding remark at the end of the dissertation summarizes the results and discusses future perspectives.

Chapter Two: Catalytic mechanism of c-di-GMP specific phosphodiesterase: A study of the EAL domain-containing RocR from *Pseudomonas aeruginosa*

2.1 INTRODUCTION

C-di-GMP (or cyclic-di-GMP) was first discovered as a regulator of cellulose synthesis in *Glucoacetobacter xylinus* and is emerging as a major bacterial second messenger (70, 138, 140, 164). The intracellular concentration of c-di-GMP is controlled by the GGDEF domain proteins with diguanylate cyclase (DGC) activity, and the EAL domain proteins with c-di-GMP specific phosphodiesterase (PDE-A) activity (138). GGDEF domains catalyze the condensation of two molecules of GTP to generate c-di-GMP while the EAL domains catalyze the hydrolysis of c-di-GMP to generate the dinucleotide 5'-pGpG, in a stereospecific manner **(Fig. 2.1)**.

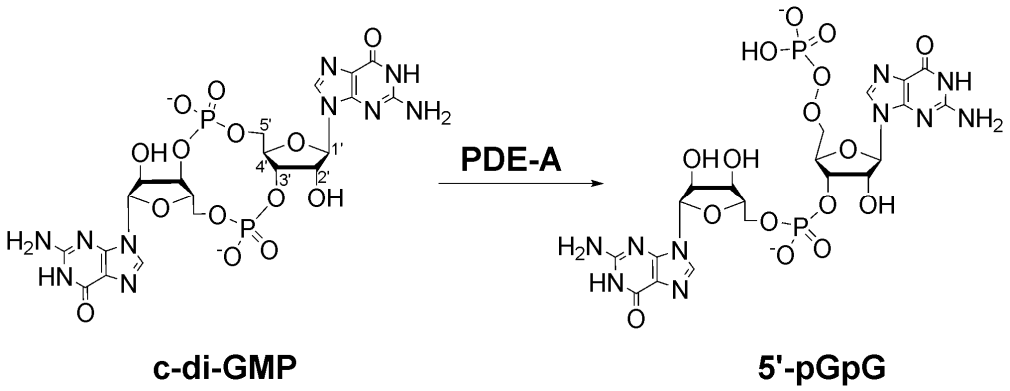


Figure 2.1 Phosphodiesterase (PDE-A) catalyzed hydrolysis of c-di-GMP to generate the linear 5'-pGpG.

Accumulating evidence suggest that EAL domain-containing proteins, including a large number of proteins that contain both the EAL and GGDEF domains, regulate a variety of cellular functions and phenotypes associated with bacterial infection. The regulation of virulence gene transcription, biofilm formation, motility and adhesion has been reported in various pathogenic bacteria. Several proteins in the human pathogen *Vibrio cholerae*, including VieA and CdgC, have been implicated in biofilm formation, motility and virulence factor production (75, 97, 172). An EAL domain protein was found to control lateral flagellar

gene expression and swarming behavior in *Vibrio parahaemolyticus* (78). In *Salmonella enterica*, the disruption of the EAL domain protein CdgR weakens bacterial resistance to hydrogen peroxide and accelerates bacterial killing of macrophages (61). In the opportunistic pathogen *P. aeruginosa*, the EAL domain-containing protein FimX controls twitching motility and biofilm formation (66, 76) and the BifA protein controls biofilm formation and swarming (85). The GGDEF-EAL di-domain protein MorA controls the timing of flagellar development and affects motility, chemotaxis, and biofilm formation (27). A systematic analysis of the GGDEF and EAL domain proteins in *P. aeruginosa* identified several other EAL domain proteins to be involved in virulence expression and biofilm formation (76, 86). Therefore, although the EAL domain proteins are not essential for the *in-vitro* viability of pathogenic bacteria (149, 86), they may be critical for the *in-vivo* survival of the pathogens in host organisms considering their roles in virulence expression and biofilm formation. This makes them potential targets for developing antibacterial agents that aim to neutralize virulence functions.

The genomes of the bacteria that contain c-di-GMP signaling network generally encode multiple EAL domain proteins. For example, the genomes of *P. aeruginosa PAO-1* and *V. cholerae* contain 21 and 32 ORFs coding for EAL domain proteins, respectively. In previous biochemical studies of EAL domain proteins, it was shown that Mg^{2+} or Mn^{2+} is required for the enzymes to hydrolyze c-di-GMP (146, 163, 165). It was also found that Zn^{2+} and Ca^{2+} can strongly inhibit 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 because the E→A mutation in two EAL domain proteins abolished their phosphodiesterase activity (28, 165). Additionally, Schmidt and coworkers suggested that other conserved motifs, including a DDFGTG motif, may be essential for the catalytic activity (146). The phosphodiesterase activity of the EAL domain of CC3396 from *Caulobacter crescentus* can be stimulated by the binding of GTP to the adjacent enzymatically inactive GGDEF domain

(28). The utilization of GGDEF domain could be a major strategy for regulating the catalytic activity of EAL domains, considering the large number of proteins that contain the GGDEF-EAL di-domain. Meanwhile, the widespread occurrence of GGDEF-EAL domains also raises the possibility that some EAL domains may function as regulatory rather than catalytic domains. However, due to limited information about the catalytic mechanism of EAL domains, such a distinction has remained speculative.

RocR was identified as a response regulator in the RocSAR (or SadARS) two-component signaling system in *P. aeruginosa* (83, 87, 141). RocSAR consists of the histidine kinase RocS1 and two response regulators RocA1 and RocR. The RocSAR system controls bacterial biofilm formation and virulence gene expression through regulating the transcription of various genes including the *cup* fimbrial gene clusters and type III secretion system (TTSS) genes (83, 86-87). Deduced from protein sequence, RocR contains an N-terminal CheY-like phosphoryl receiver domain and a C-terminal EAL domain. It was postulated that RocR negatively regulates the expression of *cup* genes by antagonizing the activity of RocA1, which is a typical response regulator with a DNA-binding domain (87). The detailed molecular mechanism for this antagonism is not known at this moment. Herein, I present biochemical data to demonstrate that the EAL domain of RocR is catalytically active with c-di-GMP specific phosphodiesterase activity. Using RocR as a model system, systematic mutagenesis in the EAL domain was carried out to probe the roles of 14 conserved polar residues in catalysis. Based on the biochemical data and aided by the crystal structure of a homologous EAL-domain protein, the functions of the conserved residues were assigned and a general base-catalyzed mechanism with the assistance of Mg^{2+} ion was proposed.

2.2 MATERIALS and METHODS

2.2.1 Cloning, expression and purification

The genomic DNA of *P. aeruginosa PAO-1* (ATCC) was isolated following standard procedures. The genes *PA3702*, *PA3947, PA0290 and PA2567* encoding WspR, RocR and PA290 and PA2567 respectively were amplified by PCR using the Expand High-Fidelity Kit (Roche). Amplified DNA fragments were cloned into the expression vector pET-26(b+) (Novagen) with compatible restriction sites. The plasmids harboring the gene and the $(His)_6$-Tag encoding sequence were transformed into *E. coli* strain BL21(DE3). For protein expression, one liter of bacterial culture (LB medium) was grown up to OD = 0.8 before induction with 0.8 mM IPTG. The culture was shaken at 16 ^{o}C for *ca.* 12 hours before being pelleted by centrifugation. The cells were lysed in 20 ml lysis buffer (20 mM Tris (pH 8.0), 500 mM NaCl, 5% glycerol, 0.1% β-ME, 0.1% Triton X-100, and 1 mM PMSF). After centrifugation of the lysate at 25,000 rpm for 30 min, the supernatant was filtered and then incubated with 2 ml of Ni^{2+}-NTA resin (Qiagen) for 1 hour at 4 ^{o}C. The resin was then washed with 50 ml of W1 buffer (lysis buffer with 20 mM imidazole) and 20 ml of W2 buffer (lysis buffer with 50 mM imidazole). The bound proteins were eluted using a step gradient method with the elution buffer containing 20 mM Tris (pH 8.0), 500 mM NaCl, 5% glycerol and 200mM, 300 mM, or 500 mM imidazole. After SDS-PAGE gel analysis, fractions with purity higher than 95% were pooled together and desalted using PD-10 column (GE-Healthcare). Proteins were first concentrated using Amicon concentrator (Millipore), and dialyzed into the storage buffer (50mM Tris buffer (pH 8.0), 250 mM NaCl, 25 mM KCl, 5 mM $MgCl_2$, 40% Glycerol, 1mM DTT). The proteins were stored at -20 ^{o}C after the measurement of protein concentration by Bradford assay. All RocR mutants were generated using the Site-Directed Mutagenesis II Kit (Stratagene) according to the manufacturer's instruction manual. Mutations were verified by sequencing using BigDye Terminator v3.1 Cycle Sequencing Kit on ABI Prism 3100 Genetic Analyzer (Applied Biosystems). Mutant

proteins were expressed, purified and stored under the same conditions as the wild type proteins. The yield of the recombinant protein is 10-15 mg/L for the wild type RocR and the mutants except for E268Q mutant (~1mg/L).

2.2.2 Size-exclusion chromatography

Gel filtration was performed at 4 °C using the AKTA FPLC system (GE Healthcare) equipped with a Superdex 200 HR 16/60 column (Amersham Biosciences). The buffer used for gel filtration comprised of 20 mM Tris-HCl (pH 8.0), 0.5 M NaCl, 5% glycerol, and 0.5 mM DTT. The molecular weight was estimated based on the standard curve generated by using ferritin, aldolase, conalbumin, ovalbumin and blue dextran as calibration standards.

2.2.3 Enzymatic synthesis of c-di-GMP

C-di-GMP was initially produced enzymatically by using WspR, the GGDEF domain-containing proteins encoded by the genes *PA3702* in *P. aeruginosa PAO-1*. The progress of c-di-GMP synthesis was monitored through HPLC by using an Agilent LC1200 system (Mobile phase: 20 mM Triethylammonium bicarbonate (pH 7.0), 10% Methanol, 0.6 mL/min) with an XDB-C18 column (4.6×150 mm). C-di-GMP was synthesized by incubating the enzymes and GTP (Sigma) at 30 ^{o}C in a 10 mL reaction mixture that contains 50 mM Tris buffer (pH 7.6), 20 mM $MgCl_2$ and 0.5 mM EDTA. Enzyme was supplemented in batches to maximize turnover. After the reaction, protein was precipitated and removed by heat treatment. Subsequently, the supernatant was filtered and concentrated using a vaccum-controlled evaporation system (Eyela) with the waterbath set at 45 ^{o}C. The concentrated supernatant was loaded onto a semi-preparative Eclipse XDB-C18 (9.4×250 mm) column for purification using the Agilent LC1200 system. Sample injection and collection of c-di-GMP fraction were automated with an autosampler (loading volume: 1 ml) and a fraction collector with 6-ml collecting vials. C-di-GMP was eluted using the same mobile phase described

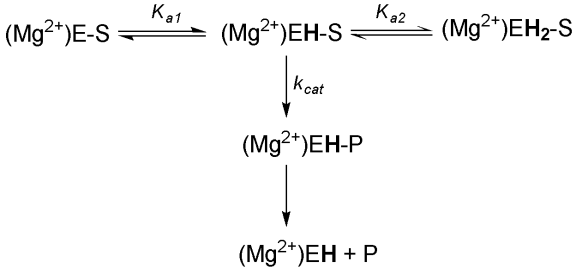
above with a flow rate of 3 ml/min. The fractions that contain c-di-GMP were pooled, concentrated by evaporation, and finally lyophilized to yield the white powder that was dissolved into 5mM Tris buffer (pH 7.0) for storage. Under this storage condition, C-di-GMP is stable for up to two years. The identity of c-di-GMP was confirmed by MALDI mass spectrometry with the MW of 690.085 (calculated MW 690.09), as well as by comparison with a standard from commercial source (Biolog). The concentration of the stock solution was determined by using a UV-Vis spectrophotometer (Shimadzu). An extinction coefficient (ε_{260}) of 26,100 OD $M^{-1}cm^{-1}$ was used for the calculation of c-di-GMP concentration (194). Subsequent development of this method involved a more stable thermophilic diguanylate cyclase, the tDGC from *Thermotoga*, that can be employed to generate hundreds of milligrams of c-di-GMP in one reaction (127). The tDGC was then used to prepare c-di-GMP for subsequent experiments.

2.2.4 Measurement of kinetic parameters

The identity of the product 5'-pGpG was confirmed by MALDI mass spectrometry with a molecular weight of 708.108 (calculated M.W. 708.11). The measurement of steady-state kinetic parameters was carried out by monitoring the formation of the product 5'-pGpG using HPLC. The standard assay was set up as: 100 mM Tris buffer (pH 8.0) with 20 mM KCl and 25 mM $MgCl_2$. Reactions were stopped by adding 1/10 volume of 1 M $CaCl_2$ and heating at 95^{o}C for 5 min. After removing the protein precipitate by centrifugation, the supernatant was loaded onto HPLC for quantification. Initial velocity at a certain substrate concentration was obtained from a series of reactions with varying incubation time. The total turnover was kept below 10% to ensure the accurate measurement of initial velocities within the linear range. Initial velocity was measured at 6-12 substrate concentrations. The kinetic parameters K_{cat} and K_m were obtained by fitting the initial velocities at various substrate concentrations to the Michaelis-Menten equation using the software Prism (GraphPad).

2.2.5 The pH dependence and data fitting

To cover the pH 6.2 - 9.5 range, 100 mM Bis-tris (pH 6.0 - 6.9), Tris (pH 7.0 – 9.0), or CHES (pH 9.0 - 9.5) buffer with 20 mM KCl and 25 mM $MgCl_2$ were used for kinetic measurement to obtain the pH profiles for RocR and mutant E352Q. The precise pH of the reaction mixture was measured using a micro pH electrode. The initial velocity (V_{max}) at two saturating c-di-GMP concentration (25 and 30 μM) was measured at various pH values. No significant difference was observed for V_{max} and only one set of the data was used for calculating the k_{cat} values used in the pH profiles. Given the observed bell-shaped pH profiles, a simplified model containing two ionizable catalytic groups was used to describe the pH dependence. (Mg^{2+})E-S, (Mg^{2+})EH-S and $(Mg^{2+})EH_2$-S represent the different ionization states of the ternary complex containing E (enzyme), Mg^{2+} and S (substrate). K_{a1} and K_{a2} are the ionization constants for the two groups that undergo ionization within the pH range of 6.2 to 9.5.



Based on the model, the bell-shaped pH profiles were fit to the equation shown below to obtain pK_{a1} and pK_{a2}, the apparent pK_a values for the two ionizable groups using the software DataFit (Oakdale Engineering).

$$\log (k_{cat}) = \log \left(\frac{k_{cat}(max)}{1 + 10^{(pK_{a2} - pH)} + 10^{(pH - pK_{a1})}} \right) \qquad \textbf{(Eq. 1)}$$

2.2.6 Structure modeling and Computational docking

The structural model for the EAL_{RocR} domain was constructed using the Swiss-Model server with the coordinates of TdEAL structure (PDB code: 2R6O) as template. The E value (3.8×10^{-31}) of the model is extremely low and more than 95% of the residues fall into the favorable region in Ramachandran plot. AutoDock 4 was used for the docking of c-di-GMP onto the receptor TdEAL (chain A) (10). Structural optimization and pK_a calculation were carried out to determine the protonation state using PDB2PQR (36, 47, 94) Formal and partial charges were then assigned by command-line tool in ADT. Mg^{2+} ion was incorporated into the PDBQT files with the charge of +1.2, +1.4 or +1.6 to represents different hydrated status of the Mg^{2+} ion. The coordinates of c-di-GMP was taken from the protein data bank (PDB ID: 2RDE). Hydrogen atoms were added using Ghemical and charges were assigned similarly as described for the receptor. Chemical affinity and electrostatics maps were computed and centered near the metal ion with $70 \times 70 \times 70$ grid points covering the putative binding cavity of the receptor with a spacing of 0.375 Å. Docking was performed using the Lamarckian genetic algorithm with the *pseudo*-Solis and Wets local search method. The preliminary efforts to dock c-di-GMP monomer and dimer into the binding pocket failed to produce any meaningful result, likely due to the closed conformation adopted by a capping helix. Instead, we used a truncated version of c-di-GMP without the two guanine bases for docking. The docking generated a single set of conformations with the ligand residing within the substrate binding pocket with a 2 Å root-mean-square deviation (RMSD). The two guanine residues were then added back to the truncated c-di-GMP. The enzyme-Mg^{2+}-substrate ternary complex was subjected to molecular dynamics simulation with GROMACS 3.3.1 using AMBER99 force field to generate the final docking model. This part of the work was done in collaboration with Yang Ye, a graduate student from the group of Prof. Lars Nordenskiold in school of biological sciences, NTU.

2.3 RESULTS

2.3.1 General characterization of RocR

Size-exclusion chromatography suggested that purified RocR **(Fig. 1.2A)** is mainly tetrameric in solution with an apparent molecular weight of ~170 kDa (calculated MW 176.4 KDa). RocR readily hydrolyzes c-di-GMP to produce 5'-pGpG **(Fig. 1.2B),** while no hydrolysis was observed for cAMP and cGMP, suggesting that the EAL_{RocR} domain is a c-di-GMP specific phosphodiesterase. Similar to other reported EAL domain proteins, Mn^{2+} ion can replace Mg^{2+} ion in catalysis, while Ca^{2+} and Zn^{2+} ions inhibit the phosphodiesterase activity. Under the experimental conditions, no substrate or product inhibition was observed for RocR. RocR catalyzes the hydrolysis of c-di-GMP with k_{cat} of 0.67 ± 0.03 s^{-1} and K_m of 3.2 ± 0.3 μM **(Table 2.1)**, which is close to the estimated c-di-GMP cellular concentration in *C. crescentus* (1.2 ± 0.11 μM) and in *Acetobacter xylinum* (5-10 μM) (28, 180).

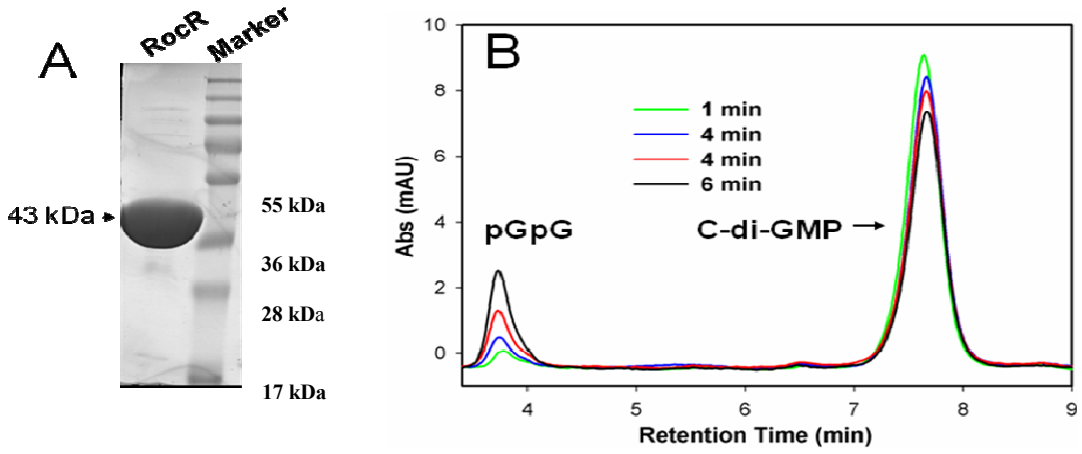


Figure 2.2 Purification and activity measurement of RocR. A. 12% SDS-PAGE image of RocR purified to homogeneity. B. HPLC traces showing time-dependent hydrolysis of c-di-GMP into 5'-pGpG. Activity is tuned to monitor turnover in linear range.

While this work was in progress, the crystal structures (PDB code: 2BAS and 2R6O) for two putative c-di-GMP specific phosphodiesterases were determined by the Midwest Center for Structural Genomics. The protein TdEAL (PDB: 2R6O) from *Thiobacillus denitrificans* contains a single EAL domain while the protein YkuI (PDB: 2BAS) from *Bacillus subtilis*

consists of an N-terminal EAL domain (EAL_{YkuI}) and a C-terminal domain with unknown function. Since neither *in-vitro* nor *in-vivo* phosphodiesterase activity has been confirmed for both proteins. We cloned, expressed and purified both proteins and found that YkuI is inactive against c-di-GMP (data not shown), while TdEAL is active **(Table 2.2)**. Together with the high sequence similarity (53%) between EAL_{RocR} and TdEAL, we constructed a structural model for EAL_{RocR} by using the structure of TdEAL as template **(Fig. 2.3)**. Overall, the model adopts the same $(\alpha/\beta)_8$ barrel fold and can be superimposed onto the crystal structure of TdEAL with a RMSD of 1.5 Å (backbone). Importantly, the side-chain conformations of the conserved residues in the active site of TdEAL are preserved in the structural model of EAL_{RocR} **(Fig. 2.3)**. The structural model should be reasonably reliable for identifying catalytic residues considering that most residues in the putative substrate binding pocket are conserved.

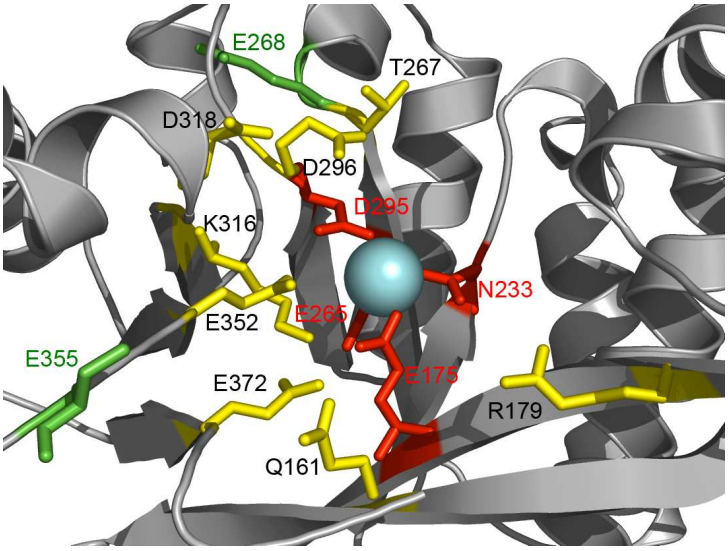


Figure 2.3 Structural model of EAL_{RocR} domain with the 14 conserved polar residues highlighted. The four residues for Mg^{2+} ion (sphere) binding are shown in red. The two distal residues are in green while the rest of the conserved residues are in yellow.

2.3.2 Site-directed mutagenesis and kinetic measurement

Sequence alignment of EAL domains (pfam 00563) in bacterial genomes revealed 14 conserved polar residues, including eight acidic residues (E^{175}, E^{265}, E^{268}, D^{295}, D^{296}, D^{318},

E^{352}, and E355), two basic residues (R^{179}, K^{316}) and four neutral residues (Q^{161}, Q^{372}, N^{233} and T^{267}). The 14 residues, which likely include the catalytically important residues such as the Glu residue (E^{175}) from the EAL signature motif, are completely conserved in EAL_{RocR} and TdEAL. As seen in the model of EAL_{RocR} (**Fig. 2.3**), most of the 14 conserved residues reside on the C-terminal ends of the central β barrel with the exception of E^{175}, R^{179}, Q^{161} and E^{265}, which are situated in the middle of β strands. E^{268} and E^{355} are located outside of the central barrel and are referred as distal residues. E^{268} is the only residue among the 14 with the side-chain carboxylate group pointing away from the central barrel.

To probe their roles in catalysis, each one of the conserved 14 residues was mutated individually to alanine. The steady-state kinetic parameters (k_{cat} and K_m) for the wild type and the 14 single mutants were measured under identical conditions. As shown in **Table 2.1**, the mutation of seven residues (Q^{161}, R^{179}, T^{267}, D^{296}, D^{318}, E^{355} and Q^{372}) caused various degrees of reduction in k_{cat}, ranging from a negligible 1.3-fold for E355A to 29.1-fold for R179A and 33.5-fold for D296A, suggesting that they only play minor or non-essential roles in catalysis. The K_m values for the mutants Q161A, R179A and D296A are 2~3-fold larger than that of the wild type RocR, indicating that they may be involved in substrate binding.

In sharp contrast, the mutation of the other seven residues, including E^{175}, N^{233}, E^{265}, E^{268}, D^{295}, K^{316} and E^{352}, reduced the activity to below measurable level (>10^5-fold decrease). Loss of protein stability does not seem to account for the decrease of activity because no noticeable precipitation or aggregation occurred during protein purification and activity assay. The profound effect of the mutations suggested that these residues play essential roles in assisting catalysis or maintaining the structure of the active site. The seven residues are referred to as essential residues from this point on. To further investigate whether the effect of the single mutations is due to structural perturbation, additional conservative mutants including E175Q, E265Q, D295N, E352Q, K316Q and E268Q were prepared. With the exception of E268Q, all of them showed slightly higher or comparable activity relative to the

corresponding alanine mutants, ruling out the possibility that the effects are caused by structural alteration and reinforcing the importance of the charge carried by the side chains. In contrast, E268Q mutant exhibits significantly higher activity relative to the E268A mutant. Kinetic measurement revealed that the turnover number (k_{cat} = 1.5 $\times 10^{-3}$ s^{-1}) for E268Q is 446-fold lower than that of the wild type RocR, in contrast to the $>10^5$-fold decrease for the E268A mutant. Given its distal location, it is likely that E^{268} may play a pivotal structural rather than catalytic role.

Table 2.1 Steady-state kinetic parameters for RocR and its mutants[a]

Enzyme	k_{cat} (s^{-1})	K_m (μM)	k_{cat}/K_m (s^{-1} μM^{-1})	Decrease in k_{cat} (fold)
RocR	0.67 ± 0.03	3.2 ± 0.3	0.21± 0.02	-
Q161A	0.13 ± 0.01	6.3 ± 1.0	(2.1 ± 0.4) × 10^{-2}	5.1
E175A	ND[b]	-	-	> 10^5
R179A	(2.3 ± 0.1) × 10^{-2}	8.0 ± 0.9	(2.9 ± 0.3) × 10^{-3}	29.1
N233A	ND	-	-	> 10^5
E265A	ND	-	-	> 10^5
T267A	(4.9 ± 0.1) × 10^{-2}	2.2 ± 0.4	(2.2 ± 0.4)× 10^{-2}	13.4
E268A	ND	-	-	> 10^5
E268Q	(1.5 ± 0.1) × 10^{-3}	0.3 ± 0.1	(4.8 ± 0.5) × 10^{-3}	446
D295A	ND	-	-	> 10^5
D296A	(2.1 ± 0.3) × 10^{-2}	8.6 ± 2.8	(2.7 ± 0.9) × 10^{-3}	33.5
K316A	ND	-	-	> 10^5
D318A	(8.0 ± 0.5) × 10^{-2}	5.9 ± 1.5	(1.4 ± 0.4) × 10^{-2}	8.4
E352A	ND	-	-	> 10^5
E352C	ND	-	-	> 10^5
E352Q	(1.1 ± 0.1) × 10^{-5}	3.8 ± 0.7	(2.9 ± 0.6) × 10^{-6}	6.1 × 10^4
E352D	(2.2 ± 0.1) × 10^{-5}	2.2 ± 0.5	(1 ± 0.2) × 10^{-5}	3.0 × 10^4
E355A	0.51 ± 0.07	2.6 ± 1.1	0.19 ± 0.09	1.3
Q372A	(8.0 ± 0.5) × 10^{-2}	1.3 ± 0.3	(6.2 ± 1.5) × 10^{-2}	8.4

[a]Conditions were 100 mM Tris buffer (pH 8.0) (23 ^oC), 20 mM KCl, 25 mM $MgCl_2$. [b]ND: not determined due to inactivity or extreme low activity (> 10^5-fold less active than wild type)

2.3.3 Recovery of the catalytic activity of inactive mutants at high Mg^{2+} concentration

Previous studies showed that Mg^{2+} or Mn^{2+} ion is indispensable for the *in-vitro* phosphodiesterase activity of EAL domains. Mg^{2+} is likely to be the metal ion utilized by EAL domains under physiological conditions. In the crystal structure of the homodimeric TdEAL, the corresponding residues of the four identified essential residues (E^{175}, E^{265}, D^{295} and N^{233}) coordinate one Mg^{2+} ion along with two water ligands in both subunits. Additionally, in one subunit (2R6O, chain B), the other two conserved residues D^{296} and D^{318} coordinate a second Mg^{2+} ion and raise the possibility of a two-metal-ion-assisted catalytic mechanism.

The binding of the Mg^{2+} ion by the four residues (E^{175}, E^{265}, D^{295} and N^{233}) in solution is supported by the observation that mutation of any of the four residues to alanine caused drastic decrease in catalytic efficiency. Additionally, we examined the catalytic activity of the four mutants E175A, E265A, D295A and N233A, along with other three inactive mutants E268A, K316A and E352A at elevated Mg^{2+} concentration. By using approximately the same amount of enzyme, we found that at elevated Mg^{2+}concentration (up to 500 mM), the catalytic activity of N233A can be fully restored, and the activity for E175A, E265A, D295A, E268A and K316A can be partially recovered, as compared to the turnover of the wild type enzyme at 10 mM Mg^{2+} concentration (**Fig. 2.4**). These results demonstrated that the Mg^{2+} binding site in the mutants can be populated again at high Mg^{2+} concentration, indicating that the lack of activity at low Mg^{2+} concentration is due to reduced binding affinity for Mg^{2+}. The different degree of recovery probably reflects the contribution of the residues in Mg^{2+} binding, with N^{233} contributing the least. Although K^{316} is not directly involved in Mg^{2+} binding, the partial recovery of activity for K316A can be rationalized given that K^{316} forms hydrogen-bonds with E^{175} and E^{265} and may be important for the positioning of the two residues for Mg^{2+} binding (**Fig. 2.3**). The partial restoration of the activity for E268A at high [Mg^{2+}] is intriguing and may reflect its indirect role in Mg^{2+} binding as we will discuss later.

Negligible recovery of activity was observed for mutant E352A, suggesting that the residue E_{352} either plays a most crucial role in metal-binding that cannot be recovered even at 500 mM Mg^{2+} or it plays another essential role other than Mg^{2+} binding.

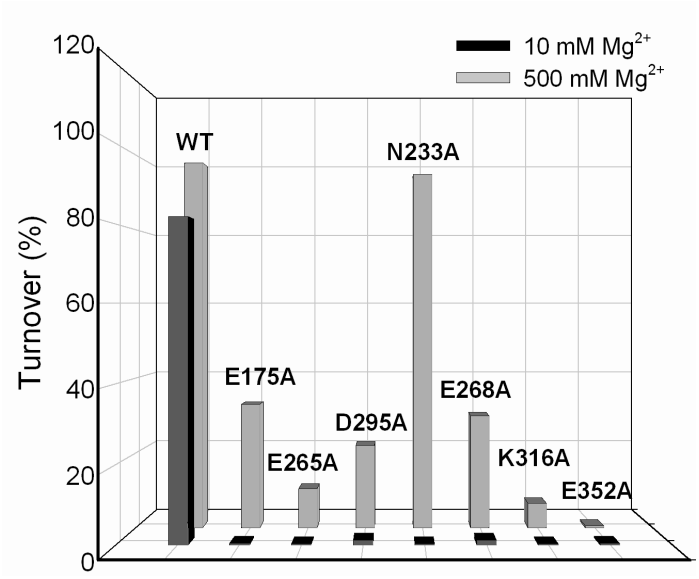


Figure 2.4 Recovery of catalytic activity at elevated Mg^{2+} concentration. (Assay conditions: 100 mM Tris buffer (pH 8.0), 20 mM KCl, ~1.0 μM of enzyme, 20 μM of c-di-GMP, 20 min incubation).

2.3.4 Identification of the general base catalyst

Phosphodiesterases generally hydrolyze substrates by using metal ion(s) and general acid/base catalyst for substrate or/and water activation. To identify the potential general acid/base catalyst, we carried out computational docking using the structure of TdEAL to explore the structure of the enzyme-Mg^{2+}-substrate complex (**Fig. 2.5A**). The relevance of the obtained docking model is supported by the following observations (**Fig. 2.5B**). First, in the model, one of the phosphates of c-di-GMP is coordinated to the essential Mg^{2+} ion with the anionic phosphate oxygen occupying a position initially occupied by a water ligand. Second, the conserved R^{63} (R^{179} in RocR) interacts with the second phosphate group of c-di-GMP in the model. The interaction between R^{63} and substrate is consistent with the elevated K_m for mutant R179A. Last and most importantly, the docking conformation of c-di-GMP predicts

that the nucleophilic attack of the Mg^{2+}-bound hydroxide ion will generate 5'-pGpG, but not 3'-pGpG as we will discuss in detail below.

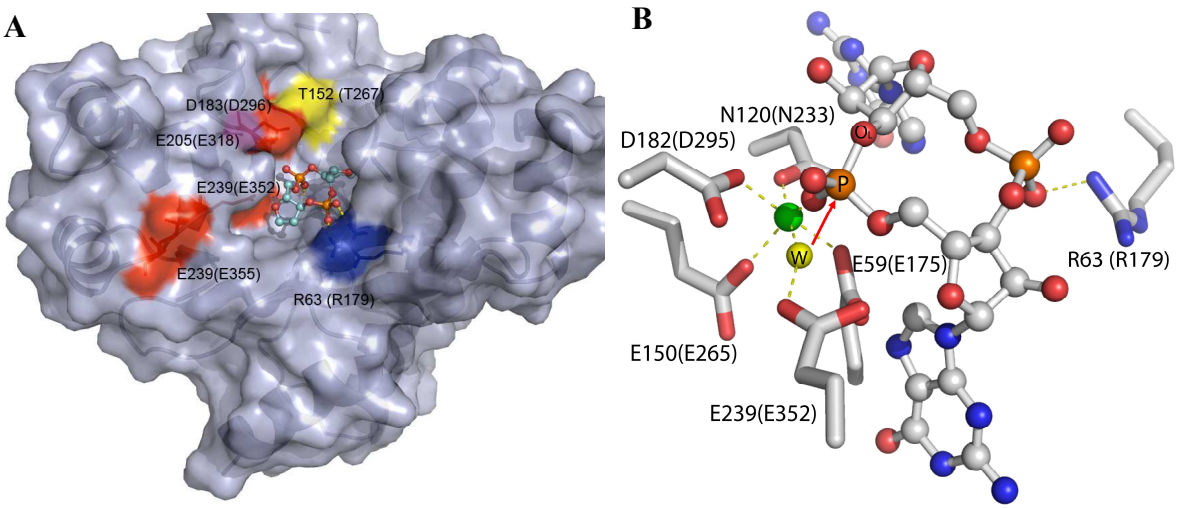


Figure 2.5 Computational docking results. The residues numbers for TdEAL are shown along with the corresponding residue numbers for EAL_{RocR} in the parentheses. A, Surface representation of TdEAL domain with the core of c-di-GMP docked in the binding pocket. B, Model of TdEAL-Mg^{2+}-substrate ternary complex generated from computational docking. The residues numbers for TdEAL are shown along with the corresponding residue numbers for EAL_{RocR} in the parentheses. The water (yellow ball) is coordinated by Mg^{2+} (green ball) and hydrogen-bonded to E^{239} (E^{352}) for deprotonation and nucleophilic attack. The water is positioned 3.0 Å away from the electrophilic phosphorus center with the leaving group (O_L) aligned linearly on the opposite side of the phosphorus.

In the crystal structure of TdEAL, a water molecule coordinated by the Mg^{2+} ion is within hydrogen-bond distance with the residues E^{239} and D^{182} (E^{352} and D^{295} in RocR) (**Fig. 2.5**). In the model of the ternary complex, the water molecule is located near the electrophilic phosphorus center of c-di-GMP with a distance of 3.0 Å. C-di-GMP is oriented in such a way that the O(water) –P(phosphate) – O_L(leaving group) angle approaches 180°. The linear alignment of the H_2O and the leaving group led us to propose a general based-catalyzed mechanism in which E^{352} (but not D^{295}) functions as a general base catalyst for deprotonating the Mg^{2+} ion-coordinated water to generate a nucleophilic hydroxide ion. To further validate the role of E^{352}, we prepared two more mutants (E352C and E352D) in addition to the E352A and E352Q mutants. Similar to E352A, the activity of E352C is so low (>10^5-fold decrease in k_{cat}) that the determination of the kinetic parameters becomes unrealistic. The lack of

activity for E352C suggested that the nucleophilic character of the side chain is unimportant given that Cys carries a strong nucleophilic group (SH). The other two mutants E352Q and E352D exhibited slightly higher activity than the E352A mutant, which allowed us to measure the kinetic parameters using high concentration of enzyme. Both mutants showed >10^4-fold reduction in k_{cat} relative to the wild type RocR (**Table 2.1**), indicating that the precise positioning of the negative charge carried by the carboxylate group is crucial for efficient catalysis.

The determination of the pH dependence of catalytic activity would yield information regarding the mechanism and the general acid/base catalyst. The measurement of the turnover numbers (k_{cat}) for RocR in the pH range of 6.2 to 9.5 revealed a bell-shaped pH profile with an ascending limb in the low pH region and a descending limb in the high pH range (**Fig. 2.6** left panel). A simple model that contains two ionizable groups was used to fit the pH dependence data as described in Methods. The limiting slope of 1.8 for the ascending limb suggests that one or two groups need to be deprotonated in the enzyme-Mg^{2+}-substrate ternary complex for efficient catalysis. The limiting slope of 0.7 was also observed for the descending limb in the high pH region. To probe the relationship between the pH dependence behavior and the ionization of E^{352}, we also determined the pH profile for the E352Q mutant (**Fig. 2.6** right panel). Although a bell-shaped curve was also observed with a similar descending slope of 0.6 for the mutant, the limiting slope for the ascending limb has changed from 1.8 for the wild type RocR to 0.7 for the mutant. The change in slope by 1.1 indicates that the observed pH dependence in the low pH region is largely attributed to the protonation/deprotonation of E^{352}. The weakened pH dependence for the mutant supports the mechanism in which the deprotonated E^{352} functions as a general base catalyst for deprotonating the water molecule. Although D^{295} may assist the deprotonation of the water molecule by influencing the pK_a of H_2O, it is unlikely that D^{295} functions as the proton acceptor (general base) given its low pK_a resulted from Mg^{2+} binding. Similar pH

dependence measurement was not carried out for D295A or D295N mutant due to the extremely low activity.

