Division of Molecular Structural Biology
Department of Medical Biochemistry and Biophysics
Karolinska Institutet, Stockholm, Sweden

STRUCTURAL STUDIES OF ENZYMES REGULATING BACTERIAL LIFE STYLE AND CELL WALL BIOGENESIS

Ming Wei Chen

Karolinska
Institutet

Stockholm 2013

ABSTRACT

Bacteria can adapt to different biotic and abiotic environments by changing their life style, switching between sessile or motile, free-living or community-bound, virulent or dormant states. Each bacterium also faces the challenges of maintaining a protective barrier while growing, replicating and responding to environmental changes. Governing these processes are enzymes involved in signalling and cell-wall biogenesis, which are the subjects of this thesis.

The second messenger cyclic di-GMP (c-di-GMP) regulates a vast array of processes such as motility, biofilm formation, virulence and cell cycle progression on transcriptional, post-transcriptional and post-translational levels. RocR from *Pseudomonas aeruginosa* is a response regulator protein containing an active EAL domain that breaks down c-di-GMP. The crystal structure of RocR was determined to 2.5 Å resolution, revealing a compact tetrameric structure with the subunits displaying two conformational states. The unique architecture allows two phospho-receiver domains to be adjacent to the EAL active sites while being exposed and available for phosphorylation events. Also, solution studies using SAXS and biochemical analyses suggest that the protein does not require large conformational changes to alter its phosphodiesterase activity, leading to a mechanistic model of signal propagation from the phosphorylation site to the EAL active site based on secondary structural changes. Tbd1265 from *Thiobacillus denitrificans* is a transmembrane protein containing a GGDEF-EAL tandem domain that can both synthesise and hydrolyse c-di-GMP. Functional studies confirmed the bifunctionality of this tandem domain and further suggested that a predicted coiled-coil region preceding the GGDEF domain is required for activity. The crystal structure of a construct comprising these two domains of Tbd1265 was determined to 3.4 Å, revealing a conformation of the GGDEF domains in the dimeric molecule that is not compatible with product-inhibition or catalysis. We propose a regulatory mechanism where Tbd1265 can adopt at least three conformations (resting, active and inhibited) based on signals from the periplasmic binding protein (PBP) domain.

MurB is an essential oxidoreductase that produces UDP-*N*-acetylmuramic acid, a precursor for peptidoglycan synthesis. The crystal structure of the ternary complex of *P. aeruginosa* MurB with $NADP^+$ and FAD revealed that the substrate channel can accommodate two distinct substrate molecules. The study also revealed a potassium ion in the active site that directly binds the substrates and can stabilise the transition state of the reaction, thus explaining the activating effect of potassium ions on MurB catalysis. The structure of the MurB ternary complex provides a useful template for the design of novel enzyme inhibitors that might be developed into promising drug candidates.

LIST OF PUBLICATIONS

I.   Ming Wei Chen, Masayo Kotaka, Clemens Vonrhein, Gérard Bricogne, Feng Rao, Mary Lay Cheng Chuah, Dmitri Svergun, Gunter Schneider, Zhao-Xun Liang and Julien Lescar. Structural insights into the regulatory mechanism of the response regulator RocR from *Pseudomonas aeruginosa* in cyclic Di-GMP signaling. J Bacteriol. 2012; 194, 4837 – 4846.

II. Chong Wai Liew, Ming Wei Chen, Chun Loong Ho, Surya Wahyu, Tobias Cornvik, Gunter Schneider, Zhao-Xun Liang and Julien Lescar. The structure of the diguanylate cyclase-phosphodiesterase tandem domain of a bacterial sensor protein suggests a model for its activation. Manuscript in progress.

III. Ming Wei Chen, Bernhard Lohkamp, Robert Schnell, Julien Lescar and Gunter Schneider. Substrate channel flexibility in *Pseudomonas aeruginosa* MurB accommodates two distinct substrates. PLoS ONE. 2013; 8, e66936.

Publication not included in this thesis:

Chong Wai Liew, Martina Nilsson, Ming Wei Chen, Huihua Sun, Tobias Cornvik, Zhao-Xun Liang and Julien Lescar. Crystal structure of the acyltransferase domain of the iterative polyketide synthase in enediyne biosynthesis. J Biol Chem. 2012; 287, 23203 – 23215.

CONTENTS

LIST OF ABBREVIATIONS

5'-pGpG	Linear di-GMP; 5'-phosphoguanylyl-(3'$\rightarrow$5')-guanosine
ASU	Asymmetric unit
C3e	Carbon-3 atom of enolpyruvyl
C4n	Carbon-4 atom of nicotinamide
cAMP	Cyclic AMP; adenosine 3',5'-cyclic monophosphate
c-di-GMP	Cyclic di-GMP; bis(3'$\rightarrow$5')-cyclic diguanylate
cGMP	Cyclic GMP, guanosine 3',5'-cyclic monophosphate
DAP	Diaminopimelic acid
DGC	Diguanylate cyclase
EcMurB	*Escherichia coli* MurB
FAD	Flavin adenine dinucleotide
FMN	Flavin mononucleotide
GMP	Guanosine 5'-monophosphate
GTP	Guanosine 5'-triphosphate
HK	Histidine kinase
IC_{50}	Half-maximal inhibitory concentration
MW	Molecular weight
NADH	β-nicotinamide adenine dinucleotide, reduced
$NADP^+$	β-nicotinamide adenine dinucleotide phosphate, oxidised
NADPH	β-nicotinamide adenine dinucleotide phosphate, reduced
NAG	*N*-acetylglucosamine
NAM	*N*-acetylmuramic acid
PaMurB	*Pseudomonas aeruginosa* MurB
PBP	Periplasmic binding protein
PCR	Polymerase chain reaction
PDB	Protein Data Bank
PDE	Phosphodiesterase
ppGpp	5'-disphosphate 3'-diphosphate guanosine
rmsd	Root mean squared deviation
SAD	Single-wavelength anomalous dispersion
SaMurB	*Staphylococcus aureus* MurB
SAXS	Small-angle X-ray scattering
SEC	Size-exclusion chromatography
UDP	Uridine 5'-diphosphate
UNAG	UDP-*N*-acetylglucosamine
UNAGEP	UDP-*N*-acetylglucosamine-enolpyruvate
UNAM	UDP- *N*-acetylmuramic acid
UTP	Uridine 5'-triphosphate

1 INTRODUCTION

1.1 CYCLIC DI-GMP IS A UBIQUITOUS SECOND MESSENGER IN BACTERIA

The amazing intricacy of Nature is well reflected in the various life processes of the Bacterial kingdom. Bacterial replication, motility, community formation, virulence and alteration of abiotic environments are among those many processes that are important to our medical and industrial endeavours. These life processes are governed by extracellular environmental cues that are relayed into the cell by second messengers, small molecules that are responsible for amplifying the cues via signal transduction cascades. First discovered serendipitously 25 years ago by Benziman et al (*1*), cyclic di-GMP (c-di-GMP) is now recognized as an ubiquitous second messenger in bacteria (Fig 1.1). It regulates transition between motile and sessile lifestyles, various stages of biofilm formation, expression of virulence and cell cycles – processes occurring in pathogens infecting humans, animals and crops, and bacteria that cause biofouling (*2*).

C-di-GMP was first discovered as an allosteric activator of cellulose synthase in the acetic acid bacterium *Gluconacetobacter xylinus* (known as *Acetobacter xylinus* at that time). Low activity of the purified enzyme fraction led to the search for a cofactor that was found to be GTP-derived with a guanine-ribose-phosphate ratio of 1:1:1 (*2-4*), and eventually identified as bis(3'$\rightarrow$5')-cyclic diguanylate. *G. xylinus* proteins that metabolize c-di-GMP were subsequently identified by the discovery of the cyclic diguanylate (*cdg*) operons (*5*). C-di-GMP was found to be synthesised by diguanylate cyclases (DGCs) consisting of the GGDEF domain, which had been previously designated based on the highly conserved Gly-Gly-Asp-Glu-Phe sequence motif. On the other hand, the phosphodiesterases (PDEs) that specifically break down c-di-GMP had a new domain architecture, named EAL for its Glu-Ala-Leu signature motif. This work pioneered further works on c-di-GMP regulation and functions. Subsequently, another c-di-GMP-specific PDE was found and named the HD-GYP domain (*6*). The structural and biochemical characteristics of these major players are described in more detail in the following section.

The advent of genomic sequencing technologies and large scale microbial genome projects revealed that all major bacterial phyla possess at least one type of the c-di-

GMP metabolising proteins, albeit not in every species (*2*). By comparing examples in individual phyla, bacteria capable of multiple life styles seem to encode more of these proteins than their relatives that are obligate parasites (*7*), suggesting that c-di-GMP signalling is important for adaptation. A good demonstration is the degeneration of genes encoding DGCs and PDEs in *Yersinia pestis*, an obligate pathogen, in comparison to homologous genes found in the non-obligate free-living close relative *Yersinia pseudotuberculosis*. Surprisingly, c-di-GMP signalling is not evident in Archaea and eukaryotes in general, with the exception of lower eukaryotes such as *Dictyostelium* (*8*).
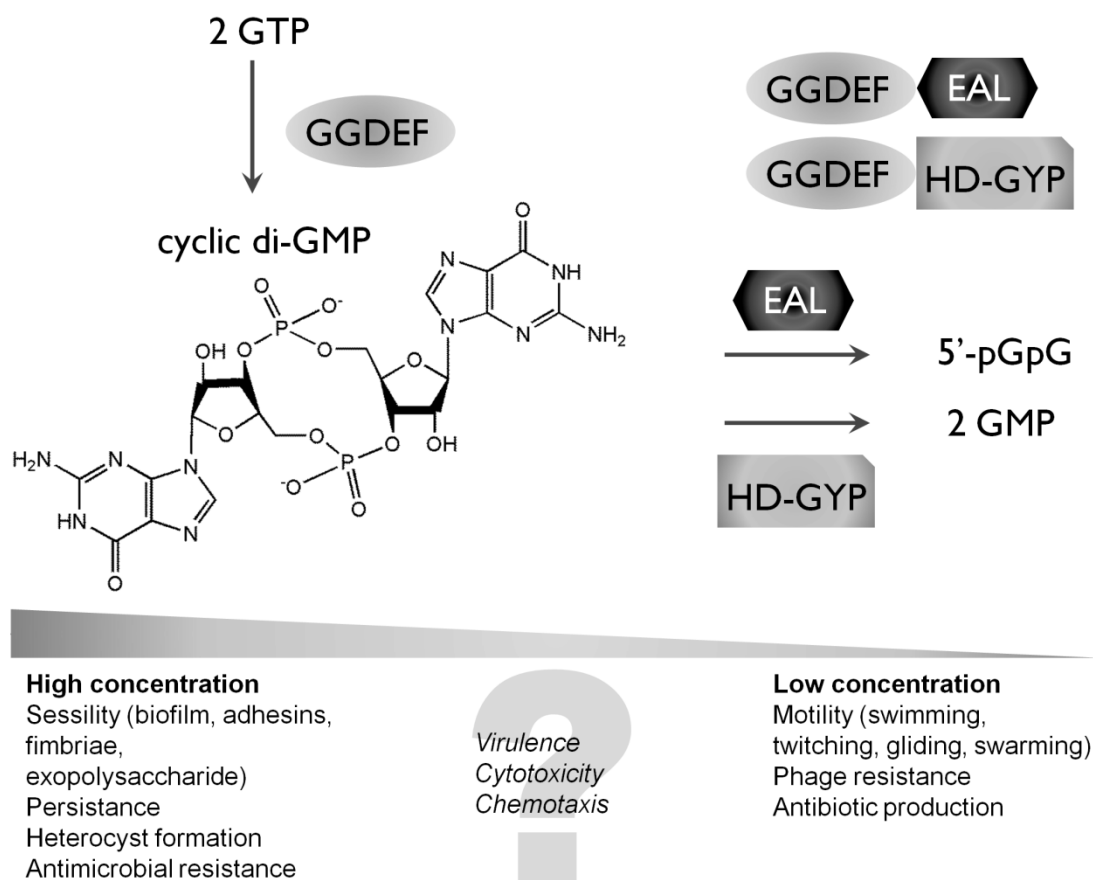
Figure 1.1. Cyclic di-GMP as a second messenger in bacteria. C-di-GMP is synthesised by GGDEF proteins acting as diguanylate cyclases and broken down by the EAL or HD-GYP phosphodiesterases. The GGDEF-EAL and GGDEF-HD-GYP tandem domain proteins may play various roles (Section 1.4). Phenotypes commonly associated with high or low concentrations of c-di-GMP are listed. Certain phenotypes, especially virulence, may be activated under either high or low c-di-GMP depending on the organism in question. Adapted from reference (*2*).

1.2 ENZYMES THAT REGULATE CYCLIC DI-GMP LEVELS: DIGUANYLATE CYCLASES AND PHOSPHODIESTERASES

The scheme of c-di-GMP metabolism in bacteria has emerged with the biochemical characterisation of the DGC and PDE enzymes involved. With two GTP molecules as precursor, GGDEF proteins catalyze the formation of c-di-GMP via a linear diguanylate intermediate (*1*). C-di-GMP can be subsequently broken down into the linear 5'-pGpG by PDE proteins; HD-GYP proteins can further hydrolyze 5'-pGpG into two GMP molecules (Fig 1.1). Sequence analysis of these proteins revealed strong conservation of signature sequence motifs amongst examples taken across microbial genomes (*9*). Each type of enzyme will be discussed in the following.

1.2.1 The GGDEF domain

Sequence analysis of the GGDEF domain indicates that it is a nucleotidyl cyclase structurally similar to the type III adenylate cyclase involved in cyclic AMP (cAMP) synthesis (*10, 11*). The predicted fold is strikingly similar to the adenylyl and guanylyl cyclases despite the low primary sequence homology. This in turn implies that the GGDEF would function as a single catalytic domain that homodimerises during catalysis; each protomer is expected to contribute a part of the complete active site located at the dimer interface. By extension, GGDEF-containing proteins would also be expected to undergo significant conformational changes upon activation or deactivation, often by rotational movements of the cyclase domain (*12*). As a result, regulatory mechanisms that physically prevent GGDEF protomers from forming the active site can be expected. Specificity of DGC proteins for GTP was established upon comparing the activities of several GGDEF proteins across different bacteria (*13*); the study also confirmed that GGDEF activity is highly regulated by sensor domains on the same polypeptide. For example, the response regulator Rrp1 from *Borrelia burgdorferi* only displays DGC activity upon phosphorylation (*13*). It was also demonstrated that GGDEF produces only c-di-GMP, and that c-di-GMP synthesis is well represented in many phylogenetic branches of bacteria.

Much of what we know about the GGDEF domain comes from the collaborative work of the Jenal and Schirmer groups on the response regulator PleD. PleD comprises a GGDEF domain preceded by two phosphoreceiver domains (i.e. REC-REC-GGDEF); the protein is localized to the cell pole of *Caulobacter crescentus*. As a phosphoreceiver protein in a two-component system, phosphorylated PleD was activated *in vitro* and
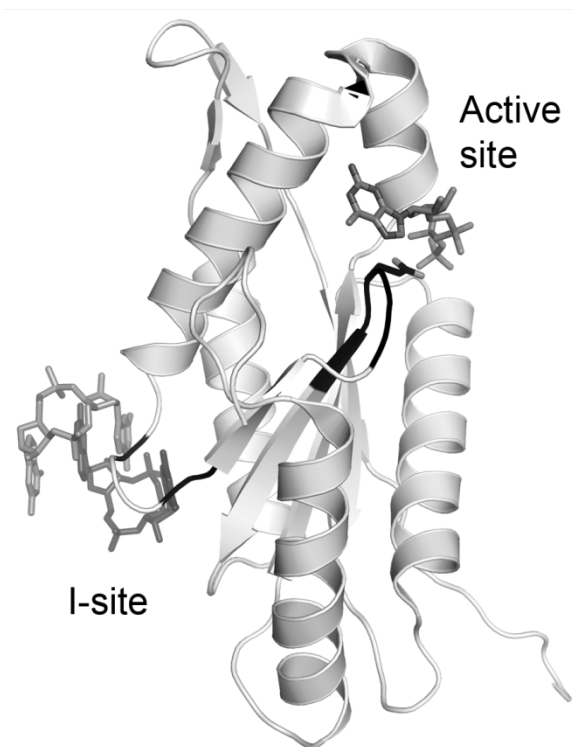
Figure 1.2. Structure of the PleD GGDEF domain. The structure of a GGDEF monomer from the PleD structure in complex with the substrate analogue GTP-α-S and the product-inhibitor c-di-GMP (PDB code 2V0N) is shown. The active site (with GGEEF motif coloured in black and the catalytic glutamate as stick model) and the inhibitory I-site (with RxxD residues in black) are located on opposite sides of the central β-sheet. GTP-α-S and c-di-GMP (grey stick models) bind to the active site and I-site, respectively. Active site magnesium ions are omitted for clarity.

catalysed c-di-GMP formation from GTP (*14*). More importantly, crystal structures were obtained for native PleD in complex with c-di-GMP (*15*) and artificially phosphorylated PleD bound to the non-hydrolysable GTP analogue GTPαS (*16*), allowing better understanding of the cyclisation reaction. PleD exists as a dimer whereby the REC-REC "stem" allows a loose association of the two GGDEF domains. The fold of the GGDEF domain is similar to adenylyl cyclase as predicted earlier (Fig 1.2). C-di-GMP was shown to bind the active site that contains the signature GG(D/E)EF motif and also an allosteric inhibitory site (I-site) that is characterised by an RxxD motif. It was predicted that the two GGDEF domains would align symmetrically along a dyad during catalysis, each contributing one GTP molecule towards the reaction. The subsequent structure of phosphorylated PleD suggested that conformational changes initiated in the REC-REC stem module promote GGDEF dimerisation that would lead to efficient catalysis; also, a two-metal mechanism employing magnesium ions similar to the adenylyl cyclase was inferred. Similar to the active site, the I-site is also formed by two GGDEF domains that are brought together by the product inhibitor c-di-GMP. The active site and I-site are approximately antipodal on a GGDEF domain (Fig 1.2), making catalysis impossible when c-di-GMP binds to the I-site and promotes a non-catalytic dimerisation mode. The same protein fold is observed in other GGDEF proteins, such as the response regulator WspR (*17, 18*), the catalytically inactive FimX and LapD (discussed in section 1.4), the zinc-regulated DgcZ (*19*), and other GGDEF crystal structures reported by structural genomics consortia. Analyses of WspR with a REC-GGDEF domain arrangement

4

support the role of protein oligomerisation and conformational changes in the regulation of DGCs (*17, 18*).

Although the structure of a catalytically competent GGDEF dimer is lacking, these works outlined two important mechanisms regulating DGC activity. Sensor domains adjacent to the GGDEF domain can initiate conformational changes according to their cognate signals; these tertiary and quaternary structural changes bring GGDEF active sites closer or apart thus changing c-di-GMP output. In addition, c-di-GMP binds to the I-site upon reaching a certain concentration, causing negative feedback inhibition that presumably imposes spatiotemporal limitations on c-di-GMP signalling.

1.2.2 The EAL domain

Work on the EAL domain started at the same time as the GGDEF domain. The pioneering efforts of the Benziman group established that *G. xylinus* EAL proteins hydrolysed c-di-GMP into the linear 5'-pGpG in a reaction that is dependent on magnesium or manganese ions and inhibited by calcium. The phosphodiesterase nature of the domain was confirmed by degrading the general PDE substrate bis(*p*-nitrophenyl) phosphate using an EAL-domain protein construct (*20*). A similar strategy employing an isolated EAL protein construct showed that it is specific for c-di-GMP but not other cyclic nucleotides such as cAMP and cGMP (*21*). As further breakdown of 5'-pGpG into GMP was too slow it was deemed irrelevant *in vivo*.

