

**Structural and functional
investigations of the *Pseudomonas*
aeruginosa YfiBNR-system**

Inauguraldissertation

zur

Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Philosophisch-Naturwissenschaftlichen Fakultät

der Universität Basel

von

Stefanie L. Kauer

aus Dürrenroth

Basel, 2017

Originaldokument gespeichert auf dem Dokumentenserver der Universität Basel

edoc.unibas.ch

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät auf
Antrag von

Prof. Dr. Tilman Schirmer

Prof. Dr. Urs Jenal

Basel, den 21.04.2015

Prof. Dr. Jörg Schibler

Dekan

Abstract

Most bacteria can exist in two fundamentally different life styles, as motile single cells and as sessile, surface-grown communities called biofilms. A key factor triggering the formation and maintenance of biofilms in a multitude of bacterial species is the second messenger bis-(3'-5')cyclic dimeric guanosine (c-di-GMP). While low intracellular levels of c-di-GMP promote planktonic behavior, where cells are generally motile and express virulence factors, increasing concentrations of c-di-GMP promote the expression of adhesive matrix components and result in multicellular behavior, biofilm formation and persistence.

The opportunistic pathogen *Pseudomonas aeruginosa* is responsible for chronic infections in the lungs of cystic fibrosis patients, a process that is eventually accompanied by the formation of small-colony variants (SCVs). The appearance of SCVs is caused by elevated levels of c-di-GMP and correlates with increased persistence of infection and antibiotic resistance. The YfiBNR-system has been previously identified as a key regulator of the SCV phenotype. While the mechanistic principles of interaction between the three proteins have been established by extensive in vivo studies, no structural information was available for any of the three proteins so far. Moreover, in vitro investigations of the effector of the system YfiN and its repressor YfiR were still missing.

In this thesis, the structure of YfiR was solved by X-ray crystallography, which revealed a dimeric assembly of the protein. Moreover, YfiR was shown to adopt a novel fold. The core of the protein is made up by a seven-stranded mixed β -sheet, flanked on the convex side by three helices, and embracing a long N-terminal α -helix with its concave side. The β -sheet

topology is 2-3-1-4-5-6-7, where strand 2 and 7 are aligned in an antiparallel fashion in regards to the rest of the sheet. The protein is stabilized by two intramolecular disulfide bonds, mediated by Cys71-110 as well as Cys145-Cys152. SEC-MALLS analysis of wildtype YfiR and a dimeric-interface mutant demonstrated that the protein is dimerizing with a high affinity in solution, via the same interface observed in the crystal structure. In a next step, a functional assay was established using membrane-bound YfiN, which allowed the investigation of YfiN activity and regulation in vitro. Production of c-di-GMP was observed in a time-dependent fashion, indicating that YfiN showed diguanylate cyclase (DGC) activity. Further evidence was provided that YfiN activity is negatively regulated by YfiR and c-di-GMP, which implied that YfiN is undergoing allosteric feedback inhibition. Based on these results, a homology model of the inactive state of YfiN was generated, which gave insight into the presumable mode of feedback inhibition, involving c-di-GMP mediated cross-linking of the GGDEF and HAMP domains. It is therefore proposed that YfiN activity is regulated either by repression upon binding of dimeric YfiR or by non-competitive product inhibition to avoid excessive substrate consumption.

Table of Contents

Abstract	iii
1 Introduction	1
1.1 <i>The bacterial second messenger c-di-GMP</i>	1
1.2 <i>The structure of c-di-GMP</i>	2
1.3 <i>Biosynthesis and degradation of c-di-GMP</i>	4
1.4 <i>Regulation of the enzymatic activity by signal input domains</i>	8
1.5 <i>C-di-GMP receptors</i>	13
1.6 <i>The YfiBNR-system from Pseudomonas aeruginosa</i>	15
1.7 <i>Aim of the thesis</i>	17
2 Material and Methods	19
2.1 <i>Cloning of constructs</i>	19
2.1.1 Cloning of YfiN	19
2.1.2 Cloning of the periplasmic YfiN constructs	19
2.1.3 Cloning of the cytoplasmic constructs from different YfiN homologs	20
2.2 <i>Protein production</i>	21
2.2.1 Protein expression	21
2.2.1.1 Expression tests of YfiN	21
2.2.1.2 Expression of YfiN	22
2.2.1.3 Expression tests of YfiN _{PAS}	22
2.2.1.4 Expression of MBP-tagged YfiN _{PAS38-159} and MBP-tagged YfiN _{PAS44-149}	23
2.2.1.5 Expression tests of YfiN _{HAMP-GGDEF} from <i>P. aeruginosa</i> , <i>P. fluorescens</i> , <i>Y. enterocoliticae</i> and <i>S. alaskensis</i>	24
2.2.1.6 Expression of YfiN _{HAMP-GGDEF} from <i>P. aeruginosa</i> , <i>P. fluorescens</i> , <i>Y. enterocoliticae</i> and <i>S. alaskensis</i>	24
2.2.2 Protein purification	25
2.2.2.1 Purification of YfiN	25
2.2.2.2 Purification of YfiN _{PAS} constructs	26
2.2.2.3 Refolding of the YfiN _{PAS44-149} construct	27
2.2.2.4 Purification of YfiN _{HAMP-GGDEF} from <i>P. aeruginosa</i> , <i>P. fluorescens</i> , <i>Y. enterocoliticae</i> and <i>S. alaskensis</i>	28
2.3 <i>Crystallization of YfiN_{HAMP-GGDEF} from P. aeruginosa and Y. enterocoliticae</i>	28
2.4 <i>Production of membranes expressing YfiN</i>	28
2.5 <i>Activity measurements</i>	29
2.5.1 FPLC-based nucleotide quantification	29
2.5.2 Phosphate sensor assay	30
2.5.3 Malachite green assay	30

2.5.4	Theoretical concepts and formulas used for the performed activity assays	31
2.5.4.1	The formula to calculate the concentration of a protein complex depending on the total protein and ligand concentrations and the Kd.....	31
2.5.4.2	General enzyme kinetics and derivation of the Michaelis-Menten Equation 33	
2.6	<i>Bioinformatical analysis</i>	36
2.6.1	Identification of YfiN homologs predicted to crystallize more readily.....	36
2.6.2	Identification and localization of the potential secondary I-site in the HAMP domain of YfiN.....	37
2.6.3	Non-competitive product inhibition model of YfiN _{HAMP-GGDEF}	37
3	Results	39
3.1	<i>Research article I (Kauer et al., in preparation)</i>	39
3.2	<i>In vitro characterization of YfiN, a key regulator of biofilm formation in Pseudomonas aeruginosa</i>	92
3.2.1	Functional characterization of membrane-bound and detergent-extracted YfiN 92	
3.2.1.1	Design and cloning of the YfiN constructs	92
3.2.1.2	Expression of the YfiN constructs	95
3.2.1.3	Optimization of the preparation protocol of membranes used for a YfiN activity assay	96
3.2.1.4	Quantification of DGC activity using membrane-bound YfiN	99
3.2.1.4.1	Non-competitive inhibition.....	99
3.2.1.4.2	The function used to quantify DGC activity using membrane-bound YfiN 100	
3.2.1.4.3	Experimental setup used to quantify DGC activity.....	102
3.2.1.5	Solubilization and purification of YfiN.....	107
3.2.1.6	Enzymatic characterization of detergent-solubilized YfiN using the malachite-green assay.....	112
3.2.1.7	Establishing the assay in a 96-well plate format using the well characterized DGC DgcZ	112
3.2.1.8	Using the established protocol to characterize the activity of detergent- solubilized YfiN.....	113
3.3	<i>Individual domains of YfiN</i>	120
3.3.1	The periplasmic PAS domain	120
3.3.1.1	Design and cloning of the YfiN _{PAS} constructs.....	120
3.3.1.2	Expression of the YfiN _{PAS} constructs	122
3.3.1.3	Purification of the YfiN _{PAS} constructs.....	124
3.3.2	The cytoplasmic HAMP and GGDEF domains.....	128
3.3.2.1	Design and cloning of the YfiN _{HAMP-GGDEF} constructs	128
3.3.2.2	Expression of the YfiN _{HAMP-GGDEF} constructs	131
3.3.2.3	Purification of the YfiN _{HAMP-GGDEF} constructs.....	132
3.3.2.4	Characterization of the degradation products by tryptic digest/MS.....	135
3.3.2.5	Initial crystallization attempts of Ps.aer. and Ye.ent. YfiN _{HAMP-GGDEF}	137
3.3.2.6	Functional characterization of Ps.aer. YfiN _{HAMP-GGDEF} by the phosphate sensor assay.....	138

3.3.2.6.1	Mechanistic principle of the phosphate sensor assay.....	138
3.3.2.6.2	Establishing the assay for DGCs using the well characterized DGC DgcZ as a reference protein	139
3.3.2.6.3	Enzymatic characterization of <i>Ps.aer.</i> YfiN _{HAMP-GGDEF}	140
3.4	<i>Bioinformatical analysis of YfiN: a model of the product-inhibited and the active state of YfiN</i>	141
3.5	<i>Discussion</i>	145
3.5.1	Activity assay of membrane-bound YfiN.....	145
3.5.2	Purification of detergent-solubilized YfiN.....	154
3.5.3	Activity assay of detergent-solubilized YfiN.....	156
3.5.4	Preliminary results of the individual domains of YfiN	157
3.5.4.1	The periplasmic PAS domain	157
3.5.4.2	The cytoplasmic HAMP and GGDEF domains	159
4	Conclusion and Outlook	161
5	Contrasting Mechanism of Inhibition of SHV5 and KPC2 β-Lactamases .	165
5.1	<i>Research article II (Kauer and Zähringer et al., in preparation)</i>	165
6	References	212
7	Acknowledgments	229

1 Introduction

1.1 The bacterial second messenger c-di-GMP

Most bacteria can exist in two fundamentally different life styles, either as free-swimming single cells or as sessile, multicellular communities forming biofilms. C-di-GMP has emerged as a regulatory mastermind to orchestrate the switch between the two states by influencing a multitude of downstream effector systems [1]. Whereas low levels of c-di-GMP promote a planktonic, motile lifestyle, increasing levels of c-di-GMP enhance the expression and production of adhesins [2],[3] exopolysaccharides [4]–[6] and attachment factors [7] and in addition curb various forms of motility, thereby leading to biofilm formation [4], [8], [9]. Moreover, c-di-GMP signaling also plays a role in bacterial virulence [10], [11] and persistence [12]–[14]. It has been shown that cells growing in biofilms are protected from phagocytosis [12], [15] and show enhanced resistance against antibiotics due to slow growth or increased persister cell populations [13], [16] compared to their free-living counterparts. Because of this tolerance against antibiotics and the host immune system, biofilm associated infections, such as chronic infections in the airways of cystic fibrosis patients caused by *Pseudomonas aeruginosa*, or fatal systemic infections after lung transplantations by *Burkholderia* species [17], are notoriously difficult to treat and thus represent a major health problem. As conventional antimicrobials cannot efficiently disperse biofilms, there is an urgent need to develop alternative measures in order to combat infections associated with biofilms. C-di-GMP mediated signaling could therefore pose a viable target for the control of biofilm formation.

1.2 The structure of c-di-GMP

C-di-GMP is a two-fold symmetrical molecule that consists of two GMP moieties forming a characteristic 12-member ribose-phosphate-ring in the center of the molecule (Figure 1-1). The first structure of an isolated cyclic dinucleotide was the one of c-di-dAMP determined by nuclear magnetic resonance (NMR) [18] and X-ray crystallography [19]. These c-di-dAMP structures were followed by crystal structure analyses of c-di-GMP [20]–[22]. They all show similar conformations for the ribosyl and phosphate moieties, suggesting a rather rigid backbone, with only the torsion angle around the glycosidic bond, and thus the position of the guanylyl base, showing conformational freedom.

C-di-GMP was shown to form a homodimer with intercalated bases and two isologous intermolecular base-phosphate hydrogen bonds (Figure 1-1 (B)). In some structures, Mg^{2+} or Co^{2+} ions are coordinated by the N-7 atoms of the two central bases, however they do not alter the structure significantly. Binding of dimeric c-di-GMP has been observed at the inhibition site of diguanylate cyclases [23]–[26], PilZ receptors [27]–[30] a response regulator [31], a riboswitch [32] as well as the active site of a diguanylate cyclase [33]. Monomeric c-di-GMP was observed in the active sites of phosphodiesterases [34]–[36], a PilZ receptor [37] in the active site of a diguanylate cyclase [23] and in the eukaryotic innate immune sensor STING [38]. In a recent study, binding of a c-di-GMP tetramer was observed to be triggering effective dimerization of the transcriptional factor BldD [39]. It is therefore not surprising that c-di-GMP has been reported to display a rich polymorphism of various oligomeric forms at concentrations higher than 1 mM [40], [41].

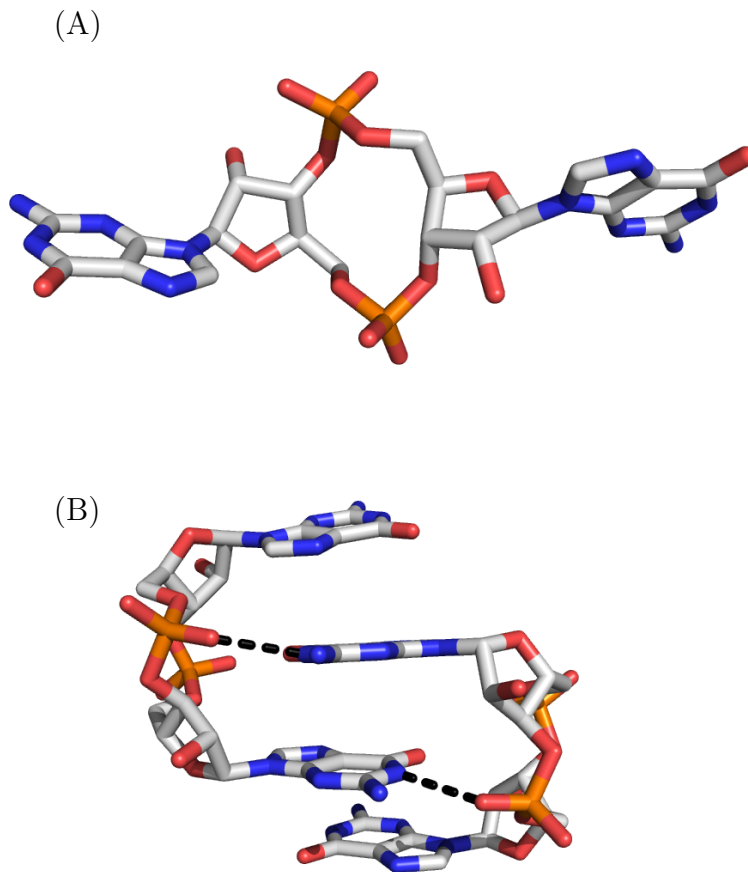


Figure 1-1: Three-dimensional structures of cyclic-di-GMP. Carbon atoms are shown in white, nitrogen in blue, oxygen in red and phosphorus in orange. **(A)** The monomeric form of c-di-GMP (from Protein Data Bank (PDB) entry 4LJ3). This state is usually binding to the active site of EAL domains. **(B)** c-di-GMP forming an intercalated dimer (from PDB entry 1w25). This form is usually seen bound to the allosteric I-site of diguanylate cyclases and PilZ domains.

In a recent NMR study, it was found that the association of the monomer to the dimeric form is in fast exchange (<milliseconds) with an equilibrium constant of about 1 mM [42]. Higher oligomers were found in the presence of cations such as K^+ at concentrations above 100 μM . Thus, at the low micromolar concentrations of c-di-GMP present in the cell and in the

absence of additional compounds that stabilize oligomers, c-di-GMP should be predominantly monomeric.

1.3 Biosynthesis and degradation of c-di-GMP

c-di-GMP is produced from two GTP molecules by diguanylate cyclases (DGC) and degraded to 5'-phosphoguanylyl-(3'-5')-guanosine (pGpG) by phosphodiesterases (PDE) (Figure 1-2).

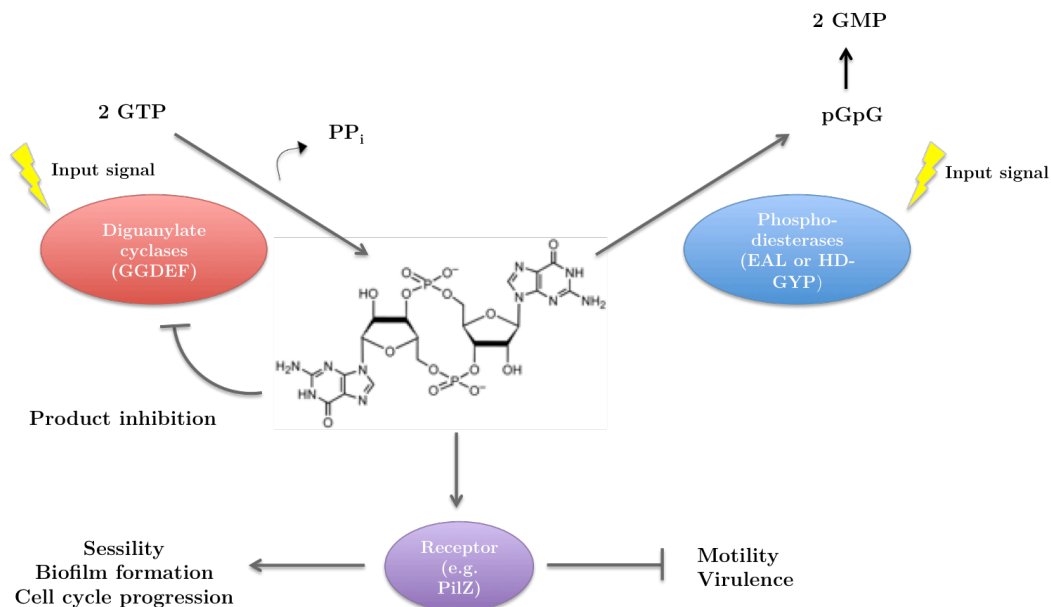


Figure 1-2: C-di-GMP signaling pathways. In the cell, c-di-GMP is generated by diguanylate cyclases (red) that carry a catalytic GGDEF domain, and is degraded by phosphodiesterases (blue) that carry either a catalytic EAL or HD-GYP domain respectively. These enzymes are regulated by internal or external input signals that are sensed by their N-terminal accessory domains. Diguanylate cyclases are subjected to non-competitive product inhibition via c-di-GMP. The interaction of c-di-GMP with downstream receptors produce an output in a range of cellular processes and functions, such as downregulation of cell motility and acute virulence on one hand and on the other stimulation of diverse biofilm-associated functions.

The catalytically active part of the DGC is the GGDEF domain that is named after the amino acid sequence motif, which constitutes an essential

part of the active site of the enzyme [43], [44]. The GGDEF domain consists of a five-stranded central β -sheet surrounded by five helices [23]. The displayed fold is similar to that of class III nucleotide cyclases and type I DNA polymerases, which implies a similar catalytic mechanism using magnesium as metal ions [43], [45], [46].

The GGDEF-sequence motif is forming a β -hairpin, where the first two residues (Gly) were shown to be important for GTP binding, while the fourth residue (Glu) is involved in metal ion coordination. The third amino acid (Asp/Glu) is crucial for catalysis and additionally plays a role in metal coordination [23], [24]. Any point mutation in this motif (except for a D to E mutation) abolishes enzymatic activity [44].

The active DGC is a dimer of two GGDEF domains, where both active sites are located at the dimer interface [23], [24]. This allows an antiparallel alignment of two GTP molecules and the formation of two intermolecular phosphodiester bonds, resulting in the production of c-di-GMP.

