Structural studies of enzymes regulating bacterial life style and cell wall biogenesis

Chen, Ming Wei

2014

Chen, M. W. (2014). Structural studies of enzymes regulating bacterial life style and cell wall biogenesis. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/59968>

<https://doi.org/10.32657/10356/59968>

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

CHEN MING WEI

CHEN MING WEI

School of Biological Sciences

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

The thesis work has been conducted within a joint PhD programme of
Doctor of Philosophy between the Nanyang Technological University of
Singapore and Karolinska Institutet, Stockholm, Sweden.

2014

This NTU thesis is adapted from the same-titled KI version (ISBN 978-91-7549-402-9), published by Karolinska Institutet and printed by Larserics Digital Print AB.

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. 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. 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. Robert Schnell, for being a wonderful scientist with the useful advice, “science is an endeavour of overtime.” I'm also grateful to Bernhard Lohkamp, who has been ever helpful to me in the unfathomable universe that is CCP4i. I'd also like to acknowledge all past and present members of the MSB and Lescar groups for their friendship, inspirations and assistance.

I would like to thank Lars Nordenskiöld 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.

To my friends – old and new, at home and abroad – I appreciate your heartfelt support. 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.

CONTENTS

1	Introduction.....	1
1.1	Cyclic di-GMP is a ubiquitous second messenger in Bacteria.....	1
1.2	Enzymes that regulate cyclic di-GMP levels: diguanylate cyclases and phosphodiesterases	3
1.2.1	The GGDEF domain	3
1.2.2	The EAL domain.....	5
1.2.3	The HD-GYP domain	7
1.3	Regulatory domains and signalling pathways that regulate DGC and PDE activity.....	8
1.3.1	RocR is a response regulator in c-di-GMP signalling.....	8
1.3.2	Sensor domains that regulate DGC and PDE activity.....	9
1.4	Enzymatic paradox? The prevalence of GGDEF-EAL tandem domains	10
1.4.1	Tandem domains containing active GGDEF and EAL.....	11
1.4.2	Half-active or completely inactive tandem proteins.....	12
1.5	Cyclic di-GMP receptors that influence bacterial activity	13
1.6	Peptidoglycan as an important structural element in the bacterial cell wall	14
1.7	Cytoplasmic enzymes that synthesise peptidoglycan precursors....	15
1.8	MurB synthesises UDP-N-acetylglucosamine-enolpyruvate.....	17
1.9	Magic bullet against MurB: an ongoing search	19
2	Aims of this thesis	21
3	Results and Discussion.....	22
3.1	Structural studies of RocR, a phosphoreceiver-EAL protein from <i>Pseudomonas aeruginosa</i> (Paper I)	22
3.1.1	Recombinant protein production of RocR-R286W, RocR-D56N and RocR-wt	22
3.1.2	Crystallisation and preparation of heavy metal derivatives	22
3.1.3	Structure determination of RocR-R286W	23
3.1.4	Open and closed conformations in the RocR tetramer.....	24
3.1.5	Comparison of the RocR variants by using small-angle X-ray scattering	25
3.1.6	Potential mechanism of RocR regulation upon phosphorylation	26
3.2	Structural studies of Tbd1265, a periplasmic sensor protein with a GGDEF-EAL cytoplasmic module (Paper II, manuscript in progress).....	28
3.2.1	Recombinant production of GGDEF-EAL constructs Tbd1265 _{DUAL} and Tbd1265 _{ΔDUAL}	28
3.2.2	In-solution oligomerisation studies of GGDEF-EAL constructs	29
3.2.3	Comparison of DGC and PDE activities of the tandem domain constructs	30
3.2.4	Crystallisation and structure determination	30
3.2.5	Proposed regulatory mechanism of Tbd1265.....	32
3.2.6	Recombinant production of the periplasmic binding protein (PBP) constructs	33
3.2.7	Crystallisation of PBP-A2 and data collection.....	34

3.3	Structural studies of the MurB enzyme from <i>Pseudomonas aeruginosa</i> (paper III).....	36
3.3.1	Recombinant protein production of PaMurB.....	36
3.3.2	Crystallisation and structure determination of the PaMurB-FAD- NADP(H) ternary complex	37
3.3.3	Crystal structure of PaMurB.....	37
3.3.4	NADPH and UNAGEP bind to the same flexible substrate channel 39	
3.3.5	Active site potassium ensures efficient hydride transfer	40
3.3.6	Implications for drug design against MurB.....	41
4	Concluding Remarks.....	42
5	References	44

SUMMARY

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 Vonnrhein, 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.

LIST OF ABBREVIATIONS

5'-pGpG	Linear di-GMP; 5'-phosphoguananylyl-(3'→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'→5')-cyclic diguanylate
cGMP	Cyclic GMP, guanosine 3',5'-cyclic monophosphate
DAP	Diaminopimelic acid
DGC	Diguanylate cyclase
EcMurB	<i>Escherichia coli</i> 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	<i>Pseudomonas aeruginosa</i> MurB
PBP	Periplasmic binding protein
PCR	Polymerase chain reaction
PDB	Protein Data Bank
PDE	Phosphodiesterase
ppGpp	5'-diphosphate 3'-diphosphate guanosine
rmsd	Root mean squared deviation
SAD	Single-wavelength anomalous dispersion
SaMurB	<i>Staphylococcus aureus</i> 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'→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 GGDEF 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 GTP α S (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

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 $\alpha 9$ and the long helix $\alpha 11$ (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 RocA1 and RocA1 that act antagonistically *in vivo* (35, 36). RocA1 contains a DNA-binding domain that activates transcription of *cupB* and *cupC* fimbriae 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 $\rightarrow$ 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

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 $\beta\alpha\beta$ 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

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 bactericide 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-dioxypyrazolidines with IC₅₀ 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 IC₅₀ 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 uncharted 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 previously 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.

panel shows the closed monomers A (black) and B (white).

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

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 $\alpha 4'$ and $\alpha 5'$ 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 $\alpha 4'$ and $\alpha 5'$, 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_{ΔDUAL} tandem domain constructs are also shown.

3.2.1 Recombinant production of GGDEF-EAL constructs Tbd1265_{DUAL} and Tbd1265_{ΔDUAL}

Two GGDEF-EAL tandem domain constructs, Tbd1265_{DUAL} (residues 289 – 758) and Tbd1265_{ΔDUAL} (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_{ΔDUAL} 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_{ΔDUAL} (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_{ΔDUAL} (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_{ΔDUAL} (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_{ΔDUAL} (Fig 4, Paper II). Thus, the N-helix which is present in Tbd1265_{DUAL} but not in Tbd1265_{ΔDUAL} 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 $\alpha 15$, loop 6 ($\beta 12$ - $\alpha 13$ loop) and the short helix $\alpha 13$ 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 $\beta 5$ - $\alpha 2$ and $\beta 3$ - $\alpha 6$ 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 $c=189.7$ Å. 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
<i>Data collection statistics</i>		
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)
$\langle I/\sigma \rangle$	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 P6₁, 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 (β 1- β 2- β 5- β 3 and β 6- β 7- β 12- β 10- β 11) packed against flanking helices α 1 and α 2, respectively. A buried FAD-binding pocket is formed by the GGG motif on the β 3- β 4 loop and the C-terminal strand β 20 contributed by domain I, along with the β 7- α 3 loop, the η 1- α 3- β 8 region, and α 4- β 12 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 α 5- α 6, while lobe 2 is close to domain I and contains the β 13- β 19- β 18- β 17 sheet and the β 14- α 7- β 15- β 16 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 Ca 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 C α 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 $\alpha 4$ and $\alpha 5$, as well as the protruding $\beta\alpha\beta\beta$ 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 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., Moskvina, O. V., and Gomelsky, M. (2005) Cyclic Diguanylate Is a Ubiquitous Signaling Molecule in Bacteria: Insights into Biochemistry of the GGDEF Protein Domain, *Journal of Bacteriology* 187, 1792-1798.
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 BeF₃⁻-modified response regulator PleD: implications for diguanylate cyclase activation, catalysis, and feedback inhibition, *Structure* 15, 915-927.

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., Griesse, 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., Vonnrhein, 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 Diguanilate 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 Diguanilate 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-Dioxypyrazolidines, 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 Tetrone 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 $\alpha 4$ - $\beta 5$ - $\alpha 5$ 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 CH₃-H₄Folate 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.

Structural Insights into the Regulatory Mechanism of the Response Regulator RocR from *Pseudomonas aeruginosa* in Cyclic Di-GMP Signaling

Vei Chen,^{a,d} Masayo Kotaka,^{a*} Clemens Vonnheim,^b Gérard Bricogne,^b Feng Rao,^{a*} Mary Lay Cheng Chuah,^a Dmitri Svergun,^c Schneider,^d Zhao-Xun Liang,^a and Julien Lescar^{a,e}

^a Biological Sciences, Nanyang Technological University, Singapore^a; Global Phasing, Ltd., Sheraton House, Castle Park, Cambridge, United Kingdom^b; European Synchrotron Radiation Facility, Grenoble, France^c; European Molecular Biology Laboratory, Hamburg Unit, EMBL c/o DESY, Hamburg, Germany^d; Department of Medical Biochemistry and Biophysics, Karolinska Institutet, Stockholm, Sweden^e; and AFMB CNRS UMR 6098, Marseille, France^e

The second-messenger cyclic di-GMP (c-di-GMP) plays a central role in the regulation of motility, virulence, and biofilm formation in many pathogenic bacteria. EAL domain-containing phosphodiesterases are the major signaling proteins responsible for the degradation of c-di-GMP and maintenance of its cellular level. We determined the crystal structure of a single mutant (W) of the response regulator RocR from *Pseudomonas aeruginosa* to show that RocR exhibits a highly unusual tetrameric structure. The four subunits are arranged around a single dyad, with the four subunits adopting two distinctly different conformations. Subunits A and B adopt a conformation with the REC domain located above the c-di-GMP binding pocket, whereas subunits C and D adopt a conformation with the REC domain swung to the side of the EAL domain. Remarkably, the access to the substrate-binding site of the EAL domains of the open subunits C and D are blocked in *trans* by the REC domains of subunits A and B, indicating that only two of the four active sites are engaged in the degradation of c-di-GMP. In conjunction with biochemical and biophysical data, we propose that the structural changes within the REC domains triggered by the phosphorylation are transmitted to the EAL domain active sites through a pathway that traverses the dimerization interfaces composed of a conserved regulatory loop and the neighboring motifs. This exquisite mechanism reinforces the crucial role of the regulatory loop and suggests that similar regulatory mechanisms may be operational in many EAL domain proteins, considering the preservation of the dimerization interface and the spatial arrangement of the regulatory domains.

Discovered as an allosteric regulator of cellulose synthase in *Acetobacter xylinus*, cyclic-di-GMP (c-di-GMP) mediates a variety of bacterial cellular functions mainly associated with the transition between a planktonic and a community-based lifestyle (7, 17, 35, 36). Dissection of the signaling network that regulate cellular c-di-GMP levels is expected to unveil molecular mechanisms underlying biofilm formation, a major factor to the persistent infections caused by many pathogenic bacteria. Cellular concentrations of c-di-GMP are controlled by two types of enzymes with opposing activities: GGDEF domain proteins with diguanylate cyclase activity and proteins containing either an EAL or a HD-GYP domain, which have c-di-phosphodiesterase (PDE) activity (38–41). While GGDEF domains catalyze the synthesis of c-di-GMP, EAL and HD-GYP domains hydrolyze c-di-GMP to generate either linear 5'-pGpG or cyclic pGpG. Many bacterial genomes contain multiple copies of genes encoding GGDEF, EAL and HD-GYP domains, an observation consistent with their prominent roles in c-di-GMP signaling. However, EAL, or HD-GYP domains usually do not function as alone proteins but in association with various regulatory proteins that modulate their enzymatic activities in response to external stimuli (4, 14, 29). A wide variety of regulatory domains sense the environmental signals from the surroundings and are fused to the three enzymatically active domains. Elucidating how the regulatory domains modulate the enzymatic activities of c-di-GMP metabolizing proteins is crucial for a better understanding of the molecular mechanism of c-di-GMP signaling. Restructuring studies on proteins containing GGDEF domain, including the response regulators PleD and WspR, suggested that

enzymatic regulation is mainly achieved through the control of their oligomerization state (9, 50). The crystal structure of the protein YkuI from *Bacillus subtilis* that contains an enzymatically inactive EAL domain and a putative regulatory PAS domain revealed the β/α barrel structure of the EAL domain (26). Based on site-directed mutagenesis studies, a regulatory mechanism was proposed for the EAL domain-containing protein RocR from *Pseudomonas aeruginosa* (31). In this mechanism, a loop of RocR [named "loop 6" to be consistent with other $(\alpha/\beta)_8$ barrel proteins, where the corresponding loop acts as a lid for substrate binding and product release] plays an important role for binding the catalytic metal ion and the substrate. More recently, Barends et al. proposed a regulatory mechanism for the BLUF photoreceptor-regulated EAL domain in the BlrP1 protein from *Klebsiella*

Received 10 April 2012 Accepted 25 June 2012

Published ahead of print 29 June 2012

Address correspondence to Julien Lescar, julien@ntu.edu.sg, or Zhao-Xun Liang, zxliang@ntu.edu.sg.

* Present address: Masayo Kotaka, Department of Physiology, University of Hong Kong, Pokfulam, Hong Kong, People's Republic of China; Feng Rao, The Solomon H. Snyder Department of Neuroscience, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA.

M.W.C. and M.K. contributed equally to this article.

Supplemental material for this article may be found at <http://jb.asm.org/>.

Copyright © 2012, American Society for Microbiology. All Rights Reserved.

doi:10.1128/JB.00560-12

TABLE 1 Data collection statistics

data set (synchrotron)	Native (NSRRC)	SeMet (NSRRC)	K ₂ PtCl ₄ (SLS PXIII)
wavelength (Å)	1.00	0.97910	1.07216
resolution (Å) ^a	30–2.50 (2.59–2.50)	30–3.95 (4.09–3.95)	20.00–4.29 (4.55–4.29)
space group	P6 ₁ 22	P6 ₁ 22	P6 ₁ 22
cell parameters a/b/c (Å)	118.8/118.8/495.1	119.1/119.1/449.5	120.3/120.3/491.1
β/γ (°)	90/90/120	90/90/120	90/90/120
no. of unique reflections	60,479	19,193	26,870
redundancy	11.9 (8.8)	20.1 (19.7)	2.43
R	61.9 (2.9)	28.7 (7.4)	10.81 (2.23)
completeness (%)	84.1 (75.0)	99.7 (98.4)	98.5 (97.0)
merge ^b	0.046 (0.495)	0.161 (0.427)	0.094 (0.579)

^aValues in parentheses indicate values in the highest resolution shell.^bmerge = $\sum |I_j - \langle I \rangle| / \sum I_j$, where I_j is the intensity of an individual reflection, and $\langle I \rangle$ is the average intensity of that reflection.

leumoniae, based on the crystal structures of the protein crystallized at different pH values. The mechanism proposed for BlrP1 involves the subtle repositioning of two catalytic metal-ions through conformational changes in the β5-α5 loop (equivalent to loop 6 in RocR), as well as in other structural motifs (2).

The response regulator RocR from the opportunistic pathogen *aeruginosa* contains an N-terminal phosphoreceiver (REC) domain and a C-terminal EAL domain that possesses c-di-GMP specific PDE activity (23, 24, 32). Based on sequence homology and *in vivo* studies, the REC domain of RocR is predicted to accept a phosphate group from the cognate membrane-bound histidine kinase sensor RocS1. Phosphorylation of RocR putatively modulates the enzymatic activity of its EAL domain and hence of the catalytic level of c-di-GMP. To understand how structural changes originating from the regulatory REC domain of RocR are transmitted to its EAL domain and hence regulate its catalytic activity, we set out to determine the crystal structure of RocR. Although the wild-type RocR could not be crystallized, a 2.5-Å crystal structure of the R286W mutant that exhibits lower catalytic activity was determined. The crystal structure reveals a highly unusual tetrameric arrangement featuring two markedly different conformations for the four subunits of RocR, with only two active sites accessible for substrate binding. The overall quaternary structures of the wild-type RocR protein and of two single mutants (R286W and D56N) are indistinguishable in solution, as assessed by small-angle X-ray scattering (SAXS) experiments. In conjunction with previously acquired biochemical and biophysical data, we propose a regulatory mechanism that involves the propagation of structural changes from the REC domain to the active site of the EAL domain through a dimerization interface and an important functional loop motif.

MATERIALS AND METHODS

Cloning, expression, purification, and crystallization. The cloning, expression, and purification of RocR was described previously (20). In brief, the RocR protein was crystallized at 18°C via the sitting-drop vapor-diffusion method by mixing 1 μl of protein with an equal volume of a solution containing 15 to 20% (wt/vol) PEG 3350 (Hampton Research), 0.1 M HEPES (pH 7.5), 0.2 M sodium tartrate, and 5% glycerol. Rod-shaped crystals measuring up to 0.3 mm in length grow over the course of 10 to 14 days. DNA sequencing revealed a mutation introduced by the PCR at position 286 of the amino acid sequence (R286W). The wild-type RocR protein fails to crystallize in the same condition as the RocR mutant, possibly because of a crystal contact established between side chains of p²⁸⁶ (subunit A) and Arg³⁷⁸ from a neighboring molecule. C-di-GMP

used for cocrystallization or soaking was synthesized enzymatically using a thermophilic DGC protein (30).

Expression, purification, and crystallization of the selenated RocR protein. The RocR-pET26b plasmid was transformed into *Escherichia coli* B834 (DE3) (Novagen). Upon reaching an optical density at 600 nm (OD₆₀₀) of 0.4, cells grown at 310 K in Luria-Bertani medium supplemented with 50 μg of kanamycin/ml were harvested by centrifugation and resuspended in selenomethionine base medium (Molecular Dimensions). The washing step was repeated, and the cells were inoculated into 2 liters of pre-aerated, prewarmed selenomethionine expression medium (Molecular Dimensions) supplemented with 40 mg of L-selenomethionine/liter and 50 μg of kanamycin/ml. Upon reaching an OD₆₀₀ of 0.6, the culture was cooled to 301 K, and protein expression was induced with IPTG (isopropyl-β-D-thiogalactopyranoside) at a final concentration of 0.1 mM. Protein purification and crystallization was carried out as described for the native protein (20), except the concentration of PEG 3350 was 10 to 15% (wt/vol).

Enzymatic assay. The phosphodiesterase activity of the wild-type RocR and mutant R286W was assessed by monitoring the formation of the product 5'-pGpG from the hydrolysis of c-di-GMP. The reaction assay was performed by incubating the enzymes (1 μM) and c-di-GMP (20 μM) at various temperatures (25, 40, 50, 60, and 70°C) in 100 mM Tris-HCl (pH 8.0), 20 mM KCl, and 25 mM MgCl₂ for 5 min. Reactions were stopped by adding 1/10 volume of 1 M CaCl₂, and the progress of c-di-GMP hydrolysis was monitored by using the Agilent LC1200 system (mobile phase of 20 mM triethylammonium bicarbonate [pH 7.0] and 5% methanol, at a rate 1 ml/min) with a XDB-C18 column.