Structures of several EAL proteins (*22-25*), including Paper I in this thesis (*26*), showed that the EAL domain is a TIM barrel (Fig 1.3), a common protein fold employed in various enzymes. A "lobe" structure composing of secondary structural elements connecting the EAL and its preceding N-terminal domain is also often seen. Curiously, the various structures also display a common dimerisation mode that could be important for regulatory purposes. The dimerisation interface consists of one long helix from each protomer, and one short helix each that align to form a compound helix, providing extensive interactions that stabilise the antiparallel EAL dimer. Works on the light sensing protein BlrP1 produced an elegant display of EAL structures across a range of pH values, suggesting that a higher pH allows optimal metal coordination bond lengths in the active site (*22*). The strong inhibitory effect of Ca^{2+} is also explained by distortion of the active site. A subsequent analysis of the isolated EAL domain from the *Thiobacillus denitrificans* Tbd1265 (*24*) (Fig 1.3) reinforced the two-

metal mechanism seen in BlrP1. When compared to some EAL domains that only show one metal ion in the absence of c-di-GMP (*26*), the Tbd1265 EAL also raises the question whether the second metal may bind only in the presence of substrate.
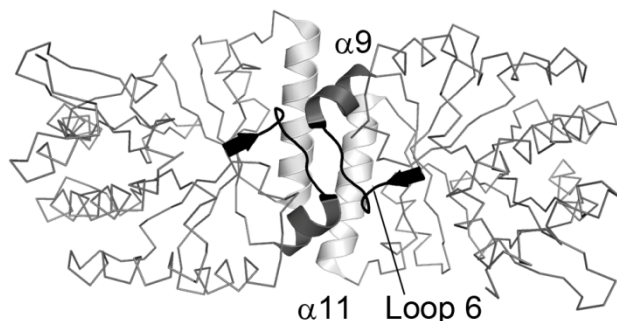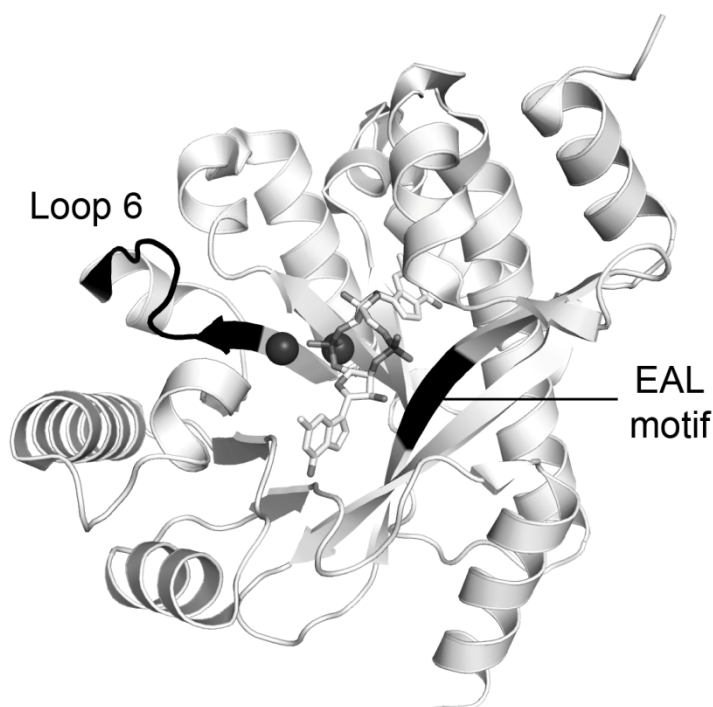
Figure 1.3. Structure of the Tbd1265 EAL domain. Top: the structure of an EAL monomer in complex with two magnesium ions (grey spheres) and the substrate c-di-GMP (stick model) is shown (PDB code 3N3T). The signature EAL motif and the catalytic Loop 6 are shown in black. All EAL domains adopt a TIM barrel architecture with the active site on the top surface where the C-termini of the β-strands are located. Bottom: EAL domains dimerise via Loop 6, the short helix α9 and the long helix α11 (numbering according to reference (*24*)).

By using the *Pseudomonas aeruginosa* response regulator RocR (PA3947) as a model protein, Rao et al studied the roles of 14 amino acid residues in the catalysis of EAL proteins (*27*). These include the E(X)L signature motif (EVL in RocR) and residues located on a putative catalytic loop ("Loop 6" as termed in RocR). Five essential residues, including E175 of the signature motif and D295 in Loop 6, were found to coordinate the catalytic magnesium ion. In addition, E352 was proposed to be a general base that is important for generating a reactive hydroxide from a metal-coordinated water molecule in the active site. Also, analysis of primary sequences of active versus inactive EAL proteins confirmed the findings, especially of the essentiality of E352. The active EAL proteins showed a conserved DFG(T/A)GYSS sequence in Loop 6; this finding was emphasised in a subsequent work where the role of Loop 6 was

investigated (*28*). Mutations introduced into Loop 6 regions of RocR and another *P. aeruginosa* EAL-containing protein PA2567 significantly lowered the proteins' PDE activities. On the other hand, reverse engineering of the degenerate Loop 6 in the PDE-inactive DGC2 protein from *G. xylinus* restored PDE activity, further corroborating the catalytic role of the loop. More intriguingly, hydrogen-deuterium exchange experiments on RocR peptides showed that Loop 6 has different solvent exposures under different PDE activity levels, suggesting that conformation changes of this region may regulate the PDE activity of EAL proteins. These results, together with the aforementioned EAL protein structures in complex with metal ions and substrate, explain the roles of conserved residues. Mechanisms regulating EAL activity however leave much to be discovered.

1.2.3 The HD-GYP domain

Biochemical and structural data of the HD-GYP domain are scarce compared to the body of literature on GGDEF and EAL domains. It was first hypothesised to be involved in c-di-GMP metabolism due to the occurrence of GGDEF-HD-GYP tandem domains, not dissimilar to the frequently observed GGDEF-EAL tandem domains. Although EAL and HD-GYP genes coexist in several analysed genomes, the prevalence of HD-GYP in the genomes seems to correlate positively with GGDEF but negatively with EAL (*6*). As examples on the extreme ends, *Escherichia coli* encodes 18 EAL proteins but no HD-GYP protein, while *Thermotoga maritima* encodes nine HD-GYP proteins but no EAL. The HD-GYP domain is one type of enzymes in the HD superfamily known as hydrolases of diverse functions (*29*); it was named after the HD and the additional GYP signature motifs.

The best known HD-GYP protein is the response regulator RpfG from *Xanthomonas campestris*, a plant pathogen that damages cruciferous vegetables. Ryan, Dow and colleagues demonstrated c-di-GMP-specific activity of RpfG *in vitro* and *in vivo*; mutations in conserved residues caused the loss of PDE activity and virulence factor expression (*30*). The one significant difference of HD-GYP compared to EAL is the full breakdown of c-di-GMP into two GMP molecules. Based on the crystal structure of the inactive protein Bd1817 from *Bdellovibrio bacteriovorus*, a catalytic mechanism involving a binuclear metal centre was proposed (*31*). Recently, crystals structures of an active GAF-HD-GYP protein from *Persephonella marina* revealed a trinuclear metal centre containing an invariant Fe(III) ion that forms a bond directly with a c-di-

GMP phosphate oxygen atom (*32*). The PDE mechanism is different from the one of EAL proteins (*24*) though more detailed roles of the HD-GYP active site elements have to be elucidated.

1.3 REGULATORY DOMAINS AND SIGNALLING PATHWAYS THAT REGULATE DGC AND PDE ACTIVITY

As c-di-GMP is a second messenger responsible for transmitting and amplifying the effects of relevant external signals, it is conceivable that DGC and PDE activities are regulated by sensor or signal receiver domains that detect those signals. Combining the common DGC and PDE domains with a variety of regulatory domains should allow recycling of a few designs in forging a complex signalling network that confers adaptability to microbes encoding them. Interestingly, the regulatory domains are almost invariably located N-terminal to the catalytic DGC and/or PDE domains (*33*). To date, the most commonly observed regulatory domains in this context are REC, PAS and GAF domains (*2*).

1.3.1 RocR is a response regulator in c-di-GMP signalling

The phosphoreceiver REC domain (CheY-like domain) is an important hallmark in two-component signalling pathways in bacteria (*34*). A canonical REC domain contains a conserved aspartate residue that can be transiently phosphorylated by the cognate transmembrane histidine kinase (HK). The HK is in turn regulated by ligand binding, conveying an extracellular signal into the cell via phosphorylation events. The phosphorylated REC domain can exert activating or inhibitory effects on its binding partner, which could be an adjacent catalytic domain that is part of the response regulator protein. The regulatory effects are caused by conformational changes resulting from changes in REC dimerisation modes upon phosphorylation. A whopping percentage of at least 5.4% of bacterial response regulator proteins contain c-di-GMP-metabolising domains (*2*), indicating the role of two-component systems in c-di-GMP signalling. A well-known example is the aforementioned REC-REC-GGDEF protein PleD, whose DGC activity is changed by phosphorylation-induced movements of the REC-REC stem region.

An example of the REC-EAL construct (the PvrR family) is the previously mentioned RocR, which is one of the two response regulators operating in the non-canonical two-component system RocSAR (Fig 1.4). RocSAR can be called a three-component

system, consisting of one sensor HK RocS1 whose cognate ligand is yet unknown, and two response regulators RocR and RocA1 that act antagonistically *in vivo* (*35, 36*). RocA1 contains a DNA-binding domain that activates transcription of *cupB* and *cupC* fimbriael genes, a process that occurs in later stages of *P. aeruginosa* biofilm formation. Meanwhile, RocR inhibits *cup* gene expression via an unknown mechanism that may involve c-di-GMP depletion. Disruption of the RocSAR system results in small colonies that have less water channels, indicating its importance in biofilm maturation and the resulting chronic infection on host tissues. The regulatory mechanism governing RocR's PDE activity and its ultimate downstream effects is thus highly interesting.
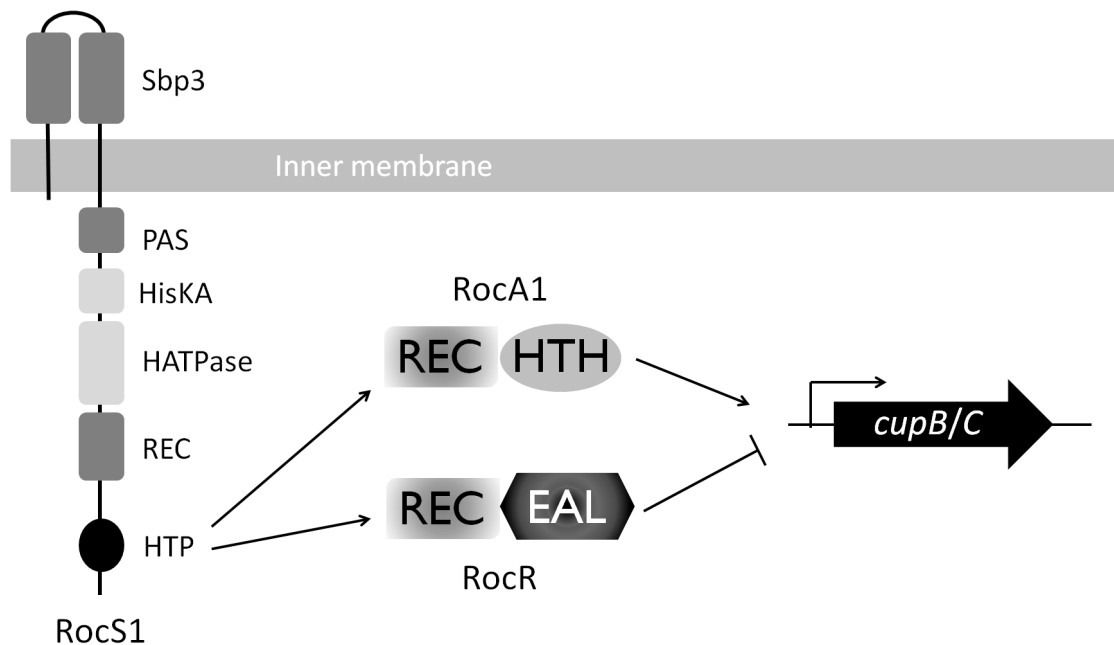
Figure 1.4. The non-canonical two-component system RocSAR from *P. aeruginosa*. This signalling system is composed of the histidine kinase RocS1 and two response regulators, RocA1 and RocR, which regulate *cupB/C* genes antagonistically. Ligand-binding events on the extracellular solute binding protein domains (Sbp3) are proposed to trigger autophosphorylation in the autokinase domain (HisKA). The phosphate is ultimately transferred to the phosphotransferase (HTP), which phosphorylates RocA1 or RocR. Phosphorylated RocA1 activates *cupB/C* gene expression via its DNA-binding domain containing a helix-turn-helix (HTH) motif. On the other hand, RocR represses *cupB/C* expression via an unknown mechanism that may involve c-di-GMP depletion.

1.3.2 Sensor domains that regulate DGC and PDE activity

PAS sensor domains are universal across all kingdoms of life, having diverged in protein sequence to detect various types of signals but maintaining a common protein structure (*37*). They can bind directly to metal ions, gases and small molecules such as carboxylic acids. Alternatively, a prosthetic group is recruited for signal detection;

flavin mononucleotide (FMN) and flavin adenine dinucleotide (FAD) are used to measure blue light and redox status in the cell, for example. The GAF domain is another universal sensor protein structurally similar to PAS, being able to bind small ligands and chromophores in a similar fashion (*38*). Aside from mediating protein-protein interactions and transducing signals, these domains may also regulate effector proteins by causing cellular localisation (*37*). C-di-GMP-metabolising domains are also coupled to more signal-specific sensor domains apart from the large REC, PAS and GAF families. A globin-coupled DGC from *Bordetella pertussis* has been reported to sense molecular oxygen via a heme cofactor and regulate biofilm formation (*39*). The EAL domain in the light-regulated PDE BlrP1 is regulated by an N-terminal BLUF domain that contains flavin mononucleotide. Light absorption by the cofactor flavin is proposed to propagate secondary structural changes that reach the Loop 6 of the adjacent EAL domain (*22*). Furthermore, the regulatory domain does not have to be the same polypeptide as the catalytic domain, as in the case of the NO sensor H-NOX that regulates DGCs in *Shewanella* (*40*). Last but not least, these regulatory effects may extend to other c-di-GMP-metabolisers thus causing a signaling cascade. For example, RpfG (a REC-HD-GYP) is proposed to bind several GGDEF proteins upon phosphorylation and thereby regulate their activities (*41*), in addition to increasing its own PDE output.

1.4   ENZYMATIC PARADOX? THE PREVALENCE OF GGDEF-EAL TANDEM DOMAINS

Survey of DGC and PDE genes in sequenced microbial genomes produces a puzzling observation that GGDEF-EAL tandems occur frequently as multi-domain constructs. It is estimated that 1/3 of all GGDEF domains and 2/3 of all EAL domains are encoded as such tandems (*2*). Just as sensor-regulatory domains almost invariably precede DGC or PDE domains in sequence, the observed tandem domains consistently adopt the GGDEF-EAL arrangement. Based on sequence analysis targeting essential residues, approximately 15% of all EAL proteins and a surprising 40% of GGDEF in GGDEF-EAL tandems are predicted to be catalytically inactive (*42*). Hence, we can envision at least three categories of tandems that play different roles: DGC- and PDE-competent, only DGC- or PDE-competent, and catalytically inactive.

1.4.1 Tandem domains containing active GGDEF and EAL

Bifunctional enzymes with opposite activities are not uncommon; the ubiquitous histidine kinases perform both kinase (autophosphorylation and phosphotransfer) and phosphatase activities (*43*). Another bacterial signaling nucleotide that responds to stress, (p)ppGpp, is both hydrolysed and synthesised (albeit more weakly) by SpoT proteins (*44*). Bifunctional GGDEF-EAL tandem domains, having retained all necessary catalytic residues in both domains throughout evolution, must be regulated by specific cues so that a GTP-wasting loop is prevented. They should be well-positioned to integrate intracellular and/or extracellular signals, changing their net c-di-GMP output in accordance to the organism's biology.

Selective reduction in catalytic activity can be achieved via two-component system signaling. Lpl0329 from *Legionella pneumophila* is a tandem protein with a REC domain, which selectively lowers DGC activity while maintaining constitutive PDE activity upon phosphorylation (*45*). The cognate HK, Lpl0330, contains a PAS domain necessary for its autophosphorylation, suggesting that the Lpl0330-Lpl0329 two-component system is regulated by a small molecule ligand. Opposing DGC-PDE activities can also be fine-tuned via interactions with a separate protein that receives the regulatory signal. The ScrABC system in *Vibrio parahaemolyticus* integrates quorum sensing and c-di-GMP signaling, controlling the switch between biofilm formation and swarming life style (*46*). The membrane-bound ScrC contains a cytoplasmic tandem module that displays predominantly DGC activity when not bound to its periplasmic binding partner, ScrB. At high cell density, ScrB binds to S-signal, an autoinducer produced by ScrA, and has increased binding affinity to ScrC. The increased ScrB-ScrC interaction propagates the extracellular signal into the cytoplasm by converting ScrC into a PDE. Another example that is medically important is MSDGC-1 (locus Rv1354c) from *Mycobacterium tuberculosis*, which is the only DGC in the pathogen. In *Mycobacterium smegmatis* which contains one orthologue, the GAF-GGDEF-EAL displays both DGC and PDE activities but only as a full-length protein. Knockout strains of *M. smegmatis* display severely impaired long-term survival under nutrient starvation, while overexpression altered the microbe's growth profile (*47*).

Tbd1265 from *T. denitrificans*, whose EAL domain structure has been solved (*24*) (PBD IDs 2R6O and 3N3T), is a transmembrane protein comprising a periplasmic binding protein (PBP) domain connected to a cytoplasmic GGDEF-EAL tandem via a

predicted single transmembrane helix. Works done in our group on the cytoplasmic fragment revealed its bifunctionality while the EAL activity has been previously established (*24*). The PBP superfamily has a common bi-lobal protein fold and binds to diverse ligands including amino acids, sugars, small inorganic ions, etc, serving as sensors that direct chemotaxis and solute uptake among other processes (*48*). A search for homologues by using *blastP* (*49*) indicated that there are only three other proteins of the PBP-GGDEF-EAL architecture, found in *Thiobacillus thioparus*, *Thioalkalivibrio* sp. ALJ17 and *Thioalkalivibrio sulfidophilus*, suggesting that this protein is specialised for life style regulation in these inorganic autotrophs. A structural outlook on how the GGDEF-EAL tandem is modulated by a periplasmic ligand would be interesting; our work in progress on Tbd1265 is presented in Paper II.

1.4.2 Half-active or completely inactive tandem proteins

Given the large percentage of inactive GGDEF domains found in tandem proteins, degenerate GGDEF domains may serve other functions than being a DGC. CC3396 from *C. crescentus* is a tandem protein containing one inactive GGDEF domain that binds GTP and activates its neighbouring EAL domain (*50*). GTP as an energy currency could thus be an important link between the organism's nutrition status and c-di-GMP signaling that responds accordingly. A combination of active GGDEF and inactive EAL can also be envisioned.

Also highly interesting are tandem proteins that are devoid of any DGC or PDE activity, being catalytically-incompetent proteins that serve sensory and regulatory roles or even functions unrelated to c-di-GMP pathways altogether. The membrane-bound protein LapD from *Pseudomonas fluorescens* contains a doubly-inactive cytoplasmic GGDEF-EAL tandem, a bridging HAMP domain, and a periplasmic domain that can interact with the extracellular protease LapG. Based on crystallographic studies (*51*), c-di-GMP binding to the degenerate EAL domain is proposed to cause large conformational changes in the LapD dimer, which is transmitted to the periplasmic domain "inside-out". This increases the periplasmic domain's affinity for LapG, effectively sequestering LapG from the surface adhesin LapA that promotes attachment to surfaces and other cells. Conversely, low c-di-GMP levels would trigger protease activities that aid biofilm dispersal. Large conformational changes that accompany c-di-GMP binding to an inactive EAL domain have also been

proposed for the protein FimX that regulates twitching motility in *P. aeruginosa* (*52, 53*).

1.5 CYCLIC DI-GMP RECEPTORS THAT INFLUENCE BACTERIAL ACTIVITY

Given the important functions of c-di-GMP signaling, a brief discussion of c-di-GMP receptors and their roles is necessary although Papers I and II do not directly relate to them. The most economical yet ingenious way to sense c-di-GMP would be employing GGDEF, EAL and HD-GYP domains, by evolving catalytically incompetent proteins that retain c-di-GMP binding sites. The allosteric inhibitory I-site in GGDEF not only allows product inhibition on DGC activity (*54*) but should also serve as a c-di-GMP binding site in inactive GGDEF domains (*2*). This strategy is seen in the *C. crescentus* response regulator PopA that promotes cell cycle progression (*55*), the *Bdellovibrio bacteriovorus* CdgA required for efficient prey invasion and predatory growth (*56*), and the *P. aeruginosa* PelD (not to be confused with PleD, an active DGC) that promotes Pel polysaccharide production and biofilm formation upon binding c-di-GMP (*57, 58*). Similarly, degenerate EAL domains are also utilised to bind and detect c-di-GMP. Two examples described in the previous section are LapD which regulates proteolysis of surface adhesions and FimX which regulates assembly of type IV pili required for twitching motility. While HD-GYP-based receptors have not yet been reported, their existence can surely be envisioned.