In order to prevent excessive substrate consumption and set an upper limit for product accumulation at the same time, most DGCs are subjected to allosteric product inhibition with an inhibition constant in the range of cellular c-di-GMP concentrations [23], [47]. This involves binding of a base-intercalated dimer of c-di-GMP molecules to a primary inhibition site (Ip) with an RxxD sequence motif (in which x denotes any amino acid) and binding to a secondary inhibition site (Is), which can come either from a regulatory domain or from the GGDEF domain of another protein monomer [23]–[25]. This allows the intercalated c-di-GMP dimer to block the movement of the GGDEF domain, which is required for formation of the

catalytically active homodimer. Despite primary sequence proximity between the A and I sites, they are located antipodal to each other, separated by a linker composed of five amino acids.

C-di-GMP is degraded to the linear dinucleotide 5'-phosphoguanylyl-(3'-5')-guanosine (pGpG) by either EAL or HD-GYP domains, named after their conserved active site residues [48], [49]. EAL-containing PDEs show high substrate specificity with a low sub-micromolar K_m [48], [50], [51] which is consistent with the low cellular concentration of c-di-GMP. These proteins require the presence of either Mg^{2+} or Mn^{2+} to function but are strongly inhibited by Ca^{2+} and Zn^{2+} [48], [49]. EAL domains exhibit a typical TIM-barrel fold, where the active site is located at the C-terminal part of the barrel. The glutamate of the EAL-sequence motif is involved in cation coordination [35].

The second class of c-di-GMP-specific PDE is made up by HD-GYP-domain proteins, which form a subfamily of the metal-dependent phosphohydrolases and are unrelated to EAL proteins [52]. In contrast to EAL proteins, these proteins break the phosphodiester bond in c-di-GMP to produce 5'-pGpG and can then further degrade it to GMP.

One interesting class of the c-di-GMP processing proteins is the so-called “composite proteins” composed of GGDEF and either one of the PDE domains, EAL or HD-GYP. The sheer number of proteins containing GGDEF-EAL tandems is large, as many as $\sim 1/3$ of all GGDEF domains and as many as $\sim 2/3$ of all EAL domains are present on the same polypeptide chain [53]. Only a few of them show bifunctionality, with both domains being active [54], [55]. One of them is the protein called ScrC,

which is responsible for the switch between the motile swarmer state and biofilm formation in *Vibrio parahaemolyticus* [56]. ScrC is composed of an N-terminal periplasmic sensor domain, followed by a GGDEF-EAL module. When expressed alone, ScrC is acting as a DGC leading to elevated intracellular c-di-GMP concentrations. In presence of its interaction partners ScrA and ScrB however, the protein is displaying PDE activity [56].

By far more common is the class of composite proteins, where one of the two domains is enzymatically inactive or catalytically incompetent [48], [50], [57], [58]. These inactive domains have evolved to carry out new functions such as binding but not processing of a substrate molecule as it was observed for an inactive GGDEF domain [48] or inactive EAL domains [59]–[61]. In the case of PdeA from *Caulobacter crescentus*, the degenerate GEDEF domain no longer shows catalytic activity but it is still able to bind GTP with a high affinity ($K_d = 4 \mu\text{M}$). This in turns stimulates the PDE activity by bringing the K_m for c-di-GMP from the physiologically irrelevant level of $\sim 100 \mu\text{M}$ to the physiologically relevant level of $0.42 \mu\text{M}$ [48]. This provides thus an example of a c-di-GMP signaling domain that has evolved beyond its role as an enzyme.

1.4 Regulation of the enzymatic activity by signal input domains

A vast majority of proteins that either generate or degrade c-di-GMP harbor N-terminal accessory domains, which are regulating the enzymatic activity of the GGDEF, EAL or HD-GYP output domains in response to various stimuli. This suggests that several environmental and cellular signals are part of the c-di-GMP signaling network [62], [63]. Typical examples of sensory domains are small molecule- and protein binding PAS and GAF domains, light sensing BLUF domains, HAMP domains, phosphoryl group accepting receiver domains (REC) and many others [53]. A majority of those domains, e.g. PAS, HAMP and REC domains, can be found in other types of sensory and signaling proteins such as the two-component histidine kinases (HK) or the methyl-accepting chemotaxis proteins (MCP).

PAS (Per-Arnt-Sim) domains are usually about 100 amino in length and are found connected to a wide range of enzymatic or nonenzymatic effector domains, such as HK, GGDEF and MCP. As these effectors are involved in diverse cellular pathways, PAS domains perform a variety of functions, such as promotion of protein-protein interactions [64], [65], signal-transfer [66] as well as sensing perceived stimuli in a direct way [67]. PAS domains have developed a remarkable plasticity in binding different ligands and cofactors over the course of evolution. The binding of small ligands to the PAS domain can either serve as a direct signal [68], or act as cofactors that are capable of sensing redox potential [69], [70] light [71], [72] or dissolved gases [73].

On the structural level PAS domains share a common three-dimensional fold [74], although they show little primary sequence identity. Two structural clades can be defined, depending on the cellular localization of the PAS domains, either in the cytoplasm or outside of the cell in the periplasm. Slight differences have been observed between the structures of cytoplasmic and periplasmic PAS domains: even though both types are characterized by a conserved β -sheet core, extracellular PAS domains are often anchored to the membrane by a long N-terminal α -helix, and most of them show a reduced set of helices between the second and the third β -strand (Figure 1-3) [74]–[76]. The term PDC (for Pho/DcuS/DctB/CitA) domain was therefore introduced by [77] for this type of PAS domain topology.

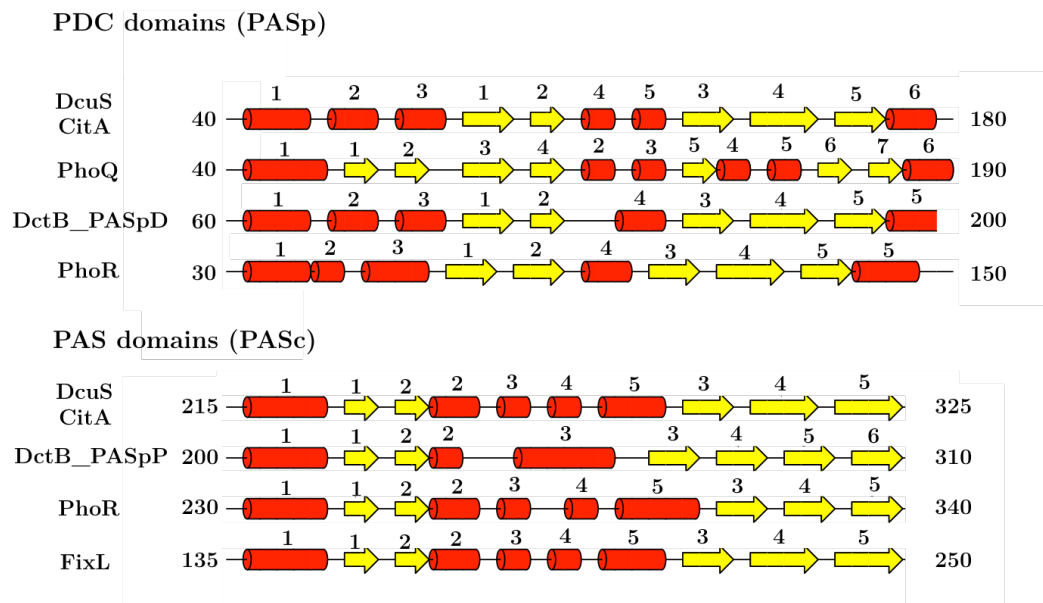


Figure 1-3: α -helical and β -sheet structures of periplasmic PASp or PDC domains in comparison to cytoplasmic PASc domains. The PDC is characterized by an „ $\alpha\alpha\alpha\beta\alpha(\alpha)\beta\beta\beta\alpha$ “ arrangement of secondary structure elements, whereas the PAS fold shows an „ $\alpha\beta\beta\alpha\alpha\alpha\beta\beta\beta$ “ arrangement. PASp and PASc stand for periplasmic and cytoplasmic PAS domains, respectively. PASpD and PASpP for distal and peripheral PAS domains. Topological arrangements refer to structures of DcuS (*Escherichia coli*), CitA (*Klebsiella pneumoniae*), PhoQ (*Salmonella typhimurium*), DctB (*Sinorhizobium meliloti*), PhoR (*Bacillus subtilis*) and FixL (*Sinorhizobium meliloti*). The figure was produced using TopDraw [159].

HAMP domains, named after their occurrence in histidine kinases, adenylyl cyclases, methyl-accepting chemotaxis proteins and some phosphatases [78], form a second type of sensory input domains. They are also found coupled to other effector domains such as GGDEF [12], [14] and GGDEF-EAL [79]. In membrane-associated signaling proteins, HAMP domains play a key mechanistic role in transmembrane signaling. They usually lie near the cytoplasmic side of the membrane, between the periplasmic sensing and cytoplasmic signaling domains of the protein, where they promote conversion of signal-induced conformational changes into behavior-controlling output signals [78], [80].

The first HAMP structure was solved by NMR, occurring in the putative protein Af1503 from *Archaeoglobus fulgidus* [81]. The structure revealed that the HAMP domain is organized as a symmetric, homodimeric parallel coiled coil. Each monomer is composed of two α -helices, AS1 and AS2, connected by a flexible linker segment of 14-15 residues in length. The same topology has been seen in many HAMP crystal structures that followed thereafter [82] [83] [84]. Coiled coils are bundles of helices that are building up a superhelix. Supercoiling results from the characteristic packing of side chains at the interface of the helices, called knobs-into-holes packing [85]. This packing mode is characterized by seven-residue sequence repeats, whose positions are labeled a–g, where residues a and d point towards the core of the bundle. The experimentally-determined Af1503 HAMP structure exhibits a “complementary x-da” packing arrangement of the helices in the bundle [81]. The two bundle structures are related by a coordinated 26° counter-rotation of each of the four helices, analogous to meshed gears in a transmission (therefore termed the “gearbox-model”). In the same study it

was suggested that the more conventional “a-d” packing arrangement might represent an alternative HAMP signaling conformation. Indeed, structures of Af1503 HAMP domain-dimerization/histidine phosphorylation (DHp) domain fusions show that the rotation of the HAMP helices results in rotation of adjacent helices of DHp [86] [87] and that this mechanism might explain the signal transduction mediated by HAMP domains in receptor histidine kinases.

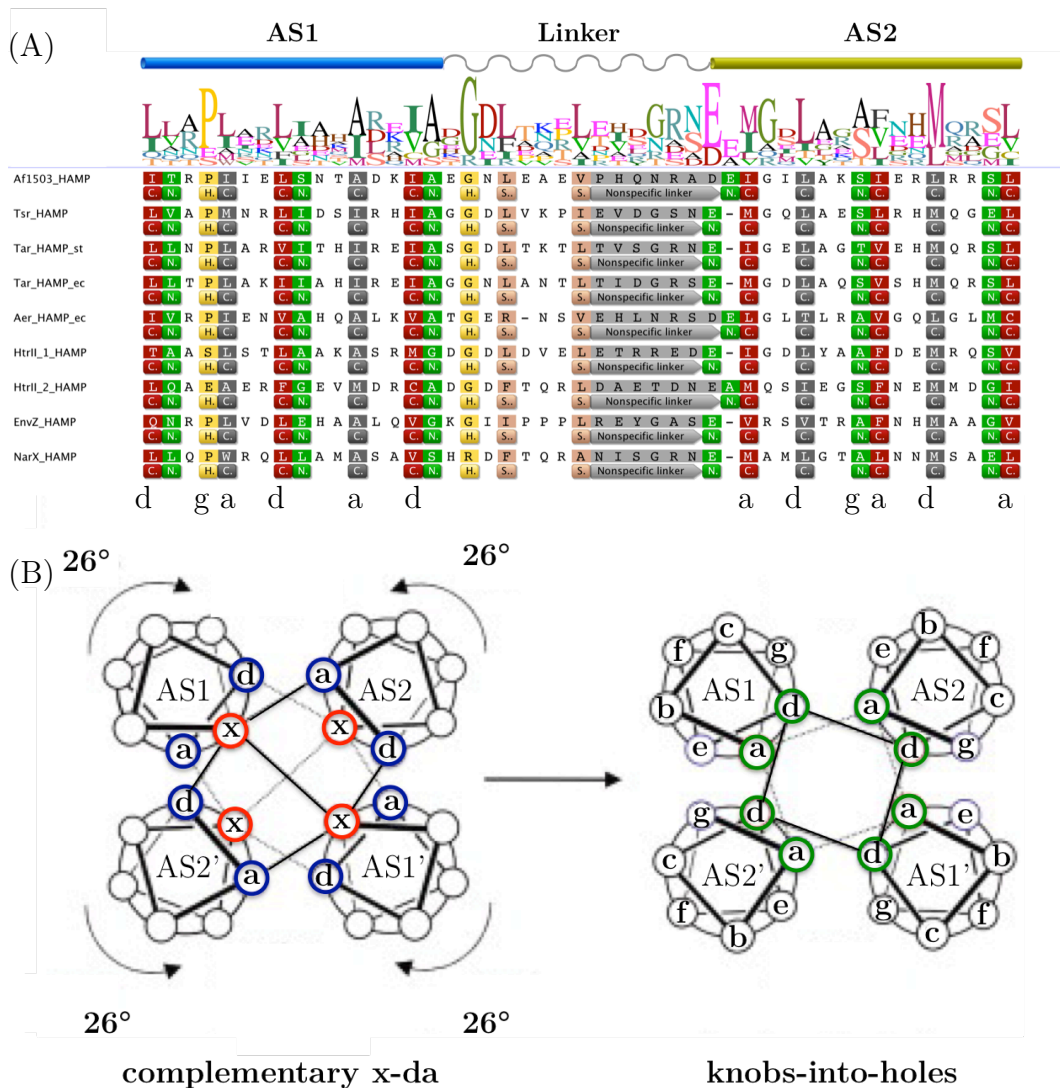


Figure 1-4: HAMP sequence features and bundle-packing arrangement. (A) Multiple sequence alignment of well characterized HAMP domains. Af1503 is from *Archaeoglobus fulgidus*; Tsr, Tar, Aer, EnvZ, and NarX are from *Escherichia coli*; Tar_st is the Tar protein of *Salmonella typhimurium*; HtrII-1 and HtrII-2 are two HAMP domains in HtrII from *Natronomonas pharaonis*. Secondary structure elements are annotated above the alignment. Residues shaded in red and grey are important for helical packing, residues shaded in green are noncritical for packing. Residues colored in yellow are highly conserved HAMP residues, whereas residues shaded in salmon are important for bundle stability. The nonspecific part of the linker region was colored in grey. (B) Schematic view of the gearbox model. Transformation from complementary x-da packing to knobs-into-holes packing is achieved by a coordinated axial rotation of all four helices by $\sim 26^\circ$.

For c-di-GMP signaling proteins, only a handful have been characterized in terms of ligand specificity and how the sensory domain conveys the conformational change triggered by the binding event to the downstream effector domain. The DGC DocC and the PDE DoSP from *E. coli* as well as the PDE AxPDE1 from *Acetobacter xylinum* have been shown to sense oxygen [73], [88]. The tandem GGDEF-EAL protein AxDGC2 also from *A. xylinum*, containing a catalytically active GGDEF domain, is responding to altered redox conditions via a non-covalently bound FAD cofactor [69]. Two phosphodiesterases, BlrP1 from *K. pneumoniae* and SL2 from *Synechococcus elongatus*, have been shown to react to light via their associated BLUF or LOV domains [35], [89]. Based on the crystal structures of BlrP1, the activation mechanism was proposed, where light absorption by the flavin molecule causes conformational changes in the BLUF domain, which are in turn propagated to the EAL active site.

Amongst the most well characterized proteins are the DGCs PleD from *C. crescentus* and WspR from *P. aeruginosa*, which both carry a N-terminal REC domain and become activated upon phosphorylation [90], [44]. Phosphorylation of the aspartate localized on the REC1 domain of PleD induces structural rearrangement in the REC1-REC2 interfaces, which in turn, allows the tight dimerization of PleD, a prerequisite for c-di-GMP production. For PleD as well as WspR, the input signals for their cognate histidine kinases remain yet to be identified.

1.5 C-di-GMP receptors

To exert its function, c-di-GMP has to bind to a downstream receptor and trigger structural or functional changes. C-di-GMP specific effectors are highly diverse, accounting for the broad range of cellular functions and processes that are controlled by c-di-GMP.

The best-studied class is the PilZ family of proteins. The name originated from the *P. aeruginosa* PilZ (PA2960) protein, which consists of the c-di-GMP binding domain only. Structural studies revealed the presence of two short stretches with a RxxxRx₂₀₋₃₀(D/N)x(S/A)xxG sequence motif, which are involved in c-di-GMP binding [28]–[30], [37], [91]. This event then triggers dramatic structural changes, where the primary binding loop wraps around the c-di-GMP molecule, bringing the second stretch into closer proximity. High c-di-GMP affinity has been observed for PilZ domain proteins, with dissociation constants being in the sub-micromolar range [5], [27], [92], [93] consistent with the cellular concentration of c-di-GMP. Interestingly, c-di-GMP can bind to the domain either as an intercalated dimer [27], [28], [30], or as a monomer [37]. It has been established that PilZ domains can occur as a standalone module or come as an attachment to other domains that generate a molecular output such as the production of cellulose [94] or alginate [5], changes in motility levels [93], [92], [95] and virulence of pathogens [92], [95].

Another class of c-di-GMP receptors is the degenerate GGDEF or EAL proteins, which lost their catalytic activity but retained the ability of c-di-GMP binding, either to the inhibiting I-site or to the degenerate EAL motif. One of these catalytically inactive GGDEF proteins that functions as a c-di-GMP receptor is the response regulator PopA from *Caulobacter crescentus*,

which recruits the cell cycle regulator CtrA to the cell pole after c-di-GMP binding to the I-site, thereby targeting CtrA for degradation [96]. Another protein called LapD is a GGDEF-EAL c-di-GMP receptor from *Pseudomonas fluorescens*, where both catalytic domains lost their activity but c-di-GMP is binding to its degenerate EAL domain with high affinity [3], [97], [98]. This binding event triggers a conformational change in LapD, leading to the recruitment of the periplasmic protease LapG, preventing it from cleaving a cell surface-bound adhesion and thereby promoting cell adhesion.

A third type of a c-di-GMP binding domain termed “GIL” has only very recently been discovered [99]. It shows an “RxGD” binding motif, which is comparable to the RxxD motif involved in c-di-GMP binding in the I-sites of the GGDEF domains. Furthermore, the region in the vicinity of the apparent c-di-GMP-binding site on the GIL domain has the same predicted secondary structure as the corresponding part of the GGDEF domains [99].

Yet another type of c-di-GMP binders comprises proteins that can not be classified or predicted as receptors for c-di-GMP as they do not have amino acid motifs that resemble previously defined c-di-GMP binding domains. FleQ is a transcription factor in *Pseudomonas aeruginosa*, involved in flagellar gene expression and exopolysachharide synthesis [100]. FleQ has been shown to bind to pel promoter DNA, leading to repression of pel transcription and that this repression is relieved by binding of c-di-GMP to FleQ.

Ultimately, riboswitches have been shown to bind to c-di-GMP via a conserved GEMM domain and serve a direct role in gene expression [101]. They generally exhibit a very high affinity for c-di-GMP with a Kd around

1 nM, which is comparable to the highest RNA:small molecule interaction known [102].

1.6 The YfiBNR-system from *Pseudomonas aeruginosa*

The opportunistic pathogen *Pseudomonas aeruginosa* is responsible for chronic infections in the lungs of patients suffering from cystic fibrosis (CF) or patients with compromised host defense mechanisms [103]. Eventually, respiratory failure caused by the infection is often the ultimate cause of death. During long-term lung colonization, *P. aeruginosa* undergoes genetic and phenotypic adaptation [104], [105], resulting in reduced virulence [106] as well as increased persistence [107]. Phenotypic adaptations lead to the appearance of small colony variants (SCV) in the sputum of CF patients, which are characterized by high auto-aggregation, attachment to surfaces, slow growth and enhanced exopolysaccharide production [108]. These persistent variants of *P. aeruginosa* constitute adaptations to the host environment and correlate with antibiotic resistance [13], [16], resistance against phagocytosis [12] and increased persister cell populations [16]. These findings suggest that the SCV phenotype may provide fitness advantages during the process of chronic lung infections and therefore plays an important role in the pathogenesis of *P. aeruginosa*.