X-ray data collection and structure determination. Data collection statistics for crystals of the native protein, for the selenomethionyl protein (SeMet) and the tetrachloroplatinate(II) derivative are given in Table 1. Heavy-metal derivative crystals were prepared using the JBScreen Heavy kit (Jena Biosciences). Native RocR crystals were soaked with 10 mM K₂PtCl₄ in a modified reservoir solution devoid of sodium tartrate for 10 min. The data sets collected at the Swiss Light Source (Paul Scherrer Institut, Switzerland, beamline PXIII) for the Pt derivative at the LIII edge, and at the National Synchrotron Radiation Research Center (Taiwan) for the selenomethionyl protein and native protein, were processed with the CCP4 program suite (6) or with the program XDS (18). Using the program SHELXD (43), four Pt sites were found. Electron density maps using either SIRAS or SAD phases calculated with the program SHARP (49) were not interpretable. However, the phases derived from the Pt derivative allowed the identification of 38 Se positions (out of a total of 40 Se), using anomalous differences collected at the selenium absorption edge. Using the program PROFESS (6), a 2-fold NCS could be located that mapped one set of Se positions onto the other. An initial map using SAD phases calculated using program SHARP could thus be averaged using 2-fold NCS and the program DM (6), yielding a partially interpretable map. Further density averaging was performed using a mask covering a

2 Refinement statistics

Parameter	Observation
Resolution (Å)	20.00–2.50 (2.57–2.50)
Observed reflections	57,265 (1,983)
Protein atoms	11,886
	0.216 (0.310)
	0.285 (0.353)
Bond length (Å)	0.015
Bond angles (°)	1.635
Strand plot (no. of residues) ^b	
Observed	1,320 (99.5)
Allowed	5 (0.4)
Forbidden	1 (0.1)

$\sum ||F_o| - |F_c|| / \sum |F_o|$, where F_o denotes the observed structure factor amplitude, F_c the structure factor amplitude calculated from the model. R_{free} is the same as R but calculated with 5% (3,044) of the randomly chosen reflections omitted from refinement. RMSD, root mean square deviation. Percentage is given in parentheses.

dimer. Program BUCANEER (8) was then used to build a partial model that contained 1414 residues. Several rounds of manual model building were then performed using program COOT (11), interspersed with refinement with program REFMAC (Table 2) (6). Interface areas and energies were calculated by the protein interfaces, surfaces, and analysis service PISA (22). Movie morphs were created by the Database of Molecular Movements (12), and figures were generated with the program PyMOL (Warren DeLano Scientific). Structure-based alignment against homologs was generated with ESPript (16). Interface residues during tetramer formation are listed in Table S1 and Fig. S1B in supplemental material.

Single-angle X-ray scattering. X-ray scattering data from the wild-type RocR and its mutants were collected according to standard procedures on beamline (34) of the EMBL Hamburg at the Deutsches Elektronen-Synchrotron using a robotic sample changer (37) and a Pilatus 1 M (Dectris, Villigen, Switzerland). The scattering patterns were recorded using a sample detector distance of 2.7 m and a wavelength of 1.0 Å, covering the range of momentum transfer $0.01 < s < 0.5 \text{ Å}^{-1}$. To minimize radiation damage, eight 15-s exposures were collected for each protein sample, and no radiation effect was observed. All of the data sets were measured for at least three solute concentrations each, in the range from 1 to 8 mg/ml. The data were normalized to the intensity of the transmitted beam, and the scattering of the buffer was subtracted. The data were scaled for concentration, the data was extrapolated to infinite dilution using PRIMUS (19) and the radius of gyration R_g . The distribution functions were computed by GNOM (44), providing maximum particle size D_{max} , and the low-resolution shape of the wild-type RocR was determined *ab initio* by DAMMIF (13). The scattering from the high-resolution models was computed by CRY SOL (45).

RESULTS

Crystal structure of the R286W mutant of RocR. The RocR protein is composed of an N-terminal regulatory REC domain connected by a linker of ~13 residues to a C-terminal EAL domain (Fig. 1A). Despite extensive screening of crystallization conditions, the wild-type RocR could not be crystallized. Instead, a mutant protein (R286W) was found to be more amenable to crystallization. Enzymatic assays show that at ambient temperatures, the R286W mutant is active with lower enzymatic activity than the wild-type protein (Fig. 1D). However, the mutant protein shares thermal stability with the wild-type RocR, recovering robust catalytic activity at elevated temperatures. The crystal structure of the R286W mutant was determined at a resolution of 2.5 Å

using SAD phasing at the Se absorption edge (see Materials and Methods and Tables 1 and 2). The asymmetric unit of the present crystal form contains one tetramer assembled around a single noncrystallographic dyad (Fig. 1B), a finding consistent with the observation that RocR forms tetramers in solution (31, 32). Unexpectedly, the four RocR subunits in the crystal structure adopt two distinct conformations (Fig. 1C and 2). Two subunits adopt a closed conformation (hereafter labeled “A” and “B”, with a root mean square deviation [RMSD] of 1.01 Å) with their REC domain situated on top of the $(\beta/\alpha)_8$ barrel of their EAL domain (Fig. 2A), while the other two subunits adopt an open conformation (“C” and “D”, RMSD = 1.06 Å) with their REC domains shifted to the side of their EAL domains (Fig. 2B). Having brought the EAL domains of a closed and an open subunit in coincidence, a 67° rotation and a translation of 29 Å is required to superimpose their REC domains (Fig. 2C). The tetrameric structure of RocR is rather unique because most EAL domain proteins form dimers through a conserved EAL-EAL dimerization interface (2, 26, 46). A dimeric form of RocR was observed in solution for a few RocR mutants with compromised enzymatic activity, suggesting that the RocR tetramer could be formed by two EAL dimers (31, 32). The crystal structure reveals a set of interactions between residues projecting from helices α_8 and α_{10} and from loop 6 that stabilize the EAL-EAL interface between subunits A and C (or their equivalents B and D, by the dyad) (Fig. 1C and see Fig. S1 in the supplemental material). This mode of interaction is also found in the dimeric EAL proteins TBD1265 from *Thiobacillus denitrificans* (46), YkuI from *Bacillus subtilis* (26), and BlrP1 from *Klebsiella pneumoniae* (2) with the same structural elements involved in stabilizing the EAL-EAL interface (see Fig. S1A in the supplemental material). However, in contrast to these proteins, the unusual tetrameric structure of RocR arises from additional intra- and intersubunit interactions between EAL and REC domains (see Fig. S1B in the supplemental material), which result in a total buried surface area of 13,820 Å². The most extensive interactions are observed between the open subunits C and D, which lead to a head-to-tail association between their respective REC and EAL domains (see Fig. S1B in the supplemental material). The REC_A and REC_B domains from the closed subunits occupy the central core of the tetramer with the putative phosphorylation sites Asp⁵⁶ pointing inside the tetramer and largely inaccessible from the solvent. Further stabilization of the tetramer is provided by interactions between REC_B and EAL_C (or their equivalent by the dyad REC_A with EAL_D). Importantly, the latter interactions give rise to the inhibition of the EAL domains in *trans* as described in details below.

Quaternary structures of wild-type and mutant RocR in solution. The observation of two markedly distinct conformations for the RocR subunits raises the possibility of the existence of a partially or fully open tetrameric structure in solution, with domains REC_A and REC_B swung outward and the catalytic sites of domains EAL_C and/or EAL_D accessible. A hypothetical fully open conformation of RocR featuring a tetramer with a 222 symmetry is shown in Fig. S2A in the supplemental material. We used the technique of SAXS to address the important question of whether the wild-type and R286W mutant proteins adopt closed structures in solution, similar to the mutant structure revealed by X-ray crystallography. The processed experimental scattering from the wild-type RocR displayed in Fig. 3A (curve 1), and the overall parameters computed from the scattering data clearly indicate that the protein is tetrameric in solution. The experimental radius of gy-

en et al.

FIG 1 Structure of the RocR R286W mutant protein. (A) Domain organization of RocR. (B) The RocR tetramer is shown from three angles separated by 90°. Its subunits are colored green (subunit A), cyan (subunit B), magenta (subunit C), and yellow (subunit D). The position of the single dyad that runs through the tetramer is indicated in each panel. The region involved in forming intermolecular EAL-EAL contacts is colored in white, and a magnified view of this interaction is shown in panel C. (C) Schematic depiction of the RocR tetramer highlighting the open conformation of its C and D subunits, which form an addle-like structure into which the REC domains of the closed subunits A and B are inserted, forming the core of the tetramer. The EAL active sites are represented as three black triangles. EAL active sites of subunits C and D are concealed (crossed circle) in *trans* by the REC domains located at the center of the tetramer. The interdomain linker is depicted as a line (dashed line if it is partially disordered in the crystal structure). (D) Enzymatic activity of the wild-type and R286W mutant RocR. (Left panel) HPLC analysis of c-di-GMP degradation in the presence of no enzyme and R286W mutant at 25 and 60°C. (Center and right panels) HPLC analysis of c-di-GMP degradation for wild type and R286W mutant across temperatures. The mutant exhibits diminished activity that is recovered at elevated temperatures. Reaction conditions are described in the experimental section.

tion $R_g = 37 \pm 1$ Å and the maximum particle size $D_{\max} = 0 \times 10$ Å fully agree with the values computed from the crystal structure of RocR ($R_g = 36.5$ Å, $D_{\max} = 116$ Å) but not with a partially or fully open tetrameric structure (see Fig. S2A in the supplemental material). Moreover, the low-resolution molecular shape determined from the experimental data neatly matches the

crystallographic model (Fig. 3B). The scattering pattern computed from the closed crystal structure fits the experimental data very well with discrepancy $\chi = 1.08$ (Fig. 3A, curve 2). In contrast, the scattering patterns computed from the putative half-open and open models (Fig. 3A, curves 3 and 4, respectively) display substantial systematic deviations from the experimental data and

A and B) Closed (subunits A and B, in cyan, in panel A) and open conformations (subunits C and D, in magenta, in panel B) that constitute the RocR tetramer. Residues from the REC and EAL active sites are represented as yellow sticks. (C) Superposition of the two conformers based on their EAL domains. The rotation (angle needed to bring their respective REC domains into coincidence) is indicated.

very high discrepancy values of $\chi = 4.7$ and 11.3, respectively. These findings suggest that, despite the altered catalytic activity exhibited by the R286W mutant, the overall closed conformation observed in the crystal structure is likely to be adopted

by both the wild-type RocR and the R286W mutant in solution. Note that the residue Trp²⁸⁶ is from the helix $\alpha 7$ of the EAL domain and situated at the REC_C-EAL_D and REC_D-EAL_C interfaces that may play a role in the regulatory mechanism, as we will discuss later. It was

SAXS analysis of the conformation of RocR in solution. (A) Experimental scattering from the wild-type protein (1, black dots), calculated scattering from the tetramer observed in the crystal (2, blue line); calculated scattering from the putative half-open and open RocR tetramers (3 and 4, red and green dashed lines). The log of the scattering intensity is plotted as a function of momentum transfer $s = 4\pi \sin(\theta)/\lambda$, where θ is the scattering angle and λ is the X-ray wavelength. (B) ab initio shape reconstructed from the wild-type RocR data (gray mesh) superimposed on the structure of the crystallographic tetramer using SUPCOMB.

en et al.

G 4 Observed conformations in the individual domains of RocR. (A) Superposition of REC_B (cyan, side chains shown as white sticks) with REC_C (yellow, side chains shown as black sticks). Phe¹⁰⁵, the “switch residue” located on strand β5', adopts the same rotamer in a buried conformation. The β3'-α3' and β4'-α4' loop regions display different conformations between the open (subunit C) and closed conformers (subunit B). (B) Close-up view of the EAL domain of RocR using a diagram representation (helices shown as cyan ribbons and β-strands as magenta arrows). Secondary structure elements of the additional lobe that protrudes over the TIM barrel are labeled. EAL active-site residues and loop 6 are represented as sticks, and the Mg²⁺ ion is represented as a green sphere. A surface representation of the EAL domain is overlaid. (C) Magnified view of the EAL active-site region with electron density displayed at a level of 3σ calculated using Fourier coefficients Fo-Fc, where the metal and one bound water were omitted from the phase calculation. (D) Close-up view of the entrance to EAL active sites in the open (yellow) and closed (cyan) conformers. The catalytically important Loop 6 displays two distinct conformations, with Glu²⁹⁶ in two rotameric conformations and Tyr³⁰¹ flipped toward opposite directions.

Previously shown that the D56N RocR mutant exhibits catalytic properties different from the wild-type RocR. Comparative SAXS experiments showed that the scattering patterns from the mutant D56N coincides with the scattering from the wild-type RocR within the experimental error (see Fig. S2B in the supplemental material), indicating that the effect of the D56N mutation on catalytic property is most likely to be exerted through local structural changes rather than global alterations to the overall quaternary structure.

REC domain and interdomain linker. The overall structure of the REC domain of RocR resembles the chemotaxis protein CheY (8). A central sheet of five parallel β-strands is flanked by helices α1', α3', and α4' on one side and α1' and α5' on the other, resulting in a (β/α)₅ topology (Fig. 4A and see Fig. S3A in the supplemental material). Canonical REC domains contain five essential residues: situated at the end of strand β1', Glu¹⁰ and Asp¹¹ are needed to coordinate a divalent metal ion that is absent here (2), whereas Asp⁵⁶ at the end of β3' is the putative residue that receives a phosphate group from the cognate histidine kinase

RocS1. Residues Ser⁸³ and Lys¹¹¹ are predicted to stabilize the phospho-aspartyl adduct and to relay conformational changes through the protein (3). In the current structure, no electron density that could account for a phosphate group is visible next to Asp⁵⁶ and RocR appears to be nonphosphorylated. Residues projecting from the face contributed by α4'-β5'-α5' are usually involved in forming specific protein-protein interactions (1, 25, 33). Residue Phe¹⁰⁵, located near its center, adopts a single rotamer conformation, with its phenyl ring buried inside the hydrophobic core. Two structural differences (overall RMSD of 2 Å) are visible between the REC domains of the open and closed conformers (Fig. 4A). Local structural remodeling is seen in the β4'-α4' loop, which undergoes conformational changes upon phosphorylation in other REC domain-containing proteins (3, 15). Moreover, the β3'-α3' loop adjacent to Asp⁵⁶ shows significant differences between the two conformers with a RMSD of 4.38 Å. Previous NMR studies on the prototypical CheY protein revealed that even unphosphorylated REC domains can sample an ensemble of confor-

s ranging from inactivated to fully activated (5, 15). Given increased accessibility of the phosphorylation site and the properties of the $\alpha 4'$ - $\beta 5'$ - $\alpha 5'$ face, the conformation adopted by EAL_A and REC_B (closed subunits) might resemble the phosphorylated state.

The interdomain linker ($\text{Q}^{130}\text{DLPRQIEVAELP}^{142}$) is visible in REC_B , where it adopts an extended conformation stabilized by interactions with Ser⁷⁴, Gly⁷⁵, and His⁷⁸ of the REC_B domain (see Fig. S1 in the supplemental material). The linker is only partially visible in subunits A, C, and D. The disordered residues from the linker are residues 135 to 138 in subunit A, residues 131 to 140 in subunit C, and residues 130 to 135 in subunit D. The intrinsic flexibility for the linker region is in agreement with the elevated B-factor levels observed by using the method of hydrogen-deuterium (H/D) exchange-coupled mass spectrometry (31).

Phosphodiesterase EAL domain. The EAL domains of RocR adopt a $(\beta/\alpha)_8$ barrel-like fold, similar to other reported EAL domain structures (Fig. 4B and see Fig. S3B in the supplemental material) (2, 26–28, 46). The structure is a variant of the TIM barrel with the $\beta 1$ strand running antiparallel to the other strands and also by the presence of additional α -helices $\alpha 1$, $\alpha 2$, and $\alpha 3$, that, together with loop $\beta 2$ - $\alpha 2$, form an additional lobe that protrudes above the barrel, near the putative c-di-GMP binding site. The catalytic residues are located at the C-terminal of the barrel, including residues that form the metal ion-binding site and the evolutionarily conserved residues of loop 6: AGYSS³⁰³ (see Fig. S3B in the supplemental material). A Mg^{2+} ion is present in each EAL active site of RocR (Fig. 4C) together with Asn²³³, Glu²⁶⁵, and Asp²⁹⁵ and a water molecule Glu¹⁷⁵ from the signature EVL motif completes the tetrahedral metal ion coordination shell. As predicted by our model (32), a water molecule is coordinated by the Mg^{2+} ion at hydrogen bonding distance from Glu³⁵² and Glu¹⁷⁵. In the crystal structures of the enzymatically active BrlP1 and TBD1265, the EAL domains were seen to bind one or two metal ions. In the high concentration of Mn^{2+} was used for cocrystallization of RocR, extensive cocrystallization and soaking experiments with Mg^{2+} , Mn^{2+} , Ca^{2+} , or Zn^{2+} in the presence or absence of c-di-GMP did not produce a RocR structure with two metal ions bound.

c-di-GMP is expected to bind to RocR in a manner similar to that of TBD1265, considering that the key residues that bind c-di-GMP are conserved among these three proteins. When we soaked c-di-GMP in the EAL active sites of RocR using the crystal structure of BrlP1 as a guide, EAL_A and EAL_B from the closed subunits highly resemble the EAL domain of BrlP1 and appear in the correct conformation for binding c-di-GMP. However, despite repeated attempts, cocrystallization and soaking experiments with c-di-GMP did not yield an enzyme-substrate complex. The structural element known to affect the catalytic properties of EAL domains is loop 6 (2, 31). Interestingly, loop 6 in the four EAL domains adopts two different conformations, as evidenced by the positioning of residue Asp²⁹⁶ and a large rotation of the phenyl ring of Tyr³⁰¹ (Fig. 4D). The observations confirm our previous proposal that loop 6 can readily undergo conformational changes that structural changes in the REC domain could be coupled to informational changes in loop 6. Another conserved function of loop 6 in EAL domains is to contribute to the dimerization of the subunits along with helix α_{10} . Despite the different conformation adopted by loop 6 in EAL_A and EAL_C (or EAL_B and EAL_D), the

loops and the two α_{10} helices form a nonsymmetric dimer interface that is stabilized by numerous polar interactions (Fig. 1C and see Fig. S1 in the supplemental material).

Inhibition of c-di-GMP binding in $\text{EAL}_{C/D}$ domains in *trans* by REC domain. One of the most striking features of the RocR crystal structure is that the quaternary structure adopted by the tetramer places the REC domains of subunits A and B against the top of the β/α barrel of the EAL domains from the subunits C and D, thus occluding their active site from the solvent (Fig. 1C and 5). The interaction between the EAL and REC domains is mainly mediated through the interaction of helices $\alpha 4'$ and $\alpha 5'$ from the REC domain. A motif consisting of three nonpolar residues (Pro⁸⁸, Ile⁸⁹, and Leu⁹⁰) located at the end of $\alpha 4'$ helix of the REC_A or REC_B domain is packed into a concave hydrophobic region in the EAL_C or EAL_D domain (Fig. 5C). The neighboring Gln⁹² residue is also seen to form hydrogen bonds with two main chain groups (Pro⁸⁷ and Ile⁸⁹) from the $\text{EAL}_{C/D}$ domain. In addition, the residue Tyr¹⁹⁵ from the lobe region is tightly sandwiched by several residues that project from helix $\alpha 5'$ of the REC domain (Fig. 5D). Together, these interactions lock REC_A and REC_B at the center of the tetrameric structure. As a result, superimposition of subunits C and D with c-di-GMP bound EAL domains of BrlP1 and TBD1265 reveals insufficient space to accommodate c-di-GMP in either EAL_C or EAL_D . It should be noted, however, that this mode of inhibition is only possible for two of the four subunits because the space at the center of the tetrameric RocR can only accommodate up to two REC domains at the same time (Fig. 1C). Importantly, when the REC_A and REC_B domains are located at the central core of the tetramer, the Asp⁵⁶ phosphorylation sites are shielded from the solvent and thus inaccessible to the cognate histidine kinase RocS1. Hence, the only Asp⁵⁶ residues that can be phosphorylated are located in the REC_C and REC_D domains.

DISCUSSION

Given the large number of two-component signaling systems, response regulators comprise a major family of signaling proteins in prokaryotes with the REC domain fused to a wide variety of DNA-binding or enzymatic domains. RocR represents a subfamily of response regulators that contain a c-di-GMP specific phosphodiesterase domain, with RocR homologues readily identified in various bacterial species (see Fig. S4 in the supplemental material). The crystal structure and SAXS results reported here revealed that RocR adopts a highly unusual quaternary structure in solution with its four subunits adopting two distinct conformations. Surprisingly, two of the four substrate-binding pockets of RocR (C and D) are not accessible to c-di-GMP and thus likely to be constitutively enzymatically inhibited. Inhibition of the enzymatic activity of the subunits C and D is achieved in *trans* by using the REC domains from the other two subunits to physically block the access of the c-di-GMP substrate. The EAL domains from the subunits A and B are likely to contain the functional active sites with open c-di-GMP binding pockets. A catalytic mechanism utilizing a single Mg^{2+} or Mn^{2+} ion was proposed for RocR previously, while a slightly different mechanism based on two metal ions was proposed for the EAL domain proteins BrlP1 and TBD1265 (2, 46). Both mechanisms share a same set of residues critical for catalysis, with the exception of Asp²⁹⁶, which is proposed to bind a second metal ion in the two-metal-ion mechanism but with a minor role in the one-metal-ion mechanism. Residue Glu³⁵² was proposed to be essential in both mechanisms, albeit with different catalytic roles. Despite our repeated

ien et al.