Another group of c-di-GMP receptors that can now be predicted by primary sequence is the PilZ domain receptors. The PilZ domain contains two sequence motifs, RxxxR and (D/N)x(S/A)xxG, separated by 20 to 30 amino acid residues. The arginine-flanked motif is necessary for c-di-GMP binding (*59*) and forms a loop around c-di-GMP, bringing other ligand-binding elements into place while also exposing potential protein interaction surfaces (*60*). This is analogous to the RxxR I-site motif in GGDEF domains. The *modus operandi* of PilZ domains can be diverse, as different c-di-GMP binding stoichiometry and oligomerisation states have been observed (*61*). Two notable PilZ-regulated proteins control flagellar motility of bacteria. YcgR from *E. coli* binds strongly to the flagellar motor complex at high c-di-GMP levels, changing the motor's conformation to reverse and slow down flagellar rotation (*62*). On the other hand, the cellulose synthase BcsA increases cellulose production and deposition under high c-di-

GMP, allowing cellulose to sterically hinder flagellar rotation (*63*). These mechanisms may facilitate biofilm formation and a sessile life style.

In addition to those predictable by bioinformatics, more c-di-GMP receptors are being uncovered at the moment. Some of these are involved in life style switching according to quorum sensing (*46*), including Clp, VpsR and CpsQ from *X. campestris*, *V. cholerae* and *V. paraheamolyticus*, respectively. They represent transcriptional regulators that act downstream of c-di-GMP-metabolising proteins controlled by autoinducer receptors. More surprisingly, RNA receptors for c-di-GMP have been found; highly conserved c-di-GMP-I ("GEMM") riboswitch domains (*64*) and c-di-GMP-II riboswitch domains (*65*) are located upstream of genes governing production of DGCs, PDEs, virulence factors, pili and flagella. Binding of c-di-GMP to these RNA regulatory elements alter mRNA stability, splicing and translation, thus exerting post-transcriptional control on the genes.

1.6 PEPTIDOGLYCAN AS AN IMPORTANT STRUCTURAL ELEMENT IN THE BACTERIAL CELL WALL

Most bacterial cells are encased by the sacculus, a crosslinked mesh of peptidoglycan (murein) biopolymers located outside the cytoplasmic or inner membrane (*66-68*). Tough but usually somewhat flexible, the peptidoglycan sacculus forms a cellular boundary that mechanically withstands turgor and provides stability against osmotic rupture. Cell death can be observed when peptidoglycan biosynthesis is impaired by antibiotics, and when degradation of peptidoglycan occurs in the presence of lysozymes, highlighting the protective role of the sacculus. Also, the well-defined size and shape of most bacterial cells are at least in part attributed to the sacculus and the enzymes that shape it; peptidoglycan biosynthesis and degradation is regulated according to stages of bacterial growth and replication. Gram-negative bacteria have a 3 – 6 nm-thick layer of peptidoglycan in the periplasmic space, while the peptidoglycan surrounding Gram-positive bacterial cell membrane is more than 10 – 20 nm thick (allowing it to be Gram-stained) and additionally contains other polymers such as teichoic acids (*69*).

Peptidoglycans can be pictured as a high molecular weight molecule, consisting of long polysaccharide glycan chains crosslinked by short peptides (*68, 70*). Each glycan chain is a polymer of alternating *N*-acetylglucosamine (NAG) and *N*-acetylmuramic acid

(NAM) bonded via β-1→4 glycosidic bonds. The lactoyl moiety of NAM is substituted by a peptide stem – an oligopeptide consisting of unusual amino acids whose sequence may be phylogenetically dependent. The most common arrangement is L-Ala-γ-D-Glu-*meso*-DAP (or L-Lys)-D-Ala-D-Ala, where DAP is 2,6-diaminopimelic acid. The carboxyl group of the position-4 D-Ala and the amino group of the DAP are usually involved in peptidoglycan linkage. The position-5 D-Ala is lost in the fully integrated peptidoglycan subunit. These structural details may vary not only across bacterial species but also in different growth conditions. The crosslinking between glycan strands also observes a general rule of linking the carboxyl group of position-4 D-Ala and the amino group of the position-3 diamino acid, forming a peptide bond directly (in most Gram-negative bacteria) or via a short oligopeptide bridge (Gram-positive species).

The formation of the peptidoglycan polymer can be divided into cytoplasmic, membrane and extracellular stages (*70*). In the cytoplasmic stage, the necessary sugars and special amino acids are synthesised and assembled into the activated precursors, UDP-*N*-acetylglucosamine (UNAG) and UDP-*N*-acetylmuramyl (UNAM) pentapeptide. These UDP-activated precursors are joined to form lipid II, a membrane-bound disaccharide pentapeptide subunit covalently linked to an undecaprenyl pyrophosphate chain. This molecule is then flipped into the outer leaflet of the cell membrane or the inner membrane. Lastly in the extracellular stage, the disaccharide pentapeptide is incorporated into a growing glycan chain by the formation of glycosidic bonds, followed by the crosslinking of peptide stems. Each stage involves a large number of proteins playing regulatory, structural and enzymatic roles.

1.7 CYTOPLASMIC ENZYMES THAT SYNTHESISE PEPTIDOGLYCAN PRECURSORS

The cytoplasmic stage of peptidoglycan biosynthesis involves an array of transferases, isomerases, oxidoreductases, etc., resulting in the formation of UNAG and UNAM-pentapeptides (*66, 67*). As the enzymes are essential for cell wall biosynthesis and hence drug targets, their mechanisms, ligand interactions and structural properties are valuable for structure-aided development of novel inhibitors. The syntheses of UNAG and UNAM-pentapeptides overlap, with the latter having more extensive reaction steps (Fig 1.5). UNAG is synthesised from fructose-6-phosphate, glutamine, acetyl-CoA and UTP via four enzymes, namely glucosamine-6-phosphate synthase (as the protein

GlmS), phosphoglucosamine mutase (as GlmM), glucosamine-1-phosphate acetyltransferase and finally NAG-1-phosphate uridyltransferase (both encoded in the bifunctional GlmU). GlmS contains two domains; the glutaminase domain hydrolyses glutamine into glutamate and ammonia, while the isomerase domain converts fructose-6-phosphate into glucosamine-6-phosphate using ammonia as the nitrogen source. The isomerisation of glucosamine-6-phosphate to glucosamine-1-phosphate is catalysed by GlmM, a hexosephosphate mutase that becomes active when a conserved serine residue is phosphorylated. Then, glucosamine-1-phosphate is acetylated by the C-terminal acetyltransferase domain of GlmU, producing NAG-1-phosphate which is subsequently uridylated by the N-terminal uridyltransferase domain.

The UNAG product resulting from the above reactions can take part in lipid II formation or be further modified to yield UNAM (Fig 1.5). The enzyme MurA transfers one enolpyruvyl group from phosphoenolpyruvate to the 3'-hydroxyl group of UNAG; enolpyruvyl transfer from phosphoenolpyruvate to a hydroxyl accompanied by phosphate release is so far a rarely seen reaction (66). The product, UDP-*N*-acetylglucosamine-enolpyruvate (UNAGEP), is subsequently reduced by the oxidoreductase MurB using nicotinamide adenine dinucleotide phosphate (NADPH) as a hydride donor to form UNAM. Since UNAG can be channelled to other pathways, the reactions catalysed by MurA and MurB represent a committed step towards peptidoglycan production. They are also the target of various natural and artificial inhibitors. The resulting UNAM then undergoes stepwise addition of amino acids to form UNAM pentapeptide by the action of the Mur ligases (MurC, MurD, MurE and MurF). The Mur ligases share a common reaction mechanism, six essential residues, ATP-binding motif and a common protein fold. MurC is responsible for the ATP-dependent addition of L-Ala to the carboxyl group in the lactoyl moiety. Then, MurD, MurE and MurF catalyse the stepwise addition of D-Glu, *meso*-DAP (most Gram-negative and Bacilli species) or L-Lys (most Gram-positive species), and the D-Ala-D-Ala dipeptide, respectively. As mentioned before, the exact amino acid composition of the peptide stem depends on bacterial species and growth conditions. Variations may also give rise to antibiotic resistance, such as vancomycin resistance in bacterial strains employing D-Ser or D-Lac at position 5. In addition to the linear pathway outlined above, several other enzymes are needed to provide necessary building blocks. The glutamate racemase MurI and the Alr/DadX alanine racemases produce D-Glu and D-Ala, respectively, while the D-Ala:D-Ala ligase Ddl catalyses the condensation between

16

two D-Ala molecules to yield the D-Ala-D-Ala dipeptide. Last but not least, *meso*-DAP can be derived from the lysine synthetic pathway.
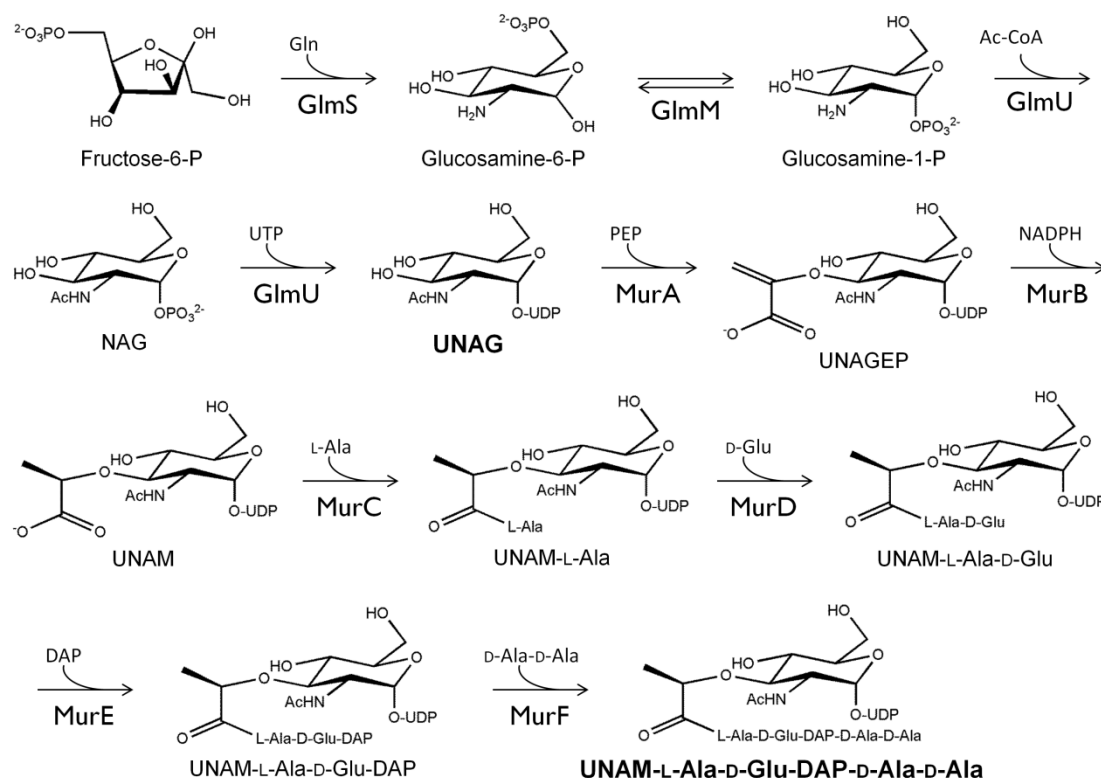
Figure 1.5. Cytoplasmic steps of peptidoglycan biosynthesis. UNAG and UNAM-pentapeptide are precursors for peptidoglycan polymers. The pentapeptide composition found in *P. aeruginosa* is shown here. Please see text (Section 1.7) for detailed description.

1.8 MurB SYNTHESISES UDP-N-ACETYLGLUCOSAMINE-ENOLPYRUVATE

The UNAGEP reductase MurB synthesises UNAM by facilitating the hydride transfer from an NADPH cosubstrate to UNAGEP (*71, 72*). All currently known examples of bacterial MurB exist and function as monomers in solution (*66, 67*). Flavin adenine dinucleotide (FAD) is required as a cofactor and binds to the protein in a 1:1 stoichiometry. A common three-domain architecture is observed in crystal structures of MurB enzymes from various species, including *E. coli* (EcMurB, PDB code 1MBT) (*73*), *Staphylococcus aureus* (SaMurB, 1HSK) (*74*), *Thermus caldophilus* (TcMurB, 2GQT) (*75*), *Listeria monocytogenes* (LmMurB, 3TX1), *V. cholerae* (VcMurB, 3I99), and *P. aeruginosa* (PaMurB, 4JAY and 4JB1) (*76*). The three domains, Domains I, II and III, are named by their order in primary sequence. Domains I and II are conserved and compose an FAD-binding didomain module that is also present in other oxidoreductases such as cytokinin dehydrogenase (*77*) and vanillyl-alcohol oxidase

(*78*). They form a shielded FAD-binding cleft by providing extensive interactions with the cofactor molecule. On the other hand, Domain III contains two lobes that form a substrate channel as first evidenced in the EcMurB-UNAGEP complex crystal structure (2MBR) (*79*). Comparison of MurB structures and sequences reveals that Domain III can be divided into two types based on the presence or absence of a tyrosine loop and a βαββ protrusion; type I MurB enzymes possess them and are exemplified by EcMurB, while type II members such as SaMurB and TcMurB lack those features. All MurB enzymes contain a proton donor residue essential for UNAGEP reduction; type I and type IIa MurB utilise a serine residue for that purpose (*74, 80, 81*) while a cysteine is used in type IIb MurB (*75*).

The reaction kinetics of MurB have been well-studied by various groups, confirming that NADPH is used as a hydride donor but not NADH (*71, 72*). A ping-pong bi bi mechanism is employed; each catalytic cycle consists of a first half reaction where NADPH is oxidised and a second half reaction where UNAGEP is reduced, giving $NADP^+$ and UNAM as end products (Fig 1.6). For so far unclear reasons, the enzyme is strongly activated by monovalent cations such as K^+ and NH_3^+ (*72, 82*). By means of isotope tracing, the 4-*pro*-S hydrogen but not the 4-*pro*-R hydrogen is abstracted during the first half reaction (*71*). In the following half reaction, the hydride is added to the C-3 atom of the enolpyruvyl group in UNAGEP (C3e), creating a carbanionic intermediate that is finally protonated at C-2 to complete the reduction (*66, 67*). The EcMurB-UNAGEP crystal structure shows a close packing of the enolpyruvyl moiety to the *si* face isoalloxazine ring of the flavin cofactor, which brings C3e close (approximately 3 Å) to the hydride-accepting N5 atom of the isoalloxazine (*79*). Thus, the observed UNAGEP binding conformation may be close to the catalytically competent enzyme-substrate complex. In contrast, no experimental structure of a MurB-FAD-NADPH complex had been published prior to our work with PaMurB (Paper III) (*76*). The binding mode of NADPH to MurB during the first half reaction had been speculated. The co-substrate could bind to the same substrate site occupied by UNAGEP, requiring the product of every half reaction to dissociate from the enzyme before the next substrate can access the active site. Alternatively, NADPH could bind to a perpendicular channel close to the protein's C-terminus. Both options result in the unusual alignment of the nictotinamide group to the *si* face of isoalloxazine, while binding to the *re* face is more common (*73*). In order for NADPH to approach the *re* face of the flavin, significant rearrangements at the Domain-II-Domain-III interface are

18

required to provide another putative NADPH site perpendicular to the UNAGEP site (*79*). The first scenario began to be favoured when solution NMR studies of the EcMurB-NADP$^+$ complex suggested that NADPH binds to the same substrate channel of UNAGEP, possibly because of rearrangements in the protein (*83, 84*). The relevance of this proposal based on NADP$^+$ is strengthened by the observation that NADP$^+$ inhibits MurB with regard to both NADPH and UNAGEP (*72*), i.e. the three molecules may occupy the same binding site. Also, NADP$^+$ can reverse the reduction of FAD in MurB (*85*), suggesting that NADP$^+$ and NADPH have a similar binding mode at least for the nicotinamide functional group. In Paper III, two crystal structures of PaMurB in complex with NADP$^+$ are presented, resolving the question of the NADPH binding site and also other mechanistic details such as metal-dependent activation.
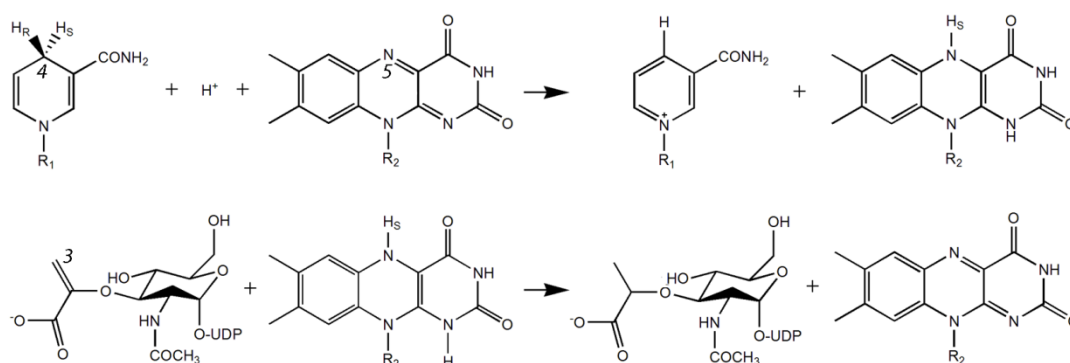
Figure 1.6. FAD- and NADPH-dependent reduction of UNAGEP by MurB enzymes. Two half reactions take place in a ping pong bi bi mechanism. NADPH is the first substrate that transfers a 4-*pro*-S hydride from the nicotinamide C4n atom to the N5 atom of the isoalloxazine. In the second half reaction, the hydride is transferred to the C3e atom of the enolpyruvyl group on UNAGEP, producing UNAM.

1.9   MAGIC BULLET AGAINST MURB: AN ONGOING SEARCH

Given the growing problem of bacterial resistance against traditional antibiotics and the essentiality of peptidoglycans to most pathogenic species, there is ongoing effort to discover natural and synthetic inhibitors against the cytoplasmic enzymes discussed above (*67, 86*). Such inhibitors would represent a different antimicrobial strategy, compared to the more traditional penicillins and cephalosporins that target the extracellular peptidoglycan enzymes. The availability of crystal structures and spectrophotometric enzyme assays for these proteins has been taken advantage of in the pursuit of structure-guided drug design and high-throughput screening. An additional advantage of targeting the Mur enzymes is the structural and functional similarity among the MurC – F ligases, which may facilitate parallel screening.

The UNAGEP reductase MurB displays distinct advantages as a drug target. It is essential for bacterial cell growth, present in both Gram-positive and Gram-negative species and yet has no eukaryotic counterpart (*87*), raising the possibility of a broad-spectrum bacteriocide with little side effect in medical and agricultural use. Bristol-Myers Squibb synthesised and reported a series of imidazolinone analogues that both inhibit MurB *in vitro* and display antibacterial activity against *S. aureus*. A group of 3,5-dioxopyrazolidines with IC_{50} values of 4 – 10 µM against methicillin-resistant *S. aureus*, penicillin-resistant *Streptococcus pneumoniae* and vancomycin-resistant *Enterococcus faecalis* were subsequently reported by Wyeth; the compounds were also shown to decrease peptidoglycan synthesis (*88*). As a result of screening against an array of Mur enzymes, naphthyl tetronic acids were also found to inhibit MurA – E with varying IC_{50} values (*89*), with excellent binding to MurB. Attempts to co-crystallise MurB with the compounds produced a crystal structure where the inhibitor occupies the UNAGEP binding site and interacts with the flavin (PDB code 2Q85). More work is however required to develop functional drugs that show good pharmacokinetics and good permeability across the bacterial cell wall.

2 AIMS OF THIS THESIS

Presently available bioinformatics and biochemical knowledge on c-di-GMP metabolism allows us to rapidly discover novel pathways in organisms of interest. However, the structural biology of biomolecules involved in c-di-GMP signalling has also vast unchartered territories. A quick search in Web of Science (Thomson Reuters) shows that 88 original research articles on c-di-GMP have been published between January 2013 to early November 2013 alone; only three and six of them are grouped as "Crystallography" and "Biophysics", respectively. More structural characterisation is needed to investigate how DGC and PDE proteins are regulated even as the focus shifts towards c-di-GMP receptors. In this thesis the structural basis of regulation of the DGC and PDE domains, especially in the case of GGDEF-EAL tandem domains that are predicted to show bifunctionality was addressed. Towards this end, the response regulator RocR and the cytoplasmic and periplasmic parts of the transmembrane protein Tbd1265 were studied. The aim was to derive mechanistic models of regulation based on a combination of X-ray crystallography, small–angle X-ray scattering and biochemical studies.