It has been shown that the appearance of the SCV phenotype is strongly linked to elevated levels of c-di-GMP [12], [109], [110]. A genetic screen in the PA01 strain for SCV-related loci, identified the YfiBNR operon, coding for the three proteins YfiB, YfiR and YfiN, which are responsible for

regulating *c*-di-GMP levels in *P. aeruginosa* in response to a still unknown environmental signal [12] [14].

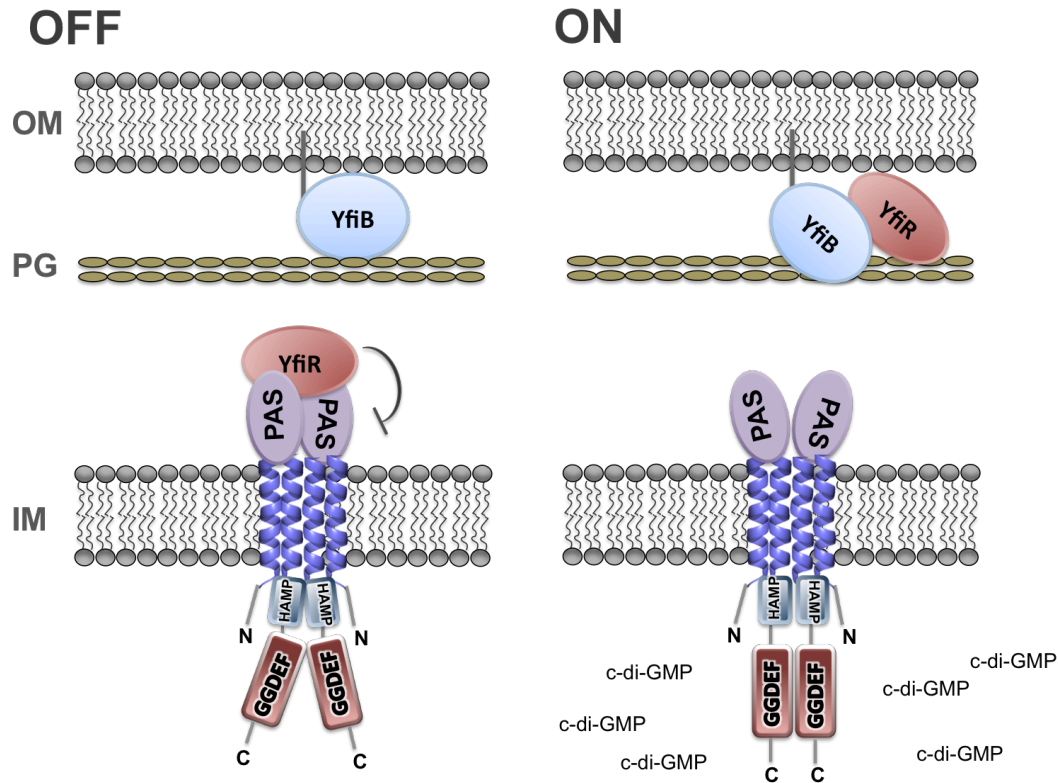


Figure 1-5: Schematic representation of the YfiB/NR system from *P. aeruginosa*. OM stands for “outer membrane”, PG for peptidoglycan and IM for “inner membrane”. YfiN is an inner-membrane protein, which is functioning as a DGC. It is repressed by interacting with YfiR, most presumably with its periplasmic PAS domain. YfiB, the outer-membrane bound OmpA/Pal-like protein activates YfiN by sequestering YfiR to the outer-membrane upon a still uncharacterized signal. YfiB also possesses a peptidoglycan-binding site, which has been shown to be crucial for YfiB function [14].

YfiN thereby constitutes the effector of the system functioning as a DGC, which is located in the inner-membrane with its N- and C-termini facing the cytosol. It is composed of three domains: a periplasmic PAS (or PDC) domain, two transmembrane helices and the cytosolic part of the protein, which encompasses a HAMP domain followed by the catalytic GGDEF domain. In order to produce *c*-di-GMP, YfiN has to dimerize and adopt a catalytically competent arrangement of the GGDEF domains. YfiR is the

second member of the system and has been shown to reside exclusively in the periplasm [12]. By analyzing the phenotypes of a series of mutations in vivo, YfiR was found to be repressing YfiN function allosterically by directly binding to the protein. In addition, YfiR was suggested to sense the redox status of the periplasm via its two highly conserved disulfide bonds, providing one way of controlling the activity of YfiR [14]. Another control mechanism involves the third component of the system YfiB, an OmpA/Pal-like outer-membrane lipoprotein with a peptidoglycan-binding site. YfiB has been shown to sequester YfiR to the outer-membrane, thereby relieving YfiN repression, thus leading to c-di-GMP synthesis. This in turn will lead to increased production of Pel and Psl exopolysachharides, which are characteristic traits of the SCV phenotype [12]. Taken all these findings into account, this establishes the YfiBNR-system as a key regulator of biofilm formation in *P. aeruginosa* and thereby the persistence of chronic infections.

1.7 Aim of the thesis

The opportunistic pathogen *Pseudomonas aeruginosa* is responsible for chronic infections in the lungs of cystic fibrosis patients. The long-term survival strategy of *P. aeruginosa* in the host environment is based on phenotypic and genetic adaptations, resulting in the progressive loss of virulence and leading to increased persistence. This process is accompanied by adaptations of colony morphology, leading to the appearance of small-colony variants (SCVs), which are characterized by auto-aggregation and enhanced exopolysaccharide production.

The YfiBNR system has been previously identified to be a key regulator of the formation of small-colony variant (SCVs) in *Pseudomonas aeruginosa*.

Major work has been done on establishing an interaction model between the three members of the system by a combination of genetic and biochemical analysis. The YfiBNR system has been shown to function via tightly controlled competition between allosteric binding sites, which are highly conserved and hydrophobic and found on all three Yfi proteins. YfiN activity was shown to be suppressed in vivo by interacting with the periplasmic protein YfiR. Release of repression is mediated by the outer-membrane (OM) lipoprotein YfiB, which upon undergoing a conformational change is sequestering YfiR to the OM. Despite the fact that the epistasis and the mechanistic principles of interaction have been determined, no structural information was available on any of the three Yfi proteins.

The aim of this thesis is therefore to investigate the interaction between the effector of the system, YfiN, and its repressor, YfiR, in more detail on a structural and functional level. Structures of the individual proteins, and potentially of the complex, would give detailed information about the residues, which are relevant for binding and allow further investigation of the allosteric binding sites. As YfiN has to dimerize in order to be active but no data is available on the oligomeric state of YfiR, in vitro characterization has to be performed in order to get information about the stoichiometry of the YfiR-YfiN complex.

On the functional side, the establishment of an assay would allow the quantification of YfiN activity in vitro. The same assay could then be used in a second step to check if YfiR-mediated repression of YfiN activity is also observed in vitro and if YfiN is undergoing non-competitive product

inhibition as it was observed for other DGCs. Taken all these investigations together, this should help to provide missing information and thereby contribute to a more complete understanding of the YfiBNR system.

2 Material and Methods

2.1 Cloning of constructs

2.1.1 Cloning of YfiN

A synthetic DNA fragment coding for YfiN from *Pseudomonas aeruginosa* was ordered from MrGene (Life Technologies) and optimized in regard to codon-usage and GC-content. This fragment was then cloned into the expression plasmid pET28a (Novagen) using restriction sites NcoI/XhoI, yielding a C-terminally hexahistidine-tagged protein. In order to obtain a N-terminally his-tagged protein, the his-tag was introduced to the protein sequence in a first step by a set of primers. In a second step, the construct was then cloned into pET28a using the NcoI/XhoI restriction sites but with a primer bearing a stop-codon after the XhoI sequence.

2.1.2 Cloning of the periplasmic YfiN constructs

Three YfiN_{PAS} constructs of different lengths have been generated by standard polymerase chain reaction using the pET28a/YfiN vector as a template. YfiN_{PAS38-159} and YfiN_{PAS44-149} were cloned into pET28a (Novagen), resulting in a C-terminally his-tagged protein, in pET28a-SUMO (modified pET28a-vector) generating an N-terminally SUMO-tagged protein, into

pMal-c2X (NEB) yielding an N-terminally MBP-tagged protein for expression in the cytoplasm, into pMal-p5x (NEB) yielding an N-terminally MBP-tagged protein for expression in the periplasm, and into pET-GSTx (pET21a, modified with the GST-coding region of pGEX-KT) generating N-terminally GST-tagged protein. All fusion-constructs were either untagged at the C-terminus or carried a his-tag. The N-terminal fusion tags SUMO, MBP and GST were cleavable by the SUMO-protease and Factor Xa respectively.

The third construct YfiN_{PAS44-154} was cloned into pET28a, generating a C-terminally his-tagged protein.

2.1.3 Cloning of the cytoplasmic constructs from different YfiN homologs

The coding region corresponding to YfiN_{HAMP-GGDEF} (residues 178-435) from *Pseudomonas aeruginosa*, was amplified by standard polymerase chain reaction using the pET28a/YfiN vector as a template, leading to a C-terminally his-tagged protein. Synthetic DNA fragments coding for YfiN_{HAMP-GGDEF} from *Pseudomonas fluorescens* (residues 174-420), *Yersinia enterocoliticae* (residues 188-448) and *Sphingopyxis alaskensis* (residues 179-408) were ordered from MrGene (Life Technologies) and optimized in regard to codon-usage and GC-content. The YfiN homologs were then cloned in the expression plasmid pET28a (Novagen) using restriction sites NcoI/XhoI, yielding C-terminally hexahistidine-tagged proteins.

2.2 Protein production

2.2.1 Protein expression

2.2.1.1 Expression tests of YfiN

The pET28a/YfiN plasmid was transformed into the *E. coli* strain BL21(DE3) (Novagen). 50 mL LB-medium supplemented with antibiotics was inoculated with overnight culture to a starting OD₆₀₀ of 0.1. The cells were grown at 37 to an OD₆₀₀ of 0.6, at which the protein expression was induced with 200 μM IPTG (isopropyl-β-D-1-thiogalactopyranoside). The temperature was then lowered to either 30 °C or 20 °C. For expression in TB-medium, a LB-starter culture was grown overnight, supplemented with antibiotics and 1% glucose to prevent leaky expression. The next day, cells were harvested by centrifugation at 5000 g and 4 °C for 15 minutes and the cell pellet was resuspended in the equivalent amount of TB-medium as the volume of the initial starter culture, supplemented with antibiotics. After a 1h incubation at 37 °C, protein expression was induced by the addition of 50 μM IPTG and the temperature was dropped to either 30 °C or 20 °C. Before and at several time points after induction samples were taken, which were centrifuged (10000g, 10min, room temperature) and the cell pellet was resuspended in lysis buffer (50 mM Tris-HCl, pH 8, 200 mM NaCl,) where the volume of buffer used corresponded to the cell density of the sample. The cells were lysed according to the BugBuster procedure (Novagen) and centrifuged (16'000g, 20 min, 4 °C) to yield a fraction containing insoluble proteins and cell debris. The supernatant containing the soluble protein fraction was loaded on an SDS-PAGE to follow YfiN expression.

2.2.1.2 Expression of YfiN

YfiN was produced using the *E. coli* BL21(DE3) strain transformed with the pET28a/YfiN plasmid. Cells were grown at 37 °C overnight in LB-medium supplemented with kanamycin (50 µg/mL) and 1% glucose to prevent protein expression. The next day, cells were harvested by centrifugation at 5000 g and 4 °C for 15 minutes and the cell pellet was resuspended in the equivalent amount of TB-medium as the volume of the initial starter culture, supplemented with kanamycin (50 µg/mL). After growing the shaking cultures at 37 °C for 1 hour, 50 µM IPTG was used to induce protein expression. The temperature was decreased to 30 °C and after 4 h of incubation, the cells were harvested by centrifugation (5000 g, 10 min, 4 °C) and the pellets were frozen at -20 °C.

2.2.1.3 Expression tests of YfiN_{PAS}

The different YfiN_{PAS} constructs were transformed into the *E. coli* strains BL21(DE3) (Novagen), BL21 StarTM (DE3) (Life Technologies), BL21-AITM (Life Technologies) and ArcticExpressTM (DE3)RIL (Agilent Technologies). For the BL21(DE3) and BL21 StarTM (DE3) strains, 50 mL LB-medium supplemented with antibiotics was inoculated with overnight culture to a starting OD₆₀₀ of 0.1. The cells were grown at 37 °C to an OD₆₀₀ of 0.6, at which the protein expression was induced with 100 µM, 250 µM or 500 µM IPTG respectively. For the expression test at lower temperatures, cells were grown at 37°C until an OD₆₀₀ of 0.5 was reached before temperature was reduced to 30 °C or 20 °C.

For the expression tests using the ArcticExpress *E. coli* strain, cells were cultivated at 37 °C to an OD₆₀₀= 0.8 and protein induction was initiated by the addition of 250 μM IPTG.

For the expression tests using the BL21-AITM strain, the cells were treated as described for BL21(DE3) and BL21 StarTM (DE3), except that the protein induction was induced by the addition of either 0.2 % or 1 % arabinose and 500 μM IPTG.

For the expression tests using TB-medium, the same protocol was applied as described in 2.2.1.1 for YfiN.

In all cases, samples were taken right before and at several time points after induction and treated as described in 2.2.1.1 to monitor protein expression levels.

2.2.1.4 Expression of MBP-tagged YfiN_{PAS38-159} and MBP-tagged YfiN_{PAS44-149}

Both constructs were produced using the *E. coli* BL21(DE3) strain transformed with the pMal-c2X/YfiN_{PAS44-149} or the pMal-c2X/YfiN_{PAS38-159} plasmids respectively. Cells were grown at 37 °C overnight in 1 L LB-medium supplemented with kanamycin (50 μg/mL) and 1% glucose to prevent protein expression. The next day, cells were harvested by centrifugation at 5000 g and 4 °C for 15 minutes and the cell pellet was resuspended in 1 L of TB-medium, supplemented with kanamycin (50 μl/mL). After growing the shaking cultures at 37 °C for 30 min, 50 μM IPTG was used to induce protein expression. For YfiN_{PAS38-159}, the temperature was decreased to 30 °C and protein expression was allowed to take place for 4 h. For YfiN_{PAS44-149}, the temperature was decreased to

20°C and protein expression was taking place overnight. Ultimately, the cells were harvested by centrifugation (5000 g, 10 min, 4°C) and stored at -20 °C for further use.

2.2.1.5 Expression tests of YfiN_{HAMP-GGDEF} from *P. aeruginosa*, *P. fluorescens*, *Y. enterocoliticae* and *S. alaskensis*

The different YfiN_{HAMP-GGDEF} constructs were transformed into the *E. coli* strains BL21(DE3) (Novagen), BL21 StarTM (DE3) (Life Technologies). 50 mL LB-medium supplemented with antibiotics was inoculated with overnight culture to a starting OD₆₀₀ of 0.1. The cells were grown at 37 °C to an OD₆₀₀ of 0.7, at which the protein expression was induced with 250 µM IPTG. The temperature was then either kept at 37 °C or lowered to 30 °C and 20 °C. In all cases, samples were taken right before and at several time points after induction and treated as described in 2.2.1.1 to monitor protein expression levels

2.2.1.6 Expression of YfiN_{HAMP-GGDEF} from *P. aeruginosa*, *P. fluorescens*, *Y. enterocoliticae* and *S. alaskensis*

The different YfiN_{HAMP-GGDEF} constructs were produced using the *E. coli* BL21(DE3) strain transformed with the pET28a/YfiN_{HAMP-GGDEF} plasmids. 1 L LB-medium supplemented with antibiotics was inoculated with 1 % v/v overnight culture. The cells were grown at 37 °C to an OD₆₀₀ of 0.6, at which the protein expression was induced with 250 µM IPTG. The temperature was decreased to 30 °C and after 5 h of incubation, the cells

were harvested by centrifugation (5000 g, 10 min, 4 °C) and the pellets were frozen at -20 °C.

2.2.2 Protein purification

2.2.2.1 Purification of YfiN

Cells were resuspended in 25 mL lysis buffer (50 mM Tris-HCl, pH 8, 200 mM NaCl), lysed with a French press at 15 000 psi and centrifuged (10'000 g, 15 min, 4 °C) to remove unbroken cells and inclusion bodies. The supernatant was ultracentrifuged (100'000 g, 30 min, 4 °C) for collection of the total membrane fraction, which was resuspended in 15 mL lysis buffer, supplemented with 21 mM DM. The solubilization process was allowed to take place for one hour with steady stirring applied. After another centrifugation step (100'000 g, 15 min, 4 °C), the supernatant was then loaded onto a 5 mL HisTrap column (GE Healthcare). After washing the column with YfiN-Ni-A buffer (50mM Tris-HCl, pH 8, 200 mM NaCl, 5 mM imidazole, 3.1 mM DM) for a length of 10 column volumes (CV), the protein was eluted with a linear gradient of imidazole from 5 to 500 mM in 10 CV. YfiN containing fractions were pooled and concentrated to 4 mL. As an increased 260/280 ratio was observed after IMAC, indicating the presence of co-eluting nucleotides, YfiN was further purified by running a size-exclusion chromatography overnight at 0.2 mL/min flowspeed using a HiLoad-26/60-Superdex-200 prep grade column (GE Healthcare) and YfiN-GF buffer (50 mM Tris-HCl, pH 8, 200 mM NaCl, 3.1 mM DM). The purification process was monitored by SDS-PAGE. Protein concentration

was determined by measuring the absorbance at 280 nm ($\epsilon_{280} = 28420 \text{ M}^{-1} \text{cm}^{-1}$).

2.2.2.2 Purification of YfiN_{PAS} constructs

For the purification of YfiN_{PAS44-149}, cells were resuspended in 25 mL lysis buffer (50 mM Tris-HCl, pH 8, 200 mM NaCl), lysed with a French press at 15 000 psi and ultracentrifuged (100'000 g, 35 min, 4 °C) to remove cellular debris and insoluble material. The supernatant was then loaded onto a 5 mL HisTrap column (GE Healthcare). After washing the column with Ni-A buffer (50mM Tris-HCl, pH 8, 200 mM NaCl, 5 mM imidazole) for a length of 10 CV, the protein was eluted with a linear gradient of imidazole from 5 to 500 mM in 10 CV. YfiN_{PAS4-149} containing fractions were pooled and concentrated to 3 mL and further purified by size-exclusion chromatography using a HiLoad-16/60-Superdex-75 prep grade column (GE Healthcare) and GF buffer (50 mM Tris-HCl, pH 8, 200 mM NaCl). The purification process was monitored by SDS-PAGE. Protein concentration was determined by measuring the absorbance at 280 nm ($\epsilon_{280} = 15470 \text{ M}^{-1} \text{cm}^{-1}$).

For MBP-tagged YfiN_{PAS38-159} and YfiN_{PAS44-149}, the cells were lysed in the same way as described above but in a lysis buffer with a different composition (50 mM Tris-HCl, pH 7.4, 200 mM NaCl, 1 mM EDTA). The supernatant was then loaded onto a 5 mL MBPTrap column (GE Healthcare). After washing the column with lysis buffer for a length of 10 column volumes (CV), the protein was eluted with a linear gradient of

maltose from 0 to 10 mM in 10 CV. YfiN_{MBP fusion protein} containing fractions were pooled and concentrated to 3 mL and further purified by size-exclusion chromatography using a HiLoad-16/60-Superdex-75 prep grade column (GE Healthcare) and GF buffer (50 mM Tris-HCl, pH 7.4, 200 mM NaCl). The purification process was monitored by SDS-PAGE. Protein concentration was determined by measuring the absorbance at 280 nm ($\epsilon_{280} = 81820 \text{ M}^{-1} \text{ cm}^{-1}$).