Fig. 5 Inhibition of *c*-di-GMP binding in *trans* by REC domains. (A) In the observed crystal structure, binding of *c*-di-GMP to the open RocR conformer (subunit C, magenta) is prevented due to physical occlusion by the REC domain of a closed conformer (subunit B, blue). (B) The *in trans* REC-EAL interaction mediated mainly through the interactions between hydrophobic residues preceding REC helix $\alpha 4'$ and an EAL hydrophobic pocket (C) and interactions between Y195 of the EAL lobe and residues projecting from helix $\alpha 5'$ (D).

tempts to cocrystallize the protein with *c*-di-GMP and high concentration of Mg^{2+} , Mn^{2+} , or Ca^{2+} ion, all of the crystal structures we obtained only contain a single metal ion per protein subunit. Because the second metal ion may only bind in the presence of *c*-di-GMP, the current structure does not allow us to discriminate between the one- or two-metal mechanisms for RocR.

Sequence comparisons of the REC domain with other canonical REC domains suggest that the key residues for phosphorylation are highly conserved. Hence, the REC domain is likely to be phosphorylated at position Asp⁵⁶ by the cognate histidine kinase RocS1, and the catalytic activity of the EAL domain could be modulated by the phosphorylation-induced structural changes. Our efforts to prepare phosphorylated forms of RocR using several small phosphate donors were unfortunately futile, as indicated by the negligible effect on enzymatic activity. The inability of the small phosphate donors to phosphorylate RocR may be due to the strong interactions between the REC and EAL domains, similar to some of the response regulators characterized by Barbieri et al. (1). Our repeated effort to crystallize the phosphate mimic BeF_3^- complexed protein for structural studies was also unsuccessful. Here we deduce a regulatory mechanism based on the current crystal structure and the results of previous biochemical and

biophysical studies. Notably, the regulatory mechanism shares some similarity with the mechanism proposed for the blue-light receptor-regulated BlrP1 (2).

Both the crystal structure and the SAXS results for the wild type and R286W and D56N mutants suggest that two REC domains are locked at the center of the tetrameric protein by a large network of specific interactions (Fig. 1 and 5). Even though the D56N mutation significantly alters the catalytic properties of RocR (31), the D56N mutant appears to adopt the same quaternary structure as the wild type and the R286W mutant, according to the SAXS experiment. Thus, given the extensive set of interactions that maintain two REC domains at the center of the RocR tetramer, the regulatory mechanism controlling the activity of the EAL domain is likely to involve local conformational changes rather than large quaternary conformational alterations such as a swinging of the centrally located REC domains to the outside of the molecule, leading to a 222 symmetric RocR tetramer (see Fig. S2A in the supplemental material). In addition, site-directed mutagenesis studies on RocR showed that the enzymatic activity of the EAL domain is very sensitive to conformational changes in a functional loop (loop 6). As revealed by the hydrogen/deuterium (H/D) ex-

Proposed regulatory mechanism. Only the relevant domains from one half of the symmetric protein are shown for clarity. Phosphorylation of Asp⁵⁶ in the REC_D domain induces local conformational changes in the REC domain (see the text). These structural changes are transmitted to the adjacent main chain through the direct contact that exists between the terminal residues Ile⁸⁹ and Leu⁹⁰ (depicted as blue spheres) and residues Phe³¹⁰ (depicted as yellow spheres) situated immediately downstream of the EAL_{D/C}. The signal is further transmitted down to the active site of the EAL domain through loop 6 that constitutes the EAL_{D/C}-EAL_{B/A} dimer interface. Based on H/D exchange-coupled mass spectrometry results reported previously, the peptides of RocR that exhibit significant conformational changes upon phosphorylation are colored in blue (REC) and red (EAL), respectively. Green spheres are located at the REC-EAL interface and may reduce the mutant's activity by disrupting the signal propagation. See also Fig. S6 in supplemental material. (Inset) Schematic view of the signal transmission pathway across the RocR tetramer. The same color code is adopted.

In experiments with the wild type and D56N mutant, the mutation causes significant structural changes in loop 6 in addition to changes in the kinetic parameters k_{cat} and K_m . From all of the information, we propose a regulatory mechanism as depicted in Fig. 6. For clarity, only the relevant domains of one half of the symmetric protein are shown and discussed here. Phosphorylation of the accessible Asp⁵⁶ in REC_D (or REC_C) triggers significant conformational changes within the REC domain. From the H/D exchange results and structural studies on other CheY-like proteins (10, 47), the greatest structural changes occur in the region that encompasses the $\alpha 4'$ and $\alpha 5'$ helices. These structural changes are propagated to the neighboring EAL domain through the direct contact between the terminal residues Ile⁸⁹ and Leu⁹⁰ of the $\alpha 5'$ helix of REC_D (Fig. 5C) and residues Phe³¹⁰ and Pro³¹¹ that follow loop 6 (amino acids 295 to 311) of the EAL_C domain (Fig. 6). Because loop 6 (EAL_C) forms multiple interactions with the adjacent loop 6 (EAL_A) at the EAL/EAL dimer interface, the structural changes incurred by one loop are likely to be transmitted to the other. Conformational changes in loop 6 would affect the catalytic efficiency of the EAL domains during binding of the c-di-GMP substrate or of the Mg²⁺ ion through desolvation of the active site required for effective catalysis. The signal transmission pathway (Fig. 6), which is initiated at the Asp⁵⁶ position of the REC_D domain and ends at the

Mg²⁺ ion in the EAL_A domain, spans an approximate distance of 48 Å and traverses the REC_D/EAL_C and EAL_C/EAL_A interfaces.

In addition to the crystal structure, several lines of evidence support the proposed regulatory mechanism and the central role of the highly conserved loop 6 (D²⁹⁶FGAGYSS³⁰²) in the mechanism. First, alteration of the conserved residues in the loop 6 region has a profound impact on catalysis. For example, the F297A mutation resulted in a 33-fold decrease in k_{cat} and a 5-fold decrease in K_m , whereas the S302A mutation gave rise to strong substrate inhibition. Replacement of residue Glu²⁶⁸ by Gln stabilizes the conformation of loop 6 and led to a reduced k_{cat} by 447-fold and K_m by 10-fold (2). These experimental observations suggest that perturbations in the conformation of loop 6 are a very effective way to regulate the catalytic efficiency of the EAL domain. Second, our previous H/D exchange studies are consistent with the proposed structural changes. Comparison of the H/D exchange patterns between RocR and the D56N mutant (the phosphorylation site) suggested that the replacement of Asp⁵⁶ by Asn induces significant structural changes in both REC and EAL domains (31). The most significant changes occur in the regions that encompass the $\alpha 4'$ and $\alpha 5'$ helices of the REC domain, loop 6, and part of the long helix $\alpha 10$ at the dimer interface. Thus, the regions undergoing structural or conformational changes as determined by H/D exchange experiments, coincide with the proposed signal propagation pathway (Fig. 6). Third, a signal transmission pathway that also involves loop 6 and other structural motifs was proposed for the EAL domain-containing protein BlrP1 (2). A superposition of the BlrP1 structure with the RocR structure shows that the regulatory BLUF and REC domains occupy similar positions close to the bottom of the α/β barrel near the $\alpha 5/\alpha 7$ helices (see Fig. S5 in the supplemental material), suggesting a conserved mode for signal transmission between the two proteins. Nonetheless, while the regulatory domain of BlrP1 only regulates the activity of the neighboring EAL domain, RocR needs to further propagate the structural changes to the second EAL domain across the EAL-EAL interface (Fig. 6 and see Fig. S5 in the supplemental material). And lastly, the regulatory mechanism can explain why the R286W mutant exhibits altered enzymatic activity. The Trp²⁸⁶ is seen at the interface of the REC_C-EAL_D and REC_D-EAL_C interfaces in the crystal structure and is part of the proposed signal transmission pathway (Fig. 6 and see Fig. S6 in the supplemental material). Hence, the Arg²⁸⁶ residue in the wild-type protein could play an important role in propagating the conformational change originated from the D56 site. The mutation R286W not only disrupts the pathway but also exerts an effect on the active site of the EAL domain through allosteric regulation, which accounts for the altered enzymatic activity of the R286W mutant.

Many bacterial genomes encode large numbers of EAL domains that are putatively regulated by a wide range of sensory domains for perceiving environmental signals. It is not known whether the EAL domain proteins share similar regulatory mechanisms. The crystal structures of BlrP1 and RocR point to an overall conservation of the mechanism controlling the enzymatic activity of their EAL domain despite the use of different regulatory domains and variation of input signals. In essence, this exquisite regulatory mechanism involves the structural remodeling of the functional loop at the EAL domain dimer interface triggered by a signal originating from a strategically positioned regulatory domain located near the $\alpha 5/\alpha 7$ helices of the EAL domain. The structural changes in the loop and its vicinity region modulate the

en et al.

talytic efficiency of the PDE domain by influencing the binding of the metal ion and c-di-GMP. It will be of great interest to see whether this mechanism is generally preserved in other EAL domain proteins, including the prevalent proteins that contain the GGDEF-EAL didomain unit.

KNOWLEDGMENTS

We thank the scientists and staff on the BL13B1 (National Synchrotron Radiation Research Center, Taiwan), PXIII (Paul Scherrer Institut, Switzerland), and X33 (DESY, Hamburg, Germany) beamlines for their expert assistance.

The laboratory of J.L. was supported by a CRP-2008 grant, a BMRC grant (08/1/22/19/589), and an ATIP grant from the CNRS. The laboratory of Z.-X.L. is supported by an ARC Tier II grant from MOE (Singapore).

REFERENCES

- Barbieri CM, Mack TR, Robinson VL, Miller MT, Stock AM. 2010. Regulation of response regulator autophosphorylation through interdomain contacts. *J. Biol. Chem.* 285:32325–32335.
- Barends TRM, et al. 2009. Structure and mechanism of a bacterial light-regulated cyclic nucleotide phosphodiesterase. *Nature* 459:1015–1018.
- Bourret RB. 2010. Receiver domain structure and function in response regulator proteins. *Curr. Opin. Microbiol.* 13:142–149.
- Chang AL, et al. 2001. Phosphodiesterase A1, a regulator of cellulose synthesis in *Acetobacter xylinum*, is a heme-based sensor. *Biochemistry* 40:3420–3426.
- Cho HS, et al. 2000. NMR structure of activated CheY. *J. Mol. Biol.* 297:543–551.
- Collaborative. 1994. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. Sect. D* 50:760–763.
- Cotter PA, Stibitz S. 2007. c-di-GMP-mediated regulation of virulence and biofilm formation. *Curr. Opin. Microbiol.* 10:17–23.
- Cowtan K. 2006. The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallogr. Sect. D Biol. Crystallogr.* 62:1002–1011.
- De N. 2008. Phosphorylation-independent regulation of the diguanylate cyclase WspR. *PLoS Biol.* 6:e67. doi:10.1371/journal.pbio.0060067.
- Djordjevic S, Goudreau PN, Xu Q, Stock AM, West AH. 1998. Structural basis for methyltransferase CheB regulation by a phosphorylation-activated domain. *Proc. Natl. Acad. Sci.* 95:1381–1386.
- Emsley P, Cowtan K. 2004. Coot: model-building tools for molecular graphics. *Acta Crystallogr. Sect. D Biol. Crystallogr.* 60:2126–2132.
- Flores S, et al. The Database of Macromolecular Motions: new features added at the decade mark. *Nucleic Acids Res.* 34:D296–D301.
- Franke D, Svergun DI. 2009. DAMMIF, a program for rapid *ab initio* shape determination in small-angle scattering. *J. Appl. Crystallogr.* 42:342–346.
- Galperin MY, Nikolskaya AN, Koonin EV. 2001. Novel domains of the prokaryotic two-component signal transduction systems. *FEMS Microbiol. Lett.* 203:11–21.
- Gardino AK, et al. 2009. Transient non-native hydrogen bonds promote activation of a signaling protein. *Cell* 139:1109–1118.
- Gouet P, Courcelle E, Stuart DI, Métoz FM. 1999. ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics* 15:305–308.
- Hengge R. 2009. Principles of c-di-GMP signaling in bacteria. *Nat. Rev. Microbiol.* 7:263–273.
- Kabsch W. 2010. Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr. Sect. D* 66:133–144.
- Konarev PV, Volkov VV, Sokolova AV, Koch MHJ, Svergun DI. 2003. PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *J. Appl. Crystallogr.* 36:1277–1282.
- Kotaka M, et al. 2009. Expression, purification and preliminary crystallographic analysis of *Pseudomonas aeruginosa* RocR protein. *Acta Crystallogr. Sect. F* 65:1035–1038.
- Kozin MB, Svergun DI. 2001. Automated matching of high- and low-resolution structural models. *J. Appl. Crystallogr.* 34:33–41.
- Krissinel E, Henrick K. 2007. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* 372:774–797.
- Kuchma SL, Connolly JP, O'Toole GA. 2005. A three-component regulatory system regulates biofilm maturation and type III secretion in *Pseudomonas aeruginosa*. *J. Bacteriol.* 187:1441–1454.
- Kulasekara HD, et al. 2005. A novel two-component system controls the expression of *Pseudomonas aeruginosa* fimbrial cup genes. *Mol. Microbiol.* 55:368–380.
- Lee SY, et al. 2001. Crystal structure of an activated response regulator bound to its target. *Nat. Struct. Biol.* 8:52–56.
- Minasov G, et al. 2009. Crystal structures of YkuL 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.
- Navarro MVAS, Bae DNN, Wang Q, Sondermann H. 2009. Structural analysis of the GGDEF-EAL domain-containing c-di-GMP receptor FimX. *Structure* 17:1104–1116.
- Navarro MVAS, et al. 2011. Structural basis for c-di-GMP-mediated inside-out signaling controlling periplasmic proteolysis. *PLoS Biol.* 9:e1000588. doi:10.1371/journal.pbio.1000588.
- Qi Y, Rao F, Luo Z, Liang Z-X. 2009. A flavin cofactor-binding PAS domain regulates c-di-GMP synthesis in *Acetobacter xylinum*. *Biochemistry* 48:10275–10285.
- Rao F, et al. 2009. Enzymatic synthesis of c-di-GMP using a thermophilic diguanylate cyclase. *Anal. Biochem.* 389:138–142.
- Rao F, et al. 2009. The functional role of a conserved loop in EAL domain-based c-di-GMP specific phosphodiesterase. *J. Bacteriol.* 191:4722–4731.
- Rao F, Yang Y, Qi Y, Liang ZX. 2008. Catalytic mechanism of c-di-GMP specific phosphodiesterase: a study of the EAL domain-containing RocR from *Pseudomonas aeruginosa*. *J. Bacteriol.* 190:3622–3631.
- Robinson VL, Wu T, Stock AM. 2003. Structural analysis of the domain interface in DrrB, a response regulator of the OmpR/PhoB subfamily. *J. Bacteriol.* 185:4186–4194.
- Roessle MW, et al. 2007. Upgrade of the small-angle X-ray scattering beamline X33 at the European Molecular Biology Laboratory, Hamburg. *J. Appl. Crystallogr.* 40:s190–s194.
- Romling U, Gomelsky M, Galperin MY. 2005. C-di-GMP: the dawning of a novel bacterial signaling system. *Mol. Microbiol.* 57:629–639.
- Ross P. 1987. Regulation of cellulose synthesis in *Acetobacter xylinum* by cyclic diguanylate. *Nature* 325:279–281.
- Round AR, et al. 2008. Automated sample-changing robot for solution scattering experiments at the EMBL Hamburg SAXS station X33. *J. Appl. Crystallogr.* 41:913–917.
- Ryan RP, et al. 2006. Cell-cell signaling in *Xanthomonas campestris* involves an HD-GYP domain protein that functions in cyclic di-GMP turnover. *Proc. Natl. Acad. Sci. U. S. A.* 103:6712–6717.
- Ryjenkov DA, Tarutina M, Moskvina OV, Gomelsky M. 2005. Cyclic diguanylate is a ubiquitous signaling molecule in bacteria: insights into biochemistry of the GGDEF protein domain. *J. Bacteriol.* 187:1792–1798.
- Schirmer T, Jenal U. 2009. Structural and mechanistic determinants of c-di-GMP signaling. *Nat. Rev. Microbiol.* 7:724–735.
- Schmidt AJ, Ryjenkov DA, Gomelsky M. 2005. The ubiquitous protein domain EAL is a cyclic diguanylate-specific phosphodiesterase: enzymatically active and inactive EAL domains. *J. Bacteriol.* 187:4774–4781.
- Schuster M, Silversmith RE, Bourret RB. 2001. Conformational coupling in the chemotaxis response regulator CheY. *Proc. Natl. Acad. Sci. U. S. A.* 98:6003–6008.
- Sheldrick G. 2010. Experimental phasing with SHELXC/D/E: combining chain tracing with density modification. *Acta Crystallogr. Sect. D* 66:479–485.
- Svergun D. 1992. Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J. Appl. Crystallogr.* 25:495–503.
- Svergun D, Barberato C, Koch MHJ. 1995. CRY SOL: a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. *J. Appl. Crystallogr.* 28:768–773.
- Tchigvintsev A, et al. 2010. Structural insight into the mechanism of c-di-GMP hydrolysis by EAL domain phosphodiesterases. *J. Mol. Biol.* 402:524–538.
- Toro-Roman A, Mack TR, Stock AM. 2005. Structural analysis and solution studies of the activated regulatory domain of the response regulator ArcA: a symmetric dimer mediated by the $\alpha 4$ - $\beta 5$ - $\alpha 5$ face. *J. Mol. Biol.* 349:11–26.
- Volz K. 1993. Structural conservation in the CheY superfamily. *J. Biol. Chem.* 32:11741–11753.
- Vonrhein C, Roversi BEP, Bricogne G. 2007. Automated structure solution with autoSHARP. *Methods Mol. Biol.* 364:215–230.
- Wassmann P, et al. 2007. Structure of BeF3—modified response regulator PleD: implications for diguanylate cyclase activation, catalysis, and feedback inhibition. *Structure* 15:915–927.

Supplemental Material

The supplemental material contains one supplementary table and six supplementary figures.

Figure S1. Protein-protein interfaces in RocR, related to Figure 1. **A.** The conserved EAL-EAL dimer interface. Superimposition of published EAL dimeric structures on the EAL interface between subunits B and D of RocR. The two EAL dimers formed in RocR (subunit B and D or their equivalent by the dyad subunits A and C) are structurally similar to other EAL dimers known to be catalytically active. A similar dimerization interface is seen, where Loop 6, $\alpha 8$ and $\alpha 10$ (helices shown as cylinders and loops as tubes) are employed. The colour code is as follows: RocR. Cyan: EAL domain of subunit B. Yellow: EAL domain of subunit D. Red sphere: Mg^{2+} in the EAL active site. TBD1265: White. (PDB code, 2r6o, r.m.s.d. of 3.4 Å with the RocR EAL domain of subunit B); YkuI-EAL: Grey. (PDB code 2w27, r.m.s.d. 4.9 Å); BlrP1-EAL: Black. (PDB code 3gfz, r.m.s.d. 5.7 Å). **B.** The buried surface areas (BSA) as well as the number of hydrogen bonds (HB), salt bridges (SB) are indicated for each interface that contributes to tetramer formation.

Figure S2. RocR solution studies by SAXS, related to Figure 5. **A.** A hypothetical model for a fully open RocR tetramer with 222 symmetry, related to Figure 5. All the REC domains are exposed at the surface of the molecule. The four EAL active sites are indicated by the presence of a c-di-GMP molecule, represented as a dumbbell. **B.** Experimental scattering from the wild type protein (1) superimposed with the scattering by the point mutants R286W (2) and D56N (3). The logarithm of the scattering intensity is plotted as a function of momentum transfer $s = 4 \pi \sin(\theta) / \lambda$, where θ is the scattering angle and λ is the X-ray wavelength.