Structural studies of MurB, an enzyme essential for peptidoglycan biosynthesis in bacteria, were ultimately aimed to facilitate structure-aided drug discovery. Despite speculations on the NADP(H) binding site of MurB, no experimental structure of the MurB-FAD-NADP(H) complex had been reported before Paper III. The primary aim of this study was to establish the binding mode of the co-substrate NADPH and compare it to the binding site of the second substrate, UDP-*N*-acetylglucosamine-enolpyruvate.

3 RESULTS AND DISCUSSION

3.1 STRUCTURAL STUDIES OF ROCR, A PHOSPHORECEIVER-EAL PROTEIN FROM *PSEUDOMONAS AERUGINOSA* (PAPER I)

The PvrR family of EAL proteins with the REC-EAL domain arrangement regulates intracellular c-di-GMP levels according to external stimuli. Its presence in various pathogenic genera such as *Bordetella*, *Vibrio* and *Pseudomonas* also suggests its role in the pathogens' adaptation to their environment. In order to elucidate the regulatory mechanism of REC-EAL proteins, we obtained the crystal structure of RocR (PA3947) from *P. aeruginosa* and compared it to solution structures acquired using small-angle X-ray scattering (SAXS). With the help of biochemical studies, we derived a mechanistic model of signal propagation from the REC domain to the EAL active site.

3.1.1 Recombinant protein production of RocR-R286W, RocR-D56N and RocR-wt

Three RocR variants have been employed in this work: wildtype RocR (RocR-wt) and the R286W (RocR-R286W) and D56N (RocR-D56N) mutants. Cloning procedures have been described for RocR-wt and RocR-R286W (*90*), where the mutant was a PCR by-product. The RocR-D56N expression construct was made previously with a quick mutagenesis protocol (*91*). Recombinant expression in *E. coli* BL21(DE3) yielded good amounts of protein that were subsequently purified and concentrated to 10 – 15 mg/mL. A selenomethionine-substituted form of RocR-R286W was also prepared by using *E. coli* B834(DE3) grown in premixed selenomethionine expression medium (Molecular Dimensions).

3.1.2 Crystallisation and preparation of heavy metal derivatives

While RocR-wt resisted crystallisation, RocR-R286W serendipitously crystallised under a PEG 3350-based condition via the sitting-drop vapour diffusion method. The crystals were solid hexagonal rods with tapered ends. Attempts to co-crystallise the protein and c-di-GMP resulted in overnucleation and small crystals that did not yield useful diffraction data. Seleniated RocR-R286W was crystallised in the same condition albeit with lowered precipitant concentration. Prior to this work, the traditional overnight-soaking method was used to obtain heavy metal derivatives but failed to yield data usable for experimental phasing. Hence, we attempted the "quick soak"

method (*92*) advocating short soaking times (10 min) under high concentrations of heavy metal compounds (10 mM). Native crystals were transferred into drops containing desired compounds but devoid of sodium tartrate, a salt found in the original reservoir solution, to perform derivatisation.

3.1.3 Structure determination of RocR-R286W

We obtained diffraction data for crystals grown with the selenomethionine-substituted protein and also from crystals containing potassium tetrachloroplatinate(II). These two sets of data were used alongside a previ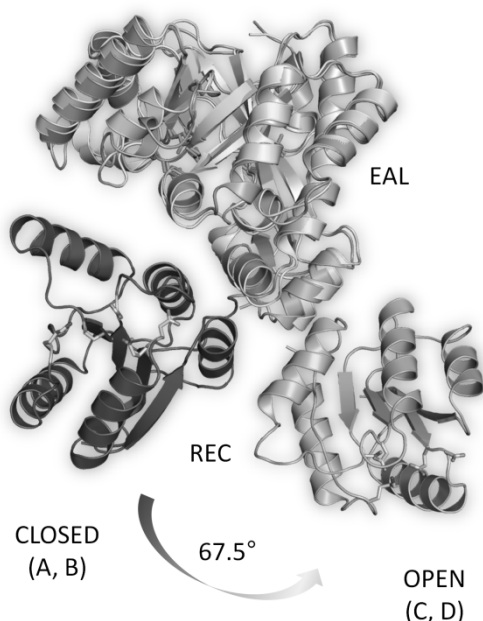ously collected native dataset in order to solve the structure by experimental phasing. Four Pt sites were found from the tetrachloroplatinate(II) dataset but the resulting electron density maps were not interpretable. However, the phases were used to identify 38 out of 40 possible Se sites in the selenomethionine data. SAD phasing at the Se absorption edge then allowed density improvement and model building with the native dataset. The protein model was refined to 2.5 Å as a tetrameric structure occupying the asymmetric unit (ASU) (Fig 3.1). The RocR tetramer can be seen as a dimer of dimers arranged around a non-crystallographic dyad. The quaternary structure is maintained by extensive interaction interfaces between the subunits, including two EAL-EAL dimerisation interfaces that are identical to previously published EAL protein structures (*93-95*). Curiously, the REC domains of monomers A and B are located at the core of the tetramer, with the phosphorylation site containing Asp-56 facing away from bulk solvent. Meanwhile REC_C and REC_D are more solvent-accessible, and thus more likely to be the target of histidine kinase RocS1. However, no electron density indicating phosphorylation was observed in any of the four REC active sites.

Figure 3.1. Overall structure of RocR. The RocR tetramer and the non-crystallographic dyad are shown (PDB code 3SY8). The left panel shows monomers C (black) and D (white) in the open conformation, while the right panel shows the closed monomers A (black) and B (white).
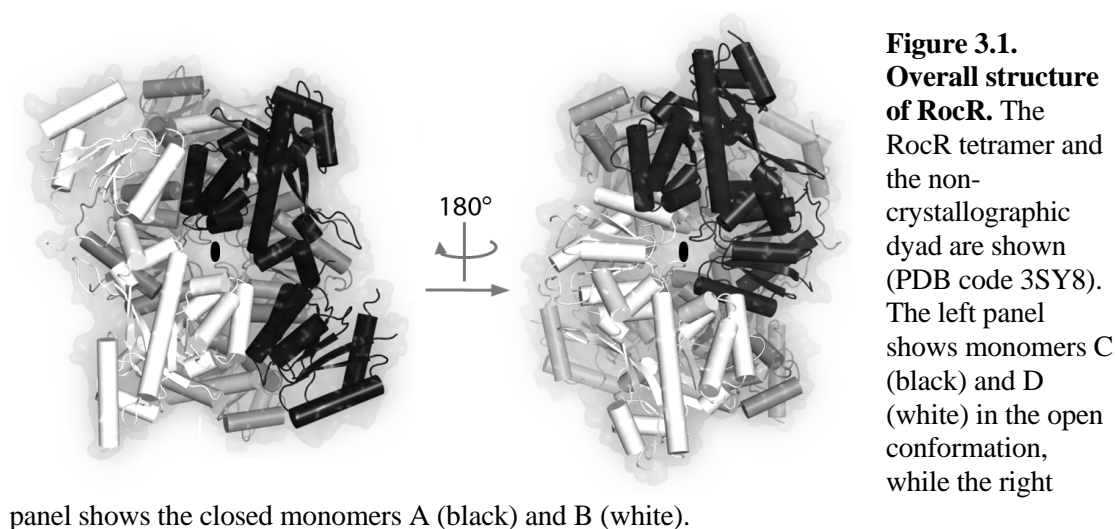
3.1.4 Open and closed conformations in the RocR tetramer

Both REC-EAL dimers (labelled AC and BD) are asymmetric as the constituent monomers adopt different conformations: monomers A and B have a "closed" conformation with the REC domain folded towards the active site of the EAL domain, while monomers C and D display an "open" conformation. Superimposing the closed and open forms of the REC-EAL monomer reveals that a large rotational and translational movement is needed to switch between the two conformations (Fig 3.2). This is made possible by the flexibility of the REC-EAL linker, which has weak and non-interpretable electron density in all but one of the monomers.

Figure 3.2. Comparison of the open and closed conformations of REC-EAL. Monomers A and B adopt the closed conformation where the REC domain is close to the apical face of the EAL barrel, while monomers C and D have REC domains that swing away. The residual rotation needed to switch between conformations is indicated.

Also, as a result of the two conformations and the tetrameric arrangement, EAL active sites of monomers C and D are occluded by REC_B and REC_A, respectively (Fig 3.3). The interaction is mediated by helices α4' and α5' of the REC domain, which make van der Waals contacts and at least one hydrogen bond with the EAL active site in question. Binding of c-di-GMP is not possible as a result. This means that the EAL dimers observed in this structure inhibit each other in *trans* at the EAL active sites of open monomers, as the pair are non-crystallographically related. In addition, this is accompanied by sequestration of REC_A and REC_B phosphorylation sites as mentioned earlier. One can envision that the inhibition can be removed if monomers A and B also adopt the open conformation, at the cost of a large conformational change and possibly breaking many protein-protein interactions.
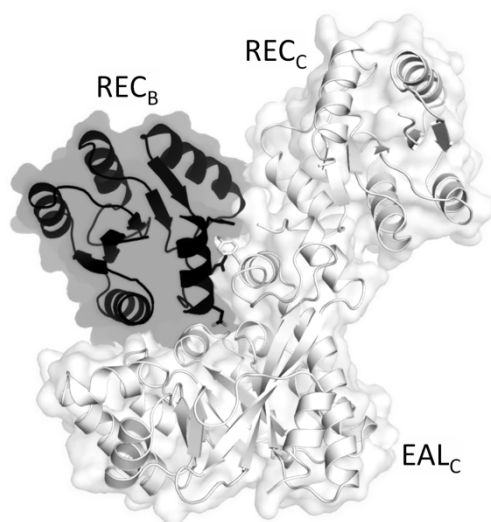
Figure 3.3. In *trans* inhibition of c-di-GMP binding. By steric occlusion, REC_B/REC_A (depicted as black) prevents EAL_C/EAL_D (in white) to function as a PDE as the EAL active site is located on the interface.

3.1.5 Comparison of the RocR variants by using small-angle X-ray scattering

Given the intriguing interplay between the subunits of RocR, solution studies of the protein were carried out in order to investigate the relationship between the RocR structure and its catalytic (PDE) activity. We carried out SAXS experiments to derive single-particle parameters of the protein such as the maximum particle size, and the averaged low-resolution shape of the protein particle. Since the crystal structure was derived from RocR-R286W, the same experiments were also performed on RocR-wt. RocR-R286W displays lowered PDE activity at room temperature but recovers at higher temperatures to be comparable to RocR-wt. More importantly, we measured RocR-D56N, a phosphorylation-site mutant which displays lowered k_{cat} compared to RocR-wt (*91*) for comparison. Surprisingly, the scattering patterns of all three RocR variants coincide, suggesting that a change in catalytic output may not require large conformational changes. Furthermore, the experimental scattering fits closely to the theoretical scattering computed from the crystal structure ($\chi = 1.08$), in stark contrast to hypothetical models where three monomers are open ("half-open" model, $\chi = 4.7$) and where all monomers are open ("open" model, $\chi = 11.3$) (Fig 3.4A). The experimental particle size also agrees with the crystal structure, with a matching particle envelope (Fig 3.4B). Taken together, the results suggest that the RocR structure is stable in solution; the EAL domains may be regulated by secondary structural changes that do not alter the overall quaternary structure of RocR.
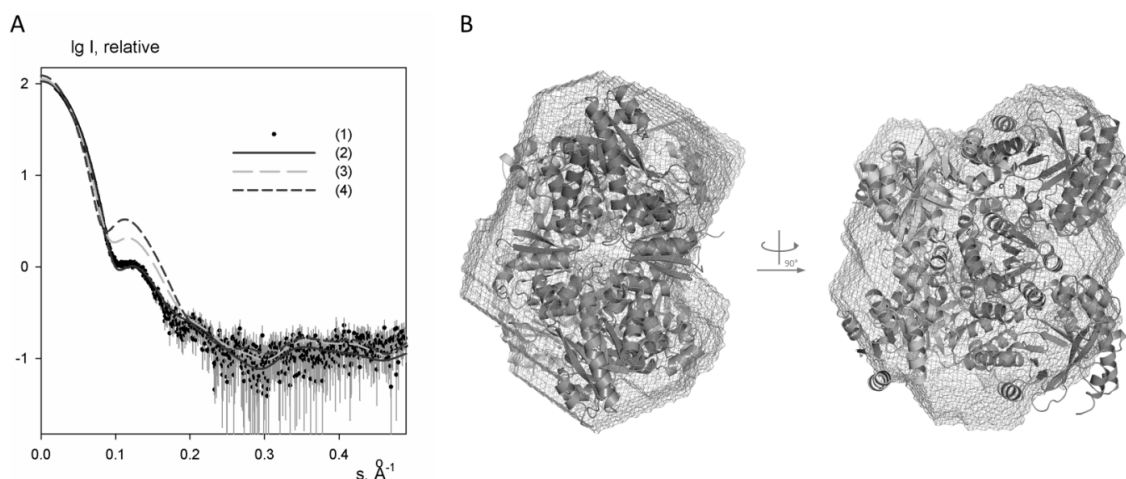
Figure 3.4. SAXS analysis of RocR in solution. (A) Comparison of RocR-wt experimental scattering (1, dots) against calculated scattering curves from the crystal structure (2, solid line overlapping with curve 1), the half-open model (3, grey dashes) and the open model (4, dark grey dashes). (B) Orthogonal views of the low-resolution protein envelope reconstructed from RocR-wt data, superimposed onto the crystal structure.

3.1.6 Potential mechanism of RocR regulation upon phosphorylation

As concluded in the previous section, large conformational changes of RocR may not be required for fine-tuning of PDE activity. However, the structure of the catalytically essential loop 6 comprising part of the EAL-EAL dimer interface (Fig 1.3) might be at play. Hydrogen-deuterium exchange studies on RocR-wt and RocR-D56N (91) revealed altered solvent exposure in loop 6 of the PDE-impaired mutant, suggesting a link between loop 6 conformation and PDE activity. Meanwhile, REC_C and REC_D are the most likely candidate for phosphorylation based on the RocR crystal structure. As the apical face of the REC_C/REC_D domain (which contains the phosphorylation site) is in contact with the dimer interface of EAL_B-EAL_D/EAL_A-EAL_C, a route of signal propagation from the REC phosphorylation site to loop 6 of the target EAL domain can be proposed. Using REC_C and EAL_B-EAL_D as an example, phosphorylation of Asp-56 could induce structural changes in helices α4' and α5', which has been observed in other REC proteins (96). This is transmitted to EAL_D via Phe-310 and Pro-311, which is connected to loop 6 (residues 295 to 305). However, EAL_D may not be the target of regulation as it is inhibited by REC_A. Instead, the structural changes are further propagated to the loop 6 of EAL_B across the EAL dimer interface, since the two loops interact. This may ultimately affect EAL_B activity through magnesium binding, c-di-GMP binding etc. The proposed mechanism would agree with our solution studies and previous hydrogen-deuterium exchange results, implicating secondary structural changes on a local scale for PDE regulation.
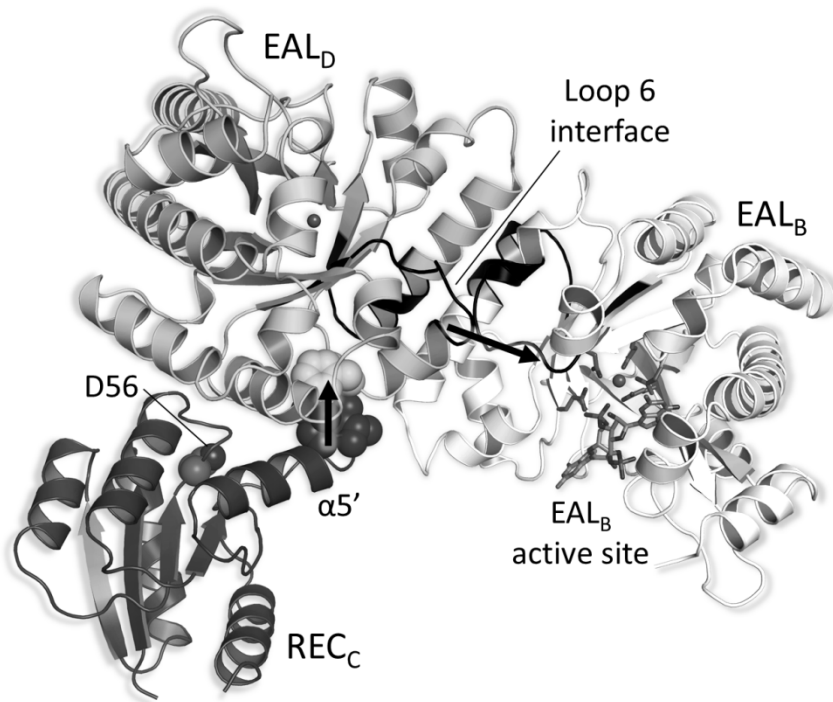
Figure 3.5. Proposed regulatory mechanism of the EAL domain by the REC domain. Phosphorylation of Asp-56 in REC_C causes structural changes to transmit from helix α5' to the adjacent Phe-310 and Pro-311 of EAL_D. The connected loop 6 then further relays the structural change to the neighboring loop 6 of EAL_B, ultimately altering the activity of EAL_B. Magnesium ions are shown as spheres while c-di-GMP (modelled in) is shown as a stick model to indicate EAL active sites.

Interestingly, the light-regulated EAL protein BlrP1 is also regulated via a sensor-regulator BLUF domain adjacent to a loop 6. Comparison of BlrP1 against RocR shows that their respective BLUF and REC domains occupy similar positions, with the signalling interfaces contacting an EAL dimer interface. Thus, this strategy may be adopted by EAL proteins of various sensor and regulatory domains.

3.2 STRUCTURAL STUDIES OF TBD1265, A PERIPLASMIC SENSOR PROTEIN WITH A GGDEF-EAL CYTOPLASMIC MODULE (PAPER II, MANUSCRIPT IN PROGRESS)

While cytoplasmic regulators of c-di-GMP such as RocR (*35, 36*) and RpfG (*30*) rely on transmembrane histidine kinases to perceive extracellular cues, some c-di-GMP-metabolising modules are fused directly to extracellular sensor domains as part of a transmembrane protein. Tbd1265 from *T. denitrificans* is such an example, consisting of an N-terminal periplasmic binding protein (PBP) domain, a transmembrane helix, a cytoplasmic α-helical region and a GGDEF-EAL tandem domain (Fig 3.6). Sequence analysis indicates that essential residues in both domains are preserved, suggesting that both DGC and PDE activities are possible. Thus, the net c-di-GMP output of this protein may be regulated by ligand binding at the PBP, which would propagate a signal transmitted across the periplasmic membrane to the cytoplasmic module. So far, no experimental protein structure has been reported for tandem domains with biochemically demonstrated bifunctionality. We acquired the crystal structure of the tandem domain where the GGDEF domains do not form a catalytically competent dimer, indicating a resting state for the DGC. By comparing structural and biochemical studies, we propose a model where PDE and DGC activities are fine-tuned based on extracellular signalling. Our ongoing work on the PBP will also be discussed.
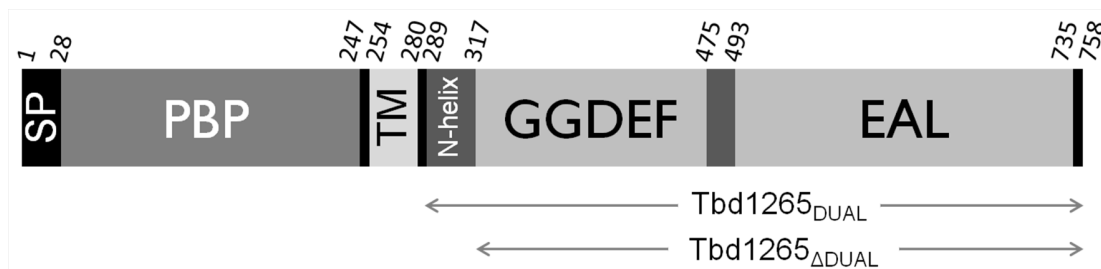
Fig 3.6. Domain organisation of Tbd1265. The protein is predicted to span the inner membrane with a single transmembrane (TM) helix. The predicted signal peptide (SP), PBP domain, cytoplasmic α-helix (N-helix), GGDEF domain, GGDEF-EAL linker and EAL domain are indicated. Construct boundaries for the $Tbd1265_{DUAL}$ and $Tbd1265_{\Delta DUAL}$ tandem domain constructs are also shown.