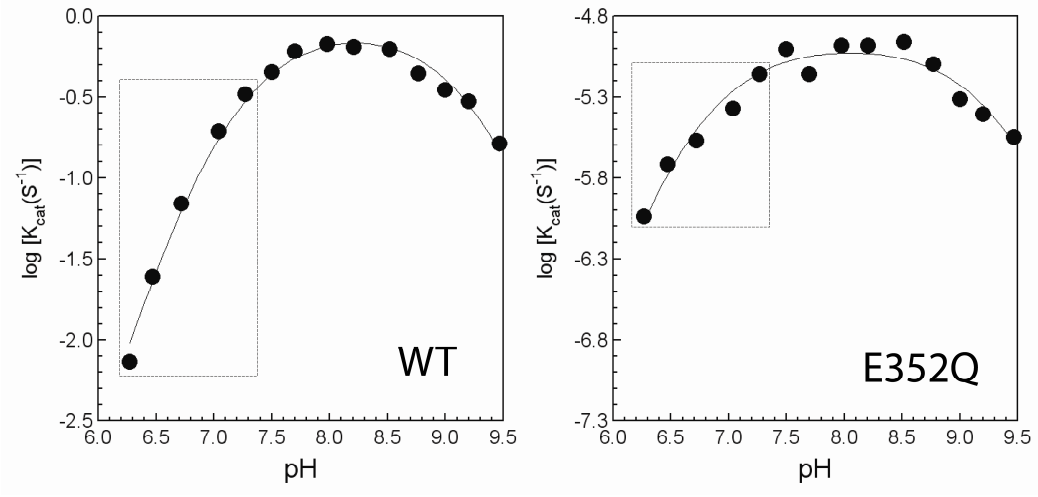


Figure 2.6 pH dependence of enzymatic activity for wild type RocR and mutant E352Q. The curves were obtained from the nonlinear-least-square fitting of the data to Eq. 1 described in Methods.

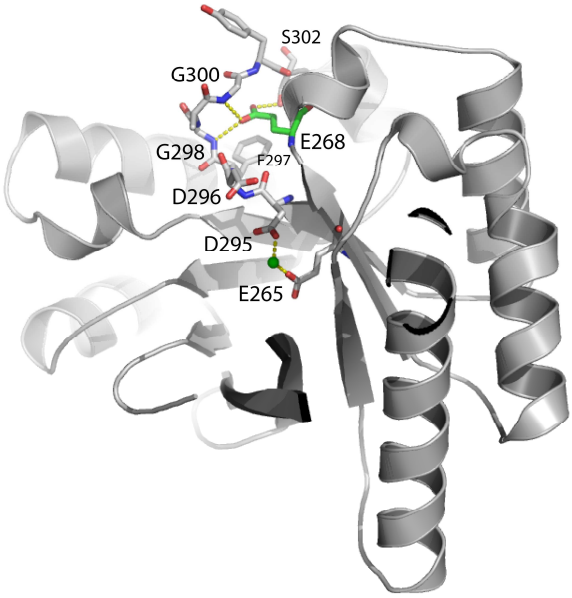


Figure 2.7 Interaction between E^{268} and the loop between β5 and α6. Residues from the loop ($F^{297}GAG^{300}YS^{302}$), E^{268} and the two neighboring residues (D^{295} and E^{265}) for Mg^{2+} ion (green ball) binding are highlighted to show that mutation of E^{268} may affect Mg^{2+} and substrate binding.

2.4 DISCUSSION

Given the large number of genes encoding EAL domain proteins in bacteria, understanding the regulation of EAL domains is essential for understanding the complex c-di-GMP signaling network. Other than cellular compartmentalization and transcriptional regulation (164), EAL domains are regulated by a wide array of regulatory domains at the protein level. The elucidation of the catalytic mechanism for EAL domains would be the first step towards understanding the protein-level regulatory mechanism. Our investigation that focused on the catalytic mechanism of the EAL domain of RocR shed light on the general catalytic mechanism of EAL domains.

2.4.1 Mg^{2+} ion-assisted catalytic mechanism

In most Mg^{2+}-dependent enzymes, Mg^{2+} is coordinated by three amino acid residues. In contrast, the crystal structure of TdEAL shows that the essential Mg^{2+} in EAL domain is coordinated by four residues, presumably with high binding affinity. The mutation of any of the four residues rendered the enzyme almost completely inactive, indicating that they are all indispensable for Mg^{2+} binding. The coordination of the Mg^{2+} ion by the four residues is further confirmed by the recovery of catalytic activity for the mutants at high [Mg^{2+}] (**Fig. 2.4**). The mutation of the Glu (E^{175}) in the EXL signature motif has been known to cause inactivity in EAL domains. Our study showed that the mutation of E^{175} to either A or Q significantly reduces the enzyme activity, in agreement with the effect of E→A mutation observed for the *V. cholerae* protein VieA (165).

Phosphodiesterases can catalyze the hydrolysis of substrate by using one, two or even three metal ions. The observation of two Mg^{2+} ions in the crystal structure of TdEAL (2R6O, Chain B) raised the possibility of a two-metal-ion catalytic mechanism as featured in polymerases and nucleases (191). However, the mutation of the two residues (D^{296}, D^{318}) for the second Mg^{2+} ion binding only caused 29 and 8-fold reduction in k_{cat}, suggesting the

binding of the second metal ion is not essential for catalysis. In addition, the large distance (8 Å) between the two Mg^{2+} ions rules out the possibility of a cooperative binuclear mechanism, in which the two Mg^{2+} ions are separated by a distance of ~4 Å and bridged by substrate or a Glu residue (191). Given the minor impact of D296A and D318A mutation on catalysis, the observed second Mg^{2+} bound at this site is likely to be a physiologically irrelevant observation that resulted from the high Mg^{2+} concentration (200 mM) used in crystallization. However, later on, a physiologically relevant second Mg^{2+} was found in the structure of the enzyme-substrate complex of some other EAL domain proteins, as discussed below.

In summary, here we propose that the essential Mg^{2+} ion plays two major roles in c-di-GMP hydrolysis: (i) it coordinates and activates substrate by polarizing the P-O bond. (ii) It lowers the pK_a of the coordinated water molecule to generate the nucleophilic hydroxide ion with the assistance of a general acid/base catalyst. Since D^{296} and D^{318} are in close proximity to c-di-GMP as seen in the model of the ternary complex, the decreases in k_{cat} and increases in K_m for D296A and D318A may reflect their roles in directing substrate binding.

2.4.2 General base-catalyzed mechanism

Phosphodiesterases generally hydrolyze their substrates by using metal ion(s) with the assistance of general-acid/base catalysts. A general base-catalyzed mechanism was first considered given the large number of conserved acidic residues in the active site. Computational docking suggested that E^{352} is the primary candidate for the general base catalyst since it is hydrogen-bonded to the Mg^{2+}-coordinated H_2O molecule poising for nucleophilic attack (**Fig. 2.5**). The attack of the hydroxide ion on the electrophilic phosphorus center would generate 5'-pGpG through a trigonal bipyramid transition state common for phosphodiesterases. Although the water is probably also hydrogen-bonded to D^{295}, one of the essential residues for Mg^{2+} binding, the likelihood of D^{295} acting as an efficient general base catalyst is small since its pK_a would be considerably lowered by the Mg^{2+}. On the other hand,

if E^{352} functions as the general base catalyst, we expect the elimination of the carboxylate group by mutation would decrease k_{cat} by more than 10^3-fold (21, 156). Indeed, the mutants E352A, E352C, E352Q all exhibit significantly reduced activity with k_{cat} of $(1.1 \pm 0.1) \times 10^{-5}$ s^{-1} for the most active E352Q. We also observed a 10^4-fold reduction in k_{cat} for the mutant E352D, which carries the carboxylate group but with a shorter side chain. The large decrease in k_{cat} for E352Q and E352D suggested that the precise positioning of the negatively-charged carboxylate group is critical for efficient hydrolysis of c-di-GMP. The observed large decrease in activity by mutating the counterpart of E^{352} to Q in PA2567 indicates the essential catalytic role of E^{352} is not restricted to RocR.

That E^{352} functions as the general base catalyst is further supported by the pH dependence study. The significant reduction in slope for the ascending limb suggests that E^{352} largely dictates the pH dependence behavior. The change in slope supports the mechanism in which E^{352} functions as a general base and proton acceptor for deprotonating the Mg^{2+}-activated water molecule. Although the limiting slopes of 0.6 and 0.7 in the descending limb for RocR and mutant E352Q raised the possibility of the participation of a general acid catalyst in catalysis, the involvement of a general acid catalyst is unlikely since no suitable candidate can be identified in the active site. The observed limiting slopes in the ascending and descending limbs for the mutant may be caused by some other pH-dependent structural changes in the ternary complex.

2.4.3 Why is E^{268} important for catalysis?

Given the distal location of the E^{268} residue, the total and partial loss of activity for mutants E268A and E268Q was initially puzzling. As seen in the structural model of EAL_{RocR} and the crystal structure of TdEAL (2RO6), the residue E^{268} (E^{153} in TdEAL), is wrapped around by the loop connecting β6 and α6 with E^{268} hydrogen-bonded to the well conserved G^{298}, G^{300} and S^{302} (**Fig. 2.7**). E^{268} also packs the hydrophobic region of its side

chain against F^{297}, another fairly conserved residue in EAL domains. The observed interaction between E^{268} and the loop residues led us to propose that E^{268} plays an important structural role in stabilizing the loop conformation. The mutation of E^{268} may cause conformational changes in the loop and results in the dislocation of the neighboring D^{295} and D^{296}, the two conserved residues involved in Mg^{2+} and substrate binding (**Fig. 2.7**). The indirect involvement of E^{268} in Mg^{2+} binding is supported by the observation that the activity of the E268A mutant can be partially restored at high Mg^{2+} concentration (**Fig. 2.4**). Although it remains to be seen whether E^{268} plays similar role in other EAL domain proteins, the importance of E^{268} for the activity of EAL domains has been observed in at least another protein. The EAL domain-containing YhjH from the cellulose-producing *S. typhimurium MAE97* was found to inhibit cellulose synthesis. The mutation of the corresponding residue of E^{268} to alanine in YhjH generated a mutant with stimulated cellulose biosynthesis, indicating that the mutation had a detrimental effect on the activity *in vivo* (150). Lastly, it is interesting to note that the DDFGTG motif proposed to be important for the activity of EAL domains contains the two residues D^{295} and D^{296}, as well as part of the loop (FGTG) discussed here (146). Given the sensitivity of the catalytic activity towards the conformation of the loop, it is intriguing to speculate whether some EAL domains are regulated through changing the conformation of the loop. This hypothesis was examined further and relevant results are presented in the next chapter.

2.4.4 Active and inactive EAL domains

EAL domains can be classified into subfamilies with different evolutionary origins according to phylogenetic analysis (86). However, it is unlikely that the subfamilies use different catalytic mechanisms because the 14 residues we examined seem to be conserved across the subfamilies. Combining the structural and biochemical data, we assigned functions for the conserved residues. Besides E^{352}, E^{268} and the four Mg^{2+}-binding residues, K^{316} and

Q^{372} are involved in the positioning of the Mg^{2+} binding residues. R^{179} and Q^{161} are involved in substrate binding by interacting with the anionic phosphate oxygen and guanine moiety of c-di-GMP respectively.

If the mechanism proposed here is indeed the general catalytic mechanism for EAL domains and the identified catalytic residues are crucial for efficient hydrolysis of c-di-GMP, then it is logical to deduce that the EAL domains lacking one or more of these residues would be catalytically inactive or inefficient. We have experimentally confirmed the importance of E^{352} in catalysis for PA2567, a protein that contains the GGDEF-EAL di-domain unit. We further examined the sequences of the EAL domains with catalytic activity characterized in either *in-vitro* or *in-vivo* assay. These domains are grouped into catalytically active and inactive (or inefficient) EAL domains as summarized in **Fig. 2.8**. All the characterized active EAL domain proteins contain the general base catalyst and Mg^{2+}-binding residues with the exception of FimX from *P. aeruginosa (76)*. Several essential residues in FimX are also absent, including the corresponding residues of N^{233} (changed to H), E^{268} (to Q), D^{295} (to S) and E^{352} (to P). Based on our mechanism, FimX should not be able to catalyze the hydrolysis with the loss of two ligands for Mg^{2+} binding and the general base catalyst. We expressed and purified the recombinant FimX and found that it exhibited very low enzymatic activity compared to RocR (at least 10^4-fold lower) (Qi Y. et al, to be unpublished). Therefore it is more appropriate to classify FimX as an inactive or inefficient EAL domain protein given its low catalytic efficiency. Besides FimX, seven of the other eight inactive EAL domains lack one or more of the essential residues. Notably, all seven of them do not contain a Glu at the corresponding E^{352} position. The EAL domain of the inactive CsrD was proposed to play a different role other than hydrolyzing c-di-GMP (162). Examination of the sequences of the homologous proteins of CsrD showed that none of them contains Glu at the corresponding position of E^{352} (162). The only "inactive" EAL domain that cannot be rationalized is DGC2, which contains all the essential residues and should be a functional phosphodiesterase domain.

Note that the protein contains a GGDEF domain and a regulatory PAS domain that may suppress the activity of EAL domain under the assay conditions (163). The studies conducted to examine the *in-vitro* activity of DGC2 and some other EAL domain proteins lacking one or two of the essential catalytic residues will be presented in Chapter 3. In addition, the YkuI and TdEAL proteins with structures solved are also included in generating the sequence alignment file; both porotein clustered in the upper half as all seven essential residues are conserved here **(Fig. 3.8)**. The TdEAL protein was found to be active by us (see table 2.2) and other group (171), whereas the YkuI protein is uncapable of hydrolyzing c-di-GMP (105), likely due to the lack of a conserved loop 6, as discussed in Chapter 3.

Overall, these observations indicate that the absence of the essential catalytic residues can serve as markers for identifying catalytically inactive EAL domains in the bacterial genomes. This has great implication considering that more than 20% of the EAL domains in bacterial genomes harbor a different residue at the corresponding position of the previously unrecognized E^{352} site. Given the large number of inactive EAL domains predicted based on this method, the elucidation of the functions of these EAL domains presents a new challenge towards understanding c-di-GMP signaling.

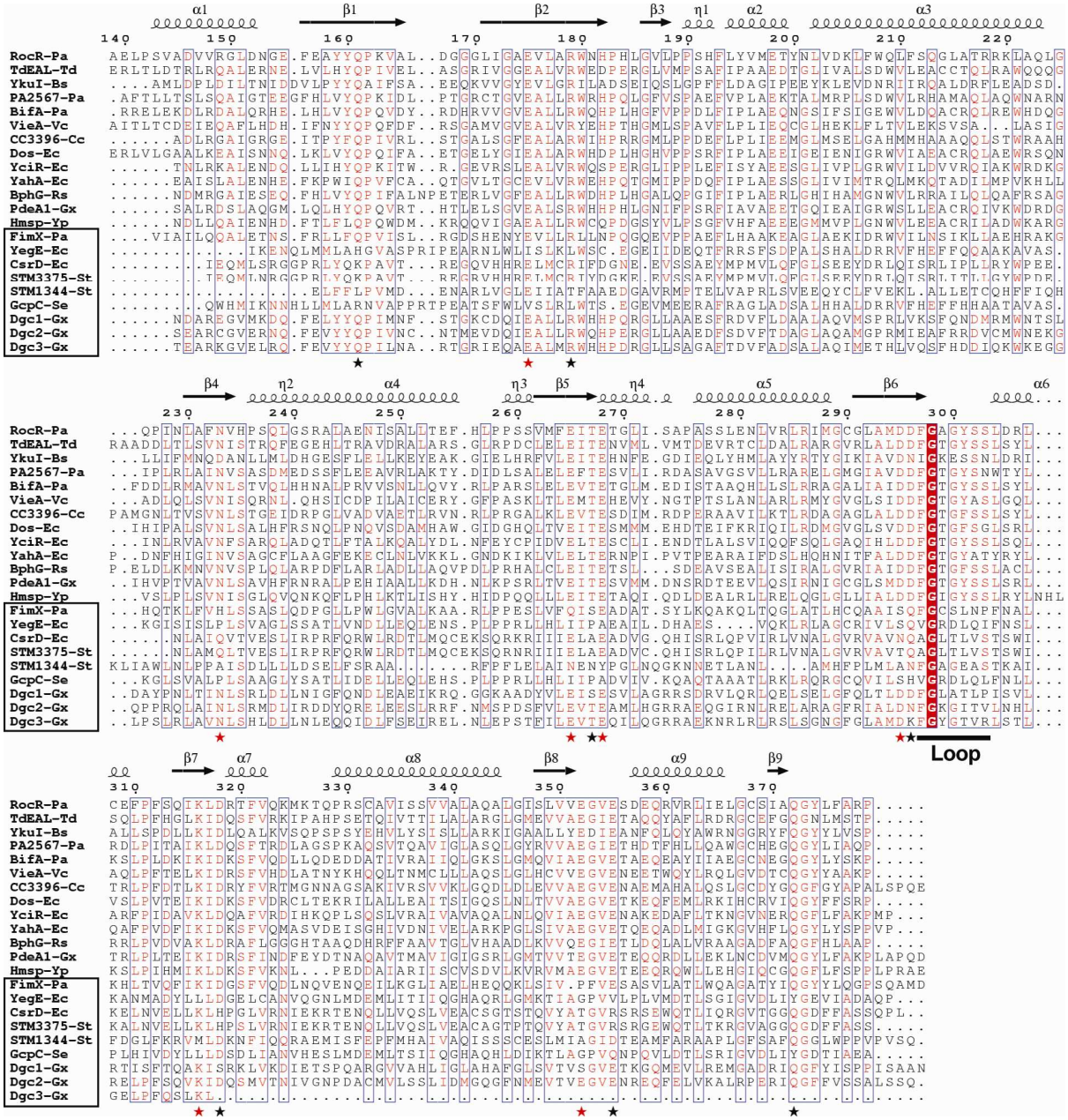


Figure 2.8 Sequence alignment of the EAL domains with catalytic activity characterized. The numbering of residues is based on the RocR sequence and the secondary structure is based on the structure of TdEAL (PDB: 2R6O). EAL_{YkuI} and TdEAL are included as active EAL domains. The inactive EAL domains are shown in the box. The residues examined in this study are indicated by black and red asterisk (essential residues). The loop that interacts with E_{268} is underlined. (PA2567, BifA and FimX from *Pseudomonas aeruginosa* (76, 85); TdEAL from *Thiobacillus denitrificans*; VieA from *Vibrio cholerae* (165); CC3396 from *Crescentus caulobacter* (28); YahA, Dos, YciR, CsrD and YegE from *Escherichia coli* (52, 146, 162, 179); BphG from *Rhodobacter sphaeroids* (169); PdeA1, DGC1, DGC2, DGC3 from *Gluconacetobacter xylinus (23, 163)* HmsP from *Yersnia pestis* (13); GcpC from *Salmonella enterica (46)*; STM1344 and STM3375 from *Salmonella typhimurium (149))* (Sequences were aligned using Multalin (http://bioinfo.genopole-toulouse.prd.fr/multalin/multalin.html) and the figure was generated using Espript 2.2 (http://espript.ibcp.fr/ESPript/ESPript/)

2.4.5 Updated mechanism

The screening for catalytically essential residues by mutating conserved polar residues, together with the guidance of substrate docking into an EAL domain structure (TdEAL, PDB id: 2R6O), had led us to propose a one-metal mechanism. In the proposed mechanism, a high-affinity active site Mg^{2+} and a general base activate a catalytic water for in-line nucleophilic attack on the phosphorus atom (132). The single Mg^{2+} was proposed to be coordinated by four highly conserved active site residues (E^{175}, N^{223}, E^{265}, D^{295}) including the glutamate from the EAL motif and a phosphate oxygen of c-di-GMP. This proposal was initially supported by the crystal structure of the EAL domain protein YkuI in complex with c-di-GMP and Ca^{2+}, despite the fact that YkuI was found to be an inactive EAL protein and the loop6 is highly degenerate (105). Meanwhile, Barends *et al* (7) proposed a two-metal ion mechanism based on the EAL protein BlrP1/M^{2+}/c-di-GMP ternary complex structures at various pH conditions (pH = 5-9). The structures revealed a second metal ion-binding site that consists of D^{295}, D^{296}, and E352 (residue numbers are based on RocR). The second metal ion was proposed to activate the nucleophilic water in collaboration with the first metal ion, which is coordinated by the same set of ligands as the Mg^{2+} in TdEAL (132) **(Fig. 2.9A)**. Most recently, the TdEAL/Mg^{2+}/c-di-GMP ternary complex structure was also solved (PDE id: 3N3T), which also revealed a second metal that is bound in a similar manner as in the BlrP1/Mn^{2+}/c-di-GMP ternary complex. Together with a mutagenesis study, a two-metal ion based mechanism was also proposed for TdEAL (171).

From these two structural studies, a two-metal mechanism seems to be operational in many EAL domain proteins, with the second metal ion being bound only in the presence of the substrate and was therefore absent in the *apo*-TdEAL structure. Given the conserved catalytic residues among active EAL domains, we re-examined the biochemical data collected on RocR in an effort to distinguish the two catalytic mechanisms. Some of the

observations discussed below seem to suggest a modified two-metal mechanism, with the second metal involved in substrate coordination rather than activation of nucleophilic water.

First, the catalytic activity of the D296A mutant protein is only moderately affected (30 fold). If D^{296} coordinates the second M^{2+} that is catalytically essential **(Fig. 2.9A)**, the alanine substitution would result in more dramatic decrease in activity, as was observed for other metal coordinating residues. This result therefore questions whether the second metal directly activates the catalytic water. In line with this argument, the *Ax*DGC2 protein from *A. xylinus*, which contains all essential catalytic residues but has an asparagine (N^{473}) at the position of D^{296} in RocR, is inactive against c-di-GMP but still retained phosphodiesterase activity against non-specific substrate bis-pNPP and 2',3'-cAMP (125, 130). In fact, *Ax*DGC2 is even more active against 2',3'-cAMP than RocR (130). These results suggest that the metal ion coordinated by D^{296} is not necessary for hydrolyzing the phosphodiester bond, but may be important to position c-di-GMP for in-line attack.

Second, the residue E^{352} functions as a general base to activate the catalytic water according to the one-metal mechanism, whereas the corresponding residue in BlrP1 (E^{359}) coordinates the second Mn^{2+} in a bidentate manner. Our activity assay at high metal concentration showed that the catalytic activity of E352A mutant is not recovered even at elevated Mg^{2+} concentration, whereas the activity of the D295A mutant can be readily recovered **(Fig. 2.4)**. According to the BlrP1 and TdEAL structures, D^{295} coordinates both metal ions; one would expect that the activity of the D295A mutant is more difficult to recover. Although it remains possible that the bidentate coordination makes E^{352} an important coordinating residue for the second Mn^{2+}, the E352Q (and E352D) seems to be inactive regardless of Mg^{2+} concentration, suggesting that E352 may not be involved in coordinateing the Mg^{2+} ion.

It should be also noted that, in contrast to our direct HPLC method to monitor pGpG formation, the recent mutagenesis study on the TdEAL protein used an indirect method

(measuring phosphate release after treating the product pGpG with alkaline phosphatase) that is not suitable for monitoring low activity (171). In the work of Tchigvintsev *et al*, a 10-20 fold lower activity was observed for D^{647} (D^{296} in RocR) and E^{703} mutants (E^{352} in RocR). The mutants were simply deemed catalytically inactive (171). However, the measurement method is incapable of differentiating the different activities among the mutants that exhibits low activity. To clarify, I measured the activity of TdEAL wildtype, D647A and E703A mutants. As can been seen from **Table 2.2**, while the activity of the D647A mutant is 78-fold lower, it is clearly measurable. The E703A and E703Q mutants are inactive even after 2-days incubation, corresponding to more than 1,000-fold lower activity. These observations are consistent with the results for RocR and suggest that D^{296} (or D^{647} in TdEAL) is dispensible for catalysis.

Table 2.2 Steady-state kinetic parameters for TdEAL and its mutants[a]

Enzyme	k_{cat} (s^{-1})	K_m (μM)	k_{cat}/K_m (s^{-1} μM^{-1})	Decrease in k_{cat} (fold)
TdEAL	$(5.1 \pm 0.3) \times 10^{-3}$	5.1 ± 0.9	$(1 \pm 0.2) \times 10^{-3}$	-
D647A	$(6.5 \pm 0.6) \times 10^{-5}$	28 ± 7.9	$(2.4 \pm 0.7) \times 10^{-6}$	78
E703A	ND[b]	-	-	$> 10^3$
E703Q	ND[b]	-	-	$> 10^3$

[a]Conditions were 100 mM Tris buffer (pH 8.0) (23 °C), 20 mM KCl, 25 mM $MgCl_2$.
[b]ND: not determined due to inactivity or undetectable activity. ($> 10^3$ fold lower than wildtype, which is our detection limit)

Taken together, these results led me to propose a slightly different two-metal mechanism for RocR in which E^{352} has dual functions: first, it coordinates the second metal in a mono-dentate manner; second, it also activates the attacking water/nucleophile as a general base. This hypothesis is consistent with the subunit A structure of the TdEAL/Mg^{2+}/c-di-GMP ternary complex (**Fig. 2.9B**), where the distance between E^{352} and the putative attacking water is 3.0 Å, significantly shorter than the distance between the second Mg^{2+} and the

putative attacking water (3.7 Å). This hypothetical mechanism is also congruent with the catalytic mechanism proposed for many other phosphotransfer enzymes such as, PvuII endonuclease (185) , RNAseH (190), and inositol 1,4-bisphosphate phosphatase (113). To distinguish the catalytic mechanisms, future work is necessary to examine the crystal structure of the wild type and E352Q mutant of RocR and the E703Q mutant of TdEAL, in complex with c-di-GMP. Special attention will be given to observe whether the second metal ion and the attacking water are still present in the E→Q mutants.

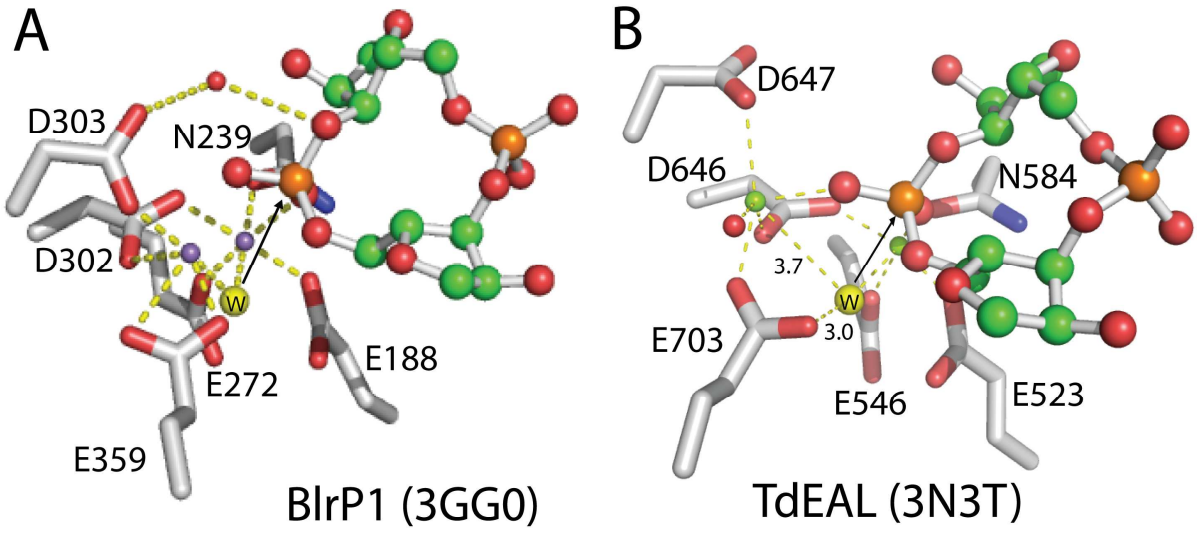


Figure 2.9 Scheme of proposed catalytic mechanism based on the structure of BlrP1(A) and TdEAL (B). The ribose and phosphate portion of c-di-GMP are represented by ball and stick. Metal ions are colored as: Mg^{2+}, green; Mn^{2+}, pink. The attacking water (W) is colored in yellow.

2.5 SUMMARY

EAL domain proteins are the major phosphodiesterases for maintaining the cellular concentration of second messenger c-di-GMP in bacteria. Given the pivotal roles of EAL domains in the regulation of many bacterial behaviors, the elucidation of their catalytic and regulatory mechanisms would contribute to the effort of deciphering the c-di-GMP signaling network. In this chapter, the data presented show that RocR, an EAL domain protein that regulates the expression of virulence genes and biofilm formation in *Pseudomonas aeruginosa PAO-1*, catalyzes the hydrolysis of c-di-GMP by using a general base-catalyzed mechanism with the assistance of Mg^{2+} ion. In addition to the five essential residues involved in Mg^{2+} binding, I propose that the essential residue E^{352} functions as a general base catalyst assisting the deprotonation of an Mg^{2+}-coordinated water to generate the nucleophilic hydroxide ion. The mutation of other conserved residues caused various degrees of changes in k_{cat} or K_m, suggesting their roles in residue positioning and substrate binding. With functions assigned for the conserved groups in the active site, I discussed the molecular basis for the lack of activity for some characterized EAL domain proteins, and the possibility of predicting the phosphodiesterase activity of the vast number of EAL domains in bacterial genomes in the light of the catalytic mechanism.

Chapter III: The functional role of a conserved loop in EAL domain-based c-di-GMP specific phosphodiesterase

3.1 INTRODUCTION

The cyclic dinucleotide c-di-GMP has emerged as a major bacterial messenger for mediating a variety of cellular functions that range from virulence expression to biofilm formation (19, 70, 138). The cellular concentration of c-di-GMP is controlled by the GGDEF domain proteins and the EAL domain proteins. GGDEF domains catalyze the synthesis of c-di-GMP from GTP, whereas EAL domains catalyze the hydrolysis of c-di-GMP to generate the linear 5'-pGpG. Although a family of HD-GYP domain proteins has also been found as c-di-GMP specific PDEs, the overwhelmingly large number of genes encoding the EAL domains in bacterial genomes suggests that the EAL domains are the major PDEs for maintaining the cellular c-di-GMP concentration. While many of the EAL domains are involved in c-di-GMP signaling by functioning as catalytically active PDEs, emerging evidence suggest that some EAL domains function as ligand or protein-binding domains without catalytic activity (112, 132, 174).

The detailed structure and catalytic mechanism of the EAL domains started to emerge recently. Specifically, the crystal structures of two EAL domains tdEAL and YkuI have been determined (PDE entries: 2BAS, 2R6O and 2W27) (105). EAL domains adopt a $(\beta/\alpha)_8$ barrel fold that contains two extended strands, including an antiparallel strand. The $(\beta/\alpha)_8$ barrel fold, first found in triosephosphate isomerase (TIM), has been observed in a diversity of enzymes that include many hydrolyases and isomerases (158). By conducting an extensive mutagenesis study as described in Chapter I, I mapped the catalytically essential residues of EAL domain protein RocR. Similar to other $(\beta/\alpha)_8$ barrel fold enzymes, the catalytic residues of the EAL domain are located at the C-terminal ends of the β-strands and within the β→α loops connecting the β-strands and α-helices. Although there is still some degree of discrepancy regarding whether the second metal ion directly activates the nucleophilic H_2O

(see discussion in chapter I), it is clear that residues coordinating the first Mg^{2+} ion and E^{352}, the residue that may function as a general base and/or coordinating the second Mg^{2+} ion, are indispensible for activity (132). The proposed catalytic mechanism can rationalize the lack of catalytic activity for most known inactive EAL domains, with the loss of enzymatic activity arisen from the absence of E^{352} and/or the residues coordinating the first Mg^{2+} ion (132).

It is well known that many $(\beta/\alpha)_8$ barrel fold enzymes contain a flexible loop between the β6 strand and α6 helix and near the active site (158). Despite the diverse reactions catalyzed by $(\beta/\alpha)_8$ barrel fold enzymes, this extended loop, often referred to as loop 6, plays an important role as functional lid for substrate sequestering, solvent exclusion and product release (74). The loop was found to facilitate substrate binding and conformational transition in tryptophan synthase (16-17), and functions as a lid for substrate sequestering during catalysis in inosine 5'-monophoshate dehydrogenase (IMPDH) (102). Notably, it was shown that the loop sways from the active site in the non-active structure of RuBP-carboxylase, but folds over to shield the active site from the solvent in the activated structure (99). Similar functions have also been proposed for loop 6 in other $(\beta/\alpha)_8$ barrel fold enzymes such as triose phosphate isomerase (TIM) and phosphoriboxyl anthranilate isomerase (74, 120, 124). Hence, it seems that the functional role of loop 6 has been well preserved in $(\beta/\alpha)_8$ barrel fold enzymes during evolution. The $(\beta/\alpha)_8$ barrel folded EAL domains also contain a seven residue loop between the β6 strand and α6 helix that seems to play a functional role in catalysis. Gomelsky and coworkers first noticed that the catalytically active EAL domains seem to contain a conserved motif that was later confirmed to contains loop 6 ($F^{297}G(T/A)GYSS^{303}$ in RocR) and two aspartate residues (D^{295}, D^{296} for RocR) in the proceeding β6 strand, residue D^{295} is for Mg^{2+} binding (132, 146). We previously noticed that mutations of the essential catalytic residues are usually accompanied by the mutation of the residues on loop 6 and the proceeding D^{296} in catalytically inactive EAL domains (132). Moreover, it was observed that the mutation of a residue interacting with loop 6 in the EAL

domain-containing RocR abolished enzymatic activity (150), which led us to postulate a critical role for loop 6 in catalysis (132). Unlike D^{295}, D^{296} is not an essential residue and is situated at the junction between β6 and loop. Structure comparison between tdEAL and YkuI suggest that the side chain of D^{296} is pulled away from the substrate c-di-GMP in YkuI, together with loop 6 (**Fig. 3.6**). Therefore, D^{296} was grouped with loop 6 rediues for the analysis presented in this chapter, unless otherwise specified.

To elucidate the precise role played by loop 6 in c-di-GMP hydrolysis in EAL domains, three EAL domain proteins (RocR, PA2567 and DGC2) were examined **(Fig. 3.1)**. The residues of Loop 6 [DFG(A/T)SYSS] in RocR and PA2567 are well conserved as observed in other catalytically active EAL domains. On the other hand, DGC2 is a cytoplasmic protein from *A. xylinus* that contains an EAL domain without c-di-GMP hydrolyzing activity(163). The EAL domain of DGC2 contains the conserved residues for Mg^{2+} binding and general base catalysis, but a highly degenerate loop 6 that consists of NFGKGITV. In this report, we present evidence to show that mutations in the loop 6 region in RocR and PA2567 have significant effect on the structure and catalysis of the EAL domain. By using the method of hydrogen-deuterium (H/D) exchange coupled mass spectrometry, we demonstrated that a remote single mutation in the phosphoreceiver domain of RocR caused correlated changes of loop 6 conformation and catalytic properties. We further show that the catalytic activity of the inactive EAL domain of DGC2 can be partially recovered by restoring loop 6. The functional roles of loop 6 in EAL domains in substrate-binding and catalysis are discussed in conjunction with the structural data for two EAL domains.

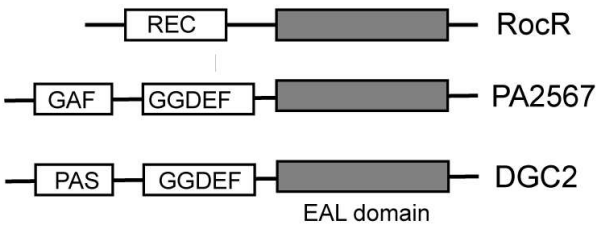


Figure 3.1 Domain organizations of the RocR and PA2567 protein from *P. aeruginosa* PAO-1, and the DGC2 protein from *A.xylinus* (*Ax*DGC2).

3.2 MATERIALS AND METHODS

3.2.1 Cloning, expression and purification

The cloning, expression and purification of RocR, PA2567 have been described in the previous Chapter. All mutants were generated using the Site- Directed Mutagenesis II Kit (Stratagene) according to the manufacturer's instruction manual. The mutant proteins were expressed and purified following the same procedure for the wild type proteins. The proteins were snap frozen in liquid N2 and stored in -80 ºC freezer after the concentration was measured by Bradford method.

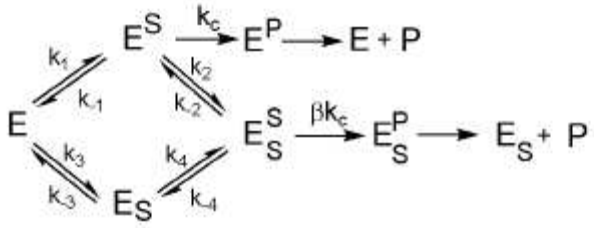
For the full length DGC2, a gene construct encoding the full-length *dgc2* gene from *A. xylinus* was commercially synthesized by GenScript Inc, amplified by PCR, and cloned into the expression vector pET26b(+) (Novagen) between the NdeI and XhoI restriction sites . The plasmid harboring the gene construct and the C-terminal $(His)_6$-Tag encoding sequence was transformed into *E. coli* strain BL21(DE3). Cell stocks were stored in 20% glycerol at -80 °C. For protein expression, 2 ml of inoculum from the cell stock was added to one liter of LB medium. Bacterial culture was grown at 37 °C up to OD = 0.8 before being induced with 0.8 mM IPTG at 16 °C for *ca.* 12 hours. After centrifugation, the pellets were lysed in 20 ml lysis buffer that contains 50 mM Tris-Cl (pH 8.0), 200 mM NaCl, 5% glycerol, 1% β-mercaptoethanol and 1 mM PMSF. After centrifugation at 25,000 rpm for 30 min, the supernatant was filtered and then incubated with 2 ml of Ni2+-NTA resin (Qiagen) for 30 min at 4 oC. The resin was washed with 50 ml of W1 buffer (lysis buffer with 20 mM imidazole) and 50 ml of W2 buffer (lysis buffer with 50 mM imidazole). The proteins were eluted using a stepped gradient method with the elution buffer containing 50 mM Tris (pH 8.0), 200 mM NaCl, 5% glycerol and 200, 300 and 500 mM imidazole. After SDS-PAGE gel analysis, fractions with purity higher than 95% were pooled together and desalted using PD-10 column (GE-Healthcare). Proteins were first concentrated using Amicon concentrator

(Millipore) to *ca* 5mg/ml (Measured by Bradford assay) before being assayed for enzymatic activity.

