Regulation of two-component signalling by a cyclic di-GMP-binding PilZ protein

Venkataramani, Prabhadevi

2016

Venkataramani, P. (2016). Regulation of two-component signalling by a cyclic di-GMP-binding PilZ protein. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/69390>

<https://doi.org/10.32657/10356/69390>

**Regulation of two-component signalling by
a cyclic di-GMP-binding PilZ protein**

Prabhadevi Venkataramani

School of Biological Sciences

2016

Regulation of two-component signalling by a cyclic di-GMP-binding PilZ protein

Prabhadevi Venkataramani

School of Biological Sciences

A thesis submitted to the Nanyang Technological
University in partial fulfilment for the degree of
Doctor of Philosophy

2016

Acknowledgements

I would like to extend my heartfelt thanks and express my sincere gratitude to my supervisor, Assoc. Prof. Liang Zhao-Xun, for all the encouragement, constant guidance, advice, patience and support he has shown me all through these 4 years.

I would like to extend a big thank you to my mentor Dr. Xu Linghui who pioneered and began this work on PilZ proteins in our lab. Without her constant support, guidance and motivation, I could not have completed my thesis. I also thank Lingyi, Grace and Wei Min for all their help and support as my group members.

I would also like to thank all our collaborators, Prof. Dr. Yang Liang, Prof. Dr. Yang Liu, Yichen Ding and Joey from SCESLSE, Prof. Dr. Yoon Ho Sup and Dr. Shin Joon from SBS and many others who have given me the opportunity to work with them.

I also thank my former and current lab members, Dr. Ho Chun Loong, Dr. Sun Huihua, Dr. Yang Lifeng, Grace, Mary, Alolika, Jamila, Ishin, Qing Wei, Lingyi, Limei, Zhen Jie, Ruijuan, and others whom I have not mentioned, for their constant encouragement and motivation.

This work wouldn't have been possible without the support of my husband, Anand, who has been the one who has always stood by me to motivate, cheer and encourage me during all my difficult times to bring out the best in me. I also would like to thank Assoc Prof. Ajai Vyas and his lab members - Wen Han and Samira, for being a great moral support to me throughout my PhD journey.

My heartfelt gratitude goes to my parents Dr. B.Venkataramani and Kamala Venkataramani who always wanted to see me achieve success and have always stood by me. They have always dreamed to see me get my doctoral degree. I would like to dedicate my thesis to them.

TABLE OF CONTENTS

Acknowledgements	2
List of Figures	7
List of Tables	9
Abbreviations	10
Abstract	11
CHAPTER 1: Introduction	12
1.1. <i>Pseudomonas aeruginosa</i>	12
1.2. Biofilm formation in <i>Pseudomonas aeruginosa</i>	14
1.3. Role of cyclic di-GMP in <i>P. aeruginosa</i>	16
c-di-GMP metabolism	17
GGDEF domain-containing proteins.....	17
GGDEF/EAL domain-containing composite proteins	21
EAL domain-containing proteins in <i>P. aeruginosa</i>	23
HD-GYP domains.....	24
1.4. c-di-GMP effectors in <i>P. aeruginosa</i>	25
PilZ domain proteins	25
Degenerate GGDEF domains	27
Degenerate EAL and HD-GYP domains.....	28
Riboswitches.....	28
Transcription factors or regulators	29
1.5. Physiological roles of c-di-GMP in <i>Pseudomonas aeruginosa</i>	29
Regulation of motility.....	29
Regulation of biofilm formation.....	30
Regulation of virulence	31
Regulation of exopolysaccharide production	32

1.6. Regulation of cellular processes in <i>Pseudomonas aeruginosa</i> by two component signalling (TCSs) and quorum sensing (QS) systems.....	34
Two-component signalling systems	34
Quorum sensing systems	37
1.7. Objectives.....	39
CHAPTER 2: Regulation of two-component signalling by c-di-GMP through a PilZ adaptor protein.....	41
2.1. Introduction	41
2.2. Materials and Methods	43
Construction of <i>P. aeruginosa</i> genomic DNA libraries	43
Bacterial two-hybrid screening.....	44
Co-immunoprecipitation (Co-IP) assay.....	44
Cloning, expression and purification of recombinant proteins.....	45
<i>In vitro</i> phosphorylation assay.....	46
Biofilm formation in flow-cell	46
RNA preparation and real-time PCR analysis	46
Swarming and twitching motility assay.....	47
Bacterial tethering analysis.....	47
Quantification of rhamnolipid production	48
Colony morphology assay	48
Transcriptome analysis by RNA-Seq	49
2.3. Results	49
PA2799 interacts with the phosphoreceiver (REC) domain of the hybrid histidine kinase SagS.....	49
The protein-protein interaction between HapZ and SagS is specific.	51
HapZ-SagS interaction is enhanced by c-di-GMP	52
Phosphotransfer between SagS and the downstream protein HptB is blocked by HapZ in a c-di-GMP-dependent manner	54

HapZ mediates surface attachment and biofilm formation.	55
HapZ transcription levels vary during different stages of biofilm formation.....	57
HapZ regulates swarming and twitching motility.	57
Colony morphology is influenced by HapZ	60
A global c-di-GMP pool may be involved in HapZ-mediated biofilm regulation	60
HapZ regulates the expression of a large number of genes related to QS, motility and other functions.	63
2.4. Discussion	65
CHAPTER 3: Protein-protein interaction between the SagS and its protein partners - HptB and HapZ.....	70
3.1. Introduction	70
3.2. Materials and Methods	73
Bacterial strains and plasmids	73
Protein expression and purification	73
NMR solution structure determination	74
NMR Titration	75
Generation of HapZ mutant library	75
Generation of REC _{SagS} mutant library	76
Bacterial two-Hybrid assay	76
Swarming motility assay	77
Colony morphology assay	77
3.3. Results	78
3.3.1. Specific protein-protein interaction between the receiver domain of SagS and HptB	78
The REC domain of SagS (REC _{SagS}) adopts a typical CheY-like structure	78
The receiver domain of SagS (REC _{SagS}) interacts with HptB.	80
The three Hpt proteins from <i>P. aeruginosa</i> adopt a typical all-helical structure.	83
A model of REC _{SagS} - HptB complex	86

3.3.2. Cyclic di-GMP dependent protein-protein interaction between the receiver domain of SagS and HapZ.....	90
Structural model of HapZ.....	90
Replacement of six residues in HapZ abolishes HapZ-REC _{SagS} interaction.....	91
Phenotypic assays validate the role of six residues of HapZ in abolishing binding with REC _{SagS}	96
3.4. Discussion	98
References	105
Appendix I	126
Appendix II.....	155
List of Publications.....	155
Conference Talks.....	155

List of Figures

Figure No.	Title	Page No.
Chapter 1		
Fig. 1	Stages of biofilm formation in <i>Pseudomonas aeruginosa</i>	15
Fig. 2	Basic c-di-GMP signalling pathway	17
Fig. 3	GGDEF domain-containing diguanylate cyclases in <i>P. aeruginosa</i>	19
Fig. 4	GGDEF/EAL domain-containing composite proteins in <i>P. aeruginosa</i>	22
Fig. 5	EAL domain-containing proteins in <i>P. aeruginosa</i>	24
Fig. 6	Types of PilZ domain proteins.	26
Fig. 7	Types of two-component signalling pathways in bacteria	36
Chapter 2		
Fig. 8	HapZ (or PA2799) interacts with SagS phosphoreceiver domain as observed in the bacterial two-hybrid screening	50
Fig. 9	c-di-GMP enhances HapZ-REC _{SagS} interaction	52
Fig. 10	c-di-GMP binding is important for HapZ-REC _{SagS} interaction	53
Fig. 11	The phosphotransfer from SagS to HptB is inhibited by HapZ and this effect is enhanced by an increasing gradient of c-di-GMP	55
Fig. 12	HapZ controls surface attachment and development of biofilm	56
Fig. 13	HapZ transcription levels vary during different stages of biofilm formation	58
Fig. 14	HapZ impacts swarming and twitching but not swimming motility	59
Fig. 15	HapZ influences colony morphogenesis	60
Fig. 16	Complexity of c-di-GMP signalling network in <i>P. aeruginosa</i>	61
Fig. 17	A global c-di-GMP pool may be involved in the HapZ-mediated biofilm regulation	62
Fig. 18	HapZ and SagS co-regulate the expression of several genes	64
Fig. 19	A model depicting the role of HapZ in an intricate two-component signalling network	67
Chapter 3		
Fig. 20	The NMR solution structure of phosphoreceiver domain of SagS	79
Fig. 21	NMR titration of REC _{SagS} and HptB	81
Fig. 22	Sequence alignment of the 11 REC domain-containing hybrid sensor histidine kinases in <i>P. aeruginosa</i>	82
Fig. 23	Comparative modeling of the histidine phosphotransfer proteins – HptA, HptB and HptC from <i>P. aeruginosa</i> using Robetta server	84
Fig. 24	Structure based sequence alignment of the three histidine phosphotransfer proteins from <i>P. aeruginosa</i> with the HPT protein YPD1 from <i>Saccharomyces cerevisiae</i> .	86

Fig. 25	Sequence alignment based comparison of residues on HptB predicted to interact with REC _{SagS} with the corresponding residues on two other histidine phosphotransfer proteins - HptA and HptC from <i>P. aeruginosa</i>	87
Fig. 26	Docking model of REC _{SagS} -HptB interaction.	88
Fig. 27	HapZ adopts a typical β barrel structure	91
Fig. 28	Residues chosen for HapZ mutant library	92
Fig. 29	Residues chosen for REC _{SagS} mutant library	93
Fig. 30	Bacterial two-hybrid screening of HapZ mutants against REC _{SagS} .	95
Fig. 31	Bacterial two-hybrid screening of REC _{SagS} mutants against HapZ	96
Fig. 32	Phenotypic assays validate the importance of the six HapZ residues in HapZ- REC _{SagS} interaction	97
Fig. 33	Model of HapZ in complex with dimeric c-di-GMP	103
Fig. 34	Model of REC _{SagS} -HapZ interaction	104
Appendix I		
Fig. S1	Agarose DNA gel shows the size of the <i>P. aeruginosa</i> gDNA fragments (the brightest band of each lane) captured by the prey vectors	126
Fig. S2	Sequence alignment of the kinase catalytic domains of CckA from <i>C. crescentus</i> with different histidine kinases in <i>P. aeruginosa</i>	126
Fig. S3	Expression of HapZ using various solubility tags	142
Fig. S4	Loading control for the <i>in vitro</i> phosphorelay experiment (Chapter 2).	143
Fig. S5	Ramachandran plot for REC _{SagS} solution structure	154

List of Tables

Table No.	Title	Page No.
Chapter 1		
Table 1	List of <i>P. aeruginosa</i> infections in humans	13
Appendix 1		
Table S1	List of bacterial strains used in the study	127
Table S2	List of plasmids used in the study	129
Table S3	List of primers used in the study	135
Table S4	Genes that are UP regulated in both $\Delta sagS$ and $\Delta hapZ$ mutants	144
Table S5	Genes that are DOWN regulated in both $\Delta sagS$ and $\Delta hapZ$ mutants	145
Table S6	Genes that are UP regulated in $\Delta hapZ$, but not in $\Delta sagS$ mutant	145
Table S7	Genes that are DOWN regulated in $\Delta hapZ$, but not in $\Delta sagS$	146
Table S8	Genes that are UP regulated in $\Delta sagS$, but not in $\Delta hapZ$	146
Table S9	Genes that are DOWN regulated in $\Delta sagS$, but not in $\Delta hapZ$	147
Table S10	Chemical shift table (ppm) for the residues of REC _{SagS}	148
Table S11	Structural statistics for final 20 REC _{SagS} structures	153

Abbreviations

c-di-GMP	Cyclic di-guanosine monophosphate
<i>P. aeruginosa</i>	<i>Pseudomonas aeruginosa</i>
<i>E. coli</i>	<i>Escherichia coli</i>
DGC	Digulanylate cyclase
PDE	Phosphodiesterase
GGDEF	“Gly-Gly-Asp-Glu-Phe” motif-containing domain with DGC activity
EAL	“Glu-Ala-Leu” motif containing domain with c-di-GMP PDE activity
HD-GYP	“His-Asp-Gly-Tyr-Pro” domain
I-site	Inhibitory site
GMP	Guanosine monophosphate
GTP	Guanosine triphosphate
c-di-AMP	Cyclic di-adenosine monophosphate
GFP	Green fluorescent protein
REC	Phosphoreceiver domain
HapZ	<u>H</u> istidine kinase <u>a</u> ssociated <u>P</u> il <u>Z</u> (PA2799)
SagS	<u>S</u> urface <u>a</u> ttachment and <u>g</u> rowth <u>s</u> ensor hybrid
HptB	<u>H</u> istidine <u>p</u> hospho <u>t</u> ransfer protein <u>B</u>
REC _{SagS}	Phosphoreceiver domain (663-786aa) of SagS
IPTG	Isopropyl-β-D-thiogalactopyranoside
Ni ²⁺ -NTA	Nickel-nitrilotriacetic acid
EDTA	Ethylenediaminetetraacetic acid
DTT	Dithiothreitol
NMR	Nuclear magnetic resonance spectroscopy
HSQC	Heteronuclear single quantum coherence spectroscopy
NOESY	Nuclear Overhauser effect spectroscopy
TOCSY	Total correlation spectroscopy
DSS	4,4-dimethyl-4-silapentane-1-sulfonic acid (NMR standard)
SOC	Super Optimal broth with Catabolite repression
RMSD	Root mean square deviation

Abstract

The bacterial messenger cyclic di-GMP (c-di-GMP) binds to a diverse range of effectors to exert its biological effect. Although free-standing PilZ proteins are by far the most prevalent among c-di-GMP effectors, their physiological functions remain largely unknown. In the present study, I found that the free-standing PilZ protein PA2799 (named as HapZ - histidine kinase associated PilZ) from the opportunistic pathogen *Pseudomonas aeruginosa* interacts directly with the phosphoreceiver (REC) domain of the hybrid histidine kinase SagS. The interaction between SagS and HapZ is further enhanced at elevated c-di-GMP levels. I demonstrated that binding of HapZ to SagS inhibits the phosphotransfer from SagS to the downstream protein HptB in a c-di-GMP-dependent manner. Consistent with the role of SagS as a motile-sessile switch and biofilm growth factor, I found that HapZ impacts surface attachment and biofilm formation most likely by regulating the expression of a large number of genes. These findings suggest a previously unknown mechanism whereby c-di-GMP mediates two-component signalling through a PilZ adaptor protein. To understand the molecular basis of the protein-protein interaction between the receiver domain of SagS (REC_{SagS}) and both HptB and HapZ, we determined the solution structure of REC_{SagS} by NMR spectroscopy to show that REC_{SagS} adopts a typical CheY-like protein fold. By structural modelling and NMR titration, we found that the interaction between REC_{SagS} and HptB involves a buried hydrophobic interface and potentially several specific hydrogen bonding residues. As some of the interacting residues are different in the homologous proteins HptA and HptC, the findings provide a mechanistic explanation for the specific recognition of SagS by HptB, but not HptA and HptC. To understand the interaction between SagS and HapZ, systematic site-directed mutagenesis and bacterial two-hybrid binding assays were performed to identify key residues involved in the protein-protein interaction. The results from the two-hybrid binding assays suggest that the only essential residues for protein-protein binding are the c-di-GMP binding residues. Based on the observations, we proposed a model where dimeric c-di-GMP mediates the interaction between HapZ and REC_{SagS} by functioning as a molecular glue. Together, the research work reveals a novel mechanism used by c-di-GMP to regulate two-component signalling through a PilZ adaptor protein and advance our understanding of the molecular mechanism of c-di-GMP signalling.

CHAPTER 1: Introduction

Throughout history, infectious diseases have been the primary cause of mortality and morbidity in humans. Although the 20th century was regarded as the era of antibiotics, the rise of the immunocompromised population has resulted in the precipitous upsurge of infectious diseases. Moreover, the most striking development is the capability of micro-organisms to evolve survival strategies and adapt to their antibiotic laden environment. Thus, we are plagued with the therapeutic challenge of combating the rising population of antibiotic resistant micro-organisms in the ‘post-antibiotic era’ (1, 2). Owing to its pervasive distribution and intrinsic resistance mechanisms to drugs, *P. aeruginosa* poses a serious threat to humans. Moreover, it has been reported to be second most commonly isolated pathogen from ICU-acquired infections (3, 4).

1.1. *Pseudomonas aeruginosa*

The genus *Pseudomonas* is one of the most ecologically significant and varied groups of bacteria. The members of this genus have a ubiquitous distribution owing to their adaptability to various environmental niches. *Pseudomonas aeruginosa* differs from the other members of the genus in its pathogenicity to humans and other mammals (Table 1) and its ability to utilize several environmental compounds as sources of energy (5). Carle Gessard succeeded in isolating the organism for the first time in 1882 and named it *Bacillus pyocyaneus* (6). In 2000, the first completed genome of the *P. aeruginosa* reference strain PAO1 was published with a size of 6.3 Mbp, containing 5570 open reading frames (7). Presently, the complete genomes of nine other clinical strains of *P. aeruginosa* such as PA14 are available, that share 80% sequence identity with the PAO1 genome (8). In the 1960s, *P. aeruginosa* emerged as a major cause of gram-negative bacteremia resulting in very high mortality rates (~90%). Although there has been an advancement in recent times due to the advent of antibiotics, the morbidity and mortality rate resulting from *P.aeruginosa* infections still continues to range between 18% to 61% (9).

Being an opportunistic pathogen, it is responsible for causing several chronic and acute diseases including nosocomial infections, urinary tract infections, chronic obstructive

Chapter 1

pulmonary disease and cystic fibrosis. Septicemia and severe tissue damage resulting from acute infections can be life threatening. During long-term persistence in chronic infections, *P. aeruginosa* develops several genetic adaptations, thus showing that it possesses the ability to choose diverse strategies during different types of infections (10).

Table 1 | List of *P. aeruginosa* infections in humans [Adapted from (11)]

Organ affected	Infection	Acute or chronic
Eye	Contact lens keratitis	Acute
Skin	Burn infection Wounds, ulcer Infections, folliculitis	Acute
Nasal sinuses	sinusitis	Chronic
Ear	External otitis	Acute/chronic
Lungs/Bronchi	Cystic fibrosis, Bronchiectasis Endobronchiolitis, Ventilator associated pneumonia,	Chronic Acute
Blood	Sepsis, neutropenic patients	Acute
Bones	Diabetic osteomyelitis in feet	Chronic
Urinary bladder	Urinary tract infection	Acute/chronic

P. aeruginosa makes a choice between initiating an acute or chronic infection based on several factors such as infection route, environmental signals, host immune status, tissue integrity and patient nutrition (12). In case of an acute infection, it can invade the lungs and disseminate into the bloodstream, causing imminent death. This is due to the production of extracellular toxins due to the activation of the Type III secretion system (TTSS) (13-16). In addition to TTSS activation, pathogenesis during acute infection is typified by the production of various virulence factors such as lipopolysaccharides (17, 18), elastases (15, 19), type IV pili (20), hydrogen cyanide (21-23) and quorum sensing (QS) molecules (24-26).

In contrast, chronic infections caused by *P. aeruginosa* are quite distinct from acute infections. Chronic infections seldom reach the bloodstream and can persist for several years. Biofilm formation can lead to chronic infections *in vivo* (27-29). This has been well documented in the case of cystic fibrosis of lungs by *P. aeruginosa* (30). This scenario is further complicated by the multi-drug resistance of biofilms in chronic infections (31). This

Chapter 1

further limits and hinders current therapeutic routes. *P. aeruginosa* utilizes multiple antibiotic-resistance mechanisms such as the presence of an AmpC β -lactamase, induced by β lactams making it resistant to ampicillin and cephalothin (32). Furthermore, it has efflux pumps such as MexAB-OprM, that makes it impervious to multiple antibiotics including chloramphenicol, novobiocin, β lactams and fluoroquinolones (33). Moreover, it can develop resistance to antibiotics such as colistin and aminoglycosides that they are not intrinsically resistant to (34). This resistant phenotype results from either the environment of the CF lung (osmotic or oxidative stress) or the selection of a phenotypic variant within the biofilm due to sub-lethal antibiotic concentrations used in treating the CF infection (35). Thus, combating resilient biofilms of *P. aeruginosa* has emerged as a major challenge in the field of microbiology.

1.2. Biofilm formation in *Pseudomonas aeruginosa*

Biofilms are structured communities characterized by an envelope of exopolysaccharide enclosing a closely juxtaposed bacterial population, adhering to a sessile or an inert surface. The extracellular matrix produced by the micro-organisms in the biofilms confers structural integrity to the biofilm (36). The extracellular matrix produced by the bacteria residing in the biofilm include various exopolysaccharides, non-fimbrial adhesins, adhesive pili and extracellular DNA (37). Bacteria residing in biofilms can be upto 1000 times more resistant to antibiotics than their planktonic counterparts (38). Community existence within biofilms provides *P. aeruginosa* additional mechanisms to survive with increased metabolic efficiency (39) and combat various stresses such as dessication, radiation etc. imposed by the surrounding environment (35). Pulmonary infections caused by *P. aeruginosa* reach a chronic stage, especially in patients suffering from cystic fibrosis owing to its tendency to form biofilms (40).

The understanding of *P. aeruginosa* biofilm development has largely stemmed from studying biofilms in flow-cell systems. Biofilm development in *P. aeruginosa* occurs in a progressive manner involving several steps which include: (1) Initial and reversible cell attachment, (2) Irreversible cell attachment (3) Microcolony formation, (4) Biofilm differentiation and maturation and (5) Biofilm dispersal (Fig. 1).

Chapter 1

The first stage in biofilm development is the initial attachment to a surface. The mechanisms by which bacterial cells reach the surface include flagellar motion, Brownian movement via diffusion or chemical attraction (41). Reversible attachment is mediated by flagellum-associated motility that aids the bacteria to surmount the repulsive forces imposed by the water-surface interface enabling their movement to the substratum (42). Microcolony formation and biofilm development requires twitching motility facilitated by the retraction and extension of type IV pili allowing the movement of cells on a surface (43-46). During biofilm maturation, mushroom-shaped microcolonies are formed with interconnected void channels which are critical for supplying nutrients and removing wastes from cells within the microcolonies (47). A mushroom shaped microcolony is comprised of a non-motile stalk population and a motile cap population. The motile population in the cap express type IV pili (48) and are increasingly resistant to membrane-targeted antibacterials (49). On the other hand, rhamnolipids and matrix-stabilizing extracellular DNA (eDNA) are confined mainly to the stalk sub-populations (48, 50).

Figure 1 | Stages of biofilm formation in *Pseudomonas aeruginosa*. The biofilm formation in *P. aeruginosa* consists of four distinct stages namely initial attachment, microcolony formation, biofilm maturation and biofilm dispersal. The biofilm formation begins when free living cells in the planktonic stage attach initially to a surface followed by irreversible attachment. The attached cells gradually form microcolonies and eventually develop to mature mushroom shaped biofilm with a stalk and a cap. During biofilm dispersal, the matrix-associated cells revert back to their free swimming state in response to environmental factors.

Chapter 1

Biofilm dispersal in *P. aeruginosa* occurs when matrix-encased sessile bacteria revert to their free-swimming form to colonize new environments. It can be mediated by various factors like nutrient availability such as oxygen (51), carbon sources (52-55) and iron (56), physiological factors, regulatory processes such as cellular c-di-GMP levels (57), exopolysaccharide lyase-facilitated biofilm matrix breakdown (58), free radical production (59, 60), surfactants (61) and a short chain fatty acid molecule *cis*-2-decenoic acid (62).

1.3. Role of cyclic di-GMP in *P. aeruginosa*

The versatile nature of *P. aeruginosa* is due to the presence of complex regulatory networks that include genes involved in the metabolism of a second messenger molecule cyclic-di (3', 5')-guanylic acid (c-di-GMP). High levels of c-di-GMP has been linked with biofilm or sessile lifestyle while low levels have been correlated with a motile lifestyle (63).

Second messengers play key roles in intracellular signal transduction in bacteria. One such second messenger, cyclic di-GMP, was serendipitously discovered 29 years ago by Benziman and his collaborators while studying the biosynthesis of cellulose as an allosteric activator of the enzyme cellulose synthase in the bacterium *Acetobacter xylinum*. C-di-GMP was identified as the cofactor responsible for activating cellulose synthase at submicromolar K_d values (64-67). Subsequently, c-di-GMP was also found in the cellulose producing soil-borne bacterium *Agrobacterium tumefaciens* (68). Although c-di-GMP exists as a stable monomer or dimer under physiological conditions, it is also capable of forming higher oligomers including octamers in the presence of cations (69).

Being unique in bacteria, c-di-GMP, has gained importance as a global regulatory molecule responsible for triggering a multitude of physiological responses. C-di-GMP mediates several cellular functions at transcriptional, translational and post-translational levels ranging from cell-cycle regulation, cell differentiation, biogenesis of extracellular components such as pili and flagella, formation and dispersion of biofilms to the synthesis of exopolysaccharides. This molecule acts as a switch between planktonic and sessile modes of lifestyles in bacteria by inhibiting various forms of motility. C-di-GMP has been recognized as an effective immunomodulator thus making it a promising vaccine adjuvant (70). In *P. aeruginosa*, c-di-GMP plays a pivotal role in regulating many cellular processes from motility, biofilm formation to virulence and exopolysaccharide production (63, 71).

Chapter 1

c-di-GMP metabolism

C-di-GMP is synthesized by the condensation of two molecules of GTP and hydrolyzed via the intermediate [5'-pGpG] to GMP in a two-step reaction. While purifying factors contributing to c-di-GMP synthesis, the genes encoding enzymes in c-di-GMP synthesis and breakdown, namely diguanylate synthases (DGCs) and phosphodiesterases (PDEs) specific to c-di-GMPs respectively were identified (Fig. 2) (64). Genetic and biochemical studies in *G. xylinum* showed that GGDEF is the catalytically active domain of DGCs and the structurally unrelated EAL and HD-GYP are the two alternative domains possessing PDE activity (72-74). *P. aeruginosa* has 40 proteins that are involved in c-di-GMP metabolism of which 17 contain GGDEF domains, 16 have both GGDEF/EAL domains, five proteins contain only EAL domains and two possess HD-GYP domains (75).

Figure 2 | Basic c-di-GMP signalling pathway. C-di-GMP is synthesized by diguanylate cyclases (DGCs) that have a conserved GGDEF motif and is degraded by phosphodiesterases (PDEs) that either have a conserved EAL or HD-GYP motif. The c-di-GMP binding receptor protein domains can be divided into five classes - degenerate GGDEF and EAL/HD-GYP, Riboswitches, PilZ and transcriptional regulators. These receptor proteins bind c-di-GMP and regulate several downstream effector pathways.

GGDEF domain-containing proteins

C-di-GMP is synthesized by DGCs that contain a highly conserved Gly-Gly-Asp-Glu-Phe (GGDEF) motif (76). The signature motif GG[D/E]EF of DGC is the active site and is

Chapter 1

implicated in GTP binding. The formation of the phosphoester bond by DGCs requires two Mg^{+2} or Mn^{+2} ions. Within the motif, the first two (Gly) residues are involved in GTP binding; the third (Asp/Glu) residue is required for catalysis and metal ion coordination while the fourth residue (Glu) is solely involved in metal ion coordination. GGDEF domains function as homodimers and catalysis is facilitated when two molecules of GTP bind to each of the GGDEF domains in a DGC dimer. Feedback product inhibition of GGDEF domains by c-di-GMP occurs at a conserved inhibitory site (I-site) that contains the RxxD motif (77).

In *P. aeruginosa*, seventeen GGDEF domain-containing proteins have been identified of which seven have transmembrane regions (75). In addition to GGDEF domains, some proteins have an additional domain implicated in signal detection such as PAS and PAC domains (heme binding domains), CHASE4 domains (Extracellular domains forming part of the CHASE family), REC domains (CheY-like response regulator domains), HAMP domains (linkers between signal transduction and receptor proteins), GAF domains (cGMP-specific phosphodiesterases, adenylyl cyclases and FhlA), and 7TMR/DISMED2 domains (a varied intracellular signalling domains containing seven transmembrane regions) (Fig. 3).

It was observed that PA14 strains overexpressing the GGDEF domain proteins PA0338, PA0847, PA1107 (RoeA), PA1120 (YfiN), PA3702 (WspR), PA4332 (SadC) and PA5487 formed noticeable pellicles in static liquid culture. Formation of pellicles is one of the diverse ways in which *P. aeruginosa* forms biofilms at the air-liquid interface, thus suggesting a hyperbiofilm (excessive biofilm) phenotype. This indicates that high c-di-GMP levels in these strains lead to enhanced attachment to surfaces evidenced by pellicle formation. This was further confirmed when insertions in genes *PA0169* (*siaD*), *PA1120*, *PA5487* led to abolishment of biofilm formation while insertions in *PA1107* and *PA3702* genes resulted in decreased biofilm formation (75).

The functions of a few GGDEF proteins have been characterized in *P. aeruginosa*. PA0169 or SiaD (SDS induced aggregation D) has a cytoplasmic GGDEF domain and forms part of the *siaABCD* operon involved in SDS (sodium dodecyl sulphate) induced cell-aggregation in a c-di-GMP dependent mode. Autoaggregation is a stress response to toxic substances that is used as a survival strategy. This occurs when the sensor protein SiaA (PA0172) detects a stress stimulus and triggers a downstream signal transduction cascade (78). Another DGC, PA1107 was found to affect biofilm formation by controlling the production of Pel

Chapter 1

polysaccharide. It has an active GGEEF motif and is capable of producing increased levels of c-di-GMP (both *in vivo* and *in vitro*). Due to its role in exopolysaccharide production, it was renamed as RoeA (regulator of exopolysaccharide A) (79).

PA1120 encodes the membrane-bound DGC YfiN or TpbB (tyrosine phosphatase related to biofilm formation B) that consists of a PAS, HAMP and a C-terminal GGDEF domain. Deletion of the *yfiN* gene was demonstrated to enhance swimming motility and increase extracellular DNA (eDNA) by ~3.2 fold (80). In one study, it was demonstrated that the tyrosine phosphatase TpbA negatively regulates c-di-GMP production by dephosphorylation of TpbB (81). YfiN is also part of the tripartite YfiBNR system and its activity is inversely regulated by periplasmic YfiR and outer membrane lipoprotein, YfiB. The crystal structure of the catalytic GGDEF domain revealed that YfiN does not undergo feedback product inhibition and requires the HAMP domain for its dimerization and catalytic activity (82).

Figure 3 | GGDEF domain-containing diguanylate cyclases in *P. aeruginosa*. *P. aeruginosa* has 17 GGDEF domain-containing DGCs of which 7 of them have transmembrane regions (PA4929 has 2 seven-transmembrane regions). In addition to GGDEF domains, some of the DGCs contain additional domains such as PAC, PAS, HAMP, GAF, CHASE4 and response regulator (REC) domains.

Chapter 1

The well-studied REC-GGEEF domain-containing protein PA3702, also known as WspR (wrinkly spreader phenotype R) is a part of an operon encoding a chemosensory signalling complex. It was discovered originally in *Pseudomonas fluorescens* SBW25 and subsequently in *P. aeruginosa* while screening for colonies displaying a wrinkling phenotype (83, 84). WspR has a characteristic wrinkly spreader phenotype and affects biofilm formation (75, 85). Mutation of the *wsp* methylesterase *wspF* leads to the production of increased c-di-GMP causing a wrinkly phenotype. This can be reversed back to a smooth phenotype by *wspR* mutation. Thus, the high levels of c-di-GMP observed in the *wspF* mutant can be attributed to WspR (86, 87). The crystal structure of the WspR-c-di-GMP complex (where c-di-GMP is bound at the I-site) indicates that the active form of WspR is a tetramer. It was also demonstrated that WspR alternates between the active dimer, c-di-GMP inhibited tetramer and inhibited elongated dimer conformations. Also, they elucidated the significance of phosphorylation in tetramer formation (88).

Another membrane-bound DGC PA4332 or SadC (surface attachment defective C) has five transmembrane regions at its N-terminus followed by an active C-terminal GGDEF domain. It forms part of the pathway involving the phosphodiesterase BifA and SadB (89). Mutations in *sadC* gene resulted in hyperswarming and defective surface attachment during biofilm formation while *sadC* overexpression increased c-di-GMP levels (79, 90). SadC plays a pivotal role in associating the Gac/Rsm cascade to c-di-GMP signalling (91). Recently, it was shown that the transmembrane (TM) domain and the α helix between its TM and GGDEF domains of SadC are critical for its DGC activity (92). Also, Liew CW, Liang ZX and Lescar observed a similar result in their study of Tbd1265 (Liew *et al* unpublished data). PA4843, also known as AdcA (AmrZ dependent cyclase A) or GcbA was found to mediate initial attachment during biofilm formation and impact swimming motility by reducing flagellar reversal rates. This was independent of viscosity and polysaccharide production (93). It also activates BdlA during biofilm development by post-translational modifications (94).

The periplasmic seven-transmembrane (7TMR-DISMED2) of the DGC protein PA4929 or NicD (nutrient induced cyclase D) can sense nutrients like glutamate, succinate and glucose (but not NO and ammonium chloride). Activation of NicD by nutrient-sensing results in its dephosphorylation and temporary enhancement of intracellular c-di-GMP levels (via its

Chapter 1

GGDEF domain). Thus eventually activates biofilm dispersion. This was further confirmed when *AnicD* biofilms were defective in dispersal upon glutamate exposure (95).

GGDEF/EAL domain-containing composite proteins

Besides 17 GGDEF domain proteins, *P. aeruginosa* also has sixteen composite proteins that contain both GGDEF and EAL domains (Fig. 4). In such cases, either one or both of these catalytic domains have been found to be in the active state (72). Ten of these GGDEF/EAL domain proteins in *P. aeruginosa* contain transmembrane regions and the functions of eight of them have been elucidated. The multi-domain protein RbdA (regulation of biofilm dispersal A) or PA0861 encodes a PAS domain in addition to GGDEF and EAL domains and is implicated in biofilm dispersal. An and colleagues showed that RbdA only possesses phosphodiesterase activity that is in turn activated by its GGDEF domain in the presence of GTP. RbdA also positively controls swimming and swarming motility and production of rhamnolipids (96). PA5017 or DipA (dispersion induced phosphodiesterase A) is another composite membrane protein mediating biofilm dispersal in *P. aeruginosa*. It has an N-terminal GAF domain followed by EAL domain at its C-terminus. DipA was found to reduce swarming motility but did not impact twitching or swimming motilities in the PA14 strain. Moreover, the typical GGDEF motif in DipA is absent and is substituted by an ASNEF motif due to which it only possesses phosphodiesterase activity, which in turn is enhanced by its GAF domain. (97).

MHYT is a seven-transmembrane sensing domain that is known to detect oxygen, carbon monoxide (CO) or nitric oxide (NO) through a coordination mediated by two copper atoms (98). Both the MHYT domain-containing GGDEF/EAL proteins have been studied in *P. aeruginosa* (Fig. 4). One of them is PA1727 that encodes the membrane-anchored MucR. It has been implicated in regulation of alginate biosynthesis. It has three MHYT domains arranged in tandem followed by GGDEF and EAL domains. MucR has been shown to exhibit DGC activity that is also necessary for its function (99). In one study, insertions in PA1727 gene resulted in biofilm reduction. On the other hand, thinner pellicles were formed on PA1727 overexpression and it displayed a very weak hyperbiofilm phenotype (75). Another composite protein described in *P. aeruginosa* with three N-terminal MHYT domains is PA3311 or NbdA (NO induced biofilm dispersal locus A) with only an active PDE activity. The phosphodiesterase activity in NbdA could be induced in the presence of NO and it is the first described PDE to mediate NO-induced biofilm dispersal (100).

Chapter 1

MorA (motility regulator A) encoded by *PA4601*, is a bi-functional GGDEF/EAL protein that was initially discovered in *Pseudomonas putida*. Flagellin development is repressed by MorA only in *P. putida*. On the other hand, MorA was found to affect the later stages of biofilm development in both *P. putida* and *P. aeruginosa* PAO1 (101). Tews and coworkers have determined the structure of the GGDEF-EAL fragment of MorA where active site formation distinguishes the two phosphodiesterase states (102). Another composite protein PA4367 or BifA (biofilm formation A) has an active phosphodiesterase activity attributed to the presence of both GGDEF and EAL domains. It has been implicated in mediating swarming motility. *ΔbifA* strain exhibited a hyperbiofilm phenotype, reduced flagellar reversals and high levels of c-di-GMP. Further, BifA was found to impact biofilm formation and swarming by functioning upstream of SadB (89).

Figure 4 | GGDEF/EAL domains-containing composite proteins in *P. aeruginosa*. *P. aeruginosa* has 16 GGDEF/EAL domain-containing proteins of which 7 of them have transmembrane regions (PA4929 has 2 seven-transmembrane regions). In addition to GGDEF domains, some of the DGCs contain additional domains such as PAC, PAS, HAMP, GAF, CHASE4 and response regulator (REC) domains.

FimX (PA4959) is a multimodular protein having an N-terminal response regulator (REC) in addition to GGDEF and EAL domains. It displays an active phosphodiesterase but no DGC activity *in vitro*. Its incomplete GGDEF domain stimulates EAL activity upon binding to

Chapter 1

GTP (103). Additionally, FimX has been shown to localize to a single pole, a process facilitated by its REC domain (103). LapD (PA1433) is a transmembrane receptor that was studied initially in *P. fluorescens* Pf01 (104). Likewise, recent studies on LapD in *P. aeruginosa* revealed that when LapD is activated by c-di-GMP, it sequesters the protease LapG. This LapG sequestration at high c-di-GMP levels allows CdrA (membrane-bound adhesin protein)-mediated cell-aggregation and facilitates biofilm formation (105).

EAL domain-containing proteins in P. aeruginosa

EAL domains are characterized by the presence of a conserved Glu-Ala-Leu (EAL) motif. In a study by Benziman and co-workers, purified PDEs from *G. xylinus* were shown to hydrolyze c-di-GMP into 5'-pGpG (linear di-GMP). This was dependent on Mg^{+2} or Mn^{+2} ions. Moreover, Ca^{+2} ions strongly inhibited this reaction. Monomeric pG was formed as the final hydrolysis product of c-di-GMP from 5'-pGpG by a different set of enzymes that possessed Ca^{+2} independent activity (65). It has been demonstrated that EAL domain proteins show phosphodiesterase (PDE) activity specific to c-di-GMP (106, 107). EAL domains were found to use a two metal catalytic mechanism for c-di-GMP hydrolysis (108). It was shown that the conserved Glu residue in EAL was involved directly in metal ion coordination (109).

Six EAL domain-containing proteins have been identified in *P. aeruginosa* (Fig 5). The aminoglycoside response regulator (Arr) or PA2818 was one of the first phosphodiesterases to be characterized. This inner-membrane PDE was demonstrated to play a role in aminoglycoside resistance during biofilm formation. Mutations in the catalytic EAL motif of Arr altered its biofilm response to tobramycin. Also, Δarr strain was susceptible to tobramycin by 100-fold as compared to the wild type PAO1 (110). Another EAL-like protein PvrR was identified as a controller of phenotypic switch from an antibiotic resistant to antibiotic susceptible form in *P. aeruginosa* PA14 and was proposed as a target for treating cystic fibrosis infections (35).

The Roc system has been implicated in upregulating the transcription of *cupB* and *cupC* genes that are part of the fimbrial gene cluster. RocR, an REC-EAL domain-containing protein, forms part of the Roc system involving two other proteins namely, RocS1 (sensor kinase) and RocA1 (DNA-binding response regulator). RocS1 binds to receiver domains of both RocR and RocA1. Also RocR was found to act antagonistically to RocA1 (111). Furthermore, Rao *et al* demonstrated the catalytic mechanism by which RocR hydrolyzes c-

Chapter 1

di-GMP aided by Mg^{2+} ions (112) and they also elucidated the importance of the conserved loop 6 region of RocR in facilitating EAL domain dimerization and binding of c-di-GMP and Mg^{2+} ions (113, 114).

Figure 5 | EAL domain-containing proteins in *P. aeruginosa*. *P. aeruginosa* has 6 EAL domain-containing proteins of which 2 of them have transmembrane regions. In addition to EAL domains, two of the PDEs PA3947 and PvrR contain response regulator (REC) domains.

HD-GYP domains

HD-GYP domains are members of a larger HD family that possess hydrolytic activities against a diverse range of substrates. Similar to the occurrence of GGDEF-EAL tandems, HD-GYP domains were also found to be tandemly associated with GGDEF domains in certain bacteria. This led to the prediction that HD-GYP domains may function as c-di-GMP specific PDEs as well (106). Mutagenesis of the two conserved residues namely His and Asp of the domain resulted in a loss of regulatory and c-di-GMP binding activity proving the hypothesis that HD-GYPs function as PDEs. Despite performing a similar function as the EAL domains, HD-GYP domains hydrolyze c-di-GMP to GMP and not 5'-pGpG (115).

P. aeruginosa encodes two HD-GYP domain-containing proteins namely, PA4108 and PA4781 and another modified YN-GYP domain protein PA2572. In addition to HD-GYP domain, PA4108 has an uncharacterized domain while PA4781 has a typical receiver domain (REC). Both these proteins have been shown to hydrolyze c-di-GMP to GMP relatively slowly in a two-step reaction involving a linear pGpG intermediate. Interestingly, PA4781 has been found to bind c-di-GMP only when its REC domain is in the phosphorylated form. Also, pGpG is an alternative substrate used by both PA4108 and phosphorylated PA4781, with PA4781 binding at a higher affinity compared to PA4108 (116). The affinity of PA4781 for pGpG was confirmed by the determination of its crystal structure. It also revealed the

Chapter 1

presence of a unique bi-metallic active pocket that does not contain iron unlike other structurally characterized HD-GYP domains (117). Functional studies on the three HD-GYP domain proteins showed an increase in c-di-GMP levels upon mutating the proteins PA4108 and PA4781, which is in agreement to their functions as phosphodiesterases. Both these proteins also affected virulence factor production, swarming and twitching motility, with the exception of PA2572 that displayed negligible effect. However all these three proteins were shown to affect biofilm architecture in a distinct manner (118).

1.4. c-di-GMP effectors in *P. aeruginosa*

Although the enzymes responsible for c-di-GMP synthesis and degradation have been characterized, there is a lack of knowledge of the mechanisms underlying the action of c-di-GMP on its target proteins or receptors. Recently, there has been a steady increase in related studies leading to the identification of various classes of c-di-GMP receptors /c-di-GMP specific effectors. On the basis of their primary sequences, the c-di-GMP receptors have been categorized into several classes. These include PilZ domain proteins, degenerate GGDEF and EAL domain receptors, riboswitches and transcription factors (Fig. 2) (119, 120).

PilZ domain proteins

PilZ was the first c-di-GMP protein receptor predicted by Amikam and Galperin as part of the protein glycosyltransferase of the cellulose synthase complex in the bacterium *G. xylinus* (121). PilZ was identified as a C-terminal 100 amino acid domain of the BcsA subunit of the glycosyltransferase enzyme. The frequent association of the BcsA domain downstream of few EAL and GGDEF domains proved it to be a good candidate as a c-di-GMP receptor. PilZ domain was named after the *P. aeruginosa* PilZ (PA2960) protein, that is responsible for pilus formation and mainly constitutes of this domain (122). PilZ was identified as the c-di-GMP receptor that was long-sought after (123-125). PilZ domain receptors have been shown to have the highest affinity for c-di-GMP (high-micromolar range) in relation to other c-di-GMP receptors, which have affinities in the low to mid-micromolar range (126-128).

X-ray crystallography and nuclear magnetic resonance (NMR) structures of several PilZ domain proteins have led to the confirmation of the prediction that two sets of conserved residues in the PilZ domain namely, RxxxR and (D/N)x(S/A)xxG, are responsible for binding c-di-GMP (121, 129-133). This also led to the revelation that c-di-GMP binds initially to the

Chapter 1

RxxxR primary binding loop that wraps around it to bring the remainder of the consensus residues into proximity. Since c-di-GMP mediates downstream signalling and regulates diverse cellular processes, PilZ domain proteins were regarded as interdomain linkers that initiate these signalling events by bringing two proteins together. The adoption of different oligomeric states by the PilZ proteins ranging from monomeric to tetrameric forms indicates varying modes of downstream signal transduction (69, 131-133).

Figure 6 | Types of PilZ domain proteins. PilZ can either exist as stand-alone PilZ proteins (52%) or as fusions with other domains such as YcgR, glycosyltransferase, cellulose synthase, GGDEF, EAL, HD-GYP, PAS, CheY-like receiver domains etc. Of these, the stand-alone PilZ domain proteins are the most predominant ones.

PilZ domain proteins can either be expressed alone as stand-alone PilZ protein domains with short N-terminal extensions or as fusions with other protein domains, including GGDEF, EAL, HD-GYP, PAS, CheY-like receiver (REC) domains, methyl accepting protein (MCP) domains, sugar transporters, adenylate/guanylate cyclases, DNA-binding domains, glycosyltransferase domains and others (121). Stand-alone PilZ comprise approximately half of all known PilZ domain proteins (Fig. 6). PilZ domains also exist in the active and inactive forms where inactive forms are unable to bind c-di-GMP.

Eight PilZ proteins have been recognized in *P.aeruginosa*, which includes the ‘PilZ’ (PA2960), from where the name PilZ originated. PA2960 was found to be an inactive PilZ protein since it demonstrated low or no affinity to c-di-GMP. This was attributed to the

Chapter 1

absence of some of the residues that form part of the PilZ signature motif (121). Pfam database searches revealed that of the eight PilZ proteins in *P.aeruginosa*, the PilZ (PA2960) exhibited a phenotype for twitching motility (122). Of the seven PilZ domain proteins that bind c-di-GMP, the functions of two di-domain PilZ proteins Alg44 and FlgZ have been elucidated. The gene *alg44* (PA3542) was assigned functionally in the synthesis of alginate by Remminghorst and Rehm (134). Subsequently, its role in the alginate synthesis complex and in the regulation of polymerization and transport of alginate was elucidated (135). The crystal structure of Alg44 in complex with dimeric c-di-GMP was reported by Howell *et al* which revealed the typical PilZ-like fold adopted by the protein and the critical residues involved in c-di-GMP binding (71). Recently, the di-domain protein FlgZ (PA3353) was demonstrated to bind to the flagellar stator protein MotC and affect swarming motility in a c-di-GMP dependent manner (136).

The NMR solution structure of the stand-alone PilZ PA4608 in the apo and holo forms (in complex with c-di-GMP) was determined independently by two research labs. The studies revealed that PA4608 has an N-terminal unstructured loop followed by a six-stranded β barrel with β sheets oriented in an antiparallel manner. One α -helix and two shorter 3^{10} helices are packed at the C-terminal end of the barrel (130, 132). The mutually intercalated c-di-GMP dimer binds at one edge of the barrel and is wrapped at its external end by the N-terminal loop. The ligand binding induces a conformational change in PA4608 due to reorientation of several residues creating an exposed negatively charged cluster. This indicated that PA4608 might bind to a downstream effector protein with a positively charged surface (132). Though studies have been done on a few PilZ proteins in *P. aeruginosa*, the functions of the di-domain PilZ protein PA2989 and remaining four stand-alone PilZ proteins namely PA0012, PA2799, PA4324 and PA4608 have not been elucidated so far.