3.2.1 Recombinant production of GGDEF-EAL constructs $Tbd1265_{DUAL}$ and $Tbd1265_{\Delta DUAL}$

Two GGDEF-EAL tandem domain constructs, $Tbd1265_{DUAL}$ (residues 289 – 758) and $Tbd1265_{\Delta DUAL}$ (residues 317-758) of the *T. denitrificans* Tbd1265 were cloned into the pET-26b vector. They differ in the presence of residues 289 to 316, which are predicted to form a long helical region at the N-terminus of the GGDEF domain ("N-helix"). The

N-helix is predicted by the programs NCOILS (*97*) and PairCoil2 (*98*) to form coiled coils, a feature that may have signalling roles. The expression constructs allowed recombinant production of the proteins in *E. coli* BL21 (DE3) with good yields. The proteins were purified with protease inhibitors added during cell lysis, and were stored in a stabilising buffer containing 500 mM NaCl and 10% v/v glycerol.

3.2.2 In-solution oligomerisation studies of GGDEF-EAL constructs

Purified $Tbd1265_{DUAL}$ and $Tbd1265_{\Delta DUAL}$ were compared by analytical size-exclusion chromatography (SEC) with and without added c-di-GMP. Both proteins showed a single peak in the chromatograms (Fig 3.7a), with $Tbd1265_{DUAL}$ (51.8 kD per monomer) eluting slightly earlier than $Tbd1265_{\Delta DUAL}$ (48.2 kD per monomer). The elution volumes suggest a dimeric state for both constructs. When c-di-GMP and magnesium chloride were added to the proteins (Fig 3.7b, c), the elution profiles shifted slightly towards lower molecular weight, suggesting that the ligand may cause a more compact protein structure.
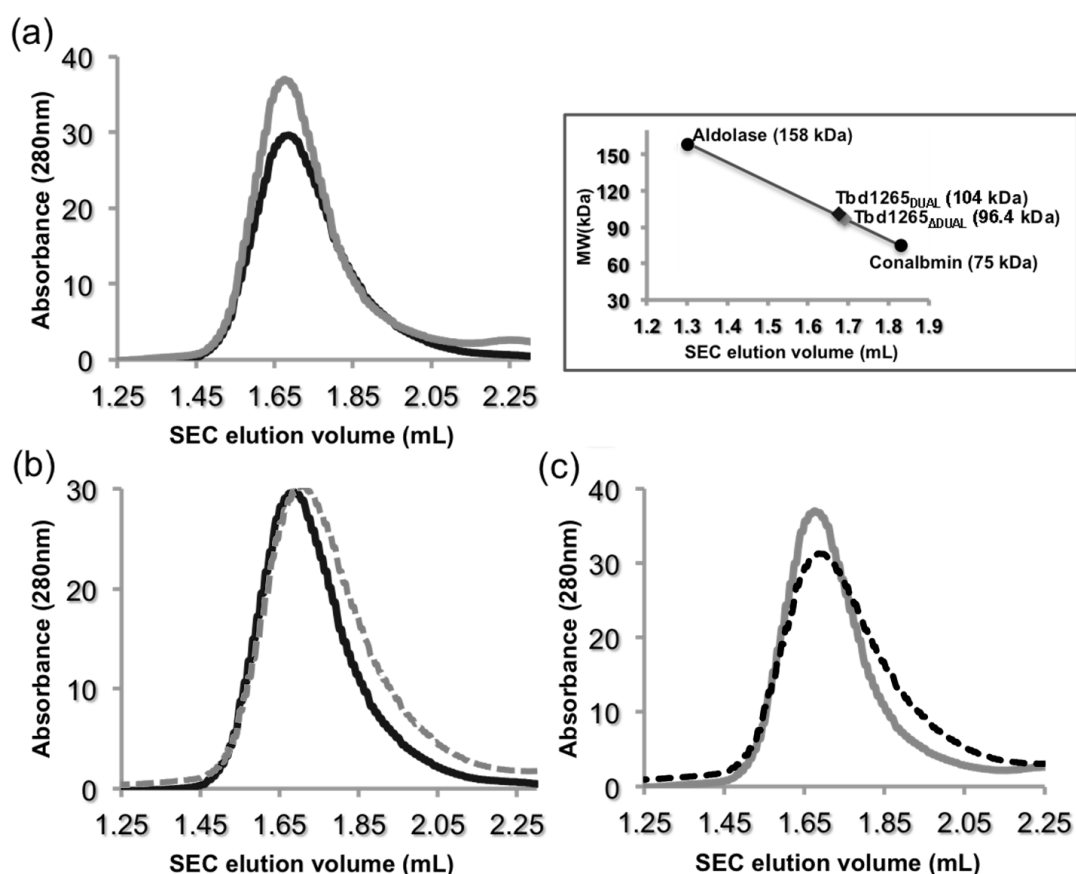
Figure 3.7. SEC profiles of the GGDEF-EAL tandem domain constructs. (a) Both $Tbd1265_{DUAL}$ (black line) and $Tbd1265_{\Delta DUAL}$ (grey line) elute as single peaks, with elution volumes corresponding to their dimeric size. Elution is slightly delayed in the presence of c-di-GMP for both $Tbd1265_{DUAL}$ (b, grey dashed line) and $Tbd1265_{\Delta DUAL}$ (c, black dashed line), suggesting a more compact structure upon ligand binding.

We then repeated the comparative experiment by using sedimentation-equilibrium analytical ultracentrifugation, which also confirmed the dimeric nature of both constructs in the presence or absence of c-di-GMP (Fig 3a – d, Paper II). Sedimentation is apparently greater for samples containing c-di-GMP (Fig 3e, f, Paper II), which may indicate a more compact protein structure as deduced from the SEC results.

3.2.3 Comparison of DGC and PDE activities of the tandem domain constructs

A dye-labelled sensor method was used to study the DGC and PDE activities of the two protein constructs in real time. The biosensor consists of the FimX EAL domain conjugated to the fluorescent dye 484-MDCC, which gives a titratable decrease with c-di-GMP binding and real-time readout of c-di-GMP concentrations as binding causes secondary structural changes in the sensor protein (*99*). The results confirm that both the GGDEF and EAL domains possess catalytic activity. Surprisingly, Tbd1265$_{DUAL}$ displays increased DGC and PDE activities compared to Tbd1265$_{\Delta DUAL}$ (Fig 4, Paper II). Thus, the N-helix which is present in Tbd1265$_{DUAL}$ but not in Tbd1265$_{\Delta DUAL}$ may play a role in regulating the tandem domain's activity, possibly by forming a coiled-coil bringing the two GGDEF domains into close proximity. It would be interesting to compare the extent of change in activities in the two domains due to the action of the N-helix; we are working on more detailed enzymatic characterisation of the tandem domains by using complementary methods.

3.2.4 Crystallisation and structure determination

Small-format crystallisation screening of the protein constructs led to an NaCl-based crystallisation condition for Tbd1265$_{DUAL}$. The rod-shaped crystals were grown in sitting drops by vapour diffusion. X-ray diffraction experiments yielded diffraction data of moderate resolution (3.4 Å), from which the crystal structure was solved by molecular replacement. This was facilitated by the previously published Tbd1265 EAL domain structure (PDB code 2R6O) (*95*) and the McR174C GGDEF domain (3ICL) which were used as search models. The structure contains five Tbd1265$_{DUAL}$ subunits per ASU, two dimers and one subunit that forms a dimer across the crystallographic twofold axis.

Each Tbd1265$_{DUAL}$ molecule shows an extended conformation where the GGDEF domain is located on top of the EAL TIM barrel (Fig 3.8a). Tbd1265$_{DUAL}$ forms a
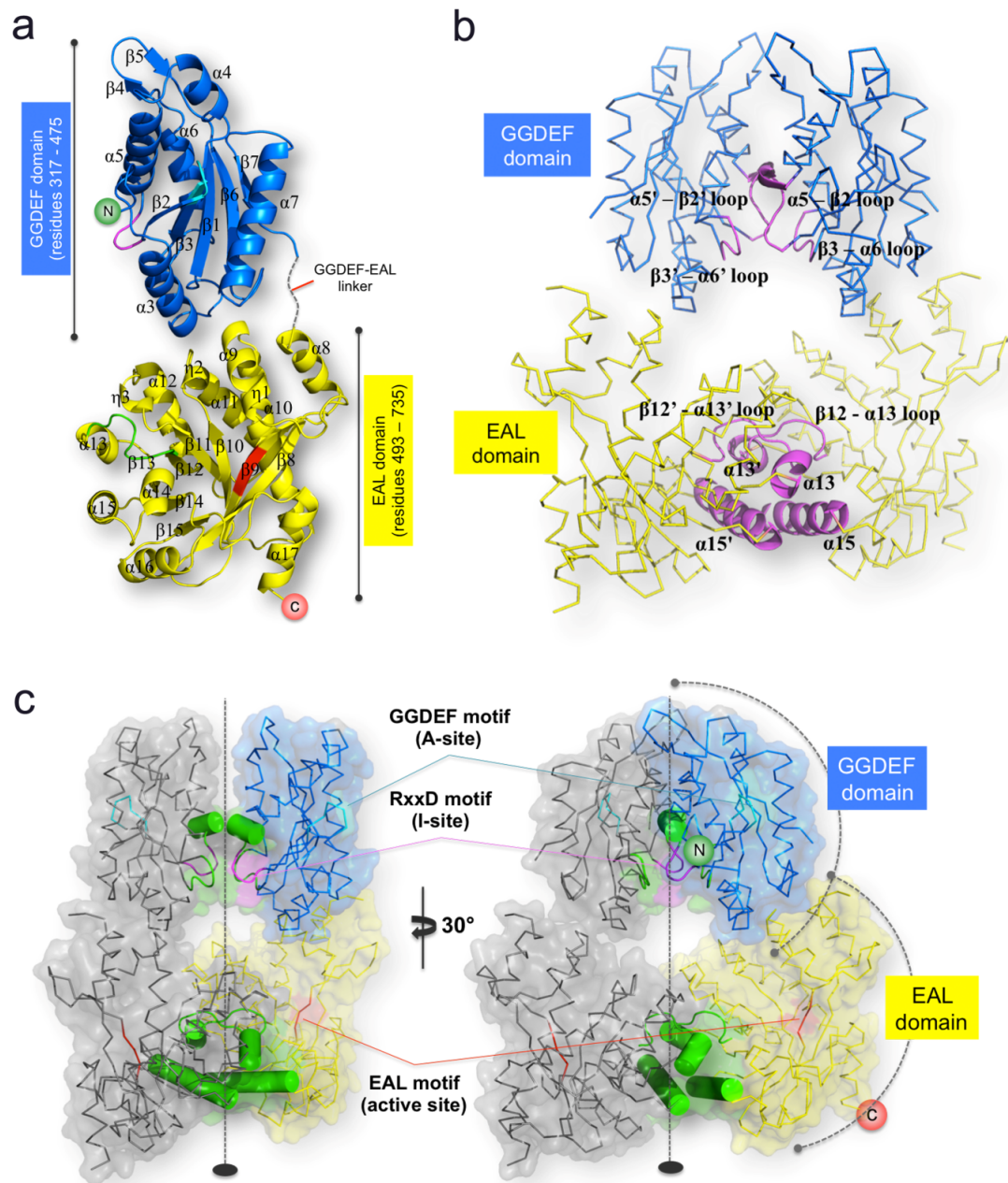
Figure 3.8. Crystal structure of Tbd1265_DUAL. (a) A monomer shown in its extended conformation. The GGDEF domain (blue) is located on the apical face of the EAL domain (yellow) TIM barrel. The interdomain linker is not modelled in this structure. (b) Structural elements involved in the Tbd1265_DUAL dimer interface are coloured magenta. The EAL domains adopt the previously observed dimerisation mode involving the catalytic loop 6 (β12-α13 loop), and helices α13 and α15. The GGDEF domains interact only loosely via the β5-α2 and β3-α6 loops. (c) The locations of catalytically relevant features are indicated. The interface area is coloured in green, while the I-sites and the EAL motifs are in magenta and red, respectively. The GGDEF active sites represented by the GGDEF motif are marked in cyan, and are antipodal to each other, making the dimer conformation DGC-incompetent.

dimer via the conserved EAL dimer interface, identical to the one seen in RocR (Paper I). The interface is formed by the long helix α15, loop 6 (β12-α13 loop) and the short helix α13 which forms a "compound helix" with its dimer counterpart (Fig 3.7b), as observed in the standalone Tbd1265 EAL structure 2R6O (*95*). Unlike the EAL domains, the GGDEF domains are only associated loosely through the β5-α2 and β3-α6 loops. The GGDEF subunits sterically occlude each other's inhibitory I-site formed by Arg-385 and Asp-388, thus preventing them from binding c-di-GMP. At the same time, the two active sites are located on opposite sides of the dimer and separated (Fig 3.7c), making this conformation catalytically incompetent. In order to achieve catalysis, both GGDEF domains must rotate about 180° and bring the active sites together. The linkers connecting the GGDEF and EAL domains (residues 476 – 492) do not show sufficient electron density to be modelled accurately, possibly because they are flexible. Electron density in the N-helix region is also ambiguous and is thus not modelled as well. It is possible that flexibility and mobility of these structural elements are required for a regulatory mechanism based on conformational changes. A crystal form that diffracts to higher resolutions may be needed in order to carry out ligand co-crystallisation or soaking experiments.

3.2.5 Proposed regulatory mechanism of Tbd1265

Based on the abovementioned biochemical studies and the crystal structure, we propose that Tbd1265 adopts at least three states whereby the DGC activity of the GGDEF domain is directly regulated by the PBP domain (Fig 3.9). This involves rotational motions of the GGDEF domain, a commonly seen regulatory mechanism amongst nucleotide cyclases (*12*). The GGDEF-EAL linker immediately downstream of the GGDEF domain must be flexible towards this end. Also, the N-helix is predicted to form a coiled coil structure which may associate and unwind according to structural changes transmitted from the PBP domain, thus controlling the relative orientation of the GGDEF subunits. In this regard the N-helix could be similar to the widespread signalling helix (S-helix) motif (*100*), although the signature ERT sequence motif is not found in N-helix. Another similar albeit distinct example of a signal transducer element is the HAMP domain (*101*), where each monomer contributes two short helices to form a four-helix bundle that propagates structural changes between neighbouring domains. Meanwhile, the extracellular ligand for Tbd1265 and how PBP exerts regulatory effects are still unknown. A PBP domain consists of two lobes which favour a tightly closed conformation when ligand-bound. However, they can undergo a large "opening"

rotation of about 50° relative to each other in the free state (*102*) possibly through multiple intermediate steps (*103, 104*). It is thus tempting to speculate that two PBP domains in a Tbd1265 dimer could push against each other in the absence of ligand, and consequently exert a pulling force on the N-helices to unwind the coiled coil. More structural and perhaps mutational studies will be needed to study the possible scenarios. Also, the extent of regulation exerted on the EAL domain still remains a question.
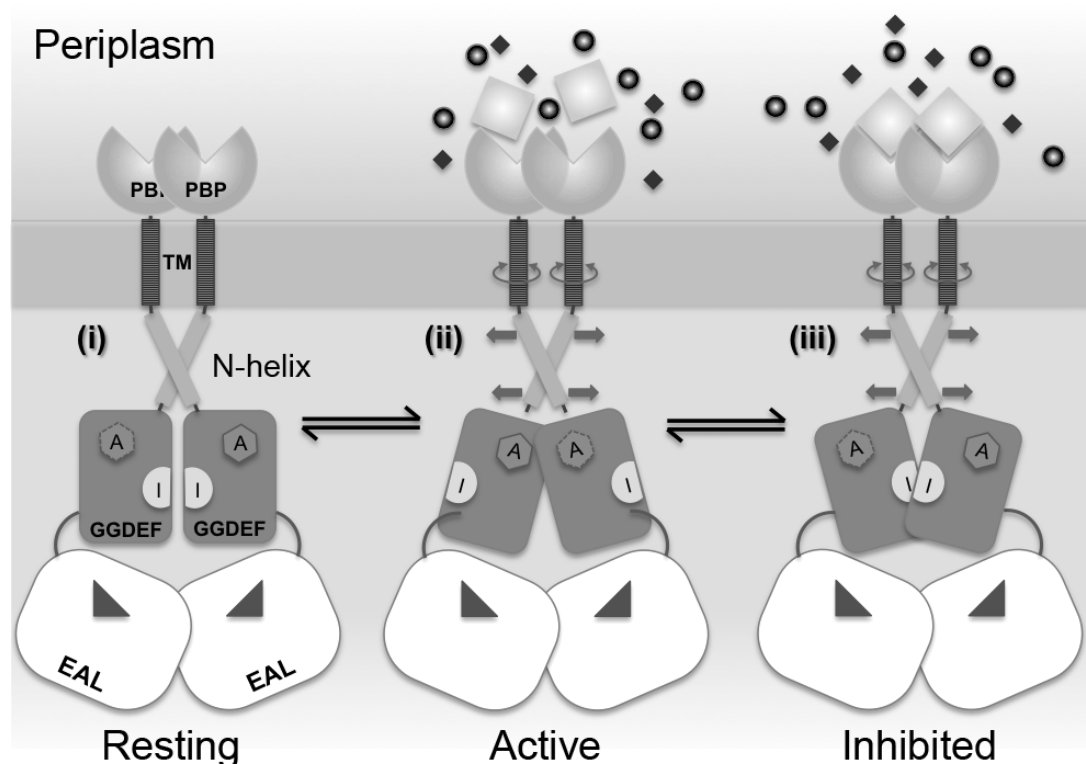
Figure 3.9. Proposed regulatory mechanism of Tbd1265. Tbd1265 may adopt at least three conformations whereby GGDEF activity is regulated. The resting state is represented by the crystal structure presented in this work, where both the active sites (labelled A) and the I-sites (labelled I) do not bind their ligands. Depending on signals originating from ligand binding events of the PBP, the N-helix coiled coil undergoes structural changes that bring the protein into an active conformation or an inhibited state.

3.2.6 Recombinant production of the periplasmic binding protein (PBP) constructs

In order to investigate the ligand-binding properties and possible conformational changes in the PBP domain, we designed and cloned 10 PBP constructs from the tbd1265 locus using *T. denitrificans* genomic DNA (ATCC 25259) as template. Expression constructs on the pNIC28-Bsa4 vector were acquired and summarised in Table 3.1. Solubility screening revealed PBP-A2 as the longest construct that has

excellent yield and solubility in *E. coli* BL21 (DE3). A pure and stable protein was attained after upscaled expression and purification.

Construct name	Starting residue	Ending residue	Expected MW with His-tag (Da)	Expected MW without His-tag (Da)
PBP-A1	Q2	R221	27069	24603
PBP-A2	Q2	R230	28007	25541
PBP-A3	Q2	W237	28897	26431
PBP-A4	V4	L223	27143	24677
PBP-A5	V4	W234	28275	25810
PBP-B1	Q2	L223	27368	24903
PBP-B2	Q2	W234	28501	26035
PBP-B3	V4	R221	26844	24378
PBP-B4	V4	R230	27782	25316
PBP-B5	V4	W237	28672	26206

Table 3.1. Constructs of PBP domain from Tbd1265.