2.2.2.3 Refolding of the YfiN_{PAS44-149} construct

For the refolding of YfiN_{PAS44-149} expressed in inclusion bodies, cells were resuspended in 25 mL lysis buffer (50 mM Tris-HCl, pH 8, 200 mM NaCl), lysed with a French press at 15 000 psi and ultracentrifuged (100'000 g, 35 min, 4 °C) to separate the soluble fraction from the insoluble material. The pellets were resuspended in lysis buffer containing 2 % Triton X-100, to solubilize the membranes, followed by a ultracentrifugation step (100'000g, 10 min, 4 °C). This procedure was repeated three times to get rid of all membranes. The purified inclusion bodies were then resuspended overnight in lysis buffer containing 8 M urea. The next day, a final ultracentrifugation step was applied and the supernatant was loaded onto a 5 mL HisTrap column for on-column refolding.

2.2.2.4 Purification of YfiN_{HAMP-GGDEF} from *P. aeruginosa*, *P. fluorescens*, *Y. enterocoliticae* and *S. alaskensis*

The purification was performed as described in 2.2.2.2 for YfiN_{PAS4-149}, except that all buffers were supplemented with 5 % glycerol and contained protease inhibitor (Roche). Protein concentration was determined by measuring the absorbance at 280 nm (*P.aer.* $\epsilon_{280}= 8480 \text{ M}^{-1}\text{cm}^{-1}$; *P.fluor.* $\epsilon_{280}= 9970 \text{ M}^{-1}\text{cm}^{-1}$; *Y.ent.* $\epsilon_{280}= 19035 \text{ M}^{-1}\text{cm}^{-1}$; *S.alas.* $\epsilon_{280}= 10095 \text{ M}^{-1}\text{cm}^{-1}$).

2.3 Crystallization of YfiN_{HAMP-GGDEF} from *P. aeruginosa* and *Y. enterocoliticae*

For protein crystallization, the vapor diffusion method was used. Initial screening was performed with commercial screens in 96-well plates using the sitting drop method. For each drop, 0.2 μL protein solution was mixed with an equal amount of reservoir buffer. The reservoir volume was 75 μL . Afterwards the trays were sealed and stored at a constant temperature of 20 $^{\circ}\text{C}$.

2.4 Production of membranes expressing YfiN

The membrane fraction for activity tests was prepared as described in 2.2.2.1, except that after the second centrifugation step, no extraction was

performed. Instead, the membrane pellet was collected, resuspended and thoroughly washed in 10 mL lysis buffer (50 mM Tris-HCl, 200 mM NaCl) by pipetting up and down, followed by a centrifugation step (10'000 g, 10 min, 4 °C). This procedure was repeated 4 times.

2.5 Activity measurements

2.5.1 FPLC-based nucleotide quantification

To test for diguanylate activity, washed membranes expressing YfiN were incubated with 2 mM GTP (Sigma) in 50 mM Tris-HCl, pH 8.0, 200 mM NaCl and 5 mM MgCl₂. 200 µl of sample were taken at different time points and the reaction was stopped by heating the sample to 80 °C for 10 min. The reaction mixture was then centrifuged (10'000 g, 2 min, 4 °C) and subsequently, 100 µL were diluted in 900 µL 5 mM NH₄HCO₃, pH 8.0, filtered (0.22 µm) and 500 µl were loaded on an ion-exchange column (ResourceQ 1 µL, GE Healthcare). The nucleotides were separated with a gradient from 0.005 M to 1 M NH₄HCO₃, pH 8.0, in 14 CV. The elution of the reaction was compared to the elution profiles of GTP and c-di-GMP (own production).

2.5.2 Phosphate sensor assay

In order to establish the phosphate sensor for quantification of DGC activity (Figure 3-30), DgcZ was used first as a well-characterized reference protein. The experiments were set up as follows: 0.5 μM phosphate sensor (PS), 130 nM pyrophosphatase (PP), different concentrations of GTP (5 μM , 10 μM , 20 μM , 50 μM and 100 μM) and 200 μM zinc-free DgcZ [111] were added sequentially to 30 μL of a 10 x stock solution of reaction buffer (50 mM Tris pH 7.5, 150 mM NaCl, 50 mM Arg/Glu, 5 mM MgCl₂) and adjusted with H₂O to yield a final volume of 300 μL . After addition of every component the resulting fluorescence was recorded by a spectrofluorometer Jasco (Jasco analytical instruments) and the reaction was allowed to take place for 14 min. After termination of the enzymatic reaction, 100 μM P_i was added to the experimental setup to fully saturate the PS.

For the characterization of the enzymatic activity of YfiN_{HAMP-GGDEF} from *P. aeruginosa* the same protocol was used as described above. 5 μM YfiN_{HAMP-GGDEF} were incubated with 50 μM GTP in reaction buffer (50 mM Tris-HCl, 200 mM NaCl, 5 mM MgCl₂) and the reaction was monitored for 20 min.

2.5.3 Malachite green assay

To measure the enzymatic activity of detergent-solubilized YfiN the Baykov assay [112] was used in a 96-well format. The setup was adapted for DGCs by replacing the alkaline phosphatase by pyrophosphatase, allowing the detection of phosphate as a readout for DGC activity. The well-characterized DGC DgcZ was again used as a reference protein to optimize the assay before using it with YfiN. The reaction mixture contained 200 nM

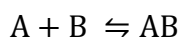
zinc-free DgcZ in 50 mM Tris-HCl, pH 8.0, 200 mM NaCl, 5mM MgCl₂, 130 nM pyrophosphatase from bakers yeast (Sigma) and either 1 μM or 10 μM GTP. The volume was adjusted with reaction buffer to 200 μL in each well. In order to stop the reaction at different time points, 40 μL of phosphate detection solution (0.55 M H₂SO₄, 0.46 mM malachite green, 2.66 mM (NH₄)₆Mo₇O₂₄, 0.04% Tween) was added to each well and the absorption was measured at 630 nm 15 min after incubation. The production of phosphate was quantified with a phosphate calibration curve.

For the characterization of the enzymatic activity of YfiN the same protocol was used as described above except that the reaction buffer was supplemented with 3.1 mM DM.

2.5.4 Theoretical concepts and formulas used for the performed activity assays

2.5.4.1 The formula to calculate the concentration of a protein complex depending on the total protein and ligand concentrations and the K_d

The formation of a protein complex (AB) by protein (A) and protein (B) can be described as follows:



where A and B are the concentrations of the free proteins and AB describes the concentration of the protein complex.

The corresponding dissociation constant is defined as:

$$Kd = \frac{[A] \times [B]}{[AB]} \quad (1)$$

The total amount of A and B are each made up by two fractions, free protein and complexed protein:

$$[A_{total}] = [A_0] = [A_{free}] + [AB], \text{ which can be rewritten as:}$$

$$[A_{free}] = [A_0] - [AB] \quad (2)$$

$$[B_{total}] = [B_0] = [B_{free}] + [AB], \text{ which can be rewritten as :}$$

$$[B_{free}] = [B_0] - [AB] \quad (3)$$

Substituting (2) and (3) in (1):

$$Kd = \frac{([A_0] - [AB]) \times ([B_0] - [AB])}{[AB]}$$

This equation can be solved for x using WolframAlpha [113].

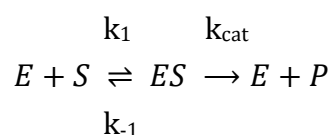
$$[AB] = 0.5 \times \left[\frac{([A_0] + [B_0] + Kd) \pm \sqrt{([A_0]^2 - 2([A_0] \times [B_0]) + 2([A_0] \times Kd) + ([B_0] \times [A_0]) + [B_0]^2 + 2([B_0] \times Kd) + (Kd^2))}}{2} \right]$$

Equation 1: Equation used to calculate the concentration of a protein complex depending on the concentrations of the total concentrations of the components and the Kd.

2.5.4.2 General enzyme kinetics and derivation of the Michaelis-Menten Equation

Enzymes are biological catalysts that accelerate or catalyze chemical reactions. The molecules at the beginning of the process upon which enzymes may act are called substrates and the enzyme converts these into molecules called products.

The substrate S binds reversibly to an enzyme E to form an enzyme-substrate complex ES . ES then in turn reacts irreversibly to generate a product P and to regenerate the free enzyme E . This system can be represented schematically as follows:



Equation 2

In the above scheme there are two different reactions. The first reaction depicted with the double arrow is a reversible reaction describing the reversible binding and dissociation of the enzyme and the substrate with the rates k_1 and k_{-1} respectively. The second reaction is an irreversible chemical reaction where the enzyme-substrate complex is converted into product and free enzyme symbolized by the single arrow. The rate of this reaction (v) is proportional to the concentration of the enzyme substrate complex $[ES]$ and the catalytic rate k_{cat} , which is a measure for how many substrate molecules are turned over per enzyme molecule per second (s^{-1}).

$$v = \frac{d[P]}{dt} = \left(-\frac{dS}{dt}\right) = k_{cat} \times [ES]$$

Equation 3

ES is an intermediate state and its concentration is therefore unknown but it can be expressed employing known values if the following assumptions are fulfilled:

- As long as **initial velocity v_0** is considered, the concentration of product can be neglected (compared to that of the substrate, thus $[P] \ll [S]$), and the concentration of substrate is in large excess over that of the enzyme ($[E] \ll [S]$).
- Mass conservation equations for the reactants state that:

$$[S_{total}] = [S_0] = [S_{free}] + [ES] + [P], \text{ which now approximates to } [S_0] = [S]$$

and

$$[E_{total}] = [E_0] = [E_{free}] + [ES]$$

During the initial phase of the reaction, as long as the reaction velocity remains constant, the reaction is in a **steady state**, resulting in ES being formed and consumed at the same rate. During this phase, the rate of association of ES equals its rate of dissociation. According to Equation 2:

$$\text{Formation of the complex: } \frac{d[ES]}{dt} = k_1 \times [E] \times [S]$$

$$\text{Dissociation of the complex: } \frac{d[ES]}{dt} = (k_{-1} + k_{cat}) \times [ES]$$

So in the **steady state**,

$$(k_{-1} + k_{cat}) \times [ES] = k_1 \times [E] \times [S]$$

Rearranging

$$\frac{(k_{-1} + k_{cat})}{k_1} = \frac{[E] \times [S]}{[ES]}$$

Express [E] in terms of [ES] and [E₀] to limit the number of unknowns:

$$[E] = [E_0] - [ES]$$

Defining Km, the Michaelis-Menten constant, as

$$Km = \frac{(k_{-1} + k_{cat})}{k_1}$$

Leading to

$$Km = \frac{([E_0] - [ES]) \times [S]}{[ES]}$$

The term [ES] can then be defined as:

$$[ES] = \frac{[E_0] \times [S]}{Km + [S]}$$

Equation 4

Substituting Eq. 4 in Eq. 3

$$v_0 = k_{cat} \times [ES] = \frac{k_{cat} \times [E_0] \times [S]}{Km + [S]}$$

Equation 5

The maximum reaction velocity, V_{max} , is reached when all enzyme sites are saturated with substrate. This is true for the condition when $[S] \gg K_m$, leading to $[S]/(K_m+[S])$ approaching 1. In this case, V_{max} can be expressed as:

$$V_{max} = k_{cat} \times [E_0]$$

Substituted in Eq. 5

$$v_0 = V_{max} \frac{[S]}{K_m + [S]}$$

Equation 6: This equation, also termed Michaelis-Menten equation, expresses the initial rate of reaction in terms of a measurable quantity, the **initial** substrate concentration. The two kinetic parameters, V_{max} and K_m , will be different for every enzyme-substrate pair.

2.6 Bioinformatical analysis

2.6.1 Identification of YfiN homologs predicted to crystallize more readily

Identification of YfiN homologs was performed by performing an NCBI-BLAST against the sequence of the full-length protein, with a filter of 40 % - 70 % sequence identity applied. Potential candidates had to have the same domain architecture as YfiN, namely a PAS, HAMP and an intact GGDEF domain plus I-site. In addition, the organism had to have a complete YfiBNR operon [14] as it is the case in *P. aeruginosa*. Xtalpred analysis [114] was then performed with the most promising candidates, and three homologs were finally selected that fulfilled all the previously stated criteria

and were predicted to crystallize more readily than YfiN from *P. aeruginosa*.

2.6.2 Identification and localization of the potential secondary I-site in the HAMP domain of YfiN

The YfiN protein sequence from *P. aeruginosa* was retrieved from the Uniprot database [115], corresponding to accession number Q9I4L5. In order to generate a multiple sequence alignment (MSA) for HAMP domains occurring exclusively in YfiN homologs, the sequence of YfiN was blasted against the nr70 database. Hits that did not cover at least 75 % of the sequence or did not represent DGC entries were omitted. Finally, 95 sequences were retrieved and aligned to the sequence of YfiN.

For the second alignment, Tar, a methyl accepting chemotaxis protein containing a HAMP domain, was blasted against the nr70 database. The resulting 100 target sequences were then aligned against YfiN, yielding a second MSA with a presumably different consensus sequence for the HAMP domains as the first one. A sequence logo was then generated using Geneious for both MSA using the sequence of YfiN as a reference.

2.6.3 Non-competitive product inhibition model of YfiN_{HAMP-GGDEF}

The structure of YfiN_{GGDEF} (pdb accession code 4iob) was superposed onto the GGDEF domain of feedback-inhibited PleD (pdb accession code 1w25). A homology model was then generated of YfiN_{HAMP} using HHpred [116] based on the HAMP domain of the osmolarity sensor protein EnvZ (E-value

6×10^{-12}). The second helix of the HAMP-model was then aligned to the last helix ($\alpha 5$) of PleD. The structural coordinates for the intercalated c-di-GMP dimer were taken from PleD and added to the theoretical model of the product-inhibited state of YfiN_{HAMP-GGDEF}. As the YfiN_{HAMP} and YfiN_{GGDEF} domains are connected by a linker, which is not seen in the structure, allowing rotational and translational freedom between the two domains, the position of the GGDEF domain was slightly adjusted in order to accommodate Arg208 from the HAMP domain as a third c-di-GMP coordinating residue.

3 Results

3.1 Research article I (Kauer et al., in preparation)

The crystal structure of YfiR is providing new insights into YfiR-mediated repression of YfiN, a key regulator of biofilm formation in *Pseudomonas aeruginosa*

Stefanie Kauer^a, Pablo Manfredi^b, Jacob G. Malone^b, Tina Jaeger^b, Timothy Sharpe^c, Urs Jenal^{b} and Tilman Schirmer^{a*}*

^a Focal Area Structural Biology and Biophysics, Biozentrum, University of Basel, Klingelbergstrasse 50/70 CH - 4056 Basel, Switzerland.

^b Focal Area Infection Biology, Biozentrum, University of Basel, Klingelbergstrasse 50/70 CH - 4056 Basel, Switzerland.

^c Biophysics Facility, Biozentrum, University of Basel, Klingelbergstrasse 50/70 CH - 4056 Basel, Switzerland.

*Corresponding authors.

Manuscript in preparation**Statement of my own contributions**

I designed and performed all the experimental parts presented in the *research article I*. I cloned, expressed, purified and crystallized YfiR. Furthermore, I have solved the crystal structure. Ultimately, I analyzed the oligomeric state of YfiR and the mutant in solution by size-exclusion chromatography coupled to multi angle laser light scattering (SEC-MALLS).

The YfiBNR system was originally investigated and characterized by Jacob Malone and Tina Jaeger, including the discovery of the activating YfiR and YfiN mutations. Pablo Manfredi performed the bioinformatical analysis and provided me with the multiple-sequence alignments used in this study. I wrote the manuscript except for the section “Bioinformatical analysis of DUF5145 proteins” (Material and Methods), which was written by Pablo Manfredi. Furthermore, I prepared and assembled all the figures and tables presented in the manuscript, except for the two figures in the supplementary data, which were produced by Pablo Manfredi.

The crystal structure of YfiR is providing new insights into YfiR-mediated repression of YfiN, a key regulator of biofilm formation in *Pseudomonas aeruginosa*

Stefanie Kauer^a, Pablo Manfredi^b, Jacob G. Malone^b, Tina Jaeger^b, Timothy Sharpe^c, Urs Jenal^b and Tilman Schirmer^a

^a Focal Area Structural Biology and Biophysics, Biozentrum, University of Basel, Klingelbergstrasse 50/70 CH - 4056 Basel, Switzerland.

^b Focal Area Infection Biology, Biozentrum, University of Basel, Klingelbergstrasse 50/70 CH - 4056 Basel, Switzerland.

^c Biophysics Facility, Biozentrum, University of Basel, Klingelbergstrasse 50/70 CH - 4056 Basel, Switzerland.

Corresponding authors:

Prof. Tilman Schirmer

Prof. Urs Jenal

Biozentrum, University of Basel

Klingelbergstrasse 70

CH-4056 Basel, Switzerland

Tel.: +41 61 267 20 89

Tel.: +41 61 267 10 84

tilman.schirmer@unibas.ch

urs.jenal@unibas.ch

Abstract

The opportunistic pathogen *Pseudomonas aeruginosa* is responsible for chronic infections in the lungs of cystic fibrosis patients, a process that is eventually accompanied by the formation of small-colony variants (SCVs). The appearance of SCVs in the sputum of patients is caused by elevated levels of c-di-GMP and correlates with increased persistence of infection and antibiotic resistance. The YfiB_{NR}-system has been previously identified as a key regulator of the SCV phenotype. The effector of this tripartite system is the inner-membrane bound diguanylate cyclase (DGC) YfiN. While the mechanistic principles of interaction between the three proteins have been established by extensive *in vivo* studies, no structural information has been available for any of the three proteins so far.

We have solved the crystal structure of YfiR from *P. aeruginosa*, which revealed that it forms a dimer and adopts a novel protein fold. By SEC-MALLS analysis, we provide evidence that YfiR is dimerizing in solution as well, via the same interface observed in the crystal structure. The data yielded a physiologically relevant dissociation constant (K_d) in the nM range. The crystal structure allowed the investigation of the proposed YfiN binding site, which revealed a highly conserved hydrophobic groove at the C-terminal end of the protein. Furthermore, as YfiR belongs to the domain of unknown function (DUF)4154 family, we investigated this class by bioinformatical means, which assigned them a regulatory role for c-di-GMP related proteins and histidine kinases.

It is therefore proposed that YfiR is regulating YfiN activity as a dimer via a conserved and hydrophobic site located at the very C-terminal end of the protein.

Introduction

Bis-(3'-5')-cyclic dimeric guanosine monophosphate (*c*-di-GMP) is an ubiquitous second messenger that regulates the switch between the motile planktonic and sedentary biofilm-associated lifestyles of bacteria [117]–[119]. *c*-di-GMP is produced from two GTP molecules by diguanylate cyclases (DGC) and degraded to 5'-phosphoguanylyl-(3'-5')-guanosine (pGpG) by phosphodiesterases (PDE). The catalytically active part of the DGC is the GGDEF domain that is named after the amino acid sequence motif, which constitutes an essential part of the active site of the enzyme [23], [43], [120]. Phosphodiesterase activity is generally associated with the presence of EAL [8], [48], [50], [121] or HD-GYP domains [52].

While low intracellular levels of *c*-di-GMP promote planktonic behavior, where cells are generally motile and express virulence factors, increasing concentrations of *c*-di-GMP promote the expression of adhesive matrix [2], [5]–[7] components and result in multicellular behavior, biofilm formation and persistence [107].

The opportunistic pathogen *Pseudomonas aeruginosa* is responsible for chronic infections in the lungs of patients suffering from cystic fibrosis [103] or patients with compromised host defense mechanisms. With progressive lung colonization, *P. aeruginosa* undergoes genetic and phenotypic adaptations to the host environment, resulting in reduced virulence [106] and increased persistence [104], [105]. Phenotypic adaptations lead to the formation of small colony variants (SCV), which are characterized by high auto-aggregation, attachment to surfaces [108], slow growth and enhanced exopolysaccharide production [108], [122]. These variants appear in sputum samples of CF patients and correlate with antibiotic resistance due to slow

growth or increased persister cell population [13], [16], resistance to phagocytosis [12] and prolonged persistence of infections. These findings suggest that the SCV phenotype may provide fitness advantages during the process of chronic lung infections and therefore plays an important role in the pathogenesis of *P. aeruginosa*.