Degenerate GGDEF domains

Several GGDEF domains that lack catalytic activity retain their ability to bind c-di-GMP via their allosteric I-sites, thus functioning as c-di-GMP receptor proteins (120). An example in *P. aeruginosa* is PelD (PA3061) that regulates the synthesis of Pel polysaccharide post-translationally in a c-di-GMP dependent manner. It has also been shown that c-di-GMP binding to PelD plays a role in facilitating pellicle formation (137, 138). Two independent studies by Howell *et al* and Nair *et al* solved the crystal structure of cytosolic region of PelD

Chapter 1

both in the apo and holo forms (in complex with c-di-GMP). The crucial residues in the I-site (RxxD) involved in c-di-GMP binding were elucidated (138-140). Also, Howell and co-workers demonstrated that the α -helical region juxtaposed between the transmembrane fragment and the soluble domain of PelD is crucial for mediating its dimerization (138).

Degenerate EAL and HD-GYP domains

Similar to degenerate GGDEF domains, degenerate EAL or HD-GYP domains have lost their catalytic activity but are able to bind c-di-GMP. FimX is an interesting example of an EAL domain protein that binds c-di-GMP and has an almost negligible catalytic activity (141). The binding of c-di-GMP to FimX was found to bring about a huge conformational change in its REC domain. Also, FimX was unable to bind c-di-GMP upon mutagenesis of its EAL motif (142). Further, the crystal structure of FimX in the apo and holo (with 5'-pGpG) for confirmed that it forms a dimer possessing a large cavity that accommodates the 5'-pGpG moiety. (103, 142-144). On the other hand, the only example of a catalytically inactive HD-GYP domain is the variant (YN-GYP) protein PA2572 from *P. aeruginosa*. Though it has an influence on biofilm formation and the synthesis of virulence factors, its functional dependence on c-di-GMP has not yet been established (118).

Riboswitches

Riboswitches are RNA aptamers that form a part of the non-coding 5'-untranslated region (UTR) of messenger RNAs. They bind to small molecular ligands to activate or repress the expression of genes at levels of transcription and translation and are capable of forming secondary structures that change upon ligand binding (145). Several metabolite binding riboswitches have been reported including those binding to certain amino acids, S-adenosylmethionine (SAM), flavin mononucleotide (FMN), adenine, guanine etc. Breaker and co-workers discovered riboswitches that bind to c-di-GMP in the Vc2 RNA of *Vibrio cholerae* and found that c-di-GMP binds to a specific element called GEMM within the riboswitches (146). A second type of riboswitch was identified in the bacterium *Clostridium difficile* where c-di-GMP binding was responsible for regulating the activity of a self-splicing ribozyme controlling the expression of a virulence gene in this pathogen (147). Moreover, c-di-GMP was found to bind to these riboswitches at nanomolar affinities. Thus far, no riboswitches have been identified in *P. aeruginosa*.

Chapter 1

Transcription factors or regulators

Certain global transcriptional regulators have been identified to respond and bind c-di-GMP. c-di-GMP controls the transcription of the genes regulated by these factors in a positive or negative manner. The first such transcriptional regulator identified was FleQ from *P. aeruginosa*. It is a member of the NtrC family and represses the expression of the *pel* operon genes by binding to the *pelA* gene promoter. The repression of *pel* expression by FleQ is relieved partially by a second protein FleN and completely upon binding of c-di-GMP to the FleQ-FleN complex (120, 148). Another transcriptional factor in *P. aeruginosa* is BrlR (Biofilm resistance locus regulator) that belongs to the MerR family of multi-drug efflux pump activating proteins. It has been demonstrated as an activator of efflux pumps and was also found to play a role in tobramycin and colistin resistance in a reciprocal manner (149-151). The DNA binding activity of BrlR was also shown to enhance by c-di-GMP binding. This in turn increased its promoter activity promoting *brlR* expression (152).

1.5. Physiological roles of c-di-GMP in *Pseudomonas aeruginosa*

Diverse ranges of phenotypes in *P. aeruginosa* are affected by the second messenger c-di-GMP. Few of the phenotypes recognized to be associated with c-di-GMP include biofilm formation, virulence and transition from sessility to motility (37, 57, 74).

Regulation of motility

P. aeruginosa can adopt several modes of motility including swimming, swarming and twitching motility. Swimming and swarming motility require the flagella while twitching motility utilizes type IV pili. C-di-GMP can regulate all these forms of motility (57, 83). The DGC SadC inhibits while the PDE BifA promotes swarming motility in *P. aeruginosa* PA14 (89, 90). Deletion of *bifA* enhances c-di-GMP levels and leads to increase in PEL polysaccharide production. However, swarming motility inhibition in *bifA* mutant was attributed to high c-di-GMP levels and was independent of Pel polysaccharide synthesis (89). Since SadC and BifA were demonstrated to regulate flagellar reversal rate rather than flagellar biosynthesis, their role was implicated in affecting swarming but not swimming motility. The PilZ protein FlgZ has also been reported to affect swarming motility by binding to the flagellar stator protein MotC in a c-di-GMP dependent mode (136).

Chapter 1

Twitching motility is mediated by type IV pili by a process of extension and retraction of a cell. Nearly 40 gene products are essential for type IV pili assembly (43, 153). *FimX* (PA4959) mutants showed reduced twitching motility and is responsible for modulating the levels of surface pili rather than the amount of pilin monomers (103, 154). FimX was also found to regulate twitching motility in response to environmental signals establishing it as a connecting link between environmental cues, motility and c-di-GMP metabolism (154).

Regulation of biofilm formation

Intracellular c-di-GMP signalling in *P. aeruginosa* plays an important role in the maintenance and maturation of biofilms (57, 83, 86). Several GGDEF/EAL domain proteins affect biofilm formation. One study investigated the impact of knockout and overexpression strains of the GGDEF and EAL domain proteins on biofilm formation (75). The GGDEF domain proteins PA1107, PA1120, PA3702 and PA5487 displayed an increased biofilm formation in the overexpression and a decreased biofilm in the knockout strains. On the contrary, the knockout strains of other GGDEF/EAL domain proteins such as PA0338, PA2133 did not show the phenotypic variation similar to the four GGDEF proteins indicating that these proteins affect biofilm formation under experimental conditions other than those used in this study. Additionally, an increase in c-di-GMP concentrations was observed upon overexpressing eight GGDEF or GGDEF/EAL domain proteins (75). The DGC SadC was shown to play a central role in the Gac/Rsm two-component signalling pathway. Moreover, its production was found to be controlled by the translational repressor protein RsmA (90). Wozniak and co-workers identified that the transcription factor AmrZ regulates biofilm development in *P. aeruginosa* in a c-di-GMP dependent manner. The GGDEF protein GcbA or AdcA (PA4843) was repressed to a great extent by AmrZ. The $\Delta adcA \Delta amrZ$ strain formed fewer microcolonies compared to the $\Delta amrZ$ strain revealing that the hyperbiofilm phenotype of $\Delta amrZ$ is attributed to AdcA (155). Also, WspR has been shown to exhibit a characteristic sub-cellular clustering upon activation by phosphorylation. This greatly enhances its DGC activity and promotes biofilm formation in *P. aeruginosa* (88, 156, 157). C-di-GMP was demonstrated to be the molecular basis of biofilm response to aminoglycoside antibiotics like tobramycin from studies on the phosphodiesterase Arr (PA2818) in *P. aeruginosa*. This tobramycin-mediated biofilm response induced by Arr was suppressed by the addition of GTP that inhibited its PDE activity (110). In addition to Arr, the c-di-GMP binding transcription factor BrlR also contributes reciprocally to colistin and tobramycin

Chapter 1

resistance in biofilms. It binds to DNA, which is enhanced in the presence of c-di-GMP and functions as an activator of multi-drug efflux pumps (150, 151).

In contrast to biofilm formation, biofilm dispersal is characterized by a rapid drop in c-di-GMP levels (53, 158). In *P.aeruginosa* PAO1, dispersal occurs in response to environmental signals like hydrogen peroxide or heavy metals. The chemotaxis transducer protein BdlA (biofilm dispersal locus A) mediates biofilm dispersal (159). This was further validated by six-fold higher levels of c-di-GMP than a mature biofilm observed in the *bdlA* mutant. Li and coworkers observed a markedly distinct virulence phenotype of dispersed cells as opposed to planktonic and biofilm cells. The *AbdlA* ($\Delta 1423$) and *AdipA* ($\Delta 5017$) strains that are deficient in dispersion exhibited weakened virulence and enhanced persistence phenotypes. This suggests that differences in virulence gene expression may be a mechanism by which dispersed cells evade host immune responses temporarily before reverting to a planktonic phenotype (160). The DGC NicD is also shown to be activated upon sensing glutamate induced dispersion leading to BdlA phosphorylation which in turn stimulates DipA PDE activity (95). In addition, the c-di-GMP degrading GGDEF/EAL domain protein RbdA (PA0861) has also been implicated in biofilm dispersal. The heme-binding PAS domain in RbdA tightly controls its phosphodiesterase activity and thus also plays a critical role in mediating biofilm dispersal (96). The precise role of another composite protein NbdA (PA3311), containing NO-sensing MHYT domains, was established in NO-mediated biofilm dispersal, where *AnbdA* strain failed to disperse in the presence of NO (100).

Regulation of virulence

Virulence and formation of biofilms are coupled processes in *P.aeruginosa* implying the role of c-di-GMP in virulence. c-di-GMP negatively impacts the type III secretion systems (118, 161). Deletion of HD-GYP domains inhibits ExoS effector protein secretion (118). Screening of *P. aeruginosa* mutant strains for their differential levels of cytotoxicity towards CHO cells revealed that PDE transposon mutants such as RocR, FimX and PvrR lacked cytotoxicity while DGC transposon mutants such as WspR displayed partial cytotoxicity. This observed pattern might be due to the fact that cytotoxicity is indirectly controlled by c-di-GMP, positively influencing Type III secretion system and negatively influencing type IV pili (75).

Chapter 1

The composite GGDEF/EAL protein MorA (PA4601) has been demonstrated to control protease secretion post-translationally through the Type II secretion system (T2SS) using its domains involved in c-di-GMP signalling in *P. aeruginosa* PAO1. The *morA* insertion mutant displayed a marked increase in the secretion of five proteases, especially elastase LasB, whose secretion increased by 40% compared to the wild type. Furthermore, this mutant also showed increased invasion efficiency in the lung fibroblast cell-line (MRC-5) compared to the wild type (162). The expression of the type Vb secretion system, involving the adhesin CdrA in *P. aeruginosa* was found to be controlled by c-di-GMP via the GGDEF-EAL composite protein LapD (PA1433) and an associated protease LapG (105).

Additionally, the DGC YfiN or TpbB that forms part of the YfiBNR system targeting exopolysaccharide production is also involved in *P. aeruginosa* virulence (163). YfiN activity is tightly repressed by the periplasmic protein YfiR by binding to YfiN via a hydrophobic interaction site. Loss of YfiR control leads to YfiN activation that in turn promotes the SCV (small colony variant) phenotype in *P. aeruginosa*. This morphotype is characterized by hyperattachment, diminished motility, hyperwrinkled colony morphotype and slow growth rates. SCV confers greater resistance to scavenging by nematodes and macrophage-mediated phagocytosis in *P. aeruginosa*. This resistance phenotype can be attributed to the protection given to individual cells by the production of copious amounts of exopolysaccharides. The persistence of *P. aeruginosa* infection over several weeks due to SCVs was demonstrated successfully in mouse models. This establishes the importance of YfiBNR system in infection persistence and associates the SCV phenotype triggered by YfiN to *P. aeruginosa* chronic infections (164).

Regulation of exopolysaccharide production

Exopolysaccharides form a vital part of the biofilm matrix in *P. aeruginosa* (165-168). Three major exopolysaccharides are synthesized by *P. aeruginosa* namely Pel, Psl and alginate (135, 169-174). Pel polysaccharide is cationic in nature comprising of acetylated linkages of 1-->4 glycosidic linkages of N-acetyl glucosamine (GlcNAc) and N-acetyl galactosamine (GalNAc). De-N-acetylation of GlcNAc and GalNAc was found to confer a positive charge to the Pel polysaccharide. Pel polysaccharide was also shown to co-localize with extracellular DNA (eDNA) in the stalk of biofilms and also to be capable of binding DNA and other anionic substances by ionic binding (175). Pel production is positively regulated by c-di-

Chapter 1

GMP at the transcriptional level by activation of the response regulator-DGC WspR (86) and that of Pel, Psl and alginate at the post-transcriptional level by the diguanylate cyclase SadC (90, 92). This effect caused by SadC is counteracted by the phosphodiesterase BifA (89). Moreover, the wrinkled colony morphology phenotype observed in the *sadC* mutant is due to the production of the Pel exopolysaccharide (86, 174). FleQ, the transcriptional regulator, upon binding c-di-GMP, promotes the transcription of *pel* and *psl* genes. In the absence of c-di-GMP, FleQ forms a complex with the ATP binding protein FleN that binds to two sites near the *pel* promoter and inhibits its transcription (148, 176). Additionally, the cationic Pel polysaccharide (175) is also regulated at the post-translational level by PelD, a degenerate GGDEF c-di-GMP receptor, which forms a part of the *pel* operon (137).

Unlike Pel, Psl polysaccharide is only synthesized in PAO1 but not in the PA14 strain of *P.aeruginosa*. Psl consists of a repeating pentamer glucose, rhamnose and mannose and is encoded by the *psl* gene locus (166, 177, 178). The Psl biosynthesis was suggested to occur by a Wzx/Wzy – dependent mechanism facilitated by the hydrolase PslG (178). Extracellular DNA (eDNA) has also been observed to interact with Psl polysaccharide forming a web-like structure in pellicle (air-liquid interface) and flow-cell biofilms (179). Irie and colleagues demonstrated that two DGCs SadC and SiaD play a role in regulating Psl production in *P. aeruginosa*. A positive feedback circuit ensues when elevated levels of c-di-GMP produced by the two DGCs leads to increased Psl production promoting biofilm formation (180).

Alginate is an anionic polysaccharide consisting of L-glucuronic acid and D-mannuronic acid that helps long-term persistence of cystic fibrosis (CF) infections in lungs. The conversion of non-mucoid to mucoid phenotype in *P. aeruginosa* during late stages of CF infections is due to the overproduction of alginate (181-184). The DGC SadC along with a hydratase OdaA and dioxygenase OdaI control to production of alginate. The c-di-GMP produced by SadC is regulated in an oxygen-dependent manner (185). Moreover, the PilZ domain-containing Alg44 was shown to act in concert with the glycosyltransferase Alg8 to mediate alginate polymerization in *P. aeruginosa* (186). Further studies revealed that a multiprotein complex consisting of Alg8, Alg44, AlgG (an epimerase) and AlgX (an acetyltransferase) play a concerted role in alginate polymerization as well as its modification (187). The induction of alginate production under anaerobic conditions is mediated by Alg44 (71, 134, 135, 186). Overexpression of few other diguanylate cyclases such as PA1107 and PA1120 and

Chapter 1

phosphodiesterases PA2133 and PA2200 have also shown to upregulate and downregulate alginate biosynthesis respectively (135). Additionally, the GGDEF/EAL protein MucR (PA1727), an active DGC, generates a local c-di-GMP pool and regulates alginate biosynthesis. The $\Delta mucR$ strain has a non-mucoid phenotype in the *P. aeruginosa* strain PDO300 that over-produces alginate. Complementation with *mucR* gene resulted in the restoration of alginate production by ~4.4 fold (99).

1.6. Regulation of cellular processes in *Pseudomonas aeruginosa* by two component signalling (TCSs) and quorum sensing (QS) systems

In addition to c-di-GMP dependent regulation, cellular processes in can be controlled by signalling systems such as two-component (TCS) and quorum sensing (QS) systems.

Two-component signalling systems

The control of the switch of *P.aeruginosa* from one lifestyle to another is mediated by regulatory systems in response to environmental stimuli. Approximately 8.4% of genes in *P. aeruginosa* are regulatory in nature, the highest among the predicted sequenced bacterial genomes (7). The majority of these regulators in *P.aeruginosa* are the two-component signalling systems (TCS) (188-193). A prototypical two component system comprises of a sensor kinase (first component) that senses and responds to extracellular signals, by undergoing autophosphorylation using ATP at a conserved histidine residue. Subsequently, it modulates the phosphorylation state of its cognate response regulator (second component) carrying an output domain at an aspartate residue mediating a specific cellular response (192, 194). The sensor histidine kinases are integral membrane proteins, consisting of a periplasmic N-terminal input domain that senses environmental stimuli and a cytosolic C-terminal output or transmitter domain, which undergoes autophosphorylation upon receiving the environmental stimulus. The transmitter domain exhibits a great degree of sequence conservation and has an invariant histidine residue that undergoes autophosphorylation using ATP and a conserved motif (NG1FG2) consisting of two glycine residues that can bind ATP (195). On the other hand, response regulators consist of an N-terminal receiver domain and a DNA binding C-terminal output domain characterized by a helix turn helix motif (189). Thus, molecular communication between the sensor protein and response regulator is mediated by a His to Asp phosphorelay (196).

Chapter 1

Although histidine kinases (HKs) are similar to Ser/Thr/Tyr kinases catalytically, phosphorylation by HKs generates phosphoramidates as opposed to phosphoesters in case of Ser/Thr/Tyr kinases. Phosphoramidates have higher negative free energy of hydrolysis than phosphoesters, which makes them well suited for phosphotransfer reactions (197-199). Phospho-His bonds exhibit acid lability and alkaline stability which aids their detection in proteins (200). His phosphorylation occurs on the N1 or N3 position on the histidine imidazole ring. Although both forms exist in proteins, the characterized HKs have only found to be phosphorylated at the N3 position (N3-phospho-His) (198, 199, 201, 202). On the contrary, aspartate phosphorylation in response regulators (RRs) generates a high energy acyl phosphate (203). Phospho-Asp has the ability to mediate long-range conformational changes which occurs in case of RR. Conformational changes in proteins are driven by harvesting the energy from the acyl phosphate bond (204). Phospho-Asp bonds in proteins are relatively difficult to detect and have been identified only in case of few RRs (205, 206). The half life of the phosphorylated-Asp protein is often limited by the autophosphatase activity of RRs (207).

Sensor histidine kinases can be classified into three subcategories namely (1) Classical sensor kinases (2) Unorthodox sensor kinases and (3) Hybrid sensor kinases (Fig. 7). Classical sensor kinases involve a direct His to Asp phosphorelay between the two components. However, unorthodox and hybrid sensor kinases involve an additional receiver domain with conserved aspartate residue at the C-terminal of the transmitter domain. This necessitates the need for an intermediate component called histidine phosphotransfer module (Hpt) to mediate the phosphorelay to the response regulator. In case of unorthodox sensors, the Hpt is part of the sensor protein whereas in hybrid sensors, the Hpt module exists as an independent protein for signal transmission. The hybrid sensor kinases are also known as ITR-type histidine-kinases (input-, transmitter- and receiver-domains) (194, 208). Overall, a multi-step phosphorelay ensues in case of unorthodox and hybrid sensor kinases.

In *P.aeruginosa* PAO1, 127 TCS family members have been identified, consisting of 63 histidine kinases and 64 response regulators (195, 209). Compared to other sequenced bacterial genomes, *P. aeruginosa* has the highest number of TCS systems. The different TCSs are involved in the regulation of multitude of cellular functions, some of which are described below. Although, most of them possess classical sensor kinases, there are certain

Chapter 1

exceptions that have an additional C-terminal output domain (210). Few of the TCSs recognized in *P.aeruginosa* include those that regulate association of extracellular appendages such as type IV pili and Cup fimbriae, exopolysaccharide production, biofilm development and maturation. A series of TCSs have a role in regulation of biofilm determinants by ensuring the expression of the pertinent components at specific stages during biofilm development. The Roc1 system (regulator of cup) was one of the first TCS described in *P.aeruginosa* PAK strain wherein Kulasekara and co-workers found that the two cup gene clusters *cupB* and *cupC* was co-regulated by the Roc1 locus consisting of one sensor kinase RocS1 and two response regulators RocA1 and RocR (111, 211).

A CLASSICAL SENSORS

B UNORTHODOX SENSORS

C HYBRID SENSORS

INPUT TRANSMITTER RECEIVER HPT RECEIVER OUTPUT

HISTIDINE KINASE

RESPONSE REGULATOR

Figure 7 | Types of two-component signalling pathways in bacteria. (A) Classical sensor pathway involving His to Asp phosphorelay. (B) Unorthodox sensor pathway involving His to Asp to His to Asp phosphorelay and (C) Hybrid sensor pathway involving His to Asp to His to Asp phosphorelay with the Hpt domain protein acting as an intermediate transducer.

The *cup* gene clusters (*cupA-E*) encode extracellular appendages that contribute to biofilm formation and modify the adhesion properties of strains that express cup fimbriae (212, 213).

Chapter 1

This same gene cluster in PA14 was named as *sadARS* and was found to regulate biofilm maturation (214). The type IV pili gene expression is modulated by a TCS system consisting of PilS (sensor kinase) and PilR (response regulator) (215) while twitching motility is regulated by a second TCS proteins, FimS (sensor) and AlgR (response regulator) (216).

The switch between cytotoxicity and biofilm formation in *P.aeruginosa* is controlled by the RetS (regulator of exopolysaccharides and type III secretion sensor) TCS pathway (217, 218). RetS is a hybrid sensor kinase that targets gene clusters encoding a type VI secretion system (T6SS) called H1-T6SS (219). While the T6SS system is greatly activated in chronic *P.aeruginosa* infections especially in cystic fibrosis patients, the Type III secretion system (T3SS) plays a critical role in acute infections (161, 219, 220). Another hybrid sensor kinase LadS (lost adherence sensor) acts antagonistically to RetS by upregulating the T3SS genes (221). Being a hybrid sensor kinase, LadS requires an intermediate phosphorelay protein to relay signals to its cognate response regulator. Recently, it has been demonstrated that LadS mediates phosphorelay by binding to the Hpt domain of the histidine kinase GacS (222). The GacS/GacA TCS regulates virulence and biofilm formation in *P.aeruginosa* by enhancing the activity of small RNAs RsmY and RsmZ (223-225) RetS, however does not require Hpt for phosphorelay (226). Interestingly, RetS and LadS pathways also target these small RNAs serving as a link between the different TCS pathways controlling biofilm formation in *P.aeruginosa* (221). Further, the intermediate hybrid sensor kinase protein HptB (PA3345) was found to intersect the GacA/GacS pathway impacting sRNA expression by a complex route involving a response regulator PA3346 (218, 227). Hsu and co-workers identified three additional hybrid sensors PA1611, PA1976 and PA2824 that mediate phosphorelay to PA3346 via HptB (228, 229). Thus, the T3SS and T6SS systems are inversely regulated by the complex RetS/LadS/GacS TCS systems that act as checkpoints facilitating change from a planktonic to a sessile or biofilm lifestyle (211).

Quorum sensing systems

Many of the virulence factors produced by *P.aeruginosa* can be regulated by cell-cell communication (quorum sensing or QS) wherein small molecules known as autoinducers synchronize bacterial population behaviour in a density dependent manner (230-232). Two well-studied QS systems are the *las* and *rhl* systems. The *las* system comprises of the regulatory protein LasR and the autoinducer compound N-(3-oxododecanoyl) homoserine

Chapter 1

lactone (3O-C₁₂-HSL) whose expression is regulated by the *lasI* gene whereas the *rhl* system is composed of the regulatory Rhl protein and the autoinducer molecule N-butyryl homoserine lactone (C₄-HSL). Both the transcriptional regulatory proteins LasR and RhlR are highly specific for their autoinducers, containing a binding site for the autoinducer molecule at their N-terminus and a DNA binding site at their C-terminus (233, 234). The *las* signalling system regulates LasB elastase expression that drives the production of other virulence factors such as LasA protease and exotoxin A (235-238). The *las* system is in turn controlled by GacA and Vfr, that is crucial for *lasR* transcription (239, 240).

On the other hand, the expression of *rhlAB* operon encoding a rhamnosyltransferase essential for rhamnolipid synthesis is controlled by the *rhl* system (234, 241-245). This system is also involved in regulation of other virulence factors such as pyocyanin, elastase, cyanide and alkaline protease (234, 239, 246, 247). The expression of *rpoS*, encoding a sigma factor (δ^S) controlling stress response genes is also governed by the *rhl* system (248, 249). Overall, the two cell-cell signalling systems are not completely independent; the *las* system controls the *rhl* system at the transcriptional and post-translational level in a hierarchy cascade (248, 250). This was demonstrated from a study where 3O-C₁₂-HSL could block C₄-HSL from binding to its transcriptional activator *rhlR* (250). Thus, cell-cell signalling mechanisms allow *P.aeruginosa* to evade host defenses by the concerted expression of virulence factors governed by bacterial density eventually leading to tissue destruction and colonization.

Besides the *rhl* and *las* quorum sensing systems, there is a *Pseudomonas* quinolone signalling (PQS) system in which alkyl quinolones (2-heptyl-3-hydroxy-4-quinolone) act as signalling molecules. The PQS system acts as a link between the *rhl* and the *las* systems and does not rely on cell-density (251). The PQS regulon is composed of nine regulatory genes (83, 252). The *pqsABCDE* operon transcription was demonstrated to be controlled by the transcriptional regulator PqsR that binds to the *pqsA* promoter that increases in the presence of the quinolone signal molecule. *pqsR* is in turn positively controlled by the *las* QS system and negatively by the *rhl* QS system (253, 254). PQS system has been established to play a role in mediating *P. aeruginosa* acute infections and virulence (255).

Thus, *P. aeruginosa* quorum sensing network comprises of three intertwined quorum-sensing systems that have a major influence on several functions in this pathogenic bacterium.

Chapter 1

Overall, quorum sensing has also been found to be crucial for biofilm formation (47). One of the major role of the cell-cell signalling systems in biofilm development includes protection, resistance to hydrogen peroxide by regulation of superoxide mutases (256) and antimicrobials like tobramycin (47, 257). These studies indicate that QS has a vital role in spatially and temporally influencing biofilm development as well as in regulation of virulence factors.

1.7. Objectives

Thus far, considerable progress has been made towards understanding the regulation of cellular processes by c-di-GMP and other signalling systems in *P.aeruginosa*. Despite the immense knowledge regarding c-di-GMP metabolizing enzymes and receptors, the molecular mechanism by which specificity in c-di-GMP signalling pathways is attained remains largely unknown. Much needs to be discovered in this regard, especially how do complex regulatory networks confer specificity to c-di-GMP signalling. Also, structure-function relationships in c-di-GMP binding proteins needs to be investigated in detail to fully understand the specificity of c-di-GMP signalling. Another important challenge is to investigate the molecular mechanisms by which c-di-GMP binding proteins mediate specific functions in *P.aeruginosa* such as biofilm development, motility, chemotaxis and switch in lifestyle in response to environmental cues. The mechanism by which c-di-GMP controls specific cellular functions can be understood by identifying a putative protein partner of a c-di-GMP binding effector protein. Further, the role of the c-di-GMP effector can be investigated by studying the cellular pathway involving its protein partner.

With these questions in mind, the first aim of this thesis was to investigate a c-di-GMP binding protein in *P.aeruginosa* of previously unknown function. Being the most predominant class of c-di-GMP binding proteins, the protein of interest chosen in this study is a stand-alone PilZ domain protein PA2799. The cellular function of PA2799 was revealed by identifying its protein partner, a sensor histidine kinase (SagS), which forms part of a two-component signalling pathway, regulating biofilm formation. It was also revealed that the interaction between PA2799 and SagS occurs in a c-di-GMP dependent manner. Further investigations on this two-component signalling pathway shed light on the regulatory functions of PA2799 in biofilm formation. On the basis of its function, PA2799 was renamed as HapZ (Histidine kinase associated PilZ). This is the first study to our knowledge reporting

Chapter 1

the function of a stand-alone PilZ domain protein HapZ that functions as a c-di-GMP adaptor regulating biofilm formation.

The second aim of this thesis was to focus on the protein-protein interaction between the phosphoreceiver domain of the sensor histidine kinase SagS, with both HptB (its downstream protein partner) and HapZ (a PilZ protein). The histidine phosphotransfer protein HptB is known to participate in two-component signalling pathways by interacting with multiple sensor histidine kinases in *P. aeruginosa*. Studying the interaction of HptB with the phosphoreceiver domain of SagS would provide insight into how the promiscuous HptB mediates specific binding with SagS and three other hybrid kinases in a complex signalling pathway. Further, the molecular basis of the involvement of a PilZ protein (HapZ) in the SagS-HptB signalling pathway was understood by investigating the binding specificity between HapZ and its cognate protein partner SagS. This can provide the basis to study interactions of other PilZ domain proteins and their protein partners. Since the HapZ-SagS interaction is dependent on c-di-GMP, investigating the mechanism by which c-di-GMP binds to HapZ would also improve our knowledge about this interaction mechanism. To answer these questions, we elucidated the NMR solution structure of REC_{SagS} and studied its interaction with HptB using HSQC titration. Further, we generated mutant libraries of HapZ and REC_{SagS} to identify the critical residues involved in the HapZ-REC_{SagS} interaction. These studies revealed the nature of binding of REC_{SagS} with HptB and HapZ. We came to understand how REC_{SagS} interacts with both HptB and HapZ in a unique and specific manner. This interaction specificity provided an understanding of how cross talks between different proteins are avoided within a complex two-component signalling network in microorganisms such as *P. aeruginosa*.

CHAPTER 2: Regulation of two-component signalling by c-di-GMP through a PilZ adaptor protein

2.1. Introduction

C-di-GMP is recognized as a central regulator influencing a broad spectrum of processes in many environmental and pathogenic bacteria. The transition from a motile to a sessile lifestyle leading to biofilm formation in bacteria is generally accompanied by a rise in c-di-GMP concentrations. The cellular concentration of c-di-GMP fluctuates in response to environmental changes owing to the contrasting activities of c-di-GMP diguanylate cyclases and phosphodiesterases. C-di-GMP mediates lifestyle changes in bacteria by binding to multitude of receptors ranging from PilZ domain proteins to riboswitches. PilZ domain proteins were the first to be discovered and represent the most predominant class of c-di-GMP receptors among others (121). PilZ proteins regulate different pathways by displaying a broad range of binding affinities for c-di-GMP (258, 259). Pultz and co-workers observed a 145-fold and a 43-fold difference in binding affinities towards c-di-GMP in case of the PilZ domain proteins in *P. aeruginosa* and *Salmonella typhimurium* respectively (258).

Free- standing PilZ domain proteins consist of a single PilZ domain (90-130 residues) and represent greater than half of the known PilZ domain proteins. The other known PilZ proteins are either di-domain or multi-domain proteins that contain a functional domain coupled to a PilZ domain. Thus far, several studies have investigated the functions of di-domain and multi-domain PilZ proteins. The di-domain PilZ protein YcgR in *E. coli* interacts with FliG and FliM, the rotor proteins in the flagella, in a c-di-GMP dependent manner. This interaction promotes a counter-clockwise motor bias affecting motility and chemotaxis in *E. coli* by a 'backstop-brake' mechanism (123, 260). The cellulose synthase subunit BcsA, containing a PilZ domain was shown to form a complex with BcsB and activate cellulose synthesis in the presence of c-di-GMP in *Rhodobacter sphaeroides* (261, 262). Similarly, the PilZ domain protein Alg44 was demonstrated to regulate alginate biosynthesis by binding to c-di-GMP in *P. aeruginosa* (71). Another well documented PilZ protein is the c-di-GMP dependent transcriptional regulator MrkH from *Klebsiella pneumoniae* which binds to a cognate DNA sequence activating MrkH expression that in turn activates the *mrkABCDF* operon resulting

Chapter 2

in increased type 3 fimbriae biosynthesis and surface attachment (263, 264). The microaerophilic bacterium *Azospirillum brasilens* adapts to varying oxygen levels by increasing its swimming velocity and decreasing flagellar reversal frequency in a c-di-GMP dependent manner. The chemotaxis and aerotaxis receptor, Tlp1 of the bacterium bears a PilZ domain that binds to c-di-GMP in response to oxygen concentrations and aids the bacterium to navigate to low oxygen niches (265). A unifying theme that has surfaced from these findings is that the PilZ domains regulate a multitude of cellular processes by binding to c-di-GMP and initiating an intra-protein conformational change to alter the enzymatic activity or binding properties of the corresponding output domains.

Although free-standing PilZ domain proteins have been implicated in pathogenesis, the function and mechanisms of these proteins have not yet been well studied as opposed to di-domain and multi-domain PilZ proteins. It was found that the free-standing PilZ proteins XC0965, XC2249 and XC3221 regulate motility, virulence and extracellular enzyme production in the plant pathogen *Xanthomonas campestris* pv. *campestris* (266, 267). Further, it was also found that the stand-alone PilZ protein PlzA binds to c-di-GMP, controls the expression of a stationary-phase global regulatory sigma factor RpoS (σ^S) and regulates motility and virulence in *Borrelia burgdorferi* (268, 269). There exists a dearth in experimental evidence supporting the broadly established assumption that free-standing PilZ proteins function as c-di-GMP adaptors by binding to their protein partners in a c-di-GMP-dependent mode. This proposition was even challenged when the PilZ protein XC_2249 from *X. campestris* pv. *campestris* formed a complex with an HD-GYP (RpfG) and GGDEF domain and influenced pilus action in a c-di-GMP independent manner (270).

The *P. aeruginosa* PAO1 c-di-GMP signalling network comprises of at least 12 c-di-GMP effector proteins. These include FimX, FleQ, BrlR, PelD, PP4395 and seven PilZ-domain proteins. FimX, a c-di-GMP receptor possessing both GGDEF and EAL domains governs twitching motility in response to environmental cues (103, 143, 154). The transcription regulator FleQ binds c-di-GMP to modulate flagellar biosynthesis and the production of extracellular polysaccharide Pel (148, 271) while the transcription regulator BrlR is a c-di-GMP receptor that contributes to antibiotic resistance functions in activator of multidrug transport in *P. aeruginosa* (150, 152). The degenerate GGDEF PelD receptor protein binds c-di-GMP, affects pellicle formation and regulates Pel polysaccharide production (138, 139). The c-di-GMP-binding PA4395, a member of the YajQ family proteins, controls gene

Chapter 2

transcription and virulence by interacting with a transcriptional regulator (272). The didomain PilZ protein Alg44, composed of an N-terminal PilZ domain and a C-terminal NolF domain is essential for the synthesis of alginate. The PilZ domain of Alg44 was shown to bind c-di-GMP both *in vivo* and *in vitro*. Moreover, Alg44 was found to The NolF domain contributes to the transport of alginate by linking the membrane and cytoplasmic parts of a transport apparatus. Single amino acid substitutions that abolished c-di-GMP binding to Alg44 affected alginate production (135). Recently, the role of the di-domain PilZ protein FlgZ (PA3353) in the regulation of swarming motility was established. FlgZ, a homologue of YcgR in *E. coli*, was shown to interact with the stator protein MotC in a c-di-GMP dependent manner and inhibit swarming motility in *P. aeruginosa* PA14 (136, 260).

Besides the eight proteins with known function, there are five c-di-GMP-binding PilZ proteins whose functions remain to be elucidated. Among the five PilZ proteins, PA0012, PA2799, PA4324 and PA4608 are free-standing PilZ proteins while PA2989 is a di-domain protein (260). Investigating the physiological role of these five PilZ proteins will substantially enhance our knowledge regarding c-di-GMP signalling in *P. aeruginosa*. Towards achieving this goal, we explored the function of the stand-alone PilZ domain protein PA2799 in *P. aeruginosa*.

2.2. Materials and Methods

Construction of P. aeruginosa genomic DNA libraries

For genomic library construction, the BamHI site was used for vector digestion and Sau3AI site for gDNA partial digestion. The vector pTRG of the BacterioMatch II Two-Hybrid system (Stratagene) was used for prey library construction. Three independent genomic DNA libraries were constructed by using the vector pTRG, and the engineered pTRGderived pTRG-A2 and pTRG-AC1. The plasmids pTRG-A2 and pTRG-AC1 were constructed by inserting the nucleotide “A” and “AC” between the 3’-end of the RNAP-alpha (992bp) gene and the BamHI site (993bp) of respectively. The insertion resulted in a shift of the polylinker site by +1 or +2 nucleotides. The vectors pTRG, pTRG-A2 and pTRG-AC1 were digested by BamHI and dephosphorylated with phosphatase. The vector digestion mixture was extracted with phenol: chloroform, precipitated and re-suspended into 1X TE buffer. The genomic DNA of *P. aeruginosa* strain PAO1 was isolated using the chloroform-

Chapter 2

cetyltrimethylammonium bromide extraction method, and partially digested by *Sau3AI*. The randomly digested fragments ranging in size from 1,000 bp to 3,000 bp were purified using NucleoSpin® Gel and PCR Clean-up kit (MACHEREY-NEGEL). Gel-purified DNA fragments were ligated to the linearized and dephosphorylated pTRG, pTRG-A2 or pTRG-AC1 vectors. The resulting ligation mixtures were precipitated and resuspended into 1XTE buffer, followed by transformation into *E. coli* XL1-Blue MRF' Kan competent cells using a standard electroporation procedure. The three libraries were collected and pooled separately as the prey libraries and stored in -80 °C freezer.

Bacterial two-hybrid screening

The plasmid pBT-PA2799 was used as the bait to probe the genomic DNA libraries using the BacterioMatch II Two-Hybrid system (Stratagene). The plasmids extracted from the cells harboring the library fragments were co-transformed into the reporter strain with the bait plasmid by electroporation. The co-transformed cells were grown on M9⁺ His-deficient medium containing 5 mM 3-AT for 48-72 h at 30 °C. Colonies that grew on these plates were selected as positive colonies. The equivalent of 10⁶ colonies were screened, as assessed by plating the transformed cells on M9⁺ His-deficient medium. Positive colonies were subsequently picked and re-streaked on M9⁺ His-deficient medium containing 5 mM 3-AT and 12.5 µg/ml streptomycin. The colonies that grew on the double selection medium were cultured in liquid medium, and used for colony PCR with pTRG F/R pairs to amplify the insert from the library plasmid. The positive colonies were grown for the preparation of plasmids for sequencing. For the specific two-hybrid binding assay, the bait and prey plasmids were co-transformed and the co-transformed mixtures were plated onto non-selective medium and incubated at 37 °C overnight. The growth of the selected co-transformed clones was verified by dotting on the selective medium by 37 °C overnight incubation. Normal growth on the selective screening medium was considered as an indicator of positive protein-protein interaction.

Co-immunoprecipitation (Co-IP) assay

To validate the interaction between HapZ and REC_{SagS} in *P. aeruginosa* by Co-IP assay, the two plasmids 6xHis-REC_{SagS}, pUCP-HA-HapZ/6xHis-REC_{SagS} (Table S2) were constructed and transformed into PAO1 cells. The strain harboring pUCP-6xHis-REC_{SagS} consistently

Chapter 2

expressed protein 6xHis-REC_{SagS} while the strain harboring pUCP-HA-HapZ/6xHis- REC_{SagS} produced both HA-HapZ and 6xHis-REC_{SagS}. The bacterial strains were cultured in LB containing 250 µg/ml carbenicillin at 37 °C overnight with shaking and subsequently sub-cultured to an absorbance (600 nm) of 1.0 at 37 °C. Cells were harvested by centrifugation (13,000 g for 10 min) and lysed with ice-cold cell lysis 1X PBS buffer [10 mM Sodium phosphate (pH 7.4), 137 mM NaCl and 2.7 mM KCl, 0.5% NP-40, 10% glycerol, 1 mM EDTA and protease inhibitor cocktail (Roche)]. The total cell lysates were clarified by centrifugation at 15,000 g for 10 min, and the supernatant was preclarified with protein G beads after being filtered by Minisart® high flow Syringe Filters. Lysates was subsequently immunoprecipitated with the protein A-Sepharose beads (EZview™ Red Anti-HA Affinity Gel, Sigma-Aldrich). Each immunoprecipitation sample containing 800 µl lysate, 20 µl of the prewashed beads with different concentrations of nucleotides (c-di-GMP, GMP or cGMP) was incubated for 1 h at 4 °C. The beads were collected by centrifugation, and washed six times with lysis buffer at 4 °C, and then the immunoprecipitated proteins were diluted with 2X sample buffer and subjected to SDS-PAGE and Western Blotting with HA antibody.

Cloning, expression and purification of recombinant proteins

The recombinant proteins HapZ, HptB and SagS₂₄₆₋₇₈₆ were expressed using *E. coli* Rosetta competent cells. Transformed cells were grown at 37°C to an absorbance (600 nm) of 0.6 in LB media containing 50 µg/ml kanamycin and induced for protein expression. Protein expression was induced using 0.5 mM isopropyl β-D-thiogalactopyranoside (IPTG) (16 °C). After 16 h, cells were harvested by centrifugation and re-suspended in Phosphate buffer pH 8.0 and lysed using French Press homogenizer at 15,000 psi. The cells were then centrifuged at 25,000 rpm for 30 minutes and supernatants were filtered and loaded onto a nickel-charged column equilibrated with KPi buffer (pH 8.0). The column was washed sequentially using W1 (KPi buffer pH 8.0 + 20 mM Imidazole) and W2 (KPi buffer pH 8.0 + 50 mM Imidazole) buffer. The 6xHis-tagged proteins were eluted from the column using elution buffers (KPi buffer pH 8.0) containing increasing concentrations of imidazole (100, 200 and 300 mM) and were analyzed using SDS-PAGE. The fractions containing the desired proteins were pooled with a purity > 95%, desalted using PD-10 column (GE Healthcare) and concentrated using Amicon concentrator (10 kDa for HptB and SagS₂₄₆₋₇₈₆ and 5 kDa for HapZ) at 4,000 rpm at 4 °C. The concentrations of the proteins were determined by Bradford method and were aliquot and stored at -80°C with the exception of HapZ, which was used without storage due

Chapter 2

to its propensity to aggregation.

In vitro phosphorylation assay

C-di-GMP was prepared using a thermophilic diguanylate cyclase as described previously (273, 274). Phosphorelay between SagS and HptB was examined by performing *in vitro* phosphorylation assay with 6xHis-tagged recombinant proteins. The reaction mixture that contains SagS₂₄₆₋₇₈₆ and HptB (1:1 ratio), ³²P-ATP and MgCl₂ was incubated at 25 °C for 30 min in reaction buffer, which is composed of 100 mM Tris-HCl (pH 8.0), 50 mM KCl, 5 mM MgCl₂, 1 mM DTT, 1 µCi of [γ -³²P] ATP (about 111 TBq/mmol, 10 mCi/ml, 3.33 µM, 3,000 ci/mmol). Phosphorylation of HptB was visualized by running 15% SDS-polyacrylamide gel electrophoresis and autoradiography. The effect of HapZ and c-di-GMP on HptB phosphorylation was examined by incubating the corresponding components with the reaction mixture.

Biofilm formation in flow-cell

The *P. aeruginosa* strains were tagged with GFP reporter by inserting the mini-Tn7-eGFP-Gmr, mini-Tn7-eCFP-Strepr and mini-Tn7-eYFP-Strepr cassette. Biofilms were grown in flow-chambers with individual channel dimensions of 1 x 4 x 40 mm. The flow-chambers were inoculated by injecting 350 µl overnight culture diluted to an absorbance (600 nm) of 0.01 into each flow channel using a small syringe. After inoculation, the flow channels were left without flow for 1 h, after which medium flow was started using a Watson Marlow 205S peristaltic pump. The mean flow velocity in the flow-chambers was 0.2 mm s⁻¹. All microscopic observations and image acquisitions were done with a Zeiss LSM780 confocal laser-scanning microscope (CLSM). Data were analyzed to generate the simulated three dimensional images using the IMARIS software package (Bitplane AG) (275).

RNA preparation and real-time PCR analysis

PAO1 cells were grown in AB minimal medium (1.51 mM (NH₄)₂SO₄; 3.37 mM Na₂HPO₄, 2.2 mM KH₂PO₄, 5 µM NaCl, 100 µM MgCl₂; 10µM CaCl₂; 0.1µM FeCl₃; 0.2% casamino acid, 0.2% glucose) overnight at 37 °C. The overnight cultures were diluted in AB minimal medium to an absorbance (600 nm) of 0.1. Aliquots of 1 ml of cultures were incubated in triplicates in 24-well plates at 37 °C with agitation till reaching the early stationary stage [Absorbance (600 nm) = 0.7-0.8]. The cells were harvested in RNAprotect Bacteria Reagent

Chapter 2

(Qiagen, Cat. 76506). The collected cells were re-suspended in TE buffer (10 mM Tris-Cl, 1 mM EDTA, pH 8.0) that contains 15 mg/ml lysozyme and incubated at room temperature for 10 min. TRIzol® Reagent (life technology, Cat. No. 15596-026) were added to the samples and the total RNA including small RNA was isolated with Direct-zol™ RNA MiniPrep (ZYMO RESEARCH). Total RNA were further digested with DNase I, and purified with RNeasy Plus Mini Kit (Qiagen). The purity and concentration of the RNA were determined by NanoDrop spectrophotometry. Elimination of contaminating DNA was confirmed via real time PCR amplification of the *cafA* gene with total RNA as template.

Quantitative reverse transcriptase PCR (qRT-PCR) was performed by a two-step method. First-strand cDNA was synthesized from total RNA by using iScript™ Select cDNA Synthesis Kit (Cat. 170-8896, Bio-Rad). The cDNA were used as template for qRT-PCR with a kit of KAPA SYBR® FAST qPCR Kit Master Mix) on an Applied Biosystems 7500 Real-Time PCR System. The gene *rplU* was used as endogenous control. Melting curve analysis was employed to verify specific single-product amplification.

Swarming and twitching motility assay

Swarming motility was analyzed on 0.5% agar M8 salt plates supplemented with 0.2% glucose, 2 mM MgSO₄ and 0.5% casamino acid. 1.5 µl of cell culture grown at 37°C overnight in LB medium were dotted on freshly prepared swarming plates and incubated for 16 hours at 37 °C. The PAO1 strain was included as a control on each swarming plate. For twitching motility assay, PAO1 cells were stab inoculated with a toothpick through a thin Difco LB agar (3 mm, 1 % Difco granulated agar) layer to the bottom of the Petri dish. After overnight growth at 37 °C, the zone of twitching was visualized by staining with Coomassie Brilliant Blue R250 and the radius of the zone was measured.