3.2.7 Crystallisation of PBP-A2 and data collection

Small-scale crystallisation screening led to several crystallisation conditions for PBP-A2. Crystal optimisation using the sitting drop vapour-diffusion method yielded diamond-shaped crystals in a succinate-based condition (crystal form A) and another where PEG-MME 2000 was a co-precipitant in addition to succinate (crystal form B). X-ray diffraction experiments were carried out at the P14 beamline at Petra III, EMBL Hamburg. Diffraction data for crystal forms A and B were collected to 2.8 and 3.4 Å, respectively, giving diffraction images of good quality (Fig 3.10). Crystal form A was indexed to space group F432 with unit cell dimensions a=190.8 Å. Crystal form B was of space group I422 with unit cell dimensions a=134.1 Å and c=189.7Å. Data collection statistics are indicated in Table 3.2. As the resolution of currently available PBP data may not be sufficient for analysing the binding site and potential ligands, we are working towards better-diffracting crystals by optimising more crystallisation conditions. Meanwhile, we also perform differential scanning fluorimetry (Thermofluor) experiments to screen for potential ligands. This would help us understand the biological context under which Tbd1265 operates, and shed light on novel pathways governing the life style of *T. denitrificans*.
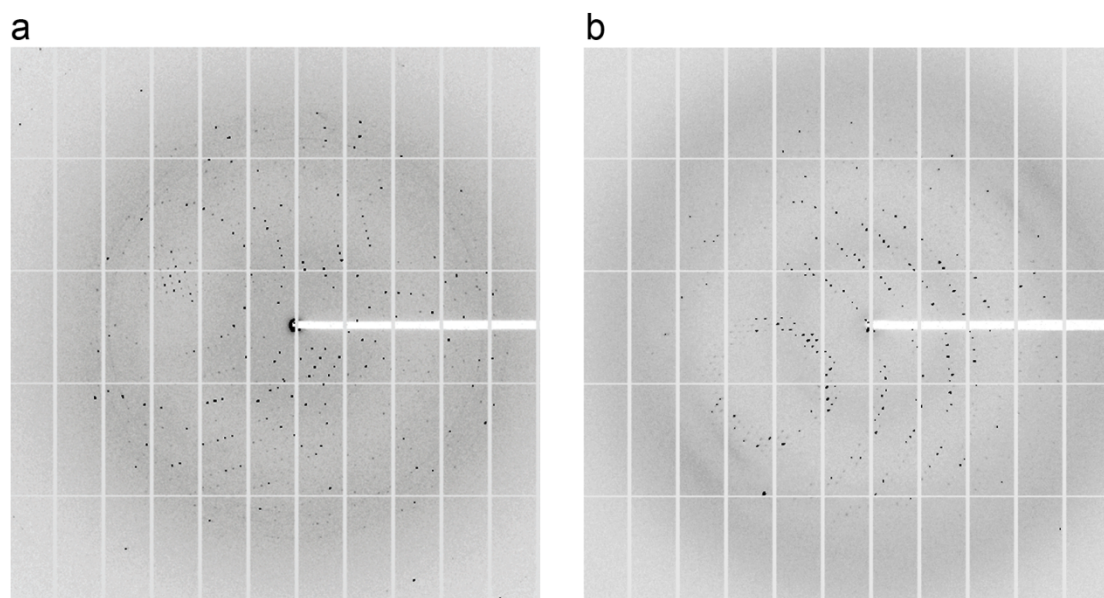
Figure 3.10. Preliminary X-ray diffraction data of PBP crystals. Diffraction images of (a) crystal form A and (b) crystal form B are shown.

Data set	Crystal form A	Crystal form B
Data collection statistics		
Beam line	P14, Petra III	P14, Petra III
Wavelength (Å)	0.9763	0.9763
Resolution (Å)	110.16 – 2.80 (2.95 – 2.80)	109.49 – 3.40 (3.58 – 3.40)
Space group	F432	I422
Cell parameters a/b/c (Å)	190.8, 190.8, 190.8	134.1, 134.1, 189.7
$\alpha/\beta/\gamma$ (°)	90.0, 90.0, 90.0	90.0, 90.0, 90.0
Unique reflections	7801 (1094)	12147 (1752)
Redundancy	8.3 (8.7)	5.9 (6.0)
$<I/\sigma>$	20.4 (1.4)	10.8 (2.9)
Completeness (%)	100.0 (100.0)	99.3 (99.6)
R_{merge}	0.083 (1.420)	0.102 (0.591)
R_{pim}	0.030 (0.507)	0.044 (0.251)
Wilson B factor (Å^2)	86.3	91.8

Table 3.2. Data collection statistics for crystal forms A and B of PBP. Values in parentheses depict the outer resolution shell.

3.3 STRUCTURAL STUDIES OF THE MURB ENZYME FROM *PSEUDOMONAS AERUGINOSA* (PAPER III)

The biosynthesis of peptidoglycan is an essential process in most pathogenic bacteria and enzymes from these pathways provide promising drug targets. MurB was therefore selected as one of the target enzymes in the the early drug discovery project AEROPATH (*105*), directed against the opportunistic pathogen *Pseudomonas aeruginosa*. We produced the protein recombinantly and obtained two crystal forms of the enzyme in complex with its cofactor FAD and the product of the first half reaction, NADP$^+$. The results allowed us to compare the binding modes of NADPH and UNAGEP to PaMurB, and to explain the activating effect of potassium ion on the oxidoreductase.

3.3.1 Recombinant protein production of PaMurB

The PaMurB gene (locus tag PA2977) was amplified by PCR from the *P. aeruginosa* PAO1 genomic DNA (ATCC 47085) and cloned into the pNIC28-Bsa4 vector, giving an expression construct with an N-terminal His-tag cleavable by TEV protease. Recombinant expression in *E. coli* BL21(DE3) and standard purification procedures afforded good amounts of protein that is stable at 25 mg/mL up to several days. The protein solution is bright yellow indicating retention of the FAD cofactor. Analytical SEC analysis shows unambiguously a single monomeric species (Fig 3.11).
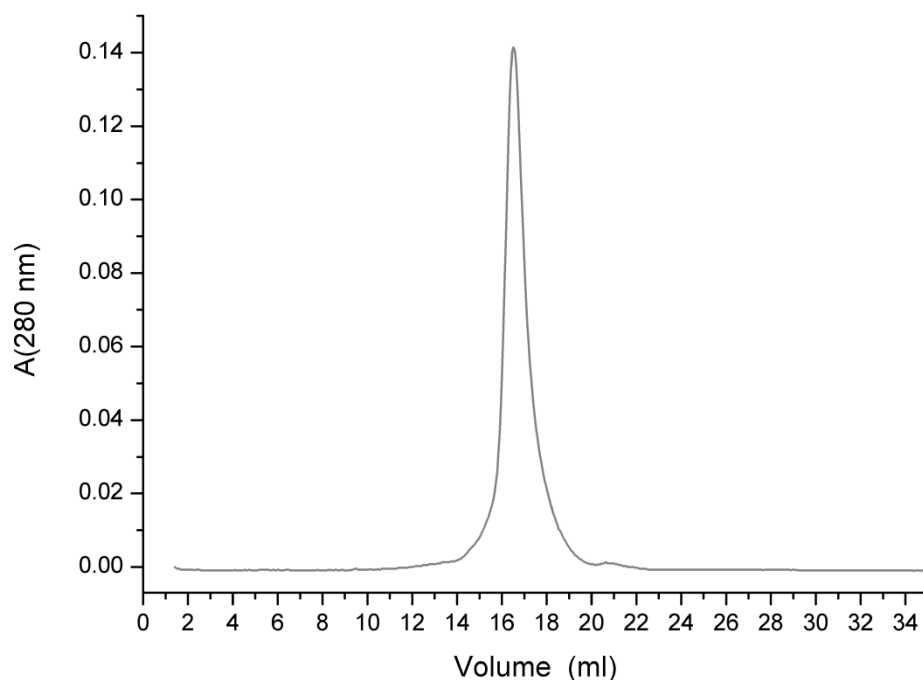

Figure 3.11. Analytical SEC analysis of PaMurB. The protein elutes as a single species on the Superdex 200 10/300 column. The elution volume corresponds to the monomeric weight of PaMurB (not shown).

3.3.2 Crystallisation and structure determination of the PaMurB-FAD-NADP(H) ternary complex

Despite its apparent stability PaMurB did not crystallise without additives. We attempted to co-crystallise the protein with UNAG and NADPH (UNAGEP was not available); the NADPH trial gave multiple crystallisation conditions with various crystal morphologies during small-format screening. Interestingly, NADPH temporarily and reversibly bleached the protein solution – a sign of FAD reduction and subsequent reoxidation. Crystal form A in the form of yellow plates was obtained with PEG 3350 as precipitant. Additional screening with $NADP^+$ sodium salt yielded crystal form B, a yellow prism grown in PEG 8000 and glycerol. Incidentally, both conditions contained a potassium salt (200 and 40 mM for conditions A and B, respectively), a point which will be discussed later. Samples of crystal forms A and B diffracted to 2.0 – 2.5 Å and were of space group C2 and $P6_1$, respectively. Diffraction patterns of crystal form A are unfortunately streaky, resulting in high R_{merge} values under repeated attempts. The two structures were solved by molecular replacement with the EcMurB structure (PDB code 2MBR) as search model. Protein models of crystal forms A and B were refined to 2.23 and 2.10 Å, and contained four and one molecule(s) per ASU, respectively.

3.3.3 Crystal structure of PaMurB

The two structures agreed with each other except in three peripheral loops involved in crystal contacts. Consistent with sequence analysis, PaMurB is a type I UNAGEP reductase similar to EcMurB, following the three-domain architecture (Fig 3.12A) of FAD-binding domains I (aa 1 – 75 and 336 – 339) and II (aa 76 – 191), and the substrate-binding domain III (aa 192 – 335). The FAD-binding module consists of two β-sheets (β1-β2-β5-β3 and β6-β7-β12-β10-β11) packed against flanking helices α1 and α2, respectively. A buried FAD-binding pocket is formed by the GGG motif on the β3-β4 loop and the C-terminal strand β20 contributed by domain I, along with the β7-α3 loop, the η1- α3- β8 region, and α4- β12 from domain II. Adjacent to the cofactor binding site is the substrate-binding domain III divided into two lobes. Lobe 1 interacts with domain II and contains helices α5-α6, while lobe 2 is close to domain I and contains the β13-β19-β18-β17 sheet and the β14-α7-β15-β16 outcrop. The substrate binding site is located between the lobes. The structure of $NADP^+$-bound PaMurB is similar to UNAGEP-bound EcMurB (2MBR, rmsd 1.3 Å based on 321 Cα atoms) but different from SaMurB and TcMurB, which are type II MurB proteins different in domain III (Fig 3.12B).
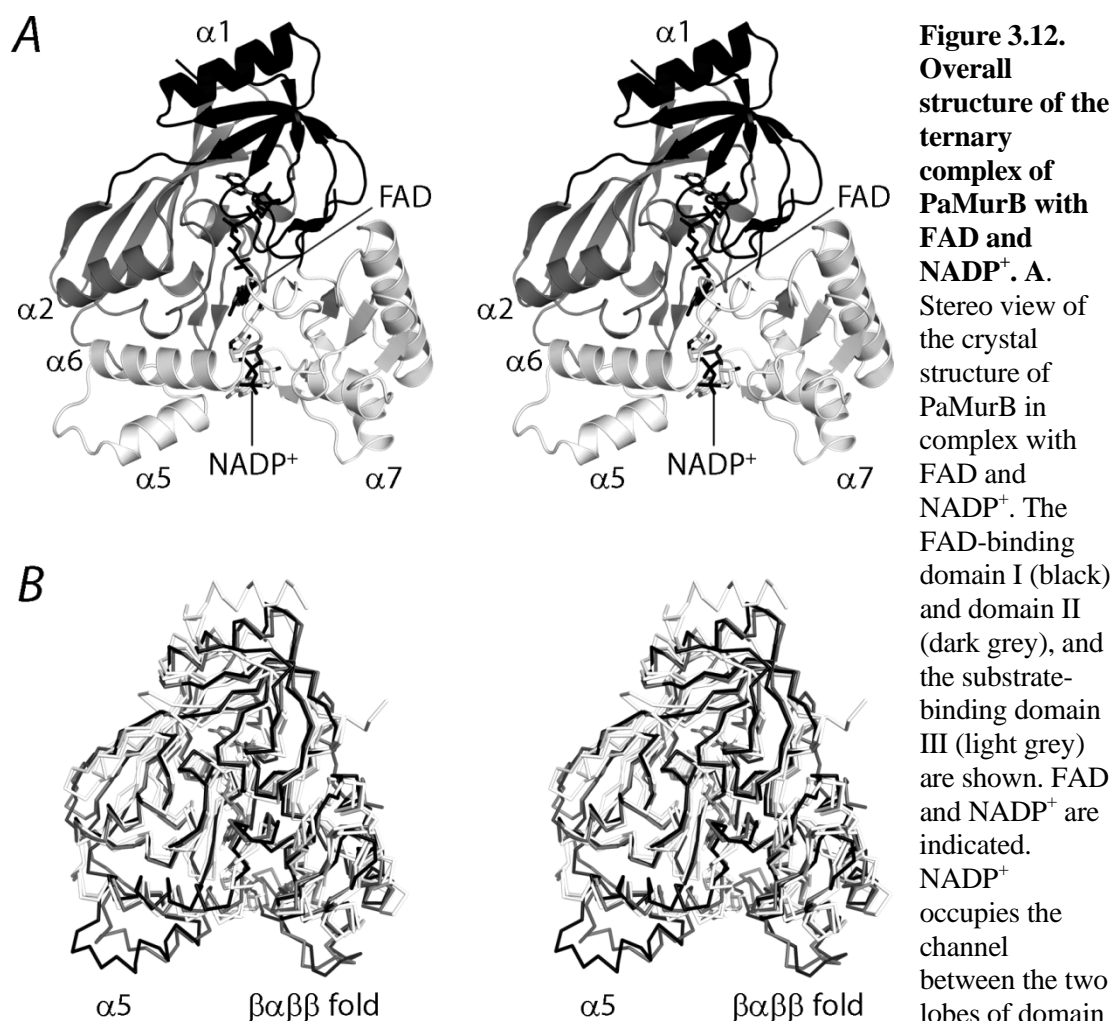
Figure 3.12. Overall structure of the ternary complex of PaMurB with FAD and NADP⁺. A. Stereo view of the crystal structure of PaMurB in complex with FAD and NADP⁺. The FAD-binding domain I (black) and domain II (dark grey), and the substrate-binding domain III (light grey) are shown. FAD and NADP⁺ are indicated. NADP⁺ occupies the channel between the two lobes of domain III (in this view: left, lobe 1; right, lobe 2). **B.** Superimposition of the Cα traces of EcMurB (grey) and type II enzymes (SaMurB and TcMurB, white) against PaMurB (black). PaMurB and EcMurB display high structural similarity. In domain III, type II MurB enzymes lack the tyrosine loop preceding helices α4 and α5, as well as the protruding βαββ fold on lobe 2.

PaMurB forms a 1:1:1 ternary complex with FAD and NADP⁺ with well-defined electron densities for both molecules in the crystal structure. The cofactor is oxidised as the isoalloxazine ring is planar, and adopts a conserved binding conformation identical to the FAD molecules in previously determined MurB structures (*73-75, 79, 80, 106*). This is likely because the amino acid residues that bind the cofactor are rather conserved themselves. The NADP⁺ is surprisingly well-stabilised and displays unambiguous electron density throughout the entire molecule (Fig 3.13). It binds in the domain III substrate channel such that the nicotinamide group is located on the *si* face of the FAD isoalloxazine, bringing the hydride-donating C4n atom to within 3 Å of the isoalloxazine N5 atom. The adenosine end of the molecule faces away from the active site. In addition, a strong residual density adjacent to the nicotinamide group was interpreted as a potassium ion according to B-factor analysis and typical coordination bond lengths (*107*) (Table S1, Paper III). The ion has a pentagonal bipyramidal

coordination sphere formed by four amino acid residues including the catalytically essential Ser-239 and Glu-335, and an oxygen atom from the nicotinamide group. Curiously, potassium can only be modelled into structure model A but not model B, probably because the higher potassium concentration in crystal form A (200 mM) is closer to physiological conditions (*108*) compared to crystal form B (only 40 mM).
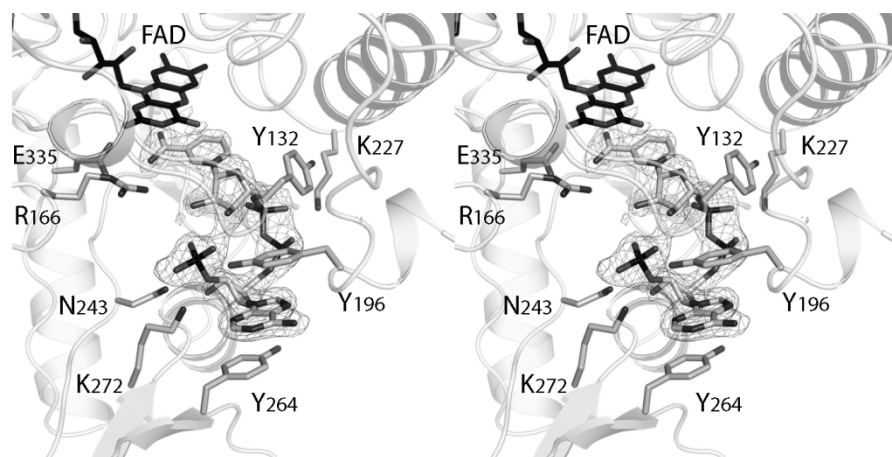

Figure 3.13. NADP$^+$ and the substrate binding site. NADP$^+$ is shown as a stick model surrounded by active site residues and FAD. The F$_o$-F$_c$ omit electron density of NADP$^+$ is contoured at 3.0 σ and indicates a well-ordered ligand. The nicotinamide ring stacks against the isoalloxazine ring system of FAD (black stick model). Residues of the binding site that form hydrogen bonds with NADP$^+$ include Tyr-132, Arg-166 and Glu-335 for the nicotinamide moiety, Lys-227 for the diphosphate backbone, and Tyr-196, Asn-243 and Lys-272 for the adenosine. The adenosine moiety is in addition stabilised by stacking interactions with Tyr-196 and Tyr-264.

3.3.4 NADPH and UNAGEP bind to the same flexible substrate channel

Structural superimposition of NADP$^+$-bound PaMurB and UNAGEP-bound EcMurB, with respect to FAD, shows that NADP(H) and UNAGEP occupy the same site between the two lobes of domain III (Fig 3.14). The nicotinamide and enolpyruvyl groups are found at the same location on the *si* face of the flavin, with C4n and C3e aligned. The pyrophosphate backbones and the nucleotide moieties of the two substrates diverge but bind within the same channel, achieved by conformational changes in the substrate-binding residues such as Tyr-196, Tyr-264, Lys-272 etc. This agrees with previous NMR studies suggesting that the MurB active site is flexible and changes are inducible in the substrate binding site (*83, 84*). An FAD-NADPH stereochemistry using the *si* face of the isoalloxazine ring has only been observed before in the methylenetetrahydrofolate reductase (*109, 110*). The first crystal structure of the ternary complex of MurB with FAD and NADPH provides first insights into the binding mode of NADPH to this enzyme class revealing the unusual FAD-NADPH *si* stereochemistry.
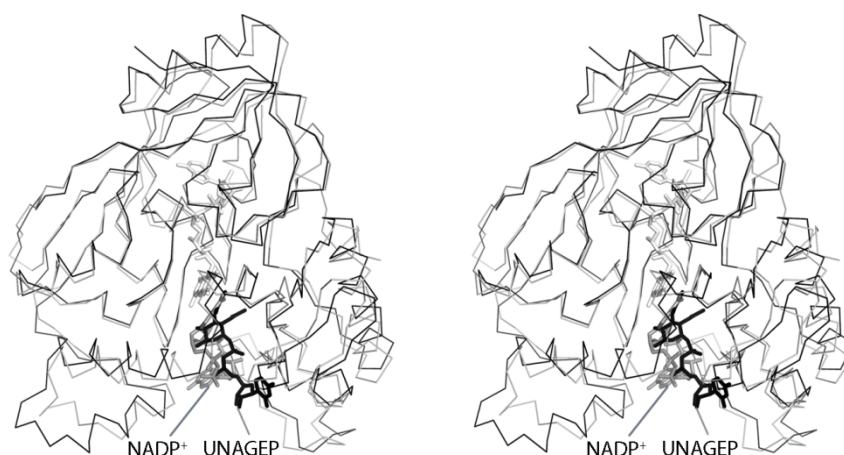
Figure 3.14. NADP⁺ shares the same substrate binding site with UNAGEP. Stereo view of PaMurB (black ribbon) and UNAGEP-bound EcMurB (grey ribbon, 2MBR) superimposed based on their FAD atomic coordinates. FAD, NADP⁺ and UNAGEP are shown as stick models in white, grey and black, respectively.

3.3.5 Active site potassium ensures efficient hydride transfer

The PaMurB-NADP⁺ structure shows that the active potassium ion and the essential residues Ser-239 and Glu-335 bind the nicotinamide group of NADP(H), bringing C4n to the isoalloxazine N5 atom within 3 Å at an angle (Fig 3.15).
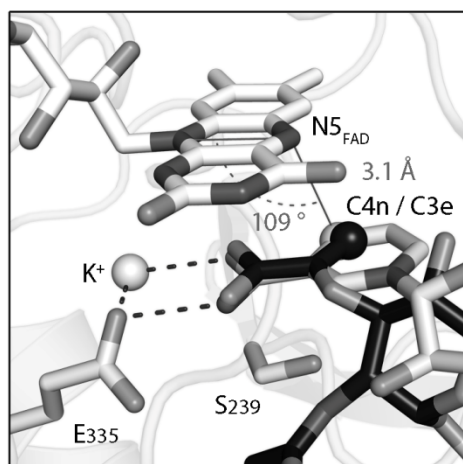
Figure 3.15. Stereochemistry of MurB and its substrates. The active site potassium ion assists in substrate orientation and binding. Superimposition of the PaMurB crystal form A structure and the EcMurB-UNAGEP complex (2MBR) based on FAD atomic coordinates shows that the C2-C3-C4 locus of NADP⁺ nicotinamide (in grey) spatially overlap with the enolpyruvyl group of UNAGEP (in black). Both substrate moieties are bound to Glu-335 and the backbone amine of Ser-239. The nicotinamide C4n atom (grey sphere), which transfers a hydride to the isoalloxazine N5 atom, coincides with the enolpyruvyl C3e (black sphere), which receives the hydride during the second half-reaction. The geometric relation of the C4n atom to the isoalloxazine is indicated. The substrate carbons are arranged in the optimal position for hydride transfer.