Figure S3. Structures of the REC, EAL domain and the inter-domain linker, related to Figure 3. **A.** The REC domain (subunit C) consists in a central β -sheet of five β -strands (yellow arrows) surrounded by five α -helices (red tubes). The five residues from the active site located at the ends of the β -strands (colored cyan) and the residue D56 (phosphorylation site, red) are labeled. E10, D10 and D56 coordinate the divalent metal ion required for phosphorylation and dephosphorylation. S83 and K111 are proposed to stabilize the phospho-aspartyl adduct and transmit signals to other parts of the protein. **B.** The EAL TIM barrel domain (subunit C) viewed from the top. Residues projecting from the EVL motif and from Loop 6 ($\beta 5$ - $\alpha 8$) are shown as sticks (colored yellow and pink). The Mg^{2+} ion is represented as a green sphere. **C.** View of the linker region (residues 130-142) in subunit B with electron density displayed at a level of 1σ calculated using Fourier coefficients $2F_o - F_c$ and phases from the refined model.

Figure S4. Sequence alignment of RocR homologs. Residues from the RocR sequence are numbered and the secondary structures elements are assigned based on the RocR structure reported here. Sequence names are abbreviated as follows: BPA, *Bordetella parapertussis* (accession number: NP_883867); BAV, *Bordetella avium* 197N (accession number: YP_786392); CVI, *Chromobacterium violaceum* ATCC 12472 (accession number: NP_902070); PSY, response regulator receiver in *Pseudomonas syringae* (accession number: EGH64345); VieA_VHA, response regulator VieA in *Vibrio harveyi* (accession number: YP_001448289); VieA_VVU, response regulator VieA in *Vibrio vulnificus* CMCP6 (accession number: NP_761989); BUR, response regulator receiver modulated diguanylate PDE in *Burkholderia* sp. Ch1-1 (accession number: 1ZP_06839714); RPI, response regulator receiver modulated diguanylate PDE in *Ralstonia pickettii* 12J (accession number: ZP_07675095). Strictly conserved residues are shaded in red and boxed. Catalytically important regions are shaded in yellow; EAL residues essential for PDE catalytic activity are marked with an asterisk. The roles of functionally important residues are indicated below the sequences. Generated by using ESPript (Gouet et al., 1999).

1

2

3 **Figure S5.** Comparison of the spatial arrangement for BlrP1 and RocR, related to Figure

4 6. The black arrows indicate the pathways for the propagation of structural changes.

5

Figure S6. Position of W286 at the REC-EAL interface, highlighted in this view rotated around the y-axis by 180° compared to **Fig. 6**. W286 (green spheres) is adjacent to residues Ile⁸⁹, Leu⁹⁰ of REC (blue spheres) and Phe³¹⁰, Pro³¹¹ of EAL (yellow spheres).

Supplementary Table 1. Protein-protein dimerization interfaces in the RocR tetramer, related to Figure 1.

Interface	Dimer	N _{res}	A (Å ²)	ΔG _s (kcal/mol)	N _{HB}	N _{SB}
A	C-D	C, 42; D, 44	1630.5	-26.0	4	6
B	A-C	A, 38; D, 39	1357.4	-16.8	14	3
	B-D	B, 37; D, 40	1354.9	-16.5	18	3
	Average		1356.1	-16.7	16	3
C	B-C	B, 26; C, 28	825.0	-10.2	7	2
	A-D	A, 28; D, 29	880.2	-10.6	4	3
	Average		852.6	-10.4	6	3
D	A-B	A, 26; B, 27	952.6	-6.6	10	5

N_{res}, number of residues buried; A, interface area; ΔG_s, solvation-energy change; N_{HB}, number of residues forming hydrogen bonds; N_{SB}, number of residues forming salt bridges. All values are calculated using the Protein interfaces, surfaces and assemblies service PISA at European Bioinformatics Institute (http://www.ebi.ac.uk/pdbe/prot_int/pistart.html).

REFERENCES

Gouet, P., Courcelle, E., Stuart, D.I., and M. J. Drenth, F. (1999). ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics* 15, 305-308.

The structure of the Diguanylate Cyclase-Phosphodiesterase tandem domain of a bacterial sensor protein suggests a model for its activation*

Chong Wai Liew¹, Ming Wei Chen¹, Chun Loong Ho¹, Surya Wahyu¹, Tobias Cornvik¹, Gunter Schneider², Zhao-Xun Liang^{1‡} & Julien Lescar^{1,3‡}

¹Division of Structural Biology and Biochemistry,
School of Biological Sciences, Nanyang Technological University, 60 Nanyang Drive,
Singapore 637551

²Department of Medical Biochemistry and Biophysics, Karolinska Institutet, Stockholm,
Sweden

³INSERM UMRS 945 "Immunité et Infection", Centre Hospitalier Universitaire Pitié-Salpêtrière
Faculté de Médecine et Université Pierre et Marie Curie
91 Bd de l'Hôpital, 75013 Paris, France

*Running title: Structure of a GGDEF-EAL tandem domain

‡ To whom correspondence should be addressed. Julien Lescar, School of Biological Sciences, Nanyang Technological University, 60 Nanyang Drive, Singapore 637551
Tel.: 65 6316 2859; Fax: 65 6791 3856; E-mail: Julien@ntu.edu.sg, zxliang@ntu.edu.sg

Abstract:

Cyclic-di-GMP (c-di-GMP) is a second messenger molecule whose levels are correlated with biofilm formation and virulence in many bacteria. To understand how the c-di-GMP abundance is controlled by external signals conveyed across the periplasmic membrane, we studied the multi-domain protein Tbd1265 from *Thiobacillus denitrificans*. Tbd1265 comprises a N-terminal periplasmic sensor domain followed by a transmembrane helix, a cytoplasmic α -helix and a diguanylate cyclase-phosphodiesterase (GGDEF-EAL) tandem domain. The latter domains regulate c-diGMP level via their opposing enzymatic activities. We report the enzymatic activity and a crystal structure for a fragment of Tbd1265 encompassing its entire GGDEF and EAL domains. The crystal structures unveil a conformation stabilized by the canonic EAL dimer with the GGDEF arranged in an inactive or "resting" conformation. The linker regions connecting the GGDEF and EAL domains are exposed to the solvent and appear flexible. The structure and biochemical data are consistent with a model where a basal level of phosphodiesterase activity is maintained, whilst diguanylate cyclase activation occurs through conformational changes propagated across the membrane via the cytoplasmic helix located N-terminal to the GGDEF domain.

INTRODUCTION

Cytoplasmic second messengers relay signals received at the cell surface across the membrane to target proteins within the cell. This allows amplification of the original input into an intracellular signal capable of eliciting biochemical and metabolic changes. Nucleotides such as cAMP, cGMP and guanosine 5'-diphosphate 3'-diphosphate (ppGpp) are well-characterized intracellular signaling molecules. More recently, bis-(3'5')-cyclic dimeric guanosine monophosphate (c-di-GMP) (1) has emerged as a second messenger involved in the regulation of several bacterial processes including motility, virulence, cell cycle progression and biofilm formation (1-3). The concentration of c-di-GMP is influenced by two types of enzymes with opposing activities: GGDEF domain proteins with diguanylate cyclase (DGC) activity and EAL domain proteins that are c-di-GMP-specific phosphodiesterases (PDEs) (4-6). Another family of PDEs is formed by proteins comprising the HD-GYP domain. Synthesis of c-di-GMP from two GTP molecules results from the cooperative action of two GGDEF domains arranged in a quaternary conformation such that their active sites (A sites) face each other and where each A-site binds one GTP molecule. A second c-di-GMP binding site (I site) functions for product inhibition of the DGC activity. Conversely, hydrolysis of the c-di-GMP phosphodiester bonds is catalyzed by EAL domains leading to 5'-pGpG or to two GMP molecules. Active EAL domains characterized structurally so far form dimers through a conserved interface that involves specific structural elements including two alpha helices and a regulatory loop (loop 6) ((26), (36) (37) (Chen et al., 2012).

Many proteins containing a GGDEF or an EAL domain have been identified in pathogenic and non-pathogenic bacteria, where they regulate c-di-GMP levels and biofilm formation (4-10). GGDEF and EAL domains found across bacterial genomes frequently belong to multi-domain proteins, where these two enzymatic modules are fused in the order GGDEF-EAL, via a short linker region. We will refer to these GGDEF-EAL dual domains as "tandem domains". Several proteins containing a tandem domain have either DGC or PDE or no enzymatic activity at all, because one or the other enzymatic module has degenerate -and hence inactive- catalytic motifs (5,7,11-14). For instance the FimX protein from *Pseudomonas aeruginosa* has inactive EAL and GGDEF domains and merely functions as a c-di-GMP cellular sensor. However, an increasing number of proteins with tandem domains having dual enzymatic activities and that can conditionally switch between DGC and PDE activity have been identified: the diguanylate cyclase 1 (*dgc-1*) from *Gluconacetobacter xylinus* (11,19), the GGDEF-EAL

proteins BphG1 from *Rhodobacter sphaeroides* (15), ScrC from *Vibrio parahaemolyticus* (17) MSDGC1 from *Mycobacterium smegmatis* (18), Rv1354c from *Mycobacterium tuberculosis* (18) and CC3396 from *Caulobacter crescentus* (7). Given the genomic abundance of tandem domain proteins and their involvement in signaling pathways of opportunistic bacterial pathogens, a better understanding on how their enzymatic activities are regulated, leading to a certain level of cellular c-di-GMP, is crucial. Most GGDEF and EAL protein containing domains, as well as tandem domain proteins, do not function as stand-alone proteins but rather in association with various regulatory and sensory domains located at their N-termini. These additional sensor domains modulate their activities in response to external stimuli (8,20,21) such as light or small ligands including O₂, NO, CO or quorum-sensing molecules (reviewed in Romling et al 2013). To gain further insight into how the enzymatic activity of tandem domain proteins, where both modules are potentially active, is regulated, we chose to study the Tbd1265 protein from *Thiobacillus denitrificans*. The Tbd1265 protein contains two functional GGDEF and EAL cytoplasmic domains arranged in tandem and an extracellular periplasmic binding domain of unknown specificity (Fig. 1). Here, we present results of biochemical analyses on the DGC and PDE activity of the GGDEF-EAL tandem domain of Tbd1265. We compare the enzymatic activities of two fragments of Tbd1265 that both include its GGDEF-EAL tandem domain but solely differs by the presence (Tbd1265_{DUAL}) or absence (Tbd1265_{ADUAL}) of a α -helix located immediately N-terminal to the GGDEF module (Fig. 1). The comparison reveals that both fragments have comparable PDE activities whilst only Tbd1265_{DUAL} has significant DGC activity, suggesting a crucial role for the α -helix for DGC catalytic activity. We determined a crystal structure for Tbd1265_{DUAL}. From these data, a picture emerges that suggests a pathway for activation of the GGDEF domain.

EXPERIMENTAL PROCEDURES

Cloning, Expression and Purification of Tbd1265_{DUAL} and Tbd1265_{ADUAL} Two fragments, encompassing residues 289 – 758 of Tbd1265 (hereafter named Tbd1265_{DUAL}) and 317 – 758 of Tbd1265 (hereafter named Tbd1265_{ADUAL}) from *Thiobacillus denitrificans*, were cloned into pET-26b (+) each with a C-terminal hexa-histidine tag. The positive recombinants were transformed into *E. coli* BL21(DE3) cells and used for overexpression. For protein expression, cells were grown in a volume of 1 L of LB

medium supplemented with 50 µg/mL kanamycin until an OD₆₀₀ of 0.5 was reached. Expression of the target proteins was induced by the addition of 0.5 mM isopropyl-β-D-thiogalactopyranoside for 17 – 20 hours. Cells were harvested by centrifugation and resuspended in lysis buffer (50 mM Na HEPES at pH 7.5, 500 mM NaCl, 5 mM imidazole, 10% (v/v) glycerol and 10 mM β-mercaptoethanol), supplemented with the Protease Inhibitor Mixture Set III, EDTA free (Calbiochem, Novagen), and 2000 units of benzonase (Merck). Cell disruption was performed by sonication on ice (using a Sonics Vibra cell at 70% amplitude, 3 s on/off for 3 min). The lysate was clarified by centrifugation at 47,000 g for 25 min at 4°C, and the supernatant was filtered through a 0.45 µm syringe filter. Filtered lysates were loaded onto 1 ml of Ni-NTA Superflow (Qiagen) in wash buffer 1 (50 mM Na HEPES pH 7.5, 500 mM NaCl, 10 mM imidazole, 10% (v/v) glycerol, 10 mM β-mercaptoethanol) and subsequently washed with (50 mM Na HEPES pH 7.5, 500 mM NaCl, 25 mM imidazole, 10% (v/v) glycerol, 10 mM β-mercaptoethanol). Bound proteins were eluted with the same buffer containing 500 mM imidazole and loaded onto a HiLoad 16/60 Superdex-200 column (GE Healthcare) pre-equilibrated with (25 mM Na HEPES pH 7.5, 500 mM NaCl, 10% (v/v) glycerol, 2mM DTT). Only the purest fractions containing Tbd1265_{DUAL} (or Tbd1265_{ADUAL}) were pooled and concentrated to a final concentration of ~15 mg/ml. The protein was aliquoted, frozen in liquid nitrogen, and stored at -80 °C until use.

Analytical size-exclusion chromatography The purified Tbd1265_{DUAL} and Tbd1265_{ADUAL} proteins (at concentrations of about 20 µM) were subjected to analytical size-exclusion chromatography using a Superdex 200 5/150 GL column (GE Healthcare). The buffer used for purification in the preparative size exclusion chromatography was used as running buffer for both native and substrate-enzyme complex. The substrate-enzyme complex mixture contained 20 µM protein and 5 mM c-di-GMP, supplemented with 10 mM MgCl₂. Runs were performed on an AKTA purifier (GE Healthcare) system at a flow rate of 0.3 mL/min.

Analytical ultracentrifugation sedimentation equilibrium Analytical ultracentrifugation sedimentation equilibrium experiments were performed on a Beckman XL-I instrument (Beckman Coulter) at 18°C. Protein samples were dialyzed overnight against the reference buffer (25 mM Na HEPES pH 7.0, 500 mM NaCl, 5% glycerol, 2 mM DTT); 20 µM c-di-GMP and 10 mM CaCl₂ were included where appropriate. Buffer density and

viscosity, and appropriate run speeds based on partial specific volume of the protein were calculated using the SEDNTERP software. Values obtained were 0.7329 ml/g for Tbd1265_{ADUAL} and 0.7318 ml/g for Tbd1265_{DUAL}, whereas the buffer density and viscosity without c-di-GMP and CaCl₂ $\rho = 1.03537$ g/ml, $\eta = 1.2937$ cP; with CaCl₂ $\rho = 1.03729$, $\eta = 1.2975$ cP (c-di-GMP was omitted from calculation). Protein samples were loaded into six-channel charcoal-filled Epon centerpieces at OD₂₈₀ of 0.3, 0.5, and 0.8 (at 1.2 cm pathlength) and centrifuged at 10,000, 12,200, and 15,000 rpm. HeteroAnalysis software was used to determine whether equilibrium had been reached. Equilibrium data sets were trimmed and optical artifacts (including spikes) were removed in SEDFIT and subsequently fitted to monomer-*N*-mer equilibrium models in SEDPHAT (22).

Enzymatic assays (PDE and DGC) Enzymatic assays for DGC and PDE activities were performed with GTP and c-di-GMP, respectively. The dye-labeled 484-MDCC biosensor was prepared according to Ho and co-workers (23). All reactions were performed in a reaction buffer consisting of 100 mM Tris-HCl at pH 8.0, 50 mM KCl, 20 mM MgCl₂. The 20 μ M of 484-MDCC biosensor protein was incubated with c-di-GMP at 2 μ M and 10 μ M at room temperature for 15 minutes. In a separate solution, 20 μ M of Tbd1265_{DUAL} (or Tbd1265_{ADUAL}) protein was prepared. The pre-incubated sensor-c-di-GMP solution was aliquoted with a total of 50 μ L into a Corning black 96-wells titer plate with clear flat bottom. To start the reaction, Tbd1265_{DUAL} (or Tbd1265_{ADUAL}) was added at equal volume to reach a final c-di-GMP concentration of 1 μ M and 5 μ M respectively. For the DGC activity assay, Tbd1265_{DUAL} (or Tbd1265_{ADUAL}) at 20 μ M, MgCl₂ at 20 mM, and 484-MDCC biosensor protein at 20 μ M were equilibrated in the reaction buffer as described above. The substrate GTP was added to a final concentration of 1 μ M and 10 μ M to start the reactions. Both DGC and PDE activity reaction were monitored using a Tecan Infinite 200 Pro multi-plate reader equipped with a fluorescence detector. Using the I-control program, the wells were excited at 419 nm and the emission wavelength was measured at emission maxima of 466 nm at a 20 seconds interval. Results were compared to the linear titration curve established, as in (25).

Enzymatic synthesis/hydrolysis of c-di-GMP – C-di-GMP was produced enzymatically by Tbd1265_{GGDEF} and subsequently hydrolyzed by Tbd1265_{EAL} domain. The progress of c-di-GMP synthesis/hydrolysis was monitored using an Agilent LC1200 system (mobile phase, 20 mM triethylammonium bicarbonate, pH 7.0, 10% methanol, 0.6 ml/ min) with

an Eclipse XDB C18 column (4.6 by 150 mm). All enzymatic reactions were performed in a final volume of 25 μ L and monitored at different time points of incubation. A reaction mixture includes 50 μ M Tbd1265_{DUAL} (or Tbd1265_{ADUAL}), 100 μ M GTP (or c-di-GMP), 10 mM MgCl₂ and reaction buffer (100 mM Tris, pH7.5 and 150 mM NaCl). The reaction mixtures were incubated at 25 °C for 1 and 2 hours. The reaction was stopped by boiling the reaction mixture for 5 min. Following centrifugation at 14,000 rpm for 4 minutes to remove the protein precipitate, the supernatant was filtered and loaded onto the analytical Eclipse XDB C18 column (4.6 by 150 mm) using an Agilent LC1200 system. Either c-di-GMP or GMP (or pGpG) was eluted using the same mobile phase described above. The standard, c-di-GMP and GTP, were diluted to 100 μ M with the reaction buffer and incubated at 25 °C for 1 and 2 hours prior to application onto a XDB C18 column. The substitution of MgCl₂ with CaCl₂ in the reaction buffer, which the reaction mixture containing 50 μ M Tbd1265_{DUAL} (or Tbd1265_{ADUAL}), 100 μ M GTP and 10 mM CaCl₂, was employed as a negative control for PDE activity.

Crystallization and Data Collection Native crystals of Tbd1265_{DUAL} were obtained after screening 672 crystallization conditions (Hampton Research), using the CyBio-Crystal Creator robot (Jena Biosciences). Native Tbd1265_{DUAL} crystals were grown in sitting drops comprising equal volumes of protein (10 mg/ml) and precipitant solutions (0.1 M Na HEPES, 2.8 M NaCl, pH 7.5). Before data collection, crystals were transferred to a cryo-protecting solution containing the precipitating solution supplemented with 20% (v/v) glycerol and cooled to 100 K in liquid N₂. Diffraction intensities were collected at the Swiss Light Source Beamline PXIII, using a Dectris Pilatus 6M detector, to a resolution of 3.4 Å for the free enzyme. Integration, scaling, and merging of the intensities were carried out using programs MOSFLM and SCALA from the CCP4 suite (24). Data collection statistics are summarized in **Table 1**.

Structure Determination and Refinement – The structure was determined using molecular replacement as implemented in the program Phaser from the CCP4 suite (25). The isolated EAL domain from Tdb1265 from *Thiobacillus denitrificans* whose structure was reported previously (PDB code: 2R6O) (26) and the GGDEF domain from McR174C (with a sequence identity of 41% for 162 aligned amino-acid residues; PDB code: 3ICL) from *Methylococcus capsulatus* were used as search probes. Model building was initiated with the program Chainsaw (27) from CCP4 suite, followed by numerous

rounds of interactive computer graphics sessions with program COOT (28).

Crystallographic refinement cycles were carried-out using REFMAC5 (24) with non crystallographic symmetry restraints between the five molecules of Tbd1265_{DUAL} present in the asymmetric unit and with the program BUSTER (29). Refinement statistics are shown in Table 1. The quality of the final model was assessed with MolProbity (30). Figures were generated using PyMOL (31).