Bacterial tethering analysis

For cell-tethering analysis, glass coverslips were precoated with flagellar antibodies prior to use and cell chambers (2.0 cm by 1.0 cm by 150 µm) were created using three layers of double-sided tape between the microscope slide and coverslips. Flagella were sheared off by passing the bacterial cells through a 34-gauge blunt-end needle four times. Cells were loaded into the cell chamber, and nontethered cells were rinsed away using LB broth. Cells were

Chapter 2

visualized using an inverted microscope (Nikon TE2000U) under a 100× objective. Videos of tethered bacteria were taken at 120 frames per second (fps) for 1 to 5 min using a complementary metal-oxide semiconductor (CMOS) camera (Thorlabs; DCC1645) (276). Following the convention, cells are considered to be rotating CW/CCW when viewed from the medium that they are tethered in. Free-swimming *P. aeruginosa* cells, with their intact flagellum, were loaded in cell chambers and observed in the middle of the chamber depth (away from the coverslip or microscope slide surface). Cells were visualized under a 40× objective, and videos of bacteria were taken at 25 fps. Bacterial swimming trajectories in two dimensions (2D) were captured using Image Pro Plus 6.3 (Media Cybernetics, Rockville, MD, USA), and images with cell outlines were obtained using ImageJ (NIH, Bethesda, MD, USA).

Quantification of rhamnolipid production

The emulsification activity assay was modified from as previously described (277). 1 ml of filtered overnight culture of the different *P. aeruginosa* strains was mixed with 1 ml of hexadecane in a glass tube. The mixture was vortexed rigorously for 2 min, and was allowed to stand for 2 h in room temperature. Emulsification activity was measured as the height of emulsion layer divided by the total height of the mixture. Experiments were performed in triplicate, and the results are shown as the mean $\pm$ s.d.

Colony morphology assay

The DGC mutant strains used for the colony morphology assay were obtained from the University of Washington *P. aeruginosa* PAO1 transposon mutant library. The 25 strains have been listed in Table S1. The 25 DGC mutants as well as PAO1 and Δ *hapZ* strains were transformed with the pUCP18-HapZ overexpression plasmid by electroporation. The colonies were obtained by following the procedure described previously (278). Colony morphology agar plates were prepared by autoclaving 1% Tryptone with 1% Bacto agar. Coomassie blue (20 μ g/ml) and Congo red (40 μ g/ml) were added to the medium after cooling it to 60 °C with or without 250 μ g/ml carbenicillin. Plates were poured (35-40 ml agar each) and allowed to dry with closed lids for a period of 24 hours. Colonies picked from single streak plates were pre-cultured in LB medium for 12 hrs. Ten microliters of the pre-culture was spotted at the centre of each plate and allowed to dry for 15 minutes. The plates were parafilmed and

Chapter 2

incubated at 25 °C for 3 to 6 days. Colony images were taken by using Canon 1000D DSLR camera fitted with 18x55 mm lens.

Transcriptome analysis by RNA-Seq

Colonies of *P. aeruginosa* PAO1 and the two mutant strains were grown overnight in ABTGC medium at 37 °C. The cultures were diluted to a starting absorbance (600nm) of 0.01 and 1 ml culture was added to a well of a 24 well plate. The plate was sealed with parafilm and incubated at 37 °C to reach stationary phase. The cultures were mixed immediately with 2 volumes of RNeasy Protect® Bacteria Reagent (Qiagen). After 5 min incubation at room temperature, samples were centrifuged and the pellets were stored at -80 °C. Bacterial cells were treated with lysozyme and total RNA was extracted with RNeasy Mini Purification kit (Qiagen). Removal of DNA was carried out by on-column DNase digestion with the RNase-free DNase Set (Qiagen). The integrity of total RNA and DNA contamination was assessed and the 16S, 23S and 5S rRNAs were removed by using the Ribo- Zero™ Magnetic Kit (Epicentre). Gene expression analysis was performed in duplicate by RNA sequencing. The rRNA-depleted RNA was fragmented to 200-300 bp fragments and the first and second strand cDNA were synthesized, followed by end repair and adaptor ligation. The libraries were sequenced using the Illumina HiSeq2000 platform with a paired-end protocol and read lengths of 100 nt. The RNA-Seq data are available in the NCBI GEO Short Read Archives (SRA No: SAMN02369270). The sequence reads were assembled and analyzed in RNA-Seq and the expression analysis application of CLC genomics Workbench 6.0 (CLC Bio, Aarhus, Denmark). The following criteria were used to filter the unique sequence reads: minimum length fraction of 0.9, minimum similarity fraction of 0.8, and maximum number of two mismatches. Data were normalized by calculating the reads per kb per million mapped reads for each gene (Mortazavi et al., 2008). ANOVA and t-test were performed on transformed data to identify the genes with significant changes in expression ($P < 0.05$, Fold change ≥ 2.0 or ≤ -2.0).

2.3. Results

PA2799 interacts with the phosphoreceiver (REC) domain of the hybrid histidine kinase SagS.

It was proposed that c-di-GMP binding free- standing PilZ proteins function as adaptor

Chapter 2

proteins by binding to their protein targets. In order to identify the protein partner of PA2799, a customized bacterial two-hybrid screening system was developed with a high-coverage *P. aeruginosa* genomic DNA library. Construction of the *P. aeruginosa* PAO1 genomic DNA library used for bacterial two-hybrid screening was done using the pTRG vector and two engineered pTRG vectors (see Materials and Methods) to augment the number of open reading frames (ORFs) covered by the library.

Figure 8 | HapZ (or PA2799) interacts with SagS phosphoreceiver domain as observed in the bacterial two-hybrid screening. (A) Organization of different domains in the hybrid sensor kinase SagS, where DHp is the dimerization & histidine phosphotransfer domain, CA is the catalytic & ATP-binding domain and REC is the phosphoreceiver domain. The varying length constructs interacting with PA2799 are shown by bars. (B) PA2799 interacts with the varying length constructs of SagS encompassing the DHp, CA and REC domains. This is indicated by the robust growth of co-transformed colonies on the selective medium containing 5 mM 3-amino-1,2,4- triazole (3-AT) (ΔN317 indicates the fragment lacking the N-terminal 317 residues; ‘-’ represents the empty pBT vector; ‘+’ represents the pBT-HapZ vector). (C) The interaction between PA2799 and the phosphoreceiver domain of SagS is specific. This was demonstrated by a specific two-hybrid screening of PA2799 (HapZ) and REC_{SagS} with multiple phosphoreceiver domains and PilZ domain proteins in *P. aeruginosa* respectively

The comprehensive PAO1 genomic library consisted of DNA fragments ranging from 1-3 kb

Chapter 2

in length (Fig. S1). The bait plasmid was designed by fusing the PA2799 gene to the bacteriophage λ repressor protein gene. PA2799 gene was screened as the bait to probe for the prey protein encoded by the PAO1 genomic DNA library. After several cycles of optimization for reducing the false-positive colonies, consistent results were obtained from the screening with more than two third of the prey plasmids containing a DNA fragment from the gene PA2824. Although most of the false positives contained shifted ORFs, many of the clones contained a DNA fragment encoding a portion of the diguanylate cyclase WspR that incorporated the entire GGDEF domain. However, further analysis using a specific two-hybrid assay revealed that PA2799 and the full length WspR were not interacting partners. PA2824 encodes the protein SagS, a membrane-bound hybrid histidine kinase consisting of a sensor domain, a dimerization & histidine phosphotransfer (DHp) domain, a catalytic & ATP-binding (CA) domain and a phosphoreceiver (REC) domain (Fig. 8A).

The protein-protein interaction was validated and the domain of SagS involved in binding PA2799 was tested by performing specific bacterial two-hybrid binding assays using a construct containing the free-standing C-terminal REC_{SagS} domain. The apparent robust growth of the colonies on the selective medium (Fig. 8B) confirmed the interaction between PA2799 and the REC_{SagS} domain suggesting that PA2799 is likely to interact with SagS through its receiver (REC) domain. Since PA2799 is the first PilZ protein found to interact with a histidine kinase, we renamed it as HapZ (Histidine kinase associated PilZ protein).

The protein-protein interaction between HapZ and SagS is specific.

The histidine-containing phosphotransfer protein-B (HptB) has been well established as a downstream phosphoryl group-receiving partner of SagS in the two-component signalling pathway (228). However, three other hybrid sensor histidine kinases (PA1976, PA1611 and PA4856 or RetS) have also been found to interact with HptB for phosphotransfer in addition to SagS (208, 227-229). Further, *P. aeruginosa* PAO1 encodes many other PilZ domain-containing proteins along with HapZ. A 'cross-binding assay' was employed to test the specificity of binding between REC_{SagS} and HapZ wherein prey plasmids encoding the REC domains of PA1976, PA1611 and PA4856 and bait plasmids encoding the PilZ proteins were constructed. Of all the prey and bait proteins tested, the bacterial two-hybrid screening could detect only the interaction between HapZ and REC_{SagS} (Fig. 8C), verifying that HapZ is possibly the only PilZ protein that interacts with SagS and that the binding between HapZ

Chapter 2

and REC_{SagS} is extremely specific. This specific recognition between PilZ adaptors and its protein partners is vital for averting crosstalk across different signalling pathways.

HapZ-SagS interaction is enhanced by c-di-GMP

It has been demonstrated in a previous study that HapZ binds to c-di-GMP with a dissociation constant (K_d) of 2.0 μ M (258). The interaction between HapZ and REC_{SagS} and the effect of c-di-GMP on SagS-HapZ interaction was further confirmed using a co-immunoprecipitation (Co-IP) assay using *P. aeruginosa* PAO1 cell lysates. Plasmids were designed to express hemagglutinin (HA)-tagged HapZ and 6xHis-tagged REC_{SagS}. It was observed that HA-HapZ co-immunoprecipitated with 6xHis-REC_{SagS} validating the interaction between SagS and HapZ (Fig. 9A, 9B). The Co-IP assay also established that the interaction between SagS-HapZ is strengthened by c-di-GMP in a dose-dependent mode within the 0-20 μ M range (Fig. 9B). A little enhancement was observed above 20 μ M c-di-GMP, which is in agreement with the binding affinity ($K_d = 2.0 \mu$ M) of HapZ for c-di-GMP.

Figure 9 | c-di-GMP enhances HapZ-REC_{SagS} interaction. (A) Co-immunoprecipitation assay validates HapZ and REC_{SagS} binding. Western blot showing cell lysates of *P. aeruginosa* PAO1 expressing either pUCP18-6xHis-REC_{SagS} and both pUCP18-HA-HapZ/6xHis-REC_{SagS}. 6xHis-REC_{SagS} and HA-HapZ co-immunoprecipitated with anti-HA beads. Western blot was done using anti-HA and anti-6xHis antibodies for testing protein expression and immunoprecipitation in the cell lysates (B) The binding between HapZ and REC_{SagS} is enhanced by c-di-GMP. Co-immunoprecipitation assay showing the impact of increasing concentrations of c-di-GMP as well as other nucleotides (GMP, cGMP & c-di-AMP) on HapZ-REC_{SagS} binding.

The structurally related nucleotides namely GMP, cyclic GMP and cyclic di-AMP had negligible effect on SagS-HapZ interaction, emphasizing that the enhancing effect of c-di-GMP is relatively specific. The enhancing effect of c-di-GMP on SagS-HapZ binding is comparable to YcgR, a di-domain PilZ protein that interacts with the flagellar motor switch

Chapter 2

complex in *E. coli* (260, 279). PilZ proteins bind c-di-GMP by using several conserved residues from the β -barrel and the flexible unstructured N-terminal loop (71, 132, 262). Sequence alignment and structural modeling studies indicate that PA2799 is expected to bind c-di-GMP using an identical set of residues both from the β -barrel and the 22-residue N-terminal loop (Fig. 10A). The interaction between HapZ and REC_{SagS} was abolished upon mutation of a conserved Arg-12 residue in the RxxxR motif of HapZ involved in c-di-GMP binding (Fig. 10B). Similarly, binding between REC_{SagS} and the HapZ- Δ 21aa construct lacking N-terminal loop could not be detected. These findings suggest that both the N-terminal loop and c-di-GMP binding are crucial for the SagS-HapZ interaction.

Figure 10 | c-di-GMP binding is important for HapZ-REC_{SagS} interaction. (A) Homology model of HapZ generated using apo-PA4608 (PDB code: 1YWU) as the template in complex with dimeric c-di-GMP. The intercalated C-di-GMP dimer is shown as a stick representation. The conserved arginine residues R12 and R16 in HapZ are shown as pink sticks and its N-terminal loop is coloured pink. (B) Mutation of the conserved Arg-12 residue (HapZ-R12A) and deletion of the N-terminal loop of HapZ (HapZ- Δ 21aa) weakens HapZ-REC_{SagS} interaction.

Our efforts to determine the nature and binding affinity of SagS-HapZ interaction using isothermal titration calorimetry (ITC) and nuclear magnetic resonance (NMR) spectroscopy were unsuccessful due of the propensity of HapZ to form insoluble aggregates at high concentration ($>100 \mu\text{M}$), which is essential for performing *in vitro* binding assays. Several approaches for purifying HapZ using different protein or peptide tags and buffer conditions were futile. Similarly, visualization of possible co-localization of HapZ and SagS *in vivo* in *P. aeruginosa* was impeded by the fact that the addition of any protein or peptide epitope tag

Chapter 2

to the C-terminus of SagS and HapZ interfered with the interaction.

Phosphotransfer between SagS and the downstream protein HptB is blocked by HapZ in a c-di-GMP-dependent manner

Several studies have demonstrated that SagS initiates a phosphorelay cascade in *P. aeruginosa* with HptB and the bifunctional PA3346 (HsbR) protein as the immediate downstream protein partners (Fig. 11A) (227-229). This multistep signalling pathway involves activation and autophosphorylation of SagS upon receiving an environmental stimulus. Further, phosphotransfer occurs from SagS to HptB, the intermediate phosphotransfer protein. Phosphorylated HptB was shown to relay this signal to a cognate response regulator PA3346 (HsbR) at its N-terminal receiver domain leading to the activation of its Ser/Thr phosphatase activity. This ultimately resulted in the dephosphorylation of PA3347 at Ser 56. Previously, it was demonstrated that the phosphorylation of PA3347 was mediated by a kinase of unknown function and was coupled to the release of an anti- σ factor that expression of genes involved in swarming and biofilm formation in *P. aeruginosa* (208).

Interaction of HapZ with the two-component sensor kinase SagS implies that HapZ can potentially regulate the SagS signalling pathway by directly binding to REC_{SagS}. The involvement of HapZ and c-di-GMP in the SagS-HptB signalling pathway was evaluated by performing an *in vitro* phosphorylation assay. When the 6xHis-tagged recombinant proteins SagS₂₄₆₋₇₈₆ and HptB (Fig. 11B and S4) were incubated in a 1:1 ratio in the presence of [γ -³²P]-ATP, phosphorylation of HptB could be detected (Fig 11C). The level of HptB phosphorylation reduced upon the addition of freshly prepared HapZ in the reaction mixture resulting in a final protein ratio of 1:1:1 (SagS/HptB/HapZ). The phosphorylation level of HptB was further decreased by c-di-GMP in a dose dependent manner (Fig. 11C). This observation substantiates the interpretation that c-di-GMP enhances the SagS-HapZ interaction. The c-di-GMP mediated effect is specific because an analogous pattern was not observed in case of GMP and other nucleotides (Fig. 11D). On the contrary, c-di-GMP did not seem to have a significant effect on HptB phosphorylation when HapZ was replaced by the HapZR12A/R16A double mutant that lacks the ability to bind c-di-GMP, (Fig. 11E), strongly suggesting that c-di-GMP exerts its effect through HapZ. Note that although a small sub-family of histidine kinases has been reported to be able to bind c-di-GMP directly (280), SagS is unlikely to bind c-di-GMP due to the absence of c-di-GMP binding residues (Fig.

Chapter 2

S2). Overall, the phosphorylation assays reinforce the fact that HapZ and c-di-GMP can hinder SagS-initiated phosphorelay cascade and that the inhibition is strengthened at high intracellular levels of c-di-GMP.

Figure 11 | The phosphotransfer from SagS to HptB is inhibited by HapZ and this effect is enhanced by an increasing gradient of c-di-GMP. (A) Schematic representation of the SagS phosphorelay pathway where phosphotransfer occurs from SagS to the downstream HptB and PA3346 (HsbR) proteins [Adapted from (227, 229)]. (B) 15% SDS-polyacrylamide (SDS-PAGE) gel showing the three proteins, HapZ, HptB and SagS₂₄₆₋₇₈₆ expressed and purified for *in vitro* phosphorylation assay and also the phosphoreceiver domain of SagS (REC_{SagS} was not used in the *in vitro* phosphorelay experiment). All the proteins expressed contained an N-terminal 6xHis tag. (C) Inhibitory effect of HapZ and c-di-GMP on HptB phosphorylation. (D) Effect of GMP on HptB phosphorylation (E) Effect of the HapZ-R12A/R16A double mutant and c-di-GMP on HptB phosphorylation.

*HapZ mediates surface attachment and biofilm formation.*¹

P. aeruginosa is notorious for its ability to form antibiotic-resistant biofilm on abiotic and biotic surfaces. SagS (surface attachment and growth sensor hybrid) has been known to control surface attachment and development of *P. aeruginosa* biofilm (228). If HapZ interacts with SagS, HapZ should also affect surface attachment and biofilm development. By using

¹ The flow-cell biofilm assay was done in collaboration with Prof. Yang Liang's lab, SCELSCE, NTU.

Figure 12 | HapZ controls surface attachment and development of biofilm. (A) Biofilm assay using flow-cells for PAO1, $\Delta hapZ$ and $hapZ^+$ strains. Flow cells were used to grow biofilms of PAO1, deletion ($\Delta hapZ$) and overexpression ($hapZ^+$) strains of HapZ. They were monitored regularly without any flow disruption using a confocal microscope. Surface attachment of bacterial cells and formation of biofilm were seen using confocal scanning laser microscopy micrographs at the end of day 3 and day 5. Scale bars (50 μm). (B). Biofilms formed by co-culturing PAO1 & $\Delta hapZ$ as well as PAO1 & $hapZ^+$ strains. Expression of different protein dyes [cyan fluorescent protein (CFP) and yellow fluorescent protein (YFP)] by the two strains being co-cultured allowed the monitoring of biofilm dominance using a confocal laser-scanning microscope.

flow cells and a confocal microscope, we studied how the deletion and overexpression of *hapZ* affects the formation of *P. aeruginosa* biofilm. As shown in Fig. 12A, the PAO1 strain formed a characteristic mushroom-shaped biofilm after five days of growth; whereas the $\Delta hapZ$ mutant strain exhibited a significant delay in early surface attachment resulting in a thin biofilm at the end. In contrast, the HapZ overexpression strain ($hapZ^+$) harbouring the pUCP18-HapZ overexpression plasmid exhibited efficient surface attachment at the early

Chapter 2

stage. However, despite the facile surface attachment, the *hapZ*⁺ strain formed a flat and fragile biofilm that completely lacked the mushroom structure at the end of five-day growth.

To reaffirm the role of HapZ in initial surface attachment and early biofilm formation, competitive biofilm assays were carried out with mixed *P. aeruginosa* populations. Co-culturing of the cyan fluorescence protein (CFP)-tagged PAO1 strain with YFP-tagged $\Delta hapZ$ strain in flow cells resulted in the final biofilm being dominated by the PAO1 strain (Fig. 12B), ascertaining that the $\Delta hapZ$ strain was defective in surface attachment. Although the *hapZ*⁺ strain was unable to form a structured biofilm, YFP-tagged *hapZ*⁺ emerged as a more dominant strain when co-cultured with PAO1, implying that the *hapZ*⁺ strain was more proficient than PAO1 in surface attachment. Overall, these findings indicate that HapZ plays a regulatory role in surface attachment and biofilm development, validating the biological role of HapZ in the SagS two-component signalling pathway.

HapZ transcription levels vary during different stages of biofilm formation

Our findings demonstrate that HapZ affects biofilm formation. To investigate whether the cellular levels of HapZ fluctuate during biofilm formation, *P. aeruginosa* PAO1 cells were grown in flow-cells and harvested at different time points (day 1, day 3 and day 5). The mRNA was isolated from these cells for real-time PCR (RT-PCR) quantification. RT-PCR revealed that HapZ transcription levels vary considerably during the five-day period of biofilm formation (Fig. 13). HapZ has high transcriptional levels before surface attachment when the cells are in the planktonic stage (0 hrs). There is a sharp drop in HapZ transcriptional levels following surface attachment (day 1). However, HapZ levels increase again along with the biofilm development and peak at day 5 during biofilm maturation. This indicates that HapZ plays an important role both during initial surface attachment and biofilm maturation.

HapZ regulates swarming and twitching motility.

Motility is essential for the survival, pathogenesis and biofilm formation in *P. aeruginosa*. *P. aeruginosa* cells explore their surrounding environments by using either swimming, swarming or twitching motility. Flagella-mediated swarming motility occurs in a coordinated manner on semi-solid surfaces and is accompanied by production of biosurfactants. *P.*

Chapter 2

Figure 13 | HapZ transcription levels vary during different stages of biofilm formation. Real-time PCR quantification of the fold change in the transcriptional levels of HapZ at various stages during biofilm formation in *P. aeruginosa* PAO1. The cells are in the planktonic stage at 0 hr and slowly progress to forming biofilms from day 1 to day 5.

aeruginosa displays a typical dendrite-like swarming pattern with radiating tendrils (281, 282). In agreement with the role of HapZ in regulating biofilm formation, we found that the deletion and overexpression of HapZ had a great impact on swarming motility. The hyper-swarming phenotype of the $\Delta hapZ$ strain could be reversed by complementation with HapZ (Fig. 14A). On the other hand, the swarming activity of the $hapZ^+$ strain was suppressed further compared to the wild type PAO1 strain, suggesting that lower levels of *hapZ* account for the swarming phenotype of $\Delta hapZ$. Further, the HapZ- $\Delta 21$ construct lacking the N-terminal loop was unable to complement the hyper-swarming phenotype of $\Delta hapZ$, confirming the significance of c-di-GMP binding for HapZ function. In addition, large fluid zones surrounding the tendrils of $\Delta hapZ$ cells were seen when the swarming plates were closely examined. It has been established that the zones of fluid are due to rhamnolipid production by *P. aeruginosa* cells (282). Hence, we speculated that the increased production of rhamnolipids may, at least partly, responsible for causing the increased swarming phenotype in the $\Delta hapZ$ strain. To investigate this, the amount of rhamnolipids produced by PAO1, $\Delta hapZ$ and $hapZ\text{-}\Delta 21$ strains were quantified using an emulsification assay. We noticed an increase in rhamnolipid production for the $\Delta hapZ$ strain and a decrease in rhamnolipid production for the $hapZ^+$ strain that supported our findings (Fig. 14B).

Chapter 2

Figure 14 | HapZ impacts swarming and twitching but not swimming motility. (A) Swarming motility patterns of PAO1, $\Delta hapZ$, $\Delta hapZ + pHapZ$, $\Delta hapZ + pHapZ\Delta 21$ and $hapZ^+$ strains. Analysis of swarming motility was performed on semi-hydrated (containing 0.5% agar) M8 medium plates. (B) Rhamnolipid emulsification assay showing differences in the quantity of rhamnolipids produced by PAO1, $\Delta hapZ$ and $hapZ^+$ strains. (C) A plot comparing the differences in radii of twitching zones of PAO1, $\Delta hapZ$, $\Delta hapZ + pHapZ$ and $\Delta hapZ + pHapZ\Delta 21$ strains (D) PAO1 spends equal time in both CW and CCW directions during swimming. A plot showing the number of PAO1 cells versus the time spent by the cells in the CW phase during the whole flagellar rotation time using bacterial tethering analysis. (E) Most of the $\Delta hapZ$ cells spend equal time in both CW and CCW directions during swimming. A plot showing the number of $\Delta hapZ$ cells versus the time spent by cells in the CW phase during the whole flagellar rotation time using bacterial tethering analysis. $\Delta hapZ$ displays a comparable swimming pattern to PAO1.

Unlike swarming motility, the flagella-independent twitching motility is mediated by the extension and retraction of surface appendages called type-IV pili (T4P) over wet surfaces which is also critical for biofilm formation in *P. aeruginosa* (43). It was revealed that HapZ controls twitching motility since a greater twitching zone was noticed in the $\Delta hapZ$ strain, which could be suppressed by HapZ, but not by HapZ- $\Delta 21$ overexpression (Fig. 14C). While it is already known that the c-di-GMP effector protein FimX regulates twitching motility in *P. aeruginosa* (154), our findings indicate that twitching motility in this bacterium can also be controlled by another effector protein HapZ in a c-di-GMP dependent mode. $\Delta hapZ$ strain in both the CW and CCW phases is comparable to the wild type PAO1 strain (Fig. 14D and 14E). This indicates that although HapZ controls swarming and twitching motility, it does not affect swimming motility.

Chapter 2

Colony morphology is influenced by HapZ

Colony morphogenesis assay is a measure of the extent of biofilm formation and is dependent on the intracellular redox state of the cells. Deficiency in the availability of electron acceptors such as oxygen, nitrate and small molecules called phenazines for *P. aeruginosa* results in the induction of colony wrinkling to compensate for redox imbalance. This maximizes the availability of oxygen to the cells, thus supporting metabolic homeostasis. Since HapZ is implicated to affect biofilm formation, we investigated whether it affects colony morphogenesis. We observed that the $\Delta hapZ$ strain formed relatively smooth colonies on Congo Red-impregnated agar plates in comparison to the PAO1 strain, and that the $hapZ^+$ mutant formed more rugose and wrinkled colonies (Fig. 15). The smooth phenotype of the $\Delta hapZ$ mutant could be reversed by complementation with the overexpressing pHapZ plasmid, but not with the pHapZ- $\Delta 21$ plasmid. The morphology of *P. aeruginosa* colony formed on solid agar surface is influenced by many factors (278, 283, 284). The differences in colony morphology indicate that HapZ may influence the biofilm formation by production of EPS, redox metabolites such as phenazine and quorum sensing (QS) signals.

Figure 15 | HapZ influences colony morphogenesis. Colony morphology of PAO1, $\Delta hapZ$, $\Delta hapZ$ + HapZ and PAO1 + HapZ strains on Congo-red agar plates. The $\Delta hapZ$ strain shows a smooth phenotype and complementation by HapZ overexpression reverts the $\Delta hapZ$ phenotype to the wild type PAO1. Hyperwrinkled colonies are formed upon HapZ overexpression in the PAO1 wild type strain.

A global c-di-GMP pool may be involved in HapZ-mediated biofilm regulation

A complex c-di-GMP signalling network regulates a plethora of cellular functions in *P. aeruginosa*. After being synthesized by diguanylate cyclases, c-di-GMP binds to effector proteins to mediate its cellular effects. A total of 40 proteins in *P. aeruginosa* PAO1 are involved in c-di-GMP metabolism. Of these, 17 are GGDEF domain-containing proteins that possess DGC activity and 16 are GGDEF/EAL domain-containing composite proteins in *P. aeruginosa* (75, 118) (Fig 16). A sub-set of these GGDEF or GGDEF/EAL proteins can be

Chapter 2

associated with an effector regulatory pathway for contributing c-di-GMP. In order to determine the GGDEF or GGDEF/EAL domain-containing protein/proteins linked to the HapZ regulatory pathway, we performed colony morphology assay by overexpressing HapZ in the deletion strains of 17 GGDEF domain-containing proteins and 8 GGDEF/EAL domain-containing composite proteins in the *P. aeruginosa* PAO1 strain.

Figure 16 | Complexity of c-di-GMP signalling network in *P. aeruginosa*. *P. aeruginosa* has a complex c-di-GMP regulatory network comprising of c-di-GMP metabolizing DGCs and PDEs and c-di-GMP effectors. The 25 GGDEF and GGDEF/EAL domain-containing proteins in PAO1 tested for affecting the c-di-GMP mediated HapZ regulation are listed in the box on the left. The 12 effector proteins in *P. aeruginosa* are listed in the box on the right where the PilZ domain-containing proteins are shown in blue except HapZ which is shown in red.

Hyperwrinkling of colonies was observed upon HapZ overexpression in the $\Delta PA3702$ ($\Delta wspR$) and $\Delta PA0861$ ($\Delta rbdA$) strains. The DGC WspR forms part of a chemosensory pathway and has an active GGDEF domain. It is known to promote biofilm formation and suppress motility in *P. aeruginosa* (83, 86, 88). Hyperwrinkling in the $\Delta wspR$ colonies overexpressing HapZ indicates that c-di-GMP synthesized by WspR GGDEF domain does not affect the HapZ-controlled pathway. Hence, HapZ continues to synthesize colony biofilms even in the absence of WspR causing hyperwrinkling. Thus, we can propose that the WspR and HapZ biofilm regulatory pathways are likely to function in an independent manner. On the other hand, RbdA is a GGDEF/EAL composite protein that has been demonstrated to function in biofilm dispersal in *P. aeruginosa*. Moreover, it has also been shown to display only phosphodiesterase (PDE) activity by *in vivo* and *in vitro* studies. The

Chapter 2

GGDEF domain of RbdA activates its PDE activity in the presence of GTP (96) due to which the $\Delta rbdA$ strain will possess higher levels of c-di-GMP compared to the wild type PAO1.

A. GGDEF only

B. GGDEF/EAL

Figure 17 | A global c-di-GMP pool may be involved in the HapZ-mediated biofilm regulation. (A) Colony morphology of deletion strains of GGDEF domain-containing proteins of *P. aeruginosa* and the effect of HapZ overexpression on the phenotype of these strains. (B) Colony morphology of deletion strains of GGDEF/EAL domain-containing proteins of *P. aeruginosa* and the effect of HapZ overexpression on the phenotype of these strains.

Chapter 2

Thus, overexpression of HapZ in the *ArbdA* strain would result in an increase in biofilm formation causing hyperwrinkling of colonies. We observed that with the exception of PA3702 (WspR) and PA0861 (RbdA), none of the 23 deletion strains displayed any significant phenotypic change after HapZ overexpression (Fig. 17A and B). This suggests that all the 23 GGDEF and GGDEF/EAL mutant strains may contribute c-di-GMP to the HapZ controlled pathway affecting biofilm formation. Thus, a global c-di-GMP pool involving 23 GGDEF and GGDEF/EAL proteins may be associated with HapZ.

***HapZ regulates the expression of a large number of genes related to QS, motility and other functions.*²**

SagS and HptB-associated signalling pathways control swarming and biofilm formation likely through regulating gene expression (229). We carried out transcriptomic analysis using RNA-Seq for PAO1, *ΔhapZ*, *ΔsagS* and the complementation strain (*ΔhapZ* + HapZ) to elucidate the functional association between SagS and HapZ and to determine the genes under HapZ regulation (229, 285). The gene expression profile of the *ΔhapZ* strain differs noticeably from PAO1 strain in the cells obtained from the stationary growth phase (Fig. 18A) The complementation (*ΔhapZ* + HapZ) strain could restore the gene expression pattern indicating that differential expression is caused by variation in HapZ levels. In contrast to PAO1, 94 genes of the *ΔhapZ* strain exhibited substantial alterations in mRNA levels (Fig. 18 B) (Table S4 and S5). Of the 61 upregulated genes (> 2.5 fold-change), (Fig 18 D and E) (Table S6), some have been shown to be critical for motility, virulence and biofilm formation. The largest change was observed for the genes *lecA* (46-fold) and *lecB* (17-fold), which encodes two lectins that are associated with cell surface adhesion. The latter (LecB) is also important for biofilm formation as the strain deficient in LecB is hampered in biofilm development (286). Under unfavourable environmental conditions, biofilm formation, motility, survival and population structure in *P. aeruginosa* are regulated by the Pseudomonas quinolone signal (PQS) (287, 288). The genes responsible for the biosynthesis of the PQS (*pqsA*, B, C, D and E) are upregulated in the *ΔhapZ* strain by 4-11 fold. The rhamnolipid biosynthetic genes *rhlA* (9.4-fold), *rhlB* (4.4-fold) and *rhlI* (2.3-fold) are also upregulated, which is in agreement with the hyper-swarming phenotype and enhanced rhamnolipid production in the *ΔhapZ* strain. Elevated expression levels were also observed for the genes that encode the MexG/H/I-OpmD efflux pump (3-8 fold) elastase (*lasB*, 6.6

² The RNA-Seq analysis was done in collaboration with Prof. Yang Liang's lab, SCELSC, NTU.

Chapter 2

fold) and three genes (*glcD/E/F*) from the central metabolic glycolate pathway.

Figure 18 | HapZ and SagS co-regulate the expression of several genes. (A) RNA-seq transcriptome analysis of $\Delta hapZ$ + pHapZ, PAO1, $\Delta hapZ$ and $\Delta sagS$ strains are depicted as heatmaps. (B) Venn diagrams depicting the number of genes regulated by HapZ and SagS both individually and together (overlapping regions). (C) Plots showing gfp-reporter assay based measurements of three up-regulated QS-controlled genes in both $\Delta hapZ$ and $\Delta sagS$ strains. (D) and (F) depict the genes whose expression is up-regulated and down-regulated in both $\Delta hapZ$ and $\Delta sagS$ strains respectively, (E) depicts the genes whose expression is up-regulated in only $\Delta hapZ$ but not in $\Delta sagS$ strain and (G) depicts the genes whose expression is down-regulated in only $\Delta hapZ$ but not in $\Delta sagS$ strain.

We validated the alteration of gene expression for the $\Delta hapZ$ mutant by using other experimental approaches. The expression levels of the QS-controlled *lasB*, *rhIA* and *pqsA* were measured for PAO1 and the $\Delta hapZ$ strains by using promoter-gfp-fusion reporters. Quantification of GFP fluorescence revealed increased transcription levels for the three key QS genes in the $\Delta hapZ$ strain (Fig. 18C). In addition, 33 genes displayed lowered expression

Chapter 2

levels in $\Delta hapZ$ strain (Fig 18G and Table S7).

We performed RNA-Seq analysis with the $\Delta sagS$ mutant strain to compare the gene regulons of HapZ and SagS. In comparison to the 94 genes of the $\Delta hapZ$ strain that exhibit increased expression levels, a total of 105 genes displayed enhanced expression levels in the $\Delta sagS$ strain (Table S8). A total of 64 co-regulated genes were identified in the regulons of SagS and HapZ (Fig. 18B). Among the 64 genes, 44 are upregulated while 20 are downregulated and none of the genes showed opposite changes in both the strains. The 44 co-upregulated genes include *lasB*, *lecA/B*, *rhlA/B*, *mexG/H/I-opmD*, *pqsA-E*, *glcD/E/F* and other genes with unknown functions (Fig. 18D). On the other hand, the main co-downregulated genes are the efflux pump genes *mexE/F*, *oprN* and several un-annotated genes (Fig. 18F). *mexE/F* and *oprN* encode the MexE/F-oprN efflux pump and *mexS* codes for an associated regulator that is known to be involved in biofilm formation, antibiotic resistance and virulence (289). Sauer and co-workers recently reported that inactivation of *sagS* affected the expression of *mexA/B*, *oprM*, *mexE/F* and *oprN* genes (290). The observed changes in the expression level for these genes in the $\Delta sagS$ strain are in agreement with their observations. Altogether, the RNA-seq revealed that the SagS and HapZ regulons share a large number of genes and provides further support for the role of HapZ in controlling gene expression within the SagS signalling pathway.

2.4. Discussion

C-di-GMP exerts some of its effects in *P. aeruginosa* by binding to several PilZ domain proteins to control unknown pathways and phenotypes. Determining the function of the PilZ effectors will considerably advance our knowledge of the role of c-di-GMP in *P. aeruginosa*. Towards this aim, our studies provided an insight into the role of one PilZ effector (HapZ) by the identification of the hybrid histidine kinase SagS as a binding partner. We demonstrated that the protein-protein interaction between SagS and HapZ is highly specific and is c-di-GMP-dependent. We showed that deletion of the *hapZ* gene had a pleiotropic effect on gene expression in *P. aeruginosa*. Our data indicates that HapZ functions as an adaptor to allow c-di-GMP to control gene expression through a two-component signalling network. The study unveils a novel mechanism whereby c-di-GMP controls two-component signalling through a discrete PilZ adaptor and establishes HapZ as a new regulatory factor in biofilm formation.

Chapter 2

How does c-di-GMP regulate gene expression through HapZ? The SagS pathway is part of a highly complex two-component signalling network consisting of multiple sensor kinases, response regulators, connectors and transcriptional regulators (Fig. 19). SagS has been reported to initiate a phosphorelay cascade through the HptB-dependent pathways and also mediate the HptB-independent BfiS/BfiR pathway (285, 291). HptB acts as a linker that accepts phosphoryl group from SagS, PA1976, PA1611 and phosphorylates RetS, a histidine kinase that interacts directly with GacS to mediate the GacS/GacA pathway (227-229). Our present understanding of this complex two-component signalling network is far from complete, with input signals for the histidine kinases and potential additional components to be identified. The way in which the histidine kinases collectively control the phosphorylation of HptB to modulate gene expression also remains to be fully understood. What seems to be clear now is that the two-component signalling network regulates the expression of a large number of genes and plays a crucial role of coordinating the timing of surface attachment and biofilm development, with the pathways activated at different time points during biofilm formation (211, 285, 292-294). Within this intricate and dynamic signalling network, we propose that HapZ enables *P. aeruginosa* to respond to additional environmental or cellular signals sensed and transmitted by c-di-GMP signalling pathways. In response to rising or falling cellular c-di-GMP concentration, HapZ interacts with SagS in a reversible manner to modulate the phosphorylation level of HptB and such response regulators as BfiR and GacA. At high c-di-GMP concentrations, formation of the SagS/HapZ/c-di-GMP ternary complex inhibits the kinase activity of SagS and phosphorylation of HptB by SagS. At low c-di-GMP concentrations, dissociation of the ternary complex frees SagS to phosphorylate its protein substrate(s).

A complete understanding of the role of c-di-GMP and HptB in this complex two-component signalling system is not feasible currently due to several reasons. Although studies have suggested that the histidine kinases are likely to be under tight control during biofilm formation (291), the input signals and the timing of activation/de-activation of the histidine kinases remain unclear. Considering that histidine kinase can exhibit either kinase or phosphatase activity depending on the conformation of its catalytic domains (295), we speculate that the phosphatase activity may be suppressed under the *in vitro* assay conditions. Hence, it remains to be determined in the future whether HapZ and c-di-GMP affect the phosphorylation of BfiS and the BfiS/BfiR pathway. SagS was also reported to control the

Chapter 2

expression of a downstream c-di-GMP-binding transcriptional regulator (BrlR) and affect cellular c-di-GMP level through unidentified DGC or PDE proteins (296). Having established the function of HapZ, the SagS signalling pathway is potentially regulated by a c-di-GMP-dependent feedback mechanism that remains to be investigated further.

Figure 19 | A model depicting the role of HapZ in an intricate two-component signalling network. When the c-di-GMP concentration rises, binding of c-di-GMP to HapZ enhances the interaction between SagS and HapZ and results in the inhibition of SagS by HapZ. Inhibition of SagS leads to extensive changes in gene expression, likely by regulating the phosphorylated state of HptB and the HKs through the PA3346/PA3347, the BfiS/R and the RetS/GacS/GacA pathways.

An important finding of our study is that the small PilZ protein HapZ is critical for the formation of mature *P. aeruginosa* biofilm. Considering the importance of biofilm in chronic infection caused by *P. aeruginosa* (297), HapZ may play an active role in *in vivo* survival and pathogenesis. Formation of biofilm involves several stages including a reversible attachment stage, an irreversible attachment stage and two maturation/development stages (298, 299). As a central regulator of biofilm formation in *P. aeruginosa*, c-di-GMP has been suggested to control different stages of biofilm formation by regulating the motile-to-sessile switch, EPS production and other functions (300). It has been suggested that SagS acts as a switch during biofilm formation by regulating the motile-to-sessile transition (228, 285). In agreement with the role of SagS as a switch, our study showed that HapZ is critical for efficient surface attachment at the early stage of biofilm formation. This was further confirmed by an increase in HapZ transcriptional levels during surface attachment and biofilm maturation.

Chapter 2

Our data further indicates that HapZ promotes surface attachment by regulating swarming motility and other functions. Swarming motility is one of the major factors that affect surface attachment; and suppressed swarming motility in a non-motile subpopulation is crucial for initial surface attachment (89, 173). The ability of *P. aeruginosa* cells to swarm relies on the production of QS signals and metabolites such as rhamnolipids (282, 301). To promote surface attachment at the early stage of biofilm formation, HapZ is likely to suppress swarming motility by controlling the production of QS signals, rhamnolipid and other protein and small-molecule factors. Maturation of *P. aeruginosa* biofilm relies on multiple regulatory pathways that control bacterial motility, EPS production, QS and rhamnolipid production. The Las, Rhl and PQS quorum-sensing systems of *P. aeruginosa* have all been implicated in biofilm development. Our transcriptomic analysis showed that HapZ controls the expression of the *pqs* genes and many other genes under the control of the Las and Rhl QS systems. The transition from the initial attachment stage to the final maturation stage during biofilm formation is characterized by extensive shift in gene expression (46) and many genetic regulators are required to synchronize the shift in gene expression. Such shift is critical for biofilm maturation, as the surface-attached *P. aeruginosa* cells need to construct the extracellular matrix and form microcolonies. We propose that HapZ is a key factor that not only promotes surface attachment but also orchestrates the transition from a “surface attachment phenotype” to a “maturation phenotype” during biofilm formation. HapZ can be considered as a pleiotropic regulator that contributes to biofilm maturation by regulating the expression of genes associated with swarming/twitching motility, QS signalling, metabolism, efflux pumps and other functions. Together with SagS and other signalling proteins, HapZ plays an important role in biofilm formation and potentially chronic persistence during *P. aeruginosa* infection by orchestrating the temporal expression of genes. It also should be noted that it was assumed previously that the trigger for the SagS-controlled motile-to-sessile switch is the input signal for SagS or other histidine kinases. Our study raises the possibility that the motile-to-sessile switching is triggered by the rise or fall of cellular c-di-GMP levels, not by the input signal of the histidine kinases.

The involvement of c-di-GMP in two-component signalling is a common phenomenon, as already evidenced by the large number of GGDEF, EAL or HD-GYP domain-containing response regulators encoded by bacterial genomes (111, 115, 157, 302, 303). By using the c-

Chapter 2

di-GMP-synthesizing or degrading response regulators, the environmental inputs sensed by histidine kinases can be translated into oscillation of cellular c-di-GMP concentration. Colony morphology studies using various DGC mutant strains in *P. aeruginosa* revealed that a global c-di-GMP pool is likely to control the HapZ regulatory pathway. One of the surprising findings from recent studies is that c-di-GMP can also control the activity and function of histidine kinases to regulate two-component signalling. For example, the bifunctional histidine kinase CckA involved in cell cycle control in *C. crescentus* binds c-di-GMP directly in the catalytic & ATP-binding domain for regulating the kinase and phosphatase activities (280). Another example is the orphan hybrid histidine kinase SgmT from *Myxococcus xanthus*. SgmT has a C-terminal non-enzymatic GGDEF domain that functions as a c-di-GMP-binding effector. The binding of c-di-GMP to the GGDEF domain is crucial for the spatial sequestration, and likely cellular function of SgmT (304). Our current study suggests that an alternative way for c-di-GMP to control histidine kinases is by using a free-standing PilZ adaptor. In principle, the use of PilZ adaptors has the advantage of allowing c-di-GMP to control multiple pathways through highly specific kinase-PilZ interaction. Because PilZ proteins exhibit a wide range of binding affinity for c-di-GMP (23), the effect of c-di-GMP on a specific pathway can be “sensed” only when the c-di-GMP concentrations reach a threshold. Moreover, creating a discrete PilZ adaptor to regulate histidine kinase is evolutionarily simpler than introducing a c-di-GMP-binding site in the kinase domain due to fewer evolutionary constricts. The ability to rapidly evolve new c-di-GMP-binding PilZ adaptors for two-component signalling pathways may be crucial for the survival of such highly adaptable bacteria as *P. aeruginosa*.

CHAPTER 3: Protein-protein interaction between the SagS and its protein partners - HptB and HapZ

3.1. Introduction

Bacteria detect and respond to various environmental stimuli and mediate precise changes in cellular processes (189, 193) by a synchronized regulation of a multitude of complex signalling pathways. One of the several regulatory mechanisms is a basic phosphorylation-mediated signal transduction known as two-component signalling (TCS) systems. These multi-step pathways control a diversity of behaviours from basic processes like metabolism, cellular motility to complex ones such as virulence, sporulation, cell-cycle regulation, development and resistance to antibiotics. Numerous TCSs have been documented in prokaryotes and eukaryotes including *Escherichia coli* (305), *Bacillus subtilis* (306-308), *Pseudomonas aeruginosa* (195), *Caulobacter crescentus* (309), *Sinorhizobium meliloti* (310), *Myxococcus xanthus* (311) and *Saccharomyces cerevisiae* (312, 313). The fundamental components of a two-component signalling (TCS) system are histidine kinases (HKs) and response regulators (RRs). When the sensor domain of a histidine kinase senses an external signal, it undergoes autophosphorylation at a conserved histidine residue and effects the appropriate cellular response by interacting and transferring the phosphate group to a conserved aspartate residue in its cognate response regulator. This induces a conformational change in the response regulator leading to an output response and subsequent alterations in gene expression (189, 194).

Some TCS pathways involve hybrid sensor kinases where the receiver (REC) domains of histidine kinases mediate phosphotransfer to the cognate response regulator via an intermediate histidine phosphotransfer protein (Hpt). The interaction between REC domain and Hpt proteins in the multistep two-component pathway is transient in nature. Furthermore, the phospho-histidiny and phospho-aspartyl bonds formed during phosphotransfer are chemically labile. This makes it difficult to study the structure and nature of interaction of REC-Hpt complexes. A few complex structures of REC and Hpt domains have been determined. These include the YPD1/SLN1 complex from *Saccharomyces cerevisiae*, Spo0F/Spo0B complex from *Bacillus subtilis* and CheY₆/CheA₃ complex from *Rhodobacter*

Chapter 3

sphaeroides. These studies demonstrate that the REC-Hpt binding interface is largely a hydrophobic patch surrounded by polar residues. This was proposed as a general feature for the interaction between REC and Hpt domain-containing proteins.

Although few REC-Hpt complexes have been studied, the mechanism of interaction between these phosphorelay partners remains unknown in the opportunistic pathogen *Pseudomonas aeruginosa*. It has been found to possess the highest number of two-component signalling systems than all other sequenced bacterial genomes. It has a total of 127 TCSs consisting of 63 histidine kinases and 64 response regulators (195, 211). Sixteen atypical kinases were identified in *P. aeruginosa* of which the receiver domains of 11 of them were not associated with an Hpt domain. Interestingly, only three such Hpt modules were recognized in *P. aeruginosa*, namely HptA, HptB and HptC. The ability of 11 atypical kinases to mediate phosphotransfer in a targeted manner using three Hpt modules requires a high degree of specific binding between the receiver domains and Hpt proteins (195). This suggests that complex regulatory strategies are employed by *P. aeruginosa* to modulate various processes and generate a specific adaptive response to environmental signals.