For the stand-alone EALDGC2 domain, an expression vector encoding aa 306 - 574 of Dgc2 was constructed by amplifying the corresponding gene fragment by high-fidelity PCR and cloning into 28(b+) (Novagen) between the NdeI and XhoI restriction sites. The plasmid harboring the gene construct and a N-terminal $(His)_6$-Tag encoding sequence was transformed into *E. coli* strain BL21(DE3). The protein was expressed and purified in the same procedure as full-length Dgc2, except in a lower-salt buffer (50 mM NaCl), since protein precipitates at higher salt concentration. The triple mutant ($N^{473}K^{476}I^{478}$ → $D^{473}T^{476}Y^{478}$) was generated using the Site-Directed Mutagenesis II Kit (Stratagene) according to the manufacturer's instruction manual. The 5-base mutations were designed in one set of primers for single round of PCR. Mutant protein was expressed and purified under the same conditions.

3.2.2 Steady-state enzymatic activity assay

The substrate c-di-GMP used for kinetic measurement was produced enzymatically by using WspR and PA290, two DGC domaincontaining proteins encoded by the genes *PA3702* and *PA0290* from *P. aeruginosa* PAO-1 as we described. The measurement of the steady-state kinetic parameters was carried out by monitoring the formation of the product 5'-pGpG (33). The kinetic parameters k_{cat} and K_m for the wild type proteins and the mutant that did not exhibited substrate inhibition were obtained by fitting the initial velocities at various substrate concentrations to the Michaelis-Menten equation using the software Prism (GraphPad). For the two mutants that exhibit substrate inhibition, we found that the kinetic data could be best fit using a model that assumes the substrate can bind to the enzyme at a productive and a non-productive (or inhibitory) binding site.



The curves were obtained by fitting the kinetic data to the following equation derived from the model based on the 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, υ is the initial velocity and [S] is the concentration of the substrate.

$$v = \frac{V_{max}\,(1/K_m + \beta[S]/\alpha K_i K_m)}{1/[S] + 1/K_m + 1/K_i + [S]/\alpha K_m K_i}$$

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

The apparatus for H/D exchange was set up following the published procedures with minor modification (65, 95, 134). Briefly, the purified RocR or D56N mutant was exchanged into 100 mM sodium phosphate buffer (pH 7.0) that contains 150 mM sodium chloride and 1 mM DTT. The peptides generated by pepsin digestion were first identified by MS/MS experiment. The H/D exchange reactions were initiated by adding 10 µl of newly thawed protein solution to 90 µl of D_2O. After incubation for various time lengths (1/3, 1, 5, 15, 30 and 60 min), the exchange reaction was rapidly quenched by lowering the pH to 2.4 and temperature to 0°C, and followed by pepsin digestion. The resulting peptides were passed through a home-made capillary reverse phase C-18 HPLC column and eluted into the Quadruple time-of-flight (TOF) mass spectrometer (Applied Biosystems/MDX Sciex QSTAR® Elite system) equipped with an IonSpray™ source. The peptides were eluted out at a flow rate of 20 µL/min using a stepped gradient (7.5, 10, 12.5, 15, 17.5, 20, 22.5, 25, 30, 40, 80% acetonitrile in 0.1% formic acid). The time from the initiation of digestion to the elution

of the last peptide was approximately 22 min. Analyst® QS software was used for spectrum analysis and data extraction and the H/D exchange data were processed by using the program HX-Express (181). 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 (65). 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 (65). The time-dependent H/D exchange data were fit by nonlinear least squares to the equation shown below.

$$Y = N - (Ae^{-k1t} + Be^{-k2t} + Ce^{-k2t})$$

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 (65, 96). The number of non-exchanging amides is calculated from the total number of the backbone amides (NH) in the peptide, excluding proline residues.

3.2.4 Bioinformatic analysis and structural modeling

The Stockholm-format alignment of 5862 sequences in EAL (PF00563) family was downloaded from PFAM database (http://pfam.sanger.ac.uk//family/EAL). Two PERL scripts were used to search for sequences that contain the essential catalytic residues (E^{656}, E^{149}, E^{447}, D^{519}, K^{570}, N^{359}) and the motif DFG[A/S/T][A/G][Y/F][S/T][S/T/G/A/N] that constitutes loop 6. Based on the search results, the 5862 sequences were divided into four groups as we discussed in the results section. For homology modeling, the structural model for the EAL domain of RocR was constructed with the coordinates of TdEAL structure (PDB entry: 2R6O) as template in Swiss-Model (4). The structural model for the RR domain of RocR was built using the structure of the RR domain of CheB (PDB entry: 1A2O) as template.

3.3 RESULTS

3.3.1 Effect of mutations in the loop 6 region

The EAL domain-containing response regulator RocR catalyzes the hydrolysis of c-di-GMP with k_{cat} of 0.67 ± 0.03 s^{-1} and K_m of 3.2 ± 0.3 μM under the enzymatic assay conditions. We previously found that the mutation of E^{268}, a highly conserved residue that was suggested to stabilize loop 6, significantly altered the catalytic activity of RocR (**Fig. 3.2 A**) (132). The mutation E268A reduced the enzyme activity to below the measurable level, whereas the mutation E268Q decreased K_m to 0.31 ± 0.1 μM and caused a 446-fold reduction in k_{cat}. The mutation of F^{297}, a residue resided on loop 6 (DF^{297}GASYSS), reduced k_{cat} to 0.02 ± 0.1 s^{-1} and K_m to 0.73 ± 0.2 μM (**Fig. 3.2 B**). Size-exclusion chromatography showed that mutants E268A, F297A, and the majority of E268Q, formed high-molecular-weight oligomers (HMWO) and eluted out in the void volume. The change in oligomeric state for the three mutants indicated that loop 6 is crucial for maintaining the quaternary structure of RocR. This is in line with the observation that loop 6 is located at the dimer interface as discussed later. Moreover, the side chain of S^{302} from the loop (DFGAGYSS^{302}) interacts with E^{268} through hydrogen bonding according to the EAL_{RocR} structural model and the tdEAL structure (**Fig. 3.2A**). Size-exclusion chromatography showed that the oligomeric state of the protein was not altered by the S302A mutation. Remarkably, the elimination of the single hydrogen bond by the mutation S302A caused the enzyme to exhibit strong substrate inhibition kinetics with an inhibition constant of 10.9 ± 3.0 μM **(Fig. 3.2B).** The observation of substrate inhibition suggested that binding of c-di-GMP is sensitive towards changes of the loop conformation. Finally, the first residue of the loop (D^{296}) may be involved in binding of c-di-GMP directly according to the structural model (132). The D296A mutation caused significant changes in catalytic parameters without altering the oligomeric structure, decreasing k_{cat} by 33.5 fold and increasing K_m to 8.6 ± 2.8 μM **(Table 3.1)**.

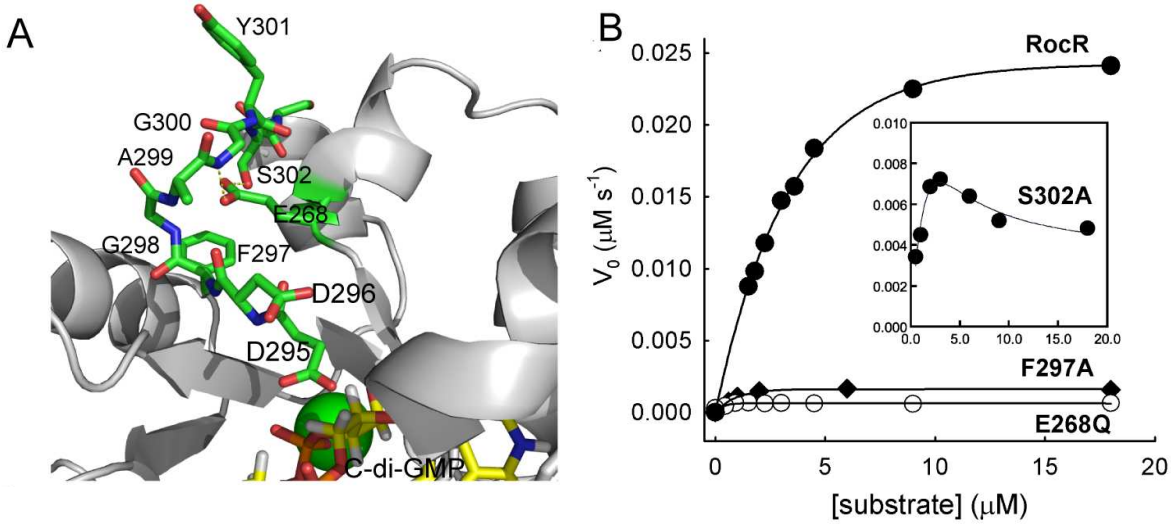


Figure 3.2 Loop 6 of RocR and the effects of mutations on catalysis. (A) Structural model of EALRocR with the residues of loop 6 and E^{268} highlighted. The hydrogen bonds formed between E^{268} and the loop residues G^{300} and S^{302} are represented by the broken lines. The Mg^{2+} ion is shown as the ball. (B) Effects of the mutations in the loop 6 region on the steady-state kinetics of RocR. The curves were generated by fitting the data to the Michaelis-Menten equation with the exception of the S302A mutant, for which the curve was generated by fitting the kinetic data to equation 1.

Table 3.1 Steady-state kinetic parameters for RocR and its mutants[a]

RocR Enzyme	k_{cat} (s^{-1})	K_m (μM)	k_{cat}/K_m (s^{-1} μM^{-1})	K_i (μM)
Wild-type	0.67 ± 0.03	3.2 ± 0.3	0.21± 0.02	
E268A	ND[b]	ND	ND	
E268Q	$(1.5 \pm 0.1) \times 10^{-3}$	0.3 ± 0.1	$(4.8 \pm 0.5) \times 10^{-3}$	
S302A	0.14 ± 0.07	5.5 ± 2.6	$(2.5 \pm 2.0) \times 10^{-3}$	10.9 ± 3.0
D296A	$(2.1 \pm 0.3) \times 10^{-2}$	8.6 ± 2.8	$(2.9 \pm 0.3) \times 10^{-3}$	
D56N/D296A	$(1.2 \pm 0.3) \times 10^{-2}$	9.4 ± 2.3	$(2.9 \pm 0.3) \times 10^{-3}$	
D56N	0.13 ± 0.01	0.2 ± 0.03	$(2.2 \pm 0.4) \times 10^{-2}$	
F297A	0.02 ± 0.01	0.7 ± 0.2	$(2.9 \pm 0.3) \times 10^{-3}$	

[a]Conditions were 100 mM Tris buffer (pH 8.0) (23 °C), 20 mM KCl, 25 mM $MgCl_2$. [b]ND: not determined due to inactivity or extreme low activity (> 10^5-fold less active than wild type)

A large number of EAL domain proteins contain GGDEF-EAL di-domain unit. One of these proteins is PA2567, a *P. aeruginosa* protein that contains a catalytically active EAL domain along with the GAF and GGDEF domains (132) (**Fig. 3.1**). Site-directed mutagenesis and kinetic measurement were performed to test whether loop 6 plays similar roles in PA2567 and other GGDEF-EAL di-domain proteins. Surprisingly, the single mutations of F^{493} and S^{498}, which are the equivalents of F^{297} and S^{302} in RocR, caused the protein to form inclusion bodies during protein expression. Meanwhile, the mutation of E^{464}, the equivalent of E^{268} in RocR, caused strong substrate inhibition at high substrate concentration, with an inhibition constant of 30.4 ± 12.0 μM (**Fig. 3.3, Table 3.2**). The mutation did not perturb the oligomeric state of the dimeric PA2567 according to size-exclusion chromatography. Overall, the results from the studies of PA2567 and RocR strongly suggested that loop 6 plays crucial roles in stabilizing the overall protein structure and participating in catalysis, likely through mediating substrate binding.

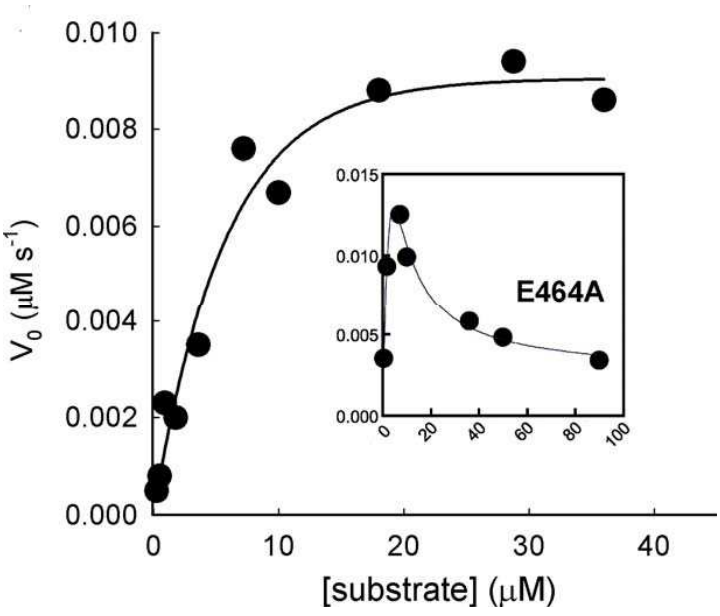


Figure 3.3 Effect of the E464A mutation on the steady-state kinetics of PA2567. The curves were generated by fitting the kinetic data of the wild-type and mutant enzyme to the Michaelis-Menten equation and equation 1, respectively.

Table 3.2 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)
Wild-type	0.39 ± 0.03	3.2 ± 0.3	$(7.5 \pm 2.0) \times 10^{-2}$	
E464A	1.1 ± 0.5	13.0 ± 5.0	$(4.8 \pm 0.5) \times 10^{-2}$	30.4 ± 12.0
S493A	ND[b]	ND	ND	
F498A	ND	ND	ND	

[a]Conditions were 100 mM Tris buffer (pH 8.0) (23 °C), 50 mM KCl, 10 mM $MgCl_2$.
[b]ND: not determined due to inactivity caused by protein instability.

3.3.2 Effect of the D56N mutation on catalytic activity and the loop 6 conformation

Since the small phosphor-donors acetyl phosphate, carbamoyl phosphate, and phosphoramidate did not phosphorylate RocR in the *in-vitro* phosphorylation assay, the effect of phosphorylation site (D^{56}) modification was probed by site-directed mutation. The mutation D56N caused significant changes in the steady-state kinetic parameters (**Table 3.1**). The most significant change is in Michaels-Menten constant K_m, which is 0.20 ± 0.03 μM for the D56N mutant, representing a 16-fold decrease from that of RocR (3.2 ± 0.3 μM). Meanwhile, the D56N mutant only exhibits a slightly smaller k_{cat} (0.13 s^{-1}) than that of RocR (k_{cat} = 0.67 s^{-1}). To probe the structural changes in the mutant that account for the small K_m, the method of hydrogen-deuterium (H/D) exchange coupled mass spectrometry was used, given that this method has been increasingly used as a sensitive tool for detecting changes in protein conformation and mobility (65). By using an established protocol as described in materials and methods, H/D exchange of RocR and the mutant D56N were performed under identical conditions. In total, 21 pairs of peptides generated from pepsin digestion were identified by MS/MS experiment and chosen for H/D exchange analysis. The 21 pairs of peptides encompass 93% of the protein sequences for RocR and the D56N mutant, including peptides 1-6 from the RR domain, peptides 8-21 from the EAL domain and peptide 7 from the linker region (**Fig. 3.4A**). The deuteration pattern of each pair of peptide from RocR and

the D56N mutant was compared by examining the time-dependent deuteration plots. While about two third of the peptides exhibited similar H/D exchange patterns, 8 out of 21 pairs of peptides exhibited notable differences in exchange rate and final deuteration level as we discuss below.

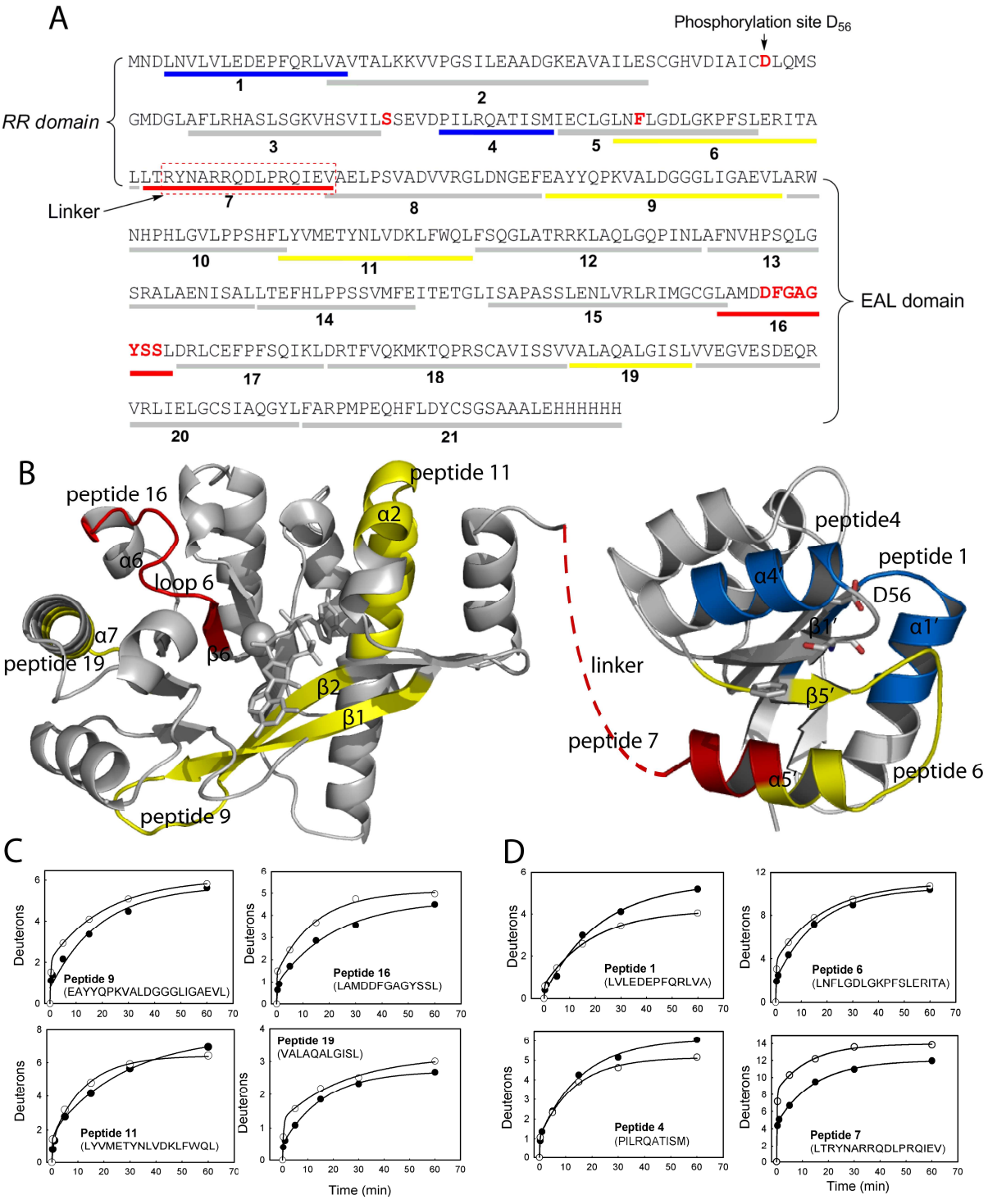


Fig. 3.4 Comparison of amide H/D exchange between RocR and the D56N mutant. (A) Sequence coverage map of RocR with the 21 peptides generated from pepsin digestion

represented by the bars below the sequence. The peptides from the mutant exhibiting decrease in deuteration are colored blue, whereas the peptides showing significant and moderate increase in deuteration are colored red and yellow, respectively. (B) Structural models of the EAL and REC domains of RocR. The peptides exhibiting decrease in deuteration are colored blue, and the peptides showing significant and moderate increase in deuteration are colored red and yellow, respectively. The linker between the RR and EAL domains is represented by the brown broken line. (C) Time-dependent H/D exchange plots for the peptides in the EAL domain that exhibit changes in deuteration (RocR [F] and D56N mutant [E]). The curves were generated by fitting the data to equation 2 to obtain the total number of incorporated deuterons (*N*) as well as the exchange rate constants *k*1, *k*2, and *k*3. (D) Time-dependent H/D exchange plots for the peptides in the phosphoreceiver (REC) domain and linker region that exhibit changes in deuteration (RocR ([F] and D56N mutant [E]). The phosphorylated site (D^{56}) and two putative residues (S^{83} and F^{105}) involved in signal transduction are shown as sticks in panel B.

For the EAL domain, 4 out of the 14 peptides in the D56N mutant exhibited a deuteration pattern different from their counterparts in RocR (**Fig. 3.4B and 4C**). The most prominent one is peptide 16, which exhibited faster exchange and higher deuteration levels in the mutant. Remarkably, peptide 16 includes four residues ($LAMD^{295}$) from the β6 strand that contains the conserved D^{295} for Mg2+ coordination and the D^{296}FGAGYSS motif that constitutes loop 6. In addition to peptide 16, two peptides (peptides 9 and 19) showed modest elevation in deuteration in the mutant. Peptide 9 contains the first two β strands of the $(\beta/\alpha)_8$ barrel and form part of the substrate binding pocket, whereas peptide 19 encompasses half of the α7 helix resided at the dimer interface. The forth peptide (peptide 11) contains helix α2 and forms a lid on the top of the substrate-binding pocket. Peptide 11 in the mutant exhibited slightly faster exchange but lower deuteration at the end of the exchange (**Fig. 3.4C**). The resulted crossover of the two exchange curves probably reflects a change in protein mobility (96). Overall, the results suggested that the D56N mutation induces rather localized structural changes in the EAL domain, with the most significant changes restricted in the region encompassing loop 6 and the end of β6 strand.

Since D^{296} is the most likely residue on the loop interacting with the substrate through a metal ion or water hydrogen bonding network, I tested whether the effect of the D56N

mutation on catalysis is exerted through D^{296}. The double mutant D56N/D296A was prepared and found to exhibit kinetic parameters similar to these of the D296A mutant, with k_{cat} of (1.2 ± 0.3) × 10-2 s-1 and an elevated K_m of 9.4 ± 2.8 μM. Together with the H/D exchange results, the resemblance of the catalytic properties of the double mutant to D296A mutant, rather than the D56N mutant, suggested that the "regulatory signal" originated from the regulatory RR domain probably exerts its effect on substrate binding and catalysis through D^{296} and the adjacent residues in loop 6.

In addition to the four peptides in the EAL domain, four peptides residing in the RR domain and linker region also exhibited altered exchange patterns (**Fig. 3.4D**). Most notably, the peptide spanning the linker region and the end of the α5' helix (peptide 7) showed a significant increase in deuteration in the mutant protein. The deuteration level for peptide 7 is already relatively high (75%) in comparison with other peptides in RocR, indicating a flexible and solvent-exposed linker region. The deuteration level of peptide 7 increased to 90% for the D56N mutant, suggesting that the solvent accessibility of the linker region has been further increased. Moreover, the two peptides (1 and 4) that contain parts of the α1' and α4' helices showed slower exchange rate and lower deuteration level in the mutant, whereas the adjacent peptide 6 that contains the β5' strand, part of α5' helix and the α5'-β5' loop shows moderate increases in deuteration and exchange rate. Overall, the changes in solvent accessibility in the RR domain indicated that the most significant structural changes caused by the mutation D56N occurred within the linker and the α4'-β5'-α5' region.

3.3.3 Recovery of the c-di-GMP specific phosphodiesterase activity of the inactive EAL domain

DGC2 is a cytoplasmic protein involved in the regulation of cellulose biosynthesis in *Acetobacter xylinum*. DGC2 contains a PAS-like domain at the N-terminus and the GGDEFEAL di-domain at the C-terminus (**Fig. 3.1**). It was found that the GGDEF domain of

DGC2 is an active diguanylate cyclase (DGC) domain capable of synthesizing c-di-GMP, and that the EAL domain is catalytically inactive for hydrolyzing c-di-GMP (163). When I cloned, expressed and purified the Dgc2 protein, serendipitously, I found that its PAS domain is associated with flavin. Subsequent collaboration work with my labmate (Qi Yaning) showed that the flavin-containing PAS domain regulates the DGC activity of the GGDEF domain (125). Examination of the sequence of DGC2 suggested that the EAL domain contains the essential general base catalyst (E^{529}) and the residues (E^{350}, N^{409}, E^{441}, and D^{472}) for Mg^{2+} binding, and thus, is a potentially active PDE domain. The stand-alone EAL domain of DGC2 ($DGC2_{306-574}$) was cloned and expressed to further test whether the activity of the EAL domain is repressed by the PAS and GGDEF domains, and was found to be inactive (**Fig. 3.5A**).

Although it contains all the catalytic residues for Mg^{2+} ion binding as well as E^{352}, the EAL domain of DGC2 contains a highly degenerate loop 6 that consists of N^{473}FGKGITVL^{481}, in contrast to the conserved DFG(T/A)GYSS loop in active EAL domains. The observed catalytic activity towards the synthetic substrates such as bis-(p-nitrophenyl) phosphate (125) indicated that the lack of activity towards c-di-GMP is probably caused by the binding of c-di-GMP in an unproductive configuration. To test whether the loss of activity is due to the degeneration of loop 6, we "restored" the loop by mutating three residues in the stand-alone EAL domain. The triple mutant ($N^{473}K^{476}I^{478} \rightarrow D^{473}T^{476}Y^{478}$) of the stand-alone EAL domain was capable of hydrolyzing c-di-GMP (**Fig. 3.5A**). Steady-state kinetic measurement showed that the triple mutant catalyzes the hydrolysis of c-di-GMP with a k_{cat} of (5.9 ± 0.2) ×10-4 s-1 and K_m of 13.0 ± 1.2 μM under the enzymatic assay conditions (**Fig. 3.5B**). The observations suggested that the degeneration of loop 6 is at least partially responsible for the lack of activity for the EAL domain of DGC2, and maybe some other EAL domain proteins as well.

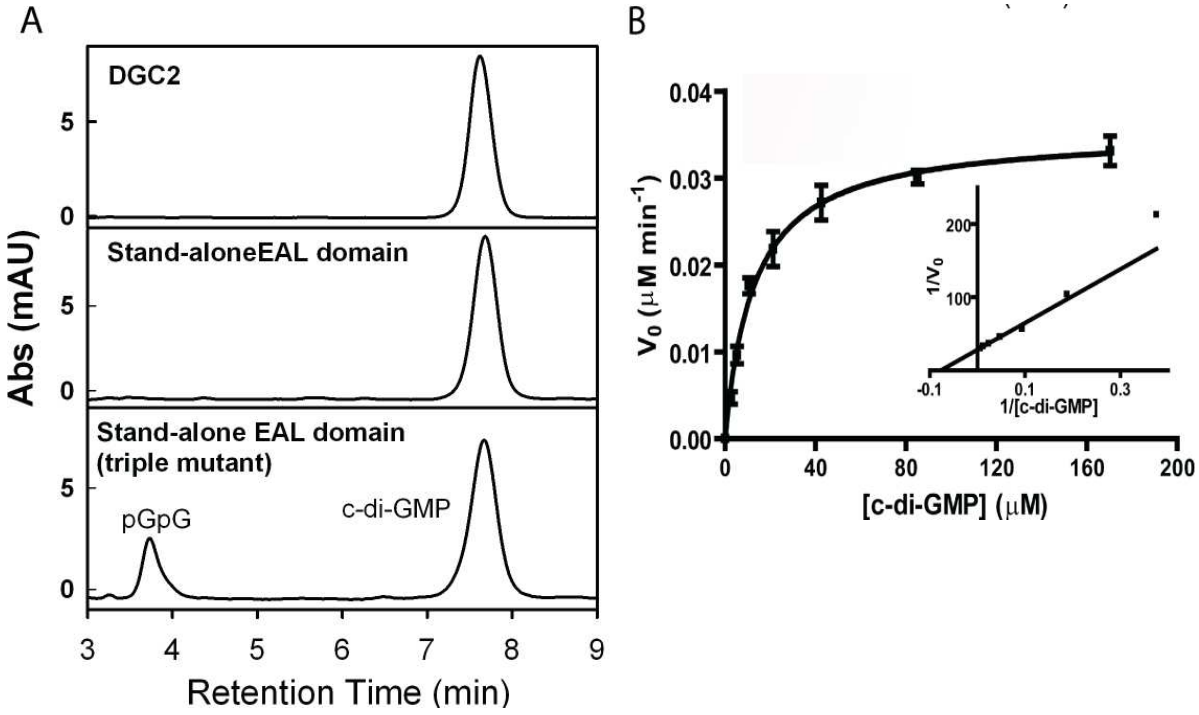


Figure 3.5 Enzymatic activity assay for the EAL domain of *A. xylinus* DGC2. (A) HPLC analysis of the enzymatic activity of the full-length *A. xylinus* DGC2, stand-alone EAL domain, and the triple mutant with three mutations ($Nn^{473}K^{476}I^{478} \rightarrow D^{473}T^{476}Y^{478}$) in the EAL domain. The enzymes were incubated with c-di-GMP and Mg^{2+} for 90 min. Abs (mAU), absorbance (milliabsorbance units). (B) Steady-state kinetics of the triple mutant of the stand-alone EAL domain. The curve was generated by fitting the data to the Michaelis-Menten equation.

3.4 DISCUSSION

It is common to find flexible loops in the vicinity of the active sites of enzymes. The loops play important roles in substrate binding and positioning, as well as shielding the substrate from the bulk solvent. Extensive structural and biochemical studies suggested a functional loop (loop 6) has been preserved during evolution in $(\beta/\alpha)_8$-barrel fold enzymes (74, 158). Despite the diverse reactions catalyzed by $(\beta/\alpha)_8$-barrel fold enzymes, the flexible loop 6 plays various functional roles assisting catalysis by undergoing significant structural or conformational changes during substrate binding and catalysis. The results presented here support an equally important role for loop 6 in the hydrolysis of c-di-GMP for the $(\beta/\alpha)_8$ barrel folded EAL domains.

3.4.1 Effect of mutations in loop 6 on catalysis

Single mutations in the loop 6 region had profound impact on the structure and enzymatic activity of RocR. The mutation of the loop residue F^{297} and the loop-stabilizing E^{268} caused the formation of HMWO and significantly changed both k_{cat} and K_m. The mutation of E^{268} and F^{297} increased the propensity of RocR to form HMWO suggested a critical structural role played by loop 6 in maintaining the quaternary structure of the protein. 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 (PDB entry: 2R6O) and EAL_{YkuI} (PDB entry: 2BAS) (105). The anti-parallel loop-loop interaction between the residues from loop 6 observed in the structures is likely to stabilize the dimeric structure. It is conceivable that the mutation of E^{268} or the residues on the loop could 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 D^{296} and S^{302} in RocR, led to significant changes in catalytic parameters without altering the oligomeric structure. The S302A mutant exhibited strong substrate inhibition with inhibition constant (K_i) of 10.9 ± 3.0

μM while D296A mutant showed greater K_m (8.6 ± 2.8 μM) and small k_{cat} (0.02 s^{-1}). The observed substrate inhibition and greater K_m strongly suggested that the loop is involved in the binding of c-di-GMP and that the perturbation of the loop can either strengthen or weaken the binding affinity for the substrate. Similar results were obtained with the GGDEF-EAL di-domain-containing PA2567. The mutation of E^{464}, the equivalent of E^{268} in RocR, resulted in strong substrate inhibition; whereas the other two mutations seemed to affect protein folding, presumably due to the disruption of dimerization. The effect of the mutations in RocR and PA2567 unequivocally established the crucial roles of loop 6 in both structure maintenance and catalysis.

3.4.2 Conformational changes of the D56N mutant

The comparison of the H/D exchange pattern between the peptides from RocR and the D56N mutant revealed intriguing structural changes in both the RR and EAL domains as an effect of a single mutation. The most significant and intriguing changes in the EAL domain are in loop 6 and the end of β6 strand **(Fig. 3.4B and 3.4C)**. The enhanced H/D exchange in the region reflects increased solvent accessibility in this region as a result of structural or conformational changes. Moderate changes were also seen near peptide 9, 11 and 19. Some of the residues of peptide 9 and 11 are located in the substrate binding pocket and it is conceivable that the changes in conformation or mobility of these residues can affect substrate binding. The slightly higher deuteration for α7 helix (peptide 19), along with the great increase in solvent accessibility in the linker region (peptide 7), probably reflects a rearrangement of the dimer interface. Such interface conformation change could be coupled to the local conformational change in loop 6, considering that part of loop 6 is located at the interface as well. Indeed, the coupling of domain rearrangement and movement of D^{295} has been suggested to be critical for controlling EAL domain activity by comparing the structures of YkuI and tdEAL (28). Together with the change in K_m caused by the mutation D56N, the

observed change in solvent accessibility suggested that the catalysis is sensitive to conformational changes in the substrate binding pocket, especially in the vicinity of loop 6. Moreover, the comparison of H/D exchange patterns also revealed significant changes in solvent accessibility in the linker and the α4'-β5'-α5' region of the RR domain. Interestingly, structural and biochemical studies have shown that the α4'-β5'-α5' region undergoes conformational changes upon activation by phosphorylation or the binding of the phosphoryl mimic beryllium fluoride (BeF_3) in RR domains (25, 54). The observed conformational changes in the α4'-β5'-α5' face raised the tantalizing possibility that the D56N mutant might be a constitutively active mutant. However, this remains speculative before the comparison between the mutant and the phosphorylated RocR proteins in the future. Another underlying implication of these observations is that the control of catalysis through loop 6 might be used by EAL domains as a regulatory mechanism.

3.4.3 Loop 6 is involved in regulating catalysis

In contrast to RocR and PA2567, the EAL domain of DGC2 represents one of the EAL domains that contain the essential residues for Mg^{2+} binding and general base catalysis, but a highly degenerate loop 6. The observation that DGC2 is active against non-physiological phosphodiesterase substrates indicated that the loss of activity towards c-di-GMP is most likely due to alteration of substrate specificity. This view is further supported by the partial recovery of activity for the triple mutant of the stand-alone EAL domain by "restoring" the degenerate loop. The triple mutant of EAL_{DGC2} exhibited catalytic activity toward c-di-GMP. Although the catalytic rate for the triple mutant is considered to be low, the K_m for the triple mutant is 13.0 ± 1.2 μM and only two to three fold higher than the K_m for RocR and PA2567, indicating that the triple mutant has reasonable binding affinity for c-di-GMP. These observations confirmed the importance of loop 6 in catalysis and further suggested that the

restoration of loop 6 enables the EAL domain to bind c-di-GMP in a productive configuration for hydrolysis, but not to improve affinity directly.

The functional roles of loop 6 in EAL domains can be rationalized in light of the structural data for TdEAL and the YkuI-substrate complex. Although its catalytic activity has not been experimentally validated, the tdEAL domain is most likely to be an active PDE domain with conserved loop 6 and the catalytic residues. By contrast, the EAL domain of YkuI, which contains a degenerate loop 6 (NIGKESS), is catalytically inactive despite its ability to bind c-di-GMP (105). The comparison of the structures of the active (TdEAL) and inactive (YkuI) EAL domain structures revealed some intriguing differences for loop 6 (**Fig. 3.6**). One of the major differences is that the "anchor" residue E^{125} (E^{268} in RocR) is no longer tucked under loop 6 in YkuI. As a result, the loop adopts a very different conformation with the loop residues shifted away from the catalytic Mg^{2+} ion. The catalytically indispensable residue D^{152} (D^{295} of RocR) in YkuI is too far away from coordinating the catalytic Mg^{2+} ion as suggested by Minasov and coworkers (105). In addition, the residue N^{153} (D^{296} in RocR), which was suggested to interact the substrate c-di-GMP based on the increased K_m of D296A mutant (132), has been pulled away from the substrate. Based on these observations, we postulate that loop 6 of EAL domain is not only crucial for stabilizing the dimer interface but also involved in the binding of Mg^{2+} ion and c-di-GMP. The conformational change in loop 6 caused by mutation or regulatory signal will impact catalysis in a several 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.

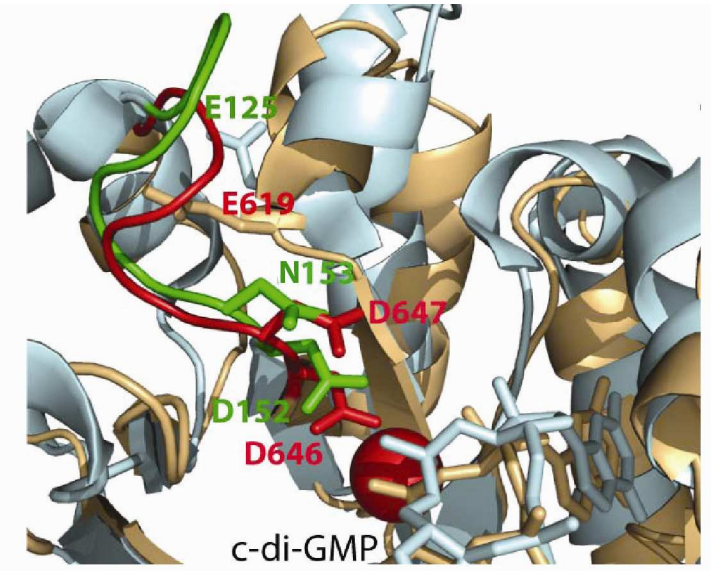


Figure 3.6 Comparison of loop 6 and part of the proceeding β6 strand in active and inactive EAL domains. Loop 6 of TdEAL (Protein Data Bank accession no. 2r6o) and YkuI (Protein Data Bank accession no. 2w27) were highlighted in orange and cyan, respectively. The Mg^{2+} ion in the TdEAL structure is shown as a green ball, and c-di-GMP is shown as a stick. The corresponding residues of D^{295}, D^{296}, and E^{268} of RocR are highlighted and labeled (D^{646}, D^{647}, and E^{619} for TdEAL; D^{152}, N^{153}, and E^{125} for YkuI).

Second, the binding affinity for c-di-GMP can be altered due to the change of the interaction between c-di-GMP and the loop. Structural and biochemical data suggested that the loop is likely to interact with c-di-GMP through the conserved Asp (Asp^{296} of RocR). Considering the large structural changes observed for loop 6 in other $(\beta/\alpha)_8$-barrel enzymes, it is also plausible that the binding of c-di-GMP in active EAL domains may induce even greater conformational change in loop 6 and result in additional interactions between the loop and c-di-GMP. Thus, mutations that either affect the direct interactions between loop 6 and substrate or impair the ability of loop 6 to undergo conformational change for effective substrate binding would affect catalysis. Second, the mutations in the loop can hamper catalysis by dislocating the essential Mg^{2+} ion-binding residue (D^{295} in RocR). The reduced binding affinity for Mg^{2+} can have catastrophic effects on the catalytic activity as demonstrated by the detrimental effect of D295N mutation in RocR. Given the sensitivity of the catalytic activity to the loop conformation and the position of the Mg^{2+} binding Asp

residue, it is intriguing to speculate whether some EAL domains are regulated through controlling the conformation of loop 6.