RESULTS

Overall structural organization of Tbd1265

The overall domain organization of Tbd1265 is illustrated in **Fig. 1**. Tbd1265 is composed of a N-terminal periplasmic binding domain (PBP) followed by a transmembrane domain that comprises a glycine-rich region and GGDEF-EAL-containing cytoplasmic module. Basically, Tbd1265 protein is a motility-to-sessility transition regulatory protein that is dependent on cyclic-di-GMP-signaling. The N-terminal periplasmic domain of Tbd1265, Tbd1265_{INPUT} domain, is a small molecule sensory domain that plays a crucial role on transducing the signal to the cytoplasmic regulatory GGDEF-EAL tandem protein. Here, we have cloned, overexpressed and purified two fragments of Tbd1265 that contain its GGDEF and EAL domains responsible for the diguanylate cyclase and phosphodiesterase activities respectively: Tbd1265_{DUAL} spanning residues 289 – 758 and Tbd1265_{ADUAL} that comprises residues 317 – 758 (**Fig. 1**). These fragments differ by the stretch of 28 residues at the N-terminal of Tbd1265_{DUAL} (residues 289-317) that are predicted to fold into α -helical coiled coil (hereafter named “*n*helix”). The presence of the N-terminal helical region plays a crucial role in the DGC catalytic activity (see below). The GGDEF domain is fused to the EAL domain via a linker region of ~15 residues (**Fig. 1**).

Multimerization properties of Tbd1265_{DUAL} and Tbd1265_{ADUAL}

Purified Tbd1265_{DUAL} (51.8 kDa) and Tbd1265_{ADUAL} (48.2 kDa) protein were obtained at a concentration of 10 mg/mL and 13 mg/mL, respectively. Size exclusion chromatography showed that both proteins eluted as single symmetrical peaks (**Fig. 2**). Comparison with protein standards reveals that both proteins are dimers in solution. To test possible influence of ligand binding on the protein multimerization state or conformation, the measurements were repeated in presence of c-di-GMP supplemented with Mg²⁺. The elution profiles of both proteins were slightly

shifted towards lower molecular weight (**Fig. 2**) indicating some changes in protein conformation upon ligand binding possibly towards a more compact structure. Since gel filtration chromatography is sensitive to protein shape, we also subjected Tbd1265_{DUAL} and Tbd1265_{ΔDUAL} to sedimentation equilibrium analytical ultracentrifugation with or without prior incubation with c-di-GMP (**Fig. 3**). Radial distribution data of Tbd1265_{DUAL} and Tbd1265_{ΔDUAL}, in the presence and absence of c-di-GMP, were fitted against several monomer-*N*-mer equilibrium models. Based on the global reduced *chi*-square values and randomness of the fitting residual, the best fit was found for a dimer in all cases. In the absence of c-di-GMP, the equilibrium association constants *K_a* were $3.234 \times 10^6 \text{ M}^{-1}$ and $3.564 \times 10^6 \text{ M}^{-1}$ for Tbd1265_{ΔDUAL} (**Fig. 3a**) and Tbd1265_{DUAL} (**Fig. 3b**), respectively, while in the presence of c-di-GMP, values of $4.736 \times 10^6 \text{ M}^{-1}$ and $1.042 \times 10^7 \text{ M}^{-1}$ were obtained for Tbd1265_{ΔDUAL} (**Fig. 3c**) and Tbd1265_{DUAL} (**Fig. 3d**) respectively. The results indicate that both Tbd1265_{DUAL} and Tbd1265_{ΔDUAL} appeared preferably as a dimer in solution. For comparison purposes, overlaid of monomer-dimer fits in the absence of c-di-GMP on top of the profile in the presence of enzyme shown that the substrate-enzyme complex sediment more for both Tbd1265_{ΔDUAL} (**Fig. 3e**) and Tbd1265_{DUAL} (**Fig. 3f**). This indicated that the substrate-enzyme complex is more compact compared to the unbound enzyme. The observation is consistent with the result obtained from the analytical size exclusion chromatography. In contrast, LapD_{DUAL}⁽³²⁾ and FimX_{DUAL}⁽³³⁾ are monomers in their native state and dimerize in a concentration-dependent manner at high protein concentration.

The *n*helix affects the DGC and PDE enzymatic activity.

To determine the synthesis and hydrolysis rates of c-di-GMP by Tbd1265_{DUAL} and Tbd1265_{ΔDUAL} in real time, we used a dye-labeled sensor technique. This technique can detect nanomolar changes in c-di-GMP concentration and is able to give insight into small differences in protein activity. The PDE assay reactions using two different substrate concentrations revealed that the EAL domain of Tbd1265_{DUAL} is slightly more active compared to the Tbd1265_{ΔDUAL} protein. In the presence of the N-helix at the N-terminal of Tbd1265_{DUAL}, an approximate 2-fold increment of enzymatic activity compared to Tbd1265_{ΔDUAL}, which is devoid of the N-helix is observed (**Fig. 4a**). The *K_m* and *V_{max}* values for the PDE activities of the EAL domains of Tbd1265_{DUAL} are 27.0 nM and 22 nM/sec and 35.1 nM and 7.2 nM/sec for Tbd1265_{ΔDUAL}. Likewise, we observed an approximate 10-fold increase in the DGC enzymatic activity for Tbd1265_{DUAL}

compared to the Tbd1265_{ΔDUAL} protein (**Fig. 4b**). Values for K_m and V_{max} of Tbd1265_{DUAL} are 126.2 nM and 64.3 nM /sec, whilst Tbd1265_{ΔDUAL} has a K_m and V_{max} value of 1.3 μM and 18.7 nM / sec. Thus, the N-helix plays an important role in regulating the enzymatic activity of the GGDEF domain and to a lesser extent of the EAL domain of the Tbd1265_{DUAL} tandem protein.

Enzymatic synthesis/hydrolysis of c-di-GMP by Tbd1265_{DUAL} and Tbd1265_{ΔDUAL}.

The DGC and PDE activity of Tbd1265_{DUAL} and Tbd1265_{ΔDUAL} were assayed using HPLC with GTP and c-di-GMP supplemented with Mg^{2+} . A total of three distinct peaks were observed corresponding to GTP (peak 1, retention time = 1.86 minutes), pGpG (peak 2, retention time = 3.69 minutes) and c-di-GMP (peak 3, retention time = 8.64 minutes). Tbd1265_{DUAL} demonstrated both DGC and PDE activities: In the presence of GTP and Mg^{2+} , its GGDEF domain was able to synthesize c-di-GMP, which was subsequently partly hydrolyzed to pGpG by the EAL domain (**Fig. X**). Moreover, incubation of Tbd1265_{DUAL} with c-di-GMP resulted in the formation of pGpG and a minor amount of GMP (**Fig. X**). The PDE activity of EAL domains is inhibited in the presence of Ca^{2+} . To verify the DGC activity of Tbd1265_{DUAL}, Ca^{2+} ions were thus added to the reaction mixture, (**Fig. X**). Only c-di-GMP was formed with no hydrolysis of c-di-GMP. Likewise, the truncated Tbd1265_{ΔDUAL} protein was endowed with PDE activity leading to the hydrolysis of c-di-GMP to form pGpG (**Fig. X**). The PDE activity was abolished in the presence of Ca^{2+} ions with no pGpG formed. However, the DGC domain of Tbd1265_{ΔDUAL} appears largely inactive: no formation of c-di-GMP was observed even after two hours incubation.

Structure determination

The crystal structure of Tbd1265_{DUAL} was determined at a resolution of 3.4 Å. Tbd1265_{DUAL} crystallized in the space group C2 with two dimers arranged along a non-crystallographic dyad (**Fig. 5c**). Each molecule can be superposed with an average r.m.s. deviation of 0.03 Å. The final model comprises residues 317 – 475 and 493 – 735 (**Fig. 5a**). Some electron density corresponding to the N-terminal helix was visible for Tbd1265_{DUAL} but its quality was not sufficient to trace residues 288 – 316 with confidence. Likewise, the linker connecting the GGDEF and EAL domains (residues 476 – 492) was omitted from the model of Tbd1265_{DUAL}. Moreover, the GGDEF-EAL linker region appears flexible; this mobility could facilitate conformational switching during

synthesis/hydrolysis of c-di-GMP as well as quaternary changes (see below). Refinement statistics are given in **Table 1**.

The GGDEF domain

The GGDEF domain of Tbd1265_{DUAL} shares a similar structure with several putative GGDEF domain-containing diguanylate cyclases (26,33-37) including PleD (38), WspR (39) and FimX (33). It consists of five helices that wrap around a five-stranded central antiparallel β -sheet (**Fig. 5a**), with two additional short β -hairpins. The topology of the GGDEF domain closely resembles the catalytic core of adenylate cyclase (40) and the 'palm' domain of DNA polymerases (41). The GGDEF domain is functionally related to these two enzymatic domains in that both catalyze the formation of 3'-5' phosphodiester bonds from two nucleotide molecules. An automatic search for structurally homologous proteins (42) reveals homology with GGDEF domains of *M. capsulatus* (PDB code 3ICL, *z* score of 28.6), despite a ~ 41% sequence identity for 162 aligned amino-acid residues. Structural similarity is also observed with LapD_{DUAL} GGDEF-EAL di-domain from *P. fluorescence* (32) (*z* score of 25.9 and ~ 24% sequence identity).

The GGDEF active site and inhibition site

A sequence alignment reveals that the GGDEF domain of Tbd1265_{DUAL} comprises residues known to be involved in binding GTP/Mg²⁺/c-di-GMP, as well as the GGDEF catalytic motif and the I-site allosteric inhibition site (**Fig. 6a**). A superposition of the nucleotide-free GGDEF domain of Tbd1265_{DUAL} with the GGDEF domain of PleD bound to GTP (or c-di-GMP) reveals the compatibility of product binding to Tbd1265_{DUAL} at the GGDEF domain active site. Unlike FimX, the GGDEF domain of Tbd1265_{DUAL} has similar GTP/Mg²⁺ binding residues as PleD and WspR (**Fig. 6b**), except Gly470 (Arg446 in PleD and Arg331 in WspR). Residues that contribute to a binding pocket for the guanosine base, in particular Asn361 and Asp370, are strictly conserved across active diguanylate cyclases (**Fig. 6b**). Only Lys466 in the GGDEF domain of Tbd1265_{DUAL} is involved in coordinating the γ -phosphate moiety of GTP if compared to Lys442/327 and Arg446/331 in PleD/WspR. These residues shape the binding sites for the γ -phosphate and the guanine base. Likewise, residues that contribute in coordinating Mg²⁺ are found similar to PleD and WspR: Asp396 in the GGDEF motif and Asp353 (**Fig. 6b**).

By comparing to the c-di-GMP bound PleD, the GGDEF domain of Tbd1265_{DUAL} revealed the possibility of c-di-GMP binding to Tbd1265_{DUAL} at the active site. In contrast to FimX, no residue in the active site of GGDEF domain in Tbd1265_{DUAL} introduced steric clashes with the di-nucleotide. In fact, few residues, like Lys358, Asn361, Asp370 and Arg392, form H-bond with the guanosine base (**Fig. 6c**). In addition to the GGDEF motif active sites, a second c-di-GMP binding site in GGDEF domain has been identified as an inhibition site (I-site) in several DGC structures (34,35,38,43). The I-site, formed by a characteristic primary site motif with the consensus R³⁸⁵xxD in Tbd1265_{DUAL}, plays a critical role in the regulation of cyclase activity by preventing excess GTP consumption and probably defines an upper boundary for the concentration of the second messenger c-di-GMP in the bacterial cell. Communication between the I-site and the A site of a GGDEF domain could cause the observed noncompetitive product inhibition (43). A superposition of Tbd1265_{DUAL} with PleD bound to c-di-GMP at the I-site shows striking residue conservation (**Fig. 6d**) suggesting a noncompetitive product inhibition activity for Tbd1265_{DUAL}.

The EAL domain

The conformation of the EAL domain in the context of the Tbd1265_{DUAL} tandem protein (or Tbd1265_{ADUAL}) is conserved compared to the isolated EAL domain reported previously (28). It adopts a (β/α)₈ TIM barrel fold with an unusual β 8 strand that runs anti-parallel with the other seven strands, a feature also present in the EAL domains of YkuI from *B. subtilis*, Blrp1 from *K. pneumonia* and FimX from *P. aeruginosa* (33,36,37). The additional lobe is composed of four helices, α 8, α 9, α 10 and α 11 (**Fig. 5a**). Helix α 8 which runs parallel to the barrel plane makes the connection with the GGDEF domain through the GGDEF-EAL linker. Interestingly, the stand-alone EAL domain of Tbd1265 forms a dimer in solution.

The structure of Tbd1265_{DUAL} reveals a resting state for the GGDEF domain

Tbd1265_{DUAL} adopts an extended conformation. The protein forms a dimer through interactions between its EAL domains (**Figs. 5b and 5c**). The crystal structure of the Tbd1265_{DUAL} reveals that the EAL domain establishes a network of hydrogen-bonds between residues projecting from helices α 13 and α 15 and from β 12 – α 13 loop as well as by hydrophobic interactions through residues projecting from helix α 15 with equivalent residues related by the crystallographic dyad. These interactions stabilize the EAL-EAL

interface between subunit A and B (or their equivalents subunit C and D, by the dyad) resulting in a total buried interface of 2090 Å (Figs. 5b and 5c). By contrast interactions established between the GGDEF domains of Tbd1265_{DUAL} do not result in the formation of a stable dimeric structure. The interactions between two GGDEF domains are limited to the $\alpha 5 - \beta 2$ loop and $\beta 3 - \alpha 6$ loop (Figs. 5b and 5c). The relative orientation of the GGDEF domains of Tbd1265_{DUAL} reveals that the conserved residues Arg385 and Asp388 of the I-site motif are buried. Hence both domains are incompetent for binding c-di-GMP (Figs. 5c and 7i). To conform to the active mode of diguanylate cyclase, each GGDEF domain has to undergo a 180° rotation (Fig. 7ii). Compared to the allosteric product binding state of the GGDEF domain of PleD (32), each GGDEF domain has to be flipped back about 10 Å with the RxxD as centre of movement in order to conform to the inhibition mode (Fig. 7iii). Thus, based on the orientation of GGDEF domains observed in Tbd1265_{DUAL}, the GGDEF domains are neither in an active conformation with both A sites facing each other nor in an allosteric product binding mode leading to feedback inhibition, but rather in an intermediate or “resting” conformation (Fig. 7i). Thus, the EAL-EAL domains form a stable dimeric interface whilst the two GGDEF domains should be able to adopt a range of different orientations leading to activation with the A site facing each other or to a feedback inhibition mode.

DISCUSSION

A minor fraction of GGDEF and EAL domains has been integrated into bacterial two-component regulatory networks that rely on phosphotransfer between a sensor histidine kinase with an extracellular input domain and a cytosolic response regulator with a dedicated output domain. A much larger fraction belongs to one-component systems, single proteins that contain input and output but lack phosphotransfer domains. Similar to FimX (33) or LapD (32), most of the one-component systems possess both types of domains in the same polypeptide chain with variation among catalytic activity preserved among protein family members. In addition, several of these function as bacterial transmembrane receptors and contain an additional HAMP domain as a juxtamembrane signal relay module, located N-terminally to the GGDEF-EAL domain (32). Within HAMP domain and GGDEF-EAL domain, the S-helix appeared as a crucial region that involved in the autoinhibition of the diguanylate cyclase enzymatic activity.

Based on the crystal structure acquired, the dimerization interface of the protein is localized at the *n*helix and XXX residue of the EAL domain. Thus, the functionality of

the protein reverts from an active DGC protein when the two helices are joined and a relatively more active EAL protein when the two helices are a part.

To convert the GGDEF domain from an inhibited to an activated state, both the GGDEF-EAL linker and the n helix play a crucial role. Firstly, the interdomain flexibility imparted by the GGDEF-EAL linker domain could be involved in switching from a resting to an active mode of diguanylate cyclization by allowing the reorientation of the GGDEF domain. However, the precise structural basis for its role remains elusive because the GGDEF-EAL linker is not ordered in most of the GGDEF-EAL crystal structures, such as LapD (32) and FimX (33).

On the other hand, it has been speculated that the n helix is an important secondary structure element not only for the stabilization of the dimeric Tbd1265_{DUAL} but also exerts a large impact on both diguanylate cyclase and phosphodiesterase activity. Based on the secondary structure prediction and coiled-coil prediction (44), these 28 residues formed a α -helical coiled-coil secondary structure that bridged the transmembrane region to the cytoplasmic Tbd1265_{DUAL}.

Conclusion

We have shown that a helix located immediately N-terminal to the GGDEF domain is a crucial determinant of the overall activity of the diguanylate cyclase activity. It acts as a relay to convey signals from the PBP region, probably by influencing the quaternary structure of the protein.

Two hypotheses emerge from this work: In the first scenario, only GGDEF domains are involved in signal regulation, given a basal phosphodiesterase activity of the EAL domain. Alternatively, both activities can vary over time depending on the output.

Acknowledgements

*This work was supported by the Ministry of Education of Singapore through the Academic Research Council (ARC) grants (Z.-X. L.). Support by BMRC (grant 08/1/22/19/589) and by the CNRS to the laboratory of JL is also acknowledged.

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) *Nature* **325**, 279-281
2. D'Argenio, D. A., and Miller, S. I. (2004) *Microbiology* **150**, 2497-2502
3. Römling, U., Gomelsky, M., and Galperin, M. Y. (2005) *Molecular Microbiology* **57**, 629-639
4. Ryjenkov, D. A., Tarutina, M., Moskvina, O. V., and Gomelsky, M. (2005) *Journal of Bacteriology* **187**, 1792-1798
5. Schmidt, A. J., Ryjenkov, D. A., and Gomelsky, M. (2005) *Journal of Bacteriology* **187**, 4774-4781
6. Tamayo, R., Tischler, A. D., and Camilli, A. (2005) *Journal of Biological Chemistry* **280**, 33324-33330
7. Christen, M., Christen, B., Folcher, M., Schauerte, A., and Jenal, U. (2005) *Journal of Biological Chemistry* **280**, 30829-30837
8. Galperin, M. Y., Nikolskaya, A. N., and Koonin, E. V. (2001) *FEMS Microbiology Letters* **203**, 11-21
9. Hickman, J. W., Tifrea, D. F., and Harwood, C. S. (2005) *Proceedings of the National Academy of Sciences of the United States of America* **102**, 14422-14427
10. Paul, R., Weiser, S., Amiot, N. C., Chan, C., Schirmer, T., Giese, B., and Jenal, U. (2004) *Genes & Development* **18**, 715-727
11. 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) *Journal of Bacteriology* **180**, 4416-4425
12. Weber, H., Pesavento, C., Possling, A., Tischendorf, G., and Hengge, R. (2006) *Molecular Microbiology* **62**, 1014-1034
13. García, B., Latasa, C., Solano, C., Portillo, F. G.-d., Gamazo, C., and Lasa, I. (2004) *Molecular Microbiology* **54**, 264-277
14. Kuchma, S. L., Brothers, K. M., Merritt, J. H., Liberati, N. T., Ausubel, F. M., and O'Toole, G. A. (2007) *Journal of Bacteriology* **189**, 8165-8178
15. Tarutina, M., Ryjenkov, D. A., and Gomelsky, M. (2006) *Journal of Biological Chemistry* **281**, 34751-34758
16. Ferreira, R. B. R., Antunes, L. C. M., Greenberg, E. P., and McCarter, L. L. (2008) *Journal of Bacteriology* **190**, 851-860
17. Boles, B. R., and McCarter, L. L. (2002) *Journal of Bacteriology* **184**, 5946-5954
18. Kumar, M., and Chatterji, D. (2008) *Microbiology* **154**, 2942-2955
19. Bae, S. O., Sugano, Y., Ohi, K., and Shoda, M. (2004) *Appl Microbiol Biotechnol* **65**, 315-322
20. Chang, A. L., Tuckerman, J. R., Gonzalez, G., Mayer, R., Weinhouse, H., Volman, G., Amikam, D., Benziman, M., and Gilles-Gonzalez, M.-A. (2001) *Biochemistry* **40**, 3420-3426
21. Qi, Y., Rao, F., Luo, Z., and Liang, Z.-X. (2009) *Biochemistry* **48**, 10275-10285
22. Vistica, J., Dam, J., Balbo, A., Yikilmaz, E., Mariuzza, R. A., Rouault, T. A., and Schuck, P. (2004) *Analytical Biochemistry* **326**, 234-256
23. Ho, C. L., Koh, S. L., Chuah, M. L. C., Luo, Z., Tan, W. J., Low, D. K. S., and Liang, Z.-X. (2011) *ChemBioChem* **12**, 2753-2758