Numerous two-component pathways have been studied in great detail in *P. aeruginosa* of which many are involved in the regulation of biofilm formation (111, 213, 218, 224). One such TCS involves a multi-step pathway where the hybrid sensor kinase SagS (PA2824) relays signals by binding and phosphorylating the intermediate histidine phosphotransfer protein B (HptB) and regulates biofilm formation (227, 229). Hsu and co-workers demonstrated that three other hybrid sensor kinases PA1611, PA1976 and PA4856 (RetS) are also involved in phosphotransfer to HptB through their receiver (REC) domains (228). SagS undergoes autophosphorylation upon sensing an environmental signal and mediates phosphotransfer through its receiver domain to an intermediate downstream protein HptB (PA3345) which in turn relays the signal to its cognate response regulator PA3346 (HsbR), thus regulating biofilm formation (Fig. 11A) (227, 229, 285). The ability of REC domains of several histidine kinases to bind to HptB indicates that they possess certain molecular features imparting specificity for binding to HptB. However, the molecular mechanism by which HptB interacts with three histidine kinases in a specific manner remains unknown.

Chapter 3

In addition to the specific interaction of SagS with HptB, our recent study in Chapter 2 unraveled that the stand-alone PilZ protein PA2799 from *P. aeruginosa* also specifically binds to the C-terminal receiver domain of SagS (REC_{SagS}) and participates in the two-component signalling pathway by blocking the phosphorelay from SagS to HptB (314). We renamed PA2799 as Histidine kinase associated PilZ (HapZ). We showed that HapZ functions as a PilZ adaptor protein to regulate two-component signalling by binding to c-di-GMP (314). RNA sequencing studies also revealed that HapZ mediates these effects by controlling the expression of a large number of genes.

Eight PilZ domain-containing proteins have been identified in *P. aeruginosa* of which seven PilZs bind c-di-GMP. These include both di-domain [Alg44, FlgZ (PA3353), PA2989] and free-standing PilZ proteins [PA0012, HapZ (PA2799), PA4324, PA4608]. Considerable progress has been made in solving the structures of PilZ domain-containing proteins in *P. aeruginosa* alone or in complex with c-di-GMP. Although studies have unravelled the structure and function of few PilZs and investigated the mode of PilZ and c-di-GMP binding, the mechanism by which a PilZ protein specifically interacts with its protein target to regulate cellular functions remains unknown (71, 129-132, 261, 262, 315, 316). Thus, understanding the molecular basis of the SagS-HapZ interaction will uncover the potential of this large family of proteins to function as adaptors and effect changes in cellular processes by binding to c-di-GMP.

In the present chapter, we probed the nature of interaction between the receiver domain of the hybrid sensor kinase SagS (REC_{SagS}) and three other hybrid kinases with the monomeric histidine phosphotransfer protein HptB in *P. aeruginosa*. Based on NMR titration data, we provide an explanation for how REC_{SagS} binds specifically to Hpt, but not HptA and HptC. Since numerous two-component systems involving Hpt domains are present in *P. aeruginosa*, this mechanism of interaction can serve as a good model for studying other REC-Hpt complexes in *P. aeruginosa* and other organisms. In addition to the REC_{SagS}-HptB interaction, we also explored the molecular basis of HapZ- REC_{SagS} interaction and the mechanism by which c-di-GMP modulates the protein-protein interaction in *P. aeruginosa*. This was done by screening a comprehensive mutant library of HapZ and REC_{SagS} for identification of residues involved in these interactions using a bacterial two-hybrid

Chapter 3

screening. The role of the residues identified from two-hybrid screening was validated using functional assays including swarming motility and colony morphology.

3.2. Materials and Methods

Bacterial strains and plasmids

The SagS (PA2824) receiver domain (residues 663-786) and HptB (PA3345) genes were amplified from the *P. aeruginosa* PAO1 genome and cloned into the pET28b (+) expression vector (Table S2 and S3). The recombinant proteins had His₆ tag on their N-terminus. The expression plasmids were transformed into *E. coli* Rosetta cells and stored at -80 °C as glycerol stocks.

Protein expression and purification

Expression and purification of ¹⁵N labelled and ¹⁵N, ¹³C labelled SagS_{663-786aa} receiver domain (REC_{SagS})

The overexpression of ¹⁵N labelled SagS_{663-786aa} receiver domain (REC) protein (REC_{SagS}) for HSQC NMR titration was carried out by growing the *E. coli* Rosetta cells carrying the pET28b-REC_{SagS} expression plasmid in 4 litres of M9 minimal medium supplemented with ¹⁵N NH₄Cl, 0.2% glucose and 50 µg/ml Kanamycin at 37 °C. Protein expression was induced overnight at 16 °C using 0.5 mM isopropyl β-D-thiogalactopyranoside (IPTG) at an absorbance (600 nm) of 0.6. Further, REC_{SagS} was purified using nickel-nitrilotriacetic acid affinity chromatography column equilibrated with 50 mM sodium phosphate buffer pH 8.0. The His₆ tagged ¹⁵N labelled protein was eluted using an imidazole gradient (20 mM to 300 mM) and analysed using SDS-PAGE. Fractions containing the overexpressed proteins were pooled, desalted using PD-10 desalting columns (GE Healthcare), concentrated using Amicon concentrator (5 kDa) and aliquot and stored in phosphate buffer pH 8.0 at -80 °C. For the purification of ¹⁵N, ¹³C labelled SagS_{663-786aa} for 3D solution structure determination, the same protocol used for overexpression of ¹⁵N labelled SagS_{663-786aa} was followed. The only difference was the addition of 0.2% ¹³C glucose to the M9 medium containing ¹⁵N NH₄Cl and 50 µg/ml kanamycin. The His-tags for ¹⁵N and ¹⁵N, ¹³C labelled SagS_{663-786aa} were retained for both HSQC and NMR solution structure determination experiments.

Chapter 3

Expression and purification of HptB

In order to perform 2D NMR titration studies with ^{15}N labelled REC_{SagS} protein, recombinant HptB (PA3345) with N-terminal His₆-tag was overexpressed using *E. coli* Rosetta cells carrying the pET28b-HptB plasmid constructs. The cells were grown in Luria Bertani medium containing 50 µg/ml Kanamycin at 37 °C to an absorbance (600 nm) of 0.6 before the induction of protein expression using 0.5 mM IPTG at 16 °C for 16 hours. Cells were harvested by centrifugation and resuspended in 50 mM phosphate buffer pH 8.0 and lysed on ice by ultrasonication for 30 mins. The cells were then centrifuged at 25,000 rpm for 30 minutes and supernatants were loaded onto a nickel-charged column equilibrated with phosphate buffer pH 8.0. The column was washed sequentially using wash buffers W1 and W2 containing 50 mM Phosphate buffer pH 8.0 and 20 mM and 50 mM Imidazole respectively. The His₆ tagged protein was eluted from the column using elution buffers (50 mM phosphate buffer pH 8.0) containing increasing concentrations of imidazole (100 mM, 200 mM and 300 mM) and were analysed using SDS-PAGE. The fractions containing the desired protein were pooled to a purity > 95%, desalted using PD-10 column (GE Healthcare) and concentrated using Amicon concentrator (5 kDa) at 4000 rpm at 4 °C. The concentration of the protein were determined by Bradford method and were aliquot and stored in 50 mM phosphate buffer pH 8.0 at -80 °C.

NMR solution structure determination

Samples were prepared for the NMR titration studies using the following combinations: (1) ^{15}N , ^{13}C REC_{SagS} (2) ^{15}N REC_{SagS} and (3) ^{15}N REC_{SagS} together with HptB. These protein samples were prepared in a buffer containing 20 mM NaPO₄, 50 mM NaCl, 1 mM DTT, 0.01% NaN₃ (w/v) and 10% D₂O (v/v) with a pH of 7.0. The samples containing two proteins were prepared in a 1:1 molar ratio (1 mM each; a total volume of 500 µl).

The NMR spectrum of ^{15}N , ^{13}C REC_{SagS} was obtained using the Bruker Avance 700 spectrometer at 298K, supplied with a cryoprobe. The processing of the spectra was carried out using NMRPipe (317) and further analysis was done using SPARKY 3.12 software (T.D Goddard and D.G. Kneller, UCSF, San Francisco, CA). Backbone assignments were done using HNCA, HN (CO) CA, HNCO, HNCACB, CBCACONH spectra. Assignments of side-chains were done using HCCH-TOCSY, ^{13}C -NOESY-HSQC, ^{15}N -NOESY-HSQC while aromatic residues were assigned using ^{13}C -NOESY-HSQC and ^{15}N -NOESY-HSQC. Proton

Chapter 3

chemical shifts were externally referenced to the NMR standard (Table S10). The ^{15}N and ^{13}C chemical shifts were referenced to DSS indirectly. Distance restraints were detected using ^{13}C -NOESY-HSQC and ^{15}N -NOESY-HSQC by a combination of manual and automated assignment using CYANA 2.1 (318). TALOS was used to calculate dihedral angle restraints from chemical shifts and hydrogen bond restraints based on protein structure (317).

NMR Titration

Two-dimensional heteronuclear single quantum coherence spectroscopy (HSQC) was used to analyze the molecular characteristics of HptB binding with REC_{SagS} domain. The HSQC spectra of the ^{15}N labelled REC_{SagS} protein was recorded on a Bruker Avance 700 spectrometer at 298K. Chemical shift perturbations were analyzed to map the HptB protein binding sites on ^{15}N REC_{SagS}. The ^{15}N HSQC spectrum showed one axis corresponding to the ^1H and other to ^{15}N heteronucleus. The spectrum yielded peaks for each unique proton linked to the heteronucleus being considered. The overlapping spectrum of free ^{15}N REC_{SagS} with the complex of ^{15}N REC_{SagS}-HptB was analyzed to identify the residues on ^{15}N REC_{SagS} involved in HptB interaction.

Generation of HapZ mutant library

Primers were designed to generate a HapZ mutant library using the plasmid pBT-HapZ as a template (Table S3). The plasmid pBT-HapZ was constructed by fusing HapZ (PA2799) gene to the bacteriophage λ repressor protein (λcI) in the pBT vector. PCR amplification using the primers listed in Table S3 was used to generate 21 single and 10 double mutants by amino acid substitution using the Quik change site-directed mutagenesis kit (Agilent Technologies Catalog #200521). The HapZ mutants were used as baits to screen their interaction with the phosphoreceiver domain of the target protein SagS (REC_{SagS}) fused between a 3'-end of the RNAP- α gene and the BamHI site of the pTRG vector. Another HapZ mutant, pBT-HapZ-T24Stop, expressing the HapZ N-terminal loop alone was constructed by designing primers to introduce a stop codon at the end of the N-terminal loop (24aa). The mutagenesis PCR products were digested with *DpnI* for 1 hour at 37°C to eliminate the methylated parental plasmid and transformed into the reporter strain (*E. coli* XL1-Blue MRF' Kan competent cells) by heat shock transformation. The cells were grown in

Chapter 3

a rich SOC (Super optimal broth with catabolite repression) medium at 30 °C for 2 hrs before being plated onto LB medium containing 25 µg/ml chloramphenicol and incubated at 30 °C. The transformed colonies were cultured in Luria Bertani medium containing 25 µg/ml chloramphenicol for 24 hours at 30 °C for plasmid extraction and confirmed by sequencing analysis.

Generation of REC_{SagS} mutant library

Primers were designed to generate a REC_{SagS} mutant library using the plasmid pTRG-REC_{SagS} as a template (Table S3). The plasmid pTRG- REC_{SagS} was constructed by fusing REC_{SagS} (PA2824₆₆₃₋₇₈₆) gene between a 3'-end of the RNAP- α gene and the BamHI site of the pTRG vector. PCR amplification using the primers listed in Table S3 was used to generate 17 single mutants and 1 double mutant by amino acid substitution using the Quik change site-directed mutagenesis kit (Agilent Technologies Catalog #200521). The REC_{SagS} mutants were used to screen their interaction with HapZ fused to the bacteriophage λ repressor protein (λ cI) in the pBT vector. The mutagenesis PCR products were digested with *DpnI* for 1 hour at 37 °C to eliminate the methylated parental plasmid and transformed into the reporter strain (*E. coli* XL1-Blue MRF' Kan competent cells) by heat shock transformation. The cells were grown in a rich medium (SOC) at 30 °C for 2 hrs before being plated onto LB medium containing 12.5 µg/ml tetracycline and incubated at 30 °C in dark. The transformed colonies were cultured in Luria Bertani medium containing 12.5 µg/ml tetracycline for 24 hours at 30 °C for plasmid extraction and confirmed by sequencing analysis.

Bacterial two-Hybrid assay

Bacterial two-hybrid screening was done using the Bacteriomatch II two-hybrid screening kit (Agilent Technologies Catalog #240065). The HapZ and REC_{SagS} mutants were co-transformed into the reporter strain (*E. coli* XL1-Blue MRF' Kan competent cells) with their target vectors pTRG-REC_{SagS} and pBT-HapZ respectively using heat shock transformation. The recovery of the cells was done initially for 2 hours at 30 °C in rich SOC medium and subsequently for another 2 hours at 37 °C in the His-dropout broth minimal medium before plating. The co-transformed cells were plated onto non-selective screening medium and incubated at 37 °C for 24 hours. The interaction of the bait and the target was validated by

Chapter 3

dotting the co-transformed colonies on selective screening medium plates supplemented with 5 mM 3-amino-1,2,4- triazole (3-AT) at 37°C. Growth on 3-AT plates was considered as an indication of protein-protein interaction. The plates were scanned every 2 hours from the 12-hour to 24-hour time points to monitor growth.

Swarming motility assay

A subset of the HapZ mutants was cloned into the overexpression vector pUCP18 using the primers listed in Table S3. The resulting pUCP18-HapZ mutants were transformed into *ΔhapZ* strain of *P. aeruginosa* PAO1 using electroporation. The pUCP18-HapZ mutants that failed to interact with REC_{SagS} in the bacterial two-hybrid screening were transformed into *ΔhapZ* strain of *P. aeruginosa* PAO1 using electroporation. Swarming motility was analyzed for these strains on 0.5% agar M8 salt plates supplemented with 0.2% glucose, 2 mM MgSO₄ and 0.5% casamino acid. 1.5 µl of cell culture grown at 37 °C overnight in LB medium were dotted on freshly prepared swarming plates and incubated for 16 hours at 37 °C. The swarming patterns of the mutants were compared to *ΔhapZ* strain overexpressing the wild-type pUCP18-HapZ plasmid.

Colony morphology assay

The colony morphology assay was performed using the *ΔhapZ* strains overexpressing different HapZ mutants by following the procedure described previously (278). Colony morphology agar plates were prepared by autoclaving 1% Tryptone with 1% Bacto agar. Coomassie blue (20 µg/ml) and Congo red (40 µg/ml) were added to the medium after cooling it to 60 °C with or without 250 µg/ml carbenicillin. Plates were poured (35-40 ml agar each) and allowed to dry with closed lids for a period of 24 hours. Colonies picked from single streak plates were pre-cultured in LB medium for 12 hours. Ten microliters of the pre-culture was spotted at the centre of each plate and allowed to dry for 15 minutes. The plates were parafilmed and incubated at 25 °C in dark for 3 to 6 days. Colony images were taken by using Canon 1000D DSLR camera fitted with 18 × 55 mm lens.z

3.3. Results

3.3.1. Specific protein-protein interaction between the receiver domain of SagS and HptB

The REC domain of SagS (REC_{SagS}) adopts a typical CheY-like structure³

The receiver domain (REC) is located at the C-terminal end of the hybrid sensor histidine kinase SagS and consists of 124 residues (663 to 786aa). We determined the NMR solution structure of REC_{SagS} revealing that it has a typical (α/β)₅ fold. The structure was based on a total of 2,405 NMR-derived distance restraints and 182 dihedral angle restraints. The ensemble of 20 low-energy structures is shown in Fig. 20A. Analysis of structural ensembles using PROCHECK-NMR showed that 79.2% and 20.6% of residues lie in the most favoured and allowed regions, respectively (Fig. S5). Structural statistics for NMR ensemble are given in Table S11. There are no distance violations greater than 0.5 Å or dihedral angle violations greater than 5°. The structures were well defined in solution. The root mean square deviation (RMSD) values for REC_{SagS} relative to the mean coordinate consisting of 20 conformers were 0.48 ± 0.07 Å for the backbone atoms and 0.89 ± 0.05 Å for the heavy atoms (Fig. 20B). The REC_{SagS} domain shares sequence and functional homology with the bacterial response regulator CheY in *E. coli* (Fig. 20B and C). The topology exhibited by REC_{SagS} is a feature shared by CheY (319) and several other response regulator domains whose structures have been solved (189, 193, 320). A few examples include the transcriptional regulator NarL from *E. coli* (321), sporulation response regulator Spo0F from *Bacillus subtilis* (306), Spo0A from *Bacillus stearothermophilus* (322), nitrogen regulatory protein NtrC (323) and methylesterase CheB (324) from *Salmonella typhimurium*, chemotaxis regulator CheY₆ from *Rhodobacter sphaeroides* (325), environmental stress response regulator SLN1 from *Saccharomyces cerevisiae* (312) and ethylene receptor ETR1 from *Arabidopsis thaliana* (326).

The REC_{SagS} domain is characterized by alternately arranged β and α strands forming a parallel central five-stranded β sheet flanked by two α helices (α 1 and α 5) on one side and three α helices (α 2, α 3 and α 4) on the other. The highly conserved aspartate residue Asp-713, representing the phosphorylation site, is located at the C-terminus of the β 3 strand in a pocket that is conserved among all the response regulator domains. Two residues, Gln-669 and Asp-

³ The ¹³C,¹⁵N-REC_{SagS} labelled protein sample was prepared by me. The solution structure of REC_{SagS} was solved by Dr. Shin Joon from Prof. Yoon Ho Sup's lab, SBS, NTU.

Figure 20 | NMR solution structure of phosphoreceiver domain of SagS. A. Ensemble structure of REC_{SagS}. Superposition of the backbone traces from the final ensemble of 20 solution structures of REC_{SagS} determined by NMR spectroscopy. Secondary structural elements are highlighted in red (α helices) and blue (β sheets). B. Ribbon representation of REC_{SagS}. NMR structure is displayed, using same colour scheme as the ensemble for secondary structural elements. The root mean square deviation (RMSD) values for REC_{SagS} relative to the mean coordinate consisting of 20 conformers were 0.48 ± 0.07 Å for the backbone atoms and 0.89 ± 0.05 Å for the heavy atoms. C. The phosphoreceiver domain of SagS has an overall $(\alpha/\beta)_5$ fold. It consists of five β sheets and five α helices arranged in an alternate manner (left). The centrally arranged β sheets are flanked on either sides by α helices. The conserved aspartate residue Asp-713 shown with a red stick representation lies at the centre in a pocket at the C-terminal end of the β_3 sheet (right).

670 implicated in metal ion binding follow the N-terminal β_1 sheet. The α_3 helix begins with a conserved Gly-721, four residues after Pro-717. The C-terminal residue of the β_4 sheet is a highly conserved Thr-743 that is followed by Ala-744 that are involved in binding to the phosphoryl group and enabling access to the phosphorylation site respectively. The Lys-765 at the C-terminal end and a moderately conserved Tyr-762 in the middle of the β_5 sheet are

Chapter 3

critical for mediating conformational changes during phosphorylation. The $\alpha 1$, $\alpha 3$ and $\alpha 5$ helices are primarily composed of hydrophobic residues and hence are amphipathic in nature.

Although the overall topology of REC_{SagS} is similar to CheY and known response regulators (306, 319, 321-324, 326), REC_{SagS} has a shorter $\alpha 5$ helix compared to CheY which is analogous to the SLN1 response regulator in *Saccharomyces cerevisiae* that also interacts with an Hpt protein YPD1 (312). The $\alpha 5$ helix is shorter by 4 residues in case of REC_{SagS} and 5 residues in case of SLN1. REC_{SagS} and SLN1 also have a C-terminal extended loop following the $\alpha 5$ helix and a longer $\alpha 2$ - $\beta 3$ loop which is absent in CheY (312). However CheY₆, a response regulator in *Rhodobacter sphaeroides* that interacts with an Hpt protein CheA₃ was observed to be quite similar to CheY in terms of the length of its $\alpha 5$ helix, C-terminal extended loop and $\alpha 2$ - $\beta 3$ loop unlike REC_{SagS} and SLN1 (325).

The receiver domain of SagS (REC_{SagS}) interacts with HptB.⁴

In order to understand the molecular basis for the specific interaction between REC_{SagS} and HptB in the two-component signalling pathway, we used NMR to characterize the heteronuclear single quantum coherence spectroscopy (HSQC) of ¹⁵N labelled REC_{SagS} with HptB. In the ¹⁵N HSQC spectrum, one axis corresponds to the ¹H and other corresponds to a heteronucleus (most often ¹⁵N or ¹³C). Thus, the spectrum yields peaks for each unique proton linked to the heteronucleus being considered. The backbone and side chain amide protons contribute to peaks in the HSQC spectra of a protein.

Chemical shift perturbations from the free REC_{SagS} domain to its complex with HptB were observed for several residues on REC_{SagS} (Fig. 21A and B). The HSQC spectra of labelled REC_{SagS} show that REC_{SagS} interacts with HptB primarily using 20 residues. They displayed a chemical shift perturbation (CSP) of greater than 0.1 ppm (Fig, 21B). The most significant changes in CSP (greater than 0.2 ppm) were observed for six of the 20 residues, indicated in red (Fig 21C and 22). Of these, Val-677 (0.64 ppm) Ala-678 (0.57 ppm) and Gln-675 (0.48 ppm) are located on its $\alpha 1$ helix, Gly-721 (0.22 ppm) is located on its $\beta 3$ - $\alpha 3$ loop while Asn-745 (0.28 ppm) and Leu-747 (0.24 ppm) are located on its $\beta 4$ - $\alpha 4$ loop (Fig. 21B and 22).

⁴ The ¹⁵NREC_{SagS} and HptB protein samples were prepared by me. The HSQC titration of REC_{SagS} with HptB as well as NMR assignments were performed by Dr. Shin Joon from Prof. Yoon Ho Sup's lab, SBS, NTU.

Figure 21 | NMR titration of REC_{SagS} and HptB. A. Overlay of ^1H - ^{15}N HSQC spectra of uniformly ^{15}N labeled REC_{SagS} domain (red) and its complex with HptB (blue) at a molar ratio of 1:1. B. Expanded view of chemical shift perturbations upon HptB binding. The perturbations of chemical shifts from free REC_{SagS} domain (red) to its complex with HptB are displayed by arrows. B. Plot of chemical shift perturbations of REC_{SagS} upon HptB binding at a molar ratio of 1:1 (REC_{SagS} : HptB). The difference of chemical shifts were calculated using the following formula, $\Delta\delta = [(\text{}^1\text{H}_{\text{free}} - \text{}^1\text{H}_{\text{bound}})^2 + (\text{}^{15}\text{N}_{\text{free}} - \text{}^{15}\text{N}_{\text{bound}})^2]^{1/2}$. The residues showing CSP more than 0.1 are labelled. C. Mapping of HptB binding site and interaction residues on REC_{SagS} based on ^1H , ^{15}N HSQC titration data with HptB. Perturbed amino acids were displayed according to degree of chemical shift perturbations: red (> 0.2), blue ($0.1 < d < 0.2$), and green (disappeared resonances).

Chapter 3

Figure 22 | Structure based sequence alignment of the 11 REC domain-containing hybrid sensor histidine kinases from *P. aeruginosa* with the REC domain of SLN1 hybrid histidine kinase from *S. cerevisiae*. The highly conserved residues are shaded red and the less conserved residues are marked in red with blue boxes. The secondary structure of REC_{SagS} based on its NMR solution structure is shown above its sequence. The four hybrid kinases involved in binding HptB are marked by brown triangles on the left. The residues on REC_{SagS} identified by HSQC titration involved in HptB binding are marked on REC_{SagS} as circles. The red circles denote the perturbed amino acids that exhibited the greatest degree of chemical shift perturbation (CSP) of greater than 0.2 ppm. The amino acids denoted by blue circles showed a CSP between 0.1 and 0.2 ppm while the amino acids denoted by green circles showed disappeared resonances. The sequence alignment was performed using ESPript 3.0 (327).

The buried hydrophobic residues (Val-677 and Ala-678) are flanked by the polar residues, Gln-675 and His-683, forming the binding interface with HptB on the $\alpha 1$ helix of REC_{SagS}. On the other hand, the residues indicated in blue showed a CSP between 0.1 and 0.2 ppm (Fig. 21C and 22). Most of these residues, namely, Val-673, Ala-692, Gly-695, Leu-696, Val-718, Leu-742, Tyr-762 and Leu-763 are hydrophobic in nature. The CSPs observed for few

Chapter 3

of these residues, especially Asp-760 and Gln-688, can be attributed to be as a result of conformational changes in REC_{SagS} upon binding HptB. Apart from the residues on REC_{SagS} that exhibited a CSP of greater than 0.1 ppm, two residues namely, Gly-680 and Leu-681 indicated in light green (Fig 21 C) showed disappeared resonances. Studies on the interactions between REC and Hpt domain proteins such as YPD1/SLN1 complex in *Saccharomyces cerevisiae*, Spo0B/Spo0F complex in *Bacillus subtilis* and CheA₃/CheY₆ complex in *Rhodobacter sphaeroides* have also demonstrated that the binding surface in REC-Hpt interactions involve a hydrophobic pocket flanked by hydrophilic residues. (312, 325, 328). This is in agreement with our NMR titration data.

Among the 11 hybrid sensor kinases identified in *P. aeruginosa* (209), three hybrid kinases [PA1611, PA1976 and PA4856 (RetS)] are known to interact with HptB in addition to REC_{SagS} (228). However, unlike SagS, PA1611 and PA1976, the sensor kinase RetS is incapable of undergoing autophosphorylation (228). Sequence alignment of all the 11 hybrid histidine kinases revealed that the REC domains of the three other hybrid kinases interacting with HptB share similar hydrophobicity around the residues Val-677 and Ala-678 as observed in REC_{SagS} (Fig. 22). Of these, Ala-678 is always a hydrophobic amino acid. Val-677 is conserved in PA1611 and RetS, while Gln-675 is conserved in PA1611. On the other hand, the polar residues Gln-675 and His-683 flanking these hydrophobic residues are more variable in nature. Similarly, Gly-721 is highly conserved among the receiver domains of all *P. aeruginosa* hybrid sensor kinases. Asn-745 is conserved in PA1611 and PA1976 while Leu-747 is conserved in PA1611 and RetS.

The three Hpt proteins from P. aeruginosa adopt a typical all-helical structure.

In order to generate a docking model of REC_{SagS}-HptB interaction, a homology model of HptB was generated with ROBETTA server (329) by comparative modelling using the Hpt protein ArcB (PDB ID: 2A0B) from *E. coli* as a template (Fig. 23). In addition to HptB, the models of two other Hpt proteins from *P. aeruginosa*, HptA and HptC, were also generated with ROBETTA server (329) by comparative modelling using the Hpt protein Ahp2 (PDB ID: 4PAC) from *Arabidopsis thaliana* as a template (Fig. 23).

The ROBETTA server performs automated protein structure prediction by either using comparative or *de novo* method. It is a much more reliable structure prediction method

Chapter 3

compared to SWISSMODEL since it is fully automated and analyses each residue in the submitted sequence using an advanced K*Sync alignment protocol to produce a high quality model. Comparative modelling approach is used by ROBETTA when an appropriate confident match to the input sequence is detected using BLAST, PSI-BLAST, FFAS03 or 3D-Jury. On the other hand, if no match is detected, predictions are based on the Rosetta fragment-insertion method. The ROBETTA server used comparative modelling approach for the three Hpt proteins from *P. aeruginosa*. In this method, the initial step of domain prediction by the server is referred to as ‘Ginzu’, which utilizes BLAST, PSI-BLAST and other approaches to detect homologous structures to the input sequence and predict domains using multiple-sequence alignment (MSA). Regions that display homology to known structures are computationally modelled whereas the unassigned regions are assumed to be domain linkers if they are less than 50 residues in length. In the final step, the remaining

Figure 23 | Comparative modeling of the histidine phosphotransfer proteins – HptA, HptB and HptC from *P. aeruginosa* using Robetta server (329). The HptB homology model (left cyan) was generated with ROBETTA server (329) by comparative modelling using the Hpt protein ArcB (PDB ID: 2A0B) from *E. coli* as a template. HptB displays an all- α helical structure and consists of five α helices arranged in a four-helix bundle. The five α -helices are labelled as α A, α B, α C, α D and α E. The conserved histidine residue His-57 representing the site of phosphorylation is located in the middle of the α C helix and is shown as a purple stick representation. Similarly, the homology models of the other two Hpt proteins, HptA and HptC from *P. aeruginosa* were also generated with ROBETTA server by comparative modelling using the Hpt protein AHP2 (PDB ID: 4PAC) from *Arabidopsis thaliana* as a template. HptA (center pink) and HptC (right yellow) consist of six and five α helices respectively arranged in a four-helix bundle. The α -helices are labelled as α A, α B, α C, α D, α E and α F. Unlike HptB and HptC, HptA has an additional short α helix preceding a longer N-terminal unstructured loop. The conserved histidine residues His-65 (light green) and His-55 (red) of HptA and HptC respectively are shown as stick representations. They are located in the middle of the α D and α C helices of HptA and HptC respectively.

unassigned regions are searched using MSA and modelled using a strong loop prediction by PSIPRED (329). The histidine phosphotransfer protein B (HptB) consists of 116 residues and

Chapter 3

participates in the two-component phosphorelay by acting as an intermediate transmitting the phosphate group from REC_{SagS} and three other hybrid kinases to the cognate response regulator PA3346 (HsbR). The homology model of HptB consists of five α helices (α A, α B, α C, α D and α E) arranged as a four-helix bundle (Fig. 23). The four-helix bundle in HptB is organized in an up-down-up-down manner with a left-handed rotation. The active histidine residue involved in phosphotransfer is His-57 located at the centre of the α C helix of HptB similar to His-64 in YPD1 and other Hpt domains like CheA-Hpt and YojN (330-332). Gly-61 in HptB is situated on the adjacent fold near His-57 and provides greater accessibility for the conserved histidine imidazole ring to the surrounding solvent by creating an empty space. This allows the histidine residue to be easily accommodated by the active pocket of the cognate receiver domains (333).

The α A helix of HptB forms a cap over the helices α B and α E and interacts with the loop connecting helices α C and α D extensively. The residue at position 60 corresponding to lysine in the α C helix of HptB is mostly basic in nature among all Hpt domains. Together with Lys-76, it constitutes a positive patch near the phosphorylation site His-57. The positive patch may neutralize the negative charge of the phosphate group attached near the histidine residue. Additionally, these positively charged residues can act as stabilizers by hydrogen bonding to oxygen atoms of the histidine phosphate group and provide acidic pockets for enhancing binding of Hpt proteins to response regulators (330). The hydrophobic residues forming the linker between the α C and α D helices are nearly conserved.

The histidine phosphotransfer protein A (HptA) or PA0991 consists of 125 residues and six α helices (α A, α B, α C, α D, α E and α F) arranged as a four-helix bundle. HptC or PA0033, on the other hand, consists of 121 residues and five α helices (α A, α B, α C, α D and α E) arranged as a four-helix bundle (Fig. 23). The four-helix bundles in both HptA and HptC are organized in an up-down-up-down manner with a left-handed rotation. Unlike HptB and HptC, HptA has a longer N-terminal unstructured loop followed by an additional short α A helix. The active histidine residue involved in phosphotransfer is His-65 located at the centre of the α D helix of HptA and His-55 in case of HptC, located at the centre of the α C helix.

A model of REC_{SagS}- HptB complex

The model of REC_{SagS}- HptB complex was built by aligning the two proteins with the 2.1Å resolution crystal structure of the YPD1/SLN1 complex in *S. cerevisiae* (PDB ID: 1OXB) using PyMOL (Fig. 26A & B) (312). The YPD1/SLN1 complex was chosen for docking because both SLN1 and YPD1 share great degree of structural similarity with REC_{SagS} and HptB respectively compared to other known structures of REC and Hpt proteins (Fig. 22 and 24). The Asp-713 and His-57 residues involved in phosphotransfer were in close proximity to each other in the docking model, which was similar to the YPD1/SLN1 complex (Fig. 26A).

Figure 24 | Structure based sequence alignment of the three histidine phosphotransfer proteins from *P. aeruginosa* with the Hpt protein YPD1 from *Saccharomyces cerevisiae*. The highly conserved residues are shaded red and the less conserved residues are marked in red with blue boxes. Secondary structure of YPD1 (PDB ID: 1QSP) is shown above its sequence in blue while the secondary structure of HptB based on its homology model is shown below its sequence in blue. The α helices of the proteins are denoted using the labels α A to α F. The sequence alignment was performed using ESript 3.0 (327). The amino acids marked with green dots represent the residues involved in hydrophobic bonding on YPD1 while the amino acids marked with purple dots represent the residues on YPD1 involved in hydrogen bonding to the REC domain containing SLN1 protein. This is based on the crystal structure of the YPD1/SLN1 complex (PDB ID: 1OXB).

The 20 interacting residues on REC_{SagS} were identified on the basis of NMR titration while the corresponding residues on HptB were predicted based on the docking model using the previously studied YPD1/SLN1 complex from *S. cerevisiae*. In the REC_{SagS}-HptB model, the six residues on REC_{SagS} that showed most significant chemical shift perturbations of greater than 0.2 ppm in the HSQC titration namely, Gln-675, Val-677, Ala-678, Gly-721, Asn-745

Chapter 3

and Leu-747 form the interaction interface with HptB on its $\alpha 1$ helix, $\beta 3$ - $\alpha 3$ and $\beta 4$ - $\alpha 4$ loops. Of these residues, Val-677 is conserved and forms part of the interaction interface in both REC_{SagS} and SLN1 REC domain. Interestingly, Val-677 is hydrophobic in the REC domains of PA1611, PA1976 and RetS that are also involved in HptB interaction. On the other hand, the other three residues, Gln-675 and Ala-678 are more specific to the REC_{SagS}-HptB interaction since the corresponding residues in SLN1 are not involved in YPD1 binding (Fig. 22). Gly-721 is a highly conserved residue among all the receiver domains of hybrid kinases in *P. aeruginosa* as well as in SLN1. Asn-745 retains its polar nature among all the hybrid kinases interacting with HptB and is conserved in PA1611 and PA1976. It is replaced by a hydrophobic phenylalanine in case of SLN1, which makes it specific to REC_{SagS}-HptB interaction. Finally, Leu-747 is hydrophobic in PA1611 and RetS but is replaced by a polar residue in PA1976, SLN1 and almost all other hybrid kinases of *P. aeruginosa* (Fig. 22). The CSPs observed for few of these residues can be due to conformational changes in REC_{SagS} upon binding HptB. This can be observed especially in case of Asp-760 and Gln-688 since they lie on the outside of the binding interface (Fig 26A and C).

Figure 25 | Sequence alignment based comparison of residues on HptB predicted to interact with REC_{SagS} with the corresponding residues on two other histidine phosphotransfer proteins - HptA and HptC from *P. aeruginosa*. The 17 residues on HptB predicted to bind to REC_{SagS} are depicted as cream circles while the three polar residues namely, Glu-21, Asp-32 and Asn-65 that are specific to the HptB- REC_{SagS} interaction are shown in red. They are most likely involved in hydrogen bonding with REC_{SagS}. The corresponding residues on the other two Hpt proteins – HptA and HptC (shown in blue) are hydrophobic in nature.

Additionally, docking of REC_{SagS} and HptB with the YPD1/SLN1 complex predicted residues (312) on HptB that might be involved in this interaction (Fig. 24, 25 and 26D).

Figure 26 | Docking model of REC_{SagS} -HptB interaction. (A) Ribbon representation of docking model of REC_{SagS} -HptB interaction generated by aligning the two proteins to the crystal structure of the YPD1/SLN1 complex from *S. cerevisiae* (PDB ID: 1OXB). According to the model, 20 residues on REC_{SagS} (yellow) make contact with HptB (cyan) as identified from the NMR HSQC titration (shown in green). Corresponding to the residues on YPD1 in the YPD1/SLN1 complex, the HptB residues that make contact with REC_{SagS} (shown in violet) are distributed primarily on its αA , αB and αC helices while few are also located on its the αA - αB loop and αD helix. The phosphotransfer residues Asp-713 on REC_{SagS} (dark blue) and His-57 (purple) on HptB are shown as stick models. (B) Surface model of REC_{SagS} -HptB interaction. (C) Ribbon representation of REC_{SagS} showing the 20 residues involved (green spheres) in HptB interaction as identified from NMR titration. Of these, the six residues that showed the greatest chemical shift perturbations are located on its $\alpha 1$ helix, $\beta 3$ - $\alpha 3$ loop and $\beta 4$ - $\alpha 4$ loop. (D) Ribbon representation of HptB showing the residues involved (violet spheres) in REC_{SagS} interaction on its αA , αB , αC and αD helices.

Since, *P. aeruginosa* has three stand-alone Hpt proteins (HptA, HptB and HptC), the sequence alignment also included HptA and HptC to determine the molecular basis of the specific interaction between HptB and REC_{SagS} (Fig. 24). Analogous to the YPD1/SLN1

Chapter 3

interaction, the REC_{SagS}-HptB docking model revealed that the HptB interaction interface primarily involves the α A, α B and α C helices. Additionally, a few residues from the α D helix and the α A- α B loop were also predicted in mediating the interaction. These residues on HptB include Val-10 and Leu-11 on the α A helix, Leu-14 on the α A- α B loop, Glu-21, Leu-25, Thr-28 and Asp-32 on the α B helix, Arg-53, Ser-58, Gly-61, Gly-62, Ser-64 and Asn-65 on the α C helix and Lys-76, Glu-79 and Arg-83 on the α D helix (Fig. 24, 25 and 26D).

Of these, Leu-25 and Gly-61 are highly conserved between YPD1 and the three stand-alone Hpt proteins in *P. aeruginosa*. Leu-11 and Thr-28 in HptB are conserved and retain their hydrophobic and polar nature respectively in HptA and HptC as well as in YPD1. On the other hand, residues corresponding to Val-10 in HptB are also hydrophobic in HptC and YPD1 but replaced by a polar residue in case of HptA. Lys-76 is also hydrophilic in all the Hpt proteins (HptA, B, C and YPD1) and is highly variable in nature. Leu-14 is similarly hydrophobic in HptA and YPD1 but replaced by a polar residue in case of HptC (Fig. 25). Interestingly, the residues Glu-21, Asp-32 and Asn-65 are unique to HptB since the analogous residues in HptA, HptC and YPD1 are hydrophobic in nature (Fig. 24 and 25). The unique nature of these three residues may likely play a role in mediating specificity to the REC_{SagS}-HptB interaction.

Analogous to the YPD1-SLN1, CheA₃-CheY₆ and Spo0F-Spo0B interactions in *S. cerevisiae*, *R. sphaeroides* and *Bacillus subtilis*, the model of REC_{SagS}-HptB complex showed that the binding interface is primarily a buried hydrophobic patch (Fig. 26A). According to the REC_{SagS}-HptB docking model, this hydrophobic interaction interface is formed between the residues Val-677, Val-673 and Ala-678 on the α 1 helix of REC_{SagS} and by Val-10 and Leu-11 on the α A helix, Leu-14 on α A- α B loop, Leu-25 and Thr-28 on the α B helix and Gly-61 and Gly-62 on the α C helix of HptB. Asn-65, unique to HptB, flanks this hydrophobic patch, imparting specificity to the interaction. The residues on HptA, HptC and YPD1 corresponding to Asn-65 are all hydrophobic in nature (Fig. 24 and 25). In addition to the hydrophobic interface, specific hydrogen bonds are formed between Asn-745 on REC_{SagS} with Lys-76, Arg-83 and Glu-79 on HptB. Interestingly, Asn-745 that is involved in hydrogen bonding on REC_{SagS} is quite conserved among all the REC domain proteins interacting with HptB (Fig. 22). Thus, these specific hydrogen bonds, in addition to the hydrophobic interface could possibly confer specificity to the REC_{SagS}-HptB complex.

3.3.2. Cyclic di-GMP dependent protein-protein interaction between the receiver domain of SagS and HapZ

Structural model of HapZ

Our multiple attempts to express and purify HapZ as a stable protein using various solubilizing tags for NMR structure determination were unsuccessful due to the tendency of HapZ to aggregate at higher concentrations (Fig. S3). We had to rely on homology model of HapZ generated by SWISSMODEL (334-336) using the NMR solution structure of PA4608 (PDB code: 1YWU) as the template (130). The apo structure of PA4608 has an N-terminal unstructured loop, an antiparallel six-stranded β barrel followed by an α helix and two 3^{10} helices. Compared to PA4608, the homology model of apo HapZ has a longer unstructured loop (22 amino acids) at its N-terminus followed by a β barrel consisting of seven β sheets arranged in an antiparallel manner (Fig. 27). Although the homology model of HapZ is similar to the general structure of PilZ domain-containing proteins such as PA4608, VCA0042 (PlzD), PP4397, Alg44 and BcsA (71, 130-132, 262), it is unique due to the lack of C-terminal α helices. An additional $\beta 7$ sheet is also present in HapZ unlike the other structurally characterized PilZ proteins like PA4608, BcsA, Alg44 and VCA0042. The conserved arginine residues constituting the c-di-GMP binding motif (RxxxR) in HapZ correspond to the positions 12 and 16 respectively in its N-terminal loop whereas the aspartate, serine and glycine residues in the second conserved c-di-GMP binding motif (DxSxxG) are located at positions 38 in its $\beta 3$ sheet and 40, 43 in its $\beta 3$ - $\beta 4$ loop.

The nature of the residue located upstream of the RxxxR motif has been shown to be crucial in deciding the stoichiometry of c-di-GMP binding to a PilZ protein. A larger hydrophobic residue such as leucine (LRxxxR), serves as a hydrophobic cap favouring the binding of monomeric c-di-GMP. This was observed in case of the PilZ domain protein VCA0042 from *Vibrio cholerae* (129, 131). On the contrary, the presence of polar residues like arginine or lysine before the motif ([Q/E][R/K]RxxxR) does not sterically hinder the binding of a second c-di-GMP molecule. (131, 337). For example, the PilZ domain protein PP4397 from *P. putida* binds to two intercalated molecules of c-di-GMP. However, substitution of the upstream residue arginine to leucine in PP4397 altered its binding stoichiometry to 1:1 although it retained its binding affinity to c-di-GMP (131). HapZ contains a polar residue (Lys-11) preceding the RxxxR motif which suggests that HapZ may bind a dimeric c-di-GMP molecule like PA4608, PP4397, BcsA and Alg44 (337).

Figure 27 | HapZ adopts a typical β barrel structure. (A) NMR structure of apo PA4608 (PDB code: 1YWU). PA4608 consists of an N-terminal unstructured loop, an antiparallel six-stranded β barrel, an α -helix and two 3^{10} helices ($\eta 1$ and $\eta 2$) at its C-terminal end (130). (B) Homology model of HapZ generated with SWISSMODEL (334-336) using PA4608 apo structure (PDB code: 1YWU) as a template. HapZ consists of a N-terminal unstructured loop and an antiparallel seven-stranded β barrel. Unlike PA4608, HapZ lacks α -helices at its C-terminal end.

Replacement of six residues in HapZ abolishes HapZ-REC_{SagS} interaction

To elucidate the molecular basis of HapZ- REC_{SagS} interaction, amino acid residues on the two proteins were chosen for mutagenesis based on sequence alignment of HapZ and REC_{SagS} with different PilZ domain proteins and REC domain-containing hybrid histidine kinases respectively in *P. aeruginosa* (Fig. 28A and 29A). Sequence alignment was done to recognize conserved and non-conserved residues on both HapZ and REC_{SagS}. A total of 50 mutants in both HapZ and REC_{SagS} were generated using site-directed mutagenesis. Sixteen non-conserved surface residues were selected in HapZ that spanned the entire protein. In addition to non-conserved residues, four conserved residues that form part of the c-di-GMP binding motif (RxxxR and DxSxxG) as well as three other nearly invariant residues identified from sequence alignment were also chosen for mutagenesis (Fig. 28B). Similarly, in REC_{SagS}, fifteen non-conserved and four invariant surface residues were selected for mutagenesis. The conserved residues include the Asp-713 that is the site of phosphorylation, Thr-743 involved in binding to the phosphoryl group, Tyr-762 that mediates conformational changes during

Chapter 3

A

B

Figure 28 | Residues chosen for HapZ mutant library. (A) Structure-based sequence alignment of PilZ domain-containing proteins using ESPrnt 3.0 (327). The secondary structures of HapZ and apo PA4608 (PDB code: 1YWU) are shown above and below their respective sequences. The highly conserved residues are shaded in red while the less conserved residues are marked in blue with blue boxes. The residues chosen for mutagenesis are indicated by triangles in the HapZ sequence. The residues that did not abolish HapZ-REC_{SagS} interaction are denoted as yellow triangles while the six residues that abolished the interaction are denoted as green triangles. (B) Homology model of HapZ displaying the residues chosen for site-directed mutagenesis. The green spheres denote the residues that did not abolish HapZ-REC_{SagS} interaction and the orange spheres indicate the six residues that abolished the interaction.

Chapter 3

A

B

Figure 29 | Residues chosen for REC_{SagS} mutant library. (A) Structure-based sequence alignment of REC domain-containing proteins in *P. aeruginosa* using ESPrpt 3.0 (327). The secondary structure of REC_{SagS} determined using NMR (Chapter 3) is shown above its sequence. The highly conserved residues are shaded in red while the less conserved residues are marked in red with blue boxes. The residues chosen for mutagenesis are indicated by yellow triangles in the REC_{SagS} sequence. (B) Structure of REC_{SagS} displaying the residues chosen for site-directed mutagenesis. The light green spheres denote the residues chosen on REC_{SagS} for site-directed mutagenesis

Chapter 3

phosphorylation and Pro-766 that forms a *cis*-peptide bond with a conserved Lys-765. The residues involved in REC_{SagS} phosphorylation were chosen since HapZ was demonstrated to block the phosphotransfer from SagS to HptB in our studies from Chapter 2. The non-conserved residues Gln-675 and His-683 were identified to interact with HptB in Chapter 3 using NMR titration. They were also chosen for mutagenesis to test whether HapZ and HptB competitively interact with REC_{SagS} (Fig. 29B).

Nearly all the chosen residues in both proteins were mutated to alanine. Alanine truncates amino acid side chains at the β carbon unlike glycine that also removes the β carbon. The retention of the β carbon minimizes the conformational changes on the protein, facilitating interpretation of results (338, 339). However, the glycine (Gly-92) residue selected in HapZ was mutated to a polar amino acid glutamate instead of alanine, since alanine is hydrophobic in nature and also to observe the effect of charge change on HapZ- REC_{SagS} interaction. In addition to alanine substitution, the charges of three other residues on HapZ namely Glu-31, Glu-67 and Glu-73 were also reversed from negative to positive (arginine) while generating their double mutants. The mutant libraries of HapZ and REC_{SagS} were screened for interaction against the wild type REC_{SagS} and HapZ proteins respectively using the bacterial two-hybrid screening assay. Mutation of six residues in HapZ abolished interaction with REC_{SagS}. Of these six residues, Arg-12 and Arg-16 form part of the RxxxR motif while Asp-38 and Ser-40 form part of the DxSxxG motif that are involved in binding c-di-GMP (Fig. 30A and 30B). This supports our findings in Chapter 2 that c-di-GMP binding to HapZ is crucial in strengthening its interaction with REC_{SagS} (314).