It has been proposed that the appearance of the SCV phenotype is strongly linked to enhanced levels of c-di-GMP [12], [45], [109], [110]. The YfiBNR-system has been previously identified as a tripartite signaling system, which is regulating c-di-GMP levels in *P. aeruginosa* in response to a still uncharacterized environmental input signal [12], [14]. YfiN thereby constitutes the effector of the system as an inner-membrane bound DGC, which consists of a periplasmic PAS domain and cytoplasmic HAMP and GGDEF domains. The YfiBNR system has been shown to function via tightly controlled competition between allosteric binding sites, which are highly conserved and hydrophobic and which are found on all three Yfi proteins [14]. YfiN activity was shown to be suppressed in vivo by interacting with the periplasmic protein YfiR [12]. Release of repression is mediated by the outer-membrane lipoprotein YfiB, which is undergoing a conformational change in response to an unidentified signal and thereby sequesters YfiR to the outer membrane. SCV formation is then caused by enhanced c-di-GMP production by YfiN, which in turn leads to elevated transcription of the *pel* and *psl* operons [12].

A key role in the regulation of biofilm formation is therefore attributed to the periplasmic protein YfiR as it shuttles between the presumable sensor YfiB in the outer membrane and the inner-membrane DGC YfiN. However, no structural or mechanical information is available on any members of the

YfiR family. According to pfam analysis, it is annotated as a member of the domain of unknown function (DUF) 4154 family [123].

In this study, we have solved the crystal structure of YfiR from *P. aeruginosa*, which revealed that YfiR is exhibiting a novel protein fold. By a combination of X-ray crystallography and SEC-MALLS analysis, we provide evidence that YfiR is dimerizing in solution using the same interface observed in the crystal structure. The Kd of dimerization was determined to be in the nM range. Furthermore, we sought to investigate the DUF5145 family bioinformatically to shed some light on their possible function in nature.

Material and Methods

Cloning, Expression and Purification

A synthetic DNA fragment coding for YfiR comprising residues 36-190, which lacks the original N-terminal signal peptide from *Pseudomonas aeruginosa*, was ordered from MrGene (Regensburg, Germany) and optimized in regard to codon-usage and GC-content. The construct was subcloned into a pET-22b (+) vector (Novagen), which provides an N-terminal pelB leader sequence resulting in export of the target protein to the periplasm, and addition of a C-terminal His6-tag to the polypeptide chain. Overexpression was performed in the *E. coli* strain BL21 Star™ (DE3). Cells were grown at 37 °C to an OD₆₀₀ of 0.6. After induction with 0.5 mM IPTG, the protein was expressed over night at 20 °C. Cells were disrupted by French Press and sonication in lysis buffer (20 mM Tris-HCl pH 8.0, 100 mM NaCl, 25 mM imidazole). After centrifugation at 100'000 g, the protein was purified by Ni-NTA affinity column (Ni-Sepharose, GE Healthcare,

Germany), eluting at approximately 370 mM imidazole. The pooled protein fractions were concentrated and further purified by size-exclusion (SEC) chromatography using a Superdex 75 16/60 column (GE Healthcare) and 20 mM Tris-HCl pH 8.0, 100 mM NaCl as running buffer.

Crystallization and Data Collection

Crystallization of YfiR was performed at room temperature by the sitting-drop vapor-diffusion technique. A 0.2 μ l + 0.2 μ l mixture of protein (10 mg/ml) and well solution (0.8 M Potassium sodium tartrate tetrahydrate, 0.1 M Tris-HCl pH 8.0, 0.5 % Polyethylene glycol monoethyl ether 5000) gave rise to small crystals of needle shape belonging to the space group P21. These crystals were used as source of microcrystals for streak-seeding, leading to the formation of larger rod-like crystals. Crystals were cryoprotected by stepwise increasing the glycerol concentration to 30 % and flash cooled in liquid nitrogen. The same protocol was applied for selenomethionine-labelled YfiR.

Crystal Structure Determination and Refinement

A SAD dataset at the selenium K-edge was collected at beamline X06DA , Swiss Light Source, Paul-Scherrer-Institute, Villigen and data processing was performed with MOSFLM/SCALA [124] [125]. At this step, a set of 5 % reflections (R_{free} set) was taken out for cross validation. Selenium positions were found with SHELXD [126], and the initial model was built by ARP/wARP [127]. Structure refinement and model building was performed using iterative cycles of phenix.refine [128] and COOT [129]. Ligand molecules were modeled in the $F_o - F_c$ difference electron density map.

Finally, water molecules were placed where the $F_o - F_c$ map exceeded 3σ and potential hydrogen bonds could be formed. Model validation was carried out with MolProbity [130]. All structure figures were produced by using DINO [131].

SEC Exclusion Chromatography (SEC) Coupled with Multiangle Light Scattering (MALS)

A Wyatt silica SEC column (4.6×300 mm, $5 \mu\text{m}$ bead, 300 \AA pore) on an Agilent 1100 series HPLC was employed for the determination of the oligomeric state of YfiR and YfiR_{D80R} at different loading concentrations. The instrument was coupled to a Wyatt miniDawn TriStar multiangle light scattering detector and a Wyatt Optilab rRex refractive index detector. The column was equilibrated for 3 h to obtain stable baseline signals from the detectors before data collection. The inter-detector delay volumes and band broadening, the light-scattering detector normalization, and the instrumental calibration coefficient were calibrated using a standard 2 mg/ml of BSA solution (Thermo Pierce) run in the same buffer, on the same day, according to standard Wyatt protocols. The absolute refractive index of the buffer was measured using the refractive index detector.

The protein sample (20 μl) was loaded on the column. All experiments were performed at 6 °C at a flow rate of 0.5 ml/min in SEC buffer (see the purification section).

The molecular weight and mass distribution of the samples were then determined using the ASTRA 5 software (Wyatt Technology).

SEC-MALS derived apparent mass values were used to derive the dimerization affinity of YfiR, assuming a fast monomer-dimer equilibrium.

According to the mass action law, the monomeric molar fraction, $x_m = [M]/[M0]$, is given, by,

$$x_m = [(8 \times [M0] \times Kd + Kd^2)^{1/2} - Kd] / (4 \times [M0])$$

with $[M0] = [M] + 2 \times [MM]$, total protein concentration at elution; $[M]$, monomer concentration; $[MM]$, dimer concentration.

This yields for the weight-averaged apparent mass (see also [132])

$$m_{apparent} = x_m \times m_{mono} + (1 - x_m) m_{dimer} = (2 - x_m) \times m_{mono} \quad (\text{Eq.1})$$

The experimental data (Figure 3) were fitted (program ProFit) to Equation 1 with a fixed m_{mono} of 18 kDa to yield the dimerization K_d .

Bioinformatical analysis of DUF5145 proteins

All sequences of single DUF4154 domain proteins (1721 protein sequences) were extracted from NCBI via the architecture service [133]. Only protein sequences that were related to genomic background information were considered. All protein sequences encoded by genes located up to two genes downstream or upstream the DUF4154 domain protein were submitted to PFAM annotation (2996 sequences) [123]. Single DUF4154 domain proteins were then categorized according to the functional annotation found in their respective neighboring genes. Prior processing by WebLogos using GENEIOUS [134], the datasets were subjected to blasclust (-L75 -b70) in order to reduce bias from overrepresented sequence sources.

Results

Structure Determination

The crystal structure of recombinant processed YfiR from *P. aeruginosa* (residues 1-150, corresponding to 36-190 of the mature protein) was solved to a resolution of 1.8 Å in space-group P2₁ with 4 monomers per asymmetric unit. Crystallographic data are given in table 1. Initial phases were obtained by single-wavelength anomalous dispersion phasing of SeMet-derivatized crystals and the resulting experimental electron density map allowed placement of the full structure using ARP/wARP.

The four subunits with similar structures (RMSD of 0.2-0.5 for all C α positions) are arranged in two identical dimers: A-B and C-D. Refinement with non crystallographic symmetry restraints yielded a final model with an R/R_{free} of 17.3/21.3. The final electron density is of high quality and the entire main-chains of the 4 subunits are well defined except for the N-terminal region from residues 35-39 in A, and 35-37 in B and D and 35-38 in C; the C-terminal region from residues 182-190 in A, 186-190 in B and 185-190 in C and D, and the purification tag in subunits B and D. These residues were therefore not included in the final model. Additionally, weak electron density was observed for several side-chains in two loop regions of the protein (142-144; 156-159). The following discussion is restricted to the A-B dimer.

Crystal Packing

The monomers within the dimer are held together by a set of 3 hydrogen bonds, 4 salt bridges (as suggested by the PISA-server [135]), as well as two van der Waals interactions, resulting in an approximate dimeric interface

area of 675 Å² (table 2). Residues belonging to the N-terminus, the loop between beta strand 1 and helix $\alpha 2$ (dimerization loop) and the loop between beta strand 3 and helix $\alpha 3$ account for most of the intermonomeric contacts within the dimers (Figure 2).

The overall structure of YfiR reveals a new fold

The entire polypeptide chain folds into a single globular domain of approximate dimensions 26 Å × 37 Å × 26 Å (Figure 1). The core of the protein is made up by a seven-stranded mixed β -sheet, flanked on the convex side by three helices, and embracing a long N-terminal α -helix with its concave side. Additionally, there is a second small α -helix present on the concave side that is mediating a crossover between $\beta 1$ and the $\beta 2$ - $\beta 3$ hairpin loop. The β -sheet topology is 2-3-1-4-5-6-7, where strand 2 and 7 are aligned in an antiparallel fashion in regards to the rest of the sheet. There is an additional small β -sheet present, formed by $\beta 0$ and $\beta 8$, located at either end of the protein. The protein is stabilized by two intramolecular disulfide bonds, mediated by Cys71-110 as well as Cys145-Cys152. The latter seems to be important to keep the loop-region between $\beta 5$ and $\beta 6$ in place.

At the very C-terminus of molecules B, C and D, the formation of an additional α -helix is observed (residues 178-185) but not in A, as the electron density of this subunit was overall of a lesser quality and therefore it was built only up to residue 181. For all four subunits, no electron density was observed for residues 186-190 but in subunits A and C placement of the following C-terminal his-tag was at least partially possible.

A search with FATCAT [136] identified that the highest structural similarity is with a 5-formyltetrahydrofolate cycloligase-related protein

(PDB entry: 1wkc, NagB/RpiA/CoA transferase-like fold) and hypoxanthine phosphoribosyltransferase (PDB entry: t1tc, PRTase-like fold). DALI [137] showed significant hits for several ABC domain or ATP binding proteins (PDB entries: 3lft, 3lkv, 3lft, rossman fold). All three folds fall into the class of α/β proteins and assume a 3-layered $\alpha/\beta/\alpha$ sandwich like YfiR. However, superposition of YfiR onto all five identified hits with an r.m.s.d of 3.1 Å, revealed that even though there are structural similarities, the resulting topology diagrams look completely different. YfiR is thus adopting a novel fold, characterized by a $\alpha/\beta/\alpha$ sandwich, displaying a strand order of 2-3-1-4-5-6-7 (Fig. 1 B).

Fast concentration-dependent YfiR monomer-dimer equilibrium

The oligomeric state of YfiR in-vitro was analyzed using SEC coupled with MALS. The data showed that the retention time and apparent mass values are dependent on the YfiR concentration (Figure 3), which is indicative of a self-association equilibrium that is fast on the time-scale of the experiment. The weight-averaged mass values were fitted to the indicated monomer-dimer equilibrium model (see “Material and Methods”). The protein is largely dimeric even at the lowest eluting concentration of 1 μ M, and nearly all proteins are forming dimers at an eluting concentration of 9 μ M. The corresponding Kd of dimerization was determined to be 139 nM.

For further functional studies, a mutant was generated with the aim to prevent dimer formation. Therefore, the salt-bridge, formed by Asp80 and Arg98, was specifically disrupted by mutating residue Asp80 to an Arginine. As indicated by the SEC-MALS chromatogram, the D80R mutant is still

forming a fast monomer-dimer equilibrium but with a dissociation constant K_d of 21 μM that is two orders of magnitude larger than for the wild-type.

Conservation Score Analysis

Conservation score analysis revealed that there are multiple conserved patches in the protein (Figure 4, A, B, C). The central beta-strands situated in the hydrophobic core of the protein are overall well conserved, as are residues involved in positioning of the disulfide-bonds and the cysteines themselves. Another well-conserved patch of the protein is clustering around the second part of the long N-terminal helix as well as β_0 and the following loop region until Pro62 (Figure 4, B, D). By far the largest conserved area is situated at the far C-terminal end of the protein, involving β_7 , α_5 , β_8 , α_6 as well as all interconnecting loop regions. Interestingly, the conserved patch around β_0 could be seen as an extension of this hyperconserved C-terminal region, forming together a large conserved patch at the opposite end of the dimeric interface. If this patch is portrayed as a surface representation it becomes evident that this is the only conserved part, that is solvent-accessible, and that the rest of the protein is rather variable in terms of amino-acid composition (Figure 4C). Zooming into this conserved region of the protein, reveals a hydrophobic platform, formed by residues Ile169, Ala170, Val176 and L181 framed by a highly conserved rim, consisting of residues R60, L166, D167, R171, G173, R175 and Pro178 (Figure 4, D, E). Overall, the rim is contributing to the hydrophobic nature of the platform, except for the few charged residues R60, D167, R171 and R175 (Figure 4 F). Interestingly, the dimeric interface is only moderately conserved. However, this can be explained by the observed types of interactions between the

monomers, which are not requiring high specificity in terms of residues that are present, such as van der Waals interactions or hydrogen bonds between main chain atoms (table 2). Furthermore, Glu140 was found to be engaging in hydrogen-bond formation with the backbone nitrogens of residues Glu77 and Tyr78, thereby forcing the dimerization loop to adopt its proper conformation (Figure 2). Interestingly, Glu140 is fully conserved and completely buried (2 % exposed) which is rather uncommon for such a highly charged amino acid.

Activating YfiR-mutations

Constitutively active YfiN variants that show reduced binding or fail to bind to YfiR altogether have been previously identified [14] (Figure 6 F). In the same study, complementary mutations in YfiR have been identified that restored wild-type colony morphology in the presence of some of the activated YfiN variants. Most of these mutations clustered either in the signal sequence region or at the C-terminal end of the protein. It was shown that signal sequence mutants led to increased expression of the protein, pointing towards a mere quantitative effect, which restored YfiR binding [14]. These mutations were therefore not further analyzed in the present study. The position of the residual mutations might help to shed some light on potentially interesting regions of the protein and therefore they were classified according to their conservation score and located in the structure (Table 3, Figure 5). Mutated residues, ranked in a moderately conserved class, were found either at the surface of the protein (K63E, Q125L and E163G) or completely buried in the core of the protein (V112A). All highly conserved residues (R60H, F151L and I169V) that showed a gain-of-function

effect are located at the C-terminal end of the protein in the highly conserved patch.

YfiN_{PAS} homology model

No structural information is available for the periplasmic domain of YfiN. Therefore, a homology model was generated using HHPred based on the template CitA, which showed a E-value of 5.7e-7 [116]. According to the model, YfiN is adopting a classical periplasmic PAS fold (Figure 6B). Mapping the conservation score of each residue onto the structure revealed that all completely conserved residues are situated in the first three helices of the protein (Figure 6A, 6C). As it was already proposed before, the putative YfiR binding site, characterized by the hydrophobic amino acid motif A-A-V-V-F, is nearly completely conserved and faces towards the aqueous environment (Figure 6D, 6E) [14]. Localization of the previously identified mutations, which lead to a constitutively active YfiN revealed four residues situated in the first helix, four in the hydrophobic stretch and one residue at the end of helix 3 (Figure 6F) [14]. Interestingly, four out of five possible positions in the potential YfiR binding site have been identified as activating YfiN mutations.

Discussion

In this study, we have successfully solved the structure of YfiR by X-ray crystallography, which revealed that YfiR is adopting a new fold as neither FATCAT nor DALI found any identical folding patterns. It is made up by a $\alpha/\beta/\alpha$ sandwich, displaying a strand order of 2-3-1-4-5-6-7 that has not been observed yet in the structural classification of proteins database (SCOP)

[138] (Figure 1). Moreover, the fact that a dimeric YfiR was found in the crystal structure, raised questions about the physiological relevance of this dimer. (Figure 2). SEC-MALLS analysis of wildtype YfiR and a dimer-interface mutant demonstrated that YfiR is also dimerizing in solution with a Kd of 139 nM, via the same interface observed in the crystal structure (Figure 3). Intracellular concentrations of YfiR have not yet been quantified to our knowledge; however considering that it is binding to a dimeric YfiN, which leads to an increase in local YfiR concentrations, the Kd is certainly in a physiologically relevant range.

Interestingly, despite the high affinity for dimerization, conservation score analysis revealed that the dimeric interface is only moderately conserved (Figure 4B). However, as the observed types of interactions between the monomers consist predominantly of mainchain-mainchain and van der Waals interactions, there is not much selective pressure on the preservations and conservation of a certain kind of side-chains (Table 2). Moreover, a highly conserved Glu (E140) was found completely buried in the protein (0.02% surface accessibility), which appeared to be positioning and stabilizing the dimerization loop by coordinating the backbone nitrogens of residues Glu77 and Tyr78. The same feature of loop stabilization by a so called “anchoring glutamate” is widespread among phosphodiesterases, where a Glu residue, e.g. E235 in YahA [36] or E268 in RocR [139], is involved in coordination of the dimerization loop $\beta 5-\alpha 5$ (loop 6) . Proper positioning of loop6 is crucial for catalytic activity and a E268A mutation in RocR rendered the protein inactive accordingly [139]. A similar importance for dimerization might be attributed to Glu140 in this study.

Besides the dimeric interface, conservation score analysis revealed an overall rather moderately conserved protein, except for one hyperconserved patch at

the C-terminal end of the protein. It has been proposed before by a combination of mutagenesis and analysis of sequence conservation that YfiR may bind to YfiN via a hydrophobic binding site located at the C-terminus [14]. We hereby provide the crystal structure to investigate the potential binding site in closer details. The well-conserved C-terminal part of YfiR, made up by the helices 5 and 6 and beta-strand 8, appears to form a sort of hydrophobic groove, which could easily accommodate another hydrophobic binding site (Figure 4D). In agreement with this hypothesis, highly conserved compensatory mutations in YfiR that are able to suppress activated YfiN are clustering in immediate proximity of this groove (Figure 5) [14]. The presumable YfiR binding site on YfiN is made up by a hydrophobic stretch of five highly conserved residues (A-A-V-V-F) localized on the surface of the periplasmic PAS domain. Even conservative mutations in this stretch abolished YfiR binding and thereby led to YfiN activation (Figure 6F) [14]. Considering the hydrophobic groove of YfiR acting as a potential binding site, it seems very appealing that YfiN is exhibiting such a highly conserved and hydrophobic stretch on the surface. It could therefore be speculated that the Ala and Val residues might be involved in the correct positioning of Phe70, which could then in turn be recognized by the hydrophobic groove of YfiR.

On the other side of the periplasmic space, YfiR is interacting with YfiB, which is an OmpA/Pal-like lipoprotein that is anchored to the outer membrane. Mutational studies on YfiB revealed a hydrophobic patch located immediately proximal to the insertion site of the protein to the membranes, which is proposed to serve as the potential YfiR binding site [14]. However, as complex structures of YfiR together with either of its interaction partners YfiB and YfiN are still missing, it remained to be

investigated, which YfiR residues are crucial for binding of YfiN or YfiB respectively.