During the oxidation of NADPH, the 4-*pro*-S hydrogen atom of the sp^3-hybridised C4n would point towards N5; the colinear alignment of C4n-H-N5 would facilitate the transfer of the hydrogen to N5. This agrees with isotope tracing studies (*71*) and explains the activating effect of potassium on MurB (*72*). The potassium both increases catalysis by increasing k_{cat} and promotes NADPH binding by lowering K_m. During the second half reaction, the C3e atom of UNAGEP occupies the same position, allowing efficient hydride transfer from N5 to complete the reaction. Meanwhile, the substrate-

binding channel allows both substrates to approach FAD without causing FAD to be exposed to solvent, which might cause the loss of the hydride ion.

3.3.6 Implications for drug design against MurB

Although the MurB substrate channel is flexible and the two substrates precisely co-align only at the reactive moieties, the comparison of UNAGEP- and NADP$^+$-bound MurB structures suggests that the binding site can be visualised as three loci (Fig 5B, Paper III). Adjacent to FAD is the substrate locus which binds the nicotinamide and enolpyruvyl moieties, followed by the middle locus that binds the sugar-diphosphate backbone. The third locus being closest to the bulk solvent is occupied by the nucleotide moieties. This could be a basis for fragment-based inhibitor design and screening. In a co-crystal structure of EcMurB bound to a naphthyl inhibitor (PDB code 2Q85) released by Wyeth (*89*), the compound also binds to the substrate channel in an extended conformation and displays a three-fragment structure. The compound's chlorophenyl group, furanone core and naphthyl group correspond to and co-localise with the natural substrates' reactive moiety, sugar-diphosphate backbone and nucleotide group, respectively. We hope that structural information of the PaMurB-NADP(H) structure, along with other MurB structures in the PDB, would facilitate the development of novel antibiotics targeting MurB specifically.

4 CONCLUDING REMARKS

This thesis focuses on the structural biology aspects of c-di-GMP signalling and cell wall biosynthesis. In particular, the regulatory mechanisms of two c-di-GMP-metabolising enzymes is emphasised from a structural point of view; X-ray crystallography, SAXS and biochemical assays have been used to complete the picture.

In order to investigate how the phosphoreceiver REC domain regulates the PvrR family of EAL proteins, the crystal structure of RocR was determined after overcoming some difficulties in experimental phasing. The RocR tetramer is in fact a dimer of dimers; due to two different monomeric conformations, the protein is folded such that two REC domains are exposed to bulk solvent and each in contact with an EAL dimer. Also, solution SAXS studies on three RocR variants with differing PDE activities suggest that altering catalytic output may not involve large conformational changes of the tetramer. Instead, a mechanism is proposed where phosphorylation of the exposed REC domains transmits secondary structural changes from the phosphorylation site to its adjacent EAL dimer, affecting the conformation of the catalytic loop 6 and thereby regulating PDE activity. Although RocR is known to regulate *cup* fimbrial genes (*35, 36*), the cognate ligand activating its kinase RocS1 and the exact effectors of RocR are unknown. The possibility of RocR interacting with other proteins related to c-di-GMP signalling, in analogy to the REC-HD-GYP protein RpfG that binds several GGDEF proteins when phosphorylated (*41*), should also be considered.

To date, no experimental structure has been published for a GGDEF-EAL tandem domain that shows bifunctionality *in vitro*. Using Tbd1265 as a subject, the second aim was to elucidate how such proteins adjust their activities according to signalling cues and how precise regulation prevents an energy-wasting loop of DGC and PDE catalysis. Activity assays indicated that the predicted coiled-coil N-helix region is important for both DGC and PDE activities. Also, a moderate-resolution crystal structure of the tandem domain construct Tbd1265$_{DUAL}$ revealed a parallel dimeric arrangement in the protein. The results suggest that N-helix may cause the GGDEF domains to rotate relative to each other, thus bringing the active sites or the inhibitory I-sites together. We are currently attempting to solve the structure of the periplasmic PBP domain and identify its ligand in order to understand how the transmembrane protein

functions as a whole. Apart from attaining crystals that diffract better, the role of N-helix should be investigated further; deletions and mutations that alter GGDEF orientations and the interhelical dissociation constant could alter the protein's activity and provide more information.

Last but not least, the crystal structure of MurB from *P. aeruginosa* in complex with FAD and $NADP^+$ was solved in an effort to study its substrate binding site. The results established that NADPH shares the same substrate channel with UNAGEP in a sequential manner during the production of UNAM. To allow the substrate functional groups of both molecules (nicotinamide and enolpyruvyl groups) to approach the buried cofactor FAD, the substrate channel can be remodelled to interact with the two distinct and extended molecules. Three loci can be visualised in the binding site and may provide a new strategy of fragment-based screening against MurB. Surprisingly, a potassium ion was identified in the active site that binds directly to the substrate, thus solving the mystery of its activating properties on MurB (*72, 82*). Our analysis focused on type I MurB enzymes exemplified by EcMurB and PaMurB; it would be interesting to see if type II MurB proteins exhibit the same flexibility in the substrate binding site.

5 ACKNOWLEDGEMENTS

I would like to take this opportunity to express my gratitude and appreciation for the wonderful people whose paths I've crossed over the past years.

A very big thank-you to my KI supervisor Gunter Schneider, for your unwavering support and guidance during my time in MSB. Your meticulous attention to details, responsible attitude in scientific writing and steadfast commitment towards bettering research infrastructures are inspiring. I would also like to thank Ylva Lindqvist for accepting me as a KI student, and leading the group together with Gunter. I really appreciate your hospitality, care and help, especially during the past few months.

My NTU supervisor, Julien Lescar, brought me into structural biology. Thank you, Julien, for taking me in since my undergraduate days, and for all the opportunities and freedom given to me. I wish I learned more physics from you! Also, many thanks to Zhao-Xun Liang, my NTU co-supervisor who introduced me to c-di-GMP signalling and advised me generously.

A million thanks to Chong Wai, a great mentor and an even greater friend. My time in the Lescar lab would never have been the same without you, and I'm happy to see you succeed in your professional and family lives. I would also like to thank Masayo, without whom the RocR project would have been impossible. Working with you was always both fun and educational, and I do wish I could join you and Chong Wai on a synchrotron trip again.

My time in KI would not have been the same without Robert Schnell, who offered me *korv* the first day we met, and who has been my personal Dr Who in MSB ever since. I will remember your advice, "science is an endeavour of overtime." I'm also grateful to Bernhard Lohkamp, the *other* Dr Who who is ever helpful in the unfathomable universe that is CCP4i. I had a good time refining the MurB structure with you.

I would like to thank all my collaborators – with you my works presented in this thesis were made possible. Thank you, Lawrence, not only for helping with enzymatic assays but also for being a loyal friend who is always there for me even across the oceans. Mary, thank you for all your generous help and I wish you all the best with the finish line. Martina, thanks for your diligence in making protein constructs for us and for bringing a sunny piece of Sweden to Singapore. Ranjana and Cecilia, I really appreciate your help with SAXS even though our effort did not bear edible fruit.

To all the past and present members of MSB, I really appreciate our time together and your friendship in moments of happiness and times of need. Dominic, you will be forever remembered as a walking encyclopedia in the Dr Mueller lab coat. I wish you all the best for your upcoming dissertation and whatever comes after. Magnus, thank you for being a helpful neighbour who displayed inspiring reductive-methylation skills. Ahmad, I can't thank you enough for your generous offers of anecdotes, assistance and chewing gum. Maria, my gratitude for helping me with all things concerning the thesis; I owe you one, and hope we'll have lunch in the sun again. Katharina, thanks for

always bringing a positive attitude to everyone, in style. Dominik, you prove that silence is indeed golden; I know who to turn to if I need help with insect cell cultures and hospital company. I also know who to approach if I need silent help with c-di-GMP issues and sauna construction, Edvard. Eva-Maria, many thanks for your kind and sharing spirit (your pumpkin bread recipe, please). Ömer, thank you for your advice and help when I needed them, and for painting romantically sad pictures of Berlin and Istanbul in my head. Jodie, thanks for all things crystallography in the lab and I wish you the very best with the next addition to your family. Doreen, thank you for our enjoyable time at Hjulet, and for baking and sharing such wonderful delights. Jason, you brought colours into the lab which we all sorely miss, but when is that follow-up paper on drop bears coming along? Rajesh, great to hear you're enjoying married life in Lisbon, and thanks for being a wonderfully helpful and entertaining neighbour. Cyprian, thanks for your stoic composure and useful insights. And thank you, Atsushi, for sharing the appreciation for Rigaku equipment and your parting advice. Shiromi, I appreciate your friendship and your infectious awesomeness.

Victoria, many thanks for your almost-magic solutions in all things administrative, and for being good company after work. Eva, thank you for constantly showing your concern and care. Alessandra, for being ever helpful especially during the start and the conclusion of my time here.

I would like to thank Lars Nordenskiold for organising the NTU-KI PhD programme which has significantly broadened my horizon. Thanks to Ute Römling for useful discussions on c-di-GMP signalling and for writing such instructive reviews that benefit all in this field. Many thanks to members of the NTU Protein Production Platform and the KI Protein Science Facility for your guidance and assistance.

Not to forget past and present members of the Lescar lab: Yee Hwa, I'm grateful for your support in experimental and administrative work. Qiwei, thanks for all the laughter. Dahai, for your inspiring work ethics and sarcastic humour. Joe, for organising various aspects of the lab and being an entertainment factor. Rong, for your clever outbursts. Yee Ling, for showering us with motherly care and food tips.

Simin and Rachel, I can't thank you enough for being such good friends who provide listening ears and helping hands. Thanks for being there for me during my recent trial. Old friends – Ruiyang, Jie Yin, Zhishan, Sam, Kuan Pern, James – I hope we will embark together on new adventures soon.

Mr Simon Lai, thank you for your selfless commitment to your students. This thesis is for you, after all these years.

My eternal gratitude to my parents, who have stood by me through thick and thin. Ming Liang, I wish you all the best in all you'll do after your graduation.

I can't acknowledge all of you in these pages, but I treasure the memories.

The leaves are falling in the stream; the River flows away.

6 REFERENCES

1. Ross, P., Weinhouse, H., Aloni, Y., Michaeli, D., Weinberger-Ohana, P., Mayer, R., Braun, S., de Vroom, E., van der Marel, G. A., van Boom, J. H., and Benziman, M. (1987) Regulation of cellulose synthesis in *Acetobacter xylinum* by cyclic diguanylic acid, *Nature 325*, 279-281.
2. Römling, U., Galperin, M. Y., and Gomelsky, M. (2013) Cyclic di-GMP: the First 25 Years of a Universal Bacterial Second Messenger, *Microbiology and Molecular Biology Reviews 77*, 1-52.
3. Ross, P., Aloni, Y., Weinhouse, C., Michaeli, D., Weinberger-Ohana, P., Meyer, R., and Benziman, M. (1985) An unusual guanyl oligonucleotide regulates cellulose synthesis in *Acetobacter xylinum*, *FEBS Letters 186*, 191-196.
4. Ross, P., Aloni, Y., Weinhouse, H., Michaeli, D., Weinberger-Ohana, P., Mayer, R., and Benziman, M. (1986) Control of cellulose synthesis *Acetobacter xylinum*. A unique guanyl oligonucleotide is the immediate activator of the cellulose synthase, *Carbohydrate Research 149*, 101-117.
5. Tal, R., Wong, H. C., Calhoon, R., Gelfand, D., Fear, A. L., Volman, G., Mayer, R., Ross, P., Amikam, D., Weinhouse, H., Cohen, A., Sapir, S., Ohana, P., and 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, *Journal of Bacteriology 180*, 4416-4425.
6. Galperin, M. Y., Natale, D. A., Aravind, L., and Koonin, E. V. (1999) A specialized version of the HD hydrolase domain implicated in signal transduction, *J Mol Microbiol Biotechnol 1*, 303-305.
7. Galperin, M. (2005) A census of membrane-bound and intracellular signal transduction proteins in bacteria: Bacterial IQ, extroverts and introverts, *BMC Microbiology 5*, 35.
8. Chen, Z.-h., and Schaap, P. (2012) The prokaryote messenger c-di-GMP triggers stalk cell differentiation in *Dictyostelium*, *Nature 488*, 680-683.
9. Galperin, M. Y., Nikolskaya, A. N., and Koonin, E. V. (2001) Novel domains of the prokaryotic two-component signal transduction systems, *FEMS Microbiology Letters 203*, 11-21.
10. Pei, J., and Grishin, N. V. (2001) GGDEF domain is homologous to adenylyl cyclase, *Proteins: Structure, Function, and Bioinformatics 42*, 210-216.
11. Sinha, S. C., and Sprang, S. R. (2006) Structures, mechanism, regulation and evolution of class III nucleotidyl cyclases, *Rev Physiol Biochem Pharmacol 157*, 105-140.
12. Linder, J. U. (2006) Class III adenylyl cyclases: molecular mechanisms of catalysis and regulation, *Cell. Mol. Life Sci. 63*, 1736-1751.
13. Ryjenkov, D. A., Tarutina, M., Moskvin, O. V., and Gomelsky, M. (2005) Cyclic Diguanylate Is a Ubiquitous Signaling Molecule in Bacteria: Insights into Biochemistry of the GGDEF Protein Domain, *Journal of Bacteriology 187*, 1792-1798.
14. Paul, R., Weiser, S., Amiot, N. C., Chan, C., Schirmer, T., Giese, B., and Jenal, U. (2004) Cell cycle-dependent dynamic localization of a bacterial response regulator with a novel di-guanylate cyclase output domain, *Genes & Development 18*, 715-727.
15. Chan, C., Paul, R., Samoray, D., Amiot, N. C., Giese, B., Jenal, U., and Schirmer, T. (2004) Structural basis of activity and allosteric control of diguanylate cyclase, *Proc Natl Acad Sci U S A 101*, 17084-17089.
16. Wassmann, P., Chan, C., Paul, R., Beck, A., Heerklotz, H., Jenal, U., and Schirmer, T. (2007) Structure of BeF_3^- -modified response regulator PleD: implications for diguanylate cyclase activation, catalysis, and feedback inhibition, *Structure 15*, 915-927.

46

17. De, N., Navarro, M. V. A. S., Raghavan, R. V., and Sondermann, H. (2009) Determinants for the Activation and Autoinhibition of the Diguanylate Cyclase Response Regulator WspR, *Journal of Molecular Biology 393*, 619-633.

18. De, N., Pirruccello, M., Krasteva, P. V., Bae, N., Raghavan, R. V., and Sondermann, H. (2008) Phosphorylation-Independent Regulation of the Diguanylate Cyclase WspR, *PLoS Biol 6*, e67.

19. Zähringer, F., Lacanna, E., Jenal, U., Schirmer, T., and Boehm, A. (2013) Structure and Signaling Mechanism of a Zinc-Sensory Diguanylate Cyclase, *Structure 21*, 1149-1157.

20. Bobrov, A. G., Kirillina, O., and Perry, R. D. (2005) The phosphodiesterase activity of the HmsP EAL domain is required for negative regulation of biofilm formation in *Yersinia pestis*, *FEMS Microbiology Letters 247*, 123-130.

21. Schmidt, A. J., Ryjenkov, D. A., and Gomelsky, M. (2005) The Ubiquitous Protein Domain EAL Is a Cyclic Diguanylate-Specific Phosphodiesterase: Enzymatically Active and Inactive EAL Domains, *Journal of Bacteriology 187*, 4774-4781.

22. Barends, T. R., Hartmann, E., Griese, J. J., Beitlich, T., Kirienko, N. V., Ryjenkov, D. A., Reinstein, J., Shoeman, R. L., Gomelsky, M., and Schlichting, I. (2009) Structure and mechanism of a bacterial light-regulated cyclic nucleotide phosphodiesterase, *Nature 459*, 1015-1018.

23. Minasov, G., Padavattan, S., Shuvalova, L., Brunzelle, J. S., Miller, D. J., Basle, A., Massa, C., Collart, F. R., Schirmer, T., and Anderson, W. F. (2009) Crystal structures of YkuI and its complex with second messenger cyclic Di-GMP suggest catalytic mechanism of phosphodiester bond cleavage by EAL domains, *J Biol Chem 284*, 13174-13184.

24. Tchigvintsev, A., Xu, X., Singer, A., Chang, C., Brown, G., Proudfoot, M., Cui, H., Flick, R., Anderson, W. F., Joachimiak, A., Galperin, M. Y., Savchenko, A., and Yakunin, A. F. (2010) Structural Insight into the Mechanism of c-di-GMP Hydrolysis by EAL Domain Phosphodiesterases, *Journal of Molecular Biology 402*, 524-538.

25. Navarro, M. V., Newell, P. D., Krasteva, P. V., Chatterjee, D., Madden, D. R., O'Toole, G. A., and Sondermann, H. (2011) Structural basis for c-di-GMP-mediated inside-out signaling controlling periplasmic proteolysis, *PLoS Biol 9*, e1000588.

26. Chen, M. W., Kotaka, M., Vonrhein, C., Bricogne, G., Rao, F., Chuah, M. L., Svergun, D., Schneider, G., Liang, Z. X., and 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*, 4837-4846.

27. Rao, F., Yang, Y., Qi, Y., and Liang, Z.-X. (2008) Catalytic Mechanism of Cyclic Di-GMP-Specific Phosphodiesterase: a Study of the EAL Domain-Containing RocR from *Pseudomonas aeruginosa*, *Journal of Bacteriology 190*, 3622-3631.

28. Rao, F., Qi, Y., Chong, H. S., Kotaka, M., Li, B., Li, J., Lescar, J., Tang, K., and Liang, Z.-X. (2009) The Functional Role of a Conserved Loop in EAL Domain-Based Cyclic di-GMP-Specific Phosphodiesterase, *Journal of Bacteriology 191*, 4722-4731.

29. Galperin, M. Y., and Koonin, E. V. (2012) Divergence and convergence in enzyme evolution, *J Biol Chem 287*, 21-28.

30. Ryan, R. P., Fouhy, Y., Lucey, J. F., Crossman, L. C., Spiro, S., He, Y.-W., Zhang, L.-H., Heeb, S., Cámara, M., Williams, P., and Dow, J. M. (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*, 6712-6717.

31. Lovering, A. L., Capeness, M. J., Lambert, C., Hobley, L., and Sockett, R. E. (2011) The structure of an unconventional HD-GYP protein from *Bdellovibrio* reveals the roles of conserved residues in this class of cyclic-di-GMP phosphodiesterases, *MBio 2*.

32. Bellini, D., Caly, D. L., McCarthy, Y., Bumann, M., An, S.-Q., Dow, J. M., Ryan, R. P., and Walsh, M. A. (2013) Crystal structure of an HD-GYP domain

cyclic-di-GMP phosphodiesterase reveals an enzyme with a novel trinuclear catalytic iron center, *Molecular Microbiology*, n/a-n/a.

33. Schirmer, T., and Jenal, U. (2009) Structural and mechanistic determinants of c-di-GMP signalling, *Nat Rev Micro 7*, 724-735.

34. Barbieri, C. M., Mack, T. R., Robinson, V. L., Miller, M. T., and Stock, A. M. (2010) Regulation of Response Regulator Autophosphorylation through Interdomain Contacts, *Journal of Biological Chemistry 285*, 32325-32335.

35. Kuchma, S. L., Connolly, J. P., and O'Toole, G. A. (2005) A Three-Component Regulatory System Regulates Biofilm Maturation and Type III Secretion in *Pseudomonas aeruginosa*, *Journal of Bacteriology 187*, 1441-1454.

36. Kulasekara, H. D., Ventre, I., Kulasekara, B. R., Lazdunski, A., Filloux, A., and Lory, S. (2005) A novel two-component system controls the expression of *Pseudomonas aeruginosa* fimbrial cup genes, *Molecular Microbiology 55*, 368-380.