In addition to the c-di-GMP motifs, the non-conserved Arg-80, and the invariant Gly-92 located in the β 5 and the β 6 sheets of the β barrel are also important for this interaction. Another mutant expressing only the N-terminal loop (Thr-24 to Stop) also abolished the HapZ- REC_{SagS} interaction. This signifies the importance of the β barrel in this interaction. Additionally, five double mutants of HapZ showed loss of interaction with REC_{SagS}. Of these, two of them were c-di-GMP binding motifs (R12AR16A and D38AS40A). However the other three double mutants had a mutation in Arg-80 residue in addition to mutations in Val-79, Glu-82 and Gly-92 respectively. Since single mutants of Val-79 and Glu-82 were not able to abolish interaction with REC_{SagS} (Fig. 29A and 29B), the effect seen in the V79AR80A and R80AE82A mutants was attributed to the Arg-80 residue. On the other hand, the loss of

Chapter 3

interaction of the R80AG92E double mutant is contributed by both Arg-80 and Gly-92 residues (Fig. 30B). Although six residues on HapZ were identified for mediating HapZ-RECSagS interaction, none of the 19 mutants in RECSagS could abolish interaction with HapZ in the bacterial two-hybrid system (Fig. 31). Since the mutation of residues Gln-675 and His-683 involved in RECSagS-HptB interaction did not affect interaction with HapZ in the bacterial two-hybrid system, it is possible that HapZ and HptB may not competitively bind to RECSagS.

Figure 30 | Bacterial two-hybrid screening of HapZ mutants against RECSagS. Protein-protein interaction was tested by checking growth of co-transformed colonies on a selective screening medium containing 5 mM 3-AT. (A) No growth in the selective medium was detected in five single mutants of HapZ namely R12A, R16A, R80A, G92E and Thr-24-Stop indicating loss of HapZ-RECSagS interaction. (B) Bacterial two-hybrid screening of HapZ double mutants against RECSagS. Loss of protein-protein interaction was observed on the selective medium in case of five double mutants namely R12AR16A, D38AS40A, V79AR80A, R80AE82A and R80AG92E.

Chapter 3

Figure 31 | Bacterial two-hybrid screening of REC_{SagS} mutants against HapZ. Protein-protein interaction was tested by checking growth of co-transformed colonies on a selective screening medium containing 5 mM 3-AT. Bacterial two-hybrid screening of REC_{SagS} mutants against HapZ. Abolishment of protein-protein interaction on the selective medium was not observed in any of the 19 mutants of REC_{SagS}.

Phenotypic assays validate the role of six residues of HapZ in abolishing binding with REC_{SagS}

In order to evaluate the importance of the six residues identified from the bacterial two-hybrid screening, two phenotypic assays were performed. Swarming motility is the flagella-mediated translocation of a bacterial population across solid or semi-solid surfaces in a coordinated manner. *P. aeruginosa* displays a characteristic dendritic (or tendrill) colony pattern along with the production of surfactants during swarming (281). In Chapter 2, we have shown that HapZ affects swarming motility. Our studies revealed that $\Delta hapZ$ strain of PAO1 exhibits a hyperswarming phenotype, which could be reversed by complementation with HapZ. On the other hand, HapZ overexpression reduces swarming to a great extent in PAO1 (314). Similarly, we assessed the ability of the six HapZ mutants that abolished interaction with REC_{SagS} to reverse the hyperswarming phenotype of the $\Delta hapZ$ strain of PAO1. The results showed that the six mutant strains were unable to complement the $\Delta hapZ$ phenotype like the wild type HapZ (Fig. 32A). Since mutation of RxxxR and DxSxxG motifs could not rescue the hyperswarming phenotype of $\Delta hapZ$, it indicates that c-di-GMP binding to HapZ is significant in mediating binding to REC_{SagS}. In addition to the c-di-GMP binding

Chapter 3

motifs, the results also validated the importance of the Arg-80 and Gly-92 residues in the HapZ- REC_{SagS} complex.

Figure 32 | Phenotypic assays validate the importance of the six HapZ residues in HapZ- REC_{SagS} interaction. (A) Swarming motility assay to confirm the role of the six HapZ mutants in the HapZ- REC_{SagS} interaction. The $\Delta hapZ$ strain exhibits a hyperswarming phenotype that can be rescued by HapZ complementation. Overexpression of HapZ in wild type PAO1 strain inhibits swarming to a greater extent. The six HapZ mutants failed to complement the hyperswarming phenotype of the $\Delta hapZ$ strain unlike HapZ, confirming their role in the HapZ- REC_{SagS} binding. (B) Colony morphology assay of HapZ mutants. Colony wrinkling is a measure to maximize oxygen availability during biofilm formation. The $\Delta hapZ$ strain displays a smooth colony phenotype that can be rescued by HapZ. Overexpression of HapZ caused hyperwrinkling of colonies. Overexpression of the six HapZ mutants that abolished binding to REC_{SagS} could not rescue the $\Delta hapZ$ phenotype suggesting their importance in biofilm formation and HapZ- REC_{SagS} binding. However, five other HapZ mutants in its N-terminal loop could complement the smooth phenotype of the $\Delta hapZ$ strain indicating that their role in this interaction is not crucial as the other six HapZ mutants.

Chapter 3

Apart from regulating swarming motility, HapZ also controls biofilm formation in *P. aeruginosa*. Colony morphology is a phenotypic assay that correlates biofilm morphological development to intracellular redox state in *P. aeruginosa*. Different electron acceptors such as oxygen, nitrate and phenazines are utilized by *P. aeruginosa* for redox control that affect biofilm community structure. Colony wrinkling is a morphological adaptation to maximize oxygen availability to cells that experience steep oxygen gradients due to both consumption and limited diffusion following biofilm formation (278). We have demonstrated in Chapter 2 that $\Delta hapZ$ colonies display a smooth phenotype that can be reversed to the PAO1 phenotype by complementation. A hyperwrinkled phenotype is observed in colonies overexpressing HapZ in PAO1, corroborating its role in biofilm formation. We evaluated the colony morphology phenotype of $\Delta hapZ$ upon complementation with the six HapZ mutants that abolished its binding with REC_{SagS} and five other HapZ mutants in its N-terminal loop. It was observed that complementation of $\Delta hapZ$ with the five N-terminal loop mutants of HapZ could reverse its phenotype like the wild type whereas the overexpression of six HapZ mutants that abolished REC_{SagS} binding in the $\Delta hapZ$ strain displayed a smooth phenotype without any complementation (Fig. 32B). This confirmed the role of the six HapZ residues and supports the importance of c-di-GMP in the HapZ-REC_{SagS} interaction.

3.4. Discussion

Two-component signalling pathways must maintain a high degree of specificity between histidine kinases and their cognate response regulators to avoid cross talk. The phosphotransfer between proteins is an important means of relaying information to achieve targeted cellular responses. A single organism can have multiple two-component systems that can be active at the same time, which raises the question of how phospho-signalling fidelity is achieved. In the current study, we investigate how the binding specificity between multiple receiver domains and an Hpt domain protein is achieved at a molecular level using structural studies in *P. aeruginosa*. The multistep two-component signalling pathway involving REC_{SagS} and HptB plays a significant role in regulating biofilm formation in *P. aeruginosa*. Previous studies have shown that HptB can interact with multiple histidine kinases. However, the precise mechanism by which monomeric HptB recognizes the REC domains of four hybrid histidine kinases remains unclear. Here we report the binding mechanism of HptB with one of its four-hybrid histidine kinases partner- REC_{SagS} and the possible basis of the specificity.

Chapter 3

To understand the promiscuous nature of this interaction, we determined the NMR solution structure of the receiver domain of SagS and recognized the important residues responsible for mediating the REC_{SagS}-HptB interaction by NMR studies. We demonstrated that REC_{SagS} exhibits a characteristic CheY-like (α/β)₅ fold (340) that is shared by several other REC domain-containing proteins such as CheY, Spo0A, Spo0F, CheB, NtrC, NarL, SLN1, CheY₆ and ETR1 belonging to the CheY superfamily (319, 322-324, 326). The positioning of the conserved aspartate residue (Asp-713) in the active site in REC_{SagS} is similar to that of CheY and other REC domains. This suggests that phosphorylation in REC_{SagS} probably occurs in the same manner as CheY. Despite being analogous to CheY, REC_{SagS} has certain distinctive features that it shares with SLN1, another REC domain-containing hybrid kinase that interacts with an Hpt protein YPD1 in *Saccharomyces cerevisiae* (330).

Studying REC-Hpt interactions have inherent difficulties owing to their transient and labile nature. However, Xu and colleagues reported the crystal structure of the complex between the hybrid sensor kinase - response regulator domain SLN1 and the Hpt protein YPD1 in *Saccharomyces cerevisiae* (312). Another example of such an interaction is between the phosphorylated CheA₃P1 and CheY₆ in *Rhodobacter sphaeroides* (325). A similar interaction mechanism was also observed in case of Spo0F/Spo0B complex in *Bacillus subtilis* (328). The mechanism of REC_{SagS}-HptB interaction was proposed on the basis of NMR titration of REC_{SagS} with HptB in combination with docking the two proteins to a previously studied REC-Hpt complex crystal structure (PDB ID: 1OXB) in *S. cerevisiae* involving the REC domain of the sensor kinase SLN1 and the Hpt protein YPD1 (312). A reliable homology model of HptB was generated using ROBETTA comparative modelling with ArcB (PDB ID: 2A0B) from *E. coli* as a template. The model of HptB was in good agreement to the general structure of Hpt proteins with a four-helix bundle fold. The 20 interacting residues on REC_{SagS} were identified on the basis of NMR titration while the corresponding residues on HptB were predicted based on the docking model using the previously studied YPD1/SLN1 complex from *S. cerevisiae*. The model of REC_{SagS}-HptB complex represented a common interaction scheme between REC domains and Hpt proteins in two-component pathways.

Analogous to other REC-Hpt interactions, the REC_{SagS}-HptB binding is also facilitated by a combination of hydrophobic interactions and hydrogen bonds. The residues on REC_{SagS}

Chapter 3

involved in hydrophobic interaction are located on its $\alpha 1$ helix. Binding of the two proteins results in a conformational change in the $\alpha 1$ helix of REC_{SagS}. Of the residues in its $\alpha 1$ helix, Val-677 is quite conserved in both REC_{SagS}-HptB as well as YPD1/SLN1 complexes. The $\alpha 1$ helix of REC_{SagS} binds to the hydrophobic residues distributed on the αA , αB and αC helices of HptB, which is consistent with the YPD1/SLN1 interaction. Although the general hydrophobic nature of the REC-Hpt interactions is conserved in case of the REC_{SagS}-HptB complex, specificity of interaction is achieved from specific hydrogen bonds between polar residues on REC_{SagS} and HptB. Of the residues involved in hydrogen bonding, Asn-745 located on the $\beta 4$ - $\alpha 4$ loop of REC_{SagS} is conserved among three of the four REC domain proteins namely, SagS, PA1611 and PA1976 that are capable of interacting with HptB. Asn-745 forms a hydrogen bond with the unique and hypervariable Lys-76 on the αD helix of HptB. Another unique residue in HptB is Asn-65, which is located on the C-terminal end of its αC helix. This is because Asn-65 from HptB is replaced by hydrophobic residues in the two other stand-alone Hpt proteins from *P. aeruginosa* (Fig. 24 and 25). However, further validation of the interacting residues identified in the REC_{SagS}-HptB model can be done in future using site-directed mutagenesis coupled with *in vitro* studies.

Altogether, it is clear that HptB is capable of interacting with multiple hybrid histidine kinase partners through a conserved hydrophobic interface. However, the specificity in the interaction between HptB with each REC containing hybrid kinases (SagS, PA1611, PA1976 and RetS) is likely achieved by a set of unique residues on HptB involved in hydrogen bonding that are not present in the other two Hpt proteins, HptA and HptC in *P. aeruginosa*. On the other hand, a relatively conserved residue involved in hydrogen bonding shared by all the four REC domains mediates specific interaction with the single HptB protein. Thus, we can suggest that hydrogen bonding has a critical role in mediating specific interactions in REC-Hpt complexes. Such interaction specificity is crucial to prevent cross talk in complex signalling networks where multiple pathways operate in concert to mediate cellular functions in organisms.

Apart from binding to HptB, we demonstrated previously (Chapter 2) that the REC domain of SagS also interacts with the free-standing PilZ protein HapZ (PA2799). We also postulated that HapZ participates in the SagS two-component signalling pathway by blocking phosphotransfer from SagS to HptB, thereby affecting biofilm formation. Although much is

Chapter 3

known about the way in which c-di-GMP binds to a PilZ protein, the precise mechanism by which a PilZ protein recognizes and binds to its target protein partner remains unclear. Understanding the mechanism that confers specificity to the interaction between a PilZ protein and its protein partner will greatly enhance our understanding of how c-di-GMP effectors function as adaptors to precisely control cellular processes.

Unfortunately, HapZ could not be purified and expressed as a stable protein despite using several solubility tags due to its propensity to form insoluble aggregates at high concentrations (Figure S3). Hence, instead of using *in vitro* studies such as NMR or Isothermal titration calorimetry (ITC), an alternative approach was used to investigate the molecular mechanism underlying the specific interaction between HapZ and REC_{SagS}. A comprehensive mutant library of HapZ and REC_{SagS} (total of 50 mutants) was generated to identify critical residues involved in this interaction. Screening these mutant libraries revealed that mutation of six residues in HapZ abolished its binding to REC_{SagS}. Four of these residues (Arg-12, Arg-16, Asp-38 and Ser-40) formed part of the c-di-GMP binding motifs RxxxR and DxSxxG. This underscored the role of c-di-GMP in facilitating HapZ-REC_{SagS} interaction. In addition to the c-di-GMP binding motif residues, two other critical residues Arg-80 and Gly-92 located in the β 6 and the β 7 sheets of the β barrel of HapZ were also identified in the screening. A HapZ mutant expressing only the N-terminal loop (Thr-24-Stop) also abolished binding with REC_{SagS} emphasizing the importance of the β barrel in this interaction. On the other hand, none of the 19 mutants in REC_{SagS} could abolish binding to HapZ in the two-hybrid system. These included mutants of two residues Gln-675 and His-683 identified in REC_{SagS}-HptB binding (Chapter 3) suggesting that HapZ and HptB may not competitively bind to REC_{SagS}.

Since HapZ was previously demonstrated to affect swarming motility and biofilm formation, phenotypic assays were carried out to validate the role of the six residues identified in HapZ in HapZ- REC_{SagS} binding. The inability of the six HapZ mutants to complement the hyperswarming phenotype of the $\Delta hapZ$ strain confirmed the significance of these residues and c-di-GMP in mediating this interaction. Additionally, colony morphology assay was also performed using the six identified HapZ mutants and five other mutants in its N-terminal loop. Biofilm formation is generally associated with a wrinkled colony phenotype (278). Overexpression of the six HapZ mutants failed to complement the smooth phenotype of the

Chapter 3

ΔhapZ strain unlike the five N-terminal loop mutants. Thus, we can propose that the six residues play a vital role in this interaction. Moreover, since increased levels of c-di-GMP are associated with biofilms, mutation of c-di-GMP binding motifs in HapZ prevents its binding to c-di-GMP, thus reducing biofilm formation and increases intracellular oxygen availability resulting in the observed smooth colony phenotype. This emphasizes that c-di-GMP binding to HapZ is crucial for HapZ- REC_{SagS} interaction.

A c-di-GMP binding model of HapZ was built based on the c-di-GMP binding motifs in HapZ. Since the residue preceding the RxxxR motif (Lys-11) in HapZ was not hydrophobic like leucine in VCA0042, we proposed that HapZ might bind to a dimeric instead of a monomeric c-di-GMP molecule similar to the binding in case of PA4608, Alg44 and PP4397 (Fig. 33) (337). The residues Arg-80 and Gly-92 on the in HapZ that were found to abolish interaction with REC_{SagS} were present on the same face as the RxxxR and DxSxxG motifs of HapZ. Of Arg-80 and Gly-92, Arg-80 is unique to HapZ in c-di-GMP binding. The involvement of such unique residues in c-di-GMP binding has been demonstrated previously in PilZ proteins such as PA4608 and Alg44 (71, 132). Although six residues were identified on HapZ, failure to detect any loss of interaction with the 19 REC_{SagS} mutants suggests a few possibilities namely (1) mutation of a single residue in REC_{SagS} is not sufficient to abolish binding with HapZ, (2) residues other than those chosen on REC_{SagS} in the present study may be involved in HapZ- REC_{SagS} interaction.

Since the residues critical for HapZ-REC_{SagS} interaction could not be identified in REC_{SagS}, we proposed a model of HapZ-REC_{SagS} complex based on the observation that the six essential binding residues in HapZ for protein-protein interaction are involved in c-di-GMP binding (Fig. 34). An important feature of this model, though remains to be fully validated, is that REC_{SagS} and HapZ interact through the dimeric c-di-GMP. Interestingly, c-di-GMP has been known to mediate homodimer interactions. One example is the binding of dimeric c-di-GMP to the transcriptional regulator VpsT from *Vibrio cholerae*. VpsT belongs to the family of response regulators involved in two-component signalling systems and consists of a receiver domain (REC) at its N-terminus and a helix-turn-helix (HTH) at its C-terminus that is involved in DNA binding. According to its crystal structure, VpsT consists of two dimerization interfaces.

Chapter 3

The first dimerization interface involves its $\alpha 1$ helix and is independent of c-di-GMP binding. On the other hand, its $\alpha 6$ helix forms the second dimerization surface, which is stabilized by the binding of self-intercalated dimeric c-di-GMP (341). Another example is the binding of the c-di-GMP tetramer to the C-terminal domain of the transcription factor BldD in *Streptomyces venezuelae* and *S. coelicolor*. The crystal structures of the BldD-c-di-GMP complex from the two *Streptomyces* species revealed that two intercalated c-di-GMP molecules (tetrameric form) interact with BldD by linking its two C-terminal protomers through the BldD motifs - Motif 1 and Motif 2. The specificity to this interaction is conferred by arginine and aspartate residues in the two motifs (342). Similarly, in case of the proposed HapZ-REC_{SagS} interaction model, a c-di-GMP dimer links the two proteins in a complex as glue. However, the HapZ-REC_{SagS} interaction might be unique since unlike VpsT and BldD, the c-di-GMP dimer interlinks different proteins instead of homodimers.

Figure 33 | Model of HapZ in complex with dimeric c-di-GMP. HapZ binds intercalated dimeric c-di-GMP using the conserved RxxxR and DxSxxG motifs and two other residues, Arg-80 that is unique to HapZ and an invariant Gly-92. The residues involved in binding c-di-GMP are shown as orange spheres. In the model, one of the c-di-GMP molecule in the dimer binds to the Arg-12, Arg-16 (RxxxR) residues in the N-terminal loop and Asp-38 (DxSxxG) residue in the $\beta 3$ sheet while the other molecule binds to Ser-40 (DxSxxG), Arg-80 and Gly-92 located on the $\beta 3$ - $\beta 4$ loop, $\beta 6$ sheet and $\beta 7$ sheet of HapZ respectively.

Thus, although we have proposed a possible model of HapZ-REC_{SagS} interaction on the basis of bacterial two-hybrid and phenotypic assays, future work would require further validation of this model. This could be done by using two alternative approaches: (1) If HapZ can be stabilized in future successfully using a solubility tag, then the construct can be used for

Chapter 3

NMR titration studies with REC_{SagS} in the presence and absence of c-di-GMP. Thus, critical interacting residues can be identified in this interaction. (2) Mutation of more than one residue on REC_{SagS} might aid in the identification of critical residues interacting with HapZ that could not be identified using single or point mutations. (3) Co-crystallization of the HapZ-REC_{SagS} complex can also be done in the presence and absence of c-di-GMP to validate the proposed model. (4) Finally, isothermal titration calorimetry (ITC) can be used to test whether REC_{SagS} has an affinity for high c-di-GMP (millimolar range) i.e. above the average cellular micromolar range.

Figure 34 | Model of REC_{SagS}-HapZ interaction. REC_{SagS} is shown in cyan, HapZ with its N-terminal unstructured loop is shown in cream while c-di-GMP is shown in yellow. The possible model of binding of HapZ with REC_{SagS} on the basis of HapZ-c-di-GMP model is shown by the cartoon representation. Based on the results, REC_{SagS} interacts with HapZ likely with dimeric c-di-GMP sandwiched between the proteins.

References

1. Lister PD, Wolter DJ, & Hanson ND (2009) Antibacterial-resistant *Pseudomonas aeruginosa*: clinical impact and complex regulation of chromosomally encoded resistance mechanisms. *Clin Microbiol Rev* 22(4):582-610.
2. Hauser AR, Meccas J, & Moir DT (2016) Beyond Antibiotics: New Therapeutic Approaches For Bacterial Infections. *Clin Infect Dis*.
3. Spencer RC (1996) Predominant pathogens found in the European Prevalence of Infection in Intensive Care Study. *Eur J Clin Microbiol Infect Dis* 15(4):281-285.
4. Silveira MC, Albano RM, Asensi MD, & Carvalho-Assef AP (2016) Description of genomic islands associated to the multidrug-resistant *Pseudomonas aeruginosa* clone ST277. *Infect Genet Evol*.
5. Williams PA & Worsey MJ (1976) Ubiquity of plasmids in coding for toluene and xylene metabolism in soil bacteria: evidence for the existence of new TOL plasmids. *J Bacteriol* 125(3):818-828.
6. Lyczak JB, Cannon CL, & Pier GB (2000) Establishment of *Pseudomonas aeruginosa* infection: lessons from a versatile opportunist. *Microbes Infect* 2(9):1051-1060.
7. Stover CK, Pham XQ, Erwin AL, Mizoguchi SD, Warren P, Hickey MJ, Brinkman FS, Hufnagle WO, Kowalik DJ, Lagrou M, Garber RL, Goltry L, Tolentino E, Westbrook-Wadman S, Yuan Y, Brody LL, Coulter SN, Folger KR, Kas A, Larbig K, Lim R, Smith K, Spencer D, Wong GK, Wu Z, Paulsen IT, Reizer J, Saier MH, Hancock RE, Lory S, & Olson MV (2000) Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an opportunistic pathogen. *Nature* 406(6799):959-964.
8. Ciofu O (2003) *Pseudomonas aeruginosa* chromosomal beta-lactamase in patients with cystic fibrosis and chronic lung infection. Mechanism of antibiotic resistance and target of the humoral immune response. *APMIS Suppl* (116):1-47.
9. Kang CI, Kim SH, Park WB, Lee KD, Kim HB, Kim EC, Oh MD, & Choe KW (2005) Risk factors for antimicrobial resistance and influence of resistance on mortality in patients with bloodstream infection caused by *Pseudomonas aeruginosa*. *Microb Drug Resist* 11(1):68-74.
10. Smith EE, Buckley DG, Wu Z, Saenphimmachak C, Hoffman LR, D'Argenio DA, Miller SI, Ramsey BW, Speert DP, Moskowitz SM, Burns JL, Kaul R, & Olson MV (2006) Genetic adaptation by *Pseudomonas aeruginosa* to the airways of cystic fibrosis patients. *Proc Natl Acad Sci U S A* 103(22):8487-8492.
11. Hoiby N, Johansen, H.K., Moser, C. and Ciofu, O. (2008) Clinical Relevance of *Pseudomonas aeruginosa*: A Master of Adaptation and Survival Strategies. *Pseudomonas: Model Organism, Pathogen, Cell Factory*, ed Rehm BHA (WILEY-VCH, Weinheim), pp 25-44.
12. Furukawa S, Kuchma SL, & O'Toole GA (2006) Keeping their options open: acute versus persistent infections. *J Bacteriol* 188(4):1211-1217.
13. Barbieri JT & Sun J (2004) *Pseudomonas aeruginosa* ExoS and ExoT. *Rev Physiol Biochem Pharmacol* 152:79-92.
14. Ghosh P (2004) Process of protein transport by the type III secretion system. *Microbiol Mol Biol Rev* 68(4):771-795.
15. Matsumoto K (2004) Role of bacterial proteases in pseudomonal and serratia keratitis. *Biol Chem* 385(11):1007-1016.

16. Vance RE, Rietsch A, & Mekalanos JJ (2005) Role of the type III secreted exoenzymes S, T, and Y in systemic spread of *Pseudomonas aeruginosa* PAO1 in vivo. *Infect Immun* 73(3):1706-1713.
17. Priebe GP, Dean CR, Zaidi T, Meluleni GJ, Coleman FT, Coutinho YS, Noto MJ, Urban TA, Pier GB, & Goldberg JB (2004) The galU Gene of *Pseudomonas aeruginosa* is required for corneal infection and efficient systemic spread following pneumonia but not for infection confined to the lung. *Infect Immun* 72(7):4224-4232.
18. Zaidi TS, Fleiszig SM, Preston MJ, Goldberg JB, & Pier GB (1996) Lipopolysaccharide outer core is a ligand for corneal cell binding and ingestion of *Pseudomonas aeruginosa*. *Invest Ophthalmol Vis Sci* 37(6):976-986.
19. Galloway DR (1991) *Pseudomonas aeruginosa* elastase and elastolysis revisited: recent developments. *Mol Microbiol* 5(10):2315-2321.
20. Zolfaghar I, Evans DJ, & Fleiszig SM (2003) Twitching motility contributes to the role of pili in corneal infection caused by *Pseudomonas aeruginosa*. *Infect Immun* 71(9):5389-5393.
21. Gallagher LA & Manoil C (2001) *Pseudomonas aeruginosa* PAO1 kills *Caenorhabditis elegans* by cyanide poisoning. *J Bacteriol* 183(21):6207-6214.
22. Rahme LG, Stevens EJ, Wolfort SF, Shao J, Tompkins RG, & Ausubel FM (1995) Common virulence factors for bacterial pathogenicity in plants and animals. *Science* 268(5219):1899-1902.
23. Rahme LG, Tan MW, Le L, Wong SM, Tompkins RG, Calderwood SB, & Ausubel FM (1997) Use of model plant hosts to identify *Pseudomonas aeruginosa* virulence factors. *Proc Natl Acad Sci U S A* 94(24):13245-13250.
24. Smith RS, Harris SG, Phipps R, & Iglewski B (2002) The *Pseudomonas aeruginosa* quorum-sensing molecule N-(3-oxododecanoyl)homoserine lactone contributes to virulence and induces inflammation in vivo. *J Bacteriol* 184(4):1132-1139.
25. Zhu H, Bandara R, Conibear TC, Thuruthyl SJ, Rice SA, Kjelleberg S, Givskov M, & Willcox MD (2004) *Pseudomonas aeruginosa* with lasI quorum-sensing deficiency during corneal infection. *Invest Ophthalmol Vis Sci* 45(6):1897-1903.
26. Smith RS & Iglewski BH (2003) *P. aeruginosa* quorum-sensing systems and virulence. *Curr Opin Microbiol* 6(1):56-60.
27. Costerton JW, Stewart PS, & Greenberg EP (1999) Bacterial biofilms: a common cause of persistent infections. *Science* 284(5418):1318-1322.
28. Darouiche RO (2004) Treatment of infections associated with surgical implants. *N Engl J Med* 350(14):1422-1429.
29. Fergie N, Bayston R, Pearson JP, & Birchall JP (2004) Is otitis media with effusion a biofilm infection? *Clin Otolaryngol Allied Sci* 29(1):38-46.
30. Singh PK, Schaefer AL, Parsek MR, Moninger TO, Welsh MJ, & Greenberg EP (2000) Quorum-sensing signals indicate that cystic fibrosis lungs are infected with bacterial biofilms. *Nature* 407(6805):762-764.
31. Chernish RN & Aaron SD (2003) Approach to resistant gram-negative bacterial pulmonary infections in patients with cystic fibrosis. *Curr Opin Pulm Med* 9(6):509-515.
32. Livermore DM (1995) beta-Lactamases in laboratory and clinical resistance. *Clin Microbiol Rev* 8(4):557-584.
33. Frederiksen B, Koch C, & Hoiby N (1999) Changing epidemiology of *Pseudomonas aeruginosa* infection in Danish cystic fibrosis patients (1974-1995). *Pediatr Pulmonol* 28(3):159-166.
34. Livermore DM (2002) Multiple mechanisms of antimicrobial resistance in *Pseudomonas aeruginosa*: our worst nightmare? *Clin Infect Dis* 34(5):634-640.

35. Drenkard E & Ausubel FM (2002) Pseudomonas biofilm formation and antibiotic resistance are linked to phenotypic variation. *Nature* 416(6882):740-743.
36. Branda SS, Vik S, Friedman L, & Kolter R (2005) Biofilms: the matrix revisited. *Trends Microbiol* 13(1):20-26.
37. Römling U, Gomelsky M, & Galperin MY (2005) C-di-GMP: the dawning of a novel bacterial signalling system. *Molecular Microbiology* 57(3):629-639.
38. Nickel JC, Ruseska I, Wright JB, & Costerton JW (1985) Tobramycin resistance of Pseudomonas aeruginosa cells growing as a biofilm on urinary catheter material. *Antimicrob Agents Chemother* 27(4):619-624.
39. Christensen BB, Haagensen JA, Heydorn A, & Molin S (2002) Metabolic commensalism and competition in a two-species microbial consortium. *Appl Environ Microbiol* 68(5):2495-2502.
40. Drenkard E (2003) Antimicrobial resistance of Pseudomonas aeruginosa biofilms. *Microbes Infect* 5(13):1213-1219.
41. van Loosdrecht MC, Norde W, & Zehnder AJ (1990) Physical chemical description of bacterial adhesion. *J Biomater Appl* 5(2):91-106.
42. O'Toole GA & Kolter R (1998) Initiation of biofilm formation in Pseudomonas fluorescens WCS365 proceeds via multiple, convergent signalling pathways: A genetic analysis. *Molecular Microbiology* 28(3):449-461.
43. Mattick JS (2002) Type IV pili and twitching motility. *Annu Rev Microbiol* 56:289-314.
44. Semmler AB, Whitchurch CB, & Mattick JS (1999) A re-examination of twitching motility in Pseudomonas aeruginosa. *Microbiology* 145 (Pt 10):2863-2873.
45. Klausen M, Heydorn A, Ragas P, Lambertsen L, Aaes-Jorgensen A, Molin S, & Tolker-Nielsen T (2003) Biofilm formation by Pseudomonas aeruginosa wild type, flagella and type IV pili mutants. *Mol Microbiol* 48(6):1511-1524.
46. Skerker JM & Berg HC (2001) Direct observation of extension and retraction of type IV pili. *Proc Natl Acad Sci U S A* 98(12):6901-6904.
47. Davies DG, Parsek MR, Pearson JP, Igleski BH, Costerton JW, & Greenberg EP (1998) The involvement of cell-to-cell signals in the development of a bacterial biofilm. *Science* 280(5361):295-298.
48. Allesen-Holm M, Barken KB, Yang L, Klausen M, Webb JS, Kjelleberg S, Molin S, Givskov M, & Tolker-Nielsen T (2006) A characterization of DNA release in Pseudomonas aeruginosa cultures and biofilms. *Mol Microbiol* 59(4):1114-1128.
49. Haagensen JA, Klausen M, Ernst RK, Miller SI, Folkesson A, Tolker-Nielsen T, & Molin S (2007) Differentiation and distribution of colistin- and sodium dodecyl sulfate-tolerant cells in Pseudomonas aeruginosa biofilms. *J Bacteriol* 189(1):28-37.
50. Lequette Y & Greenberg EP (2005) Timing and localization of rhamnolipid synthesis gene expression in Pseudomonas aeruginosa biofilms. *J Bacteriol* 187(1):37-44.
51. Applegate DH & Bryers JD (1991) Effects of carbon and oxygen limitations and calcium concentrations on biofilm removal processes. *Biotechnol Bioeng* 37(1):17-25.
52. Delaquis PJ, Caldwell DE, Lawrence JR, & McCurdy AR (1989) Detachment of Pseudomonas fluorescens from biofilms on glass surfaces in response to nutrient stress. *Microb Ecol* 18(3):199-210.
53. Gjermansen M, Ragas P, Sternberg C, Molin S, & Tolker-Nielsen T (2005) Characterization of starvation-induced dispersion in Pseudomonas putida biofilms. *Environ Microbiol* 7(6):894-906.
54. Hunt SM, Werner EM, Huang B, Hamilton MA, & Stewart PS (2004) Hypothesis for the role of nutrient starvation in biofilm detachment. *Appl Environ Microbiol* 70(12):7418-7425.

55. Sauer K, Cullen MC, Rickard AH, Zeef LA, Davies DG, & Gilbert P (2004) Characterization of nutrient-induced dispersion in *Pseudomonas aeruginosa* PAO1 biofilm. *J Bacteriol* 186(21):7312-7326.
56. Banin E, Brady KM, & Greenberg EP (2006) Chelator-induced dispersal and killing of *Pseudomonas aeruginosa* cells in a biofilm. *Appl Environ Microbiol* 72(3):2064-2069.
57. Simm R, Morr M, Kader A, Nimtz M, & Romling U (2004) GGDEF and EAL domains inversely regulate cyclic di-GMP levels and transition from sessility to motility. *Mol Microbiol* 53(4):1123-1134.
58. Allison DG, Ruiz B, SanJose C, Jaspe A, & Gilbert P (1998) Extracellular products as mediators of the formation and detachment of *Pseudomonas fluorescens* biofilms. *FEMS Microbiol Lett* 167(2):179-184.
59. Webb JS, Thompson LS, James S, Charlton T, Tolker-Nielsen T, Koch B, Givskov M, & Kjelleberg S (2003) Cell death in *Pseudomonas aeruginosa* biofilm development. *J Bacteriol* 185(15):4585-4592.
60. Barraud N, Hassett DJ, Hwang SH, Rice SA, Kjelleberg S, & Webb JS (2006) Involvement of nitric oxide in biofilm dispersal of *Pseudomonas aeruginosa*. *J Bacteriol* 188(21):7344-7353.
61. Davey ME, Caiazza NC, & O'Toole GA (2003) Rhamnolipid surfactant production affects biofilm architecture in *Pseudomonas aeruginosa* PAO1. *J Bacteriol* 185(3):1027-1036.
62. Davies DG & Marques CN (2009) A fatty acid messenger is responsible for inducing dispersion in microbial biofilms. *J Bacteriol* 191(5):1393-1403.
63. Valentini M & Filloux A (2016) Biofilms and c-di-GMP Signaling: Lessons from *Pseudomonas aeruginosa* and other Bacteria. *J Biol Chem*.
64. Ross P, Weinhouse H, Aloni Y, Michaeli D, Weinberger-Ohana P, Mayer R, Braun S, de Vroom E, van der Marel GA, van Boom JH, & Benziman M (1987) Regulation of cellulose synthesis in *Acetobacter xylinum* by cyclic diguanylic acid. *Nature* 325(6101):279-281.
65. Ross P, Aloni Y, Weinhouse H, Michaeli D, Weinberger-Ohana P, Mayer R, & Benziman M (1986) Control of cellulose synthesis *Acetobacter xylinum*. A unique guanyl oligonucleotide is the immediate activator of the cellulose synthase. *Carbohydrate Research* 149(1):101-117.
66. Ross P, Aloni Y, Weinhouse C, Michaeli D, Weinberger-Ohana P, Meyer R, & Benziman M (1985) An unusual guanyl oligonucleotide regulates cellulose synthesis in *Acetobacter xylinum*. *FEBS Lett* 186(2):191-196.
67. Ross P, Mayer R, Weinhouse H, Amikam D, Huggirat Y, Benziman M, de Vroom E, Fiddler A, de Paus P, Slidregt LA, & et al. (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(31):18933-18943.
68. Amikam D & Benziman M (1989) Cyclic diguanylic acid and cellulose synthesis in *Agrobacterium tumefaciens*. *J Bacteriol* 171(12):6649-6655.
69. Gentner M, Allan MG, Zaehring F, Schirmer T, & Grzesiek S (2012) Oligomer formation of the bacterial second messenger c-di-GMP: reaction rates and equilibrium constants indicate a monomeric state at physiological concentrations. *J Am Chem Soc* 134(2):1019-1029.
70. Römmling U, Galperin MY, & Gomelsky M (2013) Cyclic di-GMP: The first 25 years of a universal bacterial second messenger. *Microbiology and Molecular Biology Reviews* 77(1):1-52.

71. Whitney JC, Whitfield GB, Marmont LS, Yip P, Neculai AM, Lobsanov YD, Robinson H, Ohman DE, & Howell PL (2015) Dimeric c-di-GMP is required for post-translational regulation of alginate production in *Pseudomonas aeruginosa*. *J Biol Chem* 290(20):12451-12462.
72. Tal R, Wong HC, Calhoon R, Gelfand D, Fear AL, Volman G, Mayer R, Ross P, Amikam D, Weinhouse H, Cohen A, Sapir S, Ohana P, & Benziman M (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(17):4416-4425.
73. Ausmees N, Mayer R, Weinhouse H, Volman G, Amikam D, Benziman M, & Lindberg M (2001) Genetic data indicate that proteins containing the GGDEF domain possess diguanylate cyclase activity. *FEMS Microbiol Lett* 204(1):163-167.
74. Tischler AD & Camilli A (2004) Cyclic diguanylate (c-di-GMP) regulates *Vibrio cholerae* biofilm formation. *Mol Microbiol* 53(3):857-869.
75. Kulasakara H, Lee V, Brencic A, Liberati N, Urbach J, Miyata S, Lee DG, Neely AN, Hyodo M, Hayakawa Y, Ausubel FM, & Lory S (2006) Analysis of *Pseudomonas aeruginosa* diguanylate cyclases and phosphodiesterases reveals a role for bis-(3'-5')-cyclic-GMP in virulence. *Proc Natl Acad Sci U S A* 103(8):2839-2844.
76. Hecht GB & Newton A (1995) Identification of a novel response regulator required for the swarmer-to-stalked-cell transition in *Caulobacter crescentus*. *J Bacteriol* 177(21):6223-6229.
77. Paul R, Abel S, Wassmann P, Beck A, Heerklotz H, & Jenal U (2007) Activation of the diguanylate cyclase PleD by phosphorylation-mediated dimerization. *J Biol Chem* 282(40):29170-29177.
78. Klebensberger J, Birkenmaier A, Geffers R, Kjelleberg S, & Philipp B (2009) SiaA and SiaD are essential for inducing autoaggregation as a specific response to detergent stress in *Pseudomonas aeruginosa*. *Environ Microbiol* 11(12):3073-3086.
79. Merritt JH, Ha DG, Cowles KN, Lu W, Morales DK, Rabinowitz J, Gitai Z, & O'Toole GA (2010) Specific control of *Pseudomonas aeruginosa* surface-associated behaviors by two c-di-GMP diguanylate cyclases. *mBio* 1(4).
80. Sanchez-Torres V, Hu H, & Wood TK (2011) GGDEF proteins YeaI, YedQ, and YfiN reduce early biofilm formation and swimming motility in *Escherichia coli*. *Appl Microbiol Biotechnol* 90(2):651-658.
81. Ueda A & Wood TK (2009) Connecting quorum sensing, c-di-GMP, *pel* polysaccharide, and biofilm formation in *Pseudomonas aeruginosa* through tyrosine phosphatase TpbA (PA3885). *PLoS Pathogens* 5(6).
82. Giardina G, Paiardini A, Fernicola S, Franceschini S, Rinaldo S, Stelitano V, & Cutruzzola F (2013) Investigating the allosteric regulation of YfiN from *Pseudomonas aeruginosa*: clues from the structure of the catalytic domain. *PLoS ONE* 8(11):e81324.
83. D'Argenio DA, Calfee MW, Rainey PB, & Pesci EC (2002) Autolysis and autoaggregation in *Pseudomonas aeruginosa* colony morphology mutants. *J Bacteriol* 184(23):6481-6489.
84. Rainey PB & Travisano M (1998) Adaptive radiation in a heterogeneous environment. *Nature* 394(6688):69-72.
85. Malone JG, Williams R, Christen M, Jenal U, Spiers AJ, & Rainey PB (2007) The structure-function relationship of WspR, a *Pseudomonas fluorescens* response regulator with a GGDEF output domain. *Microbiology* 153(4):980-994.
86. Hickman JW, Tifrea DF, & Harwood CS (2005) A chemosensory system that regulates biofilm formation through modulation of cyclic diguanylate levels.

Proceedings of the National Academy of Sciences of the United States of America 102(40):14422-14427.

87. Guvener ZT & Harwood CS (2007) Subcellular location characteristics of the *Pseudomonas aeruginosa* GGDEF protein, WspR, indicate that it produces cyclic-di-GMP in response to growth on surfaces. *Mol Microbiol* 66(6):1459-1473.
88. De N, Pirruccello M, Krasteva PV, Bae N, Raghavan RV, & Sondermann H (2008) Phosphorylation-independent regulation of the diguanylate cyclase WspR. *PLoS biology* 6(3):e67.
89. Kuchma SL, Brothers KM, Merritt JH, Liberati NT, Ausubel FM, & O'Toole GA (2007) BifA, a cyclic-Di-GMP phosphodiesterase, inversely regulates biofilm formation and swarming motility by *Pseudomonas aeruginosa* PA14. *J Bacteriol* 189(22):8165-8178.
90. Merritt JH, Brothers KM, Kuchma SL, & O'Toole GA (2007) SadC reciprocally influences biofilm formation and swarming motility via modulation of exopolysaccharide production and flagellar function. *J Bacteriol* 189(22):8154-8164.
91. Moscoso JA, Jaeger T, Valentini M, Hui K, Jenal U, & Filloux A (2014) The diguanylate cyclase SadC is a central player in Gac/Rsm-mediated biofilm formation in *Pseudomonas aeruginosa*. *J Bacteriol* 196(23):4081-4088.
92. Zhu B, Liu C, Liu S, Cong H, Chen Y, Gu L, & Ma LZ (2016) Membrane association of SadC enhances its diguanylate cyclase activity to control exopolysaccharides synthesis and biofilm formation in *Pseudomonas aeruginosa*. *Environ Microbiol*.
93. Petrova OE, Cherny KE, & Sauer K (2014) The *Pseudomonas aeruginosa* diguanylate cyclase GcbA, a homolog of *P. fluorescens* GcbA, promotes initial attachment to surfaces, but not biofilm formation, via regulation of motility. *J Bacteriol* 196(15):2827-2841.
94. Petrova OE, Cherny KE, & Sauer K (2015) The diguanylate cyclase GcbA facilitates *Pseudomonas aeruginosa* biofilm dispersion by activating BdlA. *J Bacteriol* 197(1):174-187.
95. Basu Roy A & Sauer K (2014) Diguanylate cyclase NicD-based signalling mechanism of nutrient-induced dispersion by *Pseudomonas aeruginosa*. *Mol Microbiol* 94(4):771-793.
96. An S, Wu J, & Zhang LH (2010) Modulation of *Pseudomonas aeruginosa* biofilm dispersal by a cyclic-Di-GMP phosphodiesterase with a putative hypoxia-sensing domain. *Appl Environ Microbiol* 76(24):8160-8173.
97. Roy AB, Petrova OE, & Sauer K (2012) The phosphodiesterase DipA (PA5017) is essential for *Pseudomonas aeruginosa* biofilm dispersion. *J Bacteriol* 194(11):2904-2915.
98. Galperin MY, Gaidenko TA, Mulkidjanian AY, Nakano M, & Price CW (2001) MHYT, a new integral membrane sensor domain. *FEMS Microbiol Lett* 205(1):17-23.
99. Hay ID, Remminghorst U, & Rehm BH (2009) MucR, a novel membrane-associated regulator of alginate biosynthesis in *Pseudomonas aeruginosa*. *Appl Environ Microbiol* 75(4):1110-1120.
100. Li Y, Heine S, Entian M, Sauer K, & Frankenberg-Dinkel N (2013) NO-induced biofilm dispersion in *Pseudomonas aeruginosa* is mediated by an MHYT domain-coupled phosphodiesterase. *J Bacteriol* 195(16):3531-3542.
101. Choy WK, Zhou L, Syn CK, Zhang LH, & Swarup S (2004) MorA defines a new class of regulators affecting flagellar development and biofilm formation in diverse *Pseudomonas* species. *J Bacteriol* 186(21):7221-7228.