3.4.4 Classification of EAL domains

Comparison of the sequences of EAL domains revealed that all the characterized inactive EAL domains harbor a degenerate loop 6 (**Fig. 3.7A**). Further sequence analysis of the 5862 EAL domain found in bacterial genomes revealed that 4809 of them contain all the catalytic residues including the general base catalyst and four essential residues for Mg^{2+} coordination (**Fig. 3.7B**). Among the 4809 EAL domains, 2895 of them contain a conserved loop 6 (DFG[A/S/T][G/A][Y/F][S/A/T][S/A/G/V/T]) and 1914 of them contain a degenerate loop 6. The 2895 EAL domains are likely to be catalytically active PDE domains; whereas the functions of the 1914 EAL domains with a degenerate loop 6 could be inactive PDEs as exemplified by the EAL domains of YkuI and DGC2. The EAL domains of YkuI and DGC2 were found to be catalytically inactive toward c-di-GMP under *in-vitro* conditions (105, 125). However, it remains to be seen whether the PDE activity can be stimulated by the putative regulatory domains. In addition to the 4809 EAL domains, 1053 EAL domains were found to lack one or more of the essential catalytic residues (**Fig. 3.7B**). Among the 1053 EAL domains, the majority of them (990) do not contain a conserved loop 6. These 1053 EAL domains are most likely to be catalytically inactive. This postulation is supported by previously characterized inactive EAL domains and two recently reported EAL domain proteins (132). The EAL domain of the protein LapD from *P. fluorescens* Pf0-1 contains a highly degenerate loop 6 (QRFGGRFSM) 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 (112). 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 E^{352} in RocR and a highly degenerate loop 6 (QVGRDLQI). Recent studies

showed that the EAL domain does not exhibit PDE activity but participate in protein-protein interaction instead (174). Given the strong correlation between the divergences of loop 6 sequences and catalytic activity, the prediction of PDE activity based on the categorization described here will provide some assistance in elucidating the biological functions of the versatile EAL domains.

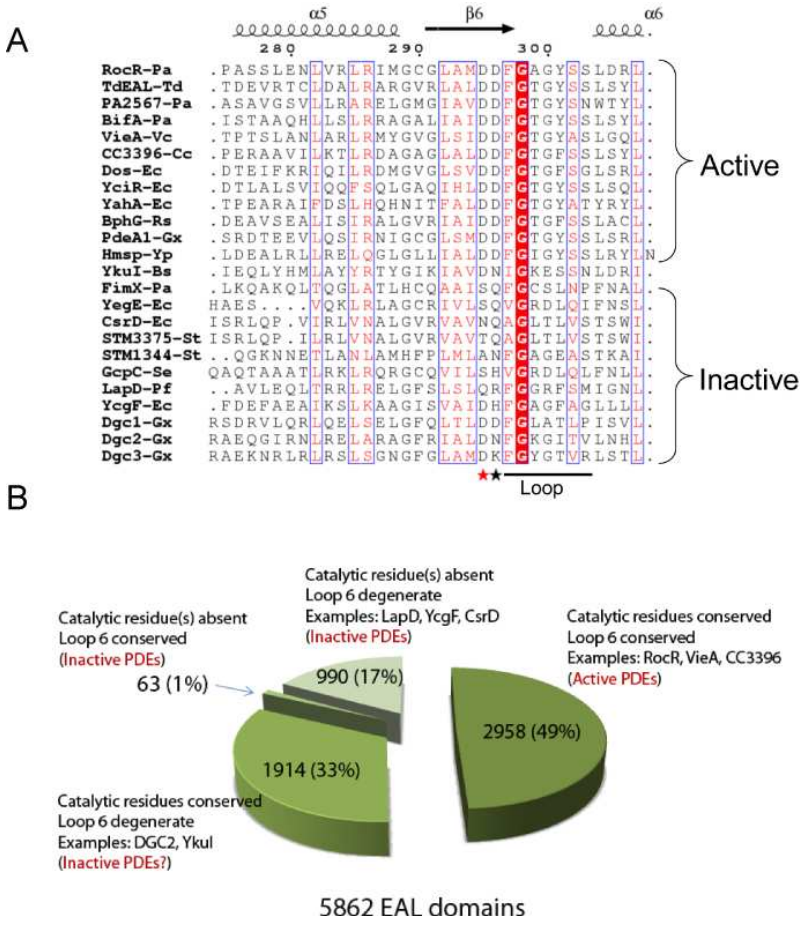


Figure 3.7 Summary of EAL domains. (A) Comparison of the sequences of the loop among characterized EAL domains. The EAL domains shown are from the following proteins: PA2567, BifA, and FimX from *Pseudomonas aeruginosa* (indicated by the Pa suffix after the hyphen and protein) (76, 84, 143); TdEAL from *Thiobacillus denitrificans* (TdEAL-Td); VieA from *Vibrio cholerae* (VieA-Vc) (165); CC3396 from *Crescentus caulobacter* (CC3396-Cc) (28); YcgF, YahA, Dos, YciR, CsrD, and YegE from *Escherichia coli* (indicated by the Ec suffix after the hyphen and protein) (52, 146, 162, 174, 179); BphG from *Rhodobacter sphaeroides* (BphG-Rs) (169); PdeA1, DGC1, DGC2, and DGC3 from *Gluconacetobacter xylinus* (indicated by the Gx suffix after the hyphen and protein) (23, 163); HmsP from *Yersinia pestis* (Hmsp-Yp) (12); GcpC from *Salmonella enterica* (GcpC-Se) (45); STM1344 and STM3375 from *Salmonella enterica* serotype Typhimurium (indicated by the St suffix after the hyphen and protein) (149); and LapD from *Pseudomonas fluorescens* Pf0-1 (LapD-Pf) (112). The sequences were aligned using MultAlin (http://bioinfo.genopole-

toulouse.prd.fr/multalin/multalin.html), and the figure was generated using ESPript 2.2 (http://espript.ibcp.fr/ESPript/ESPript/). (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.

3.4.5 A general regulatory mechanism for EAL domain?

EAL domains are generally present in proteins together with other sensory domains such as PAS, GAF, REC, BLUF. They are also frequently found in proteins as part of a GGDEF-EAL di-domain. The PDE activity of some of the EAL domains is likely to be regulated by the sensory domain. For example, the activity of the *Ax*PDEA1 protein from *A. xylinus* was positively regulated upon oxygen binding to the heme cofactor in the PAS domain (23). Oxygen also regulates the c-di-GMP PDE activity of *Ec*Dos, through binding to its heme containing PAS domain (167). For the di-domain containing protein Cc3396 from *C. crecentus*, GTP binding to the catalytically inactive GGDEF domain lowered the K_m of c-di-GMP hydrolysis from 100 μM to sub-micromolar (28), which is the physiologically relevant level (59).

Although several signals have been identified to control EAL domain activity, how the signals trigger allosteric conformational changes in the EAL domain to achieve regulation remained elusive until the recent studies on the BLUF-EAL protein BlrP1, and the REC-EAL protein RocR described here (7, 128). These two studies pointed to quaternary structure change at a common region of the EAL domain as critical for allosteric signal transduction. The region consists of a interface helix and a conserved loop 6 at the subunit interface.

The structural study of the BLUF-EAL protein BlrP1 by Barends *et al* predicts that light stimulated conformational change in the C-terminal helix of BLUF domain, which is at the inter-monomer BLUF-EAL interface, leads to a relative repositioning of the EAL-EAL dimerization interface: α6, α7, β8 and loop 6 (7). Such quaternary structural changes are believed to be important to position three catalytic residues: E^{352}, D^{295} and D^{296}. The coupled interface repositioning is made possible because of the structural proximity between the

BLUF-EAL and EAL-EAL interface (**Fig. 3.8B**). Indeed, when two interacting residues (L^{128} from BlrP1 BLUF domain, and G^{353} from BlrP1 EAL domain) were mutated to cysteines, the mutant did not respond to light and displayed constitutively higher activity.

Our characterization of the D56N mutant of RocR also revealed that the decrease in K_m correlates with a change of solvent accessibility in the loop 6 region and also part of the dimerization helix α7, as measured by H/D exchange method (128). Due to its smaller and physiologically relevant K_m (59, 128), the D56N mutant may represent a constitutive active form of RocR. The results thus suggested the same set of conformational changes triggered by different signals (light for BlrP1, phosphorylation for RocR) may serve to regulate catalytic residues E^{352}, D^{295} and D^{296}, to give rise to a constitutively active form.

To compare the inter-domain relay of conformational change between BlrP1 and RocR, a docking model showing the relative position of the REC and EAL domain was generated using Patchdock. When regions that displayed conformational change measured by H/D exchange were mapped to the highest ranked model, some interesting findings are worth noting (**Fig. 3.8A**). First, the REC-EAL interface of the RocR model involves the α4'-β5'-α5' surface of REC domain. Previous structural and biochemical studies have shown that the α4'-β5'-α5' region undergoes conformational changes upon activation by phosphorylation or the binding of the phosphoryl mimic beryllium fluoride (BeF_3) in RR domains (25, 44, 54). Thus the observed conformational change in this region upon D56N mutation probably reflects the activated state. The flip of a conserved aromatic residue (Phe or Trp) on β5 is believed to be the trigger for the activation of prototypical REC domain protein, CheY (15, 93). In our model, the corresponding Trp residue (stick in magenta) is in close proximity to the C-terminal EAL domain as well (**Fig. 3.8 A**). Second, the relative position of the N-terminal sensor domain to the catalytic EAL domain is strikingly similar between RocR model and BlrP1 structure, in agreement with the idea that the sensor domain affects the same region of EAL domain (helix α7, and the α7-β8 loop, **Fig. 3.8**). Due to the central role of α7 in

dimerization, it plays a key role in maintaining the EAL-EAL interface and conformation of loop 6 and part of β6. All the changes described here are mapped to the BlrP1 structure before (7), and are also implicated in RocR by H/D exchange experiment (128) (**Fig. 3.8A** and **Fig. 3.4**). In fact, a similar contact between accessory and catalytic EAL domain can also be found in YkuI (**Fig 3.8 C**), as has been noted by Barends et al (7). Since the N- and C-terminal end of the $(\beta/\alpha)_8$ barrel fold EAL proteins are spatially close, it is thus reasonable to further postulate that, the same relative position of the accessory domain and thus the same regulatory mechanism could be generally applicable (7), regardless of whether the sensing domain is located at the N- or C-terminus. This postulation will be borne out by more high-resolution structures of EAL domain proteins in the future.

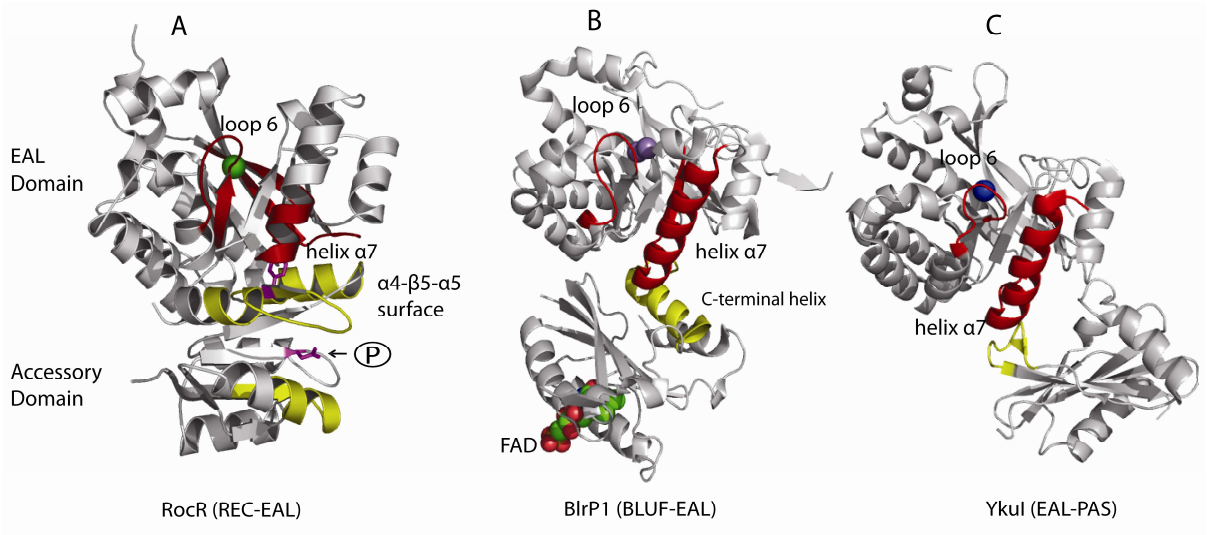


Figure 3.8 Domain interface of EAL proteins: RocR (A), BlrP1 (B), and YkuI (C). For RocR, regions displayed conformational change between wildtype and D56N mutant are colored yellow (in the REC domain) and red (in the EAL domain). For BlrP1 and YkuI, regions involved in interdomain contact are colored yellow (in the BLUF/PAS domain) and red (in the EAL domain).

3.5 SUMMARY

EAL domain-based c-di-GMP specific phosphodiesterases play important roles in bacteria by regulating the cellular concentration of the dinucleotide messenger c-di-GMP. The EAL domains belong to a family of $(\beta/\alpha)_8$ barrel fold enzymes that contain a functional active site loop (loop 6) for regulating substrate binding. In this chapter, I show that, by examining the two EAL domain-containing proteins RocR and PA2567 from *P. aeruginosa*, the catalytic activity of the EAL domains was significantly altered by mutations in the loop 6 region. The impact of the mutations ranges from apparent substrate inhibition to alteration of oligomeric structure. Moreover, the catalytic activity of the RocR was affected by mutating the putative phosphorylation site (D56N) in the phosphoreceiver domain, with the mutant exhibiting a significantly smaller K_m than that of the wild type RocR. Hydrogen-deuterium (H/D) exchange by mass spectrometry revealed that the decrease in K_m correlates with a change of solvent accessibility in the loop 6 region. I further examined the *A. xylinus* protein DGC2, which represents the proteins that contain a catalytically incompetent EAL domain with a highly degenerate loop 6. The catalytic activity of the stand-alone EAL domain of *Ax*Dgc2 could be partially recovered by restoring loop 6. In conjunction with the structural data of two EAL domains, it is reasonable to propose that loop 6 not only mediates the dimerization of EAL domain, but also controls c-di-GMP and Mg^{2+} ion binding. Importantly, sequence analysis of the 5862 EAL domains in the bacterial genomes revealed that about half of the EAL domains harbor a degenerate loop 6 and are potentially inactive EAL domains. The mutations in loop 6 may therefore represent a divergence of catalytically active and inactive EAL domains during evolution.

Chapter Four: YybT is a signalling protein that contains a cyclic dinucleotide phosphodiesterase domain and a GGDEF domain with ATPase Activity

4.1 INTRODUCTION

The cyclic dinucleotide c-di-GMP has been firmly established as a major bacterial messenger molecule in recent years, with the cellular level of c-di-GMP regulated by diguanylate cyclase and phosphodiesterase domain proteins (43, 59, 138). In contrast, the existence of the structurally similar 3', 5'-cyclic diguanylate (cyclic di-AMP or c-di-AMP) in living cells was unknown until the recent serendipitous discovery of the dinucleotide bound by the DisA protein from *Bacillus subtilis (9, 183)*. It was found that c-di-AMP was synthesized by the diadenylate cyclase (DAC) domain of DisA *via* condensation of two ATP molecules. Witte et al suggested that c-di-AMP is involved in signaling DNA damage considering that the DNA integrity scanning protein DisA scouts the chromosome for DNA double-strand breaks (183). Subsequent genomic mining revealed that the DAC domain proteins are widespread in bacteria and archaea with many of them associated with putative sensor domains (136). The wide occurrence and domain architecture of DAC domain proteins hinted that c-di-AMP may be another hidden nucleotide messenger mediating various cellular functions and phenotypes.

Currently, there is no report of c-di-AMP degrading or exporting proteins for controlling cellular c-di-AMP level. To identify potential c-di-AMP degrading proteins, we searched bacterial genomes for phosphodiesterase or phosphoesterase proteins that co-occur with the DAC domain-containing proteins *(136)*. We found that a group of multidomain proteins seem to co-occur with a subset of the DAC domain proteins that include the homologs of YojJ and YbbP from *B. subtilis*. This group of proteins (COG3387), as represented by the *B. Subtilis* protein YybT, contains two N-terminal trans-membrane (TM) helices, a region that shares minimum sequence homology with some Per-Arnt-Sim (PAS) domains, a highly modified

GGDEF domain, and a DHH/DHHA1 domain (**Fig. 4.1**). We were particularly attracted to this family of proteins because the DHH domain proteins are known to function as phosphatases or phosphoesterases hydrolyzing a wide variety of substrates ranging from inorganic pyrophosphate to single-stranded (ss) DNA (3). The DHH family proteins, which were named after the conserved Asp-His-His motif in the active site, are divided into two subfamilies based on the C-terminal sub-domain. Subfamily 1 contains a DHHA1 domain at the C-terminus, as represented by the single-stranded DNA exonuclease RecJ and 3'-phosphoadenosine 5'-phosphate (pAp) phosphatase (59, 187). Subfamily 2 contains a DHHA2 domain at the C-terminus, as exemplified by the type-II inorganic pyrophosphatase (192), yeast cytosol exopolyphosphatase (40), and the eukaryotic c(A/G)MPase protein Prune (30). The C-terminus of YybT belongs to the DHHA1 subfamily considering that it contains the conserved GGGH motif and shares moderate sequence homology with the C-termini of RecJ and pAp phosphatase (3). YybT homologs can be identified in pathogens such as *Staphylococcus aureus, Staphylococcus epidermidis* and *Clostridium difficile*. Interestingly, YybT and its homolog in *S. Aureus* (GdpP) have been speculated to be a c-di-GMP phosphodiesterase by two different groups (6, 64).

In this chapter, I present results from the study of the cytoplasmic portion of YybT from *B. Subtilis*. I demonstrate that the DHH/DHHA1 domain exhibited specific phosphodiesterase activity toward the cyclic dinucleotides c-di-AMP and c-di-GMP among the possible physiological substrates. Biochemical and bioinformatic evidence argues that the physiological substrate is likely to be c-di-AMP but not c-di-GMP. In addition, the GGDEF domain of YybT possesses unprecedented ATPase activity, rather than diguanylate cyclase (DGC) activity. The biological function of YybT and the evolutionary relationships of the DHH/DHHA1 and GGDEF domains are discussed in light of the enzymatic activities and the effect of ΔYybT mutation on sporulation.

4.2 MATERIALS and METHODS

4.2.1 Materials

Chemicals were purchased from Sigma with the exception of sodium tripolyphosphate (Alfa Aesar), ss RNA and DNA (Eurogentec), c-di-AMP (BioLog), ppGpp (TriLink) and c-di-GMP, which is synthesized using the thermophilic DGC enzyme in the lab (127).

4.2.2 Genomic DNA isolation

Bacillus subtilis strain 168 was grown in LB medium. Cells from 5 ml overnight culture was pelleted and resuspended in Tris-EDTA buffer. The suspension was incubated for one hour with 6% volume of 10% SDS and 0.6% vol. of 20 mg/ml proteinase K at 37 ^{o}C. After incubation, an equal volume of phenol/chloroform was added to the mixture. The DNA phenol mixture was spun at 14,000 rpm for 2 min. The upper aqueous phase was mixed with an equal volume of phenol/chloroform and centrifuged again. A 1/10 volume of sodium acetate (3M, pH 5.2) was then added to the upper aqueous phase, followed by 0.6 ml of isopropanol to precipitate the genomic DNA. DNA was pelleted by centrifugation and washed by using 70% ethanol. Washed DNA was re-dissolved in TE buffer for PCR cloning of the YybT gene.

4.2.3 Cloning, expression and purification

The gene (*yybT*) that encodes the full-length YybT was first amplified by PCR from the genomic DNA using the Expand High-Fidelity Kit (Roche). Subsequently, the DNA fragments that encode the cytoplasmic portion of YybT ($YybT_{84-659}$), the GGDEF-DHH/DHHA1 di-domain ($YybT_{150-659}$) and the PAS-GGDEF di-domain ($YybT_{84-303}$) were cloned into the expression vector pET-28(a+) (Novagen) between the NdeI and XhoI restriction sites. The plasmids harboring the DNA constructs and the $(His)_6$-Tag encoding sequence were transformed into *E. coli* strain BL21(DE3). The plasmid for expression of

catalytic domain of ttRecJ was a gift from Prof. K. Fukuyama and the protein was expressed and purified according to Yamagata *et al* (186). The gene encoding the DHH/DHHA1 domain of the CCA-adding tRNA nucleotidyltransferase from *Aquifex aeolicus* (aa 1-418) was custom synthesized by GeneScript, its protein expression and purification will be described elsewhere by Zhang, D., and Liang, Z-X. Expression and purification of YtqI was described in Rao *et al* (130).

For YybT protein expression, one liter of bacterial culture (LB medium) was grown to OD = 0.8 before inducing with 0.8 mM IPTG at 16 ^{o}C for ~12 hours before. After pelleting by centrifugation, the cells were lysed in 20 ml lysis buffer (50 mM Tris (pH 8.0), 150 mM NaCl, 5% Glycerol, 0.1% β-mercaptoethanol, and 1 mM PMSF). After the centrifugation at 25,000 rpm for 30 min, the supernatant was filtered and then incubated with 2 ml of Ni-NTA resin (Qiagen) for 1 hour at 4 ^{o}C. The resin was washed with 50 ml of W1 buffer (lysis buffer with 20 mM imidazole) and 20 ml of W2 buffer (lysis buffer with 50 mM imidazole). The bound proteins were eluted using a step gradient method with elution buffers containing 50 mM Tris (pH 8.0), 150 mM NaCl, 5% glycerol and 200mM, 300 mM, or 500 mM imidazole. After the analysis by SDS-PAGE, fractions with purity higher than 95% were pooled together. Size-exclusion chromatography was carried out at 4 ^{o}C using the AKTA FPLC system equipped with a Superdex 200 HR 16/60 column (Amersham Biosciences). The buffer used for gel filtration was comprised of 50 mM Tris-HCl (pH 8.0), 0.15 M NaCl, 5% Glycerol. The proteins were stored in -80 ^{o}C freezer after flashing freezing and concentration measurement by Bradford assay method.

All protein mutants were generated using the Site-Directed Mutagenesis II Kit (Stratagene) according to manufacturer's instruction manual. Mutations were verified using BigDye Terminator v3.1 Cycle Sequencing Kit on ABI Prism 3100 Genetic Analyzer (Applied Biosystems). Mutant proteins were expressed, purified and stored under the same conditions as wild type.

4.2.4 Bioinformatics and structure modeling

Sequence and genomics analysis was performed using STRING and ClustalW (71, 91). In total, 184 YybT homologs (defined as proteins share the same domain architectures) are found in bacterial genomes. The HMM logo was generated by from the alignment file by using webpage http://weblogo.berkeley.edu/logo.cgi. The structural model of the DHH_{YybT} was built using Swiss-Model sever with the structure of a DHH family protein (PDB entry: 3DMA) as template (4).

4.2.5 Substrate screening by enzymatic activity assays

P-nitrophenol-related substrates *p*NPP, bis-*p*NPP, *p*N-TMP, PEP were assayed by monitoring the formation of *p*-nitrophenol at 410 nm using a UV-Vis Spectrophotometer. A control with the protein storage buffer but not the enzyme solution was used for baseline correction. The enzymatic activity assays with nucleotides and oligos were carried out by using a HPLC system (Agilent LC1200) with a RP C-18 column as previously described for c-di-GMP activity assay (28, 29). The nucleotides and oligos tested are listed in Table S1. The activity towards RNA and ssDNA was monitored by detecting the possible products 1-mer and 2-mer nucleotide along with standards. For inorganic pyrophosphate and polyphosphate, assay was conducted using the EnzCheck Phosphate Assay Kit (Invitrogen) by measuring the release of inorganic phosphate (P_i) in solution. The assay conditions for substrate screening were: 100 mM Tris (pH8.3), 20 mM KCl, 0.5 mM $MnCl_2$, 1-2 μM enzyme, and 100 μM substrates (10 μM for 24 mer RNA and T7 promoter ssDNA, and 1 mM for pyrophosphate and tripolyphosphate). Reactions were incubated up to 2 hours for detecting enzymatic activity.

4.2.6 Metal and pH Dependence

Assay conditions used for metal screening were: 100 mM Tris (pH8.3), 20 mM KCl, 10 mM [$Metal^{2+}$], 1 μM enzyme, 20 μM c-di-AMP. For pH screening, assay buffer contains 100 mM Bis-Tris (for pH 6.7-7.3), Tris (for pH 7.3-8.8), or CHES (pH 8.8-10) with 0.5 mM $MnCl_2$ and 20 mM KCl.

4.2.7 Kinetic measurement of the DHH domain activity

The measurement of steady-state kinetic parameters were carried out by monitoring the formation of 5'-pApA/5'-pGpG (for DHH/DHHA1 domain) using HPLC. Assay buffer conditions are the same as used for substrate screening. Initial velocity at a certain substrate concentration was obtained from a series of reactions that were stopped at various time points within linear range. Initial velocity was measured at 7-8 substrate concentrations. The kinetic parameters k_{cat} and K_m were obtained by fitting the initial velocity at various substrate concentrations to Michaels-Menten equation using the software Prism (GraphPad). For ppGpp inhibition assay, the initial velocity for c-di-AMP hydrolysis (at 20 μM) was first measured in the presence of 0-2 mM ppGpp under the same reaction conditions used for steady-state kinetic measurement. The IC_{50} was obtained by fitting the data to the dose-response model using the software Prism (GraphPad). The steady-state kinetics in the presence of two ppGpp concentrations was measured with K_i obtained from data fitting.

4.2.8 Kinetic measurement of the ATPase activity

The ATPase activity of the GGDEF domain was assayed by two methods. The hydrolysis of ATP was monitored by measuring the release of P_i in solution using the EnzCheck Phosphate Assay Kit (invitrogen) and a UV-Vis spectrophotometer, or directly monitoring ADP formation using HPLC with a solvent system described by Ryjenkov et al (144). The measurement of steady-state kinetic parameters was carried out by monitoring the formation

ADP using HPLC. All kinetic measurements were carried out at 23^{o} with the assay conditions: 50 mM Tris (pH 7.5), 1 mM $MgCl_2$, 0.05 - 8 mM ATP, 2.9 μM enzyme. Competitive inhibition was carried out by in the presence of 250 μM AMPPNP. The dissociation constants (K_d) were calculated using the equation: K_d = [Inhibitor]/($K_{m(i)}/K_m$ - 1).

4.2.9 Sporulation of wild type and ΔyybT mutant in the presence of nalidixic acid

The spore formation efficiency of the wild type *B. subtilis* strain 168 and YybT mutant strain (ΔYybT) (a gift from K Kobayashi, Nara, Japan) was examined by using an established procedure (9). The ΔYybT mutant is from a complete mutant library of *B. subtilis* 168 with an insertion of pMUTIN3MCS after the codon for amino acid 88 (79). Growth and sporulation were carried out with shaking at 32 °C in casein hydrolysate medium and sporulation medium, respectively (9). Sporulation was induced by transferring the early log-phase cells from the CH medium to the sporulation medium with or without nalidixic acid at the indicated concentration. After 22 hours, cells were heated at 80 °C for 10 min and then plated with proper dilutions on LB agar plates (number of colonies < 500 per plate). The experiment was repeated three times. The sporulation efficiency was estimated by calculating sporulation (%), which equals (number of colonies with heating)/(number of colonies without heating) × 100.

4.2.10 Acid stress resistance

The procedure was adopted from a previous report (126). The wild type and ΔYybT mutant B. subtilis cells grown from an overnight culture at pH 7.0 were inoculated into LB medium (pH 4.0, 3.5, or 3.0) at final A_{600nm} = 0.1. To determine viability by colony formation, the cells were plated on LB agar plate (pH7.0) with proper dilution both prior to (T0) and after acid challenge for 2 h (T2). Relative survival was calculated as the ratio of colony-forming units/ml at time T2 to that at T0.

4.2.11 Measurement of the Cellular c-di-AMP Concentration

The procedure for measuring the cellular concentration of c-di-AMP in *B. subtilis* was essentially the same as the one reported by Simm *et al* (152). Briefly, the vegetative and sporulating cells harvested from the casein hydrolysate and sporulation medium were pelleted and resuspended in double distilled H_2O (100 mg of cells/300 µl of water). The suspension was heated at 95°C for 15 min, followed by sonication for 5 min. 99% ethanol was then added to the sample to a final concentration of 70%. After centrifugation, the supernatant was pooled, frozen, and subsequently lyophilized overnight. The precipitate was dissolved in 500 µl of double distilled H_2O, filtered, and loaded to the HPLC system. Because no clear c-di-AMP peak was observed on HPLC trace, a 4-min fraction corresponding to the retention time of c-di-AMP (±2 min) was collected from HPLC, lyophilized and redissolved into 5-10 µl of double distilled H_2O. For mass spectrometry, the sample was mixed with an equal volume of matrix (α-cyano-4-hydroxycinnamic acid, 20 mg/ml) before being spotted onto 384-well plate to crystallize. Mass spectra were recorded with a 4800 matrix-assisted laser desorption ionization time-of-flight analyzer (Applied Biosystems) in both positive and negative modes.

To determine if c-di-AMP is secreted into growth medium, overnight cultures grown in LB medium were used to inoculate 50 ml of casein hydrolysate media (OD_{600} = 0.01). The cultures were subsequently grown until OD_{600} = 0.5. Cells were then pelleted resuspended in 50 ml of chemically defined sporulation medium. Cultures were grown with shaking at 37°C for 16 hours. The supernatants from both CH medium and SM medium were collected, and treated with 0.5 mM EDTA to inhibit potential hydrolysis of c-di-AMP. The supernatants were then heated at 100 degree for 10 min, followed by centrifugation and filtration (0.2 µm). The filtrate was concentrated by 10 fold using SpeedVac, and then loaded into the HPLC system for separation. Fractions corresponding to the retention time of c-di-AMP were collected and analyzed using MALDI-MS as described above.

4.3 RESULTS

Three protein constructs that contain various portions of YybT were cloned, expressed and purified as $(His)_6$-tagged proteins (**Fig. 4.1**). $YybT_{84-659}$ spans the whole cytoplasmic portion, whereas $YybT_{150-659}$ and $YybT_{84-303}$ contain the GGDEF-DHH/DHHA1 and PAS-GGDEF segment, respectively. The recombinant proteins were purified by Ni^{2+}-NTA affinity and size-exclusion chromatography to homogeneity (**Fig. 4.2**). The purified proteins were not bound by RNA, DNA or small nucleotide ligands as detected by HPLC and spectroscopic analysis of the denatured protein solution.

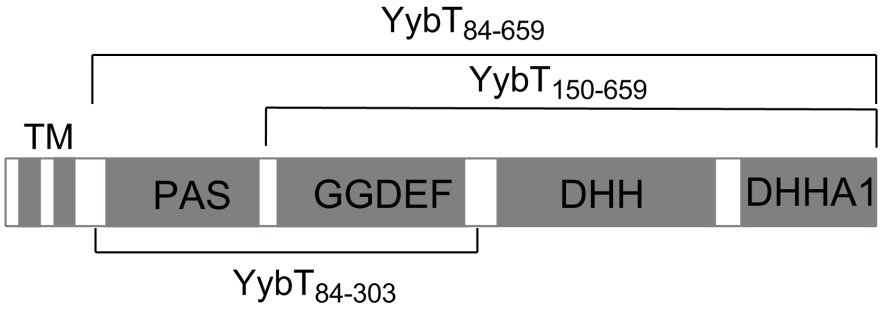


Figure 4.1 Domain architecture of YybT family proteins (COG3387) with the three protein constructs studied in this work indicated.

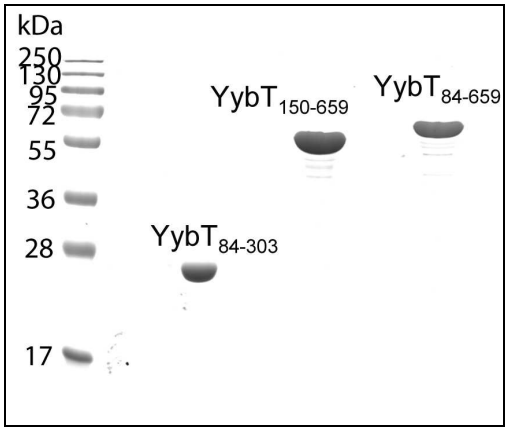


Figure 4.2 12% SDS-PAGE of purifed proteins encoding the three different YybT constructs.

4.3.1 The DHH/DHHA1 domain of YybT exhibits specific phosphodiesterase activity

The activity of YybT towards four *p*-nitrophenol substrates that are known to be easy substrates for phosphodiesterases and phosphatases was examined. In the presence of Mg^{2+} or Mn^{2+}, $YybT_{84-659}$ hydrolyzed bis-*p*-nitrophenol phosphate (bis-*p*NPP) and thymidine monophosphate *p*-nitrophenol ester (*p*N-TMP) but not *p*-nitrophenol phosphate and phosphenolpyruvate (PEP) (**Fig. 4.3B**), indicating that the enzyme is likely to be a phosphodiesterase rather than a phosphatase. This notion is further supported by the observation that YybT did not hydrolyze mono-phosphate ribonucleotides (xMPs) and deoxyribonucleotides (dXMPs) **(Appendix I)**. Moreover, YybT did not act on 3'-phosphoadenosine 5'-phosphate, suggesting that YybT differs from the DHH family phosphatase YtqI (103). YybT also did not exhibit catalytic activity against sodium tripolyphosphate and pyrophosphate, excluding it as a pyro- or polyphosphatase (39-40, 176).

After ruling out YybT as a phosphatase, the activity of YybT against a wide range of possible physiological substrates for phosphohydrolyase or phosphodiesterase was examined **(Appendix I)**. Di and tri-phosphate nucleotides (XDPs, XTPs), deoxyribonucleotides (dXDPs, dXTPs), and NADH, NADPH, NAD^+, $NADP^+$ and FAD were subjected to enzymatic assay. No phosphohydrolyase activity was observed for all the substrates tested with the exception of ATP. We found that ATP was slowly converted to ADP. This activity was attributed to the GGDEF domain rather than the DHH domain as we will discuss later. We also tested whether YybT functions as an oligoribonuclease or ribonuclease for degrading RNA or ssDNA (103, 186). Neither the ssDNA and RNA substrates, nor the sixteen oligo-ribonucleotides and oligo-deoxyribonucleotides of various lengths could be degraded by YybT in the presence of Mn^{2+} or Mg^{2+} **(Appendix I)**. Moreover, Yybt did not degrade diadenosine tetraphosphate (AppppA), triphosphate (ApppA) or pentaphosphate (AppppA), suggesting it is not a bis(5'-adenosyl) tri-, tetra or penta-phosphatase (119, 142). The small

nucleotide (p)ppGpp is a ubiquitous nucleotide messenger for signaling starvation and other stresses (123). Enzymatic assay showed that YybT did not degrade ppGpp, and thus, is not a ppGpp specific phosphohydrolyase or phosphatase.

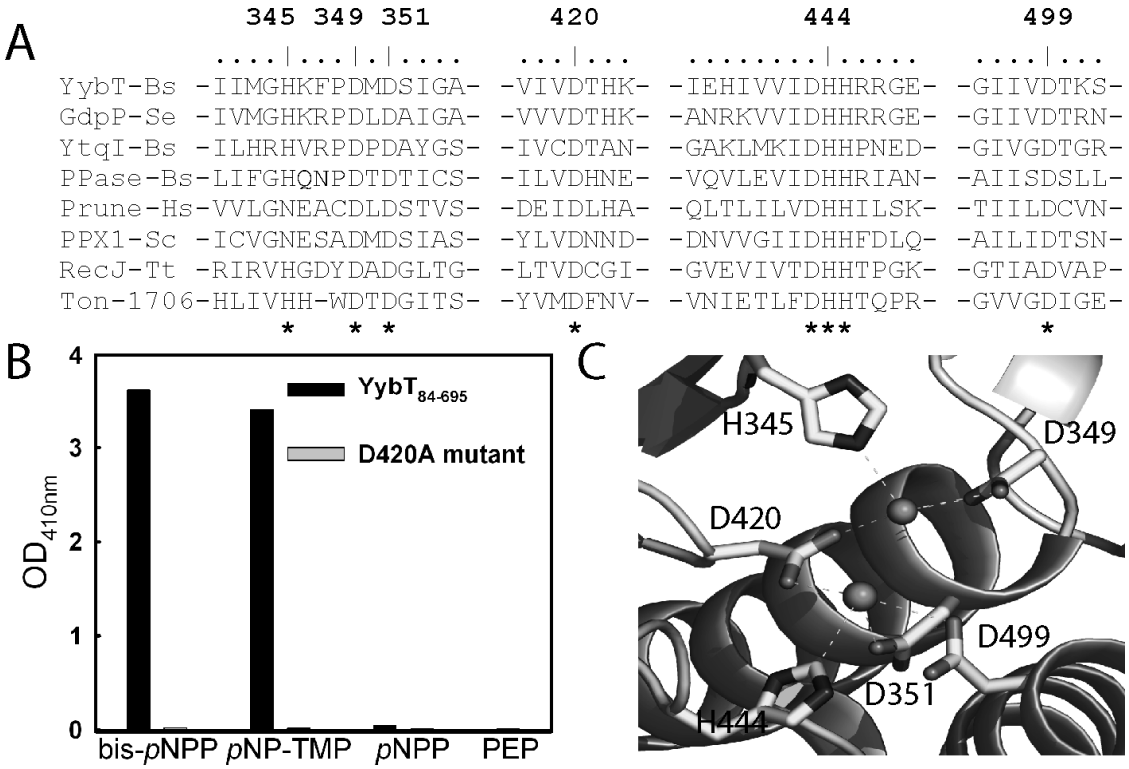


Figure 4.3 Specific phosphodiesterase activity of $YybT_{84-659}$. A. Partial alignment of DHH-DHHA1 domain sequences with the conserved residues for metal ion coordination highlighted. B. Enzymatic activity of YybT against non-physiological *p*-nitrophenol substrates. C. Structural model of DHH_{YybT} active site with the metal ions shown as balls and coordinating residues as sticks

Several naturally occurring cyclic nucleotide substrates were tested for phosphodiesterase activity. No activity was detected with 3', 5'-cAMP and cGMP even after prolonged incubation with Mg^{2+} or Mn^{2+}. In contrast, the incubation of YybT with c-di-AMP and c-di-GMP resulted in the disappearance of the cyclic dinucleotides and formation of new products. Based on HPLC analysis, YybT hydrolyzes c-di-AMP and c-di-GMP to produce exclusively the linear dinucleotides 5'-pApA and 5'-pGpG, without generating the 3'-dinucleotide (3'-pApA or 3'-pGpG) or monophosphate (AMP or GMP) (**Fig. 4.4 and 4.5**). Steady-state kinetic measurement yielded a k_{cat} of 0.55 ± 0.02 s^{-1} and Michaelis-Menten constant (K_m) of 1.3 ± 0.3 μM for c-di-AMP (**Fig. 4.4B**). In comparison, YybT exhibited a comparable k_{cat}

(0.23 ± 0.02 s^{-1}) but a much greater K_m (349 ± 55 μM) for c-di-GMP, indicating that the substrate binding pocket prefers c-di-AMP over c-di-GMP. The cyclic dinucleotide specific activity of YybT could also be detected with the construct ($YybT_{150-659}$) that contains the GGDEF-DHH/DHHA1 segment, although $YybT_{150-659}$ exhibited a ~10-fold decrease in k_{cat} (**Table 4.1**). According to sequence comparison and the structural model of the DHH_{YybT} subdomain, the residues D^{420} and D^{499} are predicted to be the metal ion-coordinating residues (**Fig. 4.3 A and C**). Enzymatic assays showed that the D420A mutant only exhibited residual activity against c-di-AMP, c-di-GMP, bis-*p*NPP and *p*N-TMP (**Fig. 4.2B**), whereas the double mutant D420A/D499A totally abolished the catalytic activity. These observations confirmed that the DHH domain is responsible for the specific phosphodiesterase activity.