24. Collaborative. (1994) *Acta Crystallographica Section D* **50**, 760-763
25. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) *Journal of Applied Crystallography* **40**, 658-674
26. 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) *Journal of Molecular Biology* **402**, 524-538
27. Stein, N. (2008) *Journal of Applied Crystallography* **41**, 641-643
28. Emsley, P., and Cowtan, K. (2004) *Acta Crystallographica Section D* **60**, 2126-2132
29. Bricogne, G. (1993) *Acta Crystallographica Section D* **49**, 37-60
30. Chen, V. B., Arendall, W. B., III, Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., and Richardson, D. C. (2010) *Acta Crystallographica Section D* **66**, 12-21
31. Delano, W. L. (2002)
32. Navarro, M. V. A. S., Newell, P. D., Krasteva, P. V., Chatterjee, D., Madden, D. R., O'Toole, G. A., and Sondermann, H. (2011) *PLoS Biol* **9**, e1000588
33. Navarro, M. V. A. S., De, N., Bae, N., Wang, Q., and Sondermann, H. (2009) *Structure (London, England : 1993)* **17**, 1104-1116
34. Chan, C., Paul, R., Samoray, D., Amiot, N. C., Giese, B., Jenal, U., and Schirmer, T. (2004) *Proceedings of the National Academy of Sciences of the United States of America* **101**, 17084-17089
35. De, N., Pirruccello, M., Krasteva, P. V., Bae, N., Raghavan, R. V., and Sondermann, H. (2008) *PLoS Biol* **6**, e67
36. Barends, T. R. M., Hartmann, E., Griesse, J. J., Beitlich, T., Kirienko, N. V., Ryjenkov, D. A., Reinstein, J., Shoeman, R. L., Gomelsky, M., and Schlichting, I. (2009) *Nature* **459**, 1015-1018
37. Minasov, G., Padavattan, S., Shuvalova, L., Brunzelle, J. S., Miller, D. J., Basler, A., Massa, C., Collart, F. R., Schirmer, T., and Anderson, W. F. (2009) *Journal of Biological Chemistry* **284**, 13174-13184
38. Wassmann, P., Chan, C., Paul, R., Beck, A., Heerklotz, H., Jenal, U., and Schirmer, T. (2007) *Structure* **15**, 915-927
39. De, N., Navarro, M. V. A. S., Raghavan, R. V., and Sondermann, H. (2009) *Journal of Molecular Biology* **393**, 619-633
40. Tesmer, J. J. G., Sunahara, R. K., Johnson, R. A., Gosselin, G., Gilman, A. G., and Sprang, S. R. (1999) *Science* **285**, 756-760
41. Doubie, S., Tabor, S., Long, A. M., Richardson, C. C., and Ellenberger, T. (1998) *Nature* **391**, 251-258
42. Holm, L. et al (2010) *Nucleic Acids Research* **38**, W545-W549
43. Christen, B., Christen, M., Paul, R., Schmid, F., Folcher, M., Jenoe, P., Meuwly, M., and Jenal, U. (2006) *Journal of Biological Chemistry* **281**, 32015-32024
44. Lupas, A., Van Dyke, M., and Stock, J. (1991) *Science* **252**, 1162-1164

Romling U, Galperin MY, Gomelsky M (2013) *Microb and Molec Biol Rev* **77**, 1-52

Figure Legends

Figure 1: Domain organization of Tbd1265 from *Thiobacillus denitrificans*. The N-terminal periplasmic binding protein domain, PBP (residues 28 – 247) is connected via transmembrane region, TM (residues 254 – 280) to the Tbd1265_{DUAL} (residues 289 – 758) with bifunctional diguanylate cyclase and phosphodiesterase activity. The predicted helix fragment (residues 289 – 317) located at the N-terminal end of GGDEF domain is indicated. The sequence of the GGDEF and EAL signature motif important for catalysis are labeled, as well as the flexible GGDEF-EAL linker region. The truncated Tbd1265_{ADUAL} (residues 317 – 758) is shown.

Figure 2: Analytical size-exclusion chromatography (SEC) profile. (a) Standard proteins (*inset*), Tbd1265_{DUAL} and Tbd1265_{ADUAL} were analyzed by size exclusion chromatography column. The elution profile in the presence of c-di-GMP in (b) Tbd1265_{DUAL} and (c) Tbd1265_{ADUAL} are indicated. *Inset*: Standard proteins (aldolase, M_r 158 KDa and conalbumin, M_r 75 kDa), Tbd1265_{DUAL} (M_r 104 kDa) and truncated Tbd1265_{ADUAL} (M_r 96.4).

Figure 3: Sedimentation equilibrium analysis of homo-dimerization of Tbd1265_{DUAL} and Tbd1265_{ADUAL} in the presence and absence of c-di-GMP. Radial distribution profile (open circles) of Tbd1265_{ADUAL} (a and c) and Tbd1265_{DUAL} (b and d) in the absence (a and b) and presence (c and d) of c-di-GMP substrate were obtained at three different speeds (1,000 rpm, *red*; 12,200 rpm, *green*; 15,000 rpm, *blue*). Data were fitted to monomer-dimer self-equilibrium model (solid *black* line, a, b, c and d) and the fitting residuals are displayed above each panel. For comparison, equivalent monomer-dimer fits in the absence of substrate are overlaid on top of the profile in the presence of enzyme (colored lines, e and f).

Figure 4: Enzymatic assays by dye-labeled 484-MDCC biosensor. (a) Diguanylate cyclase enzymatic assays. The initial velocity of Tbd1265_{DUAL} and Tbd1265_{ADUAL} measured using 1 μ M GTP and 10 μ M GTP. *Inset*: The initial velocity slope being plotted from the accumulation of c-di-GMP against time where (a) Tbd1265_{DUAL} with 10 μ M GTP (b) Tbd1265_{DUAL} with 1 μ M GTP (c) Tbd1265_{ADUAL} with 10 μ M GTP (d) Tbd1265_{ADUAL} with 1 μ M GTP. (b) Phosphodiesterase enzymatic assays. The initial velocity of Tbd1265_{DUAL} and Tbd1265_{ADUAL} measured using 1 μ M c-di-GMP and 5 μ M c-di-GMP. *Inset*: The initial velocity slope being plotted from the depletion of c-di-GMP against time where (a) Tbd1265_{DUAL} with 5 μ M c-di-GMP (b) Tbd1265_{ADUAL} with 5 μ M c-di-GMP (c) Tbd1265_{DUAL} with 1 μ M c-di-GMP (d) Tbd1265_{ADUAL} with 1 μ M c-di-GMP

Figure 5: Structural characteristic Tbd1265_{DUAL}. (a) Representation of the Tbd1265_{DUAL} monomer. Tbd1265_{DUAL} is composed of GGDEF domain (*blue*) and EAL domain (*yellow*). The protein secondary structures elements are labeled, as well as the N- and C-terminal that represented by *green* and *red* ball. The missing GGDEF-EAL linker is represented by *grey* colored dotted line. (b) Interface area of Tbd1265_{DUAL}. The secondary structure elements that involved in the dimers interface interaction are labeled and colored in *magenta*. (c) Schematic representation of dimeric Tbd1265_{DUAL}. The dimeric Tbd1265_{DUAL} is shown from two angles separated by 30°. The first monomer is colored based on its domain color in (a) and *grey* for its counter-part monomer. The interface area of two Tbd1265_{DUAL} monomers is

presented in cartoon and colored in *green*. The A-site and I-site of the GGDEF domain, and active site of EAL domain are colored in *cyan*, *magenta* and *red*, respectively.

Figure 6: Structure of the GGDEF domain of Tbd1265_{DUAL}. (a) Sequence alignment of nucleotide binding motifs in GGDEF-containing DGC proteins. GGDEF domain sequences from Tbd1265, PleD, WspR and FimX were aligned using ClustalW. The GGDEF motif and I-site are labeled. At the active site, residues involved in coordinating ion Mg²⁺ (*green*), γ -phosphate moiety of GTP (*purple*) and guanosine base of GTP (*red* star) are indicated. Whereas, residues involved in coordinating c-di-GMP at the active are indicated by *blue* colored star. (b) Close-up view of the nucleotide binding site of Tbd1265_{DUAL} GGDEF domain. The structure of nucleotide-free Tbd1265_{DUAL} GGDEF domain (*blue*) and PleD GGDEF domain bound to non-hydrolyzable GTP analog (PDB code 2v0n; *white*) were superimposed. Residues of both Tbd1265_{DUAL} and PleD involved in coordinating nucleotide and ion Mg²⁺ (*green* sphere) are shown as sticks. The residues and sequence position are indicated (*left*: Tbd1265_{DUAL}, *right*: PleD). (c – d) Cyclic di-GMP-incompatible active and I-sites of Tbd1265_{DUAL} GGDEF domain. The structures of nucleotide-free Tbd1265_{DUAL} GGDEF domain (*blue*) and the GGDEF domain of PleD bound to c-di-GMP at the active (c) and I-site (d) (PDB code 2w25; *white*) were superimposed. Cyclic di-GMP coordinating residues of PleD and corresponding residues of Tbd1265_{DUAL} are shown as sticks.

Figure 7: Proposed mechanistic model of Tbd1265 regulation. Three cartoon models for GGDEF-EAL-containing periplasmic sensory protein Tbd1265 in resting (*ii*), inhibition (*i*) and activation (*iii*) conformation are illustrated based on available structural information. The α helix, transmembrane segment (TM) and periplasmic binding protein (PBP) were modeled based on sequence alignment and structural-based prediction. The cartoon in resting conformation (*i*) resembles based on the Tbd1265_{DUAL} crystal structure. In the resting conformation, the GGDEF domains are positioned adjacent to each other but yet promote to inhibit conformation. To bind the ligand from surroundings, GGDEF domains have to undergo 180° rotation to achieve active conformation (*iii*). Upon binding, the GGDEF domains change its active conformation to inhibit conformation (*ii*).

Table 1

	Tbd1265 _{DUAL}
Data collection	
X-ray source	PXIII, SLS
Wavelength (Å)	1.000
Crystallographic parameters	
Space group	C2
Unit cell dimensions (Å)	195.12 187.94 87.31
α, β, γ (°)	90.0, 106.97, 90.0
Data collection statistics^a	
Resolution limit (Å)	37.74 – 3.4
No. of observed reflections	527,644
No. of unique reflections	74,143
Completeness (%)	94.7 (78.3)
Redundancy	7.1 (5.1)
Average $I/\sigma(I)$	14.9 (2.2)
R_{merge} (%)	0.103 (0.693)
Refinement	
R factor ($R_{\text{work}}/R_{\text{free}}$) (%)	22.9/27.8
Model contents/average B (Å ²)	
Protein atoms	12,587/53.0
Water molecules	
Root mean square deviation	
Distance (Å)	0.010
Bond angles (°)	1.2
Estimated coordinate error (Å)	0.050
Ramachandran plot^b	
Favored	
Allowed	
Outlier

Figure 3

Figure 5

Figure 6

Figure 7

Substrate Channel Flexibility in *Pseudomonas aeruginosa* MurB Accommodates Two Distinct Substrates

Ming Wei Chen^{1,2}, Bernhard Lohkamp¹, Robert Schnell¹, Julien Lescar², Gunter Schneider^{1*}

¹Department of Medical Biochemistry and Biophysics, Karolinska Institutet, Stockholm, Sweden, ²School of Biological Sciences, Nanyang Technological University, Singapore, Singapore

Abstract

Biosynthesis of UDP-*N*-acetylmuramic acid in bacteria is a committed step towards peptidoglycan production. In an NADPH- and FAD-dependent reaction, the UDP-*N*-acetylglucosamine-enolpyruvate reductase (MurB) reduces UDP-*N*-acetylglucosamine-enolpyruvate to UDP-*N*-acetylmuramic acid. We determined the three-dimensional structures of the ternary complex of *Pseudomonas aeruginosa* MurB with FAD and NADP⁺ in two crystal forms to resolutions of 2.2 and 2.1 Å, respectively, to investigate the structural basis of the first half-reaction, hydride transfer from NADPH to FAD. The nicotinamide ring of NADP⁺ stacks against the *si* face of the isoalloxazine ring of FAD, suggesting an unusual mode of hydride transfer to flavin. Comparison with the structure of the *Escherichia coli* MurB complex with UDP-*N*-acetylglucosamine-enolpyruvate shows that both substrates share the binding site located between two lobes of the substrate-binding domain III, consistent with a ping pong mechanism with sequential substrate binding. The nicotinamide and the enolpyruvyl moieties are strikingly well-aligned upon superimposition, both positioned for hydride transfer to and from FAD. However, flexibility of the substrate channel allows the non-reactive parts of the two substrates to bind in different conformations. A potassium ion in the active site may assist in substrate orientation and binding. These structural models should help in structure-aided drug design against MurB, which is essential for cell wall biogenesis and hence bacterial survival.

Citation: Chen MW, Lohkamp B, Schnell R, Lescar J, Schneider G (2013) Substrate Channel Flexibility in *Pseudomonas aeruginosa* MurB Accommodates Two Distinct Substrates. PLoS ONE 8(6): e66936. doi:10.1371/journal.pone.0066936

Editor: Beata G. Vertessy, Institute of Enzymology of the Hungarian Academy of Science, Hungary

Received: March 8, 2013; **Accepted:** May 10, 2013; **Published:** June 21, 2013

Copyright: © 2013 Chen et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported by the European Commission through the 7th Framework program, grant agreement HEALTH-F3-2008-223461, the Swedish Science Council (www.vr.se) and by the Clas Groschinsky Foundation. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: Gunter.Schneider@ki.se

Introduction

Peptidoglycan (PG) is a vital component of the cell wall in most prokaryotic organisms. The high molecular weight polymer provides osmotic stability and defines the size and shape of bacterial cells [1]. The cell wall PG of bacteria comprises a glycan chain of alternating *N*-acetyl-glucosamine and *N*-acetyl-muramic acid units, which are linked via peptide crosslinks. In Gram-negative bacteria, the formation of the high molecular weight PG polymer takes place in the periplasmic space, whereas the disaccharide-pentapeptide building blocks of PG are produced in the bacterial cytoplasm, linked to polyprenyl lipid chains and anchored to the cytoplasmic membrane [2]. Following translocation to the periplasmic side of the membrane, the PG building blocks are accessible to penicillin binding proteins, the enzymes responsible for the polymerization of PG units [2–4]. The cytoplasmic biosynthetic pathway leading to these PG-units involves ten enzymes [3,4]. UDP-*N*-acetylglucosamine (UNAG) is first synthesized from fructose-6-phosphate, acetyl-CoA and idyl triphosphate. To produce the glycan building block UDP-*N*-acetylmuramic acid (UNAM), an enolpyruvyl group is transferred from phosphoenolpyruvate to UNAG by the UNAG enolpyruvyl transferase, MurA [5]. The resulting UDP-*N*-acetylglucosamine-enolpyruvate (UNAGEP) is reduced by the UNAGEP reductase,

MurB [6], to generate UNAM. Then, the pentapeptide stem is added in consecutive steps to the lactoyl moiety of UNAM [3,4].

All MurB proteins known to date are monomeric flavoenzymes that bind strongly a stoichiometric amount of flavin adenine dinucleotide (FAD). Extensive kinetic studies on *Escherichia coli* MurB (EcMurB) revealed that nicotinamide adenine dinucleotide phosphate (NADPH), but not nicotinamide adenine dinucleotide (NADH), is utilized as the cosubstrate [6,7]. The catalytic cycle can be divided into two half reactions and follows the scheme of a ping-pong bi bi mechanism (Fig 1). NADPH transfers the 4-*pro*-S hydrogen to the isoalloxazine N5 atom during the first half-reaction [3,6] and dissociates from the enzyme. This step is followed by the binding of UNAGEP and the hydride transfer from FADH₂ to the vinyl ether of UNAGEP. The complete reaction thus converts NADPH and UNAGEP to NADP⁺ and UNAM, respectively. Crystal structures have been determined for MurB enzymes of *E. coli* (EcMurB, PDB code 1MBT) [8], *Staphylococcus aureus* (SaMurB, 1HSK) [9], *Thermus caldophilus* (TcMurB, 2GQI) [10], *Listeria monocytogenes* (LmMurB, 3TX1) and *Vibrio cholerae* (VcMurB, 3I99), which share a common three-domain architecture. Domains I and II form the FAD-binding module that is conserved in a wide range of oxidoreductases including cytokinin dehydrogenase [11] and the broad-specificity vanillyl-alcohol oxidase [12]. The substrate-binding domain III is unique but has two variations that divide MurB enzymes into two

types. Type I MurB, exemplified by the prototypical EcMurB, contains a tyrosine loop and a protruding $\beta\alpha\beta\beta$ fold. The type II enzymes lack these features on domain III and are exemplified by SaMurB and TcMurB. Furthermore, type IIa MurB utilizes a serine residue as a proton donor to the UNAGEP carbanion intermediate [9,13,14], while the type IIb group employs cysteine instead [10].

In the crystal structure of EcMurB in complex with UNAGEP (PDB code 2MBR), the enolpyruvyl moiety of UNAGEP packs against the *si* face of the isoalloxazine ring of FAD, bringing the ultimate hydride acceptor atom C3 of the enolpyruvyl (C3e) close to the isoalloxazine N5 atom [15]. However, no experimental structure of the MurB-NADP(H) complex has been reported to date. Based on the comparison between substrate-free and UNAGEP-bound EcMurB, NADPH was proposed to bind in the UNAGEP site or in a perpendicular channel close to the C-terminus of the enzyme, both resulting in the unusual binding of nicotinamide to the *si* face of isoalloxazine [8]. Alternatively, the nicotinamide ring could approach the FAD *re* face by rearranging a loop at the domain-II-domain-III interface, thus allowing NADPH to bind perpendicular to the UNAGEP site [15]. NMR studies of the EcMurB-NADP⁺ complex suggested that NADPH and UNAGEP share at least a binding subsite [16,17]. Moreover, NADP⁺ can reverse the reduction of FAD, bind to both reduced and oxidized forms of MurB with high affinity [18], and inhibit the reaction non-competitively with regard to both NADPH and UNAGEP [7].

Due to the essentiality of peptidoglycans in most pathogenic bacteria, PG biosynthetic enzymes are the target of several antibiotics. Substrate-bound structures of MurB are therefore of great value in rational design of improved and novel inhibitors of peptidoglycan biosynthesis. As part of the effort to elucidate the structures of important *Pseudomonas aeruginosa* drug targets [19], we produced and crystallized *P. aeruginosa* MurB (PaMurB; gene locus PA2977). Here, we present the crystal structure of NADP⁺-bound PaMurB, which establishes the co-localization of the NADPH and UNAGEP binding sites and supports a ping pong reaction mechanism with sequential substrate binding. The crystal structure of the ternary complex of PaMurB-FAD-NADP⁺ provides a snapshot of the first step of the catalytic cycle in this enzyme class, implying an unusual hydride transfer from nicotinamide to the *si* face of flavin.

Materials and Methods

Production and purification of PaMurB

The coding sequence for PaMurB (PA2977 gene) was amplified by PCR using the *P. aeruginosa* PAO1 genomic DNA template (ATCC 47085) and cloned into the pNIC28-Bsa4 vector (GenBank Accession No. EF198106) using the ligation-independent cloning method. The expression construct encodes a fusion protein with an N-terminal His₆-tag (H₂N-MHHHHHHSSGVDLGTENLYFQ*SM) and a TEV protease cleavage site (indicated by *). The construct was expressed in *E. coli* BL21(DE3) in 2 L LB media containing 50 µg/mL kanamycin inoculated with 20 mL of an overnight seed culture. The culture was grown with vigorous shaking to OD₆₀₀ of 0.7, chilled on ice, induced with 0.2 mM IPTG, and grown at 18°C for 20 hours. Cells were harvested by centrifugation at 3300 g and resuspended in lysis buffer (50 mM Tris-HCl pH 7.5, 0.5 M NaCl, 5% v/v glycerol). Lysate was produced by ultrasonication on ice (8 cycles of 30 s), clarified by centrifugation at 45000 g and subsequently filtered. The lysate was incubated with 1 mL Ni-NTA resin (Qiagen) in the presence of 10 mM imidazole for one hour with gentle agitation. The protein was eluted with buffer containing 200 mM imidazole. Subsequently the buffer was exchanged to 25 mM Tris-HCl pH 8.0 containing 150 mM NaCl using a PD-10 column (GE Healthcare). For His₆-tag removal 2 mM DTT and His₆-tagged TEV protease (at 1:200 protease-to-PaMurB mass ratio) were added and incubated at 20°C for 16 hours. A subsequent Ni-NTA affinity step removed uncleaved protein and the His₆-tagged TEV protease. All following steps and experiments are based on tag-free PaMurB carrying one extra serine residue in the N-terminal position. Purified PaMurB was concentrated and loaded on an S200 prep grade gel-filtration column (GE Healthcare) equilibrated with a buffer containing 25 mM Tris-HCl pH 8.0 and 150 mM NaCl, and eluted as a single peak. Pure fractions were pooled and concentrated with Vivaspin micro-concentrator units (Sartorius) to 25 mg/mL, flash-frozen in liquid nitrogen, and stored at -80°C.