102. Phippen CW, Mikolajek H, Schlaefli HG, Keevil CW, Webb JS, & Tews I (2014) Formation and dimerization of the phosphodiesterase active site of the *Pseudomonas aeruginosa* MorA, a bi-functional c-di-GMP regulator. *FEBS Lett* 588(24):4631-4636.
103. Kazmierczak BI, Lebron MB, & Murray TS (2006) Analysis of FimX, a phosphodiesterase that governs twitching motility in *Pseudomonas aeruginosa*. *Mol Microbiol* 60(4):1026-1043.
104. Newell PD, Monds RD, & O'Toole GA (2009) LapD is a bis-(3',5')-cyclic dimeric GMP-binding protein that regulates surface attachment by *Pseudomonas fluorescens* Pf0-1. *Proceedings of the National Academy of Sciences* 106(9):3461-3466.
105. Cooley RB, Smith TJ, Leung W, Tierney V, Borlee BR, O'Toole GA, & Sondermann H (2016) Cyclic Di-GMP-Regulated Periplasmic Proteolysis of a *Pseudomonas aeruginosa* Type Vb Secretion System Substrate. *J Bacteriol* 198(1):66-76.
106. Galperin MY, Nikolskaya AN, & Koonin EV (2001) Novel domains of the prokaryotic two-component signal transduction systems. *FEMS Microbiol Lett* 203(1):11-21.
107. Galperin MY (2004) Bacterial signal transduction network in a genomic perspective. *Environ Microbiol* 6(6):552-567.
108. Salter EA & Wierzbicki A (2007) The mechanism of cyclic nucleotide hydrolysis in the phosphodiesterase catalytic site. *J Phys Chem B* 111(17):4547-4552.
109. Barends T, Hartmann E, Griese J, Beitlich T, Kirienko N, Ryjenkov D, Reinstein J, Shoeman R, Gomelsky M, & Schlichting I (2009) Structure and mechanism of a bacterial light-regulated cyclic nucleotide phosphodiesterase. *Nature* 459(7249):1015-1018.
110. Hoffman LR, D'Argenio DA, MacCoss MJ, Zhang Z, Jones RA, & Miller SI (2005) Aminoglycoside antibiotics induce bacterial biofilm formation. *Nature* 436(7054):1171-1175.
111. Kulasekara HD, Ventre I, Kulasekara BR, Lazdunski A, Filloux A, & Lory S (2005) A novel two-component system controls the expression of *Pseudomonas aeruginosa* fimbrial cup genes. *Mol Microbiol* 55(2):368-380.
112. Rao F, Yang Y, Qi Y, & Liang ZX (2008) Catalytic mechanism of cyclic di-GMP-specific phosphodiesterase: A study of the EAL domain-containing RocR from *Pseudomonas aeruginosa*. *Journal of Bacteriology* 190(10):3622-3631.
113. Chen MW, Kotaka M, Vonrhein C, Bricogne G, Rao F, Chuah ML, Svergun D, Schneider G, Liang ZX, & Lescar J (2012) Structural insights into the regulatory mechanism of the response regulator RocR from *Pseudomonas aeruginosa* in cyclic Di-GMP signaling. *J Bacteriol* 194(18):4837-4846.
114. Rao F, Qi Y, Chong HS, Kotaka M, Li B, Li J, Lescar J, Tang K, & Liang ZX (2009) The functional role of a conserved loop in EAL domain-based cyclic di-GMP-specific phosphodiesterase. *J Bacteriol* 191(15):4722-4731.
115. Ryan RP, Fouhy Y, Lucey JF, Crossman LC, Spiro S, He Y-W, Zhang L-H, Heeb S, Cámara M, Williams P, & Dow JM (2006) Cell-cell signaling in *Xanthomonas campestris* involves an HD-GYP domain protein that functions in cyclic di-GMP turnover. *Proceedings of the National Academy of Sciences* 103(17):6712-6717.
116. Stelitano V, Giardina G, Paiardini A, Castiglione N, Cutruzzola F, & Rinaldo S (2013) C-di-GMP hydrolysis by *Pseudomonas aeruginosa* HD-GYP phosphodiesterases: analysis of the reaction mechanism and novel roles for pGpG. *PLoS ONE* 8(9):e74920.
117. Rinaldo S, Paiardini A, Stelitano V, Brunotti P, Cervoni L, Fernicola S, Protano C, Vitali M, Cutruzzola F, & Giardina G (2015) Structural basis of functional diversification of the HD-GYP domain revealed by the *Pseudomonas aeruginosa*

- PA4781 protein, which displays an unselective bimetallic binding site. *J Bacteriol* 197(8):1525-1535.
118. Ryan RP, Lucey J, O'Donovan K, McCarthy Y, Yang L, Tolker-Nielsen T, Dow JM, Ryan RP, Lucey J, O'Donovan K, McCarthy Y, Yang L, Tolker-Nielsen T, & Dow JM (2009) HD-GYP domain proteins regulate biofilm formation and virulence in *Pseudomonas aeruginosa*. *Environmental Microbiology* 11(5):1126-1136.
 119. Krasteva PV, Giglio KM, & Sondermann H (2012) Sensing the messenger: The diverse ways that bacteria signal through c-di-GMP. *Protein Science* 21(7):929-948.
 120. Ryan RP, Tolker-Nielsen T, & Dow JM (2012) When the PilZ don't work: Effectors for cyclic di-GMP action in bacteria. *Trends in Microbiology* 20(5):235-242.
 121. Amikam D & Galperin MY (2006) PilZ domain is part of the bacterial c-di-GMP binding protein. *Bioinformatics* 22(1):3-6.
 122. Alm RA, Boder AJ, Free PD, & Mattick JS (1996) Identification of a novel gene, pilZ, essential for type 4 fimbrial biogenesis in *Pseudomonas aeruginosa*. *Journal of Bacteriology* 178(1):46-53.
 123. Ryjenkov DA, Simm R, Römling U, & Gomelsky M (2006) The PilZ domain is a receptor for the second messenger c-di-GMP: The PilZ domain protein YcgR controls motility in enterobacteria. *Journal of Biological Chemistry* 281(41):30310-30314.
 124. Christen B, Christen M, Paul R, Schmid F, Folcher M, Jenoe P, Meuwly M, & Jenal U (2006) Allosteric control of cyclic di-GMP signaling. *Journal of Biological Chemistry* 281(42):32015-32024.
 125. Pratt JT, Tamayo R, Tischler AD, & Camilli A (2007) PilZ domain proteins bind cyclic diguanylate and regulate diverse processes in *Vibrio cholerae*. *J Biol Chem* 282(17):12860-12870.
 126. Weinhouse H, Sapir S, Amikam D, Shilo Y, Volman G, Ohana P, & Benziman M (1997) c-di-GMP-binding protein, a new factor regulating cellulose synthesis in *Acetobacter xylinum*. *FEBS Lett* 416(2):207-211.
 127. Paul R, Weiser S, Amiot NC, Chan C, Schirmer T, Giese B, & Jenal U (2004) Cell cycle-dependent dynamic localization of a bacterial response regulator with a novel di-guanylate cyclase output domain. *Genes Dev* 18(6):715-727.
 128. Schmidt AJ, Ryjenkov DA, & Gomelsky M (2005) The ubiquitous protein domain EAL is a cyclic diguanylate-specific phosphodiesterase: enzymatically active and inactive EAL domains. *J Bacteriol* 187(14):4774-4781.
 129. Benach J, Swaminathan SS, Tamayo R, Handelsman SK, Folta-Stogniew E, Ramos JE, Forouhar F, Neely H, Seetharaman J, Camilli A, & Hunt JF (2007) The structural basis of cyclic diguanylate signal transduction by PilZ domains. *EMBO Journal* 26(24):5153-5166.
 130. Ramelot TA, Yee A, Cort JR, Semesi A, Arrowsmith CH, & Kennedy MA (2007) NMR structure and binding studies confirm that PA4608 from *Pseudomonas aeruginosa* is a PilZ domain and a c-di-GMP binding protein. *Proteins: Structure, Function and Genetics* 66(2):266-271.
 131. Ko J, Ryu KS, Kim H, Shin JS, Lee JO, Cheong C, & Choi BS (2010) Structure of PP4397 reveals the molecular basis for different c-di-GMP binding modes by Pilz domain proteins. *J Mol Biol* 398(1):97-110.
 132. Habazettl J, Allan MG, Jenal U, & Grzesiek S (2011) Solution structure of the PilZ domain protein PA4608 complex with cyclic di-GMP identifies charge clustering as molecular readout. *Journal of Biological Chemistry* 286(16):14304-14314.
 133. Shin JS, Ryu KS, Ko J, Lee A, & Choi BS (2011) Structural characterization reveals that a PilZ domain protein undergoes substantial conformational change upon binding to cyclic dimeric guanosine monophosphate. *Protein Science* 20(2):270-277.

134. Remminghorst U & Rehm BH (2006) Alg44, a unique protein required for alginate biosynthesis in *Pseudomonas aeruginosa*. *FEBS Lett* 580(16):3883-3888.
135. Merighi M, Lee VT, Hyodo M, Hayakawa Y, & Lory S (2007) The second messenger bis-(3'-5')-cyclic-GMP and its PilZ domain-containing receptor Alg44 are required for alginate biosynthesis in *Pseudomonas aeruginosa*. *Molecular Microbiology* 65(4):876-895.
136. Baker AE, Diepold A, Kuchma SL, Scott JE, Ha DG, Orazi G, Armitage JP, & O'Toole GA (2016) A PilZ domain protein FlgZ mediates c-di-GMP-dependent swarming motility control in *Pseudomonas aeruginosa*. *J Bacteriol.*
137. Lee VT, Matewish JM, Kessler JL, Hyodo M, Hayakawa Y, & Lory S (2007) A cyclic-di-GMP receptor required for bacterial exopolysaccharide production. *Molecular Microbiology* 65(6):1474-1484.
138. Whitney JC, Colvin KM, Marmont LS, Robinson H, Parsek MR, & Howell PL (2012) Structure of the Cytoplasmic Region of PelD, a Degenerate Diguanylate Cyclase Receptor That Regulates Exopolysaccharide Production in *Pseudomonas aeruginosa*. *Journal of Biological Chemistry* 287(28):23582-23593.
139. Li Z, Chen JH, Hao Y, & Nair SK (2012) Structures of the PelD cyclic diguanylate effector involved in pellicle formation in *Pseudomonas aeruginosa* PAO1. *J Biol Chem* 287(36):30191-30204.
140. Marmont LS, Whitney JC, Robinson H, Colvin KM, Parsek MR, & Howell PL (2012) Expression, purification, crystallization and preliminary X-ray analysis of *Pseudomonas aeruginosa* PelD. *Acta Crystallogr Sect F Struct Biol Cryst Commun* 68(Pt 2):181-184.
141. Gao R & Stock AM (2009) Catalytically incompetent by design. *Structure* 17(8):1038-1040.
142. Qi Y, Chuah MLC, Dong X, Xie K, Luo Z, Tang K, & Liang Z-X (2011) Binding of Cyclic Diguanylate in the Non-catalytic EAL Domain of FimX Induces a Long-range Conformational Change. *Journal of Biological Chemistry* 286(4):2910-2917.
143. Navarro MV, De N, Bae N, Wang Q, & Sondermann H (2009) Structural analysis of the GGDEF-EAL domain-containing c-di-GMP receptor FimX. *Structure* 17(8):1104-1116.
144. Robert-Paganin J, Nonin-Lecomte S, & Rety S (2012) Crystal structure of an EAL domain in complex with reaction product 5'-pGpG. *PLoS ONE* 7(12):e52424.
145. Barrick JE & Breaker RR (2007) The distributions, mechanisms, and structures of metabolite-binding riboswitches. *Genome Biol* 8(11):R239.
146. Sudarsan N, Lee ER, Weinberg Z, Moy RH, Kim JN, Link KH, & Breaker RR (2008) Riboswitches in Eubacteria Sense the Second Messenger Cyclic Di-GMP. *Science* 321(5887):411-413.
147. Lee ER, Baker JL, Weinberg Z, Sudarsan N, & Breaker RR (2010) An allosteric self-splicing ribozyme triggered by a bacterial second messenger. *Science* 329(5993):845-848.
148. Hickman JW & Harwood CS (2008) Identification of FleQ from *Pseudomonas aeruginosa* as a c-di-GMP-responsive transcription factor. *Molecular Microbiology* 69(2):376-389.
149. Chambers JR & Sauer K (2013) The MerR-like regulator BrlR impairs *Pseudomonas aeruginosa* biofilm tolerance to colistin by repressing PhoPQ. *J Bacteriol* 195(20):4678-4688.
150. Liao J & Sauer K (2012) The MerR-like transcriptional regulator BrlR contributes to *Pseudomonas aeruginosa* biofilm tolerance. *J Bacteriol* 194(18):4823-4836.

151. Liao J, Schurr MJ, & Sauer K (2013) The MerR-like regulator BrlR confers biofilm tolerance by activating multidrug efflux pumps in *Pseudomonas aeruginosa* biofilms. *J Bacteriol* 195(15):3352-3363.
152. Chambers JR, Liao J, Schurr MJ, & Sauer K (2014) BrlR from *Pseudomonas aeruginosa* is a c-di-GMP-responsive transcription factor. *Mol Microbiol* 92(3):471-487.
153. Mattick JS, Whitchurch CB, & Alm RA (1996) The molecular genetics of type-4 fimbriae in *Pseudomonas aeruginosa*--a review. *Gene* 179(1):147-155.
154. Huang B, Whitchurch CB, & Mattick JS (2003) FimX, a multidomain protein connecting environmental signals to twitching motility in *Pseudomonas aeruginosa*. *J Bacteriol* 185(24):7068-7076.
155. Jones CJ, Ryder CR, Mann EE, & Wozniak DJ (2013) AmrZ modulates *Pseudomonas aeruginosa* biofilm architecture by directly repressing transcription of the *psl* operon. *J Bacteriol* 195(8):1637-1644.
156. Huangyutitham V, Güvener ZT, & Harwood CS (2013) Subcellular Clustering of the Phosphorylated WspR Response Regulator Protein Stimulates Its Diguanylate Cyclase Activity. *mBio* 4(3).
157. De N, Navarro MV, Raghavan RV, & Sondermann H (2009) Determinants for the activation and autoinhibition of the diguanylate cyclase response regulator WspR. *J Mol Biol* 393(3):619-633.
158. Thormann KM, Duttler S, Saville RM, Hyodo M, Shukla S, Hayakawa Y, & Spormann AM (2006) Control of formation and cellular detachment from *Shewanella oneidensis* MR-1 biofilms by cyclic di-GMP. *J Bacteriol* 188(7):2681-2691.
159. Morgan R, Kohn S, Hwang SH, Hassett DJ, & Sauer K (2006) BdlA, a chemotaxis regulator essential for biofilm dispersion in *Pseudomonas aeruginosa*. *J Bacteriol* 188(21):7335-7343.
160. Li Y, Petrova OE, Su S, Lau GW, Panmanee W, Na R, Hassett DJ, Davies DG, & Sauer K (2014) BdlA, DipA and induced dispersion contribute to acute virulence and chronic persistence of *Pseudomonas aeruginosa*. *PLoS Pathog* 10(6):e1004168.
161. Moscoso JA, Mikkelsen H, Heeb S, Williams P, & Filloux A (2011) The *Pseudomonas aeruginosa* sensor RetS switches type III and type VI secretion via c-di-GMP signalling. *Environ Microbiol* 13(12):3128-3138.
162. Ravichandran A, Ramachandran M, Suriyanarayanan T, Wong CC, & Swarup S (2015) Global Regulator MorA Affects Virulence-Associated Protease Secretion in *Pseudomonas aeruginosa* PAO1. *PLoS ONE* 10(4):e0123805.
163. Malone JG, Jaeger T, Manfredi P, Dotsch A, Blanka A, Bos R, Cornelis GR, Haussler S, & Jenal U (2012) The YfiBNR signal transduction mechanism reveals novel targets for the evolution of persistent *Pseudomonas aeruginosa* in cystic fibrosis airways. *PLoS Pathog* 8(6):e1002760.
164. Malone JG, Jaeger T, Spangler C, Ritz D, Spang A, Arrieumerlou C, Kaeffer V, Landmann R, & Jenal U (2010) YfiBNR mediates cyclic di-GMP dependent small colony variant formation and persistence in *Pseudomonas aeruginosa*. *PLoS Pathog* 6(3):e1000804.
165. Boyd A & Chakrabarty AM (1995) *Pseudomonas aeruginosa* biofilms: role of the alginate exopolysaccharide. *J Ind Microbiol* 15(3):162-168.
166. Friedman L & Kolter R (2004) Two genetic loci produce distinct carbohydrate-rich structural components of the *Pseudomonas aeruginosa* biofilm matrix. *J Bacteriol* 186(14):4457-4465.
167. Friedman L & Kolter R (2004) Genes involved in matrix formation in *Pseudomonas aeruginosa* PA14 biofilms. *Mol Microbiol* 51(3):675-690.

168. Meluleni GJ, Grout M, Evans DJ, & Pier GB (1995) Mucoid *Pseudomonas aeruginosa* growing in a biofilm in vitro are killed by opsonic antibodies to the mucoid exopolysaccharide capsule but not by antibodies produced during chronic lung infection in cystic fibrosis patients. *J Immunol* 155(4):2029-2038.
169. Jackson KD, Starkey M, Kremer S, Parsek MR, & Wozniak DJ (2004) Identification of *psl*, a locus encoding a potential exopolysaccharide that is essential for *Pseudomonas aeruginosa* PAO1 biofilm formation. *J Bacteriol* 186(14):4466-4475.
170. Overhage J, Schemionek M, Webb JS, & Rehm BH (2005) Expression of the *psl* operon in *Pseudomonas aeruginosa* PAO1 biofilms: *PslA* performs an essential function in biofilm formation. *Appl Environ Microbiol* 71(8):4407-4413.
171. Vasseur P, Vallet-Gely I, Soscia C, Genin S, & Filloux A (2005) The *pel* genes of the *Pseudomonas aeruginosa* PAK strain are involved at early and late stages of biofilm formation. *Microbiology* 151(Pt 3):985-997.
172. Matsukawa M & Greenberg EP (2004) Putative exopolysaccharide synthesis genes influence *Pseudomonas aeruginosa* biofilm development. *J Bacteriol* 186(14):4449-4456.
173. Caiazza NC, Merritt JH, Brothers KM, & O'Toole GA (2007) Inverse regulation of biofilm formation and swarming motility by *Pseudomonas aeruginosa* PA14. *Journal of Bacteriology* 189(9):3603-3612.
174. Lee VT, Matewish JM, Kessler JL, Hyodo M, Hayakawa Y, & Lory S (2007) A cyclic-di-GMP receptor required for bacterial exopolysaccharide production. *Mol Microbiol* 65(6):1474-1484.
175. Jennings LK, Storek KM, Ledvina HE, Coulon C, Marmont LS, Sadovskaya I, Secor PR, Tseng BS, Scian M, Filloux A, Wozniak DJ, Howell PL, & Parsek MR (2015) *Pel* is a cationic exopolysaccharide that cross-links extracellular DNA in the *Pseudomonas aeruginosa* biofilm matrix. *Proc Natl Acad Sci U S A* 112(36):11353-11358.
176. Baraquet C, Murakami K, Parsek MR, & Harwood CS (2012) The *FleQ* protein from *Pseudomonas aeruginosa* functions as both a repressor and an activator to control gene expression from the *pel* operon promoter in response to c-di-GMP. *Nucleic Acids Res* 40(15):7207-7218.
177. Ma L, Lu H, Sprinkle A, Parsek MR, & Wozniak DJ (2007) *Pseudomonas aeruginosa* *Psl* is a galactose- and mannose-rich exopolysaccharide. *J Bacteriol* 189(22):8353-8356.
178. Baker P, Whitfield GB, Hill PJ, Little DJ, Pestrak MJ, Robinson H, Wozniak DJ, & Howell PL (2015) Characterization of the *Pseudomonas aeruginosa* Glycoside Hydrolase *PslG* Reveals That Its Levels Are Critical for *Psl* Polysaccharide Biosynthesis and Biofilm Formation. *J Biol Chem* 290(47):28374-28387.
179. Wang S, Liu X, Liu H, Zhang L, Guo Y, Yu S, Wozniak DJ, & Ma LZ (2015) The exopolysaccharide *Psl*-eDNA interaction enables the formation of a biofilm skeleton in *Pseudomonas aeruginosa*. *Environ Microbiol Rep* 7(2):330-340.
180. Irie Y, Borlee BR, O'Connor JR, Hill PJ, Harwood CS, Wozniak DJ, & Parsek MR (2012) Self-produced exopolysaccharide is a signal that stimulates biofilm formation in *Pseudomonas aeruginosa*. *Proc Natl Acad Sci U S A* 109(50):20632-20636.
181. Evans LR & Linker A (1973) Production and characterization of the slime polysaccharide of *Pseudomonas aeruginosa*. *J Bacteriol* 116(2):915-924.
182. Hentzer M, Teitzel GM, Balzer GJ, Heydorn A, Molin S, Givskov M, & Parsek MR (2001) Alginate overproduction affects *Pseudomonas aeruginosa* biofilm structure and function. *J Bacteriol* 183(18):5395-5401.

183. Bjarnsholt T, Jensen PO, Fiandaca MJ, Pedersen J, Hansen CR, Andersen CB, Pressler T, Givskov M, & Hoiby N (2009) *Pseudomonas aeruginosa* biofilms in the respiratory tract of cystic fibrosis patients. *Pediatr Pulmonol* 44(6):547-558.
184. Ciofu O, Lee B, Johannesson M, Hermansen NO, Meyer P, Hoiby N, & Scandinavian Cystic Fibrosis Study C (2008) Investigation of the *algT* operon sequence in mucoid and non-mucoid *Pseudomonas aeruginosa* isolates from 115 Scandinavian patients with cystic fibrosis and in 88 in vitro non-mucoid revertants. *Microbiology* 154(Pt 1):103-113.
185. Schmidt A, Hammerbacher AS, Bastian M, Nieken KJ, Klockgether J, Merighi M, Lapouge K, Poschgan C, Kolle J, Acharya KR, Ulrich M, Tummler B, Unden G, Kaever V, Lory S, Haas D, Schwarz S, & Doring G (2016) Oxygen-dependent regulation of c-di-GMP synthesis by SadC controls alginate production in *Pseudomonas aeruginosa*. *Environ Microbiol*.
186. Oglesby LL, Jain S, & Ohman DE (2008) Membrane topology and roles of *Pseudomonas aeruginosa* Alg8 and Alg44 in alginate polymerization. *Microbiology* 154(Pt 6):1605-1615.
187. Fata Moradali M, Donati I, Sims IM, Ghods S, & Rehm BH (2015) Alginate Polymerization and Modification Are Linked in *Pseudomonas aeruginosa*. *mBio* 6(3):e00453-00415.
188. Hoch JA (2000) Two-component and phosphorelay signal transduction. *Curr Opin Microbiol* 3(2):165-170.
189. Stock AM, Robinson VL, & Goudreau PN (2000) Two-component signal transduction. *Annu Rev Biochem* 69:183-215.
190. Mascher T, Helmann JD, & Unden G (2006) Stimulus perception in bacterial signal-transducing histidine kinases. *Microbiol Mol Biol Rev* 70(4):910-938.
191. Gao R, Mack TR, & Stock AM (2007) Bacterial response regulators: versatile regulatory strategies from common domains. *Trends Biochem Sci* 32(5):225-234.
192. Mitrophanov AY & Groisman EA (2008) Signal integration in bacterial two-component regulatory systems. *Genes Dev* 22(19):2601-2611.
193. West AH & Stock AM (2001) Histidine kinases and response regulator proteins in two-component signaling systems. *Trends in Biochemical Sciences* 26(6):369-376.
194. Gao R & Stock AM (2009) Biological insights from structures of two-component proteins. *Annu Rev Microbiol* 63:133-154.
195. Rodrigue A, Quentin Y, Lazdunski A, Mejean V, & Foglino M (2000) Two-component systems in *Pseudomonas aeruginosa*: why so many? *Trends Microbiol* 8(11):498-504.
196. Mizuno T (1997) Compilation of all genes encoding two-component phosphotransfer signal transducers in the genome of *Escherichia coli*. *DNA Res* 4(2):161-168.
197. Stock JB, Stock AM, & Mottonen JM (1990) Signal transduction in bacteria. *Nature* 344(6265):395-400.
198. Wylie D, Stock A, Wong CY, & Stock J (1988) Sensory transduction in bacterial chemotaxis involves phosphotransfer between Che proteins. *Biochem Biophys Res Commun* 151(2):891-896.
199. Ninfa AJ, Ninfa EG, Lupas AN, Stock A, Magasanik B, & Stock J (1988) Crosstalk between bacterial chemotaxis signal transduction proteins and regulators of transcription of the Ntr regulon: evidence that nitrogen assimilation and chemotaxis are controlled by a common phosphotransfer mechanism. *Proc Natl Acad Sci U S A* 85(15):5492-5496.
200. Weigel N, Powers DA, & Roseman S (1982) Sugar transport by the bacterial phosphotransferase system. Primary structure and active site of a general

- phosphocarrier protein (HPr) from *Salmonella typhimurium*. *J Biol Chem* 257(23):14499-14509.
201. Surette MG, Levit M, Liu Y, Lukat G, Ninfa EG, Ninfa A, & Stock JB (1996) Dimerization is required for the activity of the protein histidine kinase CheA that mediates signal transduction in bacterial chemotaxis. *J Biol Chem* 271(2):939-945.
 202. Amemura M, Makino K, Shinagawa H, & Nakata A (1990) Cross talk to the phosphate regulon of *Escherichia coli* by PhoM protein: PhoM is a histidine protein kinase and catalyzes phosphorylation of PhoB and PhoM-open reading frame 2. *J Bacteriol* 172(11):6300-6307.
 203. Anthony RS & Spector LB (1972) Phosphorylated acetate kinase. Its isolation and reactivity. *J Biol Chem* 247(7):2120-2125.
 204. Jencks WP (1980) The utilization of binding energy in coupled vectorial processes. *Adv Enzymol Relat Areas Mol Biol* 51:75-106.
 205. Sanders DA, Gillece-Castro BL, Stock AM, Burlingame AL, & Koshland DE, Jr. (1989) Identification of the site of phosphorylation of the chemotaxis response regulator protein, CheY. *J Biol Chem* 264(36):21770-21778.
 206. Sanders DA, Gillece-Castro BL, Burlingame AL, & Koshland DE, Jr. (1992) Phosphorylation site of NtrC, a protein phosphatase whose covalent intermediate activates transcription. *J Bacteriol* 174(15):5117-5122.
 207. Hess JF, Oosawa K, Kaplan N, & Simon MI (1988) Phosphorylation of three proteins in the signaling pathway of bacterial chemotaxis. *Cell* 53(1):79-87.
 208. Lin CT, Huang YJ, Chu PH, Hsu JL, Huang CH, & Peng HL (2006) Identification of an HptB-mediated multi-step phosphorelay in *Pseudomonas aeruginosa* PAO1. *Research in Microbiology* 157(2):169-175.
 209. Chen YT, Chang HY, Lu CL, & Peng HL (2004) Evolutionary analysis of the two-component systems in *Pseudomonas aeruginosa* PAO1. *J Mol Evol* 59(6):725-737.
 210. Grebe TW & Stock JB (1999) The histidine protein kinase superfamily. *Adv Microb Physiol* 41:139-227.
 211. Mikkelsen H, Sivaneson M, & Filloux A (2011) Key two-component regulatory systems that control biofilm formation in *Pseudomonas aeruginosa*. *Environ Microbiol* 13(7):1666-1681.
 212. Vallet I, Diggle SP, Stacey RE, Camara M, Ventre I, Lory S, Lazdunski A, Williams P, & Filloux A (2004) Biofilm formation in *Pseudomonas aeruginosa*: fimbrial cup gene clusters are controlled by the transcriptional regulator MvaT. *J Bacteriol* 186(9):2880-2890.
 213. Vallet I, Olson JW, Lory S, Lazdunski A, & Filloux A (2001) The chaperone/usher pathways of *Pseudomonas aeruginosa*: identification of fimbrial gene clusters (cup) and their involvement in biofilm formation. *Proc Natl Acad Sci U S A* 98(12):6911-6916.
 214. Kuchma SL, Connolly JP, & O'Toole GA (2005) A three-component regulatory system regulates biofilm maturation and type III secretion in *Pseudomonas aeruginosa*. *J Bacteriol* 187(4):1441-1454.
 215. Hobbs M, Collie ES, Free PD, Livingston SP, & Mattick JS (1993) PilS and PilR, a two-component transcriptional regulatory system controlling expression of type 4 fimbriae in *Pseudomonas aeruginosa*. *Mol Microbiol* 7(5):669-682.
 216. Whitchurch CB, Alm RA, & Mattick JS (1996) The alginate regulator AlgR and an associated sensor FimS are required for twitching motility in *Pseudomonas aeruginosa*. *Proc Natl Acad Sci U S A* 93(18):9839-9843.
 217. Yahr TL & Greenberg EP (2004) The genetic basis for the commitment to chronic versus acute infection in *Pseudomonas aeruginosa*. *Mol Cell* 16(4):497-498.

218. Goodman AL, Kulasekara B, Rietsch A, Boyd D, Smith RS, & Lory S (2004) A signaling network reciprocally regulates genes associated with acute infection and chronic persistence in *Pseudomonas aeruginosa*. *Dev Cell* 7(5):745-754.
219. Mougous JD, Cuff ME, Raunser S, Shen A, Zhou M, Gifford CA, Goodman AL, Joachimiak G, Ordonez CL, Lory S, Walz T, Joachimiak A, & Mekalanos JJ (2006) A virulence locus of *Pseudomonas aeruginosa* encodes a protein secretion apparatus. *Science* 312(5779):1526-1530.
220. Potvin E, Lehoux DE, Kukavica-Ibrulj I, Richard KL, Sanschagrin F, Lau GW, & Levesque RC (2003) In vivo functional genomics of *Pseudomonas aeruginosa* for high-throughput screening of new virulence factors and antibacterial targets. *Environ Microbiol* 5(12):1294-1308.
221. Ventre I, Goodman AL, Vallet-Gely I, Vasseur P, Soscia C, Molin S, Bleves S, Lazdunski A, Lory S, & Filloux A (2006) Multiple sensors control reciprocal expression of *Pseudomonas aeruginosa* regulatory RNA and virulence genes. *Proceedings of the National Academy of Sciences of the United States of America* 103(1):171-176.
222. Chambonnier G, Roux L, Redelberger D, Fadel F, Filloux A, Sivaneson M, de Bentzmann S, & Bordi C (2016) The Hybrid Histidine Kinase LadS Forms a Multicomponent Signal Transduction System with the GacS/GacA Two-Component System in *Pseudomonas aeruginosa*. *PLoS Genet* 12(5):e1006032.
223. Parkins MD, Ceri H, & Storey DG (2001) *Pseudomonas aeruginosa* GacA, a factor in multihost virulence, is also essential for biofilm formation. *Mol Microbiol* 40(5):1215-1226.
224. Brencic A, McFarland KA, McManus HR, Castang S, Mogno I, Dove SL, & Lory S (2009) The GacS/GacA signal transduction system of *Pseudomonas aeruginosa* acts exclusively through its control over the transcription of the RsmY and RsmZ regulatory small RNAs. *Mol Microbiol* 73(3):434-445.
225. Lapouge K, Schubert M, Allain FH, & Haas D (2008) Gac/Rsm signal transduction pathway of gamma-proteobacteria: from RNA recognition to regulation of social behaviour. *Mol Microbiol* 67(2):241-253.
226. Records AR & Gross DC (2010) Sensor kinases RetS and LadS regulate *Pseudomonas syringae* type VI secretion and virulence factors. *J Bacteriol* 192(14):3584-3596.
227. Bordi C, Lamy M-C, Ventre I, Termine E, Hachani A, Fillet S, Roche B, Bleves S, Méjean V, Lazdunski A, & Filloux A (2010) Regulatory RNAs and the HptB/RetS signalling pathways fine-tune *Pseudomonas aeruginosa* pathogenesis. *Molecular Microbiology* 76(6):1427-1443.
228. Hsu JL, Chen HC, Peng HL, & Chang HY (2008) Characterization of the histidine-containing phosphotransfer protein B-mediated multistep phosphorelay system in *Pseudomonas aeruginosa* PAO1. *Journal of Biological Chemistry* 283(15):9933-9944.
229. Bhuwan M, Lees HJ, Pengs HL, & Chang HY (2012) Histidine-containing phosphotransfer protein-B (HptB) regulates swarming motility through partner-switching system in *Pseudomonas aeruginosa* PAO1 strain. *Journal of Biological Chemistry* 287(3):1903-1914.
230. Fuqua C, Winans SC, & Greenberg EP (1996) Census and consensus in bacterial ecosystems: the LuxR-LuxI family of quorum-sensing transcriptional regulators. *Annu Rev Microbiol* 50:727-751.
231. Gray KM (1997) Intercellular communication and group behavior in bacteria. *Trends Microbiol* 5(5):184-188.

232. Kleerebezem M, Quadri LE, Kuipers OP, & de Vos WM (1997) Quorum sensing by peptide pheromones and two-component signal-transduction systems in Gram-positive bacteria. *Mol Microbiol* 24(5):895-904.
233. Fuqua WC, Winans SC, & Greenberg EP (1994) Quorum sensing in bacteria: the LuxR-LuxI family of cell density-responsive transcriptional regulators. *J Bacteriol* 176(2):269-275.
234. Latifi A, Winson MK, Foglino M, Bycroft BW, Stewart GS, Lazdunski A, & Williams P (1995) Multiple homologues of LuxR and LuxI control expression of virulence determinants and secondary metabolites through quorum sensing in *Pseudomonas aeruginosa* PAO1. *Mol Microbiol* 17(2):333-343.
235. Passador L, Cook JM, Gambello MJ, Rust L, & Iglewski BH (1993) Expression of *Pseudomonas aeruginosa* virulence genes requires cell-to-cell communication. *Science* 260(5111):1127-1130.
236. Gambello MJ & Iglewski BH (1991) Cloning and characterization of the *Pseudomonas aeruginosa* lasR gene, a transcriptional activator of elastase expression. *J Bacteriol* 173(9):3000-3009.
237. Gambello MJ, Kaye S, & Iglewski BH (1993) LasR of *Pseudomonas aeruginosa* is a transcriptional activator of the alkaline protease gene (apr) and an enhancer of exotoxin A expression. *Infect Immun* 61(4):1180-1184.
238. Seed PC, Passador L, & Iglewski BH (1995) Activation of the *Pseudomonas aeruginosa* lasI gene by LasR and the *Pseudomonas* autoinducer PAI: an autoinduction regulatory hierarchy. *J Bacteriol* 177(3):654-659.
239. Reimmann C, Beyeler M, Latifi A, Winteler H, Foglino M, Lazdunski A, & Haas D (1997) The global activator GacA of *Pseudomonas aeruginosa* PAO positively controls the production of the autoinducer N-butyryl-homoserine lactone and the formation of the virulence factors pyocyanin, cyanide, and lipase. *Mol Microbiol* 24(2):309-319.
240. Albus AM, Pesci EC, Runyen-Janecky LJ, West SE, & Iglewski BH (1997) Vfr controls quorum sensing in *Pseudomonas aeruginosa*. *J Bacteriol* 179(12):3928-3935.
241. Pearson JP, Passador L, Iglewski BH, & Greenberg EP (1995) A second N-acylhomoserine lactone signal produced by *Pseudomonas aeruginosa*. *Proc Natl Acad Sci U S A* 92(5):1490-1494.
242. Ochsner UA, Koch AK, Fiechter A, & Reiser J (1994) Isolation and characterization of a regulatory gene affecting rhamnolipid biosurfactant synthesis in *Pseudomonas aeruginosa*. *J Bacteriol* 176(7):2044-2054.
243. Winson MK, Camara M, Latifi A, Foglino M, Chhabra SR, Daykin M, Bally M, Chapon V, Salmond GP, Bycroft BW, & et al. (1995) Multiple N-acyl-L-homoserine lactone signal molecules regulate production of virulence determinants and secondary metabolites in *Pseudomonas aeruginosa*. *Proc Natl Acad Sci U S A* 92(20):9427-9431.
244. Ochsner UA & Reiser J (1995) Autoinducer-mediated regulation of rhamnolipid biosurfactant synthesis in *Pseudomonas aeruginosa*. *Proc Natl Acad Sci U S A* 92(14):6424-6428.
245. Ochsner UA, Fiechter A, & Reiser J (1994) Isolation, characterization, and expression in *Escherichia coli* of the *Pseudomonas aeruginosa* rhlAB genes encoding a rhamnosyltransferase involved in rhamnolipid biosurfactant synthesis. *J Biol Chem* 269(31):19787-19795.
246. Brint JM & Ohman DE (1995) Synthesis of multiple exoproducts in *Pseudomonas aeruginosa* is under the control of RhlR-RhlI, another set of regulators in strain PAO1 with homology to the autoinducer-responsive LuxR-LuxI family. *J Bacteriol* 177(24):7155-7163.

247. Pearson JP, Pesci EC, & Iglewski BH (1997) Roles of *Pseudomonas aeruginosa* las and rhl quorum-sensing systems in control of elastase and rhamnolipid biosynthesis genes. *J Bacteriol* 179(18):5756-5767.
248. Latifi A, Foglino M, Tanaka K, Williams P, & Lazdunski A (1996) A hierarchical quorum-sensing cascade in *Pseudomonas aeruginosa* links the transcriptional activators LasR and RhIR (VsmR) to expression of the stationary-phase sigma factor RpoS. *Mol Microbiol* 21(6):1137-1146.
249. Loewen PC & Hengge-Aronis R (1994) The role of the sigma factor sigma S (KatF) in bacterial global regulation. *Annu Rev Microbiol* 48:53-80.
250. Pesci EC, Pearson JP, Seed PC, & Iglewski BH (1997) Regulation of las and rhl quorum sensing in *Pseudomonas aeruginosa*. *J Bacteriol* 179(10):3127-3132.
251. McKnight SL, Iglewski BH, & Pesci EC (2000) The *Pseudomonas* quinolone signal regulates rhl quorum sensing in *Pseudomonas aeruginosa*. *J Bacteriol* 182(10):2702-2708.
252. Gallagher LA, McKnight SL, Kuznetsova MS, Pesci EC, & Manoil C (2002) Functions required for extracellular quinolone signaling by *Pseudomonas aeruginosa*. *J Bacteriol* 184(23):6472-6480.
253. McGrath S, Wade DS, & Pesci EC (2004) Dueling quorum sensing systems in *Pseudomonas aeruginosa* control the production of the *Pseudomonas* quinolone signal (PQS). *FEMS Microbiol Lett* 230(1):27-34.
254. Wade DS, Calfee MW, Rocha ER, Ling EA, Engstrom E, Coleman JP, & Pesci EC (2005) Regulation of *Pseudomonas* quinolone signal synthesis in *Pseudomonas aeruginosa*. *J Bacteriol* 187(13):4372-4380.
255. Bala A, Chhibber S, & Harjai K (2014) *Pseudomonas* quinolone signalling system: a component of quorum sensing cascade is a crucial player in the acute urinary tract infection caused by *Pseudomonas aeruginosa*. *Int J Med Microbiol* 304(8):1199-1208.
256. Hassett DJ, Ma JF, Elkins JG, McDermott TR, Ochsner UA, West SE, Huang CT, Fredericks J, Burnett S, Stewart PS, McFeters G, Passador L, & Iglewski BH (1999) Quorum sensing in *Pseudomonas aeruginosa* controls expression of catalase and superoxide dismutase genes and mediates biofilm susceptibility to hydrogen peroxide. *Mol Microbiol* 34(5):1082-1093.
257. Bjarnsholt T, Jensen PO, Burmolle M, Hentzer M, Haagensen JA, Hougen HP, Calum H, Madsen KG, Moser C, Molin S, Hoiby N, & Givskov M (2005) *Pseudomonas aeruginosa* tolerance to tobramycin, hydrogen peroxide and polymorphonuclear leukocytes is quorum-sensing dependent. *Microbiology* 151(Pt 2):373-383.
258. Pultz IS, Christen M, Kulasekara HD, Kennard A, Kulasekara B, & Miller SI (2012) The response threshold of *Salmonella* PilZ domain proteins is determined by their binding affinities for c-di-GMP. *Mol Microbiol* 86(6):1424-1440.
259. Ho CL, Chong KS, Oppong JA, Chuah ML, Tan SM, & Liang ZX (2013) Visualizing the perturbation of cellular cyclic di-GMP levels in bacterial cells. *J Am Chem Soc* 135(2):566-569.
260. Paul K, Nieto V, Carlquist WC, Blair DF, & Harshey RM (2010) The c-di-GMP binding protein YcgR controls flagellar motor direction and speed to affect chemotaxis by a "backstop brake" mechanism. *Mol Cell* 38(1):128-139.
261. Morgan JL, Strumillo J, & Zimmer J (2013) Crystallographic snapshot of cellulose synthesis and membrane translocation. *Nature* 493(7431):181-186.
262. Morgan JL, McNamara JT, & Zimmer J (2014) Mechanism of activation of bacterial cellulose synthase by cyclic di-GMP. *Nat Struct Mol Biol* 21(5):489-496.
263. Wilksch JJ, Yang J, Clements A, Gabbe JL, Short KR, Cao H, Cavaliere R, James CE, Whitchurch CB, Schembri MA, Chuah ML, Liang ZX, Wijburg OL, Jenney AW,

- Lithgow T, & Strugnell RA (2011) MrkH, a novel c-di-GMP-dependent transcriptional activator, controls *Klebsiella pneumoniae* biofilm formation by regulating type 3 fimbriae expression. *PLoS Pathog* 7(8):e1002204.
264. Tan JW, Wilksch JJ, Hocking DM, Wang N, Srikhanta YN, Tauschek M, Lithgow T, Robins-Browne RM, Yang J, & Strugnell RA (2015) Positive autoregulation of mrkHI by the cyclic di-GMP-dependent MrkH protein in the biofilm regulatory circuit of *Klebsiella pneumoniae*. *J Bacteriol* 197(9):1659-1667.
 265. Russell MH, Bible AN, Fang X, Gooding JR, Campagna SR, Gomelsky M, & Alexandre G (2013) Integration of the Second Messenger c-di-GMP into the Chemotactic Signaling Pathway. *mBio* 4(2).
 266. McCarthy Y, Ryan RP, O'Donovan K, He YQ, Jiang BL, Feng JX, Tang JL, & Dow JM (2008) The role of PilZ domain proteins in the virulence of *Xanthomonas campestris* pv. *campestris*. *Mol Plant Pathol* 9(6):819-824.
 267. Yang F, Tian F, Chen H, Hutchins W, Yang CH, & He C (2015) The *Xanthomonas oryzae* pv. *oryzae* PilZ Domain Proteins Function Differentially in Cyclic di-GMP Binding and Regulation of Virulence and Motility. *Appl Environ Microbiol* 81(13):4358-4367.
 268. Pitzer JE, Sultan SZ, Hayakawa Y, Hobbs G, Miller MR, & Motaleb MA (2011) Analysis of the *Borrelia burgdorferi* cyclic-di-GMP-binding protein PlzA reveals a role in motility and virulence. *Infect Immun* 79(5):1815-1825.
 269. He M, Zhang JJ, Ye M, Lou Y, & Yang XF (2014) Cyclic Di-GMP receptor PlzA controls virulence gene expression through RpoS in *Borrelia burgdorferi*. *Infect Immun* 82(1):445-452.
 270. Ryan RP, McCarthy Y, Kiely PA, O'Connor R, Farah CS, Armitage JP, & Dow JM (2012) Dynamic complex formation between HD-GYP, GGDEF and PilZ domain proteins regulates motility in *Xanthomonas campestris*. *Molecular Microbiology* 86(3):557-567.
 271. Matsuyama BY, Krasteva PV, Baraquet C, Harwood CS, Sondermann H, & Navarro MV (2016) Mechanistic insights into c-di-GMP-dependent control of the biofilm regulator FleQ from *Pseudomonas aeruginosa*. *Proc Natl Acad Sci U S A* 113(2):E209-218.
 272. An SQ, Caly DL, McCarthy Y, Murdoch SL, Ward J, Febrer M, Dow JM, & Ryan RP (2014) Novel cyclic di-GMP effectors of the YajQ protein family control bacterial virulence. *PLoS Pathog* 10(10):e1004429.
 273. Rao F, Pasunooti S, Ng Y, Zhuo W, Lim L, Liu AW, & Liang ZX (2009) Enzymatic synthesis of c-di-GMP using a thermophilic diguanylate cyclase. *Analytical Biochemistry* 389(2):138-142.
 274. Pasunooti S, Surya W, Tan SN, & Liang Z-X (2010) Sol-gel immobilization of a thermophilic diguanylate cyclase for enzymatic production of cyclic-di-GMP. *Journal of Molecular Catalysis B: Enzymatic* 67(1-2):98-103.
 275. Tolker-Nielsen T & Sternberg C (2011) Growing and analyzing biofilms in flow chambers. *Curr Protoc Microbiol* Chapter 1:Unit 1B 2.
 276. Silverman M & Simon M (1974) Flagellar rotation and the mechanism of bacterial motility. *Nature* 249(452):73-74.
 277. Yang L, Rybtke MT, Jakobsen TH, Hentzer M, Bjarnsholt T, Givskov M, & Tolker-Nielsen T (2009) Computer-aided identification of recognized drugs as *Pseudomonas aeruginosa* quorum-sensing inhibitors. *Antimicrob Agents Chemother* 53(6):2432-2443.

278. Dietrich LE, Okegbe C, Price-Whelan A, Sakhtah H, Hunter RC, & Newman DK (2013) Bacterial community morphogenesis is intimately linked to the intracellular redox state. *J Bacteriol* 195(7):1371-1380.
279. Boehm A, Kaiser M, Li H, Spangler C, Kasper CA, Ackermann M, Kaefer V, Sourjik V, Roth V, & Jenal U (2010) Second messenger-mediated adjustment of bacterial swimming velocity. *Cell* 141(1):107-116.
280. Lori C, Ozaki S, Steiner S, Bohm R, Abel S, Dubey BN, Schirmer T, Hiller S, & Jenal U (2015) Cyclic di-GMP acts as a cell cycle oscillator to drive chromosome replication. *Nature* 523(7559):236-239.
281. Kearns DB (2010) A field guide to bacterial swarming motility. *Nature Reviews Microbiology* 8(9):634-644.
282. Caiazza NC, Shanks RMQ, & O'Toole GA (2005) Rhamnolipids modulate swarming motility patterns of *Pseudomonas aeruginosa*. *Journal of Bacteriology* 187(21):7351-7361.
283. Gupta R & Schuster M (2012) Quorum sensing modulates colony morphology through alkyl quinolones in *Pseudomonas aeruginosa*. *BMC Microbiol* 12:30.
284. Starkey M, Hickman JH, Ma L, Zhang N, De Long S, Hinz A, Palacios S, Manoil C, Kirisits MJ, Starner TD, Wozniak DJ, Harwood CS, & Parsek MR (2009) *Pseudomonas aeruginosa* rugose small-colony variants have adaptations that likely promote persistence in the cystic fibrosis lung. *J Bacteriol* 191(11):3492-3503.
285. Petrova OE & Sauer K (2011) SagS contributes to the motile-sessile switch and acts in concert with BfiSR to enable *Pseudomonas aeruginosa* biofilm formation. *Journal of Bacteriology* 193(23):6614-6628.
286. Tielker D, Hacker S, Loris R, Strathmann M, Wingender J, Wilhelm S, Rosenau F, & Jaeger KE (2005) *Pseudomonas aeruginosa* lectin LecB is located in the outer membrane and is involved in biofilm formation. *Microbiology* 151(Pt 5):1313-1323.
287. Ha DG, Merritt JH, Hampton TH, Hodgkinson JT, Janecek M, Spring DR, Welch M, & O'Toole GA (2011) 2-Heptyl-4-quinolone, a precursor of the *Pseudomonas* quinolone signal molecule, modulates swarming motility in *Pseudomonas aeruginosa*. *J Bacteriol* 193(23):6770-6780.
288. Haussler S & Becker T (2008) The pseudomonas quinolone signal (PQS) balances life and death in *Pseudomonas aeruginosa* populations. *PLoS Pathog* 4(9):e1000166.
289. Tian ZX, Mac Aogain M, O'Connor HF, Fargier E, Mooij MJ, Adams C, Wang YP, & O'Gara F (2009) MexT modulates virulence determinants in *Pseudomonas aeruginosa* independent of the MexEF-OprN efflux pump. *Microb Pathog* 47(4):237-241.
290. Gupta K, Marques CN, Petrova OE, & Sauer K (2013) Antimicrobial tolerance of *Pseudomonas aeruginosa* biofilms is activated during an early developmental stage and requires the two-component hybrid SagS. *J Bacteriol* 195(21):4975-4987.
291. Petrova OE & Sauer K (2009) A novel signaling network essential for regulating *Pseudomonas aeruginosa* biofilm development. *PLoS Pathogens* 5(11).
292. Petrova OE & Sauer K (2010) The Novel Two-Component Regulatory System BfiSR Regulates Biofilm Development by Controlling the Small RNA rsmZ through CafA. *Journal of Bacteriology* 192(20):5275-5288.
293. Petrova OE & Sauer K (2012) Sticky situations: key components that control bacterial surface attachment. *J Bacteriol* 194(10):2413-2425.
294. Kong W, Chen L, Zhao J, Shen T, Surette MG, Shen L, & Duan K (2013) Hybrid sensor kinase PA1611 in *Pseudomonas aeruginosa* regulates transitions between acute and chronic infection through direct interaction with RetS. *Mol Microbiol* 88(4):784-797.