We therefore sought to further investigate the DUF4154 family of proteins by bioinformatical analysis. All sequences of single DUF4154 domain proteins (1721 protein sequences) were extracted from NCBI via the architecture service. The protein sequences encoded by genes located up to two genes downstream or upstream the DUF4154 domain proteins were then submitted to PFAM annotation (2996 sequences). This revealed that the family is found most often together with c-di-GMP related proteins, followed by Plug, HAMP, histidine kinases and OmpA-like proteins (Table S1). We then categorized the DUF4154 domains in respect to their neighboring genes and generated sequence logos for each of them in order to get information about which residues might be involved in interactions with the aforementioned proteins (Figure 7). As YfiR is interacting with YfiN and YfiB, comparison of the DUF4154 + cdG and DUF4154 + OmpA sequence logos were of primary interest to this study (Figure 7A, 7B). Overall, the sequence conservation appears rather similar between the two groups, however the strength of global conservation seems to be more pronounced for DUF4154 proteins interacting with OmpA than with c-di-GMP related proteins. Significant differences can be found in the conservation of the second disulfide bond of YfiR, consisting of C145 and C152 and R60, which all appear completely conserved in DUF4154 proteins interacting with OmpA-like proteins in contrast to the c-di-GMP group (Figure 7A, 7B). From YfiR knockout studies we know, however, that the second disulfide bond is of lesser importance for protein stability, as a corresponding mutant was still able to complement a Δ YfiR-strain

(unpublished results, Jenal *et al.*). In contrast, the first disulfide bond, consisting of Cys71 and Cys110, is absolutely crucial for correct functioning of YfiR, as a corresponding mutant failed to complement a Δ YfiR-strain due to protein misfolding (unpublished results, Jenal *et al.*).

Of even bigger interest may be the strong conservation of residue Arg60 in OmpA-like proteins in comparison to the c-di-GMP group, which might constitute a YfiB specific interaction site. Moreover, the same residue has been identified before as a compensatory mutation in YfiR (R60H), which was able to suppress an activated YfiN (Table 3, Figure 4D) [14]. Initially, it was proposed that the R60H mutation was a gain-of-function mutation of YfiR in terms of YfiN binding, however it seems that, in fact, it was a loss-of-function mutation in terms of binding to YfiB. This finding seems reasonable as loss-of-function mutations are usually occurring at a higher frequency than their gain-of-function counterparts. In conclusion it is proposed that the YfiN-YfiR interaction, which is exhibiting a 2:2 complex, is in a dynamic equilibrium, with fractions of bound and unbound YfiN and YfiR in the periplasm. Rather than directly competing with the YfiN-YfiR interaction, activated YfiB would deplete the periplasm from unbound YfiR, thereby shifting the equilibrium to unbound and therefore active YfiN (Figure 8). The stoichiometry of the YfiB-YfiR complex, however, still remains to be investigated.

In addition to regulating the activity of c-di-GMP related proteins, the DUF4154 family was also found in the vicinity of genes coding for histidine kinases. Interestingly, several fusion events have been reported, where the DUF4154 domain is located at the N-terminal part of the polypeptide chain of histidine kinases, presumably regulating enzymatic affinity

intramolecularly [123]. The sequence logo of this class looks rather different than for c-di-GMP proteins, likely implying another mode of binding and regulation (Figure 7A, 7C). Ultimately, DUF4154 proteins are often occurring in the proximity of genes coding for plug domains, which are associated with outer membrane transporters. The sequence logo of this class also proposes a mechanism of interaction, which is substantially different from the one observed for c-di-GMP related proteins (Figure 7D). This led us to propose that the DUF4154 family is presumably involved in the regulation of either histidine kinases or c-di-GMP related proteins but most likely via different modes of interaction, which remain to be elucidated.

Figure 1:

Overall structure of *P.aeruginosa* YfiR₃₆₋₁₉₀.

(A) Cartoon representation with secondary structure elements, disulfide bonds and chain termini labeled. The numbers noted at the N- and C-terminal end designate the terminal amino acid residues of the molecule observed in the crystal. The protein was colored according to a red to yellow gradient.

(B) A topology diagram of the secondary structure of YfiR. The beta-sheet strands are shown in red, α -helices in yellow, disulfide bonds in green and connecting loops in black. Alpha helices lying on the concave side of the sheet are marked in bold, whereas the ones lying on the convex side are labeled in grey. The start of alpha helix 6 is observed in subunits B, C and D but not in A as residues 182-185 could not be modeled in that chain.

Figure 1

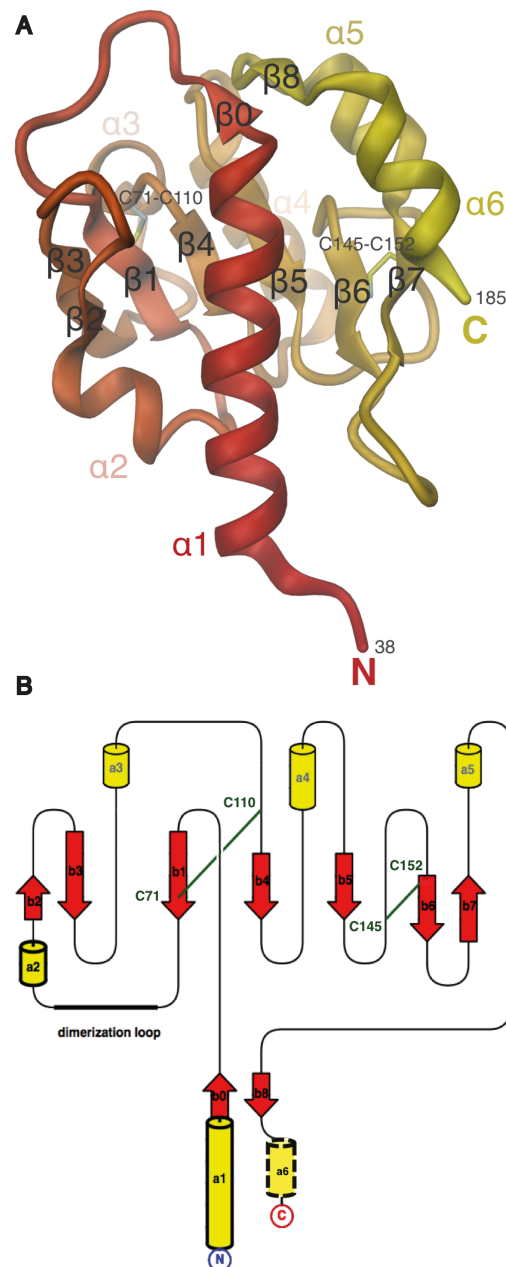


Figure 2:

Dimeric assembly of YfiR molecules in the crystal.

(A) The four YfiR molecules observed in the asymmetric unit associate to form two dimers, each of which is related by a noncrystallographic two-fold (shown in magenta). The subunits are distinguished by color.

(B) View along the local two-fold onto the dimeric interface. The same color code was used as in (A). Side-chains of residues engaging in molecular interactions are shown in sticks and were labeled for one monomer only, except for residues 38 and 39 that were labeled in the symmetry-related molecule (marked by a ') as no density was observed in the other chain.

For residue T76 the backbone atoms are shown in sticks as well as they are engaging in mainchain-mainchain interactions. All hydrogen bonds are colored in grey, while van der Waals interactions are colored in magenta.

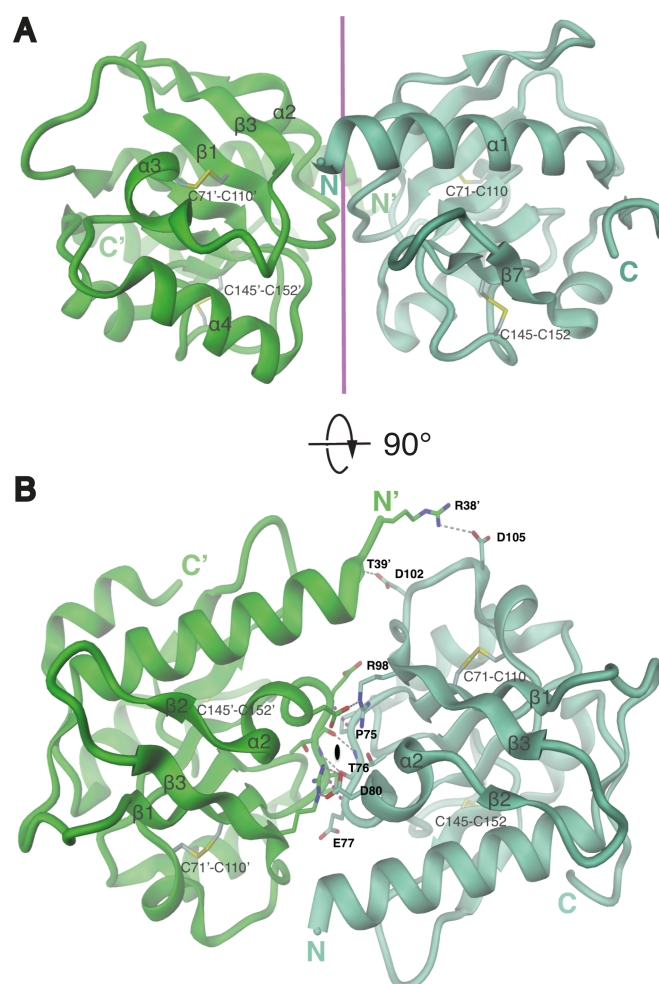
Figure 2

Figure 3:

A set of MALS data (weight-average molecule mass) acquired at various loading concentrations for wild-type YfiR (blue) and YfiR_{D80R} (red). The data were fitted with a dynamic monomer-dimer model in fast exchange, with the values of the pure species set to their nominal values (horizontal lines).

A SEC-chromatogram of YfiR_{D80R} at different loading concentrations (green, blue, red) is shown as inset. Continuous lines represent the dRI signal (left axis), broken lines the MALS derived apparent mass values (right axis).

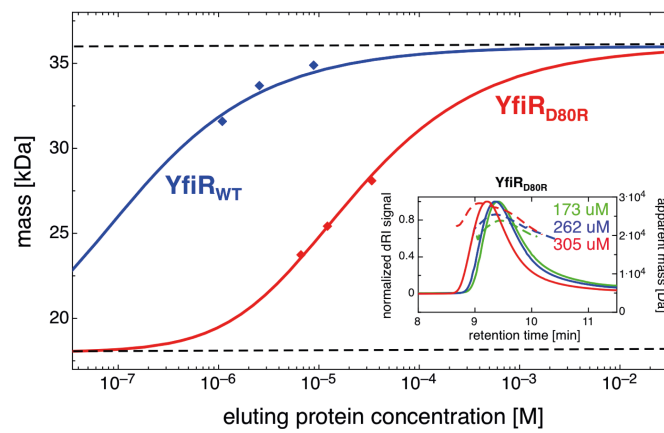
Figure 3

Figure 4:

(A) Sequence of YfiR from *P. aeruginosa* (UNIPROT accession name Q9I4L4) with HMM logo based on a non-redundant set of **59** homologous DUF4154 protein sequences. Only DUF4154 proteins were chosen that occur together with YfiN and YfiB as neighboring genes in a radius of +/- 2 positions. Every tenth residue is labeled with its number. Residues that are highly conserved (class 8 and 9 according to ConSurf) and exposed (= / > than 40 % surface accessibility) were labeled according to their properties: charged amino acid (**•**), hydrophobic nature (**#**), polar (**◆**) and structural importance (**+**). E140, which stabilizes the dimerization loop was labeled with a (**★**). Residues E140, I169 and V176 had a surface accessibility that was lower than 40%, however they were still shown for completeness reasons. The secondary structure elements observed in the structure are shown schematically.

(B) The YfiR dimer was colored according to the conservation score of each residue, using the same scale and coloring scheme as ConSurf. The residues labeled in (A) are shown as sticks.

(C) Surface representation of (A), where the surface was colored as well according to the conservation of each residue.

(D) Same view as in (B) but rotated by 90 degrees around the local two-fold. Only monomer B is shown. All residues that are labeled in (A) are shown as sticks.

(E) Same view as in (C) but rotated by 100 degrees around the local two-fold.

(F) Same view as (E) but showing the hydrophobicity mapped onto the surface.

Figure 4

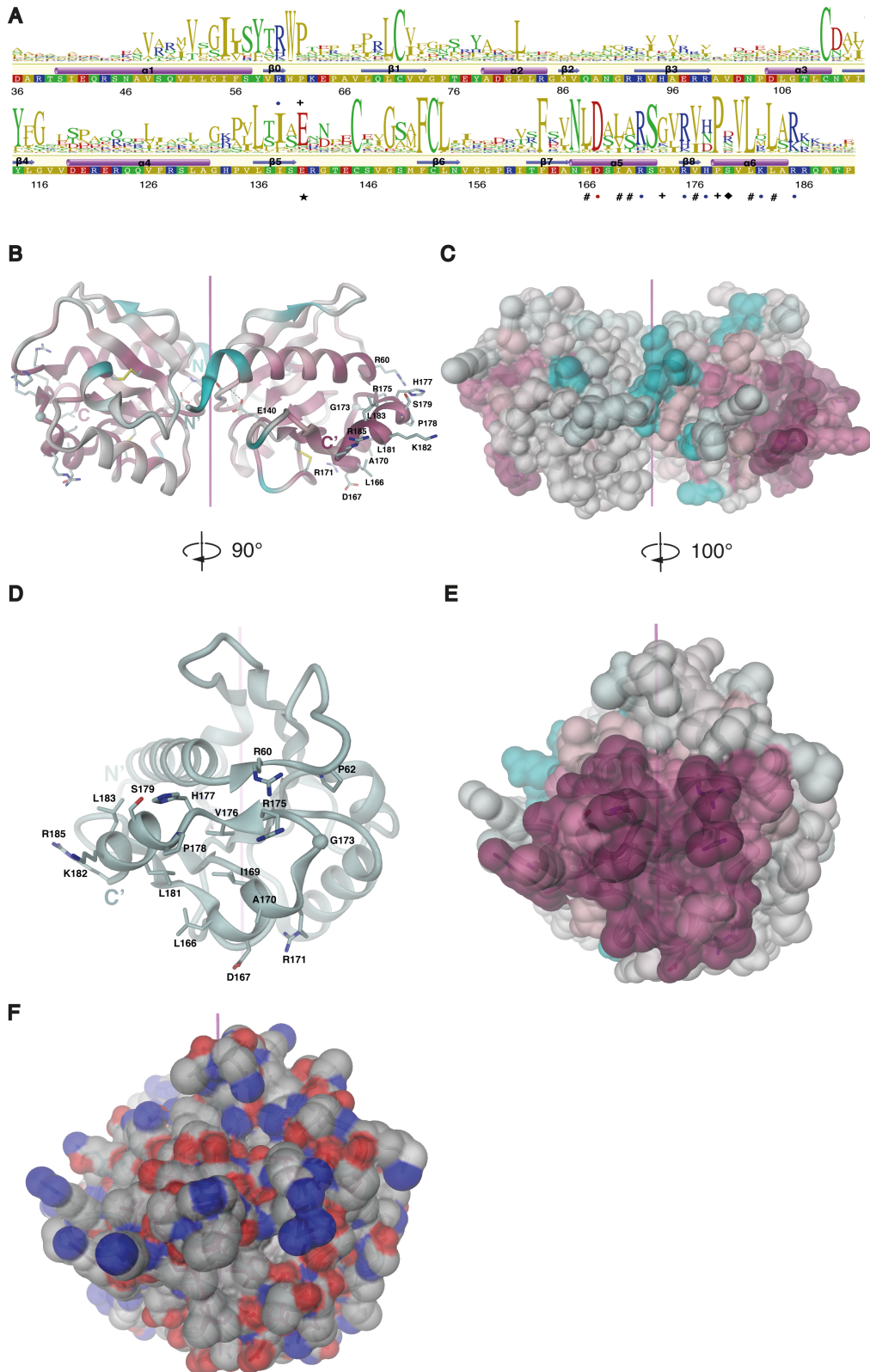


Figure 5:

The positions of previously identified activating YfiR mutations are shown within the structure [14]. The mutated residues are shown in sticks. The YfiN gain-of-function backgrounds in which the complementary YfiR mutations have been identified are noted in subscript (Table 3).

Figure 5

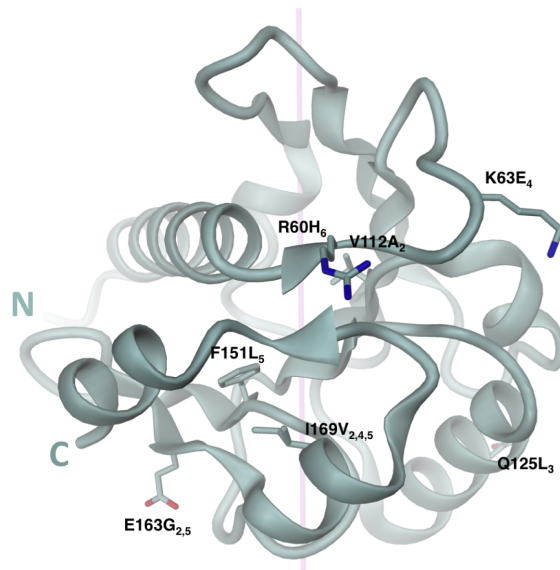
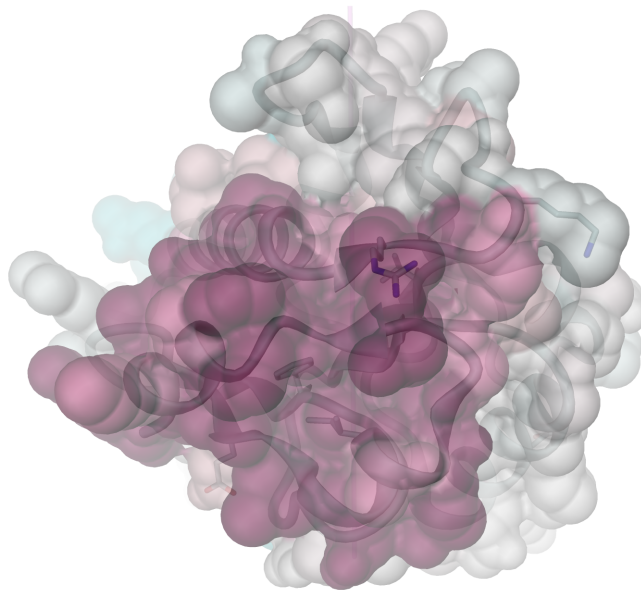
A**B**

Figure 6:

(A) Sequence of YfiN_{PAS} from *P. aeruginosa* (UNIPROT accession name Q9I4L5) with HMM logo based on a non-redundant set of 38 homologous YfiN sequences. Only YfiN proteins were chosen that occur together with DUF4154 proteins and YfiB as neighboring genes in a radius of +/- 2 positions. Every tenth residue is labeled with its number. Residues that are completely conserved (class 9 according to ConSurf) and exposed (= / > than 40 % surface accessibility) were labeled according to their properties: charged amino acid (•), hydrophobic nature (#), polar (◆) and structural importance (+). The secondary structure elements were predicted by Jpred and are shown schematically [140].

(B) A homology model of the YfiN_{PAS} dimer was generated based on the template CitA (PDB entry 1p0z). The two subunits are distinguished by color. The numbers noted at the N- and C-terminal end designate the terminal amino acid residues of the homology model.

(C) The YfiN_{PAS} dimer was colored according to the conservation score of each residue, using the same scale and coloring scheme as ConSurf. The residues labeled in (A) are shown as sticks.

(D) Same view as in (C) but rotated by 90 degrees around the local two-fold. Only monomer A is shown. All residues that are labeled in (A) are shown as sticks.

(E) Surface representation of (D), where the surface was colored according to the conservation of each residue.

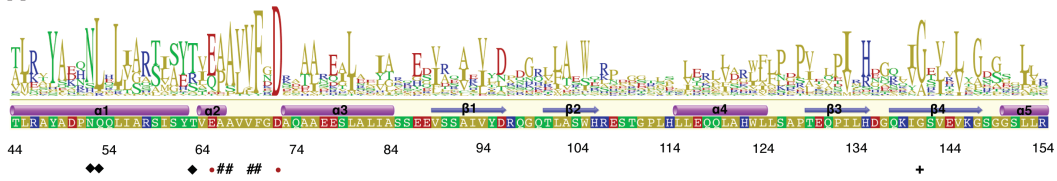
(F) Same view as in (D) but rotated by 20 degrees around x. The positions of the previously identified activating YfiN mutations are shown within the structure [14]. The mutated residues are shown in sticks.