Crystallization of PaMurB

PaMurB at 25 mg/mL concentration with 2 mM unbuffered NADPH (Sigma-Aldrich) or 2 mM Tris-(pH 8.0)-buffered NADP⁺ sodium salt (Sigma-Aldrich) was used to set up 96-well format

Figure 1. MurB-catalyzed FAD- and NADPH-dependent reduction of UNAGEP to UNAM. In the first half reaction (upper panel) the 4-*pro-S* hydrogen and two electrons from the reduced nicotinamide C4 atom are transferred to the N5 atom of the isoalloxazine resulting in NADP⁺ and reduced FAD. In the second half reaction (lower panel) the hydride is transferred to the C3 atom of the UNAGEP enolpyruvyl group, reducing the enolpyruvyl to a lactoyl group and thus converting UNAGEP to UNAM. doi:10.1371/journal.pone.0066936.g001

tallization trials. A Mosquito crystallization robot (ITP Tech Ltd, Melbourn UK) and commercially available tallization screening reagents (JSCG+ and PACT from gen, SaltRX from Hampton Research) were used. Two different crystal forms were obtained and used for X-ray diffraction data collection. During crystal optimization, crystal form A was obtained by mixing 1 μ L of the protein solution mg/ml PaMurB, 2 mM unbuffered NADPH) with 1 μ L of precipitant (0.1 M bis-tris propane pH 7.0, 0.2 M sodium potassium tartrate, 15% w/v PEG 3350), followed by equilibration in 0.5 mL of reservoir solution via vapour diffusion. Crystals in these drops were cryoprotected with a solution containing 15% w/v PEG3350 and 2 mM NADPH before flash-freezing. Crystal form B was obtained from a condition containing 10 mM potassium phosphate, 20% v/v glycerol, 16% w/v PEG 300 and 2 mM Tris-buffered NADP⁺ sodium salt and were frozen directly without addition of a cryoprotectant.

Data collection and processing

X-ray diffraction data for crystal forms A and B were collected at the European Synchrotron Radiation Facility (ESRF) beamline ID14-4 and ID14-1, respectively. Data were indexed and integrated with iMosflm [20], and subsequently scaled using XSCALE [21] (for form A) and AIMLESS [22] (for form B) from the CCP4 package [23]. Form A crystals are of space group C2, with cell dimensions $a = 153.7$ Å, $b = 154.3$ Å, $c = 64.5$ Å and $\beta = 102.3^\circ$, while form B crystals belong to space group P6₁ with parameters $a = 88.7$ Å and $c = 100.2$ Å. We could not detect signs of twinning in these crystals using the twin tests [24] implemented in CTRUNCATE. Crystals of form A showed a somewhat streaky and anisotropic diffraction pattern, reflected in a relatively high R_{merge} value. The X-ray diffraction data statistics are given in Table 1.

Structure solution, refinement and analysis

The structure of PaMurB was solved by molecular replacement using MOLREP [24] with the atomic coordinates of the complex EcMurB with UNAGEP (PDB code 2MBR), stripped of all ligands and water molecules as search model. Molecular replacement solutions for data of crystal forms A and B were obtained independently with the search model, and contained four monomers per asymmetric unit, respectively (Table 1). Models were built by alternating rounds of model-building in COOT [25] and refinement in REFMAC5 [26]. The protein models of crystal forms A and B were refined to 2.23 and 2.10 Å, respectively. Where one model displayed ambiguous electron density, model-building was aided by referring to the other model. TLS analysis was performed for both protein models using the SMD server [27,28]. The results were used for restrained refinement with TLS in REFMAC5, with 7 segments for crystal form B, resulting in an improvement of $R_{\text{work}}/R_{\text{free}}$ of 2.0%/%. Two TLS segments per monomer were employed during final refinement of crystal form A. Space groups were validated using ZANUDA from the CCP4 suite [23]. The quality of the models was assessed using MolProbity [29].

Superimposition of structural models was performed in COOT; the SSM [30] method was used for superimposing protein backbones while the LSQ method was used when ligand-aided structural alignment was required. Structure-aided multiple sequence alignment, secondary structure assignment and generation of Figure 2 was performed using STRAP [31], STRIDE [32] and ESPript [33], respectively. The ligand-protein interaction plots were generated by LIGPLOT+ [34] and manually inspected. The remaining figures were made with PyMOL

Table 1. Data collection and structure refinement statistics.

Data set	Crystal form A	Crystal form B
<i>Data collection statistics</i>		
Beam line	ID14-4, ESRF	ID14-1, ESRF
Wavelength (Å)	0.9762	0.9334
Resolution (Å)	58.38–2.23 (2.35–2.23)	50.11–2.10 (2.16–2.10)
Space group	C2	P6 ₁
Cell parameters a/b/c (Å)	153.7, 154.3, 64.5	88.7, 88.7, 100.2
$\alpha/\beta/\gamma$ (°)	90.0, 102.3, 90.0	90.0, 90.0, 120.0
Unique reflections	69794 (9862)	26203 (2179)
Redundancy	3.0 (2.6)	7.6 (7.6)
I/σ	3.8 (1.6)	13.1 (2.9)
Completeness (%)	98.2 (95.7)	100.0 (100.0)
R_{merge}	0.215 (0.686)	0.109 (0.666)
R_{pim}	0.143 (0.507)	0.064 (0.390)
Wilson B factor (Å ²)	20.0	21.6
<i>Refinement statistics</i>		
Data range (Å)	54.33–2.23 (2.29–2.23)	44.39–2.10 (2.16–2.10)
Used reflections	66255	24851
R_{work}	0.222 (0.348)	0.180 (0.272)
R_{free}	0.258 (0.367)	0.236 (0.320)
No. of monomers/ASU	4	1
No. of protein atoms	10580	2732
No. of FAD molecules	4	1
No. of NADP ⁺ molecules	4	1
No. of water molecules	340	164
No. of K ⁺	4	–
<i>Average B-factors (Å²)</i>		
Protein	33.1	42.5
FAD	14.8	25.4
NADP ⁺	25.2	30.0
Water molecules	27.1	41.0
K ⁺ ions	19.0	–
<i>R.m.s. deviations</i>		
Rmsd bond length (Å)	0.014	0.019
Rmsd bond angles (°)	1.700	2.125
<i>Ramachandran plot (% residues)</i>		
Allowed	1319 (98.1%)	322 (95.8%)
Generously allowed	25 (1.9%)	13 (3.9%)
Disallowed	0 (0.0%)	1 (0.3%)

Values in parentheses are for the highest-resolution shell, ASU: Asymmetric unit. doi:10.1371/journal.pone.0066936.t001

(www.pymol.org). The crystallographic data and models for crystal form A and B were deposited with the Protein Data Bank, accession codes 4JAY and 4JB1, respectively.

Results

Structure determination and quality of the models

The crystal structure of PaMurB in complex with NADP⁺ was determined in two different crystal forms by molecular replacement. Crystal form A belongs to the space group C2 with four molecules per asymmetric unit, while crystal form B belongs to P6₁

Figure 2. Structure-based sequence alignment of PaMurB against three MurB enzymes with experimentally determined crystal structures. Each of the three enzymes represents one type of MurB. Annotations of secondary structure are based on PaMurB and are colored red for FAD-binding domain I, blue for domain II and pink for substrate-binding domain III. Strictly conserved residues are shaded in red and conserved regions are boxed; the GXG motif important for FAD-binding is indicated and shaded in yellow. The sequences compared are *Pseudomonas aeruginosa* MurB (PA2977), *Escherichia coli* MurB (EC(I), type I, PDB code 2MBR), *Staphylococcus aureus* MurB (SA(IIa), type IIa, PDB code 1HSK), and *Thermus caldophilus* MurB (TC(IIb), type IIb, PDB code 2GQT). Residues involved in cofactor- and substrate-binding are indicated: ◆, residues that interact with FAD in both PaMurB and EcMurB; ▲, equivalent residues that interact with UNAGEP in EcMurB; □, residues that interact with NADP⁺; ○, residues that interact with both UNAGEP and NADP⁺; *, residues that coordinate the putative catalytic metal ion.
doi:10.1371/journal.pone.0066936.g002

with one molecule per asymmetric unit; the solutions are in agreement with solvent content analysis by using Matthews coefficient [35]. The models were refined to $R_{\text{work}}/R_{\text{free}}$ of 22.3%/25.8% and 18.2%/23.4% using data to the resolution of 2.23 and 2.10 Å, respectively. Ligands and the polypeptide chains are well-defined except the first two residues of the N-termini (Ser-0, Met-1). Also for crystal form A, the resulting electron density maps were of good quality, most likely due to the presence of four-fold non-crystallographic symmetry. Statistics for the refinement and the structural models are provided in Table 1. Ala-83, located on the β 6- β 7 turn, is a Ramachandran plot outlier in crystal form B with well-defined electron density, possibly because of crystal contacts.

PaMurB is a type I UNAGEP reductase

Based on the sequence alignment, PaMurB is expected to be a type I UNAGEP reductase, similar to EcMurB (Fig 2). Size-exclusion chromatography indicated a monomeric state in solution (Fig. S1). Consistent with this finding, an analysis of oligomeric assembly by the PISA server [36] indicated that PaMurB exists also as a monomer in both crystal forms. The two crystal structures are identical in protein and ligand conformations, except three loops involved in crystal contacts on one side of the protein (β 6- β 7, β 12- α 5 and α 5- α 6) with an overall root-mean-squared difference, rmsd, of 0.8 Å, based on 336 C α atoms. In the following results and figures are based on crystal form B unless otherwise noted.

PaMurB consists of three domains: the FAD-binding domains I (aa 1-75 and 336-339) and II (aa 76-191), and the substrate-binding domain III (aa 192-335) (Fig 3A). The structural framework of the FAD-binding module is provided by two β -

ets ($\beta 1$ - $\beta 2$ - $\beta 5$ - $\beta 3$ and $\beta 6$ - $\beta 7$ - $\beta 12$ - $\beta 10$ - $\beta 11$) packed against king helices $\alpha 1$ and $\alpha 2$, respectively. The buried cofactor binding pocket is formed by the GXG motif on the $\beta 3$ - $\beta 4$ strand- $\alpha 2$ (Fig 2) and the C-terminal strand $\beta 20$ contributed by main I, and the $\beta 7$ - $\alpha 3$ loop, the $\eta 1$ - $\alpha 3$ - $\beta 8$ region, and $\alpha 4$ - $\beta 12$ n domain II. The substrate-binding domain III is adjacent to FAD pocket and can be visualized as two lobes (Fig 3). Lobe 1 consists of the linker originating from domain II and helices $\alpha 5$ - $\alpha 6$, racting with domain II. Lobe 2 is adjacent to domain I and ures the $\beta 13$ - $\beta 19$ - $\beta 18$ - $\beta 17$ sheet and the prominent $\beta 14$ - $\alpha 7$ - $\beta 16$ outcrop. The substrate binding site, located between the es, is more shallow and solvent-accessible than the cofactor site. The structures of substrate-bound PaMurB and EcMurB (PDB e 2MBR) are very similar (rmsd 1.3 Å based on 321 C α atoms) (Fig 3B), consistent with the moderately high sequence identity of % (Fig. 2). Compared to EcMurB, PaMurB has a longer N-ninus that forms a partially disordered strand, and a shorter C-ninus that completes the fold of domain I (Fig 2, Fig 3A). ctural alignment of PaMurB with EcMurB (type I), SaMurB e IIa) and TcMurB (type IIb) highlights the presence of the β - $\alpha 5$ loop and the consequential displacement of $\alpha 5$, and the truding $\beta\alpha\beta\beta$ fold ($\beta 14$ - $\alpha 7$ - $\beta 15$ - $\beta 16$), which are not observed ype IIa/IIb enzymes (Fig 3B, Fig 2).

Figure 3. Overall structure of the ternary complex of PaMurB with FAD and NADP⁺. **A.** Stereo view of the crystal structure of MurB with bound FAD and NADP⁺. The enzyme is shown in cartoon representation and comprises FAD-binding domain I (red) and domain II (blue), and the substrate-binding domain III (pink). FAD and NADP⁺ are shown as yellow and cyan stick models, respectively. NADP⁺ occupies the channel between the two lobes of domain III (in this view: lobe 1; right, lobe 2). Relevant secondary structure elements are labeled. **B.** Stereo view of the superimposition of the C α traces of MurB (olive green), SaMurB (orange) and TcMurB (red) against EcMurB (black). PaMurB and EcMurB display highly similar structures in their respective NADP⁺- and UNAGEP-bound complexes. In domain III, the II MurB enzymes lack the tyrosine loop preceding helices $\alpha 4$ and as well as the protruding $\beta\alpha\beta\beta$ fold on lobe 2. doi:10.1371/journal.pone.0066936.g003

Conserved FAD binding in PaMurB

PaMurB, FAD and NADP⁺ form a ternary complex at a 1:1:1 ratio. The FAD cofactor has well-defined electron density, occupying the same binding site in a conformation identical to the one previously reported in other MurB structures [8–10,13,15,37] (Fig S2 and Fig S3). FAD is stabilized by multiple hydrogen bonds and van der Waals interactions with protein main chain and side chain atoms in the cofactor site. The binding site residues are essentially identical to those observed in the EcMurB structure (Fig S2). The isoalloxazine ring system is planar, suggesting that FAD is in its oxidized form. The yellow protein was immediately bleached upon addition of NADPH but recovered the yellow color, suggesting that reduction by NADPH is reversible under aerobic conditions.

NADP(H) and UNAGEP share the same binding site

In the substrate binding cleft in domain III strong continuous electron density was observed that corresponds well to a bound NADP⁺ molecule (Fig 4A). NADP⁺ displays an extended “S” conformation where the 2-hydroxyl group of the nicotinamide ribose moiety interacts with the 2'-phosphate group of adenosine. The nicotinamide amide group is located on the *si* face of the isoalloxazine ring of FAD, and forms hydrogen bonds with the conserved Ser-239 and Glu-335 residues. The two rings are nearly parallel and juxtaposed together, with the nicotinamide C4n atom in close proximity to the isoalloxazine N5 atom (3.1 Å and 2.8 Å in crystal forms A and B, respectively). This suggests that the nicotinamide group is in a position that favors the first half-reaction (Fig 1). The phosphate-ribose backbone is stabilized by various active site residues and water bridges. Tyr-132 and Arg-166 interact with the nicotinamide ribose, while the phosphate backbone is stabilized by Lys-227 (Fig 4B). The phosphoribose in the adenosine moiety forms hydrogen bonds with Tyr-196, Asn-243 and Lys-272. The adenosine moiety, which is located at the entrance of the substrate channel, adopts the *anti* conformation with the adenine ring forming π - π stacking interactions with Tyr-196 and Tyr-264 from lobe 1 and the $\beta\alpha\beta\beta$ outcrop of lobe 2, respectively (Fig 4A&B).

Superimposition of the structure of the PaMurB ternary complex with UNAGEP-bound EcMurB (PDB code: 2MBR; rmsd 0.19 Å for all FAD atoms) reveals that NADP⁺ occupies the same binding site as UNAGEP, between the two lobes of the substrate-binding domain III (Fig 5A). The nicotinamide moiety occupies the position of the enolpyruvyl group of UNAGEP. The C4n atom of the nicotinamide ring, which is oxidized by FAD in the first half-reaction, is approximately aligned with C3e of the enolpyruvyl which in turn would receive the hydride from FADH₂ during the second half-reaction (Fig 1). This is thus an overlapping sub-site for both substrates. However, the positions of the two molecules diverge at the pyrophosphate backbone and the non-reactive nucleotide moieties point in opposite directions (Fig 5B). Although the overall protein structure of the two complexes is preserved, the divergent binding of the two substrates causes structural changes in the PaMurB domain III; compared to the EcMurB-UNAGEP complex, the two lobes of domain III move away from the binding cleft. An analysis on domain movements using the program DynDom [38,39] also reveals bending and rotational motions of the Tyr-196 loop and helices 5 and 6, as NADP⁺ binds closer to this lobe. Tyr-196 adopts a different rotamer, stabilizing NADP⁺ through a hydrogen bond with the adenosyl 2'-phosphate, and stacking interactions with the NADP⁺ adenine ring and Tyr-264 (Fig 4A). In addition, Lys-272 rotates inwards to form hydrogen bonds with the nicotinamide ring. Gln-288, which interacts with the diphosphate backbone of UNAGEP,

Figure 4. NADP⁺ and the substrate binding site. **A.** Stereo view of NADP⁺ bound in the substrate channel of PaMurB. NADP⁺ is depicted as a cyan stick model. The $F_o - F_c$ omit electron density of NADP⁺, contoured at 3.0 σ , indicates tight binding of the ligand. The nicotinamide ring stacks against the isoalloxazine ring system of FAD (shown in yellow). 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 stabilized by stacking interactions with Tyr-196 and Tyr-264. **B and C.** Comparison of protein residues interacting with NADP⁺ and UNAGEP (based on EcMurB, PDB code 2MBR). NADP⁺ and UNAGEP are shown as ball-and-stick models in cyan and green, respectively. Red radiating lines around ligand atoms indicate van der Waals contacts, while green dashes represent potential hydrogen bonds. Ball-and-stick models are shown for binding site residues that provide polar interactions but not those involved in van der Waals interactions only. Orange shading highlights binding site residues that are conserved and involved in binding both substrates. doi:10.1371/journal.pone.0066936.g004

adopts a different conformation and does not contact NADP⁺ because of the divergent binding mode (Fig 4C). A significant change in van der Waals contacts is also observed (Fig 4B&C). Although the non-reactive moieties do not overlap completely, the substrate channel formed by the two lobes of domain III can be visualized as three loci or subsites arranged from the active site interior to the channel opening (Fig 5B). Each locus is respectively occupied by the substrate moiety, pyrophosphate backbone and the nucleotide portion of either NADP(H) or UNAGEP.