295. Albanesi D, Martin M, Trajtenberg F, Mansilla MC, Haouz A, Alzari PM, de Mendoza D, & Buschiazzo A (2009) Structural plasticity and catalysis regulation of a thermosensor histidine kinase. *Proc Natl Acad Sci U S A* 106(38):16185-16190.
296. Gupta K, Liao J, Petrova OE, Cherny KE, & Sauer K (2014) Elevated levels of the second messenger c-di-GMP contribute to antimicrobial resistance of *Pseudomonas aeruginosa*. *Molecular Microbiology* 92(3):488-506.
297. Costerton JW (2001) Cystic fibrosis pathogenesis and the role of biofilms in persistent infection. *Trends Microbiol* 9(2):50-52.
298. Sauer K, Camper AK, Ehrlich GD, Costerton JW, & Davies DG (2002) *Pseudomonas aeruginosa* displays multiple phenotypes during development as a biofilm. *J Bacteriol* 184(4):1140-1154.
299. O'Toole G, Kaplan HB, & Kolter R (2000) Biofilm formation as microbial development. *Annu Rev Microbiol* 54:49-79.
300. Boyd CD & O'Toole GA (2012) Second messenger regulation of biofilm formation: breakthroughs in understanding c-di-GMP effector systems. *Annu Rev Cell Dev Biol* 28:439-462.
301. Shrout JD, Chopp DL, Just CL, Hentzer M, Givskov M, & Parsek MR (2006) The impact of quorum sensing and swarming motility on *Pseudomonas aeruginosa* biofilm formation is nutritionally conditional. *Molecular Microbiology* 62(5):1264-1277.
302. Galperin MY (2006) Structural Classification of Bacterial Response Regulators: Diversity of Output Domains and Domain Combinations. *Journal of Bacteriology* 188(12):4169-4182.
303. Aldridge P, Paul R, Goymer P, Rainey P, & Jenal U (2003) Role of the GGDEF regulator PleD in polar development of *Caulobacter crescentus*. *Mol Microbiol* 47(6):1695-1708.
304. Petters T, Zhang X, Nesper J, Treuner-Lange A, Gomez-Santos N, Hoppert M, Jenal U, & Sogaard-Andersen L (2012) The orphan histidine protein kinase SgmT is a c-di-GMP receptor and regulates composition of the extracellular matrix together with the orphan DNA binding response regulator DigR in *Myxococcus xanthus*. *Molecular Microbiology* 84(1):147-165.
305. Iuchi S & Weiner L (1996) Cellular and molecular physiology of *Escherichia coli* in the adaptation to aerobic environments. *J Biochem* 120(6):1055-1063.
306. Madhusudan, Zapf J, Whiteley JM, Hoch JA, Xuong NH, & Varughese KI (1996) Crystal structure of a phosphatase-resistant mutant of sporulation response regulator Spo0F from *Bacillus subtilis*. *Structure* 4(6):679-690.
307. Hoch JA (1993) Regulation of the phosphorelay and the initiation of sporulation in *Bacillus subtilis*. *Annu Rev Microbiol* 47:441-465.
308. Perego M (1998) Kinase-phosphatase competition regulates *Bacillus subtilis* development. *Trends Microbiol* 6(9):366-370.
309. Domian IJ, Quon KC, & Shapiro L (1996) The control of temporal and spatial organization during the *Caulobacter* cell cycle. *Curr Opin Genet Dev* 6(5):538-544.
310. Tuckerman JR, Gonzalez G, & Gilles-Gonzalez MA (2001) Complexation precedes phosphorylation for two-component regulatory system FixL/FixJ of *Sinorhizobium meliloti*. *J Mol Biol* 308(3):449-455.
311. Kaplan HB & Plamann L (1996) A *Myxococcus xanthus* cell density-sensing system required for multicellular development. *FEMS Microbiol Lett* 139(2-3):89-95.
312. Xu Q, Porter SW, & West AH (2003) The yeast YPD1/SLN1 complex: insights into molecular recognition in two-component signaling systems. *Structure* 11(12):1569-1581.

313. Porter SW, Xu Q, & West AH (2003) Ssk1p response regulator binding surface on histidine-containing phosphotransfer protein Ypd1p. *Eukaryot Cell* 2(1):27-33.
314. Xu L, Venkataramani P, Ding Y, Liu Y, Deng Y, Yong GL, Xin L, Ye R, Zhang L, Yang L, & Liang ZX (2016) A Cyclic di-GMP-binding Adaptor Protein Interacts with Histidine Kinase to Regulate Two-component Signaling. *J Biol Chem* 291(31):16112-16123.
315. Fujiwara T, Komoda K, Sakurai N, Tajima K, Tanaka I, & Yao M (2013) The c-di-GMP recognition mechanism of the PilZ domain of bacterial cellulose synthase subunit A. *Biochem Biophys Res Commun* 431(4):802-807.
316. Li TN, Chin KH, Liu JH, Wang AH, & Chou SH (2009) XC1028 from *Xanthomonas campestris* adopts a PilZ domain-like structure without a c-di-GMP switch. *Proteins* 75(2):282-288.
317. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, & Bax A (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* 6(3):277-293.
318. Guntert P (2004) Automated NMR structure calculation with CYANA. *Methods Mol Biol* 278:353-378.
319. Volz K & Matsumura P (1991) Crystal structure of *Escherichia coli* CheY refined at 1.7-Å resolution. *Journal of Biological Chemistry* 266(23):15511-15519.
320. Robinson VL, Buckler DR, & Stock AM (2000) A tale of two components: a novel kinase and a regulatory switch. *Nat Struct Biol* 7(8):626-633.
321. Baikalov I, Schroder I, Kaczor-Grzeskowiak M, Grzeskowiak K, Gunsalus RP, & Dickerson RE (1996) Structure of the *Escherichia coli* response regulator NarL. *Biochemistry* 35(34):11053-11061.
322. Lewis RJ, Brannigan JA, Muchova K, Barak I, & Wilkinson AJ (1999) Phosphorylated aspartate in the structure of a response regulator protein. *J Mol Biol* 294(1):9-15.
323. Volkman BF, Nohaile MJ, Amy NK, Kustu S, & Wemmer DE (1995) Three-dimensional solution structure of the N-terminal receiver domain of NTRC. *Biochemistry* 34(4):1413-1424.
324. Djordjevic S, Goudreau PN, Xu Q, Stock AM, & West AH (1998) Structural basis for methylesterase CheB regulation by a phosphorylation-activated domain. *Proc Natl Acad Sci U S A* 95(4):1381-1386.
325. Bell CH, Porter SL, Strawson A, Stuart DI, & Armitage JP (2010) Using structural information to change the phosphotransfer specificity of a two-component chemotaxis signalling complex. *PLoS Biol* 8(2):e1000306.
326. Muller-Dieckmann HJ, Grantz AA, & Kim SH (1999) The structure of the signal receiver domain of the *Arabidopsis thaliana* ethylene receptor ETR1. *Structure* 7(12):1547-1556.
327. Robert X & Gouet P (2014) Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res* 42(Web Server issue):W320-324.
328. Zapf J, Sen U, Madhusudan, Hoch JA, & Varughese KI (2000) A transient interaction between two phosphorelay proteins trapped in a crystal lattice reveals the mechanism of molecular recognition and phosphotransfer in signal transduction. *Structure* 8(8):851-862.
329. Kim DE, Chivian D, & Baker D (2004) Protein structure prediction and analysis using the Robetta server. *Nucleic Acids Res* 32(Web Server issue):W526-531.
330. Xu Q & West AH (1999) Conservation of structure and function among histidine-containing phosphotransfer (HPt) domains as revealed by the crystal structure of YPD1. *J Mol Biol* 292(5):1039-1050.

331. Mourey L, Da Re S, Pedelacq JD, Tolstykh T, Faurie C, Guillet V, Stock JB, & Samama JP (2001) Crystal structure of the CheA histidine phosphotransfer domain that mediates response regulator phosphorylation in bacterial chemotaxis. *J Biol Chem* 276(33):31074-31082.
332. Rogov VV, Bernhard F, Lohr F, & Dotsch V (2004) Solution structure of the Escherichia coli YojN histidine-phosphotransferase domain and its interaction with cognate phosphoryl receiver domains. *J Mol Biol* 343(4):1035-1048.
333. Kato M, Mizuno T, Shimizu T, & Hakoshima T (1997) Insights into multistep phosphorelay from the crystal structure of the C-terminal HPt domain of ArcB. *Cell* 88(5):717-723.
334. Arnold K, Bordoli L, Kopp J, & Schwede T (2006) The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics* 22(2):195-201.
335. Biasini M, Bienert S, Waterhouse A, Arnold K, Studer G, Schmidt T, Kiefer F, Gallo Cassarino T, Bertoni M, Bordoli L, & Schwede T (2014) SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res* 42(Web Server issue):W252-258.
336. Kiefer F, Arnold K, Kunzli M, Bordoli L, & Schwede T (2009) The SWISS-MODEL Repository and associated resources. *Nucleic Acids Res* 37(Database issue):D387-392.
337. Chou SH & Galperin MY (2016) Diversity of Cyclic Di-GMP-Binding Proteins and Mechanisms. *J Bacteriol* 198(1):32-46.
338. Cunningham BC & Wells JA (1989) High-resolution epitope mapping of hGH-receptor interactions by alanine-scanning mutagenesis. *Science* 244(4908):1081-1085.
339. Wells JA (1991) Systematic mutational analyses of protein-protein interfaces. *Methods Enzymol* 202:390-411.
340. Volz K (1993) Structural conservation in the CheY superfamily. *Biochemistry* 32(44):11741-11753.
341. Krasteva PV, Fong JC, Shikuma NJ, Beyhan S, Navarro MV, Yildiz FH, & Sondermann H (2010) Vibrio cholerae VpsT regulates matrix production and motility by directly sensing cyclic di-GMP. *Science* 327(5967):866-868.
342. Tschowri N, Schumacher MA, Schlimpert S, Chinnam NB, Findlay KC, Brennan RG, & Buttner MJ (2014) Tetrameric c-di-GMP mediates effective transcription factor dimerization to control Streptomyces development. *Cell* 158(5):1136-1147.

Appendix I

Figure S1. Agarose DNA gel shows the size of the *P. aeruginosa* gDNA fragments (the brightest band of each lane) captured by the prey vectors.

CckA	--ADGDTA I EVSDDGPGIPPDVMGK I DPFFTTKP----VGE G TGLGLATV V GIVKQSD
SagS	----AEGV R ISVRDTGIGIAQEALDR I FQPFTQADAGITRQYGGTGLGLALT R KLCEAMQ
BfiS	---PGGL L CLWVRDQGPVPADQLEH I ETPFHTS-----KPEGLGLGLS M SR S IVEGFG
PA1611	LDHDVLW L TC A VHDSGIGISPERLEH M DAFQQADSSISRRYGGTGLGLA I A R TLAERM G
PA1976	----RQRL S IEVWDTGVGIAADKLGE I FQEFKRGESQRCHQDRGLGLGLA I V D KIAR M LG
RetS	-QGETPRL R IAVQDSGHFPDAKEREAL L TAELHSGD F LSASKLGSHLGL I IA R QLVRL M G
LadS	-THDGVAL R VEVIDTGIGFDMAAGSD L YQRFVQADSS L TRGYGGLGIGLALC R KLVEL L G

CckA	G I HVHSRPNEGA A F R I F LPVYEAPA-----
SagS	G E LTVESTVGLGSL F S V GLPLAPVSP--PLQAL--PLRGRVIAQCSANSGLAQL L QTWLP
BfiS	G T LEAELPADGG-----
PA1611	G T LQAESKEGSGST F T L E IPLPFQ S --PAHR-----
PA1976	HRVRVGS L PGKGSC F AIEVPLARHAP--RSRAE--P-----
RetS	G E FGIQSGSSQGT T L S L T LPLDPQQLENPTADLDG P LQGARLLVDDNETCRKVLVQ Q CS
LadS	G E LTHESRPGQGS R F L L R L QLTQPAQ--GLAPP--P-----

Figure S2. Sequence alignment of the kinase catalytic domains of CckA from *C. crescentus* with different histidine kinases in *P. aeruginosa*. The residues involved in c-di-GMP binding in CckA are shaded in cyan and the corresponding residues in histidine kinases that do not show similarity with the c-di-GMP binding residues are shaded in yellow.

Table S1. List of bacterial strains used in the study

S.No	Bacterial strain	Description	Source
1	BacterioMatch II reporter strain	Host strain used for library screening and detection of protein-protein interactions in BacterioMatch II 2-hybrid system.	Stratagene
2	<i>E. coli</i> BL21 (DE3)	Cells enabling high-level expression of heterologous proteins in <i>E. coli</i> .	Stratagene
3	<i>E. coli</i> Rosetta cell	Cells enabling high-level expression of heterologous proteins in <i>E. coli</i> .	Stratagene
4	<i>P. aeruginosa</i> PAO1	Wild type strain for transposon mutation	PA two-allele library [Ref (a)]
5	PW5685(<i>ΔhapZ</i>)	PA2799-C09::ISlacZ/hah, ISlacZ/hah Transposon inserted after nt 192 of PA2799 (300nt)	PA two-allele library [Ref (a)]
6	PW5729(<i>ΔsagS</i>)	SagS-D02::ISlacZ/hah, ISlacZ/hah Transposon inserted after nt 1203 of SagS (2361nt)	PA two-allele library [Ref (a)]
7	PW1288 (<i>ΔPA0169</i>)	PA0169-F06::ISphoA/hah, ISphoA/hah Transposon inserted after nt 385 of PA0169 (708nt)	PA two-allele library [Ref (a)]
8	PW1531 (<i>ΔPA0290</i>)	PA0290-B05::ISphoA/hah, ISphoA/hah Transposon inserted after nt 496 of PA0290 (972nt)	PA two-allele library [Ref (a)]
9	PW1627 (<i>ΔPA0338</i>)	PA0338-E12::ISphoA/hah, ISphoA/hah Transposon inserted after nt 585 of PA0338 (1131nt)	PA two-allele library [Ref (a)]
10	PW2544 (<i>ΔPA0847</i>)	PA0847-G10::ISphoA/hah, ISphoA/hah Transposon inserted after nt 988 of PA0847 (2208 nt)	PA two-allele library [Ref (a)]
11	PW3000 (<i>ΔPA1107</i>)	PA1107-A03::ISphoA/hah, ISphoA/hah Transposon inserted after nt 496 of PA1107 (1197 nt)	PA two-allele library [Ref (a)]
12	PW3023 (<i>ΔPA1120</i>)	PA1120-G01::ISlacZ/hah, ISlacZ/hah Transposon inserted after nt 304 of PA1120 (1308nt)	PA two-allele library [Ref (a)]
13	PW4247 (<i>ΔPA1851</i>)	PA1851-H02::ISlacZ/hah, ISlacZ/hah Transposon inserted after nt 25 of PA1851 (1206nt)	PA two-allele library [Ref (a)]
14	PW5641 (<i>ΔPA2771</i>)	PA2771-E11::ISlacZ/hah, ISlacZ/hah Transposon inserted after nt 492 of PA2771 (1026nt)	PA two-allele library [Ref (a)]
15	PW5818 (<i>ΔPA2870</i>)	PA2870-B06::ISlacZ/hah, ISlacZ/hah Transposon inserted after nt 75 of PA2870 (1578nt)	PA two-allele library [Ref (a)]
16	PW6316 (<i>ΔPA3177</i>)	PA3177-G07::ISphoA/hah, ISphoA/hah Transposon inserted after nt 377 of PA3177 (924 nt)	PA two-allele library [Ref (a)]
17	PW6632 (<i>ΔPA3343</i>)	PA3343-H12::ISphoA/hah, ISphoA/hah Transposon inserted after nt 993 of PA3343 (1170 nt)	PA two-allele library [Ref (a)]
18	PW7264 (<i>ΔPA3702</i>)	PA3702-H07::ISphoA/hah, ISphoA/hah Transposon inserted after nt 993 of PA3702 (1170 nt)	PA two-allele library [Ref (a)]
19	PW8315 (<i>ΔPA4332</i>)	PA4332-C04::ISphoA/hah, ISphoA/hah Transposon inserted after nt 647 of PA4332 (1464 nt)	PA two-allele library [Ref (a)]
20	PW8424 (<i>ΔPA4396</i>)	PA4396-D06::ISphoA/hah, ISphoA/hah	PA two-allele library

		Transposon inserted after nt 328 of PA4396 (1101 nt)	[Ref (a)]		
21	PW9146 (<i>ΔPA4843</i>)	PA4843-D06::ISphoA/hah , ISphoA/hah Transposon inserted after nt 231 of PA4843 (1629 nt)	PA two-allele library [Ref (a)]		
22	PW9303 (<i>ΔPA4929</i>)	PA4929-F02::ISlacZ/hah , ISlacZ/hah Transposon inserted after nt 323 of PA4929 (2043nt)	PA two-allele library [Ref (a)]		
23	PW10281(<i>ΔPA5487</i>)	PA5487-E07::ISphoA/hah , ISphoA/hah Transposon inserted after nt 263 of PA5487 (2016 nt)	PA two-allele library [Ref (a)]		
24	PW1521 (<i>ΔPA0285</i>)	PA0285-C07::ISphoA/hah , ISphoA/hah Transposon inserted after nt 956 of PA0285 (2283 nt)	PA two-allele library [Ref (a)]		
25	PW2060 (<i>ΔPA0575</i>)	PA0575-A03::ISphoA/hah , ISphoA/hah Transposon inserted after nt 1900 of PA0575 (3738 nt)	PA two-allele library [Ref (a)]		
26	PW2570 (<i>ΔPA0861</i>)	PA0861-A03::ISphoA/hah , ISphoA/hah Transposon inserted after nt 706 of PA0861 (2457 nt)	PA two-allele library [Ref (a)]		
27	PW4043 (<i>ΔPA1727</i>)	PA1727-C01::ISlacZ/hah , ISlacZ/hah Transposon inserted after nt 487 of PA1727 (2058 nt)	PA two-allele library [Ref (a)]		
28	PW4568 (<i>ΔPA2072</i>)	PA2072-B02::ISlacZ/hah , ISlacZ/hah Transposon inserted after nt 279 of PA2072 (2595 nt)	PA two-allele library [Ref (a)]		
29	PW8372 (<i>ΔPA4367</i>)	PA4367-C03::ISlacZ/hah , ISlacZ/hah Transposon inserted after nt 522 of PA4367 (2064 nt)	PA two-allele library [Ref (a)]		
30	PW8754 (<i>ΔPA4601</i>)	PA4601-C02::ISlacZ/hah , ISlacZ/hah Transposon inserted after nt 2905 of PA4601 (4248 nt)	PA two-allele library [Ref (a)]		
31	PW9424 (<i>ΔPA5017</i>)	PA5017-A01::ISphoA/hah , ISphoA/hah Transposon inserted after nt 1059 of PA5017 (2700 nt)	PA two-allele library [Ref (a)]		
32	<i>hapz</i> ⁺	The PAO1 strain of <i>P. aeruginosa</i> overexpressing the pUCP18-HapZ plasmid	This study		

Table S2. List of plasmids used in the study

S.No	Plasmid	Description	Source
1	pBT	Empty bait plasmid used for detection of protein-protein interaction in BacterioMatch II 2-hybrid system, Cm ^r	Agilent Technologies; catalog no. 240065
2	pBT-PA2960	A DNA fragment encoding full-length PA2960(1-119aa/end) was cloned into pBT	This study
3	pBT-HapZ	A DNA fragment encoding full length of PA2799 (1-99aa/end) cloned into pBT vector at EcoRI/BamHI site, Cm ^r	This study
4	pBT-PA2989	A DNA fragment encoding full length of PA2989 (1-254aa/end) cloned into pBT vector at BamHI/BgIII site, Cm ^r	This study
5	pBT-PA4324	A DNA fragment encoding full length of PA4324 (1-119aa/end) cloned into pBT vector at EcoRI/BgIII site, Cm ^r	This study
6	pBT-PA0012	A DNA fragment encoding full length of PA0012 (1-88aa/end) cloned into pBT vector at EcoRI/BgIII site, Cm ^r	This study
7	pBT-PA4608	A DNA fragment encoding full length of PA4608 (1-125aa/end) cloned into pBT vector EcoRI/BamHI, Cm ^r	This study
8	pBT-LGF2	Positive interaction control plasmid in BacterioMatch II 2-hybrid system, encoding dimerization domain of Gal4 transcriptional activator protein	Stratagene
9	pBT-HapZ-ΔN21aa	A DNA fragment encoding truncated PA2799 (22-99aa/end) cloned into pBT vector at EcoRI/BamHI site, Cm ^r	This study
10	pBT-HapZ-Y8A		
11	pBT-HapZ-E10A	Plasmid was made by site-directed mutagenesis with pBT-HapZ as template	This study
12	pBT-HapZ-K11A		
13	pBT-HapZ-R12A		
14	pBT-HapZ-F14A		
15	pBT-HapZ-R16A		
16	pBT-HapZ-L19A		
17	pBT-HapZ-E31A		
18	pBT-HapZ-L35A		
19	pBT-HapZ-L45A		
20	pBT-HapZ-Q49A		
21	pBT-HapZ-D56A		
22	pBT-HapZ-R59A		

23	pBT-HapZ-E67A	Plasmid was made by site-directed mutagenesis with pBT-HapZ as template	This study
24	pBT-HapZ-E73A		
25	pBT-HapZ-R77A		
26	pBT-HapZ-V79A		
27	pBT-HapZ-R80A		
28	pBT-HapZ-E82A		
29	pBT-HapZ-L84A		
30	pBT-HapZ-G92E		
31	pBT-HapZ-R12AR16A		
32	pBT-HapZ-D38AS40A		
33	pBT-HapZ-E31AE67A	Plasmid was made by site-directed mutagenesis with pBT-HapZ as template	This study
34	pBT-HapZ-E31AE73A		
35	pBT-HapZ-E31RE73A		
36	pBT-HapZ-E31AE67R		
37	pBT-HapZ-E31AE73R		
38	pBT-HapZ-V79AR80A		
39	pBT-HapZ-R80AE82A		
40	pBT-HapZ-R80AG92E		
41	pBT-HapZ-T24Stop	Plasmid was made by site-directed mutagenesis with pBT-HapZ as template. The Thr-24 was mutated to a stop codon so that the construct only expressed the N-terminal loop of HapZ	This study
42	pTRG	Prey plasmid used for detection of protein-protein interaction in BacterioMatch II 2-hybrid system, Tetr ^r	Stratagene
43	pTRGA	A pTRG-derived mutation was made by insertion of the nucleotides 'A' between the end of C-terminal of RNAP-alpha (992bp) and BamHI site (993bp) of vector pTRG	This study
44	pTRGAC	A pTRG-derived mutation was made by insertion of the nucleotides 'AC' between the end of C-terminal of RNAP-alpha (992bp) and BamHI site (993bp) of vector pTRG	This study
45	Δ N317 _{sagS}	Original clone from pTRGA library with truncated <i>sagS</i> (317-786aa/end) fused to the C-terminal of RNAP-alpha in pTRGA vector, Tetr ^r	This study
46	Δ N377 _{sagS}	Original clone from pTRGA library with truncated <i>sagS</i> (377-786aa/end) fused to the C-terminal of RNAP-alpha in pTRGA	This study

		vector, Tetr ^r	
47	Δ N468 _{sagS}	Original clone from pTRGAC1 library with truncated <i>SagS</i> (468-786aa/end) fused to the C-terminal of RNAP-alpha in pTRGAC1 vector Tetr ^r	This study
48	pTRG-REC _{sagS}	A DNA fragment encoding C-terminal REC domain (663-786aa) of <i>SagS</i> cloned into pTRG vector at BamHI/XhoI site, Tetr ^r	This study
49	pTRG-REC _{PA1976}	A DNA fragment encoding C-terminal REC domain (757-881aa) of PA1976 cloned into pTRG vector at BamHI/XhoI site, Tetr ^r	This study
50	pTRG-REC _{PA1611}	A DNA fragment encoding C-terminal REC domain (515-651aa/end) of PA1611 cloned into pTRG vector BamHI/XhoI site, Tetr ^r	This study
51	pTRG-REC _{RetS}	A DNA fragment encoding C-terminal REC domain (808-942aa/end) of <i>RetS</i> cloned into pTRG vector BamHI/XhoI site, Tetr ^r	This study
52	pTRG-Gal11P	Positive interaction control plasmid in BacterioMatch II two-hybrid system, encoding a domain of mutant form of Gal11 protein, Tetr ^r	Stratagene
53	pTRG-REC _{sagS} -N671A		
54	pTRG-REC _{sagS} -P672A		
55	pTRG-REC _{sagS} -Q675A		
56	pTRG-REC _{sagS} -H683A		
57	pTRG-REC _{sagS} -W690A	Plasmid was made by site-directed mutagenesis with pTRG-REC _{sagS} as template	
58	pTRG-REC _{sagS} -E693A		
59	pTRG-REC _{sagS} -D713A		
60	pTRG-REC _{sagS} -N715A		
61	pTRG-REC _{sagS} -P717A		
62	pTRG-REC _{sagS} -R733A		
63	pTRG-REC _{sagS} -T743A		

64	pTRG-REC _{sagS} -R755A		
65	pTRG-REC _{sagS} -Y762A	Plasmid was made by site-directed mutagenesis with pTRG-REC _{sagS} as template	This study
66	pTRG-REC _{sagS} -P766A		
67	pTRG-REC _{sagS} -H768A		
68	pTRG-REC _{sagS} -D770A		
69	pTRG-REC _{sagS} -R778A		
70	pTRG-REC _{sagS} -E704AH705A	Plasmid was made by site-directed mutagenesis with pTRG-REC _{sagS} as template	This study
71	pET-28(b)	His -tag protein expression vector, Kan ^r	Addgene
72	pET28b-SagSL	A DNA fragment encoding residues 246-786aa of SagS into pET-28(b) at NdeI/XhoI ; plasmid was used for His ₆ -tagged SagS _{246-786aa} protein expression in <i>E.coli</i>	This study
73	pET28b-REC _{SagS}	A DNA fragment encoding residues 663-786aa of SagS cloned into pET-28(b) at NdeI/ XhoI site ; plasmid was used for His ₆ -tagged protein expression in <i>E.coli</i>	This study
74	pET28b-HptB	A DNA fragment encoding full length of PA3345 cloned into pET-28(b) at NdeI/XhoI site ; plasmid was used for His ₆ -tagged PA3345 protein expression in <i>E.coli</i>	This study
75	pET28b-HapZ	A DNA fragment encoding full length of PA2799 cloned into pET-28(b) at NdeI/XhoI site; plasmid was used for His ₆ -tagged PA2799 protein expression in <i>E.coli</i>	This study
76	pET28b-HapZ-R12AR16A	The plasmid was made by site-directed mutagenesis of pET28b-HapZ	This study
77	pGEX-4T1-HapZ	A DNA fragment encoding full length of PA2799 cloned into pGEX-4T1 at BamHI/XhoI site; plasmid was used for N-terminal GST-tagged PA2799 protein expression in <i>E.coli</i>	This study (Vector from GE healthcare)
78	pOPTHM_His ₆ MBP-HapZ	A DNA fragment encoding full length of PA2799 cloned into pOPTHM_His ₆ MBP at XbaI/ EcoRI site; plasmid was used for N-terminal His ₆ MBP-tagged PA2799 protein expression in <i>E.coli</i>	This study (Vector provided by Prof. Dr. Gao Yonggui's lab)
79	pTRX-His ₆ -HapZ	A DNA fragment encoding full length of PA2799 cloned into pTRX using LIC (ligase independent cloning) by HpaI vector digestion; plasmid was used for C-terminal	This study (Vector from Addgene)

His ₆ TRX-tagged PA2799 protein expression in <i>E.coli</i>			
80	pET28b-poly _{Lys} -HapZ	A DNA fragment encoding full length of PA2799 cloned into pET28b at NdeI/XhoI site with an N-terminal His ₆ tag and 6 lysine residues were added at the C-terminus of HapZ sequence using specific primers; plasmid was used for poly _{Lys} His ₆ -tagged PA2799 protein expression in <i>E.coli</i>	This study
81	pSUMO-poly _{Lys} -HapZ	A DNA fragment encoding full length of PA2799 cloned into pSUMO at BsaI/XhoI site with an N-terminal SUMO tag and 6 lysine residues were added at the N-terminus of HapZ using specific primers; plasmid was used for poly _{Lys} SUMO-tagged PA2799 protein expression in <i>E.coli</i>	This study (Vector provided by Prof. Dr. Yoon Ho Sup's lab)
82	pUCP18	<i>E.coli</i> - <i>P.aeruginosa</i> shuttle expression vector with P _{lac} , Amp ^r , Car ^r	Addgene
83	pUCP18-HapZ	A DNA fragment encoding full length of PA2799 (1-99aa/end) was cloned into pUCP18 at EcoRI/BamHI site under control of P _{lac}	This study
84	pUCP18-HapZ-ΔN21	A DNA fragment encoding truncated PA2799 (22-99aa/end) was cloned into pUCP18 at EcoRI/BamHI site under control of P _{lac} .	This study
85	pUCP18-HapZ-Y8A		
86	pUCP18-HapZ-E10A		
87	pUCP18-HapZ-K11A		
88	pUCP18-HapZ-R12AR16A		
89	pUCP18-HapZ-F14A	The plasmid was made by site-directed mutagenesis of pUCP18-HapZ	This study
90	pUCP18-HapZ-L19A		
91	pUCP18-HapZ-D38AS40A		
92	pUCP18-HapZ-R80A		
93	pUCP18-HapZ-G92E		

94	pUCP18-6xHis-REC _{SagS}	A DNA fragment (containing a Shine-Dalgarno sequence) encoding REC domain (663-786aa) of SagS with 6xHis tag at the N-terminal under control of P _{lac} promoter was cloned at EcoRI/BamHI site of pUCP18; This plasmid was used for 6xHis-REC _{SagS} protein expression (as a control) in <i>P. aeruginosa</i> .	This study
95	pUCP-HA-HapZ/6xHis-REC _{SagS}	6xHis tag at the N-terminal following P _{lac} promoter and a Shine-Dalgarno sequence was fused with a translational enhancer, a Shine-Dalgarno sequence upstream of REC _{SagS} and a HA tag (YPYDVPDYA) fused to N-terminal of HapZ. This fragment was cloned at EcoRI/BamHI site of pUCP18. This plasmid was used for 6x His-REC _{SagS} and HA-HapZ co-expression in <i>P. aeruginosa</i> . This was synthesized in pUC57 by Genescript and sub-cloned into pUCP18	Genescript, USA
96	pUCP-P _{HapZ} -gfp	A DNA fragment covering the upstream region of PA2799 (3156257-3156502), <i>P. aeruginosa</i> mPAO1, complete genome) was cloned between the EcoRI and BamHI sites of pMH305 [Ref (b) (c)].	This study

^a Cmr^r, chloramphenicol resistance

Amp^r, ampicillin resistance

Tetr^r, tetracycline resistance

Kan^r, kanamycin resistance

Car^r, carbenicillin resistance.

^b aa, amino acids

^cThe antibiotic concentrations used were as follows for *E. coli*: ampicillin at 200 µg/ml, kanamycin at 50 µg/ml, chloramphenicol at 12.5µg/ml, tetracycline at 25 µg/ml, and streptomycin at 300 µg/ml. and for *P. aeruginosa*, carbenicillin at 200 µg/ml, tetracycline at 100 µg/ml.

Ref (a) Jacobs *et al*, Comprehensive transposon mutant library of *Pseudomonas aeruginosa*. *Proc. Natl. Acad. Sci. USA*. 2003, 100:14339-44.

Ref (b) Bordi *et al*, Regulatory RNAs and the HptB/RetS signalling pathways fine-tune *Pseudomonas aeruginosa* pathogenesis. *Mol Microbiol*. 2010; 76: 1427–1443.

Ref (c) Christophe *et al*, Fluorescence-based reporter for gauging cyclic di-GMP levels in *Pseudomonas aeruginosa*. *Appl Environ Microbiol*. 2012, 78:5060-9.

Ref(d) An S, *et al* (2010) Modulation of *Pseudomonas aeruginosa* Biofilm Dispersal by a Cyclic-Di-GMP Phosphodiesterase with a Putative Hypoxia-Sensing Domain. *App. Environ. Microbiol*. 76(24):8160-8173.

Table S3. List of primers used in the study

S.No	Primer	Sequence	Application
1	pTRG-F(A)	AAACCAGAGGCGGCC <u>A</u> GGATCCGCGGCCGCAA GAATTCAGTC	For insertion mutation of pTRGA
2	pTRG-R(T)	TGCCGGCGCGGATCC <u>T</u> GGCCGCCTCTGGTTTCTC TTCTTTC	For insertion mutation of pTRGA
3	pTRG-F(AC)	AAACCAGAGGCGGCC <u>A</u> CGGATCCGCGGCCGCA AGAATTCAGTC	For insertion mutation of pTRGAC
4	pTRG-R(GT)	TGCCGGCGCGGATCC <u>G</u> TGGCCGCCTCTGGTTTCT CTTCTTTC	For insertion mutation of pTRGAC
5	PA2799EcoR-F	CCG <u>G</u> AATTCATGCAGCCCATCGAGCACAAC	For construction of pBT-PA2799
6	PA2799BamH-R	CGC <u>G</u> GATCCCTCAGTGGATCGCGACGATGCTC	For construction of pBT-PA2799
7	SagS-663-F	AAACCAGAGGCGGCC <u>G</u> GATCCACCCGGGTCCTG CTGGTGGAGGACAC	For construction of pTRG-REC _{SagS}
8	SagS-786R	ATTAATTAATTAATTACTCGAGCTAGTCGCTCGC GGTGAGCGGGCAC	For construction of pTRG-REC _{SagS}
9	pTRG1976- 757Bam-F	AAACCAGAGGCGGCC <u>G</u> GATCCCTCGCGGGTGTGG GTGCTGGAC	For construction of pTRG-REC _{PA1976}
10	pTRG1976- 881Xho-R	ATTAATTAATTAATTACTCGAGTCAGACCGAGG CTTCGCGCTC	For construction of pTRG-REC _{PA1976}
11	pTRG1611- 515Bam-F	AAACCAGAGGCGGCC <u>G</u> GATCCCAGGAAATCCTC CTGGTCGAGGAC	For construction of pTRG-REC _{PA1611}
12	pTRG1611- 651Xho-R	ATTAATTAATTAATTACTCGAGTCATTCCGGCTC CCCTCGTCCGG	For construction of pTRG-REC _{PA1611}
13	pTRGRetS- 808Bam-F	AAACCAGAGGCGGCC <u>G</u> GATCCCTCCGGATCCTC GTCGCCGAGGAC	For construction of pTRG-REC ₄₈₅₆
14	pTRGRetS- 942Xho-R	ATTAATTAATTAATTACTCGAGTCAGGAGGGCA GGGCGTCGCCCTG	For construction of pTRG-REC ₄₈₅₆
15	pTRG-F	TGGCTGAACAACTGGAAGCT	For PCR amplification of inserts in pTRG- derived plasmids
16	pTRG-R	ATTCGTCGCCCCGCCATAA	For PCR amplification of inserts in pTRG- derived plasmids
17	HapZ-Y8A-F	ATCGAGCACAAACG <u>C</u> AGCGAGAAGCGC GAC	For construction of pBT-HapZ-Y8A and pUCP18-HapZ-Y8A plasmid
18	HapZ-Y8A-R	CGCTTCTCGCTG <u>G</u> CGTTGTGCTCGATGGGCTG	For construction of pBT-HapZ-Y8A and pUCP18-HapZ-Y8A plasmid

19	HapZ-E10A-F	CACAACTACAGCG <u>GCG</u> AAGCGCGACTTC ATC	For construction of pBT-HapZ-E10A and pUCP18-HapZ-E10A plasmid
20	HapZ-E10A-R	GAAGTCGCGCTTC <u>GCG</u> GCTGTAGTTGTGCTCGAT	For construction of pBT-HapZ-E10A and pUCP18-HapZ-E10A plasmid
21	HapZ-K11A-F	AACTACAGCGAG <u>GCG</u> GCGCGACTTCATC CGC	For construction of pBT-HapZ-K11A and pUCP18-HapZ-K11A plasmid
22	HapZ-K11A-R	GATGAAGTCGCG <u>GCG</u> CCTCGCTGTAGTTGTGCTC	For construction of pBT-HapZ-K11A and pUCP18-HapZ-K11A plasmid
23	HapZ-R12A-F	CAACTACAGCGAGAAG <u>GCC</u> GACTTCATCCGCAT GAAGCTC GAC	For construction of pBT-HapZ-R12A plasmid
24	HapZ-R12A-R	CATGCGGATGAAGTC <u>GGC</u> CTTCTCGCTGTAGTTG TGCTCGATG	For construction of pBT-HapZ-R12A plasmid
25	HapZ-F14A-F	GAGAAGCGCGAC <u>GCC</u> ATCCGCATGAAG CTCG	For construction of pBT-HapZ-F14A and pUCP18-HapZ-F14A plasmid
26	HapZ-F14A-R	CTTCATGCGGAT <u>GCG</u> GTCGCGCTTCTC GCTGTA	For construction of pBT-HapZ-F14A and pUCP18-HapZ-F14A plasmid
27	HapZ-R16A-F	GCCGACTTCATC <u>GCC</u> ATGAAGCTCGAC	For construction of pBT-HapZ-R16A plasmid
28	HapZ-R16A-R	GGCGATGAAGTC <u>GGC</u> CTTCTCGCTGTA GTTGTG	For construction of pBT-HapZ-R16A plasmid
29	HapZ-R12AR16A-F	CTACAGCGAGAAG <u>GCC</u> GACTTCATC <u>GCC</u> ATGAAGCTCGAC	For construction of pBT-HapZ-R12AR16A and pUCP18-R12AR16A plasmid
30	HapZ-R12AR16A-R	GAGCTTCATG <u>GCG</u> GATGAAGTC <u>GGC</u> CTT CTCGCTGTAGTTGTG	For construction of pBT-HapZ-R12AR16A and pUCP18-R12AR16A plasmid
31	HapZ-L19A-F	CATCCGCATGAAG <u>GCC</u> GACACCCCG GTCACC	For construction of pBT-HapZ-L19A and pUCP18-HapZ-L19A plasmid
32	HapZ-L19A-R	CCGGGGTGTC <u>GCG</u> CTTCATGCGGAT GAAGTCG	For construction of pBT-HapZ-L19A and pUCP18-HapZ-L19A plasmid
33	HapZ-E31A-F	CTGCGAAGGCCGCG <u>GCG</u> TCGGTGGCC CTG	For construction of pBT-HapZ-E31A plasmid
34	HapZ-E31A-R	GGCCACCGAC <u>GCG</u> GCGGCCTTCGCAGCG	For construction of pBT-HapZ-E31A

			plasmid
35	HapZ-L35A-F	GAGTCGGTGGCC <u>CG</u> GTGCCGGGACCTG TCC	For construction of pBT-HapZ-L35A plasmid
36	HapZ-L35A-R	CAGGTCCCGGCAC <u>CG</u> CGGCCACCGACTC GC	For construction of pBT-HapZ-L35A plasmid
37	HapZ-L45A-F	CCAGCGGGCTG <u>GCG</u> CTGGAAGCGCAA AGC	For construction of pBT-HapZ-L45A plasmid
38	HapZ-L45A-R	GCGCTTCCAG <u>CG</u> CCAGCCCGCTGGC GGAC	For construction of pBT-HapZ-L45A plasmid
39	HapZ-Q49A-F	CTGCTGGAAGCG <u>GCA</u> AGCGAACTGGCT CCGGG	For construction of pBT-HapZ-Q49A plasmid
40	HapZ-Q49A-R	GCCAGTTCGCT <u>TGCC</u> GCTTCCAGCAG CAGCCC	For construction of pBT-HapZ-Q49A plasmid
41	HapZ-D56A-F	CTGGCTCCGGGCG <u>CCC</u> CACCTGCGGGTG AGC	For construction of pBT-HapZ-D56A plasmid
42	HapZ-D56A-R	CCCGCAGGTG <u>GGC</u> GCCCGGAGCCAG TTCGC	For construction of pBT-HapZ-D56A plasmid
43	HapZ-R59A-F	GGCGACCACCTG <u>GCG</u> GTGAGCATCCCG TCC	For construction of pBT-HapZ-R59A plasmid
44	HapZ-R59A-R	CGGGATGCTCAC <u>CG</u> CCAGGTGGTCGCC CGGAG	For construction of pBT-HapZ-R59A plasmid
45	HapZ-E67A-F	CCGTCCACCCAC <u>GCG</u> ACGCTCACCGGC CTG	For construction of pBT-HapZ-E67A plasmid
46	HapZ-E67A-R	GGTGAGCGT <u>CGC</u> GTGGGTGGACGGGAT GCTC	For construction of pBT-HapZ-E67A plasmid
47	HapZ-E73A-F	CACCGGCCTG <u>GCA</u> CTGGAAGCCCGCGTG	For construction of pBT-HapZ-E73A plasmid
48	HapZ-E73A-R	GGCTTCCAGT <u>GCC</u> CAGGCCGGTGAGCGT	For construction of pBT-HapZ-E73A plasmid
49	HapZ-R77A-F	GAACTGGAAGCC <u>CC</u> GTGGTGCGCTGC GAAC	For construction of pBT-HapZ-R77A plasmid
50	HapZ-R77A-R	GCAGCGCACCAC <u>GCG</u> GGCTTCCAGTTC CAG	For construction of pBT-HapZ-R77A plasmid
51	HapZ-V79A-F	GAAGCCCGCGTGG <u>CG</u> GCCTGCGAACGC CTCGGCC	For construction of pBT-HapZ-V79A plasmid
52	HapZ-V79A-R	GAGGCGTTTCGCAGG <u>CGC</u> CACGCGGGC TTCCAGTTC	For construction of pBT-HapZ-V79A plasmid
53	HapZ-R80A-F	CCGCGTGGTGG <u>CCT</u> GCGAACGCCTC	For construction of pBT-HapZ-R80A and pUCP18-HapZ-