Table 4.1 Kinetic parameters for the phosphodiesterase activity of the DHH_{YybT} domain.

Protein	Substrate	k_{cat} (s^{-1})	K_m (μM)	k_{cat}/K_m (s^{-1} μM^{-1})
$YybT_{84-659}$	c-di-AMP	0.55 ± 0.02	1.3 ± 0.24	0.42 ± 0.08
$YybT_{84-659}$	c-di-GMP	0.23 ± 0.02	349 ± 55	(6.6 ± 1.1) × 10^{-4}
$YybT_{150-659}$	c-di-AMP	0.074 ± 0.002	3.0 ± 0.22	0.024 ± 0.002
$YybT_{150-659}$	c-di-GMP	0.021 ± 0.002	199 ± 39	(1.1 ± 0.23) × 10^{-4}

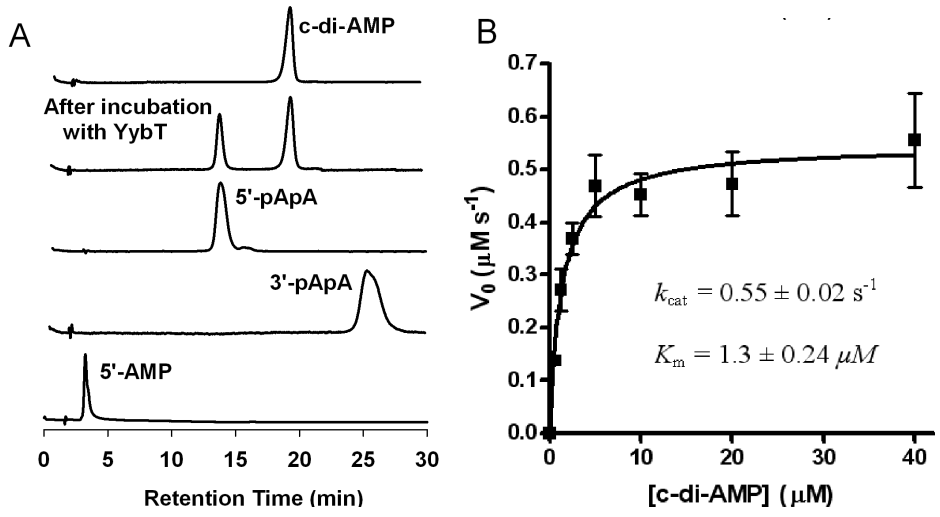


Figure 4.4 Degradation of c-di-AMP by YybT. A. HPLC analysis of the hydrolysis of c-di-AMP by $YybT_{84-659}$. Assay conditions are described in materials and methods. B. Steady-state kinetic analysis of the hydrolysis of c-di-AMP catalyzed by $YybT_{84-659}$.

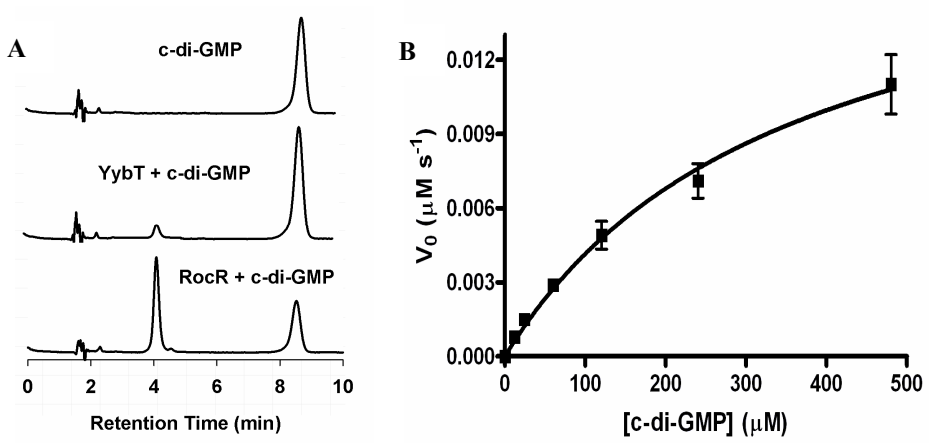


Figure 4.5 YybT catalyzes the hydrolysis of c-di-GMP. A. HPLC analysis of the hydrolysis of c-di-GMP by YybT and the EAL domain-containing RocR. B. Steady-state kinetic analysis of c-di-GMP hydrolysis (Reaction conditions: 100mM Tris, pH 8.3, 0.5mM MnCl2, 20mM KCl, 1 μM YybT, 10 – 430 μM substrate).

In addition, YybT also hydrolyzed 2',3'-cAMP and other 2',3'-cyclic nucleotides into nucleotide 3'-phosphate . However, the K_m was high (15 mM) and 2',3'-cyclic nucleotides are known to be unstable (100, 130). Furthermore, six PDEs belonging to three different domain families all showed 2',3'-cAMP hydrolysis activity, with the catalytic efficiency against 2',3'-cAMP correlated with that against the known non-specific PDE susbtrate bis-pNPP (129). It was therefore concluded that the hydrolysis of 2',3'-cAMP by YybT is non-specific and not physiologically relevant (130).

In another direction, I also ruled out the possibility that specific c-di-AMP hydrolysis is a shared property for DHH/DHHA1 domain containing proteins. Other than YybT, there are three other families of DHH/DHHA1 domain proteins, exemplified by the single-stranded DNA specific nuclease RecJ from thermus thermophilus (ttRecJ) (186), the CCA-adding tRNA nucleotidyltransferase from *Aquifex aeolicus* (named *Aa*_DHHA1 here) (173), and the stand alone DHH/DHHA1 domain protein YtqI (103) (**Fig. 4.6**). These proteins were prepared and their activities against c-di-AMP and 2',3'-cAMP were examined.

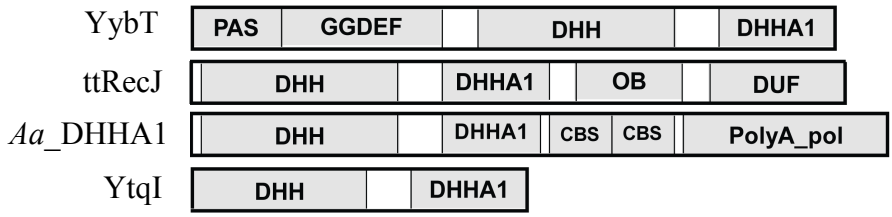


Figure 4.6 Dommain composition of YybT (cytoplasmic portion) and other DHHA1 subdomain containing proteins. DUF: domain of unknown function.

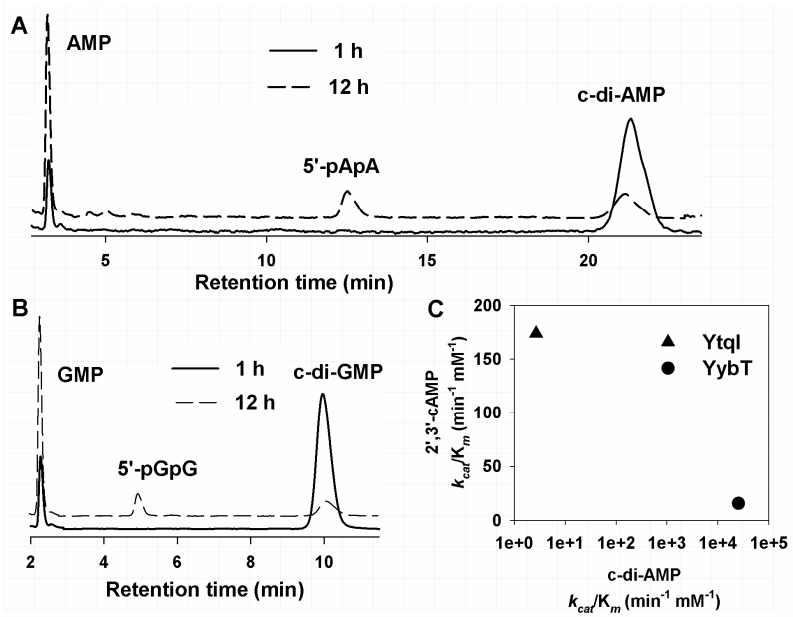


Figure 4.7. The hydrolysis of c-di-AMP (**A**) and c-di-GMP (**B**) by YtqI. **C.** The catalytic efficiency of YybT and YtqI against 2',3'-cAMP and c-di-AMP.

While ttRecJ and *Aa*_DHHA1 showed hydrolytic activity against the nonspecific PDE substrate 2',3'-cAMP, they are inactive against c-di-AMP. For YtqI, c-di-AMP hydrolysis was observed (**Fig. 4.7**), but with important distinctions when compared to YybT. First, the enzyme hydrolyzed c-di-AMP and c-di-GMP at the same speed (**Fig. 4.7A, B**), whereas YybT is much more active against c-di-AMP. Second, the prominent hydrolysis products were AMP and GMP, with 5'-pApA and 5'-pGpG as minor intermediates (**Fig. 4.7A, B**). The pattern is reminiscent of nonspecific hydroysis of cyclic-di-nucleotide by the snake venom phosphodiesterase (165). Last, the catalytic efficiency of c-di-AMP hydrolysis by YtqI is 1600-fold lower compared to YybT, despite its 2',3'-cAMP hydrolysis efficiency being 65

fold higher than YybT (**Fig. 4.7C**). Thus, specific hydrolysis of c-di-AMP into 5'-pApA is a unique property of YybT.

4.3.2 Metal and pH dependence of the phosphodiesterase activity

The phosphodiesterase activity of YybT is strictly dependent on divalent metal ions such as Mn^{2+}, Mg^{2+}, Ni^{2+} and Co^{2+} ions (**Fig. 4.8A**). Meanwhile, Zn^{2+} ion inhibits the catalytic activity regardless of the presence of Mn^{2+}or Mg^{2+}; whereas Ca^{2+} only inhibits the catalysis in the presence of Mg^{2+}, but not Mn^{2+} (**Fig. 4.8A**). Mn^{2+} and Mg^{2+} are most likely the physiological metal ion for YybT considering that DHH family proteins tend to use Mn^{2+} or Mg^{2+} for catalysis (186). To identify the physiological metal ion for YybT, the Mg^{2+} and Mn^{2+} ion dependent catalytic activity was examined (**Fig. 4.8B**). Mn^{2+} ion exhibits a bell shaped profile with the optimal contraction in the range of 10 μM - 10 mM, with the activity strongly inhibited at high Mn^{2+} concentration (≥ 10 mM). By contrast, the catalytic activity remained low for Mg^{2+} at 1 mM [Mg^{2+}] and did not increase significantly until [Mg^{2+}] reached 10 mM. Since the cellular Mn^{2+} and Mg^{2+} concentrations are around 10 μM and 1 mM respectively in *B. subtilis* (178), the results support Mn^{2+} ion as the physiological metal ion for the DHH/DHHA1 domain of YybT. The pH dependence study of the catalytic activity showed a strong preference for alkaline conditions with a pH optimum between 8.5 and 9.0 (**Fig. 4.8C**), which seems to be consistent with a two-metal-ion assisted catalytic mechanism (106).

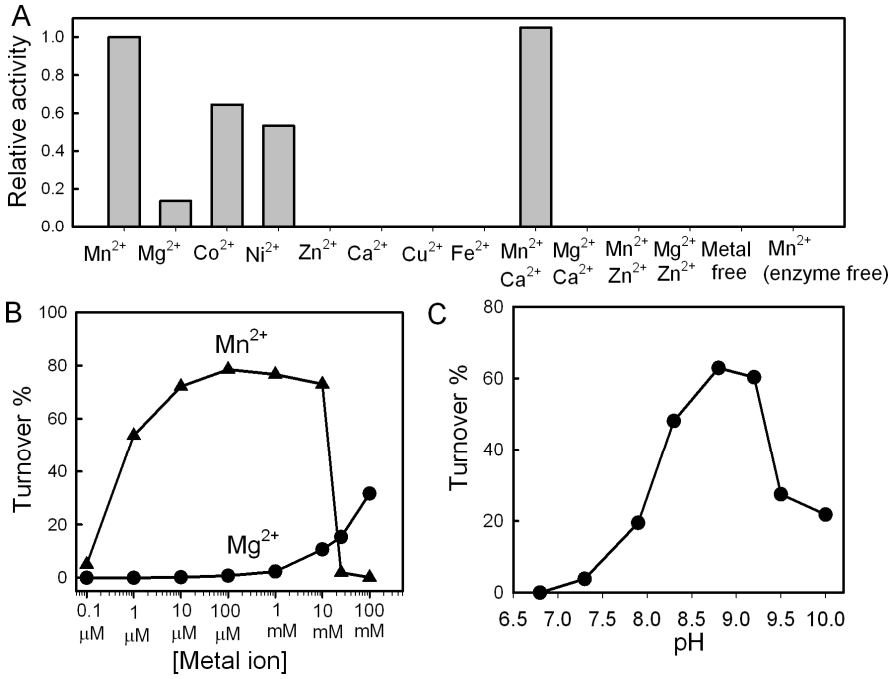


Figure 4.8 Metal and pH dependence of c-di-AMP degradation by $YybT_{84-659}$. **A.** Metal dependence of YybT-catalyzed c-di-AMP hydrolysis. The enzymatic assay conditions are described in Experimental Procedures. **B.** Mn^{2+} and Mg^{2+} concentration dependence of c-di-AMP hydrolysis. **C.** pH dependence of DHH_{YybT} domain activity**.**

4.3.3 The GGDEF domain of YybT binds and hydrolyzes ATP

Initial enzymatic assays suggested that the GGDEF domain of YybT is neither a functional diguanylate cyclase domain for converting GTP into c-di-GMP, nor a GTPase for hydrolyzing GTP. During substrate screening for the DHH/DHHA1 domain of YybT, we noticed that the concentration of ATP decreases with time in the presence of YybT and Mg^{2+} (or Mn^{2+}). HPLC analysis revealed that the depletion of ATP was accompanied by the formation of ADP, suggesting that YybT catalyzes the hydrolysis of ATP. Surprisingly, the D420A/D499A double mutation in the DHH domain did not affect ATP hydrolysis, hinting that DHH/DHHA domain is not responsible for the ATPase activity. Meanwhile, sequence alignment showed that YybT contains a highly modified GGDEF domain that lacks the signature GGDEF motif and the guanosine-binding residues critical for the DGC activity (22). To test whether the degenerate GGDEF domain is responsible for the ATPase activity, Protein $YybT_{84-303}$ that contains the PAS-GGDEF di-domain was prepared **(Fig. 4.1)**. While

the purified $YybT_{84-303}$ protein did not show any activity against GTP and other tri-phosphate nucleotides, it hydrolyzed ATP to yield exclusively ADP (**Fig. 4.9A**). Enzymatic assay with the EnzCheck phosphate assay kit confirmed that the inorganic phosphate was released into the solution, suggesting that the GGDEF is not a kinase that transfers the phosphate onto a substrate (**Fig. 4.10**). Steady-state kinetic measurement yielded a k_{cat} of 0.59 ± 0.03 min^{-1} and K_m of 0.90 ± 0.12 mM for the ATPase in the presence of Mg^{2+} ion (**Fig. 4.9B**). The metal ion Mg^{2+} or Mn^{2+} is required for the ATPase activity with Mg^{2+} as the preferred metal ion, in contrast to the metal preference for the DHH/DHHA1 domain (**Fig. 4.9C**). To further provide evidence for the preferred binding of ATP by the divergent GGDEF domain, we measured the dissociation constant (K_D) for the non-hydrolyzing ATP analogue adenylyl imidodiphosphate (AMPPNP) by competitive inhibition experiments (**Fig. 4.9B**). The measurement revealed a K_D of 44.0 ± 14.6 μM for AMPPNP. The presence of the AMPPNP did not exert any significant effect on the DHH/DHHA1 domain activity, suggesting that the GGDEF domain is not an ATP-binding regulatory domain. Finally, while the signature GGDEF motif and the residues for guanosine group binding are absent, the residues that coordinate the Mg^{2+} ions in orthodox GGDEF domains (33, 114), with one in the signature GGD/EEF motif and another at the end of β-strand 1, seem to be conserved in $GGDEF_{YybT}$ (E^{225} and D^{183}) (**Fig. 4.9D and Fig. 4.11A, B**). The enzymatic assay showed that the mutation of the conserved residues D^{183}, E^{225} and the conserved γ-phosphate binding residue R^{291} reduced the catalytic activity to below 5% (**Fig. 4.9C**). Although the structural motifs for the binding of the adenosine group remain to be defined, the mutagenesis results suggested that the metal ion and phosphate binding sites in the GGDEF domain may be conserved despite the divergence of function.

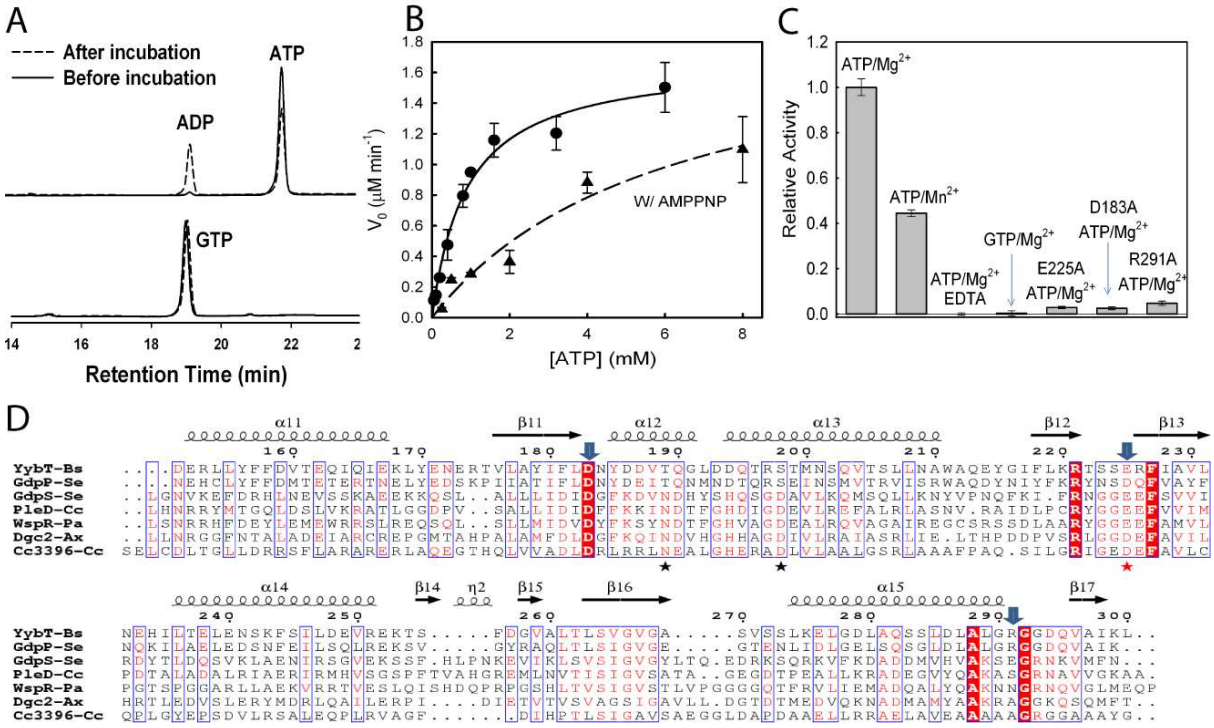


Figure 4.9 ATPase activity of the $GGDEF_{YybT}$ domain. **A.** HPLC analysis of ATP hydrolysis catalyzed by $YybT_{84-303}$. **B.** Steady-state kinetic measurement of ATPase domain activity in the presence and absence of non-hydrolysable ATP analog. **C.** Relative ATPase activity of $YybT_{84-303}$ and its mutants in the presence of Mg^{2+} and Mn^{2+}. **D.** Sequence comparison of the $GGDEF_{YybT}$ domain with some characterized GGDEF domains. The three residues mutated in this study are indicated by the arrows, whereas the residues for guanosine binding in orthodox GGDEF domains are indicated by asterisks. GdpP (*S. aureus*) is a homolog of YybT. The GGDEF domain of GdpS (*S. aureus*) exhibits residual GTPase activity (64). The GGDEF domain of PleD (*C. crescentus*), WspR (*P. aeruginosa*) and Dgc2 (*A. xylinum*) function as diguanylate cycles (22, 125, 132). The GGDEF domain of Cc3396 binds GTP as a regulatory domain (28).

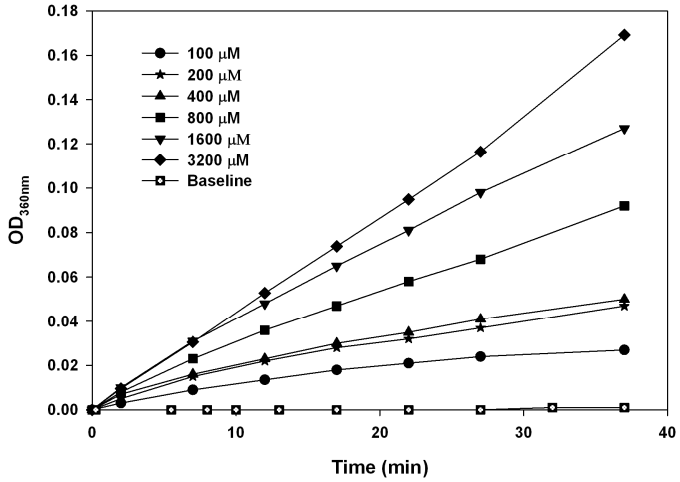


Figure 4.10 Hydrolysis of ATP by $YybT_{84-303}$ measured using the Enz Check Phosphate assay at various ATP concentrations. Assay condition: 50 mM Tris (pH7.5), 1 mM $MgCl_2$, 100-3200 μM ATP, 1 U/ml purine nucleoside phosphrylase, 200 μM MESG (substrate 2-amino-6-mercapto-7-methyl-purine). Reaction was initiated with addition of enzyme at final

concentration of 2.88 µM. A parallel control assay with the protein storage buffer instead of enzyme was used for baseline correction.

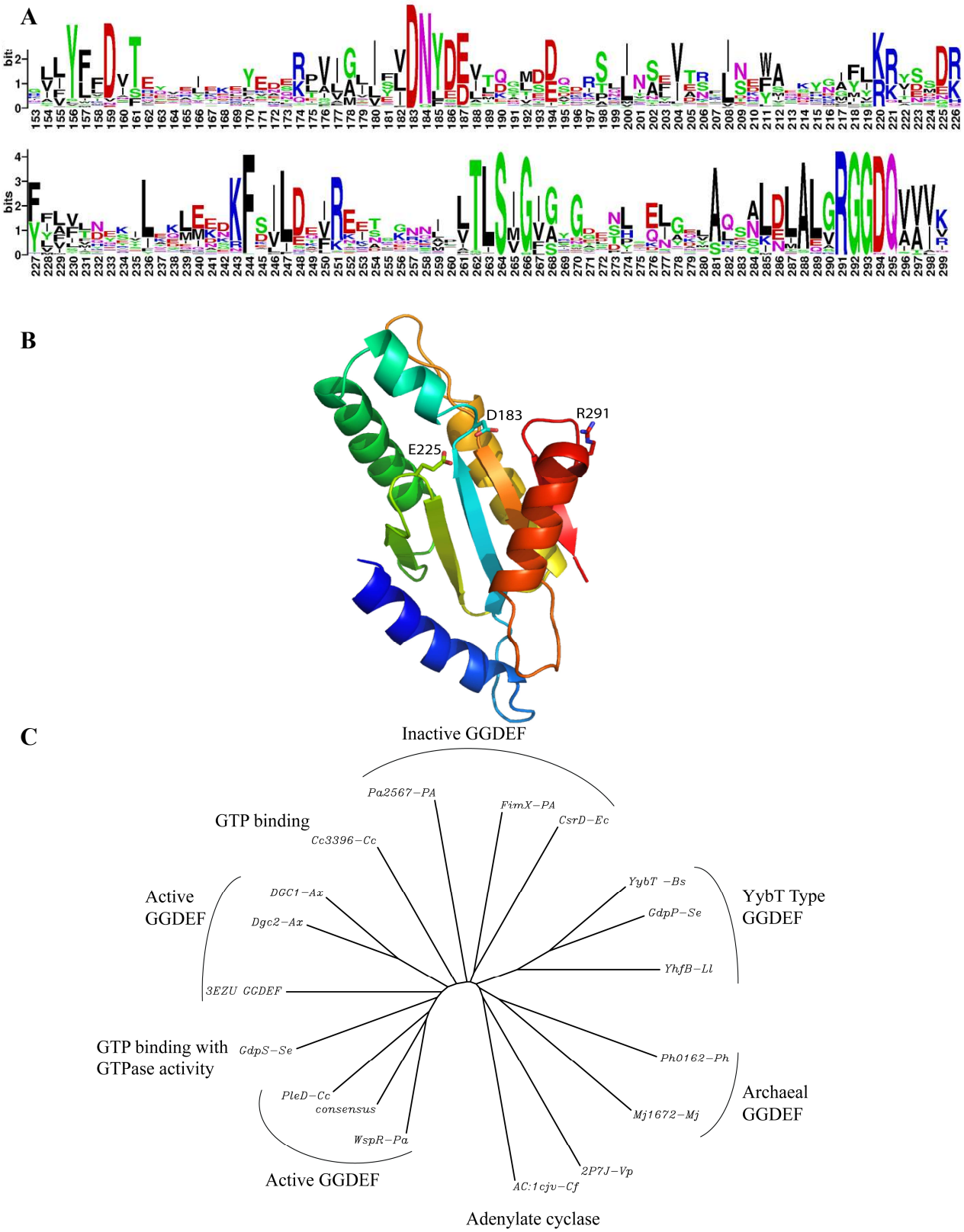


Figure 4.11 Summary of the $GGDEF_{YybT}$ domain. A. HMM logo generated from the alignment of 185 YybT homologs for the $GGDEF_{YybT}$. B. Structural model of the $GGDEF_{YybT}$ domain generated

using the PleD structure (2V0N) as a template. The putative residues for metal and phosphate binding are shown as sticks. C. Unrooted Neighbor-joining tree generated from sequence alignment of various class of GGDEF domains. Ax, Acetobacter xylinus; Bs, Bacillus subtilis; Cf, cannis familiaris; Cc, caulobacter crecentus; Ec, Escherichia, coli; Ll, Lactobacillus lactis; Mj, Methanococcus jannachir; Pa, Pseudomonas aeruginosa; Ph, pyrococcus horikoshii; Se, Straphylococcus epidermidis; Vp, Vibrio parahaemolyticus.

4.3.4 Inhibition of the phosphodiesterase activity by ppGpp

Phenotype screening of the mutants of *Lactococcus lactis* revealed that the YybT family protein llmg1816 is involved in acid and starvation stress resistance, with the Δllmg1816 mutant strain exhibiting higher surviving rate under low pH and starvation conditions (126). The same group also demonstrated that treatment of *L. lactis* with stringent response inducers, which resulted in high level of the stress alarmon (p)ppGpp, conferred improved acid resistance. It is well known that (p)ppGpp level is significantly elevated during stringent responses, resulting in the binding of (p)ppGpp to its protein targets such as RNA polymerase and inosine monophosphate dehydrogenase (IMPDH) (68, 126, 182). To test whether the YybT protein is a previously unknown target for (p)ppGpp, enzymatic assays of the phosphodiesterase and ATPase activities were performed in the presence of ppGpp.

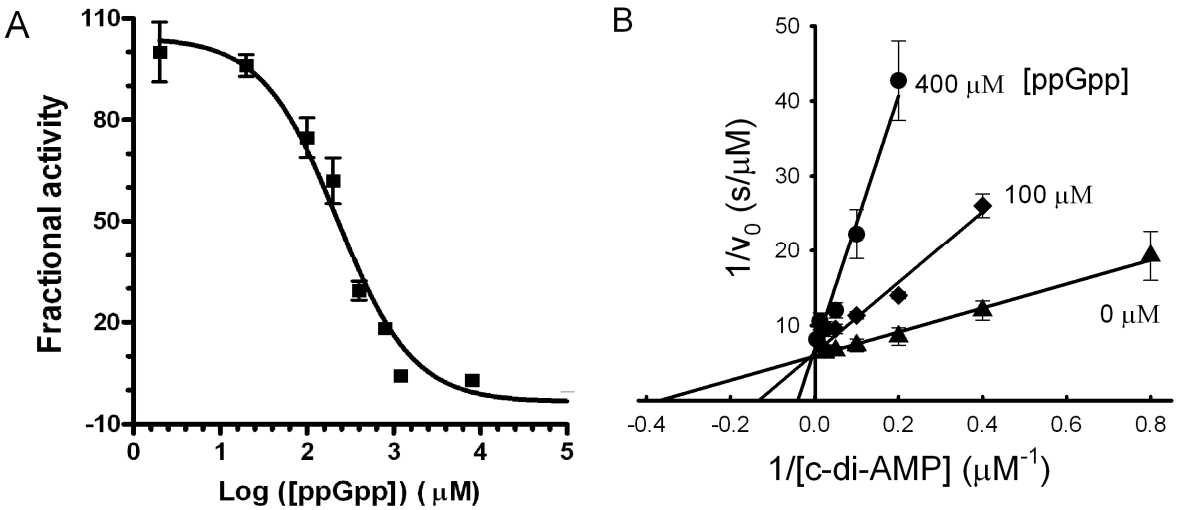


Figure 4.12 Inhibition of the phosphodiesterase activity by ppGpp. A. ppGpp concentration dependence of the c-di-AMP hydrolyzing activity. The IC_{50} value is determined from the midpoint of the plot. B. Double-reciprocal plots of c-di-AMP hydrolyzing activity at various ppGpp concentrations. The experimental conditions are described in Material and Methods.

Remarkably, while ppGpp did not inhibit or stimulate the ATPase activity, it significantly suppressed the c-di-AMP and c-di-GMP hydrolyzing activity of the DHH/DHHA1 domain. An apparent IC_{50} value of 234 μM was derived from the activity vs. [ppGpp] plot at 20 μM [c-di-AMP] (**Fig. 4.12A**). Initial velocity measurement at various ppGpp concentrations indicated that the inhibition is competitive in nature, with an inhibition constant (K_i) of 35.9 ± 7.2 μM (**Fig. 4.12B**). Considering that bacterial cells can rapidly accumulate ppGpp up to mM level under starvation conditions (123, 155), the inhibition of the DHH domain by ppGpp is likely to be physiologically relevant. No such inhibition effect was observed for other cyclic or linear nucleotides, or on the c-di-GMP hydrolyzing activity of the EAL domain proteins such as RocR (128, 132), indicating that the inhibition of the DHH/DHHA1 domain by ppGpp is rather specific.

4.3.5 Effect of ΔyybT mutation on acid and DNA damage resistance

Lactobacillus lactis contains an yybT-like gene (llmg1816). The Δllmg1816 mutant was found to be more resistant under acid stress and also to exhibit enhanced long term (2-day) survival (126). I found that the Δ*yybT* mutant strain of *B. Subtilis* is also more resistant than the wildtype against acid stress (**Fig 4.13A**). Because the c-di-AMP synthesizing DisA has been proposed to signal DNA damage (183), we further tested whether YybT plays a role in mediating DNA damage resistance. No significant survival advantage or disadvantage was observed when the wild-type *B. subtilis* 168 and ΔYybT mutant were treated with the topoisomerase inhibitor nalidixic acid or γ-ionizing irradiation, suggesting that YybT does not play a significant role in DNA repair during the normal development process. However, when we compared the sporulating cells, a significant difference was observed between the wild type and mutant strains. As shown in **Fig. 4.13B**, when both the wild type and *yybT* mutant cells were induced to enter sporulation without exposure to nalidixic acid, the spore

formation efficiency was comparable, suggesting that the gene knock-out does not significantly affect sporulation. However, in the presence of nalidixic acid, the mutant strain consistently exhibited more colonies, which were derived from the spores that survived the heat treatment. The discrepancy in spore formation efficiency widens between the wildtype and the *yybT* mutant with increasing nalidixic acid concentration. This is in contrast to the observation for the ΔdisA mutant, which formed less spores than the widl type strain (9).

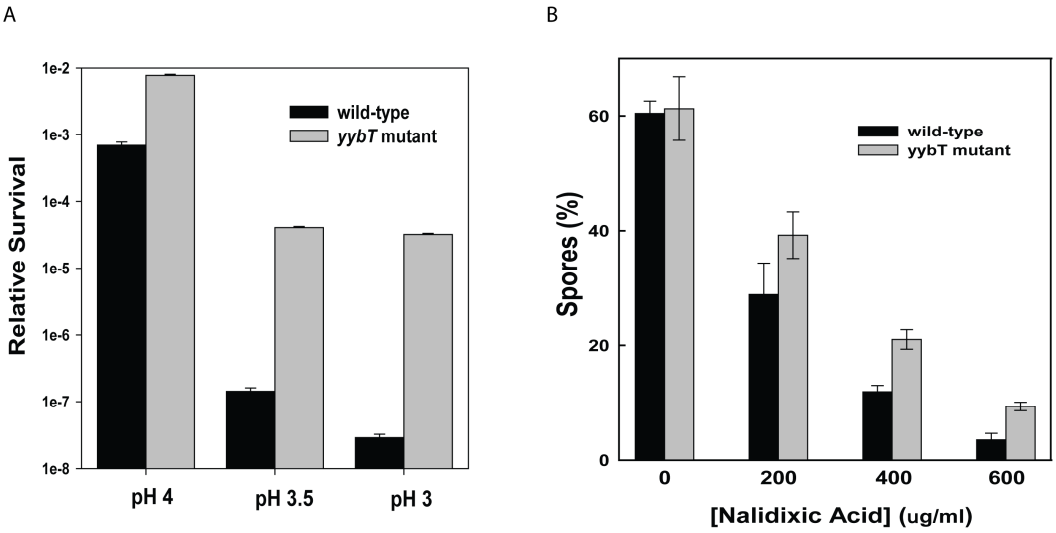


Figure 4.13 Acid and DNA damage resistance. A. relative survival of wild type and ΔyybT strain under acid stress conditions. B. Spore formation efficiency for the wild type and ΔyybT mutant in the presence of DNA damage-inducing agent narlidixic acid. The spore (%) was obtained as described in Material and Methods.

4.3.6 Isolation of c-di-AMP from *B. Subtilis*

To test whether YybT affects the global c-di-AMP level, I set out to measure the cellular concentration of c-di-AMP for wild type and ΔYybT strains during vegetative growth and sporulation. Intriguingly, although I could readily detect the c-di-AMP standard in the spiked sample and c-di-GMP in *E. coli* and *P. aeruginosa* using the mass spectrometry method described by Simm et al (152), significant c-di-AMP peak above the background in both the wild type and mutant strains was not observable after repeated attempts. The results led us to suspect that the global concentration of c-di-AMP in *B. subtilis* is rather low.

4.4 DISCUSSION

4.4.1 The c-di-AMP specific PDE activity

DHH family proteins function as phosphatases or phosphodiesterases hydrolyzing a wide variety of substrates ranging from pyrophosphate to ssDNA. The enzymatic activity assays suggested that YybT possesses different substrate specificity from other known DHH family proteins. The biochemical data suggested that YybT is a unique cyclic dinucleotide specific phosphodiesterase for hydrolyzing the cyclic dinucleotides c-di-GMP and c-di-AMP, but not 3',5'-cAMP or cGMP. The substrate specificity is consistent with the low sequence homology shared by the substrate-binding DHHA1 subdomains between YybT and other DHH family proteins. C-di-GMP is a ubiquitous second messenger in bacteria and is known to be degraded by the EAL domain and HD-GYP domain phosphodiesterases (143, 146). Although YybT is able to degrade c-di-GMP under the experimental conditions, c-di-GMP is unlikely to be the physiological substrate for YybT for two major reasons. First, the cellular c-di-GMP concentration is in the low- or sub-μM range according to the estimated cellular c-di-GMP concentration (60, 104, 150). In addition, the K_m values of the c-di-GMP degrading EAL domain proteins are also known to be in the low- or sub-μM range (28, 128, 132). Hence, the elevated K_m (349 ± 54.6 μM) for c-di-GMP indicates that the DHH domain is incapable of hydrolyzing c-di-GMP efficiently under physiological conditions. Second, considering that the YybT family proteins (COG3387) can be found in the bacteria (e.g. *Streptococcus* and *Staphylococcus*) that do not seem to harbour the c-di-GMP signalling network as judged by the absence of catalytically active EAL/HD-GYP and GGDEF domain proteins (64), it seems that the YybT proteins are unlikely to be involved in c-di-GMP signalling. Hence, the physiologically irrelevant K_m and presence of YybT in the bacteria lacking the c-di-GMP network strongly disfavor c-di-GMP as the cellular substrate for YybT.