PaMurB contains a potassium ion in the active site

A strong residual electron density in the vicinity of the nicotinamide ring was modeled as a metal ion in crystal form A

(Fig 6A). The identity of the cation was inferred by alternative rounds of refinement using a water molecule, a sodium or potassium ion (present in the mother liquor) and a nickel ion (possibly derived from the Ni-NTA affinity chromatography step) as possible ligand. Refinement with potassium gave the most satisfactory results in terms of B factors (Table S1) and difference densities. Sodium and water were ruled out as a possibility because of the abnormally low B factors after refinement and in the case of sodium also the mismatch in ligand-metal distances [40]. The metal site is located at the active site in the vicinity of the nicotinamide and the isoalloxazine rings of NADP⁺ and FAD, respectively. The potassium ion displays a pentagonal bipyramidal coordination geometry with two main chain oxygen atoms (Ala-

Figure 5. NADP⁺ shares the same substrate binding site with UNAGEP. **A.** Stereo view of PaMurB (black ribbon) and UNAGEP-bound EcMurB (green ribbon, 2MBR) superimposed based on their FAD atomic coordinates. FAD, NADP⁺ and UNAGEP are depicted as stick models in yellow, red and green, respectively. NADP and UNAGEP occupy the same substrate channel located between the lobes of domain III. **B.** Co-localization of NADP(H) and UNAGEP substrate moieties on the *si* face of the FAD isoalloxazine ring system. The reactive moieties of the two substrates align well after superimposing the structures of PaMurB-NADP⁺ and EcMurB-UNAGEP (PDB code 2MBR) based on the FAD atomic coordinates. However, the non-reactive parts of the ligands diverge in the binding site, which can be visualized as three loci. S: the substrate moiety that reacts with FAD. B: the backbone region consisting of sugar and diphosphate. N: the non-reactive nucleotide moiety, which shows the greatest deviation. The remodeling of the binding site according to individual substrates is discussed in the text. doi:10.1371/journal.pone.0066936.g005

side chain carboxyl oxygen atom from Glu-335, side chain oxygen atom of Asn-57 and the O7n oxygen atom of the nicotinamide ring as ligands. The interactions between Glu-335, nicotinamide and the metal ion position the nicotinamide in 3.1 Å from the isoalloxazine N5, and 109.3° from the isoalloxazine plane (defined as the C4n-N5-N10 angle) (Fig 6B). A strikingly close alignment between the enolpyruvyl from EcMurB and the nicotinamide is also observed upon superimposing the two structures, with a positional difference of 0.4 Å between the reactive C3e and C4n atoms. Crystal form B also shows electron density at this position, but refinement suggests that this site is occupied by a water molecule or a metal ion at lower occupancy (Table S1). Potassium binding apparent in only crystal form A could be due to the five times higher potassium concentration in crystallization liquor for form A compared to form B crystals. It is noteworthy that the potassium concentration used to grow crystal form A is closer to physiological concentrations of this metal ion in the bacterial cell [41] than the significantly lower concentration in form B crystals.

Discussion

Hydride transfer occurs via the isoalloxazine *si* face

The crystal structure of the PaMurB-NADP⁺ complex presented in this work allows the comparison of NADP(H)- and UNAGEP-bound MurB structures. As proposed based on NMR [16,17] and kinetic studies [7], the binding of NADPH and UNAGEP to MurB occurs in the same binding site and is mutually exclusive. The nicotinamide ring of NADP(H) occupies the same position as the enolpyruvyl group of UNAGEP, on the *si* face of the FAD isoalloxazine ring (Fig 5, Fig 6B). In the PaMurB ternary complexes, the nicotinamide C4n atom is located close to the N5 atom of the isoalloxazine ring (2.8–3.1 Å) and the 4-*pro*-S hydrogen atom of the *sp*²-hybridized C4n atom would point towards the N5 atom receiving the hydride ion (Fig 6B). The co-linear alignment of C4n, the 4-*pro*-S hydrogen and the isoalloxazine N5 atoms facilitates hydride transfer in the first half-reaction of the enzyme, the reduction of FAD. Thus, the conformation of NADP⁺ we observed most likely resembles the actual reactive complex. The UNAGEP enolpyruvyl C3e atom in the MurB-UNAGEP complex occupies the same position as the nicotinamide C4n atom, suitable for the hydride transfer in the second half-reaction. The flexibility

Figure 6. Potassium binding site in PaMurB and its proposed role in catalysis. **A.** Structure of the potassium binding site in PaMurB crystal form A. The coordination sphere is formed by the carboxamide oxygen of NADP⁺ nicotinamide in addition to two side chain oxygens and two main chain oxygens from the protein. The potassium ion (gold sphere) and its $F_o - F_c$ omit difference density contoured at 5.0 σ (green mesh) are shown. **B.** The active site potassium ion assists in substrate orientation and binding. Superimposition of the PaMurB crystal form A structure and the EcMurB-UNAGEP complex (PDB code 2MBR) based on FAD atomic coordinates shows that the C2-C3-C4 locus of NADP⁺ nicotinamide (in cyan) spatially overlap with the enolpyruvyl group of UNAGEP (in green). Both substrate moieties are bound to Glu-335 and the backbone amine of Ser-239. The nicotinamide C4n atom (cyan sphere), which transfers a hydride to the isalloxazine N5 atom, coincides with the enolpyruvyl C3e (green sphere), which receives the hydride during the second half-reaction. The geometric relation of the C4n atom to the isalloxazine is indicated. In synergy with the isalloxazine ring, Glu-335 and Ser-239, the potassium ion positions NADPH and UNAGEP such that the substrate carbons are in the optimal position for hydride transfer.

doi:10.1371/journal.pone.0066936.g006

of the substrate-binding domain accommodates the distinct substrates enabling hydride transfer to the isalloxazine (Fig 4, Fig 5), without any requirement for significant re-arrangements of FAD, which might increase exposure of the hydride to bulk solvent. The difficulty in the previous NMR study to assign residues in the substrate binding site may reflect this dynamic nature [17]. As seen in EcMurB, Arg-166 and Arg-224 interact with the N1-C2-O2 and the N5 loci of isalloxazine, a common feature of flavoenzymes thought to stabilize the reduced FAD [42].

The NADPH-FAD stereochemistry utilizing the isalloxazine *si* face is rare and has only one known precedent, the NADH-dependent methylenetetrahydrofolate reductase (MetTHFR) that also employs the ping pong mechanism [43,44]. Comparison of NADH- and methylenetetrahydrofolate-bound MetTHFR crystal structures (PDB codes 1ZPT and 1ZP4, respectively) reveals sharing of the substrate-binding cleft and alignment of hydride donor and acceptor substrate atoms, similar to the MurB-substrate complexes, despite having a different nicotinamide orientation relative to the FAD cofactor and a different NADH conformation.

Potassium activates MurB by facilitating substrate orientation and binding

Monovalent cations have been shown to activate MurB [45], and potassium increased MurB activity more efficiently compared to other alkali metal ions and ammonium [7]. The K_m for both NADPH and UNAGEP was lowered and k_{cat} was increased with increasing potassium concentration. Hill plot analysis also suggested that one potassium ion is involved for each active site [45]. The observation of a bound potassium ion in the crystal structure of PaMurB (Fig 6A) provides a structural and chemical explanation for these findings. Superposition of the structure of crystal form A of PaMurB with the coordinates of the EcMurB-UNAGEP complex (PDB code 2MBR) shows that the carboxyl group of the substrate would be in coordination distance to the potassium ion. In the seminal paper by Hogle and colleagues [15] it was also proposed that the water molecule (W624 in PDB file 2MBR) observed at this position in the *E. coli* enzyme could potentially represent a metal site. Together with the conserved

active site residues Ser-239 and Glu-335, which interact with both NADP⁺ nicotinamide and UNAGEP enolpyruvyl (Fig 4), the potassium ion orientates both substrate groups to facilitate MurB catalysis. Direct coordination of the enolpyruvyl group of the UNAGEP substrate would contribute to transition state stabilization during hydride transfer from reduced FAD to the substrate and thus explain the activation of the enzyme by monovalent ions. Also, the first step of the reaction, the reduction of FAD by NADPH is influenced by the lowered K_m for the co-substrate NADPH observed in the presence of potassium ions [7,45], possibly due to the direct coordination of the nicotinamide carboxamide group to the metal ion.

Discrimination against NADH is mediated by Tyr-196 and Lys-272

NADP(H) differs from NAD(H) only by the substitution of the adenosine 2'-OH by 2'-phosphate. Despite the similarity, conversion of UNAGEP to UNAM is undetectable in the presence of NADH [7]. In the structure of the ternary complex of PaMurB, Lys-272 forms hydrogen bonds with the 2'-phosphate and the N3 atom of the adenine moiety (Fig 4). Furthermore, in addition to stabilizing the adenine ring by stacking interactions, Tyr-196 also points towards the 2'-phosphate and forms a hydrogen bond through the phenolic hydroxyl group. These interactions may be important for achieving a catalytically competent ternary complex and orientating the nicotinamide to the correct catalytic site location. NADPH analogs lacking 2'- or 3'-phosphate i.e. 2',3'-cyclic NADP⁺, NAD⁺ and NADH were unable to bind as substrates or inhibitors, while 3'-NADPH was an inefficient substrate [7], presumably because of the loss of stabilizing interactions with Tyr-196 and Lys-272.

Conclusions

The crystal structure of the ternary complex of PaMurB-FAD-NADP⁺ provides direct evidence of overlapping substrate binding sites for NADPH and UNAGEP. Hydride transfer occurs to and from the *si* face of the FAD isalloxazine ring system, with the

for carbon atom C4n of NADPH and the acceptor carbon m C3e of UNAGEP in the same position, approximately 3 Å from the N5 atom of the isoalloxazine ring of FAD. A potassium ion in the active site is proposed to assist in substrate binding and utilization of the UNAGEP transition state. Several important binding site residues are also found to provide the basis for substrate discrimination against NADH. Importantly, a three-point strategy targeting three loci in the substrate channel could prove invaluable in designing new inhibitors against MurB in *P. aeruginosa* as has been demonstrated for EcMurB [46].

Supporting Information

Figure S1 PaMurB elutes as a single monomeric species in size-exclusion chromatography. Purified PaMurB was loaded on the Superdex 200 10/300 column (GE Healthcare) equilibrated with protein buffer (25 mM Tris-Cl pH 8.0, 150 mM NaCl). The protein elutes as a single peak at 16.5 ml volume. The molecular mass of PaMurB calculated based on the calibration curve (insert) is 37.5 kDa, that agrees well with the value calculated from the sequence, 37.7 kDa.

(TIF)

Figure S2 Conservation of the MurB FAD-binding site. FAD binds to PaMurB and EcMurB (PDB code 2MBR) in an identical conformation, involving highly conserved hydrogen bonding and 1 der Waals contacts between the two proteins. FAD molecules are shown as purple ball-and-stick models. Hydrogen bonds and 1 der Waals interactions are indicated as green dashes and red

radiating lines, respectively. Conserved interacting partners are circled.

(TIF)

Figure S3 F_o-F_c omit electron density of FAD. FAD and NADP⁺ are depicted as stick models in yellow and cyan, respectively. The electron density is contoured at 3.0 σ and depicted as a green mesh.

(TIF)

Table S1 Analysis of the putative metal site in PaMurB. A water molecule or one of the three different metal ions, Na⁺, K⁺ or Ni²⁺ were tested to model the residual electron density near the nicotinamide carboxamide group of NADP⁺. All metal ions were refined with their appropriate charges specified. Comparison of the resulting atomic B factor with that of coordinating oxygen atoms and donor-acceptor distances suggest that potassium binds to the site in crystal form A.

(DOCX)

Acknowledgments

We gratefully acknowledge access to synchrotron radiation at the ESRF, Grenoble, France and thank the staff for support at the beamlines.

Author Contributions

Conceived and designed the experiments: GS RS MWC JL. Performed the experiments: MWC RS BL. Analyzed the data: MWC RS BL JL GS. Wrote the paper: MWC GS. Contributed to writing the manuscript: BL RS JL.

References

- Vollmer W, Blanton D, De Pedro MA (2008) Peptidoglycan structure and architecture. *FEMS Microbiol Rev* 32: 149–167.
- van Heijenoort J (1998) Assembly of the monomer unit of bacterial peptidoglycan. *Cell Mol Life Sci* 54: 300–304.
- Barrecau H, Kovac A, Boniface A, Sova M, Gobec S, et al. (2008) Cytoplasmic steps of peptidoglycan biosynthesis. *FEMS Microbiol Rev* 32: 168–207.
- Gautam A, Vyas R, Tewari R (2011) Peptidoglycan biosynthesis machinery: A rich source of drug targets. *Crit Rev Biotechnol* 31: 295–336.
- Brown ED, Vivas EI, Walsh CT, Kolter R (1995) MurA (MurZ), the enzyme that catalyzes the first committed step in peptidoglycan biosynthesis, is essential in *Escherichia coli*. *J Bacteriol* 177: 4194–4197.
- Benson TE, Marquardt JL, Marquardt AG, Etzkorn FA, Walsh CT (1993) Overexpression, purification, and mechanistic study of UDP-N-acetylenolpyruvylglucosamine reductase. *Biochemistry* 32: 2024–2030.
- Dhalla AM, Yanchunas JJ Jr, Ho HT, Falk PJ, Villafranca JJ, et al. (1995) Steady-state kinetic mechanism of *Escherichia coli* UDP-N-acetylenolpyruvylglucosamine reductase. *Biochemistry* 34: 5390–5402.
- Benson TE, Walsh CT, Hogle JM (1996) The structure of the substrate-free form of MurB, an essential enzyme for the synthesis of bacterial cell walls. *Structure* 4: 47–54.
- Benson TE, Harris MS, Choi GH, Ciadella JL, Herberg JT, et al. (2001) A structural variation for MurB: X-ray crystal structure of *Staphylococcus aureus* UDP-N-acetylenolpyruvylglucosamine reductase (MurB). *Biochemistry* 40: 2340–2350.
- Kim MK, Cho MK, Song HE, Kim D, Park BH, et al. (2007) Crystal structure of UDP-N-acetylenolpyruvylglucosamine reductase (MurB) from *Thermus caldophilus*. *Proteins* 66: 751–754.
- Bae E, Bingman CA, Bitto E, Aceti DJ, Phillips GN (2008) Crystal structure of *Arabidopsis thaliana* cytokinin dehydrogenase. *Proteins* 70: 303–306.
- Mattevi A, Fraaije MW, Mozzarelli A, Olivi L, Coda A, et al. (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.
- Benson TE, Walsh CT, Hogle JM (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.
- Benson TE, Walsh CT, 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.
- Benson TE, Filtman DJ, Walsh CT, Hogle JM (1995) An enzyme-substrate complex involved in bacterial cell wall biosynthesis. *Nat Struct Mol Biol* 2: 644–653.
- Farmer BT (1996) Localizing the NADP(+) binding site on the MurB enzyme by NMR. *Nat Struct Biol* 3: 995–997.
- Constantine KL, Mueller L, Goldfarb V, Wittekind M, Metzler WJ, et al. (1997) Characterization of NADP(+) binding to perdeuterated MurB: Backbone atom NMR assignments and chemical-shift changes. *J Mol Biol* 267: 1223–1246.
- Axley MJ, Fairman R, Yanchunas JJ Jr, Villafranca JJ, Robertson JG (1997) Spectroscopic properties of *Escherichia coli* UDP-N-acetylenolpyruvylglucosamine reductase. *Biochemistry* 36: 812–822.
- Moynic L, Schnell R, McMahon SA, Sandoval T, Boulkerou WA, et al. (2013) The AEROPATH project targeting *Pseudomonas aeruginosa*: crystallographic studies for assessment of potential targets in early-stage drug discovery. *Acta Crystallogr Sect F Struct Biol Cryst Commun* 69: 25–34.
- Battye TGG, Kontogiannis L, Johnson O, Powell HR, Leslie AGW (2011) iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr D Biol Crystallogr* 67: 271–281.
- Evans P (2006) Scaling and assessment of data quality. *Acta Crystallogr D Biol Crystallogr* 62: 72–82.
- Evans P (2011) An introduction to data reduction: space-group determination, scaling and intensity statistics. *Acta Crystallogr D Biol Crystallogr* 67: 282–292.
- Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, et al. (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* 67: 235–242.
- Vagin A, Teplyakov A (1997) MOLREP: an automated program for molecular replacement. *J Appl Crystallogr* 30: 1022–1025.
- Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66: 486–501.
- Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* 53: 240–255.
- Painter J, Merritt EA (2006) TLSMD web server for the generation of multi-group TLS models. *J Appl Crystallogr* 39: 109–111.
- Painter J, Merritt EA (2006) Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. *Acta Crystallogr D Biol Crystallogr* 62: 439–450.
- Davis IW, Leaver-Fay A, Chen VB, Block JN, Kapral GJ, et al. (2007) MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res* 35: W375–383.
- Krisinel E, Henrick K (2004) Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr D Biol Crystallogr* 60: 2256–2268.
- Gille C, Frommel C (2001) STRAP: editor for STRuctural Alignments of Proteins. *Bioinformatics* 17: 377–378.

32. Heinig M, Frishman D (2004) STRIDE: a web server for secondary structure assignment from known atomic coordinates of proteins. *Nucleic Acids Res* 32: W500–502.
33. Gouet P, Robert X, Courcelle E (2003) ESPript/ENDscript: Extracting and rendering sequence and 3D information from atomic structures of proteins. *Nucleic Acids Res* 31: 3320–3323.
34. Laskowski RA, Swindells MB (2011) LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. *J Chem Inf Model* 51: 2778–2786.
35. Kantardjicff KA, Rupp B (2003) Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein nucleic acid complex crystals. *Protein Sci* 12: 1865–1871.
36. Krissinel E, Henrick K (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol* 372: 774–797.
37. Harris MS, Herberg JT, Cialdella JL, Martin JP, Benson TE, et al. (2001) Crystallization and preliminary X-ray analysis of UDP-N-acetylenolpyruvylglucosamine reductase (MurB) from *Staphylococcus aureus*. *Acta Crystallogr D Biol Crystallogr* 57: 1032–1035.
38. Hayward S, Berendsen HJC (1998) Systematic analysis of domain motions in proteins from conformational change: New results on citrate synthase and T4 lysozyme. *Proteins* 30: 144–154.
39. Hayward S, Kitao A, Berendsen HJC (1997) Model-free methods of analyzing domain motions in proteins from simulation: A comparison of normal mode analysis and molecular dynamics simulation of lysozyme. *Proteins* 27: 425–437.
40. Harding M (2002) Metal-ligand geometry relevant to proteins and in proteins: sodium and potassium. *Acta Crystallogr D Biol Crystallogr* 58: 872–874.
41. Shabala L, Bowman J, Brown J, Ross T, McMeekin T, et al. (2009) Ion transport and osmotic adjustment in *Escherichia coli* in response to ionic and non-ionic osmotica. *Environ Microbiol* 11: 137–148.
42. Fraaije MW, Mattevi A (2000) Flavoenzymes: diverse catalysts with recurrent features. *Trends Biochem Sci* 25: 126–132.
43. Sumner JS, Matthews RG (1992) Stereochemistry and mechanism of hydrogen transfer between NADPH and methylenetetrahydrofolate in the reaction catalyzed by methylenetetrahydrofolate reductase from pig liver. *J Am Chem Soc* 114: 6949–6956.
44. Pejchal R, Sargeant R, Ludwig ML (2005) Structures of NADH and CH₃-H₄Folate Complexes of *Escherichia coli* Methylenetetrahydrofolate Reductase Reveal a Spartan Strategy for a Ping-Pong Reaction. *Biochemistry* 44: 11447–11457.
45. Taku A, Anwar RA (1973) Biosynthesis of uridine diphospho-N-acetylmuramic acid. IV. Activation of uridine diphospho-N-acetylenolpyruvylglucosamine reductase by monovalent cations. *J Biol Chem* 248: 4971–4976.
46. Mansour TS, Caufield CE, Rasmussen B, Chopra R, Krishnamurthy G, et al. (2007) Naphthyl Tetrone Acids as Multi-Target Inhibitors of Bacterial Peptidoglycan Biosynthesis. *Chemmedchem* 2: 1414–1417.

Figure S1

Figure S2

Figure S3

Table S1.

Crystal form	Atomic B factor of model atom at occ=1.0 ($\text{\AA}^2$)				Atomic B factor of ligand atom* ($\text{\AA}^2$)	Metal-ligand distance* ($\text{\AA}$)
	H ₂ O	Na ⁺	K ⁺	Ni ²⁺		
A (4 mol/asu)	2.3, 2.0, 2.0, 2.6	5.3, 3.8, 5.6, 6.1	20.0, 16.3, 19.7, 19.9	37.7, 33.1, 36.8, 36.7	O δ_{N57} = 22.4 O $_{\text{A237}}$ = 17.8 O $_{\text{S239}}$ = 19.1 O $_{\text{E335}}$ = 37.1 O $_{\text{7NADP}}$ = 17.3	O δ_{N57} = 2.7 O $_{\text{A237}}$ = 2.7 O $_{\text{S239}}$ = 2.7 O $_{\text{E335}}$ = 2.3 O $_{\text{7NADP}}$ = 3.0
B (1 mol/asu)	27.3	39.4	62.1	-	O δ_{N57} = 26.0 O $_{\text{A237}}$ = 29.3 O $_{\text{S239}}$ = 27.4 O $_{\text{E335}}$ = 31.6	O δ_{N57} = 2.8 O $_{\text{A237}}$ = 2.7 O $_{\text{S239}}$ = 2.9 O $_{\text{E335}}$ = 2.3

* averaged value of 4 NCS copies in the case of crystal form A.