			R80A plasmid
54	HapZ-R80A-R	GCGTTCGCAG <u>GGCC</u> ACCACGCGGGCTTC	For construction of pBT-HapZ-R80A and pUCP18-HapZ-R80A plasmid
55	HapZ-G92E-F	CGGCATAGCCTGGAGCTGAGCATCGTC GCG	For construction of pBT-HapZ-G92E and pUCP18-HapZ-G92E plasmid
56	HapZ-G92E-F	CGATGCTCAG <u>CTCC</u> AGGCTATGCCGGCC GAG	For construction of pBT-HapZ-G92E and pUCP18-HapZ-G92E plasmid
57	HapZ-T24Stop-F	GACACCCCGGTCTAGCATCCGCTGCGA AGG	For construction of pBT-HapZ-T24Stop plasmid
58	HapZ-T24Stop-R	GCAGCGGATGCTAGACCGGGGTGTCGAGC	For construction of pBT-HapZ-T24Stop plasmid
59	HapZ-D38AS40A-F	CTGTGCCGCG <u>CCC</u> CTGGCCGCCAGCGGG CTGCTG	For construction of pBT-HapZ-D38AS40A and pUCP18-HapZ-D38AS40A plasmid
60	HapZ-D38AS40A-R	CCCGCTGGCG <u>GGC</u> CAGGGCGCGGCACAG GGCCAC	For construction of pBT-HapZ-D38AS40A and pUCP18-HapZ-D38AS40A plasmid
61	HapZ-E31R-F	CTGCGAAGGCCGCGGTCGGTGGCCCTG	For construction of pBT-HapZ-E31RE73A plasmid
62	HapZ-E31R-R	GGCCACCGA <u>CCGG</u> CGGCCTTCGCAGCG	For construction of pBT-HapZ-E31RE73A plasmid
63	HapZ-E67R-F	CCGTCCACCCAC <u>CGG</u> ACGCTCACCGGCCTG	For construction of pBT-HapZ-E31AE67R plasmid
64	HapZ-E67R-R	GGTGAGCGT <u>CCGG</u> TGGGTGGACGGGATGCTC	For construction of pBT-HapZ-E31AE67R plasmid
65	HapZ-E73R-F	CACCGGCCTG <u>CGA</u> CTGGAAGCCCGCGTG	For construction of pBT-HapZ-E31AE73R plasmid
66	HapZ-E73R-R	GGCTTCCAGT <u>TCG</u> CAGGCCGGTGAGCGT	For construction of pBT-HapZ-E31AE73R plasmid
67	HapZ-EcoR-Δ21aa-F	CCGGAATTCCGACACCCCGGTCACCATCCGC	For construction of pBT-HapZ-Δ21aa and pUCP18-HapZ-Δ21aa plasmid
68	HapZ-EcoR-Δ21aa-R	CGCGGATCCTCAGTGGATCGCGACGATGCTC	For construction of pBT-HapZ-Δ21aa and pUCP18-HapZ-Δ21aa plasmid
69	REC _{SagS} -N671A-F	GGTGGAGGACG <u>CCCC</u> GGTCAACCAGTTGGTC	For construction of pTRG-REC _{SagS} -N671A plasmid
70	REC _{SagS} -N671A-R	CTGGTTGACCGGGGCGTCCTCCACCAGCAG	For construction of

			pTRG- REC _{SagS} - N671A plasmid
71	REC _{SagS} -P672A-F	GGAGGACAAC <u>GCGG</u> TCAACCAGTTGGTCGCC	For construction of pTRG- REC _{SagS} - P672A plasmid
72	REC _{SagS} -P672A-R	CAACTGGTTGAC <u>GCGG</u> TGTCCTCCACCAGC	For construction of pTRG- REC _{SagS} - P672A plasmid
73	REC _{SagS} -Q675A-F	CCCGGTCAAC <u>GCGG</u> TGGTCGCCAAGGGCCTG	For construction of pTRG- REC _{SagS} - Q675A plasmid
74	REC _{SagS} -Q675A-R	CTTGCGGACCAAC <u>GCGG</u> TGACCGGGTTGTCC	For construction of pTRG- REC _{SagS} - Q675A plasmid
75	REC _{SagS} -H683A-F	GGGCCTGCTG <u>GCCA</u> AAGCTCGGCTGCCAG	For construction of pTRG- REC _{SagS} - H683A plasmid
76	REC _{SagS} -H683A-R	GCAGCCGAGCTT <u>GGCC</u> CAGCAGGCCCTTGG	For construction of pTRG- REC _{SagS} - H683A plasmid
77	REC _{SagS} -W690A-F	GCTGCCAGGTAG <u>GCGG</u> ATCGCCGAACACGGGC	For construction of pTRG- REC _{SagS} - W690A plasmid
78	REC _{SagS} -W690A-R	GTGTTCGGCGAT <u>GCGG</u> TACCTGGCAGCCGAGC	For construction of pTRG- REC _{SagS} - W690A plasmid
79	REC _{SagS} -E693A-F	GTATGGATCGCC <u>GCGG</u> CACAGGGCTGAATGCCCTG	For construction of pTRG- REC _{SagS} - E693A plasmid
80	REC _{SagS} -E693A-R	CATTCAGCCCGTG <u>GCGG</u> GCGATCCATACCTGGC	For construction of pTRG- REC _{SagS} - E693A plasmid
81	REC _{SagS} - E704AH705A-F	GATGCTGGAG <u>GCGG</u> CTCCCATCGACCTGG	For construction of pTRG- REC _{SagS} - E704AH705A plasmid
82	REC _{SagS} - E704AH705A-R	GGTCGATGGG <u>GCGG</u> CCTCCAGCATCTTCAG	For construction of pTRG- REC _{SagS} - E704AH705A plasmid
83	REC _{SagS} -D713A-F	GGTCCTGATG <u>GCGG</u> CTGCAACATGCCAGTGATG	For construction of pTRG- REC _{SagS} - D713A plasmid
84	REC _{SagS} -D713A-R	CTGGCATGTTGCAG <u>GCGG</u> CATCAGGACCAGGTC	For construction of pTRG- REC _{SagS} - D713A plasmid
85	REC _{SagS} -N715A-F	GATGGACTGCG <u>GCGG</u> CATGCCAGTGATGGACGG	For construction of pTRG- REC _{SagS} - N715A plasmid
86	REC _{SagS} -N715A-R	CATCACTGGCATG <u>GCGG</u> CAGTCCATCAGGAC	For construction of pTRG- REC _{SagS} - N715A plasmid
87	REC _{SagS} -P717A-F	CTGCAACATG <u>GCGG</u> AGTGATGGACGGCTACGAAG	For construction of pTRG- REC _{SagS} - P717A plasmid
88	REC _{SagS} -P717A-R	GGCGACCATCACT <u>GCGG</u> CATGTTGCAGTCCATC	For construction of pTRG- REC _{SagS} - P717A plasmid
89	REC _{SagS} -R733A-F	CGACAGCGGAG <u>GCGG</u> CCTGGGGCGGCCTGCCG	For construction of

			pTRG- REC _{SagS} - R733A plasmid
90	REC _{SagS} -R733A-R	GGCCGCCCCAGGCTCCGCTGTGCGGG	For construction of pTRG- REC _{SagS} - R733A plasmid
91	REC _{SagS} -T743A-F	CATCGCGCTGGCCGCCAACGCCCTGCCGG	For construction of pTRG- REC _{SagS} - T743A plasmid
92	REC _{SagS} -T743A-R	CAGGGCGTTGGCGGCCAGCGCGATGATCG	For construction of pTRG- REC _{SagS} - T743A plasmid
93	REC _{SagS} -R755A-F	CGAACGTTGTGCTGCCGCCGGCATGGACG	For construction of pTRG- REC _{SagS} - R755A plasmid
94	REC _{SagS} -R755A-R	CATGCCGGCGGCAGCACAACTTCGCGCTCG	For construction of pTRG- REC _{SagS} - R755A plasmid
95	REC _{SagS} -Y762A-F	CATGGACGACGCCCTGGCCAAGCCTTTCC	For construction of pTRG- REC _{SagS} - Y762A plasmid
96	REC _{SagS} -Y762A-R	GGCTTGGCCAGGGCGTCGTCCATGCCGG	For construction of pTRG- REC _{SagS} - Y762A plasmid
97	REC _{SagS} -P766A-F	CCTGGCCAAGGCTTTCCACCGCGATGAACTG	For construction of pTRG- REC _{SagS} - P766A plasmid
98	REC _{SagS} -P766A-R	CATCGCGGTGGAAAGCCTTGGCCAGGTAGTC	For construction of pTRG- REC _{SagS} - P766A plasmid
99	REC _{SagS} -H768A-F	CCAAGCCTTTCCGCCCGCGATGAACTGAAAGCC	For construction of pTRG- REC _{SagS} - H768A plasmid
100	REC _{SagS} -H768A-R	CAGTTCATCGCGGGCGAAAGGCTTGGCCAGG	For construction of pTRG- REC _{SagS} - H768A plasmid
101	REC _{SagS} -D770A-F	CTTCCACCGCGCTGAACTGAAAGCCATCC	For construction of pTRG- REC _{SagS} - D770A plasmid
102	REC _{SagS} -D770A-R	GGCTTTCAGTTCAGCGCGGTGGAAAGGCTTGG	For construction of pTRG- REC _{SagS} - D770A plasmid
103	REC _{SagS} -R778A-F	GCCATCCTCGACGCTTGGTGCCCGCTCACC	For construction of pTRG- REC _{SagS} - R778A plasmid
104	REC _{SagS} -R778A-R	GCGGGCACCAAGCGTCGAGGATGGCTTTC	For construction of pTRG- REC _{SagS} - R778A plasmid
105	pET-HapZ-Nde-F	CTGGTGCCGCGCGGCAGCCATATGCAGCCCATC GAGCACAAC	For construction of pET28b-HapZ plasmid
106	pET-HapZ-XhoI-R	GTGGTGGTGGTGGTGTCTGAGTCAGTGGATCGC GACGATGCTC	For construction of pET28b-HapZ and pGEX-4T1-HapZ plasmid
107	pUCP-HapZ- EcoR-F	CCGGAATTCCATGCAGCCCATCGAGCACAAC	For construction of pUCP18-HapZ plasmid
108	pUCP-HapZ- BamH-R	CGCGGATCCTCAGTGGATCGCGACGATGCTC	For construction of pUCP18-HapZ

			plasmid
109	pET-248SagS-F	CTGGTGCCGCGCGGCAGCC <u>CATATGGT</u> AGAGATC GAACAGCGCCGGGAGG	For construction of pET28b-SagSL plasmid
110	pET-663SagS-F	AAACCAGAGGCGGCCG <u>GATCC</u> ACCCGGGTCTG CTGGTGGAGGACAC	For construction of pET28b-REC _{SagS} plasmid
111	pET-786SagS-F	GTGGTGGTGGTGGT <u>GCTCGAG</u> CTAGTCGCTCGCG GTGAGCGGGCAC	For construction of pET28b-SagSL and pET28b-REC _{SagS} plasmid
112	pET-HptB-Nde-F	CTGGTGCCGCGCGGCAGCC <u>CATATG</u> TCCGCGCCGC ATCTCGATG	For construction of pET28b-HptB plasmid
113	pET-HptB-Xho-F	GTGGTGGTGGTGGT <u>GCTCGAG</u> TCAGCGATAGCG CTGACGTTCC	For construction of pET28b-HptB plasmid
114	HapZ-MBP-EcoR-F	CCGGA <u>ATTCC</u> ATGCAGCCCATCGAGCACAAC	For construction of pOPTHM_His ₆ - MBP-HapZ plasmid
115	HapZ-MBP-Xba-R	CGC <u>TCTAGAT</u> CAGTGGATCGCGACGATGCTC	For construction of pOPTHM_His ₆ - MBP-HapZ plasmid
116	HapZ-GST-Bam-F	CCGGG <u>ATTCC</u> ATGCAGCCCATCGAGCACAAC	For construction of pGEX-4T1-HapZ plasmid
117	pTRX-HapZ-LICv3-F	TTTAAGAAGGAGATATAGTTCATGCAGCCCATCG AGCACAAC	For construction of pTRX-His ₆ -HapZ plasmid
118	pTRX-HapZ-LICv3-R	GGATTGGAAGTAGAGGTTCTCGTGGATCGCGAC GATGCTCAG	For construction of pTRX-His ₆ -HapZ plasmid
119	poly _{Lys} -HapZ-NdeI-F	CCGCATATGATGCAGCCGATTGAACA	For construction of pET28b-poly _{Lys} - HapZ plasmid
120	poly _{Lys} -HapZ-XhoI-R	CGCCTCGAGTTATTTTTTTTTTTTTTTAGAGGTTC T	For construction of pET28b-poly _{Lys} - HapZ plasmid
121	pSUMO-poly _{Lys} -HapZ-BsaI-F	GCGGTCTCTAGGTAAAAAATAATGCAGC CCATC	For construction of pSUMO-poly _{Lys} - HapZ plasmid
122	pSUMO-poly _{Lys} -HapZ-XhoI-R	CGCCTCGAGTTAATGGATGGCGACAAC	For construction of pSUMO-poly _{Lys} - HapZ plasmid

Figure S3 | Expression of HapZ using various solubility tags. HapZ formed insoluble aggregates when expressed along with an N-terminal His₆ tag. Five other tags namely His₆-MBP (MBP: Maltose binding protein), Trx-His₆ (Trx: Thioredoxin), SUMO-poly_{Lys} (SUMO: Small Ubiquitin-like Modifier, poly_{Lys}: 6 lysine residues), GST (GST: Glutathione S-transferase) and His₆-poly_{Lys} were tested to assess whether HapZ is stable with any of these tags. However, with the exception of GST, none of the other tags could stabilize HapZ. Unfortunately, GST-HapZ was unable to bind to REC_{SagS} when tested in the phosphorelay experiment due to which it could not be used for *in vitro* studies.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
SagS₂₄₆₋₇₈₆	+	-	+	+	+	+	+	+	+	+	+	+	+	+
HptB	-	+	+	+	+	+	+	+	+	+	+	+	+	+
HapZ	-	-	+	+	+	+	+	+	+	+	+	-	-	-
HapZR12A/R16A	-	-	-	-	-	-	-	-	-	-	-	+	+	+
C-di-GMP (μM)	-	-	-	2	5	20	50	-	-	-	-	-	10	50
GMP (μM)	-	-	-	-	-	-	-	2	5	20	50	-	-	-

Figure S4 | Loading control for the *in vitro* phosphorelay experiment (Chapter 2). SDS-PAGE gel of the different reaction mixtures constituted for the phosphorelay experiment. The proteins SagS₂₄₆₋₇₈₆, HptB, HapZ or HapZR12AR16A were loaded in 1:1:1 ratio in the various reactions. Different concentrations of the two nucleotides c-di-GMP and GMP were also added to the reaction mixtures. Aliquots of these reaction mixtures were loaded on the 15% SDS-PAGE gel as a loading control before the addition of [γ -³²P]-ATP.

Table S4. Genes that are UP regulated in both *ΔsagS* and *ΔhapZ* mutants.

Gene	Locus_Tag	Fold Change for <i>ΔhapZ</i>	Padj value	Fold Change for <i>ΔsagS</i>	Padj value
<i>PA0122</i>	PA0122	12.60563	3.38E-203	15.36549	1.33E-228
<i>PA0123</i>	PA0123	3.941959	5.19E-42	5.504636	5.20E-65
<i>pqsA</i>	PA0996	7.888052	9.37E-63	7.543361	3.13E-59
<i>pqsB</i>	PA0997	11.25075	1.18E-63	9.53725	2.58E-53
<i>pqsC</i>	PA0998	4.118644	1.40E-28	3.338863	9.70E-20
<i>pqsD</i>	PA0999	6.783603	1.69E-63	6.17092	2.54E-57
<i>pqsE</i>	PA1000	10.66266	2.95E-71	8.162358	7.91E-54
<i>phnA</i>	PA1001	4.09217	4.49E-24	3.333577	4.32E-17
<i>PA1221</i>	PA1221	4.042764	5.32E-84	3.195601	1.67E-57
<i>PA1869</i>	PA1869	3.841414	3.91E-14	4.292616	5.16E-18
<i>PA2069</i>	PA2069	4.772683	5.08E-31	6.416301	2.75E-47
<i>hcnA</i>	PA2193	4.257261	0.001483	4.226306	0.001476
<i>PA2274</i>	PA2274	3.352512	4.61E-07	8.323989	1.05E-23
<i>PA2563</i>	PA2563	2.651982	7.11E-29	5.408048	2.44E-87
<i>PA2564</i>	PA2564	2.720697	7.09E-41	6.24032	6.32E-140
<i>PA2565</i>	PA2565	2.865173	3.64E-20	6.468926	2.91E-64
<i>PA2566</i>	PA2566	3.642291	1.20E-88	5.879509	6.38E-159
<i>lecA</i>	PA2570	46.21417	0	75.13671	0
<i>ppyR</i>	PA2663	3.165283	0.000899	2.510229	0.014769
<i>fhp</i>	PA2664	3.998109	8.96E-21	3.713203	1.97E-19
<i>cpg2</i>	PA2787	3.602466	8.34E-56	2.566022	1.08E-28
<i>PA2788</i>	PA2788	4.24936	1.41E-104	3.443247	2.76E-76
<i>lecB</i>	PA3361	17.34817	3.28E-167	19.58909	3.75E-177
<i>PA3362</i>	PA3362	2.991626	4.56E-20	3.254447	7.61E-24
<i>rhlB</i>	PA3478	4.397117	3.11E-42	5.701496	1.36E-62
<i>rhlA</i>	PA3479	9.457885	1.33E-103	13.75423	2.73E-148
<i>PA3520</i>	PA3520	3.926651	7.13E-13	6.153162	4.09E-25
<i>lasB</i>	PA3724	6.666506	7.99E-65	14.28647	1.80E-121
<i>recJ</i>	PA3725	2.720722	1.85E-22	3.300272	7.00E-34
<i>PA3734</i>	PA3734	2.887577	5.85E-30	4.562142	4.51E-66
<i>PA4078</i>	PA4078	3.011322	9.29E-15	3.188276	2.90E-16
<i>PA4141</i>	PA4141	6.695079	7.32E-116	8.402511	2.00E-145
<i>PA4142</i>	PA4142	2.577345	2.19E-08	2.914675	1.10E-11
<i>mexG</i>	PA4205	8.764616	1.09E-38	26.88714	1.05E-95
<i>mexH</i>	PA4206	7.113256	2.17E-16	25.70526	9.59E-44
<i>mexI</i>	PA4207	3.022729	4.41E-17	8.980392	2.77E-74
<i>opmD</i>	PA4208	3.210544	6.14E-07	8.50483	2.59E-28
<i>PA4384</i>	PA4384	2.907779	1.25E-08	3.461626	3.57E-13

PA4591	PA4591	2.858705	3.42E-05	3.785468	1.76E-09
PA5352	PA5352	3.877959	1.71E-07	26.75801	8.64E-51
glcF	PA5353	4.046208	0.000298	25.66649	1.88E-17
glcE	PA5354	3.277831	7.08E-06	20.82219	3.65E-31
glcD	PA5355	2.795769	0.045587	17.75682	1.03E-09
PA5460	PA5460	2.588757	5.24E-33	3.469486	1.20E-55

Table S5. Genes that are DOWN regulated in both *AsagS* and *ΔhapZ* mutants.

Gene	Locus_Tag	Fold Change for <i>ΔhapZ</i>	Padj value	Fold Change for <i>AsagS</i>	Padj value
PA1743	PA1743	-0.1202	1.47E-08	-0.17715	1.54E-07
PA1744	PA1744	-0.26458	4.30E-13	-0.20998	2.04E-19
PA1942	PA1942	-0.01542	9.82E-205	-0.01632	4.88E-205
PA1970	PA1970	-0.00525	1.57E-294	-0.00417	0
PA2485	PA2485	-0.36032	1.48E-13	-0.17334	7.75E-36
PA2486	PA2486	-0.18244	8.09E-20	-0.08112	4.64E-36
PA2490	PA2490	-0.03583	3.20E-17	-0.45834	0.006252
mexS	PA2491	-4.4E-05	0	-0.05382	0
mexE	PA2493	-0.00224	1.74E-119	-0.04144	3.04E-80
mexF	PA2494	-0.00048	3.15E-274	-0.18249	1.57E-92
oprN	PA2495	-0.00427	3.36E-94	-0.11344	1.45E-45
PA2758	PA2758	-0.20296	6.58E-60	-0.21589	6.78E-62
PA2759	PA2759	-0.00789	0	-0.0077	0
ibpA	PA3126	-0.45636	0.000127	-0.38204	2.22E-06
PA3229	PA3229	-0.00403	0	-0.00496	0
PA3230	PA3230	-0.01915	0	-0.0168	0
PA4354	PA4354	-0.23009	1.24E-24	-0.2349	1.16E-26
PA4622	PA4622	-0.28198	2.16E-12	-0.30073	8.28E-12
PA4623	PA4623	-0.00232	0	-0.0023	0
PA4881	PA4881	-0.00795	4.69E-76	-0.01038	2.61E-76

Table S6. Genes that are UP regulated in *ΔhapZ*, but not in *AsagS* mutant.

Gene	Locus_Tag	Fold Change	Padj value
PA0187	PA0187	2.506166	1.78E-15
PA0226	PA0226	2.837317	2.42E-07
PA0227	PA0227	3.055158	3.53E-09
pcaF	PA0228	2.792759	1.16E-11
PA0242	PA0242	5.288861	3.58E-45
PA0243	PA0243	3.603449	1.63E-22

<i>PA0244</i>	PA0244	5.789267	1.93E-40
<i>aroQ2</i>	PA0245	5.572544	1.64E-16
<i>PA2184</i>	PA2184	2.581332	6.38E-10
<i>PA2488</i>	PA2488	5.326211	3.86E-42
<i>PA3441</i>	PA3441	3.131677	0.000129
<i>PA3446</i>	PA3446	2.533547	0.000902
<i>PA3677</i>	PA3677	4.759271	1.11E-35
<i>PA3678</i>	PA3678	2.832901	4.37E-14
<i>PA3931</i>	PA3931	2.639322	0.024436
<i>PA4170</i>	PA4170	3.582022	4.38E-26
<i>csaA</i>	PA3221	3.359666	4.20E-08

Table S7. Genes that are DOWN regulated in *ΔhapZ*, but not in *ΔsagS*.

Gene	Locus_Tag	Fold Change	Padj value
<i>PA0720</i>	PA0720	-0.35701	9.89E-09
<i>PA0722</i>	PA0722	-0.33145	6.02E-23
<i>PA1332</i>	PA1332	-0.36425	1.78E-32
<i>PA1506</i>	PA1506	-0.3822	0.019803
<i>pcrH</i>	PA1707	-0.32303	0.014073
<i>PA1957</i>	PA1957	-0.22386	0.038307
<i>PA2489</i>	PA2489	-0.01774	3.29E-35
<i>mexT</i>	PA2492	-0.00693	5.40E-84
<i>PA2496</i>	PA2496	-0.03012	1.00E-20
<i>PA2497</i>	PA2497	-0.0084	2.02E-70
<i>PA2498</i>	PA2498	-0.02923	1.07E-21
<i>PA2499</i>	PA2499	-0.04802	5.47E-13
<i>ampDh3</i>	PA0807	-0.39319	2.20E-46

Table S8. Genes that are UP regulated in *ΔsagS*, but not in *ΔhapZ*.

Gene	Locus_Tag	Fold Change	Padj value
<i>PA0269</i>	PA0269	2.507724	4.11E-08
<i>cbpD</i>	PA0852	3.343366	4.22E-51
<i>PA1784</i>	PA1784	3.037464	2.31E-69
<i>metE</i>	PA1927	3.122865	5.81E-10
<i>PA2031</i>	PA2031	3.569073	1.68E-16
<i>PA2146</i>	PA2146	2.64476	6.61E-35
<i>catA</i>	PA2507	3.630083	7.37E-06
<i>catC</i>	PA2508	3.970969	1.31E-07
<i>catB</i>	PA2509	4.536784	1.37E-06

<i>catR</i>	PA2510	5.686352	2.47E-08
<i>PA2511</i>	PA2511	5.799262	1.68E-22
<i>antA</i>	PA2512	17.98502	0.000136
<i>antB</i>	PA2513	2.779289	0.002573
<i>antC</i>	PA2514	2.626344	0.000403
<i>xylL</i>	PA2515	15.91726	0.000325
<i>xylZ</i>	PA2516	3.583439	0.005172
<i>PA2682</i>	PA2682	4.919637	2.16E-13
<i>snr1</i>	PA3032	2.611049	1.54E-34
<i>fdx2</i>	PA3809	3.559731	8.80E-18
<i>hscB</i>	PA3811	2.552585	2.78E-26
<i>iscA</i>	PA3812	3.6551	4.11E-42
<i>iscU</i>	PA3813	3.762651	5.41E-68
<i>PA4133</i>	PA4133	3.944041	1.10E-49
<i>PA4134</i>	PA4134	4.879722	6.36E-08
<i>PA4139</i>	PA4139	4.341379	3.39E-54
<i>rmlC</i>	PA5164	3.329034	1.62E-13
<i>rubA1</i>	PA5351	6.632369	6.80E-14
<i>lasA</i>	PA1871	3.113196	2.17E-39
<i>chiC</i>	PA2300	4.52418	1.05E-30
<i>PA1914</i>	PA1914	5.668759	1.73E-09
<i>PA2030</i>	PA2030	3.629538	2.34E-38
<i>PA2588</i>	PA2588	3.284435	2.49E-45
<i>rhII</i>	PA3476	3.278509	2.88E-52
<i>PA4128</i>	PA4128	2.57487	5.59E-09
<i>iscS</i>	PA3814	3.911427	3.25E-48
<i>iscR</i>	PA3815	3.853353	8.69E-67
<i>rmlD</i>	PA5162	2.880304	3.68E-13

Table S9. Genes that are DOWN regulated in *AsagS*, but not in *ΔhapZ*.

Gene	Locus_Tag	Fold Change	Padj value
<i>PA0433</i>	<i>PA0433</i>	-0.22878	2.12E-13
<i>exbD2</i>	<i>PA0694</i>	-0.11718	0.026334
<i>fahA</i>	<i>PA2008</i>	-0.39336	8.83E-23
<i>PA3278</i>	PA3278	-0.36026	1.32E-16

Table S10. Chemical shift table (ppm) for the residues of REC_{SagS}

Residue	Ca	Ha	Cβ	Hβ	Others
Thr663	64.354	4.011	69.356	4.107	C 173.515; Cγ2 23.244; H 8.066; N 121.459; Hγ2 1.151
Arg664	55.999	4.800	30.024		Hβ2 1.925; Hβ3 1.841; Hδ2 3.270; Hδ3 3.100; Hγ2 1.883 Hγ3 1.836; N 129.879 Cδ 42.835; Cγ 26.129; H10.032
Val 665	61.018	4.866	35.304	1.851	C 173.771; Cγ1 22.565; Cγ2 20.995; H 9.316; N 126.677; Hγ1 0.829; Hγ2 0.669
Leu666	52.762	5.079	44.47		C 172.655; Cδ1 26.899; Cδ2 23.917; H 8.433; Hβ2 1.784; Hβ3 1.043; Hγ 1.195; Hδ1 0.551; Hδ2 0.623; N 127.542
Leu667	53.366	4.684	44.133		C 173.047; Cδ1 27.758; Cγ 27.674; H 9.275; Hβ2 1.793; Hβ3 0.756; Hγ 1.356; Hδ 0.676; N 131.188
Val668	59.757	4.854	31.155	2.265	C 173.609; Cγ1 22.503; Cγ2 21.521; H 9.106; Hγ1 0.494; Hγ2 0.543; N 129.819
Glu669	55.477	4.463	31.885		Cγ 33.482; Hβ 2.402; Hβ3 2.331; Hγ2 1.952; Hγ3 1.812
Asn671	51.172	4.969	39.712		Hβ2 2.959; Hβ3 2.752; Hδ21 6.978; Hδ22 6.945; Nδ2 114.225
Pro672	64.94	3.894	32.415		C 178.687; Cδ 51.392; Cγ 27.443; Hβ2 2.23; Hβ3 1.883; Hδ2 4.166; Hδ3 3.769; Hγ21.981; Hγ3 1.902
Val673	65.922	3.582	31.567	1.926	C 178.122; Cγ1 22.15; Cγ2 20.703; H 7.558; N 118.891; Hγ1 0.905; Hγ2 0.793
Asn674	55.053	4.283	37.47		C 178.295; Hβ 2.721; Hβ3 2.662
Gln675	59.118	3.54	28.654		C 177.085; Cγ 33.047; H 8.221; Hβ2 2.225; Hβ3 1.691; Hε21 7.786; Hε22 6.937; Hγ2 2.407 Hγ3 1.873; N 118.84; Nε2 109.233
Leu676	58.061	3.836	41.737		C 180.502; Cδ1 24.551; Cδ2 23.538; Cγ 26.774; H 7.214; Hβ2 1.74; Hβ3 1.569; Hγ 1.657; Hδ1 0.802; Hδ2 0.747; N 117.467
Val677	65.996	3.54	32.25	1.883	C 179.039; Cγ1 22.256; Cγ2 20.787; H 7.82; Hγ1 0.951; Hγ2 0.769; N 120.511
Ala678	55.476	3.706	18.567	1.159	C 178.861; H 7.969; N 120.719
Lys679	60.466	3.498	32.243		C 177.474; Cδ 29.718; Cε 41.646; Cγ 24.835; H 24.364; Hβ2 1.661; Hβ3 1.584; Hγ2 0.948; Hγ3 0.876; Hδ 1.398; Hε 2.671
Gly680	47.268				C 176.693; H 7.65; Hα2 3.795; Hα 3.765; N 104.139
Leu681	57.743	3.972	43.106		C 178.841; Cδ1 25.867; Cδ2 23.397; H 7.449; Hβ2 1.764; Hβ3 1.152; N 121.439; Hδ1 0.69; Hδ2 0.757
Leu682	58.088	3.701	42.091		C 179.585; Cδ1 25.99; Cδ2 25.111; 27.693; H 8.313; Hβ2 1.799; Hβ3 0.85; Hδ1 0.587; Hδ 0.786
His683	58.797	4.514	30.188		C 180.296; H 8.587; Hβ2 3.17; Hβ3 3.075; N 119.408
Lys684	58.873	4.037	32.005		C 178.014; Cε 42.014; Cγ 25.536; H 7.713; Hβ2 1.994; Hβ3 1.902; Hγ2 1.593; Hγ3 1.484; Hε 2.891; N 120.79
Leu685	54.506	4.286	41.968		C 176.302; Cδ1 26.198; Cδ2 22.811; Cγ 26.644; H 7.394; Hβ2 1.675; Hβ3 1.646; Hγ 1.684; Hδ1 0.687; Hδ2 0.772; N 118.308
Gly686	46.331				C 174.631; H 7.735; Hα2 4.062; Hα3 3.7; N 107.076
Cys687	58.017	4.888	29.487		C 174.727; H 7.625; Hβ2 2.543; Hβ3 2.249; N 115.443
Gln688	55.064	4.277	30.87		C 174.477; Cγ 34.057; H 8.702; Hβ2 1.799; Hβ3 1.715; Hε21 7.026; Hε22 6.717; Hγ2 2.018; Hγ3 1.809; N 122.822; Nε2 111.456

Val689	61.142	4.888	35.222	1.486	C 174.44; C γ 1 23.815; C γ 2 21.712; H 8.211; H γ 1 0.553; H γ 2 0.543; N 123.276
Trp690	57.845	4.432	32.361		C173.992; H 9.031; H β 2 2.828; H β 3 2.664; H δ 1 6.607; H ϵ 1 9.513; N 130.755; N ϵ 1 128.131
Ile691	59.656	5.074	40.089	1.526	C 174.776; C δ 1 13.299; C γ 1 28.118; C γ 2 17.166; H 8.517; H γ 12 1.334; H γ 13 0.815; H δ 1 0.585; H δ 2 0.552; N 121.554
Ala692	49.952	4.58	22.8	1.169	C 175.563; N 127.753; H 8.925
Glu693	56.631	4.267	30.015	1.825	C 174.228; C γ 36.356; H 9.239; H γ 1.999; N 118.334
His694	54.965	5.478	32.376		C 177.607; H 6.949; H β 2 3.595; H β 3 3.317; N 107.367
Gly695	48.008				C 173.54; H 8.476; H α 2 3.503; H α 3 3.332; N 104.704
Leu696	57.962	4.214	42.037		C 179.805; C δ 1 24.335; H 7.777' H β 2 1.782; H β 3 1.521; H δ 1 0.811; N 122.02
Asn697	54.342	4.641	37.117		C 177.726; H 8.138; H β 2 2.657; H β 3 2.068; H δ 21 7.518; H δ 22 6.803; N 119.131; N δ 2 109.966
Ala698	55.457	3.576	18.306	1.21	C 177.517; H 7.844; N 123.43
Leu699	58.182	3.564	41.148		C 179.118; C δ 1 25.837; C δ 2 23.465; C γ 27.333; H 7.271; H β 2 1.823; H β 3 1.47; H γ 1.812; H δ 1 0.805; H δ 2 0.714; N 116.097
Lys700	58.87	3.958	32.295	1.813	C 179.328; C δ 28.977; C ϵ 42.077; C γ 24.629; H 7.046; H δ 2 1.595; H δ 3 1.543; H γ 2 1.457; H γ 3 1.319; H ϵ 2.857; N 117.241
Met701	58.369	4.079	33.292		C 178.345; C ϵ 17.016; C γ 33.201; H 7.905; H β 2 2.573; H β 3 2.497; H γ 2 2.045; H γ 3 1.55; H ϵ 1.906; N 117.926
Leu702	56.623	3.694	40.87		C 177.889; C δ 1 24.73; C δ 2 23.143; C γ 25.931; H 7.587; H β 2 1.035; H β 3 0.255; H γ 0.948; H δ -0.183; H δ 2 -0/11; N 118.202
Glu703	57.849	4.211	29.983		C 177.965; C γ 36.165; H 7.129; H β 2 2.145; H β 3 2.097; H γ 2.529; H γ 3 2.343; N 113.911
Glu704	57.022	31.97 9	3.95		C 175.711; C γ 36.126; H 7.149; H β 2 1.473; H β 3 1.323; H γ 2 2.064; H γ 3 1.744; N 116.099
His705	52.542	4.719	30.486		H 7.502; H β 2 1.934; H β 3 1.889; H δ 2 6.493; N 117.063
Pro706	62.559	4.315	30.923		C 174.952; C δ 50.302; C γ 27.632; H β 2 2.014; H β 3 1.748; H δ 2 3.58; H δ 3 3.516; H γ 2 2.016; H γ 3 1.901
Ile707	57.821	3.838	39.973	1.479	C 174.542; C δ 1 10.116; C γ 1 26.021' C γ 2 20.631; H 7.279; H γ 12 1.337; H γ 13 0.748; H δ 1 0.215; H δ 2 0.633; N 123.507
Asp708	55.697	4.708	44.009		C 174.732; H 9.194; H β 2 2.855; H β 3 2.664; N 122.745
Leu709	55.563	4.339	45.033		C 172.978; C δ 1 26.483; C δ 2 25.885; C γ 26.728; H 7.274; H β 2 1.63; H β 3 1.186; H γ 1.165; H δ 1 0.525; H δ 2 0.66; N 118.69
Val710	60.301	4.813	34.063	1.926	C 173.983; C γ 1 22.887; C γ 2 21.686; H 8.303; H γ 1 0.662; H γ 2 0.629; N 124.49
Leu711	53.123	4.708	42.898		C 173.29; C δ 1 27.178; C δ 2 23.067; C γ 26.512; H 8.901; H β 2 1.72; H β 3 1.16; H γ 1.256; H δ 1 0.558; H δ 2 0.423; N 128.517
Met712	53.498	4.681	36.226		C 173.646; C ϵ 16.666; C γ 31.407; H 9.218; H ϵ 1.539; N 125.257
Asp713	55.506	4.811	41.886		H 8.337; H β 2 3.076; H β 3 2.805; N 128.878
Pro717	62.428	4.34	32.612		C 175.568; C δ 50.335; C γ 26.83; H β 2.224; H β 3 1.907; H δ 2 3.639; H δ 3 3.497; H γ 2 1.988; H γ 3 1.968

Val718	69.291	2.867	29.572	1.975	C γ 1 21.499; C γ 2 1.347; H 8.201; H γ 1 0.538; H γ 2 0.332; N 111.417
Met719	56.246	4.186	35.667	2.077	C 174.265; C ϵ 16.49; C γ 31.181; H γ 2 2.527; H γ 3 2.338; H ϵ 1.854
Asp720	53.834	4.42	40.683	2.885	C 175.35; H 8.731; N 127.365
Gly721	46.743	3.853			C 176.704; H 8.381; N 103.873
Tyr722	60.297	3.975	36.789		C 178.255; H 7.669; H β 2 3.145; H β 3 3.011; H δ 6.978; N 123.487
Glu723	58.632	4.048	29.167		C 178.931; C γ 35.238; H 8.826; H β 2 2.093; H β 3 1.772; H γ 2 2.381; H γ 3 2.348; N 121.568
Ala724	55.615	3.993	18.485	1.183	C 179.272; H 8.287; N 120.13
Thr725	68.552	3.427	68.245	4.145	C 175.572; C γ 2 21.074; H 7.625; H γ 2 0.926; N 114.225
Arg726	59.973	3.909	30.19		C 178.772; C δ 43.375; C γ 27.493; H 7.985; H β 2 1.981; H β 3 1.866; H δ 2 3.192; H δ 3 3.166; H γ 2 1.825; H γ 3 1.607; N 122.24
Gln727	59.762	3.952	27.695		C γ 33.984; H 8.08; H β 2 2.113; H β 3 2.054; H ϵ 21 7.883; H ϵ 22 6.512; H γ 2 2.519; H γ 3 2.342; N 116.915; N ϵ 2 111.39
Ile728	66.221	3.66	38.362	1.854	C 180.25; C δ 1 13.754; C γ 1 29.66; C γ 2 17.378; H 8.437; H γ 12 1.773; H γ 13 0.368; H δ 1 0.274; H γ 2 0.818; N 121.993
Arg729	57.74	3.956	29.473		C 179.341; C δ 42.139; C γ 26.822; H 8.96; H β 2 2.139; H β 3 1.926; H δ 3.106; H δ 3 3.04; H γ 2 2.021; H γ 3 1.697; N 119.434
Asp730	56.3	4.403	40.617		C 177.815; H 8.708; H β 2 2.672; H β 3 2.564; N 119.014
Ser731	60.747	4.097	63.918		C 176.119; H 7.537; H β 2 4.211; H β 3 3.834; N 113.094
Gly732	46.131				C 174.678; H 7.491; H α 2 4.03; H α 3 3.794; N 109.048
Arg733	56.79	3.834	30.454		C 176.048; C δ 43.231; C γ 26.426; H 7.582; H β 2 0.478; H β 3 0.294; H δ 2 2.732; H δ 3 2.495; H γ 2 0.916; H γ 3 0.848; N 119.232
Trp734	56.677	4.802	28.341		C 177.455; H 8.404; H β 2 3.312; H β 3 2.841; H δ 1 7.112; H ϵ 1 9.671; H γ 2 6.924; N 119.378; N ϵ 127.892
Gly735	47.376				C 175.632; H 8.255; H α 2 3.737; H α 3 3.681; N 109.446
Gly736	45.227	3.882			C 174.167; H 8.815; N 112.86
Leu737	53.204	4.414	44.691	1.636	C δ 1 26.957; C δ 2 23.309; C γ 26.435; H 7.29; H γ 1.213; H δ 1 0.476; H δ 2 0.615; N 124.244
Pro738	61.705	4.635	31.208		C 175.917; C δ 51.908; C γ 28.417; H β 2 1.728; H β 3 1.45; H δ 2 4.367; H δ 3 3.978; H γ 2 2.341; H γ 3 2.136
Ile739	60.83	4.353	40.34	1.379	C 173.676; C δ 1 14.325; C γ 2 17.563; H 9.044; H γ 12 1.582; H γ 13 0.404; H δ 1 0.622; H δ 2 0.542; N 124.039
Ile740	57.23	3.829	37.173	1.899	C 174.566; C δ 1 10.241; C γ 1 27.949; C γ 2 17.688; H 8.935; H γ 12 1.142; H γ 13 0.616; H δ 1 -0.211; H δ 2 0.383; N 129.159
Ala741	52.202	4.577	22.105	1.458	C 176.219; H 8.509; N 129.94
Leu742	52.536	5.464	42.81		C 176.547; C δ 1 25.655; C δ 2 24.38; C γ 26.35; H 8.126; H β 2 2.06; H β 3 1.068; H γ 1.383; H δ 1 0.765; H δ 2 0.626; N 121.075
Thr743	59.308	4.727	70.067	4.214	C 173.003; C γ 2 20.936; H 8.031; H γ 2 0.923; N 114.195
Ala744	52.322	4.54	19.746	1.373	C 177.358
Asn745	52.89	4.55	39.11		C 173.56; H 8.449; H β 2 2.786; H β 3 2.611; H δ 21 7.609; H δ 22 6.807; N 116.58; N δ 113.636

Ala746	52.09	4.259	18.884	1.205	C 177.514; H 8.472; N 123.957
Leu747	53.291	4.546	41.81	1.595	Cδ1 25.289; Cδ2 22.634; Cγ 27.274; H 8.305; Hγ 1.594; Hγ1 0.862; Hγ2 0.867; N 123.31
Pro748	65.574	4.099	31.823		C 178.131; Cδ 50.485; Cγ 27.702; Hβ2 2.278; Hβ3 1.887; Hδ2 3.785; Hδ3 3.3716; Hγ2 2.058; Hγ3 1.958
Asp749	55.4	4.4	39.519		C 177.869; H 8.508; Hβ2 2.592; Hβ3 2.53; N 115.429
Glu750	59.031	3.941	30.126		C 178.327; Cγ36.202; H 7.295; Hβ2 2.114; Hβ3 1.908; Hγ 2.173; N 121.503
Arg751	60.226	3.501	29.396	1.56	C 178.966; Cδ 43.292; Cγ 27.603; H 8.058; Hδ2 2.828; Hδ3 2.795; Hγ2 1.237; Hγ3 0.964; N 118.86
Glu752	59.027	4.044	28.78		C 178.855; Cγ 35.817; H 7.89; Hβ 1.996; Hβ3 1.915; Hγ2 2.253; Hγ3 2.204; N 118.743
Arg753	59.795	3.922	30.245		C 180.104; Cδ 43.437; Cγ 27.851; H 7.79; Hβ2 1.861; Hβ3 1.64; Hγ2 1.54; Hγ3 1.229; Hδ 2.826; N 121.459
Cys754	63.892	4.294	27.1		C 176.806; H 8.303; Hβ2 3.114; Hβ3 2.76; N 118.421
Arg755	59.033	4.213	29.647	1.855	C 180.678; Cδ 42.84; Cγ 26.769; H 8.023; Hδ2 3.205; Hδ3 3.077; Hγ2 1.701; Hγ3 1.521; N 122.016
Ala756	54.64	3.988	17.909	1.375	C 179.027; H 8.368; N 122.341
Ala757	53.236	4.207	19.603	1.756	C 176.528; H 7.414; N 118.061
Gly758	44.839				C 175.277; H 7.448; Hα2 4.252; Hα3 3.624; N 100.259
Met759	57.445	4	33.149		C 175.652; Cε 18.838; Cγ 34.858; H 7.876; Hβ2 2.142; Hβ3 1.756; Hγ2 2.525; Hγ3 2.241; Hε 1.545; N 119.691
Asp760	57.075	4.178	43.232		C 174.757; H 8.985; Hβ2 1.79; Hβ3 0.99; N 120.966
Asp761	51.914	4.37	44.167		C 173.902; H 7.372; Hβ2 2.586; Hβ3 2.31; N 113.731
Tyr762	56.339	5.392	42.144		C 172.038; H 9.463; Hβ2 2.826; Hβ3 2.736; Hδ 6.836; N 120.768
Leu763	53.771	4.155	45.454		C 173.696; Cδ1 25.914; Cδ2 24.163; Cγ 26.214; H 9.017; Hβ2 1.36; Hβ3 0.859; Hγ 1.184; Hδ1 0.148; Hδ2 0.513; N 124.828
Ala764	51.155	4.712	19.365	1.193	C 176.633; H 8.253; N 127.218
Lys765	53.465	4.181	33.554		Cδ 29.738; Cε 41.147; Cγ 25.161; H 7.909; Hβ2 1.498; Hβ3 1.365; Hδ2 1.485; Hδ3 1.40; Hγ2 1.254; Hγ3 0.981; Hε2 2.599 Hε3 2.561; N118.114
Pro766	62.043	4.2	34.443		C 175.072; Cδ 50.24; Cγ 24.753; Hβ2 2.137; Hβ3 1.75; Hδ 3.531; Hδ3 3.322; Hγ 1.765
Phe767	54.94	4.855	40.334		C 174.825; H 7.454; Hβ2 3.068; Hβ3 2.831; Hδ 6.995; N 116.87
His768	55.05	4.681	32.535		H 8.756; Hβ2 3.182; Hβ3 3.09; N 120.297
Arg769	61.078	3.678	30.333		C 177.366; Cδ 43.355; Cγ 26.664; Hβ2 1.898; Hβ3 1.702; Hγ2 1.548; Hγ3 1.5; Hδ 3.07
Asp770	57.761	4.303	39.614	2.603	C 179.44; H 9.718; N 118.507
Glu771	66.955	4.044	29.679		C 179.227; Cγ 36.81; H 7.717; Hβ2 2.241; Hβ3 1.962; Hγ2 2.308; Hγ3 2.208; N 120.779
Leu772	57.479	3.839	41.707		C 177.767; Cδ1 25.965; Cδ2 22.772; Cγ 26.797; H 7.768; Hβ2 1.629; Hβ3 1.305; Hγ 1.3; Hδ1 0.579; Hδ2 0.482; N 120.468
Lys773	59.633	3.714	32.184		C 176.959; Cδ 28.845; Cε 42.215; Cγ 24.96; H 8.646; Hβ2 1.895; Hβ3 1.782; Hδ2 1.647; Hδ3 1.503; Hγ2 1.413; Hγ3 1.347; Hε 2.922; N 119.842

Ala774	55.112	4.078	17.699	1.408	C 180.953; H 7.41; N 118.383
Ile775	61.87	3.955	36.097	2.361	C 177.61; Cδ1 9.942; Cγ1 28.567; Cγ2 17.895; H 7.195 Hγ12 1.693; Hγ13 1.249; Hγ1 0.714; Hγ2 0.788; N 119.295
Leu776	58.446	3.744	40.864		C 178.896; Cδ1 26.847; Cδ2 23.146; Cγ 26.972; H 8.076; Hβ2 1.801; Hβ3 1.387; Hγ 1.61; Hδ1 0.649; Hδ2 0.616; N 121.688
Asp777	56.685	4.056	40.104		C 177.594; H 8.6; Hβ 2.153; Hβ3 2.427; N 117.181
Arg778	57.931	3.818	30.838		C 177.896; Cδ 43.134; Cγ 26.344; H 7.335; Hβ2 1.655; Hβ3 1.452; Hδ2 2.782; Hδ3 2.746; Hγ2 0.553; Hγ3 0.309; N 117.497
Trp779	58.218	4.517	31.176		C 176.419; H 7.941; Hβ2 3.316; Hβ3 2.858; Hδ1 7.37; Hε1 11.352; HZ2 7.22; N 116.27; Nε1 130.083
Cys780	56.737	4.83	28.143		H 8.396; Hβ2 3.021; H3β 2.736; N 118.782
Pro781	63.498	4.318	31.934		C 177.465; Cδ 50.329; Cδ 27.296; Hβ2 2.229; Hβ3 1.802; Hδ2 3.436; Hδ3 3.182; Hδ2 1.857; Hδ3 1.813
Leu782	55.304	4.22	42.094		C 177.689; Cδ1 24.834; Cδ2 23.491; Cγ 26.923; H 8.363; Hβ2 1.524; Hβ3 1.427; Hγ 1.525; Hδ1 0.779; Hδ2 0.711; N 121.551
Thr783	61.465	4.216	69.946	4.083	C 173.957; Cγ2 21.56; H 7.967; Hγ 1.048; N 114.485
Ala784	52.372	4.273	19.521	1.289	C 177.315; H 8.162; N 126.503
Ser785	58.117	4.354	64.092		C 173.393; H 8.212; Hβ2 3.759; Hβ3 3.717; N 115.372
Asp786	55.938	4.282	42.143		H 7.933; Hβ2 2.558; Hβ3 2.451; N 127.622

Table S11. Structural statistics for final 20 REC_{SagS} structures

Number of NOE constraints	
All	2311
Intra residues $ i-j =0$	589
Sequential, $ i-j =1$	619
Medium-range, $1< i-j <5$	396
Long-range, $ i-j \geq 5$	707
Number of Hydrogen Bond Constraints	94
Number of Dihedral Angle Constraints	182
Number of Constraint Violations ($>0.5\text{\AA}$)	0
Number of Angle Violations ($>5^\circ$)	0
Average target function	1.03 ± 0.10
RMSD for residue N-terminal helix-loop-helix domain(2-73) ^a	
Backbone	$0.48 \pm 0.07\text{\AA}$
Heavy Atoms	$0.89 \pm 0.05\text{\AA}$
Ramachandran Plot ^b	
Most Favored Regions	79.2%
Additionally Allowed Regions	20.6%
Generously Allowed Regions	0.1%
Disallowed Regions	0.1%

^aNumber of structures used in rmsd calculation was 20.

^bRamachandran analysis was performed using PROCHECK-NMR program (ref).

79.2% in most favored regions
 20.6% in additionally allowed regions
 0.1% in generously allowed regions
 0.1% in disallowed regions

Figure S5 | Ramachandran plot of REC_{Sags}. Ramachandran plot of REC_{Sags} showing sterically allowed ϕ and ψ angles. The most favored regions are shaded blue, the additionally allowed regions are shaded dark purple, the generously allowed regions are shaded light purple while the disallowed regions are shaded white.

Appendix II

List of Publications

(*Equal contribution)

1. Xu, L. *, **Venkataramani, P***, Ding, Y., Liu, Y., Deng, Y., Yong, G. L., Xin, L., Ye, R., Zhang, L., Yang, L. and Liang, Z-X. (2016) A Cyclic di-GMP-binding Adaptor Protein Interacts with Histidine Kinase to Regulate Two-component Signaling. *The Journal of Biological Chemistry* **291(31)**:16112-16123.
2. **Venkataramani, P.** and Liang, Z-X. (2016) Enzymatic synthesis of c-di-GMP using a thermophilic diguanylate cyclase In c-di-GMP Signaling: Methods and Protocols. *Methods in Molecular Biology*. (Accepted).
3. Xu, L., Xin, L., Zeng, Y., Yam, J. K. H., Ding, Y., **Venkataramani, P.**, Cheang, Q. W., Yang, X., Tang, X., Zhang, L-H. , Chiam, K-H. , Yang, L. and Liang, Z-X. (2016) A Cyclic di-GMP-binding Adaptor Protein Interacts with a chemotaxis methyltransferase to control flagellar motor switching. *Science Signaling* **9**: ra102.

Conference Talks

1. **American Society for Microbiology (ASM) Conference on *Pseudomonas* 2015.**
Washington D.C., USA
8th-12th September 2015
Delivered a talk on ‘Regulation of two-component signalling by a cyclic-di-GMP binding PilZ protein’ and was awarded Travel Grant from ASM for attending the conference.