(G) Surface representation of (F) where the surface was colored according to the conservation of each residue.

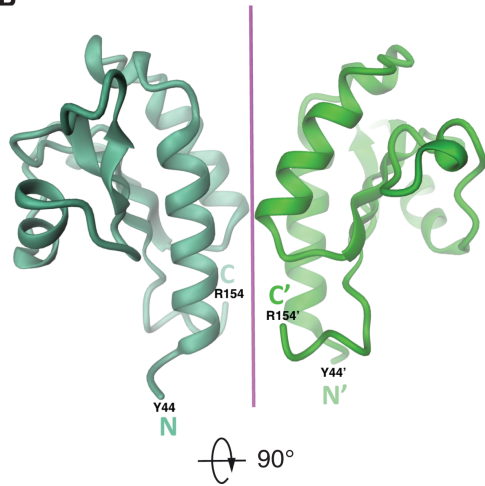
Results

Figure 6

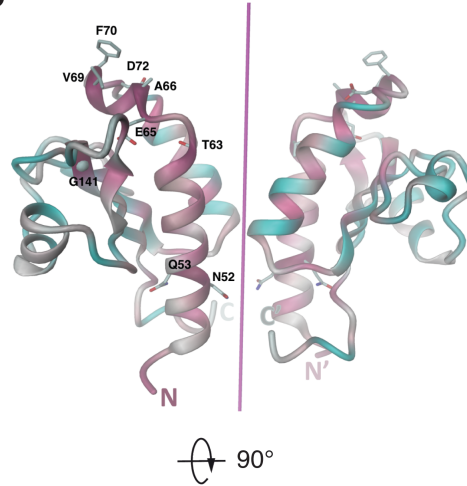
A



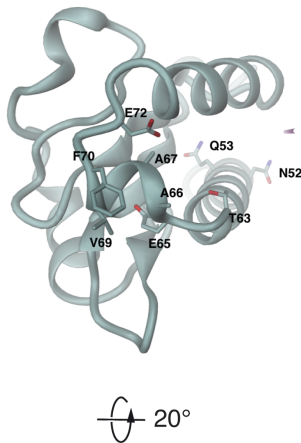
B



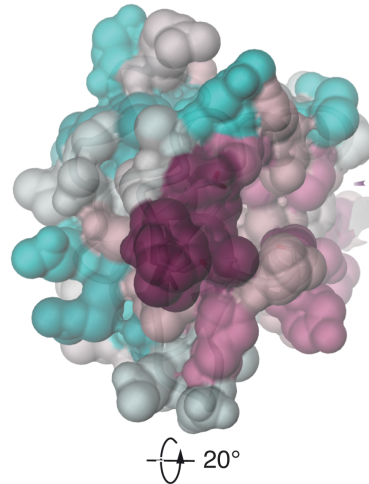
C



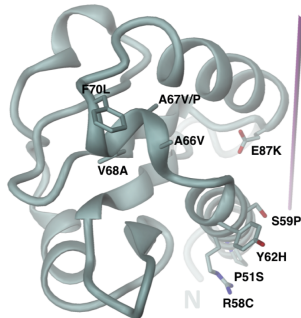
D



E



F



G

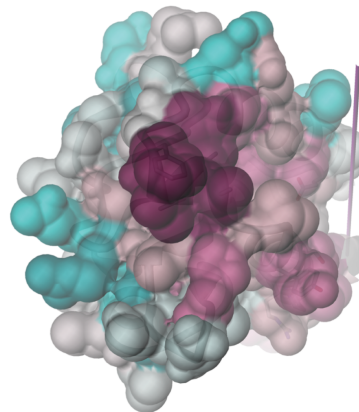


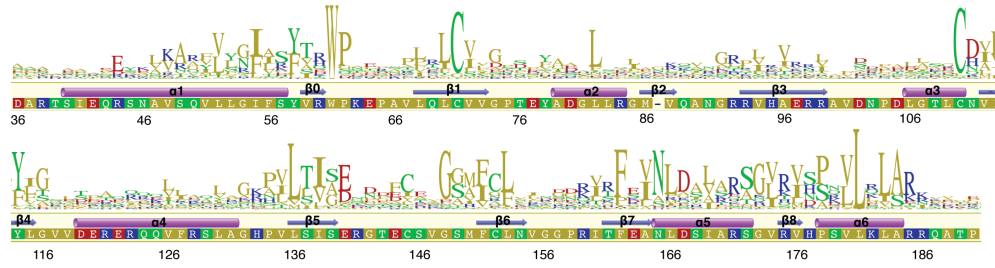
Figure 7

Multiple sequence alignment of DUF4154 proteins that occur with genes coding for either c-di-GMP related proteins (A), OmpA-like proteins (B), histidine kinases (C) or PLUG (D) that are localized up to two positions upstream or downstream of the DUF4154 proteins. Prior generation of the WebeLogos by Geneious [134], the datasets were subjected to blastclust (-L75 -b70) in order to reduce bias from overrepresented sequence sources. Every tenth residue is labeled with its number. The secondary structure elements observed in the structure are shown schematically.

Figure 7

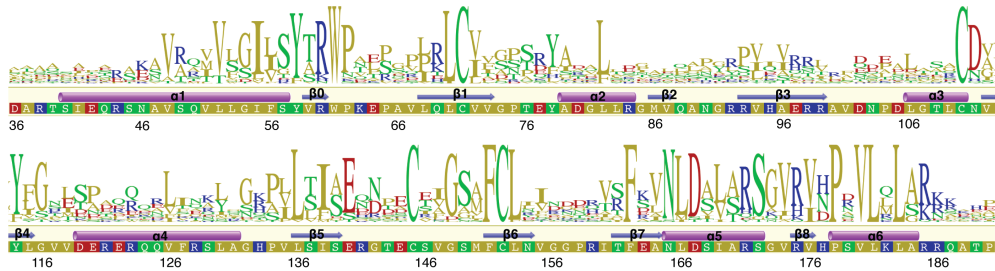
A

DUF4154 + cdG sequence logo



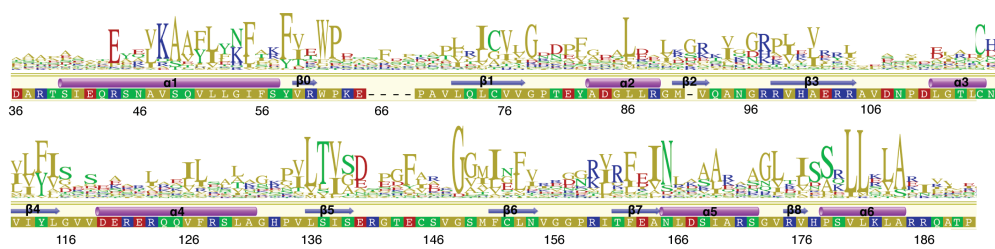
B

DUF4154 + OmpA sequence logo



C

DUF4154 + HK sequence logo



D

DUF4154 + PLUG sequence logo

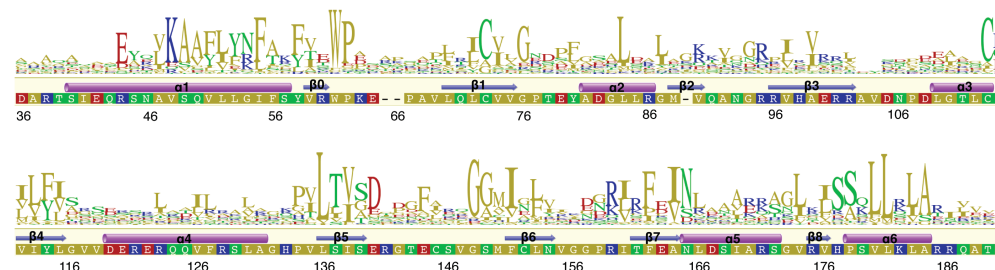


Figure 8:

The YfiBNR system from *P.aeruginosa*. The three proteins, the individual domains and the nucleotides are shown schematically. OM stands for “outer membrane”, PG for “peptidoglycan layer” and IM for “inner membrane”. YfiN constitutes the effector of the system as a diguanylate cyclase that is integral to the inner membrane. In the on state, the OmpA/Pal-like lipoprotein YfiB is binding to the periplasmic repressor protein YfiR, thereby sequestering it to the outer membrane and activating YfiN. YfiB also has a peptidoglycan-binding site, which was shown to be necessary for its function. In the off state, YfiB is undergoing a conformational change, thereby hiding its YfiR binding site, which results in the release of YfiR to the periplasm and YfiN inactivation.

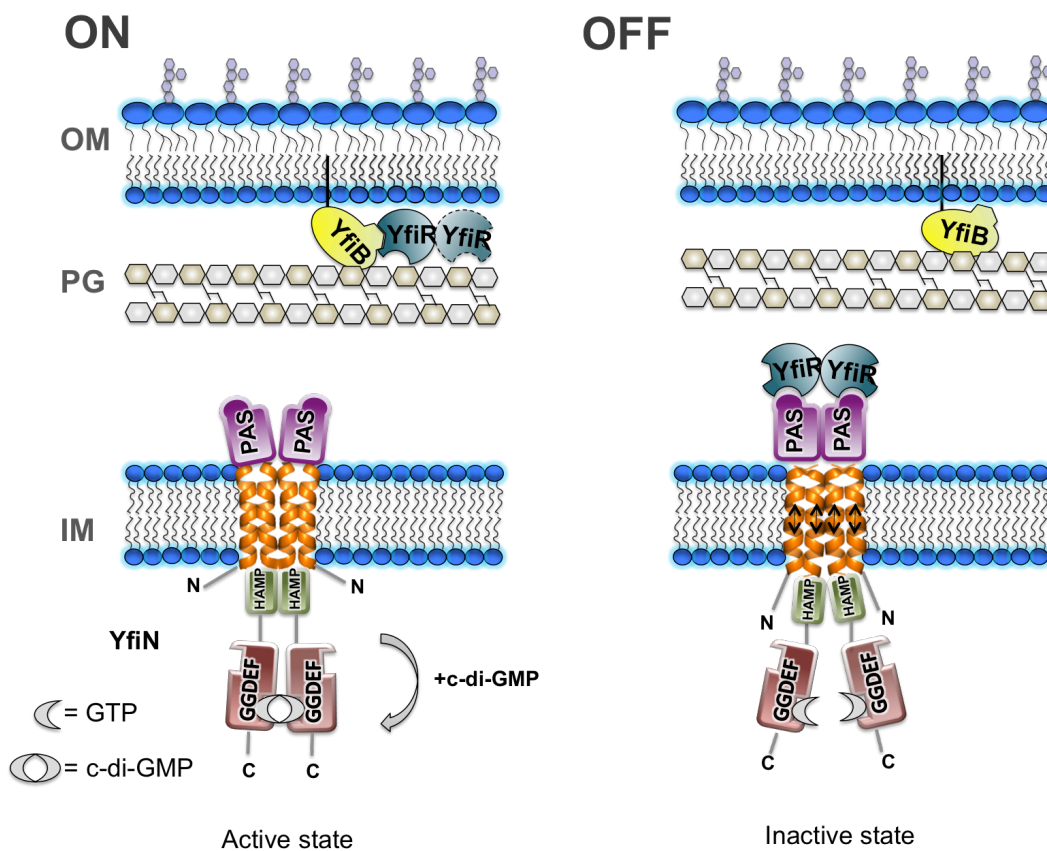


Table 1:

Table 1. Data collection and refinement statistics.	
yfiR-SeMet	
Wavelength [Å]	0.9791
Resolution range [Å]	47.0 - 1.8 (1.9 - 1.8)
Space group	P 1 21 1
Unit cell [Å, °]	69.7 68.1 72.0 90.0 111.4 90.0
Total reflections	346046 (15069)
Unique reflections	54458 (3992)
Multiplicity	6.4 (3.8)
Completeness [%]	94.0 (70.0)
Mean I/sigma(I)	17.1 (5.0)
Wilson B-factor	18.4
R-merge [%]	8.4 (21.4)
R-meas [%]	9.1 (24.7)
CC1/2 [%]	99.7 (94.6)
CC* [%]	99.9 (98.6)
Reflections used in refinement	54452 (3991)
Reflections used for R-free	2560 (183)
R-work [%]	17.3 (21.6)
R-free [%]	21.3 (28.3)
CC(work) [%]	96.0 (92.7)
CC(free) [%]	94.3 (87.0)
Number of non-hydrogen atoms	5017
macromolecules	4668
ligands	38
Protein residues	598
RMS(bonds) [Å]	0.011
RMS(angles) [°]	1.22
Ramachandran favored [%]	98
Ramachandran allowed [%]	2.2
Ramachandran outliers [%]	0
Rotamer outliers [%]	0.2
Clashscore	4.2
Average B-factor [Å ²]	27.7
macromolecules	27.5
ligands	29.7
solvent	30.5
Number of TLS groups	35
Statistics for the highest-resolution shell are shown in parentheses.	

Table 2

A list of all interactions found in the dimeric interface of YfiR. The distances as well as the type of interactions are indicated. mc stands for main chain, whereas sc denotes side chain.

Table 2

Subunit A	Subunit B	Distance [Å]	Type of interaction
PRO 75 [CA]	GLU 77 [CA]	4.1	van der Waals
GLU 77 [CA]	PRO 75 [CA]	4.3	van der Waals
THR 76 [N]	THR 76 [O]	3	mc-mc h-bond
THR 76 [O]	THR 76 [N]	3	mc-mc h-bond
ASP 80 [OD1]	ARG 98 [NH2]	3	electrostatic
ASP 80 [OD2]	ARG 98 [NE]	2.8	electrostatic
ARG 98 [NE]	ASP 80 [OD2]	2.9	electrostatic
ASP 102 [OD2]	THR 39 [OG1]	2.9	sc-sc h-bond
ASP 105 [OD2]	ARG 38 [NH1]	3.5	electrostatic

Table 3

A list of all the compensatory mutations in YfiR with their corresponding conservation score and the fraction of solvent-accessible area noted. The preceding activating mutations in YfiN are marked in the second column.

Table 3

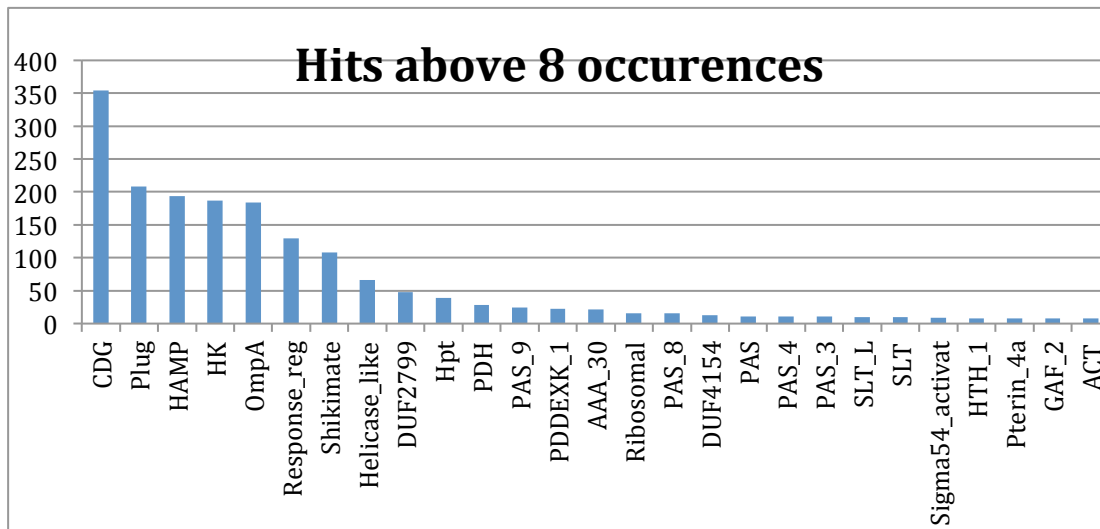
Mutant nr.	YfiN allele	Compensatory YfiR residue change	Conservation score of mutated residue	Fraction of solvent-accessible area
1	A67P	Q187R	5	-
2	V68A	I169V	9	0.12
		E163G	5	0.51
		Q187R	5	-
		V112A	6	0
3	G162S	Q125L	1	0.27
4	D204N	K63E	6	0.91
		I169V	9	0.12
5	A226T	E163G	5	0.51
		F151L	9	0
		I169V	9	0.12
6	L228F	R60H	9	0.46

Supplementary tables

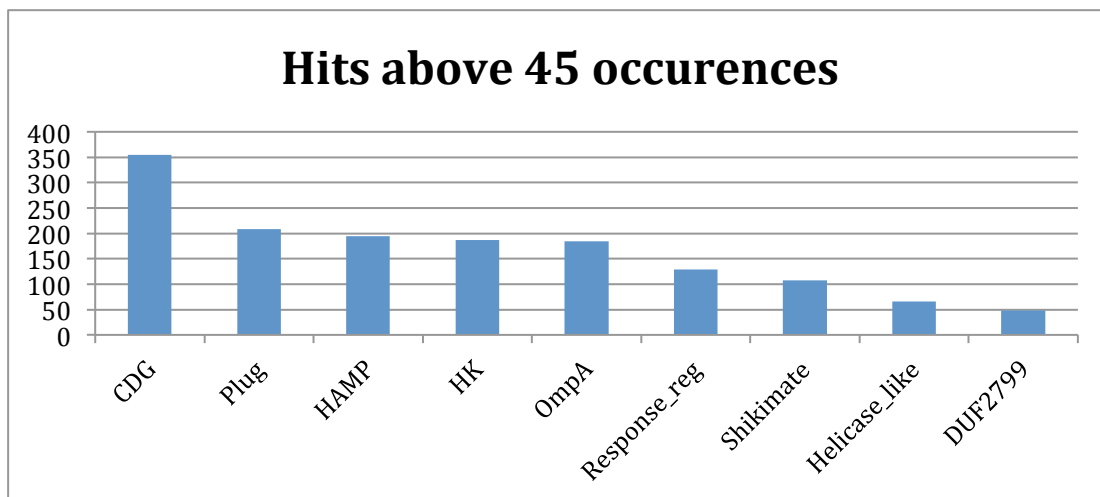
Table S1

A table illustrating all the proteins, which are found 2 positions upstream or downstream of the 1721 DUF4154 members. Domains and proteins occurring 8 times or more are indicated in (A), whereas in (B) only hits that are occurring 45 times or more are shown.