On the other hand, the biochemical and genetic data seem to support c-di-AMP as the cellular substrate for YybT. First, the small K_m of 1.3 ± 0.3 μM toward c-di-AMP indicates that YybT can degrade c-di-AMP efficiently even if each *B. subtilis* cell only contains 200-600 molecules of c-di-AMP, which correspond to a cellular concentration of low- or sub-μM range. Second, YybT hydrolyzes c-di-AMP to generate exclusively 5'-pApA (not 3'-pApA or AMP), suggesting preferred binding of cyclic dinucleotide over linear dinucleotide. The crystal structures of several DHH domains revealed that the C-terminal subdomain (DHHA1 or DHHA2) is largely responsible for the binding of the substrate (40, 166, 186). Sequence alignment of YybT and DHH family proteins revealed that while the DHH subdomain is fairly conserved with the putative metal-ion binding residues (**Fig. 4.3A**), the DHHA1 subdomain does not share significant sequence homology with any characterized DHH family proteins. Sequence comparison also revealed that the Arg residues critical for the binding of polyphosphate, RNA or ssDNA in other DHHA1 domains (e.g. RecJ and YtqI) are not conserved in YybT. In contrast, only one conserved Arg (R^{606}) was found in YybT homologs (**Fig. 4.14**) and it is likely to interact with the phosphate group of c-di-AMP according to our structural model. Accordingly, we showed that all other three families of DHH/DHA1 domain containing proteins either do not hydrolyze c-di-AMP, or hydrolyze c-di-AMP in a manner analogous to non-specific PDE **(Fig 4.7)**. Hence, the uniqueness of the DHHA1 sub-domain of YybT seems to be in agreement with its unprecedented substrate specificity. Third, the only putative phenotype mediated by c-di-AMP is the signalling of DNA damage. The mutation of DisA significantly retards sporulation in the presence of a DNA damaging agent. In sharp contrast, the Δ*yybT* mutant exhibited higher spore formation efficiency than the wildtype (**Fig. 4.11**). Fourth, DAC domains can be found in 275 bacterial or archaeal species with many of them belonging to the phylums of firmicutes and ε-proteobacteria; whereas YybT family proteins can be found in 123 bacterial species that

mainly consist of firmicutes **(Appendix II)**. Importantly, for all the 123 bacterial species that contain the YybT family protein, at least one DAC domain protein can be identified. Based on these considerations, we propose that YybT is involved in c-di-AMP signalling by hydrolyzing the cyclic dinucleotide, resembling the role of the EAL domain in c-di-GMP signalling. More characterization of YybT homologs from other species is necessary to generalize this claim. Meanwhile, the absence of YybT protein in some DAC domain protein-containing bacteria seems to hint the existence of other families of c-di-AMP specific phosphodiesterases.

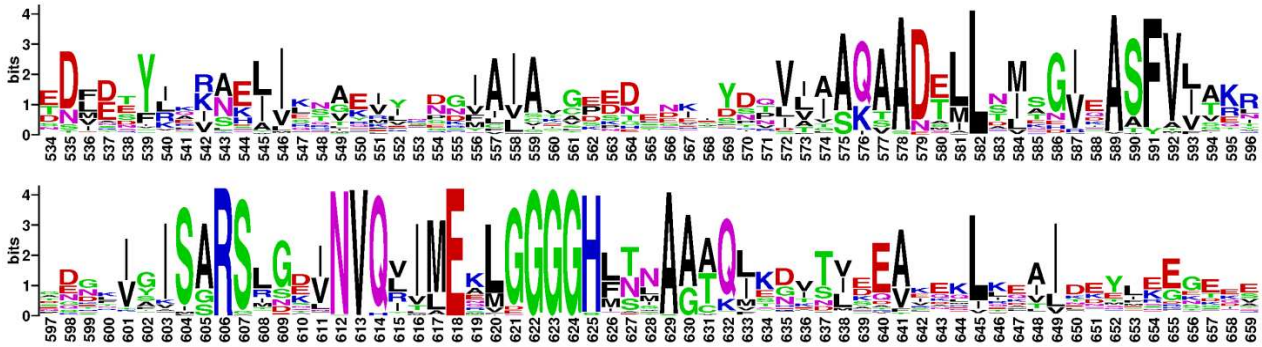


Figure 4.14 HMM logo generated from the alignment of 185 YybT homologs for the DHHA1YybT subdomain.

4.4.2 The ATPase activity of the GGDEF domain

It has been noticed that YybT family proteins contain a highly modified GGDEF domain with unknown function (42, 64). GGDEF domains are homologs of adenylate cyclase (AC) domains and mainly known to be DGC domains for synthesizing c-di-GMP from GTP (144, 153). A large number of catalytically inactive GGDEF domains that lack the signature GG(D/E)EF motif have been identified including two divergent GGDEF domains that can still bind GTP despite the lack of DGC activity. Jenal and co-workers reported a regulatory GGDEF domain that binds GTP and stimulates the activity of the adjacent EAL domain (28); while O'Gara and co-workers recently reported a GGDEF domain with residual GTPase activity (64). In contrast to these GGDEF domains, our study revealed that the GGDEF

domain of YybT specifically binds and hydrolyzes ATP but not GTP. Sequence comparison suggested that the metal ion binding residues in $GGDEF_{YybT}$ are probably conserved, but the residues for the guanosine group binding are absent (**Fig. 4.9D**). Based on the mutagenesis results, we speculate that ATP is probably bound at approximately the same site as GTP in the typical GGDEF domains, though the structural motifs that confer the binding specificity remain to be identified. Interestingly, phylogenetic analysis revealed that the GGDEF domain of YybT resembles adenylate cyclase domains more than the typical GGDEF domains (**Fig. 4.11C**), providing further support for the notion that GGDEF and adenylate cyclase domains were evolved from the same ancestor (73). Considering that the YybT family proteins are only found in an early branch of the evolutionary tree, the ATPase activity exhibited by the $GGDEF_{YybT}$ domain seems to suggest an early divergence of function for the GGDEF domains during evolution. Lastly, the intrinsic ATPase activity of the GGDEF domain is relatively low, reminiscent of the basal activity of the AAA^{+}- ATPases prior to the stimulation by small-molecule or protein partners (175). It is possible that the PAS domain may function as a sensor domain for perceiving and transducing a signal to regulate the activity of the GGDEF and/or DHH/DHHA1 domain.

4.4.3 Regulation of the PDE activity by ppGpp

The nucleotide ppGpp has been known as a bacterial and plant messenger that binds RNA polymerase and other protein targets for regulating gene expression and other cellular activities during stringent responses (24). Rallu *et al* showed that high (p)ppGpp level conferred better acid resistance in *L. Lactis* (126). We and Rallu *et al* also showed that YybT loss-of-function mutants are more resistant to acid stress in *B. subtilis* and *L. Lactis*. The observed inhibition of the DHH/DHHA1 domain by ppGpp suggests a direct link between the ppGpp and c-di-AMP signalling networks, and implies that high (p)ppGpp level induced acid

resistance may be due to, at least in part, regulation of the activity of the DHH/DHHA1 domain. The measured IC_{50} and K_i values indicate that the hydrolysis of c-di-AMP will be fully suppressed during stringent response with the ppGpp level raised above 1 mM. Interestingly, the interference of the ppGpp and c-di-GMP signalling networks has been also reported recently (14, 118). The inhibition of c-di-AMP hydrolysis by ppGpp seems to represent another example of cross-talking between two nucleotide signalling networks.

In addition to the novel enzymatic activities exhibited by YybT, the opposite effects of the Δ*disA* and Δ*yybT* mutations on sporulation in the presence of a DNA-damaging agent point towards contrasting roles played by the proteins in DNA-damage resistance. The higher sporulation efficiency for the Δ*yybT* mutant indicates that the mutant strain is more resistant against DNA damage, whereas the acid resistance experiment showed that the Δ*yybT* mutant strain is more tolerant to acid stress, consistent with the observation for the *L. Lactis* mutant (126). Together, the acid and DNA damage resistance exhibited by the mutant suggest that YybT functions as a signalling protein for coordinating stress sensing and cellular responses. Because acid stress can also cause chromosomal DNA damage in bacteria (29, 72), the observed acid and DNA damage resistance may be highly related on the molecular level. Lastly, considering that the DHH/DHHA1 domain of YybT is an efficient c-di-AMP phosphodiesterase under *in vitro* conditions and that the c-di-AMP synthesizing DisA is implicated in signalling DNA damage, it is tempting to conclude that the effect of the ΔyybT mutation is exerted through c-di-AMP. However, despite the *in-vitro* activity towards c-di-AMP, it remains to be seen whether c-di-AMP is the physiological substrate of YybT, it also remains to be fully established in the future whether c-di-AMP is truly a signalling messenger that regulates the phenotypes associated with stress responses. The failure to detect c-di-AMP in both the vegetative and sporulating *B. subtilis* cells seems to hint that even if the cyclic-nucleotide is at play, it may only function at a local level.

4.5 SUMMARY

The cyclic dinucleotide cyclic-di-AMP (c-di-AMP) synthesized by the diadenylate cyclase (DAC) domain was recently discovered as a messenger molecule for signaling DNA breaks in *Bacillus subtilis*. By searching bacterial genomes, we identified a family of DHH/DHHA1 domain proteins (COG3387) that co-occur with a subset of the DAC domain proteins. In this chapter, data presented show that the *B. subtilis* protein YybT, a member of the COG3387 family proteins, exhibits specific phosphodiesterase activity towards cyclic dinucleotides. The DHH/DHHA1 domain hydrolyzes c-di-AMP and c-di-GMP to generate the linear dinucleotides 5'-pApA and 5'-pGpG. The data suggests c-di-AMP to be the physiological substrate for YybT given the physiologically relevant Michaelis-Menten constant (K_m) and the presence of YybT family proteins in the bacteria lacking c-di-GMP signaling network. The stringent alarmone ppGpp was also found to be a strong competitive inhibitor of the DHH/DHHA1 domain, suggesting a direct link between ppGpp and c-di-AMP signaling. Moreover, the atypical GGDEF domain of YybT exhibits unexpected ATPase activity, distinct from the common diguanylate cyclase activity for GGDEF domains. Finally, evidence was provided to suggest that YybT participates in DNA quality control during acid stress and sporulation through the comparison of the wild type and Δ*yybT* mutant strains. Together, the results revealed unexpected signal sensing mechanism, unprecedented substrate specificity and enzymatic activity for the two catalytic domains of YybT and suggested a role for YybT in c-di-AMP signaling.

Chapter Five: The PAS Domain of the YybT family proteins binds heme and regulates the phosphodiesterase domain activity

5.1 INTRODUCTION

In Chapter 4, i have characterized the enzymatic activities of *Bacillus subtilis* YybT protein, a member of the COG 3387 family of transmembrane proteins that are widely distributed among firmicutes such as pathogenic *Staphylococcus aureus* and *Listeria monocytogene* (42). Although the biological function of this family of proteins is still unknown, gene knockout experiments have revealed that the COG3387 family proteins are involved in the regulation of various phenotypes in several organisms. Disruption of the llmg1816 (*Ll_yybT*) gene through transposon insertion renders *Lactococcus lactis* more acid tolerant (126). Knock-out of the *gcp* (*Sm_yybT*) gene results in abnormal biofilm formation in *Streptococcus mutans* (188). In *Staphylococcus aureus*, disruption of the *SA0014* (*Sa_yybT*) gene abolishes the secretion of the virulence factor hemolysin (18), suggesting that protein *Sa*YybT may play a critical role in infection and *in-vivo* survival of this pathogen. A separate study also identified *SA0014* as an *in-vivo* survival gene, whose mutation seems to attenuate virulence during murine infection (10, 121). Our study of *yybT* from *B. subtilis* also showed that the knockout mutant is more resistant towards acid and DNA damage caused by the bacteriastatic antibiotic nalidixic acid (131). Together, these results indicate that YybT is likely a stress signaling protein and a plausible target for designing antibiotics.

Our study on YybT revealed that the highly degenerate GGDEF domain possesses residual ATPase activity, rather than the diguanylate cyclase (DGC) activity associated with the typical GGDEF domain (Chapter Four). The enzymatic assay also showed that the DHH/DHHA1 domain possesses phosphodiesterase activity towards c-di-AMP and c-di-GMP, with a significantly smaller Michaels-Menten constant (K_m) towards c-di-AMP (131).

YybT family proteins also contain a putative PAS domain immediately following the trans-membrane helix. Almost nothing is known about the function of the PAS domain given the low sequence identity shared with any characterized PAS domains. The PAS domain, named after three eukaryotic proteins initially discovered containing this fold: Period, Aryl hydrocarbon receptor nuclear translocator, and Single-minded (63), is ubiquitously present in all three kingdoms of life as sensor or protein-protein interaction domains (58, 107, 170). The core PAS fold consists of a central five-strand anti-parallel β-sheet and three or four flanking α-helices. The sensory PAS domains generally contain a hydrophobic pocket to accommodate the signaling ligand (147) or cofactors such as heme (49), FAD (125), FMN (108) and *para*-hydroxcinnamate (62). In the most studied heme or flavin-bound PAS domains, it is through the direct interaction between the cofactor and signal that PAS domains sense changes in oxygen concentration, light intensity, voltage and redox potential (49, 107, 170).

Here I present the unexpected discovery that the PAS domain of YybT is fully capable of binding a heme cofactor despite the absence of an invariant histidine as the proximal ligand in the PAS domain. Spectroscopic characterization of the heme-reconstituted YybT and a thermophilic homolog (*Gt*YybT) in the oxidized, reduced and the nitric oxide (NO), carbon monoxide (CO) and cyanide (CN^-) ligated forms reveals a heme-binding environment distinct from other heme-binding PAS domains. Enzymatic activity measurement showed that the enzymatic activity of the DHH/DHHA1 and GGDEF domains of YybT can be modulated by the binding of the heme and the coordination of the small ligands to the heme iron. Kinetic measurements suggest that ligand binding stimulated PDE activity by increasing k_{cat} and decreasing K_m; The stimulatory effect of ligand binding is specific for cyclic-di-nucleotide substrate, as no activation was found for the hydrolysis of the non-specific PDE substrate 2',3'-cAMP. Together with the phenotypes associated with the YybT family proteins, the

results presented here suggest that the PAS domain may represent a unique subfamily of heme-binding proteins without using a proximal histidine or cysteine ligand for heme coordination.

5.2 MATERIALS AND METHODS

5.2.1 Materials

MAHMA NONOate, DEA-NONOate, sodium nitroprusside (SNP), the CO donor Tricarbonyldichlororuthenium (II) dimer (CORM2), potassium cyanide (KCN) and potassium ferricyanide [$Fe(CN)_6^{3-}$], hemin chloride, N_ω-Nitro-L-arginine methyl ester hydrochloride (L-NAME), and thrombin were from Sigma; spermine NONOate was from Caymen chemical. Hydrogen peroxide (30%) is from BDH. Gas cylinders for CO (99.5%), N_2 (99.9%) and O_2 (99.99%) tanks are from Air Products. Other materials are obtained the same as described previously (131).

5.2.2 Protein Expression and Purification

The expression and purification of the various *Bs*YybT fragments had been described previously (131). The gene encoding the YybT homolog from *Geobacillus thermodenitrificans* (*Gt*$YybT_{55-658}$) was purchased from GeneScript and cloned into PET28(a+) (Novagen) between the NdeI and XhoI restriction sites, resulting in the N-terminal His_6-tagged recombinant construct *Gt*$YybT_{55-658}$. The plasmids were used to transform *Escherichia coli* strain BL21 (DE3). Subsequent expression and purification procedures were the same as used for *Bs*YybT (131). To confirm whether the PAS domain alone can bind heme, we made three constructs (*Gt*$YybT_{55-303}$, *Gt*$YybT_{55-162}$, and *Gt*$YybT_{83-162}$) by mutating the codon encoding residue 304 and 163 to stop-codons. Site-directed mutagenesis was performed as before (128). All three proteins were associated with heme. Protein *Gt*$YybT_{55-162}$ has better stability and monodispersity in gel-filtration chromatography. It was therefore used for NMR studies. To prepare *Gt*$YybT_{55-162}$ for NMR studies, 1 ml overnight culture was used to inoculate 800 ml M9 minimal medium (composition for 1 litre: 12.8 g Na_2HPO_4, 3.0 g KH_2PO_4, 0.5 g NaCl, 1 g $^{15}NH_4Cl$, 2 g ^{13}C-glucose, 0.494 g

$MgSO_4 \cdot 7H_2O$, 0.0152 g $CaCl_2 \cdot 7H_2O$, 0.01 g thiamine, 0.01 g $FeSO_4 \cdot 7H_2O$). Bacterial culture (LB medium) was grown to OD = 0.8 before inducing with 0.6 mM IPTG. The culture was shaken at 18^{o}C for ~16 hours before it was pelleted by centrifugation. The cells were lysed in 20 ml lysis buffer [50 mM Sodium Phosphate (pH 6.5), 150 mM NaCl, 0.1% β-mercaptoethanol, and 1 mM PMSF]. After the centrifugation at 25,000 rpm for 30 min, the supernatant was filtered and then incubated with 2 ml of Ni-NTA resin for 1 hour at 4 ^{o}C. The resin was washed with 50 ml of W1 buffer (lysis buffer with 20 mM imidazole) and 50 ml of W2 buffer (lysis buffer with 50 mM imidazole). The bound proteins were eluted using a step gradient method with elution buffers containing 50 mM Sodium Phosphate (pH 6.5), 150 mM NaCl, and 200mM, 300 mM, or 500 mM imidazole. After the analysis by SDS-PAGE, fractions with purity higher than 95% were pooled together. Size-exclusion chromatography was carried out at 4 ^{o}C using the AKTA FPLC system equipped with a Superdex 75 HR 16/60 column (Amersham Biosciences). The buffer used for gel filtration comprised of 50 mM Sodium Phosphate (pH 6.5), 150 mM NaCl and 2x Protease inhibitor cocktail (Calbiochem). The proteins were stored in -80 ^{o}C freezer after snap freezing and concentration measurement by Bradford assay method. The structure determination by NMR is ongoing collaboration work with the group of Prof. Konstantin Pervushin, and is not included in this thesis.

5.2.3 Pyridine hemochrome assay

To determine the type of heme that was bound by purified *BsYybT* and *Gt*YybT, pyridine hemochrome assay was performed according to Berry and Trumpower (11). Briefly, 0.5 mL aliquot of the sample was added to the stock solution (200 mM NaOH; 40% (v/v) pyridine; 3 µL of 0.1 M $K_3Fe(CN_6)$), followed by thorough mixing. The oxidized spectrum was recorded (wavelength 800nm-200nm) several times till stable using a Shimadzu UV-1700

spectrophotometer. Sodium dithionite (2-5mg) was added to the cuvette and the spectrum of the reduced pyridine hemochromes was recorded. To estimate the percentage of hemin incorporation, the absorbance difference at the selected wavelength between the reduced and the oxidized spectra was used to calculate the concentration of heme based on extinction coefficients as described (11).

5.2.4 Heme reconstitution.

The hemin chloride stock (50mM) in 0.1 M NaOH was slowly titrated into purified protein (50 μM) with gentle agitation for ~4 hours with the final heme/protein ratio of 5:1. The unbound heme was then separated from protein by size-exclusion chromatography using an AKTA FPLC system (131). Fractions corresponding to *holo*-protein were pooled and first dialyzed into a series of PBS buffer (pH7, 6, 5.5). The protein was then exchanged back to the original buffer (50 mM Tris (pH 8.0), 150 mM NaCl, and 5% glycerol) by dialysis.

5.2.5 UV-Vis spectroscopy

Unless otherwise noted, all absorption spectra were measured with 200 μl holo-YybT (10 -20 μM) in 50mM Tris (pH 8.0), 150 mM NaCl, at 23 °C in a sealed quartz cuvette. Ferrous deoxy-YybT was obtained after the addition of 1 mM sodium dithionite (DTH). The NO donors stocks (100 mM) were made in nitrogen-saturated 0.01 M NaOH in anaerobic chamber and stored at -80 °C in aliquots. The CO donor (CORM2) was prepared in DMSO as a 100 mM stock and stored at -80 °C. Following the exposure to NO donors (100 μM), CO donor (100 μM), KCN (1 mM), or oxygen bubbles, the spectra of the various forms of YybT were recorded with a UV-Vis spectrophotomerter.

5.2.6 Phosphodiesterase activity assay

All the enzymatic reactions were carried out in a Coy anaerobic chamber, as used previously (125). The assay conditions and steady-state kinetic measurement for *apo*-YybT were reported previously. Assay for the ferric holo-protein in the presence and absence of potassium cyanide (1 mM) were performed following the conditions previously used for apo-YybT (131). The effect of redox state and gases binding on enzymatic activity was measured in the reaction buffer containing 100 mM Tris-HCl (pH 7.3), 20 mM KCl and 0.5 mM $MnCl_2$. 100 µl protein (0.5 mg/ml) was split into two vials, followed by the addition of 0.5 µl 100 mM of sodium dithionite to one portion and 0.5 µl of 100 mM NaCl to another. To remove residual dithionite that may interfere with the reaction, the reduced and mock-treated proteins were buffer-exchanged by passing through pre-equilibrated desalting spin columns (Pierce). To initiate the reaction, 5 µl of the reduced or oxidized proteins were added to 195 µl reaction buffer containing 5 µM c-di-AMP. For the NO and CO-ligated protein forms, the buffer was saturated with N_2 or CO gases through extensive bubbling. CO gas was used for activity assay because the CO donor compound CORM2 decomposes into a divalent ruthenium salt that inhibited enzyme activity. The effect of NO binding was evaluated by the addition of 100 µM freshly prepared DEA-NONOate. Reactions were stopped at various time points by the addition of 1/10 vol of 0.5M EDTA. The progress of c-di-AMP hydrolysis was monitored using the same Agilent LC1200 system as before (131).

5.2.7 ATPase activity detection

The ATPase reaction was perfomed as described previously (131). Reactions were carried out at 23°C with the assay conditions: 50 mM Tris (pH 7.5), 1 mM $MgCl_2$, 1 mM ATP, 4 µM enzyme. The HPLC method of ATPase activity measuremen was carried out suing reverse-phase C-18 column, and the mobile phase used was: 100 mM KH_2PO_4 (pH 6.5), 10 mM

Tetrabutyl ammonium bromide (TBAB), 5.5% Acetonitrile, 1.5 ml/min. In this solvent system, the rentention times for ADP and ATP are 5.0 min and 12.0 min, respectively.

5.2.8 Survival of wild type and ΔyybT mutant under stress conditions

The survival of the wild type *B. subtilis* strain 168 and YybT mutant strain (ΔYybT) were examined using an assay similar to the one for acid resistance conducted previously (131). Cells were grown at 37 °C in casein hydrolysate (CH) medium overnight. The culture was then inoculated in CH medium at OD_{600} = 0.01 and grown for four hours in the presence/absence of the stress reagents, after which serial dilutions were plated in triplicate for cfu counting. For nitrosative stress, 200 μM DEA-NONOate or sodium nitropruside (SNP) was added to the culture every hour, based on the finding that continuous rather than one-time addition has stronger effect (109). For cyanide stress, HCl and KCN were added at a final concentration of 100 μM each, and the tubes were then sealed. For heme stress, the method was adapted from Stauff *et al* (157), where 20 μM heme was added during inoculation. For CO stress, 50 μM CORM2 was added to culture that has already grown for 4 hours. For H_2O_2 stress, the method was adapted from Gusarov *et al* (55), where, after 4 hr of initital growth, 10 mM H_2O_2 was added and cultures were plated after 30 minutes of H_2O_2 exposure. Where indicated, the NO synthase (NOS) inhibitor L-NAME (500 μM), which effectively inhibits bacillus NOS (148), was added 10 min prior to the addition of H_2O_2, and the NO donor DEA-NONOate (50 μM) was added 10 seconds prior to the addition of H_2O_2.

5.3 RESULTS

5.3.1 Purified YybT proteins bind heme

In the previous study of the function of the GGDEF domain and DHH/DHHA1 domain of *Bs*YybT, we have cloned three protein constructs for enzymatic assays (**Fig. 5.1A**). Surprisingly, we noticed that the concentrated PAS-GGDEF di-domain containing construct (*Bs*YybT$_{84-303}$) has a faint brownish appearance. Absorption spectroscopy showed an absorbance peak at 412 nm and two other smaller peaks at 533 and 566 nm (**Fig. 5.1B**). Although this spectrum pattern is reminiscent of heme protein, the absorbance was previously attributed to a co-purified *E. coli* heme protein. However, when we expressed the cytoplasmic portions of the thermophilic homolog of YybT (*Gt*YybT$_{55-658}$ and *G*tYybT$_{55-304}$) from the thermophilic firmicutes *Geobacillus thermodenitrificans*, the solution of the protein purified by metal affinity and gel-exclusion chromatography exhibit a more intense yellow-brown color. The absorption spectra of GtYybT$_{55-658}$ and GtYybT$_{55-304}$ exhibit identical Soret and α/β peaks as *Bs*YybT$_{84-303}$ (**Fig. 5.1B**). The $A_{412\ nm}/A_{278\ nm}$ ratio did not seem to decrease for the protein with increased homogeneity, suggesting that heme is associated with the YybT protein rather than a contaminating protein. Notably, the $A_{412\ nm}/A_{278\ nm}$ ratio for *Gt*YybT$_{55-304}$ is greater than that of the *Bs*YybT$_{84-303}$, indicating a high occupancy of the heme for the recombinant thermophilic protein.

YybT family proteins contain a PAS domain that is known to bind heme and other cofactors. The likelihood of heme-binding by PAS domain is also supported by the observation that the protein construct *Bs*YybT$_{162-659}$ does not bind the heme. To verify that the heme is indeed bound by the PAS domain, we cloned the stand-alone PAS domain of *Bs*YybT and *Gt*YybT. Although the insolubility of the PAS_{BsYybT} protein construct (*Bs*YybT$_{55-162}$) prevented further study, the stand-alone PAS domain of GtYybT (*Gt*YybT$_{55-}$

$_{162}$) was purified with heme associated (**Fig. 5.1B**). The $A_{412\ nm}/A_{278\ nm}$ ratio for *Gt*YybT$_{55-162}$ is comparable to that of *Gt*YybT$_{55-304}$.

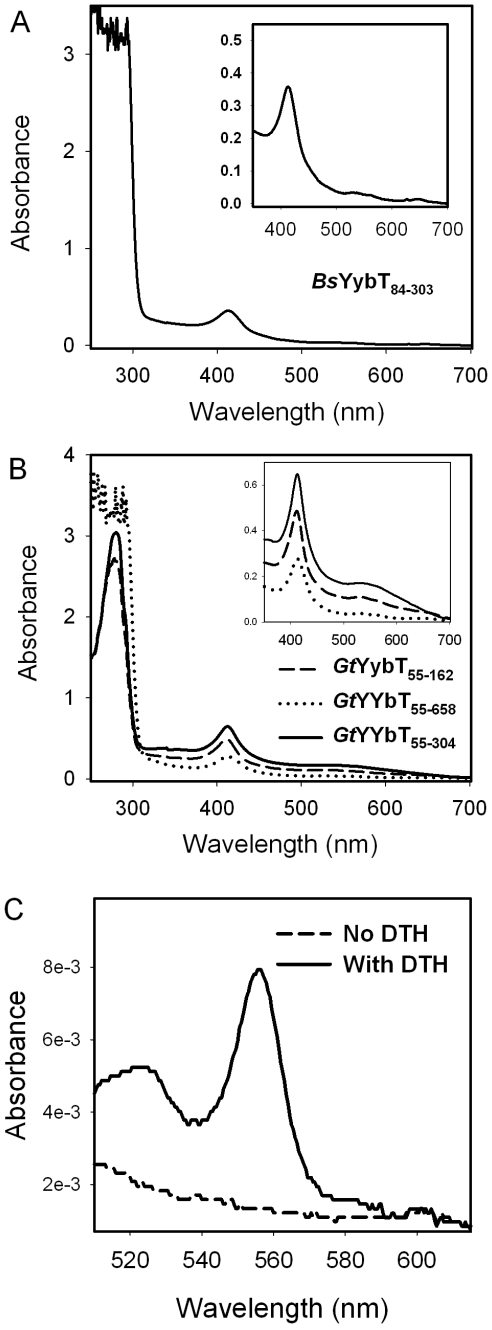


Figure 5.1 YybT PAS domain contains a heme-binding site. UV-Vis absorption spectra of purified YybT proteins. **A.** Absorption spectrum of BsYybT$_{84-303}$. **B.** Absorption spectra of the three *Gt*YybT protein constructs. **C.** Spectrum of associated heme assayed by pyridine hemochrome method.

The heme bound by *Gt*YybT$_{55-162}$ was confirmed to be heme b by the pyridine hemochrome method (11), with diagnostic absorption peak at 557 nm and 525 nm when reduced with dithionite (**Fig. 5.1C**). These results demonstrate that the PAS domain of YybT family proteins is capable of binding heme b, though the amount of heme associated with the purified proteins is sub-stoichiometric. As quantified by the pyridine hemochrome assay, the ratio of heme to protein ranged from 0.5 % for *Bs*YybT$_{84-303}$ to 5% for *Gt*YybT$_{55-162}$. Supplementing heme or the heme precursor 4-amino levulinate to cell culture or cell lysate did not significantly increase the percentage of the *holo*-protein.

5.3.2 Effect of mutations on heme binding

The core PAS domains of YybT family proteins contain only 70-80 residues, significantly shorter than the ~100 residue canonical PAS domains. The serendipitous finding of heme binding is unexpected as there is no histidine in the PAS domain of *Bs*YybT. Furthermore, sequence alignment of the YybT homologs does not identify any highly conserved residue that is suitable to coordinate iron (**Fig. 5.2A**), except an Asn (N^{100} for *Bs*YybT) that is a signature residue conserved among all PAS domains (38). The sequence alignment further highlights that the C-terminal portion of PAS$_{YybT}$, namely helix α3, the linker between α3 and β3, strand β4 and β5, are shortened in comparison to the canonical heme-binding PAS domain protein *Ec*Dos or *Bj*FixL (53, 89) **(Fig. 5.2A)**.

The canonical heme-binding pocket usually lies in a deep hydrophobic pocket made up by the five beta strands and a long flanking alpha helix (107). Secondary structure prediction and 3-D models showed the same topology and fold for PAS$_{YybT}$ (**Fig. 5.2B**). Furthermore, a cluster of hydrophobic residues, including M^{88}, L^{103}, F^{107}, L^{139}, F^{144}, and F^{157}, are located in the pocket region (**Fig. 5.2B**). These positions are conserved among YybT homologs as bulky hydrophobic residues **(Fig. 5.2A)**. To determine whether this region influence heme-binding,

a mutant protein *Gt*YybT$_{55-162}$-F107A was prepared and its heme protein ratio was examined. Compared to wildtype, the mutant protein has similar yield, but with a ~90% reduction in heme content (**Fig. 5.2C**), which underscores the importance of F^{107} and the hydrophobic region in heme-binding. Gel-filtration chromatography showed overlapped dimeric peaks for both wild-type and mutant proteins, suggesting that the mutation does not pertube tertiary structure or dimer interface. The results suggest that heme binds into a well-defined pocket in the PAS domains, but the proximal ligand remain to be identified.

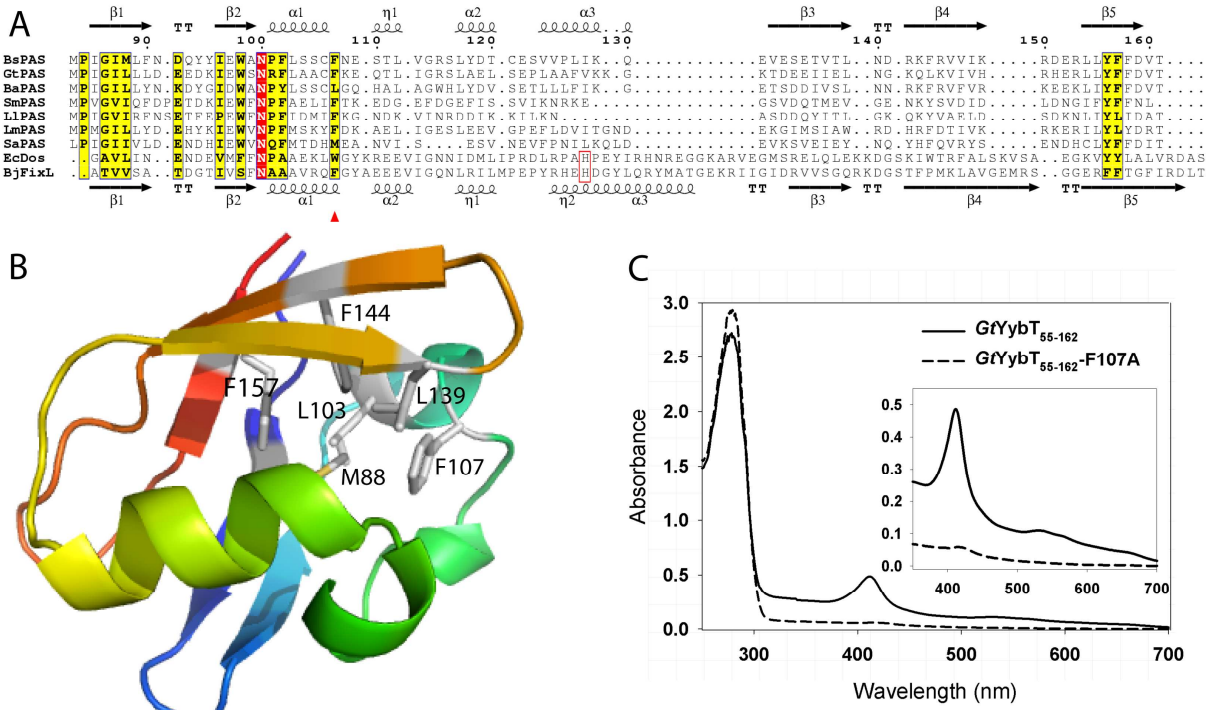


Figure 5.2 Effect of PAS domain mutation on heme binding. **A.** Sequence alignment of the PAS domain among YybT homologs and two heme binding PAS domains (*Ec*Dos and *Bj*FixL). Sequences were aligned by using ClustalW. Residue numbering is based on *Bs*YybT. The position of residue F^{107}, which was replaced by Ala, is highlighted by the red triangle. The axial His ligands in *Ec*Dos and *Bj*FixL are marked by red box. Bs, *Bacillus subtilis*; Ba, *Bacillus anthracis*; Gt, *Geobacillus thermodenitrificans*; Ll, *Lactooccus lactis*; Lm, *Listeria monocytogene*; Sa, *Staphylococcus aureus*; Sm, *Streptococcus mutans*; The sequence was aligned using CLUSTALW and the figure was generated using Espript 2.2. **B.** Homology model of BsYybT PAS domain generated by I-Tasser. The model with top score was selected. The conserved hydrophobic residues are shown as sticks. **C.** Absorption spectrum of GtYybT$_{55-162}$ wildtype and F107A mutant.

5.3.3 Reconstitution of YybT proteins

Given the low heme occupancy of the purified proteins, reconstitution methods were tested to prepare *holo*-proteins. We used a reconstitution method that involves the titration of alkaline heme solution into the protein solution and gradual dialysis for pH adjustment. After the removal of the excess heme by size-exclusion chromatography, the reconstituted proteins exhibited similar spectra as the purified *Bs*YybT$_{84\text{-}303}$ and *Gt*YybT$_{55\text{-}162}$, with absorbance maxima around 412, 533 and 566 nm (**Fig. 5.3**). The protein to heme ratio and the size of the shoulder peak at 365 nm varied with protein constructs. The ratio of protein to heme is close to 1:1 for the full-length protein constructs *Bs*YybT$_{84\text{-}659}$ (hereafter named *holo*-YybT) and *G*tYybT$_{55\text{-}658}$ according to pyridine hemochrome assay. The protein to heme ratio is significantly lower for *Gt*YybT$_{55\text{-}162,}$ which indicates that heme binding is somehow weakened in the free-standing PAS domain (**Fig. 5.3**). As a control, the GGDEF-DHH/DHHA1 di-domain containing *Bs*YybT$_{150\text{-}659}$ failed to bind any heme in the reconstitution, ruling out the non-specific binding of heme by the two catalytic domains. In addition, $(His)_6$-tag free protein prepared by thrombin cleavage of *Gt*YybT$_{55\text{-}162}$ is fully capable of binding heme, indicating that the histidine residues of the $(His)_6$-tag are not involved in the coordination of the heme.

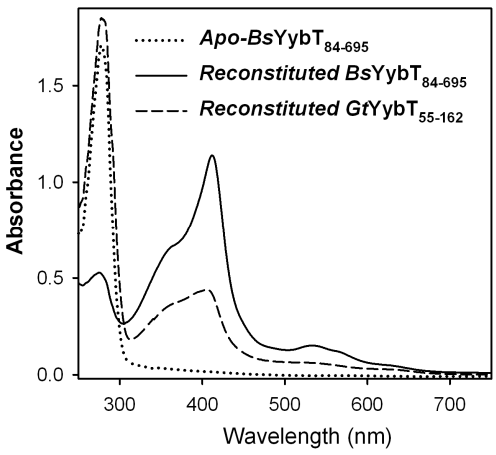


Figure 5.3 UV-Vis absorption spectra of apo-*Bs*YybT$_{84\text{-}695}$ and reconstituted *Bs*YybT$_{84\text{-}695}$ and *Gt*YybT$_{55\text{-}162}$. Hemin (ferric) was used for reconstitution without any reducing agent, the spectra shown here thus represents heme in the ferric state.

5.3.4 Absorption Spectra of *holo*-YybT

For heme proteins, the peak wavelengths of the two π→π* electronic transitions (one is the soret band, the other is the β/α band) are sensitive to the redox, spin state and coordination environment of the heme iron. In general, heme with a penta-coordinated iron has a soret band between 350 to 400 nm, whereas a band above 400 nm indicates hexa-coordinated heme (154). Additionally, a small band between 600-650 nm is characteristic of charge-transfer between the porphyrin ring and the high-spin Fe^{3+} (8). Examination of the spectra of the ferrous and ferric will yield valuable information about the heme binding mode. Since heme-binding PAS domains usually function as sensor domain for O_2, nitric oxide (NO), carbon monoxide (CO) and redox signal, we also examined the spectroscopic changes caused by the binding of the gaseous NO and CO ligands as well as the strong-field ligand cyanide (CN^-). Cyanide is usually considered as a physiologically irrelevant molecule, but the soil bacterium *P. fluorescens* secrets significant amount of cyanide to its natural habitat where *B. subtilis* also dwells (57).