37. Henry, J. T., and Crosson, S. (2011) Ligand-Binding PAS Domains in a Genomic, Cellular, and Structural Context, *Annual Review of Microbiology 65*, 261-286.

38. Ho, Y.-S. J., Burden, L. M., and Hurley, J. H. (2000) Structure of the GAF domain, a ubiquitous signaling motif and a new class of cyclic GMP receptor, *EMBO J 19*, 5288-5299.

39. Wan, X., Tuckerman, J. R., Saito, J. A., Freitas, T. A. K., Newhouse, J. S., Denery, J. R., Galperin, M. Y., Gonzalez, G., Gilles-Gonzalez, M.-A., and Alam, M. (2009) Globins Synthesize the Second Messenger Bis-(3′–5′)-Cyclic Diguanosine Monophosphate in Bacteria, *Journal of Molecular Biology 388*, 262-270.

40. Plate, L., and Marletta, Michael A. (2012) Nitric Oxide Modulates Bacterial Biofilm Formation through a Multicomponent Cyclic-di-GMP Signaling Network, *Molecular Cell 46*, 449-460.

41. Ryan, R. P., McCarthy, Y., Andrade, M., Farah, C. S., Armitage, J. P., and Dow, J. M. (2010) Cell–cell signal-dependent dynamic interactions between HD-GYP and GGDEF domain proteins mediate virulence in *Xanthomonas campestris*, *Proceedings of the National Academy of Sciences 107*, 5989-5994.

42. Seshasayee, A. S. N., Fraser, G. M., and Luscombe, N. M. (2010) Comparative genomics of cyclic-di-GMP signalling in bacteria: post-translational regulation and catalytic activity, *Nucleic Acids Research 38*, 5970-5981.

43. Wolanin, P., Thomason, P., and Stock, J. (2002) Histidine protein kinases: key signal transducers outside the animal kingdom, *Genome Biology 3*, reviews3013.3011 - reviews3013.3018.

44. Potrykus, K., and Cashel, M. (2008) (p)ppGpp: Still Magical?, *Annual Review of Microbiology 62*, 35-51.

45. Levet-Paulo, M., Lazzaroni, J.-C., Gilbert, C., Atlan, D., Doublet, P., and Vianney, A. (2011) The Atypical Two-component Sensor Kinase Lpl0330 from *Legionella pneumophila* Controls the Bifunctional Diguanylate Cyclase-Phosphodiesterase Lpl0329 to Modulate Bis-(3′-5′)-cyclic Dimeric GMP Synthesis, *Journal of Biological Chemistry 286*, 31136-31144.

46. Srivastava, D., and Waters, C. M. (2012) A Tangled Web: Regulatory Connections between Quorum Sensing and Cyclic Di-GMP, *Journal of Bacteriology 194*, 4485-4493.

47. Bharati, B. K., Sharma, I. M., Kasetty, S., Kumar, M., Mukherjee, R., and Chatterji, D. (2012) A full-length bifunctional protein involved in c-di-GMP turnover is required for long-term survival under nutrient starvation in *Mycobacterium smegmatis*, *Microbiology 158*, 1415-1427.

48. Tam, R., and Saier Jr, M. H. (1993) Structural, functional, and evolutionary relationships among extracellular solute-binding receptors of bacteria, *Microbiological Reviews 57*, 320-346.

49. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, *Nucleic Acids Res 25*, 3389-3402.

50. Christen, M., Christen, B., Folcher, M., Schauerte, A., and Jenal, U. (2005) Identification and Characterization of a Cyclic di-GMP-specific

Phosphodiesterase and Its Allosteric Control by GTP, *Journal of Biological Chemistry 280*, 30829-30837.

51.   Navarro, M. V. A. S., Newell, P. D., Krasteva, P. V., Chatterjee, D., Madden, D. R., O'Toole, G. A., and Sondermann, H. (2011) Structural Basis for c-di-GMP-Mediated Inside-Out Signaling Controlling Periplasmic Proteolysis, *PLoS Biol 9*, e1000588.

52.   Navarro, M. V. A. S., De, N., Bae, N., Wang, Q., and Sondermann, H. (2009) Structural Analysis of the GGDEF-EAL Domain-Containing c-di-GMP Receptor FimX, *Structure 17*, 1104-1116.

53.   Qi, Y., Chuah, M. L. C., Dong, X., Xie, K., Luo, Z., Tang, K., and 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*, 2910-2917.

54.   Christen, B., Christen, M., Paul, R., Schmid, F., Folcher, M., Jenoe, P., Meuwly, M., and Jenal, U. (2006) Allosteric Control of Cyclic di-GMP Signaling, *Journal of Biological Chemistry 281*, 32015-32024.

55.   Duerig, A., Abel, S., Folcher, M., Nicollier, M., Schwede, T., Amiot, N., Giese, B., and Jenal, U. (2009) Second messenger-mediated spatiotemporal control of protein degradation regulates bacterial cell cycle progression, *Genes & Development 23*, 93-104.

56.   Hobley, L., Fung, R. K. Y., Lambert, C., Harris, M. A. T. S., Dabhi, J. M., King, S. S., Basford, S. M., Uchida, K., Till, R., Ahmad, R., Aizawa, S.-I., Gomelsky, M., and Sockett, R. E. (2012) Discrete Cyclic di-GMP-Dependent Control of Bacterial Predation versus Axenic Growth in *Bdellovibrio bacteriovorus*, *PLoS Pathog 8*, e1002493.

57.   Lee, V. T., Matewish, J. M., Kessler, J. L., Hyodo, M., Hayakawa, Y., and Lory, S. (2007) A cyclic-di-GMP receptor required for bacterial exopolysaccharide production, *Molecular Microbiology 65*, 1474-1484.

58.   Whitney, J. C., Colvin, K. M., Marmont, L. S., Robinson, H., Parsek, M. R., and Howell, P. L. (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*, 23582-23593.

59.   Ryjenkov, D. A., Simm, R., Römling, U., and 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*, 30310-30314.

60.   Habazettl, J., Allan, M. G., Jenal, U., and 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*, 14304-14314.

61.   Ko, J., Ryu, K.-S., Kim, H., Shin, J.-S., Lee, J.-O., Cheong, C., and Choi, B.-S. (2010) Structure of PP4397 Reveals the Molecular Basis for Different c-di-GMP Binding Modes by Pilz Domain Proteins, *Journal of Molecular Biology 398*, 97-110.

62.   Paul, K., Nieto, V., Carlquist, W. C., Blair, D. F., and Harshey, R. M. (2010) The c-di-GMP Binding Protein YcgR Controls Flagellar Motor Direction and Speed to Affect Chemotaxis by a "Backstop Brake" Mechanism, *Molecular Cell 38*, 128-139.

63.   Zorraquino, V., García, B., Latasa, C., Echeverz, M., Toledo-Arana, A., Valle, J., Lasa, I., and Solano, C. (2013) Coordinated Cyclic-Di-GMP Repression of *Salmonella* Motility through YcgR and Cellulose, *Journal of Bacteriology 195*, 417-428.

64.   Sudarsan, N., Lee, E. R., Weinberg, Z., Moy, R. H., Kim, J. N., Link, K. H., and Breaker, R. R. (2008) Riboswitches in Eubacteria Sense the Second Messenger Cyclic Di-GMP, *Science 321*, 411-413.

65.   Lee, E. R., Baker, J. L., Weinberg, Z., Sudarsan, N., and Breaker, R. R. (2010) An Allosteric Self-Splicing Ribozyme Triggered by a Bacterial Second Messenger, *Science 329*, 845-848.

66.   Barreteau, H., Kovac, A., Boniface, A., Sova, M., Gobec, S., and Blanot, D. (2008) Cytoplasmic steps of peptidoglycan biosynthesis, *FEMS Microbiology Reviews 32*, 168-207.

67.   Gautam, A., Vyas, R., and Tewari, R. (2011) Peptidoglycan biosynthesis machinery: A rich source of drug targets, *Critical Reviews in Biotechnology 31*, 295-336.

68.   Vollmer, W., Blanot, D., and De Pedro, M. A. (2008) Peptidoglycan structure and architecture, *FEMS Microbiology Reviews 32*, 149-167.

69.   Egan, A. J. F., and Vollmer, W. (2013) The physiology of bacterial cell division, *Annals of the New York Academy of Sciences 1277*, 8-28.

70.   Typas, A., Banzhaf, M., Gross, C. A., and Vollmer, W. (2012) From the regulation of peptidoglycan synthesis to bacterial growth and morphology, *Nat Rev Micro 10*, 123-136.

71.   Benson, T. E., Marquardt, J. L., Marquardt, A. C., Etzkorn, F. A., and Walsh, C. T. (1993) Overexpression, purification, and mechanistic study of UDP-*N*-acetylenolpyruvylglucosamine reductase, *Biochemistry 32*, 2024-2030.

72.   Dhalla, A. M., Yanchunas, J., Jr., Ho, H. T., Falk, P. J., Villafranca, J. J., and Robertson, J. G. (1995) Steady-state kinetic mechanism of *Escherichia coli* UDP-*N*-acetylenolpyruvylglucosamine reductase, *Biochemistry 34*, 5390-5402.

73.   Benson, T. E., Walsh, C. T., and Hogle, J. M. (1996) The structure of the substrate-free form of MurB, an essential enzyme for the synthesis of bacterial cell walls, *Structure 4*, 47-54.

74.   Benson, T. E., Harris, M. S., Choi, G. H., Cialdella, J. I., Herberg, J. T., Martin, J. P., and Baldwin, E. T. (2001) A structural variation for MurB: X-ray crystal structure of *Staphylococcus aureus* UDP-*N*-acetylenolpyruvylglucosamine reductase (MurB), *Biochemistry 40*, 2340-2350.

75.   Kim, M. K., Cho, M. K., Song, H. E., Kim, D., Park, B. H., Lee, J. H., Kang, G. B., Kim, S. H., Im, Y. J., Lee, D. S., and Eom, S. H. (2007) Crystal structure of UDP-*N*-acetylenolpyruvylglucosamine reductase (MurB) from *Thermus caldophilus*, *Proteins-Structure Function and Bioinformatics 66*, 751-754.

76.   Chen, M. W., Lohkamp, B., Schnell, R., Lescar, J., and Schneider, G. (2013) Substrate Channel Flexibility in *Pseudomonas aeruginosa* MurB Accommodates Two Distinct Substrates, *PLoS ONE 8*, e66936.

77.   Bae, E., Bingman, C. A., Bitto, E., Aceti, D. J., and Phillips, G. N. (2008) Crystal structure of *Arabidopsis thaliana* cytokinin dehydrogenase, *Proteins-Structure Function and Bioinformatics 70*, 303-306.

78.   Mattevi, A., Fraaije, M. W., Mozzarelli, A., Olivi, L., Coda, A., and van Berkel, W. J. H. (1997) Crystal structures and inhibitor binding in the octameric flavoenzyme vanillyl-alcohol oxidase: the shape of the active-site cavity controls substrate specificity, *Structure 5*, 907-920.

79.   Benson, T. E., Filman, D. J., Walsh, C. T., and Hogle, J. M. (1995) An enzyme-substrate complex involved in bacterial cell wall biosynthesis, *Nat Struct Mol Biol 2*, 644-653.

80.   Benson, T. E., Walsh, C. T., and Hogle, J. M. (1997) X-ray crystal structures of the S229A mutant and wild-type MurB in the presence of the substrate enolpyruvyl-UDP-*N*-acetylglucosamine at 1.8-Å resolution, *Biochemistry 36*, 806-811.

81.   Benson, T. E., Walsh, C. T., and Massey, V. (1997) Kinetic characterization of wild-type and S229A mutant MurB: Evidence for the role of Ser 229 as a general acid, *Biochemistry 36*, 796-805.

82.   Taku, A., and Anwar, R. A. (1973) Biosynthesis of uridine diphospho-*N*-acetylmuramic acid. IV. Activation of uridine diphospho-*N*-acetylenolpyruvylglucosamine reductase by monovalent cations, *Journal of Biological Chemistry 248*, 4971-4976.

83.   Constantine, K. L., Mueller, L., Goldfarb, V., Wittekind, M., Metzler, W. J., Yanchunas, J., Robertson, J. G., Malley, M. F., Friedrichs, M. S., and Farmer, B. T. (1997) Characterization of NADP(+) binding to perdeuterated MurB: Backbone atom NMR assignments and chemical-shift changes, *Journal of Molecular Biology 267*, 1223-1246.

84. Farmer, B. T. (1996) Localizing the NADP(+) binding site on the MurB enzyme by NMR, *Nat. Struct. Biol. 3*, 995-997.

85. Axley, M. J., Fairman, R., Yanchunas, J., Jr., Villafranca, J. J., and Robertson, J. G. (1997) Spectroscopic properties of *Escherichia coli* UDP-*N*-acetylenolpyruvylglucosamine reductase, *Biochemistry 36*, 812-822.

86. Bugg, T. D. H., Braddick, D., Dowson, C. G., and Roper, D. I. (2011) Bacterial cell wall assembly: still an attractive antibacterial target, *Trends in Biotechnology 29*, 167-173.

87. Bronson, J. J., DenBleyker, K. L., Falk, P. J., Mate, R. A., Ho, H.-T., Pucci, M. J., and Snyder, L. B. (2003) Discovery of the first antibacterial small molecule inhibitors of MurB, *Bioorganic & Medicinal Chemistry Letters 13*, 873-875.

88. Yang, Y., Severin, A., Chopra, R., Krishnamurthy, G., Singh, G., Hu, W., Keeney, D., Svenson, K., Petersen, P. J., Labthavikul, P., Shlaes, D. M., Rasmussen, B. A., Failli, A. A., Shumsky, J. S., Kutterer, K. M. K., Gilbert, A., and Mansour, T. S. (2006) 3,5-Dioxopyrazolidines, Novel Inhibitors of UDP-*N*-Acetylenolpyruvylglucosamine Reductase (MurB) with Activity against Gram-Positive Bacteria, *Antimicrobial Agents and Chemotherapy 50*, 556-564.

89. Mansour, T. S., Caufield, C. E., Rasmussen, B., Chopra, R., Krishnamurthy, G., Morris, K. M., Svenson, K., Bard, J., Smeltzer, C., Naughton, S., Antane, S., Yang, Y., Severin, A., Quagliato, D., Petersen, P. J., and Singh, G. (2007) Naphthyl Tetronic Acids as Multi-Target Inhibitors of Bacterial Peptidoglycan Biosynthesis, *ChemMedChem 2*, 1414-1417.

90. Kotaka, M., Dutta, S., Lee, H. C., Lim, M. J. M., Wong, Y., Rao, F., Mitchell, E. P., Liang, Z.-X., and Lescar, J. (2009) Expression, purification and preliminary crystallographic analysis of *Pseudomonas aeruginosa* RocR protein, *Acta Crystallographica Section F 65*, 1035-1038.

91. Rao, F., Qi, Y., Chong, H. S., Kotada, M., Li, B., Lescar, J., Tang, K., and Liang, Z.-X. (2009) The functional role of a conserved loop in EAL domain-based c-di-GMP specific phosphodiesterase, *J. Bacteriol. 191*, 4722-4731.

92. Sun, P. D., Radaev, S., and Kattah, M. (2002) Generating isomorphous heavy-atom derivatives by a quick-soak method. Part I: test cases, *Acta Crystallographica Section D 58*, 1092-1098.

93. Barends, T. R. M., Hartmann, E., Griese, J. J., Beitlich, T., Kirienko, N. V., Ryjenkov, D. A., Reinstein, D. A., Shoeman, R. I., Gomelsky, M., and Schlichting, I. (2009) Structure and mechanism of a bacterial light-regulated cyclic nucleotide phosphodiesterase, *Nature 459*, 1015-1018.

94. Minasov, G., Padavattan, S., Shuvalova, L., Brunzelle, J. S., Miller, D. J., Basle, A., Massa, C., Collart, F. R., Schirmer, T., and Anderson, W. F. (2009) Crystal structures of YkuI and its complex with second messenger c-di-GMP suggests catalytic mechanism of phosphodiester bond cleavage by EAL domains, *J. Biol. Chem. 284*, 13174-13184.

95. Tchigvintsev, A., Xu, X., Singer, A., Chang, C., Brown, G., Proudfoot, M., Cui, H., Flick, R., Anderson, W. F., Joachimiak, A., Galperin, M. Y., Savchenko, A., and Yakunin, A. F. (2010) Structural Insight into the Mechanism of c-di-GMP Hydrolysis by EAL Domain Phosphodiesterases, *J. Mol. Biol. 402*, 524-538.

96. Toro-Roman, A., Mack, T. R., and Stock, A. M. (2005) Structural Analysis and Solution Studies of the Activated Regulatory Domain of the Response Regulator ArcA: A Symmetric Dimer Mediated by the α4-β5-α5 Face, *Journal of Molecular Biology 349*, 11-26.

97. Lupas, A., Van Dyke, M., and Stock, J. (1991) Predicting coiled coils from protein sequences, *Science 252*, 1162-1164.

98. McDonnell, A. V., Jiang, T., Keating, A. E., and Berger, B. (2006) Paircoil2: improved prediction of coiled coils from sequence, *Bioinformatics 22*, 356-358.

99. Ho, C. L., Koh, S. L., Chuah, M. L. C., Luo, Z., Tan, W. J., Low, D. K. S., and Liang, Z.-X. (2011) Rational Design of Fluorescent Biosensor for Cyclic di-GMP, *ChemBioChem 12*, 2753-2758.

100. Anantharaman, V., Balaji, S., and Aravind, L. (2006) The signaling helix: a common functional theme in diverse signaling proteins, *Biol Direct 1*, 25.

101. Parkinson, J. S. (2010) Signaling Mechanisms of HAMP Domains in Chemoreceptors and Sensor Kinases, *Annual Review of Microbiology 64*, 101-122.

102. Oh, B. H., Pandit, J., Kang, C. H., Nikaido, K., Gokcen, S., Ames, G. F., and Kim, S. H. (1993) Three-dimensional structures of the periplasmic lysine/arginine/ornithine-binding protein with and without a ligand, *Journal of Biological Chemistry 268*, 11348-11355.

103. Björkman, A. J., and Mowbray, S. L. (1998) Multiple open forms of ribose-binding protein trace the path of its conformational change, *Journal of Molecular Biology 279*, 651-664.

104. Sooriyaarachchi, S., Ubhayasekera, W., Park, C., and Mowbray, S. L. (2010) Conformational Changes and Ligand Recognition of *Escherichia coli* D-Xylose Binding Protein Revealed, *Journal of Molecular Biology 402*, 657-668.

105. Moynie, L., Schnell, R., McMahon, S. A., Sandalova, T., Boulkerou, W. A., Schmidberger, J. W., Alphey, M., Cukier, C., Duthie, F., Kopec, J., Liu, H., Jacewicz, A., Hunter, W. N., Naismith, J. H., and Schneider, G. (2013) The AEROPATH project targeting *Pseudomonas aeruginosa*: crystallographic studies for assessment of potential targets in early-stage drug discovery, *Acta Crystallographica. Section F, Structural Biology and Crystallization Communications 69*, 25-34.

106. Harris, M. S., Herberg, J. T., Cialdella, J. I., Martin, J. P., Benson, T. E., Choi, G. H., and Baldwin, E. T. (2001) Crystallization and preliminary X-ray analysis of UDP-*N*-acetylenolpyruvylglucosamine reductase (MurB) from *Staphylococcus aureus*, *Acta Crystallographica. Section D, Biological Crystallography 57*, 1032-1035.

107. Harding, M. (2002) Metal-ligand geometry relevant to proteins and in proteins: sodium and potassium, *Acta Crystallographica. Section D, Biological Crystallography 58*, 872-874.

108. Shabala, L., Bowman, J., Brown, J., Ross, T., McMeekin, T., and Shabala, S. (2009) Ion transport and osmotic adjustment in *Escherichia coli* in response to ionic and non-ionic osmotica, *Environmental Microbiology 11*, 137-148.

109. Pejchal, R., Sargeant, R., and Ludwig, M. L. (2005) Structures of NADH and CH_3-H_4Folate Complexes of *Escherichia coli* Methylenetetrahydrofolate Reductase Reveal a Spartan Strategy for a Ping-Pong Reaction, *Biochemistry 44*, 11447-11457.

110. Sumner, J. S., and Matthews, R. G. (1992) Stereochemistry and mechanism of hydrogen transfer between NADPH and methylenetetrahydrofolate in the reaction catalyzed by methylenetetrahydrofolate reductase from pig liver, *Journal of the American Chemical Society 114*, 6949-6956.