Ferric YybT - The reconstituted *holo*-YybT is in the ferric form, with the soret peak at 412 nm, the β band at 533 nm, followed by a shoulder α band at 566 nm (**Table 5.1**, **Fig 5.4A**), characteristic of a hexa-coordinated heme protein. Absorbance peak was also observed at 639 nm (**Fig. 5.3**), suggesting that the heme iron is in its high-spin state. The small peak above 600 nm was also observed in the spectrum of purified *Bs*YybT$_{84-303}$, *Gt*YybT$_{55-162}$, and *Gt*YybT$_{55-303}$ **(Fig. 5.1A)**, suggesting that the reconstituted protein recapitulated the property of natural heme binding. No change in the soret or β/α band was observed when NO donors (Spermine, MAHMA, and DEA-NONOate) or CO donor (CORM2) were supplied to the protein solution (**Fig. 5.4A**), indicating that no binding occurred. Together with the Fe^{2+}-YybT spectrum shown below, the results also showed that NO is unable to reduce ferric-

YybT, unlike other heme protein such as the dissimilative nitrate respiration regulator DNR (48) and CLOCK (98). Addition of KCN immediately caused a persistent red shift of the soret peak to 417 nm, and the merge of α and β bands into a single peak at 536 nm (**Fig. 5.4A**), suggesting fast formation of Fe^{3+}-CN complex.

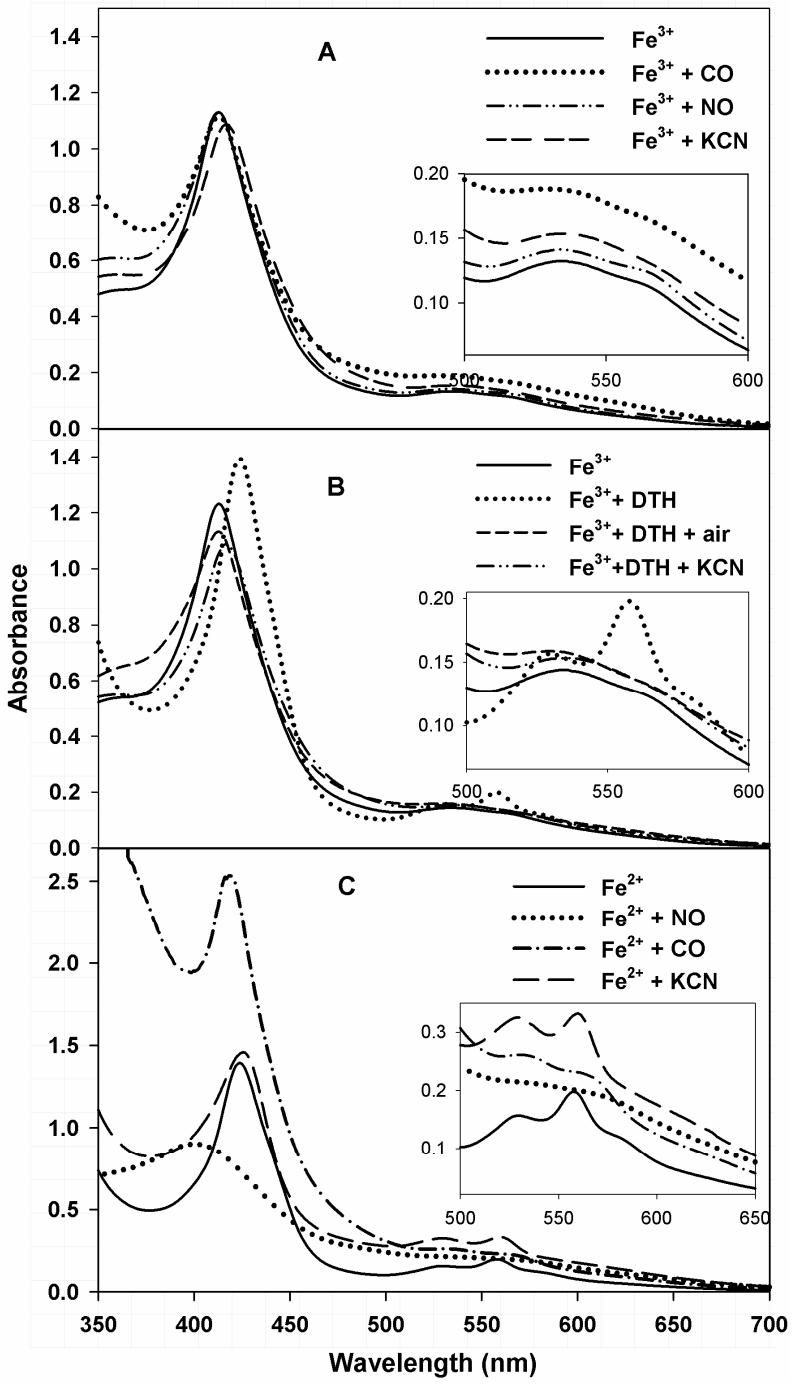


Figure 5.4 Absorption spectra of *holo*-YybT proteins at the Ferric (**A**) and Ferrous state (**B, C**), in the presence of various potential ligands.

Table 5.1 UV-Vis peak positions for hemoproteins in the ferric and ferrous states.

Protein	Heme state	Soret (nm)	α/ β (nm)	Others (nm)	Ref
***Bs*YybT**	Fe^{3+}-	412	566/533	639	This work
	Fe^{2+}-	423	557/528		
sGC	Fe^{3+}-	391	506/538	634	(160)
	Fe^{2+}-	431	562		(159)
DNR	Fe^{3+}-	417	566/533		(48)
	Fe^{2+}-	426	560/530		
HRI	Fe^{3+}-	418	538		(67)
	Fe^{2+}-	426	560/531		
E75	Fe^{3+}-	425	569/542	650/750	(101, 133)
	Fe^{2+}-	426	559/531		
RcoM	Fe^{3+}-	420	574/539		(77)
	Fe^{2+}-	426	562/531		
***Ax*PDEA1**	Fe^{3+}-	Not reported			(23)
	Fe^{2+}-	431	573/550		
***Bj*FixL**	Fe^{3+}-	395	509	645	(51)
	Fe^{2+}-	437	556		
***Ec*Dos**	Fe^{3+}-	417	564/530		(145)
	Fe^{2+}-	428	563/532		

Table 5.2 UV-Vis absorption maxima for hemoproteins in the presence of NO, CO and O_2

Protein	Heme state	Soret (nm)	α/ β (nm)	Remark	Ref
BsYybT	Fe^{3+}-CN^-	417	536	NO sensor	This work
	Fe^{2+}- NO	**398**	568/532		
	Fe^{2+}-CO	418	567/530		
	Fe^{2+}-CN^-	426	561/529		
	Fe^{2+} + O_2	Immediate oxidation			
sGC	Fe^{3+}-CN	418	538	NO sensor	(160) (159)
	Fe^{2+}-NO	**398**	572/537		
	Fe^{2+}-CO	423	567/541		
	Fe^{2+} + O_2	No binding or oxidation in 24 hr			
DNR	**Fe^{2+}-NO**	**389**	~550	NO sensor	(48)
	Fe^{2+}-CO	421	570/540		
	Fe^{2+} + O_2	Oxidation within mins			
HRI	**Fe^{2+}-NO**	**398**	538	NO sensor	(67)
	Fe^{2+}-CO	421	575/544		
	Fe^{2+} + O_2	Oxidation with one minute			
E75	**Fe^{2+}-NO**	**391**	569/541	NO sensor	(133)
	Fe^{2+}-CO	421	569/540		
	Fe^{2+} + O_2	Oxidation in hours			
RcoM	Fe^{2+}-NO	421	578/545	CO sensor	(77)
	Fe^{2+}-CO	423	570/540		
	Fe^{2+} + O_2	Oxidation in 2 hours			
*Ec*Dos	Fe^{2+}-NO	417	579/541	O_2 sensor	(145)
	Fe^{2+}-CO	419	563/537		
	Fe^{2+}-O_2	423	570/540		
*Ax*PDEA1	Fe^{2+}-NO	Not reported		O_2 sensor	(23)
	Fe^{2+}-CO	Same as EcDos			
	Fe^{2+}-O_2	Same as EcDos			
*Bj*FixL	Fe^{2+}-O_2	419	560/545	O_2 sensor	(51)
	Fe^{2+}-CO	427	560/548		
*Ec*Dos	Fe^{2+}-NO	417	579/541	O_2 sensor	(145)
	Fe^{2+}-CO	419	563/537		
	Fe^{2+}-O_2	423	570/540		

Ferrous YybT - treatment with DTT reduced *holo*-YybT to its Fe^{2+} state, which is evidenced by a change in the absorption spectrum: the soret band is red shifted to 423 nm; the β and α bands are more distinct and blue shifted to 528 nm and 557 nm, respectively; and the band at 639 nm disappeared (**Fig. 5.4B**). The resulted spectrum is consistent with a hexa-coordinated low-spin heme. When this reduced sample was exposed to atmospheric oxygen, or when oxygen was bubbled into the protein solution, the absorption spectrum reverted back to that of Fe^{3+}-YybT within seconds (soret, β, and α bands at 412nm, 566nm, and 532 nm, respectively) (**Fig. 5.4B**), suggesting that Fe^{2+}-YybT cannot bind oxygen but is instead easily oxidized. In line with this, treatment of the air-exposed sample with 500 μM $Fe(CN)_6^{3-}$, an oxidant that effectively oxidize ferrous heme iron (88, 159), did not result in observable changes in peak positions. To further rule out the possibility that O_2 binding to ferrous heme happen to shift the spectrum such that it looks identical to the spectrum of Fe^{3+}-YybT, we employed the characteristic binding of cyanide to Fe^{3+}-YybT described above. Addition of KCN to the air-exposed sample induced red shift of soret band to 417 nm and merge of α and β bands to give a single band at 535 nm, yielding spectrum identical to the addition of KCN to Fe^{3+}-YybT (**Fig. 5.4B**). In comparison, addition of KCN to the reduced sample (Fe^{2+}) immediately resulted in a red shift of the soret peak from 423 nm to 426 nm, with accompanying changes in the spectrum for the β /α bands to 561/529 nm (**Fig. 5.4C)**, suggesting that Fe^{2+}-YybT forms a complex with CN^- that is different from Fe^{3+}-YybT-CN^-. We thus conclude that the air-exposed sample is in the oxidized Fe^{3+} form. The rapid auto-oxidation suggests that the PAS domain of YybT is very different from some of the O_2 sensing heme-PAS proteins that exhibit resistance to O_2 oxidation (23, 51, 145) **(Appendix III)**. Cyanide binding is mostly examined for heme proteins in the ferric state, and cyanide usually does not bind ferrous heme or binds weakly. For instance, ferrous heme in DosT and sGC is not capable of forming a complex with CN^- (31, 88). The binding of cyanide to

ferrous-YybT is thus an interesting observation, and is further validated through the activity assay, which will be described later.

Addition of NO and CO donors to the ferrous YybT immediately caused shifts and changes in the Soret and α/β bands, with the β band being the major peak and the α band only a shoulder peak (**Fig. 5.4C, Table 5.2**), suggesting binding of the ligands to the ferrous iron. The NO and CO complexes are stable as the spectra remained unchanged even when exposed to air for more than 60 min. Notably, the changes in optical spectrum caused by CO and NO binding differ drastically. The binding of CO shifts the Soret peak from 423 nm to 418 nm and results in less intense α/β bands, indicating that the CO complex is likely to contain a hexa-coordinated low-spin heme. In comparison, the binding of NO results in a broad and less intense Soret band at 398 nm and ill-defined α and β bands. The 398 nm Soret band suggests that the NO-complex likely contains a penta-coordinated heme that resembles those of the NO sensors, including heme-nitric oxide/oxygen binding domain (HNOX) protein soluble guanylate cyclase (sGC) (34, 159), the nuclear receptor E75 (133), the CRP-FNR superfamily transcription factor DNR (20, 48), and the heme-regulated eukaryotic initiation factor 2α kinase (HRI-ND) (67, 177). In addition, it is common among the five proteins that the extinction coefficients decrease upon NO binding and increase upon CO binding.

Taken together, the spectroscopic studies suggest that the ferric YybT binds cyanide, but not NO or CO, whereas the ferrous YybT binds all three ligands and is prone to oxidation by O_2. The spectral features indicate that the Fe^{2+}-CO and Fe^{2+}-NO complexes contain hexa-coordinated and penta-coordinated heme iron respectively. We also observed similar spectroscopic changes with *Gt*YybT$_{55\text{-}162}$, which is naturally associated with heme, indicating that the heme environment of the free-standing PAS domain is comparable, and also that the reconstituted holo-YybT retained the properties of natural PAS domain.

5.3.5 Effect of heme and ligand binding on the phosphodiesterase activity

We previously noticed a difference in PDE activity between the heme free form of *Bs*YybT$_{84-659}$ (hereafter named *apo*-YybT) and *Bs*YybT$_{150-659}$ that lacks the PAS domain (131). Compared to *Bs*YybT$_{150-659}$, *apo*-YybT exhibits a 7.4 fold higher k_{cat}, and a 2.3 fold lower K_m (**Table 5.3**), suggesting that the PAS domain may play a regulatory role in modulating the enzyme activities. In light of the heme binding property of the PAS domain, the effect of heme binding and ligand coordination on the catalytic activity of the DHH/DHHA1 and GGDEF domains was examined.

Table 5.3 Kinetic parameters for c-di-AMP hydrolysis by *Bs*YybT at different heme state

Protein	k_{cat} (s^{-1})	K_m (μM)	k_{cat}/K_m (s^{-1} μM^{-1})	*reference*
YybT$_{150-659}$	0.074 ± 0.002	3.0 ± 0.22	0.024 ± 0.002	(131)
Apo-YybT	0.55 ± 0.02	1.3 ± 0.24	0.42 ± 0.08	(131)
Holo-YybT	0.024 ± 0.003	16 ± 4.0	0.0015 ± 0.0004	This work
Holo-YybT + 1mM KCN	0.083 ± 0.003	8.9 ± 1.0	0.009 ± 0.001	This work

Effect of heme binding and redox change. Although the physiological substrate of the DHH/DHHA1 domain is unknown at this moment, c-di-AMP was used as the substrate for *in-vitro* enzymatic assay given the efficient hydrolysis of the cyclic dinucleotide catalyzed by the DHH/DHHA1 domain (131). Compared to *apo*-YybT, a significant decrease of about 15 fold in initial velocity activity was already evident for *holo*-YybT (**Fig. 5.5A**). Subsequent kinetics measurement revealed a k_{cat} of 0.024 ± 0.003 s^{-1} and a K_m of 16 ± 4 μM, which are 23 fold lower and 12 fold higher than those of apo-YYbT, resulting in an overall 276-fold decrease in catalytic efficiency (**Table 5.3**). Hence, heme binding in the PAS domain exerts an inhibitory effect on the catalytic activity of the DHH/DHHA1 domain. We next examined whether redox state of the heme iron affects the catalytic activity by measuring the phosphodiesterase activity of ferrous and ferric YybT. No statistically significant difference in the rate of c-di-AMP

hydrolysis was observed (**Fig. 5.5B**) and the turnover progress curve overlapped (**Fig. 5.5A**), suggesting that YybT is unlikely to be a direct sensor of redox potential.

Effect of cyanide binding. Binding of CN^- induced a 5-fold increase in activity for ferric-YybT and a relatively lower but significant 2-fold increase in activity for ferrous-YybT (**Fig. 5.5A, B**); whereas CN^- had no influence on the reaction progress of *apo*-YybT (**Fig. 5.5A)**. To rule out the possibility that the effect of cyanide on ferrous-YybT is due to the presence of auto-oxidized enzyme, I examined the activity of Fe^{2+}-YybT in the presence of excess DTH and observed similar increase in activity. The lower magnitude of stimulation for ferrous-YybT seems to be consistent with the milder shift of absorption spectrum (**Fig. 5.4A, 5.4C, Table 2**).

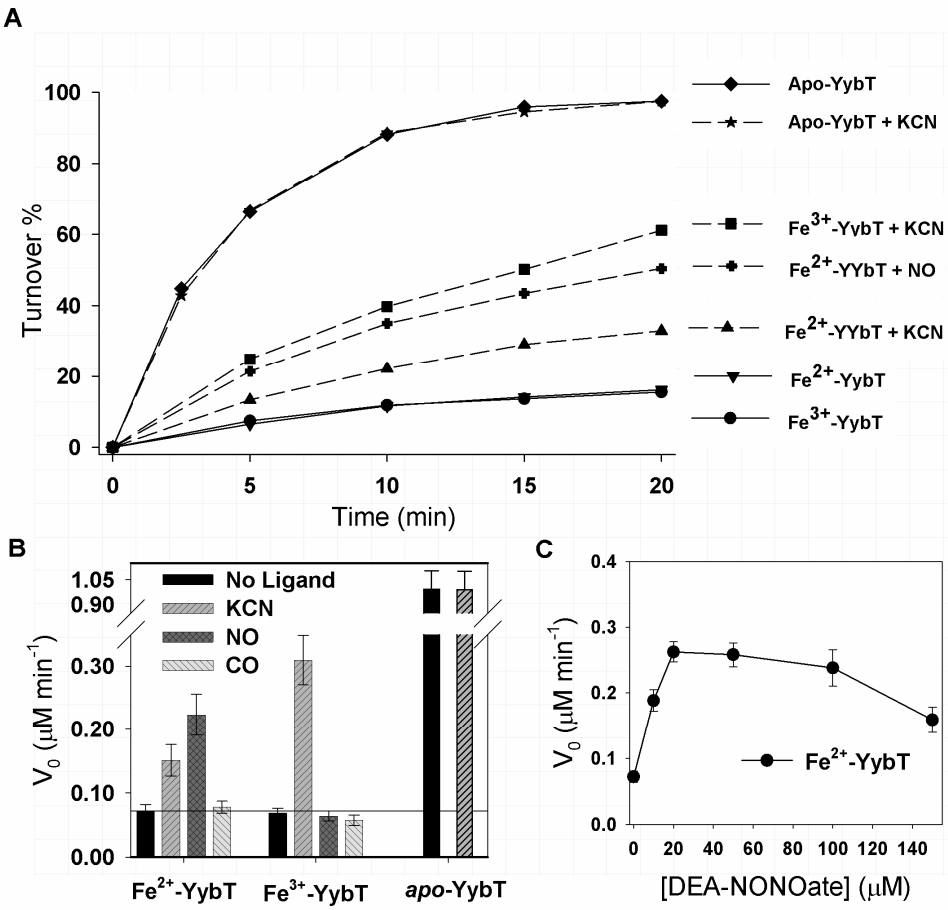


Figure 5.5 The effect of ligand binding on YybT PDE activity. **A,** the time dependant progress curve of c-di-AMP turnover. **B**, Initial velocity of c-di-AMP (5 μM) hydrolysis. **C**, PDE activity of Fe^{2+}-YybT in the presence of different concentrations of NO-donor.

Effect of NO and CO binding. The NO- and CO- complexes showed drastically different spectra, we therefore compared the activity of ferrous YybT in the presence/absence of NO and CO. Ferric YybT was included as a control to detect potential non-specific effect since it does not bind NO or CO. As shown in **Fig. 5.5A and 5.5B**, NO binding increased the activity of ferrous-YybT by 3-fold, whereas CO had a negligible effect. The stimulatory effect of NO was observed for all the three NO donor tested: DEA-NONOate, sMAHMA-NONOate and Spermine-NONOate. The stimulatory effect of NO is dependent on the concentration of the NO donor, with a saturating concentration of 20 μM observed for DEA-NONOate (**Fig. 5.5C**). The observation further supports the presence of specific binding site(s) for NO in ferrous YybT, presumably the heme iron center. In contrast to the ferrous YybT, no significant effect of CO and NO on ferric YybT was observed (**Fig. 5.5B**), which is consistent with the conclusion that CO and NO only coordinate to the ferrous iron. This observation further suggests that the stimulatory effect is exerted through binding to ferrous heme iron (**Fig. 5.5B**).

5.3.6 Kinetic measurements and specificity of ligand-binding induced PDE activation

To probe whether the increase in activity is due to increase in k_{cat} or decrease in K_m, steady-state kinetic parameters were measured. The NO donors prefers low pH for gas release and their effect was measured in pH7.3, which does not allow direct comparison with kinetic parameters of holo-YybT that was measured at pH 8.3. In comparison, the lack of non-specific effect/reactivity, the ease of use, and precise control of actual concentration makes CN^- a more preferable tool for detailed characterization of the mechanism of PDE activation. Addition of cyanide resulted in a 3.5-fold increase in k_{cat} and a 2-fold decrease in K_m (**Fig. 5.6A, Table 5.3**), suggesting an overall tighter interaction between enzyme and substrate.

I then probed whether the activation is specific for c-di-AMP as a substrate. To this end, we examined the hydrolysis against c-di-GMP and 2',3'-cAMP. C-di-GMP is another cyclic-dinucleotide that can also be hydrolyzed by YybT, with a much higher K_m (131). 2',3'-cAMP is a recently identified non-specific PDE substrate that can be hydrolyzed by many metal-dependant PDEs (including YybT), as long as it is accessible to nucleophilic attack by metal-activated water/hydroxide in the active site. (130). Addition of KCN had no effect on the rate of 2',3'-cAMP hydrolysis, but increased the rate of c-di-GMP and c-di-AMP hydrolysis by 2-fold, and 3.5-fold, respectively (**Fig. 5.6B**), suggesting that the allosteric effect is not due to the generation of a better attacking nucleophile, in which case hydrolysis of 2',3'-cAMP would also be increased. We speculate that conformational change in active site residues that specifically interact with the cyclic-di-nucleotide, rather than active site residues involved in coordinating catalytic metal ion, is the underlying mechanism for this activation. This is also consistent with a change both in k_{cat} and K_m (**Fig. 5.6A**). Moreover, the higher magnitude of activation of c-di-AMP hydrolysis compared to that of c-di-GMP hydrolysis seems to be consistent with the notion that c-di-AMP is the physiological substrate and the PDE active site is designed to effectively regulate its level.

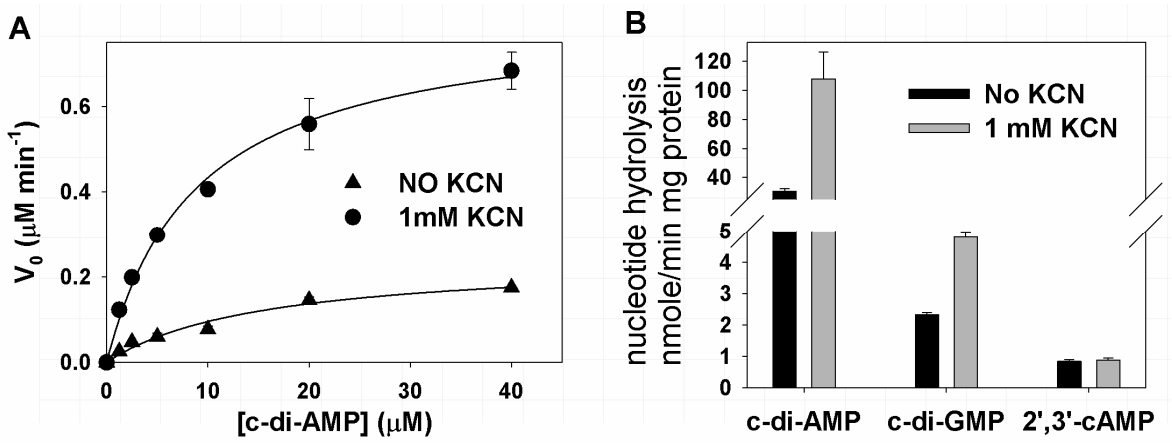


Figure 5.6 Effect of cyanide on the kinetics and specifitiy of PDE activity. **A.** Steady-state kinetic measurement of PDE activity of Fe^{3+}-YybT in the presence and absence of KCN. **B.** PDE activity of Fe^{3+}-YybT against three substrates (20 μM) in the presence and absence of KCN.

5.3.7 Effect of heme on ATPase activity

Unlike the orthodox GGDEF domains, the GGDEF domain of YybT is highly degenerate and does not possess the diguanylate cyclase (DGC) activity. We previously found that the degenerate GGDEF domain of the *apo*-YybT$_{84\text{-}303}$ exhibits unexpected ATPase activity with a small k_{cat} of 0.59 ± 0.03 min^{-1}(131). The binding of the heme to both the *Bs*YybT$_{84\text{-}303}$ and *Bs*YybT$_{84\text{-}659}$ protein constructs further suppresses the ATPase activity to a residual level that is close to detection limit of the HPLC method (**Fig. 5.7**).

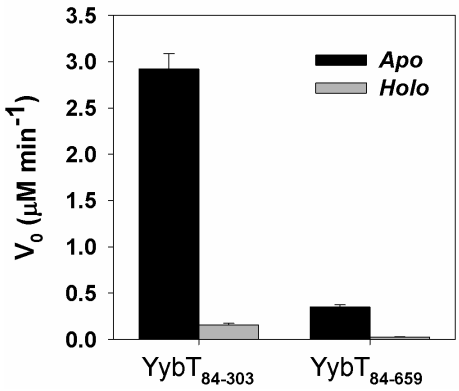


Figure 5.7 Initial velocities of the ATPase actity. The reaction rate of *apo*- and *holo*-forms of *Bs*YybT$_{84\text{-}303}$ and *Bs*YybT$_{84\text{-}659}$ are compared.

5.3.8 Wildtype and ΔyybT mutant survival under stress conditions

To probe the physiological function of heme-sensing in YybT, I compared the stress resistance of wildtype and ΔyybT mutant in the presence or absence of nitrosative, oxidative, cyanide, heme, and CO stress (**Fig. 5.8A**). Except SNP, all compounds affected cell survival to various extents. However, no significant difference was observed between wildtype and mutant, except a ~2.5 fold higher resistance against heme stress for the ΔyybT mutant. In sharp contrast, the ΔYybT mutant had a consistently more than 100-fold lower level of survival in the presence of H_2O_2 (**Fig. 5.8B**). In Bacillus, exogenous and endogenous NO is involved in immediate protection against hydrogen-peroxide (55), oxidative stress generated by macrophage (148), and bactericidal antibiotics that are known to generate reactive oxygen

species (56, 80). I therefore hypothesized that YybT is a downstream target for the protective effect of NO. Similarly to the protective effect observed by Gusarov *et al* (55), exposure to NO for 10 seconds before the addition of H_2O_2 increased the survival of both wildtype and mutant (**Fig. 5.8B**). But the mutant survival rate is still 33-fold lower than wildtype, suggesting that the protective effect of NO is partially lost in the absence of YybT. Indeed, NO retained most of the protective effect in a strain deficient in vegetative catalase, the major downstream target of NO identified by Gusarov et al (55), suggesting the presence of other target(s). In the presence of the NOS inhibitor L-NAME, we observed further decrease in survival for both wildtype and mutant, validating the protective effect of endogenous NO. Interestingly, while the wild-type survival descreased by 20-fold, the YybT mutant survival decreased 2.3-fold (**Fig. 5.8B**). The relatively mild detrimental effect of NOS inhibition in the YybT mutant is consistent with the notion that the protective effect of endogenous NO was partially exerted through YybT.

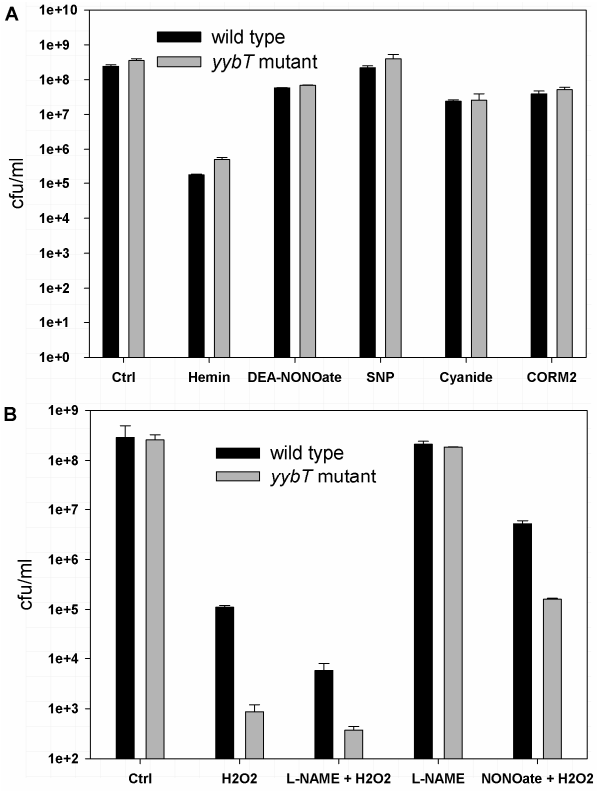


Figure 5.8 Colony forming efficiency of wildtype and *yybT* mutant in the presence of various stress-inducing compounds.

5.4 DISCUSSION

The unusual PAS_{YybT} has only 70-80 residues but retained secondary structure elements. Examination of predicted secondary structure suggested that the C-terminal portion is significantly shortened. Surprisingly, I identified its function as heme-binding, like many other canonical PAS domains. This finding and the subsequent enzymatic assays suggest that heme binding itself or NO binding to heme maybe the physiological signal sensed by the PAS domain in YybT. The NO sensing role is further discussed in conjunction with the in vivo studies.

5.4.1 Heme sensing function of YybT

Heme binding exerts a strong inhibitory effect on YybT activity by reducing the k_{cat}/K_m towards c-di-AMP by more than 200-fold. Such large effect raises the possibility that the PAS domain of YybT may function as a direct heme sensor. Intriguingly, a link between heme intake and YybT proteins had been established by an earlier *in-vivo* study. Burnside *et al* demonstrated that the knockout of *Sa_yybT* (*SA0014*) gene disrupted hemolysin secretion in the pathogenic *S. aureus* (18). Hemolysin is known as a virulence factor or exotoxin for the lysis of host blood cells to obtain nutrients such as heme iron from the host, which would be consistent with the observed function of YybT in *in-vivo* survival (10, 121). The approximately 2.5 fold higher survival rate for the *yybT* mutant in the presence of heme stress also implicates a role for YybT in the heme sensing process **(Fig. 5.8A)**. However, this difference is relatively small compared to the difference in the presence of H_2O_2 **(Fig. 5.8A)**.

5.4.2 Effect of ligand binding on spectra and activity of YybT

Heme-containing PAS domains are well-known sensor domains for perceiving redox signal as well as the gaseous ligand such as O_2 and CO (50). Examples of the redox or O_2-sensing PAS domains include the protein *Ax*PDEA1 from *A. xylinum* (23), *Ec*Dos from *E.*

coli (145)*, and Bj*FixL from *B. japonicum* (23). The heme PAS domain of the prokaryotic transcription factor RocM and the mammalian transcription factor NPAS2 are proposed to function as CO sensors (35, 77). The high-resolution crystal structures of some of the heme-binding PAS domains reveal an invariant histidine coordinating the heme iron centre as the fifth or proximal ligand (53, 89). The use of histidine for heme coordination is also seen in many other non-PAS fold heme-binding signaling protein such as HNOX (117), GAF (26), CRP-FNR (90), Globin (193), SCHIC (110) and HRI-ND (67).

Given the absence of a conserved histidine residue, the observed heme-binding capability of the PAS domain of YybT family proteins is unexpected. The spectral features of ferric and ferrous *Bs*YybT are consistent with a six-coordinated heme. The ferrous form is prone to O_2 oxidation, which prevents the formation of long-lasting oxy-ferrous complex as observed for *Ax*PDEA-1 and some other PAS domains. Coordination of the gaseous ligands NO and CO to the ferrous YybT led to the formation of the Fe^{2+}-NO and Fe^{2+}-CO complexes as evidenced by the changes in absorption spectrum. With the characteristic Soret band at 398 nm, the spectrum of the Fe^{2+}-NO complex is indicative of a penta-coordinated heme that was also observed for several heme-containing proteins for NO sensing that include sGC, DNR, HRI and E75 **(Table 5.2)** (20, 34, 67, 133). The binding of CO causes a ~ 5 nm blue-shift in the Soret band and results in a spectrum that is consistent with a hexa-coordinated heme. The hexa-coordinated Fe^{3+} and Fe^{2+}-CO heme in YybT implies that the heme may be coordinated by one or two endogenous ligands. The identity of the endogenous ligand is unknown at this moment. Observation of the penta-coordinated Fe^{2+}-NO complex suggests that NO-binding triggers the cleavage of the coordination interaction between the iron and the endogenous ligand, similar to the models proposed for sGC and other NO sensor proteins. In addition to the NO-sensing proteins, the transition from hexa- to penta-coordination upon NO-binding is also observed for RrCooA, CLOCK, CytC' and NPAS2 (2, 98, 111, 135) **(Appendix III)**.

The coordination of NO or CN^- to the ferrous heme iron influences the phosphodiesterase activity of the DHH/DHHA1 domain, whereas the effect of CO is negligible. An increase in PDE activity of 3- or 5-fold was observed for NO and CN^- respectively. NO is a signaling molecule in both prokaryotes and eukaryotes and the activity of some proteins is regulated by NO through the NO-sensing domains. Although there has been no report of heme-PAS domains as NO sensors, at least four protein domains using a heme cofactor to sense NO have been characterized (20, 34, 48, 67, 133, 159, 177). Compared to the effect of NO on the PDE activity of YybT, the guanylate cyclase activity of sGC was increased by several hundred fold in the presence of NO (159), the kinase activity of HRI is increased by 5-fold (67), the transcription activation function of DNR is augmented by ~4 fold in a luciferase reported assay (20), and the transcription repressor activity of E75 is significantly reduced in the presence of NO (133). Together with the similar spectral features shared by PAS_{YybT} and NO sensor proteins, the observed regulation of enzymatic activity by NO suggests that the function of NO sensing could be another possible function of the PAS domain.

5.4.2 The involvement of NO in the oxidative stress sensing function of YybT

The cytoplasmic space of the bacterial cell is generally a reducing environment. For instance, in *B subtilis*, the abundance of small thiols such as bacillithiol and cysteine give the cells strong reducing power (41). Nonetheless, many specific oxidative and/or nitrosative stress sensing systems have been identified and studied in detail for bacteria including *B. subtilis* (195). We find that ΔyybT mutant significantly impaired the resistance against H_2O_2 (**Fig. 5.8B**), but not against nitrosative stress (**Fig. 5.8A**), suggesting a previously unknown role for this protein in redox sensing. This seemingly paradoxical finding can be rationalized considering that NO is well known to be involved in hydrogen-peroxide resistance in bacillus (55), and in resisting reactive oxygen species generated from antibiotics or host cells (56,

148). Furthermore, loss of NO synthase rendered the bacteria more susceptible to H_2O_2 (55, 148), validating that endogenous NO played a role in combating H_2O_2 stress. The possibility that YybT is a NO responsive signaling protein is consistent with two observations: (1) the protective effect of NO was less efficient in *yybT mutant*; (2) the detrimental effect of L-NAME was weaker in ΔyybT mutant (**Fig. 5.8B**). Since YybT is amenable to redox change, it may be oxidized in the presence H_2O_2. Oxidized ferric-YybT does not respond to NO, which may therefore shut off the NO signaling pathway mediated by YybT. On the other hand, the tight NO binding may prevent YybT from oxidization by oxidative stress. In line with this, the protective effect of NO against H_2O_2 stress observed by Nudler and colleagues is only effective when NO is applied first (55).

In contrast to the susceptibility to H_2O_2, we and Rallu *et al* previously observed a survival advantage of YybT mutant towards acid stress in both *B. subtilis* and *L. lactis* (126, 131). That observation was consistent with the findings that the stress alarmone ppGpp competitively inhibits the PDE activity of YybT (131), and that the rise of (p)ppGpp level confers *L. lactis* stronger acid resistance (126). The observed correlation between low PDE activity and acid resistance implied that acid resistance induced by high (p)ppGpp level may be due to, at least in part, regulation of the PDE activity of YybT. Interestingly, H_2O_2 decreased ppGpp concentration in the YybT-containing firmicute pathogen *Enterococcus faecalis* (189), which could be another route to increase YybT activity in response of H_2O_2. Take together, it seems that low YybT PDE activity correlates with better acid resistance, whereas high YybT PDE activity correlates with better H_2O_2 resistance. Notably, NO and ppGpp regulate the PDE activity of YybT through different mechanisms (allosteric vs competitive) and towards different directions (activation vs inhibition), suggesting that the concentration of YybT PDE substrate needs to be carefully modulated to generate different response towards different stress, a feature reminiscent of a second messenger.

5.5 SUMMARY

The YybT family proteins (COG3387) in *Bacillus* and other Firmicutes contain a predicted Per-ARNT-Sim (PAS) domain with unknown function. In this chapter, I show that two representative members of the YybT family bind heme unexpectedly through the PAS domain despite the absence of any obvious endogenous ligand for heme coordination. UV-Vis spectroscopic characterization of the YybT proteins revealed that the six-coordinated ferrous heme form is prone to oxidation by oxygen. The ferrous heme iron of YybT proteins forms a six-coordinated complex with carbon monoxide, but a penta-coordinated complex with nitric oxide as indicated by the characteristic Soret band at 398 nm. Enzymatic assays showed that the binding of the heme suppresses the phosphodiesterase activity of the C-terminal DHH/DHHA1 domain by reducing the k_{cat}/K_m by more than 200-fold. Binding of NO or cyanide to the ferrous or ferric YybT proteins stimulate the PDE activity by 3- and 5-fold, respectively, whereas the change of redox state or CO binding did not seem to affect the PDE activity. *In-vivo* studies showed that the *B. subtilis* Δ*yybT* mutant exhibits more than 100-fold lower survival rate when challenged with H_2O_2 stress, while NO is known to play an important role in combating H_2O_2 mediated cell killing. Taken together, the results revealed a distinct heme-binding PAS domain and suggest that the PAS domain may function as regulatory domain by sensing heme or NO in stress signaling.

CONCLUDING REMARKS

The cyclic di-nucleotides are emerging as important signaling molecules in prokaryotes, in paralell to the role of cyclic mono-nucleotides in eukaryotes. While the presence of c-di-GMP network is ubiquitous and its messenger role is firmly established by now, the role of c-di-AMP is still largely unexplored. My work on RocR and other EAL domain phosphodiesterases represents the first effort to elucidate the structure-function relationship in c-di-GMP specific phosphodiesterases. The work has successfully identified essential residues for catalysis, as well as the involvement of a conserved loop and domain interface in regulation. Establishment of the detailed catalytic mechanism for RocR requires crystal structures, which is now being actively pursued through a collaboration work with a structural biology group. The recent discovery of c-di-AMP led us to identify a family of proteins that contain a potential c-di-AMP specific phosphodiesterase domain. Our *in-vitro* assay results suggest that the DHH/DHHA1 domain of the YybT family proteins could be a c-di-AMP phosphodiesterase. YybT also happens to contain a heme-PAS sensor domain, with the activity of the DHH/DHHA1 domain modulated by the binding of heme or coordination of nitric oxide (NO) to the heme iron. Although the biological functions of the domains remain to be fully established by *in-vivo* studies, the stress signaling role of the YybT proteins is evidenced by the comparison of phenotypes of the wild type and mutant strains. The findings raise the exciting possibility that YybT may be a drug target for designing antibacterial drugs to combat pathogens such as *Staphylococcus aureus* and *Listeria monocytogenes*.

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. **Andrew, C. R., E. L. Green, D. M. Lawson, and R. R. Eady.** 2001. Resonance Raman studies of cytochrome c' support the binding of NO and CO to opposite sides of the heme: implications for ligand discrimination in heme-based sensors. Biochemistry **40:**4115-22.
3. **Aravind, L., and E. V. Koonin.** 1998. A novel family of predicted phosphoesterases includes Drosophila prune protein and bacterial RecJ exonuclease. Trends Biochem. Sci. **23:**17-9.
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. **Ausmees, N., R. Mayer, H. Weinhouse, G. Volman, D. Amikam, M. Benziman, and M. Lindberg.** 2001. Genetic data indicate that proteins containing the GGDEF domain possess diguanylate cyclase activity. FEMS Microbiol Lett **204:**163-7.
6. **Barbe, V., S. Cruveiller, F. Kunst, P. Lenoble, G. Meurice, A. Sekowska, D. Vallenet, T. Wang, I. Moszer, C. Medigue, and A. Danchin.** 2009. From a consortium sequence to a unified sequence: the Bacillus subtilis 168 reference genome a decade later. Microbiology **155:**1758-75.
7. **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.
8. **Bartalesi, I., I. Bertini, K. Ghosh, A. Rosato, and P. Turano.** 2002. The unfolding of oxidized c-type cytochromes: the instructive case of Bacillus pasteurii. J Mol Biol **321:**693-701.
9. **Bejerano-Sagie, M., Y. Oppenheimer-Shaanan, I. Berlatzky, A. Rouvinski, M. Meyerovich, and S. Ben-Yehuda.** 2006. A checkpoint protein that scans the chromosome for damage at the start of sporulation in *Bacillus subtilis*. Cell **125:**679-690.
10. **Benton, B. M., J. P. Zhang, S. Bond, C. Pope, T. Christian, L. Lee, K. M. Winterberg, M. B. Schmid, and J. M. Buysse.** 2004. Large-scale identification of genes required for full virulence of Staphylococcus aureus. J Bacteriol **186:**8478-89.
11. **Berry, E. A., and B. L. Trumpower.** 1987. Simultaneous determination of hemes a, b, and c from pyridine hemochrome spectra. Anal Biochem **161:**1-15.
12. **Bobrov, A. G., O. BKirillina, and R. D. Perry.** 2005. The phosphodiesterase activity of the HmsP EALdomain is required for nagativr regulation of biofilm formation in Yersinia pestis. FEMS Microbiol. Lett. **247:**123-130.
13. **Bobrov, A. G., V. Kirillina, and R. D. Perry.** 2007. Regulation of biofilm formation in *Yersinia Pestis*. Adv. Exp. Med. Biol. **603:**201-210.
14. **Boehm, A., S. Steiner, F. Zaehringer, A. Casanova, F. Hamburger, D. Ritz, W. Keck, M. Ackermann, T. Schirmer, and U. Jenal.** 2009. Second messenger signaling governs Escherichia coli biofilm induction upon ribosomal stress. Mol. Microbiol. **72:**1500-1516.
15. **Bourret, R. B.** 2010. Receiver domain structure and function in response regulator proteins. Curr Opin Microbiol **13:**142-9.
16. **Brzovic, P. S., C. C. Hyde, E. W. Miles, and M. F. Dunn.** 1993. Characterization of the functional role of a flexible loop in the .alpha.-subunit of tryptophan synthase from *Salmonella typhimurium* by rapid-scanning, stopped-flow spectroscopy and site-directed mutagenesis. Biochemistry **32:**10404-10413.

17. **Brzovic, P. S., Y. Sawa, C. C. Hyde, E. W. Miles, and M. F. Dunn.** 1992. Evidence that mutations in a loop region of the alpha-subunit inhibit the transition from an open to a closed conformation in the tryptophan synthase bienzyme complex. J. Biol. Chem. **267:**13028-13038.
18. **Burnside, K., A. Lembo, M. de Los Reyes, A. Iliuk, N. T. Binhtran, J. E. Connelly, W. J. Lin, B. Z. Schmidt, A. R. Richardson, F. C. Fang, W. A. Tao, and L. Rajagopal.** 2010. Regulation of hemolysin expression and virulence of Staphylococcus aureus by a serine/threonine kinase and phosphatase. PLoS One **5:**e11071.
19. **Camilli, A., and B. L. Bassler.** 2006. Bacterial small-molecule signaling pathways. Science **311:**1113-1116.
20. **Castiglione, N., S. Rinaldo, G. Giardina, and F. Cutruzzola.** 2009. The transcription factor DNR from Pseudomonas aeruginosa specifically requires nitric oxide and haem for the activation of a target promoter in Escherichia coli. Microbiology **155:**2838-44.
21. **Cha, J., and D. S. Auld.** 1997. Site-directed mutagenesis of the active site glutamate in human matrilysin: investigation of its role in catalysis. Biochemistry **36:**16019-16024.
22. **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.
23. **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.
24. **Chatterji, D., and A. Kumar Ojha.** 2001. Revisiting the stringent response, ppGpp and starvation signaling. Curr. Opin. Microbiol. **4:**160-165.
25. **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.
26. **Cho, H. Y., H. J. Cho, Y. M. Kim, J. I. Oh, and B. S. Kang.** 2009. Structural insight into the heme-based redox sensing by DosS from Mycobacterium tuberculosis. J Biol Chem **284:**13057-67.
27. **Choy, W. K., L. Zhou, C. K. 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:**7221-8.
28. **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.
29. **Cotter, P. D., and C. Hill.** 2003. Surviving the acid test: responses of gram-positive bacteria to low pH. Microbiol. Mol. Biol. Rev. **67:**429-453.
30. **D'Angelo, A., L. Garzia, A. Andre, P. Carotenuto, V. Aglio, O. Guardiola, G. Arrigoni, A. Cossu, G. Palmieri, L. Aravind, and M. Zollo.** 2004. Prune cAMP phosphodiesterase binds nm23-H1 and promotes cancer metastasis. Cancer Cell **5:**137-49.
31. **Dai, Z., and E. M. Boon.** 2010. Engineering of the heme pocket of an H-NOX domain for direct cyanide detection and quantification. J Am Chem Soc **132:**11496-503.
32. **De, N.** 2008. Phosphorylation-independent regulation of the diguanylate cyclase WspR. PLoS Biol. **6:**e67.
33. **De, N., M. Pirruccello, P. V. Krasteva, N. Bae, R. V. Raghavan, and H. Sondermann.** 2008. Phosphorylation-independent regulation of the diguanylate cyclcase WspR. PLoS Biology **6:**e67.
34. **Derbyshire, E. R., and M. A. Marletta.** 2009. Biochemistry of soluble guanylate cyclase. Handb Exp Pharmacol**:**17-31.

35. **Dioum, E. M., J. Rutter, J. R. Tuckerman, G. Gonzalez, M. A. Gilles-Gonzalez, and S. L. McKnight.** 2002. NPAS2: a gas-responsive transcription factor. Science **298:**2385-7.
36. **Dolinsky, T. J., J. E. Neilsen, J. A. MacCammon, and N. A. Baker.** 2004. PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations Nucl. Acids Res. **32:**665-667.
37. **Duerig, A.** 2009. Second messenger-mediated spatiotemporal control of protein degradation regulates bacterial cell cycle progression. Genes Dev. **23:**93-104.
38. **Etzkorn, M., H. Kneuper, P. Dunnwald, V. Vijayan, J. Kramer, C. Griesinger, S. Becker, G. Unden, and M. Baldus.** 2008. Plasticity of the PAS domain and a potential role for signal transduction in the histidine kinase DcuS. Nat Struct Mol Biol **15:**1031-9.
39. **Fabrichniy, I. P., L. Lehtio, A. Salminen, A. B. Zyryanov, A. A. Baykov, R. Lahti, and A. Goldman.** 2004. Structural studies of metal ions in family II pyrophosphatases: the requirement for a Janus ion. Biochemistry **43:**14403-11.
40. **Fabrichniy, I. P., L. Lehtio, M. Tammenkoski, A. B. Zyryanov, E. Oksanen, A. A. Baykov, R. Lahti, and A. Goldman.** 2007. A trimetal site and substrate distortion in a family II inorganic pyrophosphatase. J. Biol. Chem. **282:**1422-31.
41. **Gaballa, A., G. L. Newton, H. Antelmann, D. Parsonage, H. Upton, M. Rawat, A. Claiborne, R. C. Fahey, and J. D. Helmann.** 2010. Biosynthesis and functions of bacillithiol, a major low-molecular-weight thiol in Bacilli. Proc Natl Acad Sci U S A **107:**6482-6.
42. **Galperin, M. Y.** 2005. A census of membrane-bound and intracellular signal transduction proteins in bacteria: Bacterial IQ, extroverts and introverts. BMC Microbiol. **5**.
43. **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.
44. **Gao, R., and A. M. Stock.** 2010. Molecular strategies for phosphorylation-mediated regulation of response regulator activity. Curr Opin Microbiol **13:**160-7.
45. **Garcia, B.** 2004. Role of the GGDEF protein family in Salmonella cellulose biosynthesis and biofilm formation. Mol. Microbiol. **54:**264-277.
46. **Garcia, B., C. Latasa, C. Solano, F. G. Portillo, C. Gamazo, and I. Lasa.** 2004. Role of the GGDEF protein family in *Salmonella* cellulose biosynthesis and biofilm formation. Mol. Microbiol. **54:**264-277.
47. **Garrett, G. M., D. S. Goodsell, R. S. Halliday, R. Huey, W. E. Hart, R. K. Belew, and A. J. Olson.** 1998. Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function J. Comput. Chem. **19:**1639-1662.
48. **Giardina, G., S. Rinaldo, K. A. Johnson, A. Di Matteo, M. Brunori, and F. Cutruzzola.** 2008. NO sensing in Pseudomonas aeruginosa: structure of the transcriptional regulator DNR. J Mol Biol **378:**1002-15.
49. **Gilles-Gonzalez, M.-A., and G. gonzalez.** 2004. Singal transduction by heme-containing PAS-domain proteins. J. Appl. Physiol. **96:**774-783.
50. **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.
51. **Gilles-Gonzalez, M. A., G. Gonzalez, M. F. Perutz, L. Kiger, M. C. Marden, and C. Poyart.** 1994. Heme-based sensors, exemplified by the kinase FixL, are a new class of heme protein with distinctive ligand binding and autoxidation. Biochemistry **33:**8067-73.
52. **Girgris, H. S., Y. Liu, W. S. Ryu, and S. Tavazoie.** 2007. A comprehensive genetic characterization of bacterial motility. PLOS Genet. **3:**1644-1660.

53. **Gong, W., B. Hao, S. S. Mansy, G. Gonzalez, M. A. Gilles-Gonzalez, and M. K. Chan.** 1998. Structure of a biological oxygen sensor: a new mechanism for heme-driven signal transduction. Proc Natl Acad Sci U S A **95:**15177-82.
54. **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.
55. **Gusarov, I., and E. Nudler.** 2005. NO-mediated cytoprotection: instant adaptation to oxidative stress in bacteria. Proc Natl Acad Sci U S A **102:**13855-60.
56. **Gusarov, I., K. Shatalin, M. Starodubtseva, and E. Nudler.** 2009. Endogenous nitric oxide protects bacteria against a wide spectrum of antibiotics. Science **325:**1380-4.
57. **Haas, D., and G. Defago.** 2005. Biological control of soil-borne pathogens by fluorescent pseudomonads. Nat Rev Microbiol **3:**307-19.
58. **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.
59. **Hengge, R.** 2009. Principles of c-di-GMP signalling in bacteria. Nat. Rev. Micro. **7:**263-273.
60. **Hickman, J. W., and C. S. Harwood.** 2008. Identification of FleQ from *Pseudomonas aeruginosa* as a c-di-GMP-responsive transcription factor. Mol. Microbiol. **69**.
61. **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.
62. **Hoff, W. D., P. Dux, K. Hard, B. Devreese, I. M. Nugteren-Roodzant, W. Crielaard, R. Boelens, R. Kaptein, J. van Beeumen, and K. J. Hellingwerf.** 1994. Thiol ester-linked p-coumaric acid as a new photoactive prosthetic group in a protein with rhodopsin-like photochemistry. Biochemistry **33:**13959-62.
63. **Hoffman, E. C., H. Reyes, F. F. Chu, F. Sander, L. H. Conley, B. A. Brooks, and O. Hankinson.** 1991. Cloning of a factor required for activity of the Ah (dioxin) receptor. Science **252:**954-8.
64. **Holland, L. M., S. T. O'Donnell, D. A. Ryjenkov, L. Gomelsky, S. R. Slater, P. D. Fey, M. Gomelsky, and J. P. O'Gara.** 2008. A *Staphylococcal* GGDEF domain protein regulates biofilm formation independently of cyclic dimeric GMP. J. Bacteriol. **190:**5178-5189.
65. **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.
66. **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.
67. **Igarashi, J., A. Sato, T. Kitagawa, T. Yoshimura, S. Yamauchi, I. Sagami, and T. Shimizu.** 2004. Activation of heme-regulated eukaryotic initiation factor 2alpha kinase by nitric oxide is induced by the formation of a five-coordinate NO-heme complex: optical absorption, electron spin resonance, and resonance raman spectral studies. J Biol Chem **279:**15752-62.
68. **Ikehara, K., H. Okada, K. Maeda, A. Ogura, and K. Sugae.** 1984. Accumulation of *relA* gene-independent ppGpp in *Bacillus subtillis* vegetative cells upon temperature shift-down. . J. Biochem. **95:**895-897.
69. **Jenal, U.** 2004. Cyclic di-guanosine-monophosphate comes of age: a novel secondary messenger involved in modulating cell surface structures in bacteria? Current Opinion in Microbiology **7:**185-191.

70. **Jenal, U., and J. Malone.** 2006. Mechanisms of cyclic-di-GMP signaling in bacteria. Annu. Rev. Genet. **40:**385-407.
71. **Jensen, L. J., M. Kuhn, M. Stark, S. Chaffron, C. Creevey, J. Muller, T. Doerks, P. Julien, A. Roth, M. Simonovic, P. Bork, and C. von Mering.** 2009. STRING 8--a global view on proteins and their functional interactions in 630 organisms, p. D412-416, vol. 37.
72. **Jeong, K. C., K. F. Hung, D. J. Baumler, J. J. Byrd, and C. W. Kaspar.** 2008. Acid stress damage of DNA is prevented by Dps binding in *Escherichia col*i O 157:H7. BMC Biochemistry **8:**181.
73. **Jimin Pei, N. V. G.** 2001. GGDEF domain is homologous to adenylyl cyclase. Proteins: Structure, Function, and Genetics **42:**210-216.
74. **Joseph, D., G. A. Petsko, and M. Karplus.** 1990. Anatomy of a conformational change: hinged "lid" motion of the triosephosphate isomerase loop. Science **249:**1425-1428.
75. **Karatan, E., D. T. R., and P. I. Watnick.** 2005. NspS, a predicted polyamine sensor, mediates activation of *Vibrio cholerae* biofilm formation by norspermidine. J. Bacteriol. **187:**7434-7443.
76. **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.
77. **Kerby, R. L., H. Youn, and G. P. Roberts.** 2008. RcoM: a new single-component transcriptional regulator of CO metabolism in bacteria. J Bacteriol **190:**3336-43.
78. **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.
79. **Kobayashi, K. e. a.** 2003. Essential Bacillussubtilis genes. Proc. Natl Acad. Sci. USA **100:**4678-4683.
80. **Kohanski, M. A., D. J. Dwyer, B. Hayete, C. A. Lawrence, and J. J. Collins.** 2007. A common mechanism of cellular death induced by bactericidal antibiotics. Cell **130:**797-810.
81. **Kolb, A., S. Busby, H. Buc, S. Garges, and S. Adhya.** 1993. Transcriptional regulation by cAMP and its receptor protein. Annu Rev Biochem **62:**749-95.
82. **Krasteva, P. V., J. C. Fong, N. J. Shikuma, S. Beyhan, M. V. Navarro, F. H. Yildiz, and H. Sondermann.** 2010. Vibrio cholerae VpsT regulates matrix production and motility by directly sensing cyclic di-GMP. Science **327:**866-8.
83. **Kuchma, S., J. P. Connolly, and G. A. O'Toole.** 2005. A three-component regulatory system regulates biofilm maturation and type III secretion in *Pseudomanas aeruginosa*. J. Bacteriol. **187:**1441-1454.
84. **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.
85. **Kuchma, S. L., K. M. Brothers, J. H. Merritt, N. T. Liberati, F. M. Ausubel, and G. A. O'Toole.** 2007. BifA, a c-di-GMP phosphodiesterase, Inversely Regulates Biofilm Formation and Swarming Motility by *Pseudomonas aeruginosa* PA14. J. Bacteriol. **189:**8165-8178.
86. **Kulasakara, H., V. Lee, A. Brencic, N. Liberati, J. Urbach, S. Miyata, D. G. Lee, A. N. Neely, M. Hyodo, Y. Hayakawa, F. M. Ausubel, and S. Lory.** 2006. Analysis of *Pseudomonas aeruginosa* diguanylate cyclases and phosphodiesterases reveals a role for bis-(3'-5')-cyclic-GMP in virulence. Proc. Natl. Acad. Sci. USA **103:**2839-2844.

87. **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.
88. **Kumar, A., J. C. Toledo, R. P. Patel, J. R. Lancaster, Jr., and A. J. Steyn.** 2007. Mycobacterium tuberculosis DosS is a redox sensor and DosT is a hypoxia sensor. Proc Natl Acad Sci U S A **104:**11568-73.
89. **Kurokawa, H., D. S. Lee, M. Watanabe, I. Sagami, B. Mikami, C. S. Raman, and T. Shimizu.** 2004. A redox-controlled molecular switch revealed by the crystal structure of a bacterial heme PAS sensor. J Biol Chem **279:**20186-93.
90. **Lanzilotta, W. N., D. J. Schuller, M. V. Thorsteinsson, R. L. Kerby, G. P. Roberts, and T. L. Poulos.** 2000. Structure of the CO sensing transcription activator CooA. Nat Struct Biol **7:**876-80.
91. **Larkin, M. A., G. Blackshields, N. P. Brown, R. Chenna, P. A. McGettigan, H. McWilliam, F. Valentin, I. M. Wallace, A. Wilm, R. Lopez, J. D. Thompson, T. J. Gibson, and D. G. Higgins.** 2007. Clustal W and Clustal X version 2.0. Bioinformatics **23:**2947-2948.
92. **Lee, E. R., J. L. Baker, Z. Weinberg, N. Sudarsan, and R. R. Breaker.** 2010. An allosteric self-splicing ribozyme triggered by a bacterial second messenger. Science **329:**845-8.
93. **Lee, S. Y., H. S. Cho, J. G. Pelton, D. L. Yan, E. A. Berry, and D. E. Wemmer.** 2001. Crystal structure of activated CheY - Comparison with other activated receiver domains. J. Biol. Chem. **276:**16425-16431.
94. **Li, H., A. D. Robertson, and J. H. Jensen.** 2005. Very fast empirical prediction and rationalization of protein pKa values. Protein Sci: Struct. Func. Bioinfo. **61:**704-721.
95. **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.
96. **Liang, Z. X., I. Tsigos, T. Lee, V. Bouriotis, K. A. Resing, N. G. Ahn, and J. P. Klinman.** 2004. Evidence for increased local flexibility in psychrophilic alcohol dehydrogenase relative to its thermophilic homologue. Biochemistry **43:**14676-14683.
97. **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.
98. **Lukat-Rodgers, G. S., C. Correia, M. V. Botuyan, G. Mer, and K. R. Rodgers.** 2010. Heme-Based Sensing by the Mammalian Circadian Protein CLOCK. Inorg Chem **49:**6349-65.
99. **Lundqvist, T., and G. Schneider.** 1991. Crystal structure of activated ribulose-1,5-bisphosphate carboxylase complexed with its substrate, ribulose-1,5-bisphosphate. J. Biol. Chem. **266:**12604-12611.
100. **Marsh, F. J., P. Weiner, J. E. Douglas, P. A. Kollman, G. L. Kenyon, and J. A. Gerlt.** 1980. Theoretical calculations of hydrolysis energies of "high-energy" molecules. 3. Theoretical calculations on the geometric destabilization of 3',5'- and 2',3'-cyclic nucleotides. Journal of the American Chemical Society **102:**1660-1665.
101. **Marvin, K. A., J. L. Reinking, A. J. Lee, K. Pardee, H. M. Krause, and J. N. Burstyn.** 2009. Nuclear receptors homo sapiens Rev-erbbeta and Drosophila melanogaster E75 are thiolate-ligated heme proteins which undergo redox-mediated ligand switching and bind CO and NO. Biochemistry **48:**7056-71.
102. **McMillan, F. M., M. Cahoon, A. White, L. Hedstrom, G. A. Petsko, and D. Ringe.** 2000. Crystal structure at 2.4 A resolution of *Borrelia burgdorferi* inosine 5'-

monophosphate dehydrogenase: evidence of a substrate-induced hinged-lid motion by loop 6. Biochemistry **39:**4533-4542.

103. **Mechold, U., G. Fang, S. Ngo, V. Ogryzko, and A. Danchin.** 2007. YtqI from Bacillus subtilis has both oligoribonuclease and pAp-phosphatase activity. Nucleic Acids Res. **35:**4552-61.
104. **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.
105. **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.
106. **Mitic, N., S. J. Smith, A. Neves, L. W. Guddat, L. R. Gahan, and G. Schenk.** 2006. The catalytic mechanisms of binuclear metallohydrolases. Chem. Rev. **106:**3338-3363.
107. **Moglich, A., R. A. Ayers, and K. Moffat.** 2009. Structure and signaling mechanism of Per-ARNT-Sim domains. Structure **17:**1282-94.
108. **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:**112-26.
109. **Moore, C. M., M. M. Nakano, T. Wang, R. W. Ye, and J. D. Helmann.** 2004. Response of Bacillus subtilis to nitric oxide and the nitrosating agent sodium nitroprusside. J Bacteriol **186:**4655-64.
110. **Moskvin, O. V., S. Kaplan, M. A. Gilles-Gonzalez, and M. Gomelsky.** 2007. Novel heme-based oxygen sensor with a revealing evolutionary history. J Biol Chem **282:**28740-8.
111. **Mukaiyama, Y., T. Uchida, E. Sato, A. Sasaki, Y. Sato, J. Igarashi, H. Kurokawa, I. Sagami, T. Kitagawa, and T. Shimizu.** 2006. Spectroscopic and DNA-binding characterization of the isolated heme-bound basic helix-loop-helix-PAS-A domain of neuronal PAS protein 2 (NPAS2), a transcription activator protein associated with circadian rhythms. FEBS J **273:**2528-39.
112. **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.
113. **Patel, S., L. Yenush, P. L. Rodriguez, R. Serrano, and T. L. Blundell.** 2002. Crystal structure of an enzyme displaying both inositol-polyphosphate-1-phosphatase and 3'-phosphoadenosine-5'-phosphate phosphatase activities: a novel target of lithium therapy. J Mol Biol **315:**677-85.
114. **Paul, R., S. Abel, P. Wassmann, A. Beck, H. Heerklotz, and U. Jenal.** 2007. Activation of the diguanylate cyclase pleD by phosphorylation-mediated dimerization. J. Biol. Chem. **282:**29170-29177.
115. **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 Dev **18:**715-27.
116. **Pei, J., and N. V. Grishin.** 2001. GGDEF domain is homologous to adenylyl cyclase. Proteins **42:**210-6.
117. **Pellicena, P., D. S. Karow, E. M. Boon, M. A. Marletta, and J. Kuriyan.** 2004. Crystal structure of an oxygen-binding heme domain related to soluble guanylate cyclases. Proceedings of the National Academy of Sciences of the United States of America **101:**12854-12859.

118. **Pesavento, C., and R. Hengge.** 2009. Bacterial nucleotide-based second messengers. Curr. Opin. Microbiol. **12:**170-176.
119. **Plateau, P., M. Fromant, F. Kepes, and S. Blanquet.** 1987. Intracellular 5',5'-dinucleoside polyphosphate levels remain constant during the *Escherichia coli* cell cycle. J. Bacteriol. **169:**419-422.
120. **Pompliano, D. L., A. Peyman, and J. R. Knowles.** 1990. Stabilization of a reaction intermediate as a catalytic device: definition of the functional role of the flexible loop in triosephosphate isomerase. Biochemistry **29:**3186-3194.
121. **POPE RC, B. B., BOND S, BLAIS J, GALAZZO JL, LEE M, BUYSSE J;.** 2001. A Compound Synthetic Lethality Screen to Identify Inhibitors of Conditional Essential Bacterial Gene Products. Interscience Conference on Antimicrobial Agents and Chemotherapy **41** abstract no. F-2126.
122. **Potrykus, K., and M. Cashel.** 2008. (p)ppGpp: still magical? Annu Rev Microbiol **62:**35-51.
123. **Potrykus, K., and M. Cashel.** 2008. (p)ppGpp: still magical? Annu. Rev. Microbiol. **62:**35-51.
124. **Priestle, J. P., M. G. GrÃ¼tter, J. L. White, M. G. Vincent, M. Kania, E. Wilson, T. S. Jardetzky, K. Kirschner, and J. N. Jansonius.** 1987. Three-dimensional structure of the bifunctional enzyme N-(5'-phosphoribosyl)anthranilate isomerase-indole-3-glycerol-phosphate synthase from *Escherichia coli*. Proc. Natl. Acad. Sci. USA **84:**5690-5694.
125. **Qi, Y., F. Rao, Z. Luo, and Z. X. Liang.** 2009. A flavin cofactor-binding PAS domain regulates c-di-GMP synthesis in AxDGC2 from Acetobacter xylinum. Biochemistry **48:**10275-85.
126. **Rallu, F., A. Gruss, S. D. Ehrilich, and E. Maguin.** 2000. Acid- and multistress-resistant mutants of *Lactococcus lactis*: identification of intracellular stress signals. Mol. Microbiol. **35:**517-528.
127. **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.
128. **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.
129. **Rao, F., Y. Qi, E. Murugan, S. Pasunooti, and Q. Ji.** 2010. 2',3'-cAMP hydrolysis by metal-dependent phosphodiesterases containing DHH, EAL, and HD domains is non-specific: Implications for PDE screening. Biochem Biophys Res Commun **398:**500-5.
130. **Rao, F., Y. Qi, E. Murugan, S. Pasunooti, and Q. Ji.** 2010. 2',3'-cAMP hydrolysis by metal-dependent phosphodiesterases containing DHH, EAL, and HD domains is non-specific: Implications for PDE screening. Biochem Biophys Res Commun.
131. **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. J Biol Chem **285:**473-82.
132. **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.
133. **Reinking, J., M. M. Lam, K. Pardee, H. M. Sampson, S. Liu, P. Yang, S. Williams, W. White, G. Lajoie, A. Edwards, and H. M. Krause.** 2005. The Drosophila nuclear receptor e75 contains heme and is gas responsive. Cell **122:**195-207.
134. **Resing, K. A., A. N. Hoofnagle, and N. G. Ahn.** 1999. Modeling deuterium exchange behavoir of ERK2 using pepsin mapping to probe seconardy structure. Am. Soc. Mass. Spec. **10:**685-702.

135. **Reynolds, M. F., R. B. Parks, J. N. Burstyn, D. Shelver, M. V. Thorsteinsson, R. L. Kerby, G. P. Roberts, K. M. Vogel, and T. G. Spiro.** 2000. Electronic absorption, EPR, and resonance raman spectroscopy of CooA, a CO-sensing transcription activator from R. rubrum, reveals a five-coordinate NO-heme. Biochemistry **39:**388-96.
136. **Romling, U.** 2008. Great times for small molecules: c-di-AMP, a second messenger candidate in Bacteria and Archaea. Sci. Signal. **1:**pe39.
137. **Romling, U., and D. Amikam.** 2006. Cyclic di-GMP as a second messenger. Curr. Opion. Microbiol. **9:**218-228.
138. **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.
139. **Ross, P.** 1987. Regulation of cellulose synthesis in Acetobacter xylinum by cyclic diguanylate. Nature **325:**279-281.
140. **Ross, P., R. Mayer, H. Weinhouse, D. Amikam, Y. Huggirat, M. Benziman, E. de Vroom, A. Fidder, 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.
141. **Ruer, S., S. Stender, A. Filloux, and S. de Bentzmann.** 2007. Assembly of Fimbrial Structures in Pseudomonas aeruginosa: Functionality and Specificity of Chaperone-Usher Machineries. J. Bacteriol. **189:**3547-3555.
142. **Ruiz, A., C. Hurtado, J. M. Ribeiro, A. Sillero, and M. A. Gunther Sillero.** 1989. Hydrolysis of bis(5'-nucleosidyl) polyphosphates by Escherichia coli 5'-nucleotidase. J. Bacteriol. **171:**6703-6709.
143. **Ryan, R. P., Y. Fouhy, J. F. Lucey, L. C. Crossman, S. Spiro, Y.-W. He, L.-H. Zhang, S. Heeb, M. Camara, P. Williams, and J. M. Dow.** 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.
144. **Ryjenkov, D. A., M. Tarutina, O. V. Moskvin, 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.
145. **Sasakura, Y.** 2002. Characterization of a direct oxygen sensor heme proein from Escherichia coli. Effects of the heme redox states and mutations at the heme-binding site on catalysis and structure. J. Biol. Chem. **277:**23821-23827.
146. **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.
147. **Sevvana, M., V. Vijayan, M. Zweckstetter, S. Reinelt, D. R. Madden, R. Herbst-Irmer, G. M. Sheldrick, M. Bott, C. Griesinger, and S. Becker.** 2008. A ligand-induced switch in the periplasmic domain of sensor histidine kinase CitA. J Mol Biol **377:**512-23.
148. **Shatalin, K., I. Gusarov, E. Avetissova, Y. Shatalina, L. E. McQuade, S. J. Lippard, and E. Nudler.** 2008. Bacillus anthracis-derived nitric oxide is essential for pathogen virulence and survival in macrophages. Proc Natl Acad Sci U S A **105:**1009-13.
149. **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.
150. **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.

151. **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.
152. **Simm, R., M. Morr, U. Remminghorst, M. Andersson, and U. Römling.** 2009. Quantitative determination of cyclic diguanosine monophosphate concentrations in nucleotide extracts of bacteria by matrix-assisted laser desorption/ionization-time-of-flight mass spectrometry. Anal. Biochem. **386:**53-58.
153. **Sinha, S. C., and S. R. Sprang.** 2007. Structures, mechanism, regulation and evolution of class III mucleotidyl cycases. Rev. Physiol. Biochem. Pharmacol. **157:**105-140.
154. **Smulevich, G., F. Neri, M. P. Marzocchi, and K. G. Welinder.** 1996. Versatility of heme coordination demonstrated in a fungal peroxidase. Absorption and resonance Raman studies of Coprinus cinereus peroxidase and the Asp245-->Asn mutant at various pH values. Biochemistry **35:**10576-85.
155. **Srivatsan, A., and J. D. Wang.** 2008. Control of bacterial transcription, translation and replication. Curr. Opin. Microbiol. **11:**100-105.
156. **Stanford, N. P., S. E. Halford, and G. S. Baldwin.** 1999. DNA cleavage by the EcoRV restriction endonuclease: pH dependence and proton transfers in catalysis. J. Mol. Biol. **288:**105-116.
157. **Stauff, D. L., and E. P. Skaar.** 2009. Bacillus anthracis HssRS signaling to HrtAB regulates heme resistance during infection. Mol Microbiol.
158. **Sterner, R., and B. Hocker.** 2005. Catalytic versatility, stability and evolution of the $(\beta/\alpha)_8$ -barrel enzyme fold. Chem. Rev. **105:**4038-4055.
159. **Stone, J. R., and M. A. Marletta.** 1994. Soluble guanylate cyclase from bovine lung: activation with nitric oxide and carbon monoxide and spectral characterization of the ferrous and ferric states. Biochemistry **33:**5636-40.
160. **Stone, J. R., R. H. Sands, W. R. Dunham, and M. A. Marletta.** 1996. Spectral and ligand-binding properties of an unusual hemoprotein, the ferric form of soluble guanylate cyclase. Biochemistry **35:**3258-62.
161. **Sudarsan, N.** 2008. Riboswitches in eubacteria sense the second messenger cyclic di-GMP. Science **321:**411-413.
162. **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.
163. **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.
164. **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.
165. **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.
166. **Tammenkoski, M., V. M. Moiseev, M. Lahti, E. Ugochukwu, T. H. Brondijk, S. A. White, R. Lahti, and A. A. Baykov.** 2007. Kinetic and mutational analyses of the major cytosolic exopolyphosphatase from Saccharomyces cerevisiae. J. Biol. Chem. **282:**9302-11.
167. **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.

168. **Tao, F., Y. W. He, D. H. Wu, S. Swarup, and L. H. Zhang.** 2010. The cyclic nucleotide monophosphate domain of Xanthomonas campestris global regulator Clp defines a new class of cyclic di-GMP effectors. J Bacteriol **192:**1020-9.
169. **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.
170. **Taylor, B. L., and I. B. Zhulin.** 1999. PAS domains: internal sensors of oxygen, redox potential, and light. Microbiol. Mol. Biol. Rev. **63:**479-506.
171. **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.** 2010. Structural insight into the mechanism of cyclic di-GMP hydrolysis by EAL domain phosphodiesterases. J Mol Biol.
172. **Tischler, A. D., and A. Camilli.** 2004. Cyclic diguanylate (c-di-GMP) regulates Vibrio cholerae biofilm formation. Mol. Microbiol. **53:**857-869.
173. **Tomita, K., S. Fukai, R. Ishitani, T. Ueda, N. Takeuchi, D. G. Vassylyev, and O. Nureki.** 2004. Structural basis for template-independent RNA polymerization. Nature **430:**700-4.
174. **Tschowri, N., and R. Hengge.** 2009. The BLUF-EAL protein YcgF acts as a direct anti-repressor in a blue light response of E.coli. Genes Dev. **23:**522-534.
175. **Tucker, P. A., and L. Sallai.** 2007. The AAA+ superfamily - a myriad of motions. Curr. Opin. Stuct. Biol. **17:**641-652.
176. **Ugochukwu, E., A. L. Lovering, O. C. Mather, T. W. Young, and S. A. White.** 2007. The crystal structure of the cytosolic exopolyphosphatase from *Saccharomyces cerevisiae* reveals the basis for substrate specificity. J. Mol. Biol. **371:**1007-21.
177. **Uma, S., B. G. Yun, and R. L. Matts.** 2001. The heme-regulated eukaryotic initiation factor 2alpha kinase. A potential regulatory target for control of protein synthesis by diffusible gases. J Biol Chem **276:**14875-83.
178. **Waldron, K. J., and N. J. Robinson.** 2009. How do bacterial cells ensure that metalloproteins get the correct metal? Nat. Rev. Microbiol. **7:**25-35.
179. **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.
180. **Weinhouse, H., S. Sapir, D. Amikam, Y. Shilo, G. Volman, P. Ohana, and M. Benziman.** 1997. c-di-GMP-binding protein, a new factor regulating cellulose synthesis in *Acetobacter xylinum* FEBS Lett. **416:**207-211.
181. **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.
182. **Wendrich, T. M., and M. A. Marahiel.** 1997. Cloning and characterization of a relA/spoT homologue from Bacillus subtillis. Mol. Microbiol. **26:**65-79.
183. **Witte, G., S. Hartung, K. Büttner, and K.-P. Hopfner.** 2008. Structuralbiochemistry of a bacterial checkpoint protein reveals diadenylate cyclase activity regulated by DNA recombination intermediates. Mol. Cell **30:**167-178.
184. **Woodward, J. J., A. T. Iavarone, and D. A. Portnoy.** 2010. c-di-AMP secreted by intracellular Listeria monocytogenes activates a host type I interferon response. Science **328:**1703-5.
185. **Xie, F., S. H. Qureshi, G. A. Papadakos, and C. M. Dupureur.** 2008. One- and two-metal ion catalysis: global single-turnover kinetic analysis of the PvuII endonuclease mechanism. Biochemistry **47:**12540-50.

186. **Yamagata, A., Y. Kakuta, R. Masui, and K. Fukuyama.** 2002. The crystal structure of exonuclease RecJ bound to Mn2+ ion suggests how its characteristic motifs are involved in exonuclease activity. Proc. Natl. Acad. Sci. USA **99:**5908-12.
187. **Yamagata, A., R. Masui, Y. Kakuta, S. Kuramitsu, and K. Fukuyama.** 2001. Overexpression, purification and characterization of RecJ protein from Thermus thermophilus HB8 and its core domain. Nucleic Acids Res. **29:**4617-24.
188. **Yan, W., T. Qu, H. Zhao, L. Su, Q. Yu, J. Gao, and B. Wu.** 2010. The effect of c-di-GMP (3'-5'-cyclic diguanylic acid) on the biofilm formation and adherence of Streptococcus mutans. Microbiol Res **165:**87-96.
189. **Yan, X., C. Zhao, A. Budin-Verneuil, A. Hartke, A. Rince, M. S. Gilmore, Y. Auffray, and V. Pichereau.** 2009. The (p)ppGpp synthetase RelA contributes to stress adaptation and virulence in Enterococcus faecalis V583. Microbiology **155:**3226-37.
190. **Yang, W.** 2008. An equivalent metal ion in one- and two-metal-ion catalysis. Nat Struct Mol Biol **15:**1228-31.
191. **Yang, W., J. Y. Lee, and M. Nowotny.** 2006. Making and breaking mucleic acids: Two-Mg2+ catalysis and substrate specificity. Mol. Cell. **22:**5-13.
192. **Young, T. W., N. J. Kuhn, A. Wadeson, S. Ward, D. Burges, and G. D. Cooke.** 1998. Bacillus subtilis ORF yybQ encodes a manganese-dependent inorganic pyrophosphatase with distinctive properties: the first of a new class of soluble pyrophosphatase? Microbiol. **144:**2563-71.
193. **Zhang, W., and G. N. Phillips.** 2003. Structure of the Oxygen Sensor in Bacillus subtilis: Signal Transduction of Chemotaxis by Control of Symmetry. Structure **11:**1097-1110.
194. **Zhang, Z., B. L. Gaffney, and R. A. Jones.** 2004. c-di-GMP displays a monovalent metal ion-dependent polymorphism J. Am. Chem. Soc. **126:**16700-16701.
195. **Zuber, P.** 2009. Management of oxidative stress in Bacillus. Annu Rev Microbiol **63:**575-97.