A



B



References

- [1] U. Römling, M. Gomelsky, and M. Y. Galperin, “C-di-GMP: The dawning of a novel bacterial signalling system,” *Molecular Microbiology*, vol. 57, pp. 629–639, 2005.
- [2] R. D. Monds, P. D. Newell, R. H. Gross, and G. A. O’Toole, “Phosphate-dependent modulation of c-di-GMP levels regulates *Pseudomonas fluorescens* Pf0-1 biofilm formation by controlling secretion of the adhesin LapA,” *Mol. Microbiol.*, vol. 63, pp. 656–679, 2007.
- [3] P. D. Newell, C. D. Boyd, H. Sondermann, and G. A. O’Toole, “A c-di-GMP effector system controls cell adhesion by inside-out signaling and surface protein cleavage,” *PLoS Biol.*, vol. 9, 2011.
- [4] O. Kirillina, J. D. Fetherston, A. G. Bobrov, J. Abney, and R. D. Perry, “HmsP, a putative phosphodiesterase, and HmsT, a putative diguanylate cyclase, control Hms-dependent biofilm formation in *Yersinia pestis*,” *Mol. Microbiol.*, vol. 54, pp. 75–88, 2004.
- [5] M. Merighi, V. T. Lee, M. Hyodo, Y. Hayakawa, and S. Lory, “The second messenger bis-(3’-5’)-cyclic-GMP and its PilZ domain-containing receptor Alg44 are required for alginate biosynthesis in *Pseudomonas aeruginosa*,” *Mol. Microbiol.*, vol. 65, pp. 876–895, 2007.
- [6] V. T. Lee, J. M. Matewish, J. L. Kessler, M. Hyodo, Y. Hayakawa, and S. Lory, “A cyclic-di-GMP receptor required for bacterial exopolysaccharide production,” *Mol. Microbiol.*, vol. 65, pp. 1474–1484, 2007.
- [7] D. A. D’Argenio, M. W. Calfee, P. B. Rainey, and E. C. Pesci, “Autolysis and autoaggregation in *Pseudomonas aeruginosa* colony morphology mutants,” *J. Bacteriol.*, vol. 184, pp. 6481–6489, 2002.
- [8] R. Simm, M. Morr, A. Kader, M. Nimtz, and U. Römling, “GGDEF and EAL domains inversely regulate cyclic di-GMP levels and transition from sessility to motility,” *Mol. Microbiol.*, vol. 53, pp. 1123–1134, 2004.
- [9] A. Nakhamchik, C. Wilde, and D. A. Rowe-Magnus, “Cyclic-di-GMP regulates extracellular polysaccharide production, biofilm formation, and rugose colony development by *Vibrio vulnificus*,” *Appl. Environ.*

- Microbiol.*, vol. 74, pp. 4199–4209, 2008.
- [10] H. Kulasakara, V. Lee, A. Brencic, N. Liberati, J. Urbach, S. Miyata, D. G. Lee, A. N. Neely, M. Hyodo, Y. Hayakawa, F. M. Ausubel, and S. Lory, “Analysis of *Pseudomonas aeruginosa* diguanylate cyclases and phosphodiesterases reveals a role for bis-(3'-5')-cyclic-GMP in virulence,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 103, pp. 2839–2844, 2006.
- [11] A. G. Bobrov, O. Kirillina, D. A. Ryjenkov, C. M. Waters, P. A. Price, J. D. Fetherston, D. Mack, W. E. Goldman, M. Gomelsky, and R. D. Perry, “Systematic analysis of cyclic di-GMP signalling enzymes and their role in biofilm formation and virulence in *Yersinia pestis*,” *Mol. Microbiol.*, vol. 79, pp. 533–551, 2011.
- [12] J. G. Malone, T. Jaeger, C. Spangler, D. Ritz, A. Spang, C. Arrieumerlou, V. Kaever, R. Landmann, and U. Jenal, “YfiBNR mediates cyclic di-GMP dependent small colony variant formation and persistence in *Pseudomonas aeruginosa*,” *PLoS Pathog.*, vol. 6, 2010.
- [13] S. Haussler, B. Tummler, H. Weissbrodt, M. Rohde, and I. Steinmetz, “Small-colony variants of *Pseudomonas aeruginosa* in cystic fibrosis,” *Clin Infect Dis*, vol. 29, pp. 621–625, 1999.
- [14] J. G. Malone, T. Jaeger, P. Manfredi, A. Dötsch, A. Blanka, R. Bos, G. R. Cornelis, S. Häussler, and U. Jenal, “The YfiBNR signal transduction mechanism reveals novel targets for the evolution of persistent *Pseudomonas aeruginosa* in cystic fibrosis airways,” *PLoS Pathog.*, vol. 8, 2012.
- [15] J. G. Leid, C. J. Willson, M. E. Shirtliff, D. J. Hassett, M. R. Parsek, and A. K. Jeffers, “The exopolysaccharide alginate protects *Pseudomonas aeruginosa* biofilm bacteria from IFN-gamma-mediated macrophage killing,” *J. Immunol.*, vol. 175, pp. 7512–7518, 2005.
- [16] A. L. Spoering and K. Lewis, “Biofilms and planktonic cells of *Pseudomonas aeruginosa* have similar resistance to killing by antimicrobials,” *J. Bacteriol.*, vol. 183, pp. 6746–6751, 2001.
- [17] S. Häussler, C. Lehmann, C. Breselge, M. Rohde, M. Classen, B. Tümmeler, P. Vandamme, and I. Steinmetz, “Fatal outcome of lung transplantation in cystic fibrosis patients due to small-colony variants of the *Burkholderia cepacia* complex,” *Eur. J. Clin. Microbiol. Infect. Dis.*, vol. 22, pp. 249–253, 2003.

- [18] M. J. Blommers, C. A. Haasnoot, J. A. Walters, G. A. van der Marel, J. H. van Boom, and C. W. Hilbers, "Solution structure of the 3'-5' cyclic dinucleotide d(pApA). A combined NMR, UV melting, and molecular mechanics study.," *Biochemistry*, vol. 27, pp. 8361–8369, 1988.
- [19] C. A. Frederick, M. Coll, G. A. van der Marel, J. H. van Boom, and A. H. Wang, "Molecular structure of cyclic deoxydiadenylic acid at atomic resolution.," *Biochemistry*, vol. 27, pp. 8350–8361, 1988.
- [20] M. Egli, R. V Gessner, L. D. Williams, G. J. Quigley, G. A. van der Marel, J. H. van Boom, A. Rich, and C. A. Frederick, "Atomic-resolution structure of the cellulose synthase regulator cyclic diguanylic acid.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 87, pp. 3235–3239, 1990.
- [21] Y. Guan, Y. G. Gao, Y. C. Liaw, H. Robinson, and A. H. J. Wang, "Molecular-Structure of Cyclic Diguanylic Acid at 1 Angstrom Resolution of 2 Crystal Forms - Self-Association, Interactions with Metal Ion/Planar Dyes and Modeling Studies," *J. Biomol. Struct. Dyn.*, vol. 11, pp. 253–276, 1993.
- [22] Y. C. Liaw, Y. G. Gao, H. Robinson, G. M. Sheldrick, L. A. Sliedregt, G. A. van der Marel, J. H. van Boom, and A. H. Wang, "Cyclic diguanylic acid behaves as a host molecule for planar intercalators.," *FEBS Lett.*, vol. 264, pp. 223–227, 1990.
- [23] C. Chan, R. Paul, D. Samoray, N. C. Amiot, B. Giese, U. Jenal, and T. Schirmer, "Structural basis of activity and allosteric control of diguanylate cyclase.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 101, pp. 17084–17089, 2004.
- [24] P. Wassmann, C. Chan, R. Paul, A. Beck, H. Heerklotz, U. Jenal, and T. Schirmer, "Structure of BeF₃-Modified Response Regulator PleD: Implications for Diguanylate Cyclase Activation, Catalysis, and Feedback Inhibition," *Structure*, vol. 15, pp. 915–927, 2007.
- [25] N. De, M. Pirruccello, P. V. Krasteva, N. Bae, R. V. Raghavan, and H. Sondermann, "Phosphorylation-independent regulation of the diguanylate cyclase WspR," *PLoS Biol.*, vol. 6, pp. 0601–0617, 2008.
- [26] N. De, M. V. A. S. Navarro, R. V. Raghavan, and H. Sondermann, "Determinants for the Activation and Autoinhibition of the Diguanylate Cyclase Response Regulator WspR," *J. Mol. Biol.*, vol. 393, pp. 619–633, 2009.

-
- [27] D. A. Ryjenkov, R. Simm, U. Römling, and M. Gomelsky, “The PilZ domain is a receptor for the second messenger c-di-GMP: The PilZ domain protein YcgR controls motility in enterobacteria,” *J. Biol. Chem.*, vol. 281, pp. 30310–30314, 2006.
- [28] J. Ko, K. S. Ryu, H. Kim, J. S. Shin, J. O. Lee, C. Cheong, and B. S. Choi, “Structure of PP4397 reveals the molecular basis for different c-di-GMP binding modes by pilz domain proteins,” *J. Mol. Biol.*, vol. 398, pp. 97–110, 2010.
- [29] J. Habazettl, M. G. Allan, U. Jenal, and S. Grzesiek, “Solution structure of the PilZ domain protein PA4608 complex with cyclic di-GMP identifies charge clustering as molecular readout,” *J. Biol. Chem.*, vol. 286, pp. 14304–14314, 2011.
- [30] J. S. Shin, K. S. Ryu, J. Ko, A. Lee, and B. S. Choi, “Structural characterization reveals that a PilZ domain protein undergoes substantial conformational change upon binding to cyclic dimeric guanosine monophosphate,” *Protein Sci.*, vol. 20, pp. 270–277, 2011.
- [31] P. V Krasteva, J. C. N. Fong, N. J. Shikuma, S. Beyhan, M. V. A. S. Navarro, F. H. Yildiz, and H. Sondermann, “*Vibrio cholerae* VpsT regulates matrix production and motility by directly sensing cyclic di-GMP,” *Science*, vol. 327, pp. 866–868, 2010.
- [32] K. D. Smith, S. V Lipchock, T. D. Ames, J. Wang, R. R. Breaker, and S. A. Strobel, “Structural basis of ligand binding by a c-di-GMP riboswitch,” *Nat. Struct. Mol. Biol.*, vol. 16, pp. 1218–1223, 2009.
- [33] C.-Y. Yang, K.-H. Chin, M. L.-C. Chuah, Z.-X. Liang, A. H.-J. Wang, and S.-H. Chou, “The structure and inhibition of a GGDEF diguanylate cyclase complexed with (c-di-GMP)₂ at the active site,” *Acta Crystallogr. D. Biol. Crystallogr.*, vol. 67, no. Pt 12, pp. 997–1008, Dec. 2011.
- [34] G. Minasov, S. Padavattan, L. Shuvalova, J. S. Brunzelle, D. J. Miller, A. Baslé, C. Massa, F. R. Collart, T. Schirmer, and W. F. Anderson, “Crystal structures of Ykul and its complex with second messenger cyclic Di-GMP suggest catalytic mechanism of phosphodiester bond cleavage by EAL domains,” *J. Biol. Chem.*, vol. 284, pp. 13174–13184, 2009.
- [35] T. R. Barends, E. Hartmann, J. J. Griese, T. Beitlich, N. V Kirienko, D. A. Ryjenkov, J. Reinstein, R. L. Shoeman, M. Gomelsky, and I. Schlichting, “Structure and mechanism of a bacterial light-regulated

- cyclic nucleotide phosphodiesterase,” *Nature*, vol. 459, pp. 1015–1018, 2009.
- [36] A. Sundriyal, C. Massa, D. Samoray, F. Zehender, T. Sharpe, U. Jenal, and T. Schirmer, “Inherent regulation of EAL domain-catalyzed hydrolysis of second messenger cyclic di-GMP,” *J. Biol. Chem.*, vol. 289, pp. 6978–6990, 2014.
- [37] J. Benach, S. S. Swaminathan, R. Tamayo, S. K. Handelman, E. Folta-Stogniew, J. E. Ramos, F. Forouhar, H. Neely, J. Seetharaman, A. Camilli, and J. F. Hunt, “The structural basis of cyclic diguanylate signal transduction by PilZ domains,” *EMBO J.*, vol. 26, pp. 5153–5166, 2007.
- [38] D. L. Burdette, K. M. Monroe, K. Sotelo-Troha, J. S. Iwig, B. Eckert, M. Hyodo, Y. Hayakawa, and R. E. Vance, “STING is a direct innate immune sensor of cyclic di-GMP,” *Nature*, vol. 478, pp. 515–518, 2011.
- [39] N. Tschowri, M. A. Schumacher, S. Schlimpert, N. B. Chinnam, K. C. Findlay, R. G. Brennan, and M. J. Buttner, “Tetrameric c-di-GMP mediates effective transcription factor dimerization to control *Streptomyces* development,” *Cell*, vol. 158, no. 5, pp. 1136–1147, Aug. 2014.
- [40] Z. Zhang, B. L. Gaffney, and R. A. Jones, “c-di-GMP displays A monovalent metal ion-dependent polymorphism,” *J. Am. Chem. Soc.*, vol. 126, pp. 16700–16701, 2004.
- [41] Z. Zhang, S. Kim, B. L. Gaffney, and R. A. Jones, “Polymorphism of the signaling molecule c-di-GMP,” *J. Am. Chem. Soc.*, vol. 128, pp. 7015–7024, 2006.
- [42] M. Gentner, M. G. Allan, F. Zaehring, T. Schirmer, and S. Grzesiek, “Oligomer formation of the bacterial second messenger c-di-GMP: Reaction rates and equilibrium constants indicate a monomeric state at physiological concentrations,” *J. Am. Chem. Soc.*, vol. 134, pp. 1019–1029, 2012.
- [43] R. Paul, S. Weiser, N. C. Amiot, C. Chan, T. Schirmer, B. Giese, and U. Jenal, “Cell cycle-dependent dynamic localization of a bacterial response regulator with a novel di-guanylate cyclase output domain,” *Genes Dev.*, vol. 18, pp. 715–727, 2004.
- [44] J. G. Malone, R. Williams, M. Christen, U. Jenal, a. J. Spiers, and P. B. Rainey, “The structure-function relationship of WspR, a

- Pseudomonas fluorescens* response regulator with a GGDEF output domain,” *Microbiology*, vol. 153, pp. 980–994, 2007.
- [45] J. W. Hickman, D. F. Tifrea, and C. S. Harwood, “A chemosensory system that regulates biofilm formation through modulation of cyclic diguanylate levels,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 102, pp. 14422–14427, 2005.
- [46] D. A. Ryjenkov, M. Tarutina, O. V. Moskvina, and M. Gomelsky, “Cyclic diguanylate is a ubiquitous signaling molecule in bacteria: Insights into biochemistry of the GGDEF protein domain,” *J. Bacteriol.*, vol. 187, pp. 1792–1798, 2005.
- [47] B. Christen, M. Christen, R. Paul, F. Schmid, M. Folcher, P. Jenoe, M. Meuwly, and U. Jenal, “Allosteric control of cyclic di-GMP signaling,” *J. Biol. Chem.*, vol. 281, pp. 32015–32024, 2006.
- [48] M. Christen, B. Christen, M. Folcher, A. Schauerte, and U. Jenal, “Identification and characterization of a cyclic di-GMP-specific phosphodiesterase and its allosteric control by GTP,” *J. Biol. Chem.*, vol. 280, pp. 30829–30837, 2005.
- [49] R. Tamayo, A. D. Tischler, and A. Camilli, “The EAL domain protein VieA is a cyclic diguanylate phosphodiesterase,” *J. Biol. Chem.*, vol. 280, pp. 33324–33330, 2005.
- [50] A. J. Schmidt, D. A. Ryjenkov, and M. Gomelsky, “The ubiquitous protein domain EAL is a cyclic diguanylate-specific phosphodiesterase: Enzymatically active and inactive EAL domains,” *J. Bacteriol.*, vol. 187, pp. 4774–4781, 2005.
- [51] F. Rao, Y. Yang, Y. Qi, and Z. X. Liang, “Catalytic mechanism of cyclic di-GMP-specific phosphodiesterase: A study of the EAL domain-containing RocR from *Pseudomonas aeruginosa*,” *J. Bacteriol.*, vol. 190, pp. 3622–3631, 2008.
- [52] M. Y. Galperin, D. A. Natale, L. Aravind, and E. V. Koonin, “A specialized version of the HD hydrolase domain implicated in signal transduction,” *Journal of molecular microbiology and biotechnology*, vol. 1, pp. 303–305, 1999.
- [53] A. S. N. Seshasayee, G. M. Fraser, and N. M. Luscombe, “Comparative genomics of cyclic-di-GMP signalling in bacteria: Post-translational regulation and catalytic activity,” *Nucleic Acids Res.*, vol. 38, pp. 5970–5981, 2010.

- [54] B. R. Boles and L. L. McCarter, “Vibrio parahaemolyticus scrABC, a novel operon affecting swarming and capsular polysaccharide regulation,” *J. Bacteriol.*, vol. 184, pp. 5946–5954, 2002.
- [55] M. Tarutina, D. A. Ryjenkov, and M. Gomelsky, “An unorthodox bacteriophytochrome from *Rhodobacter sphaeroides* involved in turnover of the second messenger c-di-GMP,” *J. Biol. Chem.*, vol. 281, pp. 34751–34758, 2006.
- [56] R. B. R. Ferreira, L. C. M. Antunes, E. P. Greenberg, and L. L. McCarter, “Vibrio parahaemolyticus ScrC modulates cyclic dimeric GMP regulation of gene expression relevant to growth on surfaces,” *J. Bacteriol.*, vol. 190, pp. 851–860, 2008.
- [57] H. Weber, C. Pesavento, A. Possling, G. Tischendorf, and R. Hengge, “Cyclic-di-GMP-mediated signalling within the sigma network of *Escherichia coli*,” *Mol. Microbiol.*, vol. 62, pp. 1014–1034, 2006.
- [58] R. Tamayo, S. Schild, J. T. Pratt, and A. Camilli, “Role of cyclic Di-GMP during El Tor biotype *Vibrio cholerae* infection: Characterization of the in vivo-induced cyclic Di-GMP phosphodiesterase CdpA,” *Infect. Immun.*, vol. 76, pp. 1617–1627, 2008.
- [59] B. I. Kazmierczak, M. B. Lebron, and T. S. Murray, “Analysis of FimX, a phosphodiesterase that governs twitching motility in *Pseudomonas aeruginosa*,” *Mol. Microbiol.*, vol. 60, pp. 1026–1043, 2006.
- [60] M. V. A. S. Navarro, N. De, N. Bae, Q. Wang, and H. Sondermann, “Structural Analysis of the GGDEF-EAL Domain-Containing c-di-GMP Receptor FimX,” *Structure*, vol. 17, pp. 1104–1116, 2009.
- [61] Y. Qi, M. L. C. Chuah, X. Dong, K. Xie, Z. Luo, K. Tang, and Z. X. Liang, “Binding of cyclic diguanylate in the non-catalytic EAL domain of FimX induces a long-range conformational change,” *J. Biol. Chem.*, vol. 286, pp. 2910–2917, 2011.
- [62] M. Y. Galperin, A. N. Nikolskaya, and E. V. Koonin, “Novel domains of the prokaryotic two-component signal transduction systems,” *FEMS Microbiology Letters*, vol. 203, pp. 11–21, 2001.
- [63] M. Y. Galperin, “Bacterial signal transduction network in a genomic perspective,” *Environmental Microbiology*, vol. 6, pp. 552–567, 2004.
- [64] J. Lee, D. R. Tomchick, C. A. Brautigam, M. Machius, R. Kort, K. J.

- Hellingwerf, and K. H. Gardner, "Changes at the KinA PAS-A dimerization interface influence histidine kinase function," *Biochemistry*, vol. 47, pp. 4051–4064, 2008.
- [65] X. Ma, N. Sayed, P. Baskaran, A. Beuve, and F. Van Den Akker, "PAS-mediated dimerization of soluble guanylyl cyclase revealed by signal transduction histidine kinase domain crystal structure," *J. Biol. Chem.*, vol. 283, pp. 1167–1178, 2008.
- [66] Y. Oka, T. Matsushita, N. Mochizuki, P. H. Quail, and A. Nagatani, "Mutant screen distinguishes between residues necessary for light-signal perception and signal transfer by phytochrome B," *PLoS Genet.*, vol. 4, 2008.
- [67] B. L. Taylor and I. B. Zhulin, "PAS domains: internal sensors of oxygen, redox potential, and light.," *Microbiol. Mol. Biol. Rev.*, vol. 63, pp. 479–506, 1999.
- [68] S. Reinelt, E. Hofmann, T. Gerharz, M. Bott, and D. R. Madden, "The structure of the periplasmic ligand-binding domain of the sensor kinase CitA reveals the first extracellular pas domain," *J. Biol. Chem.*, vol. 278, pp. 39189–39196, 2003.
- [69] Y. Qi, F. Rao, Z. Luo, and Z. X. Liang, "A flavin cofactor-binding PAS domain regulates c-di-GMP synthesis in AxDGC2 from *Acetobacter xylinum*," *Biochemistry*, vol. 48, pp. 10275–10285, 2009.
- [70] E. B. Purcell, C. A. McDonald, B. A. Palfey, and S. Crosson, "An analysis of the solution structure and signaling mechanism of LovK, a sensor histidine kinase integrating light and redox signals," *Biochemistry*, vol. 49, pp. 6761–6770, 2010.
- [71] E. B. Purcell, D. Siegal-Gaskins, D. C. Rawling, A. Fiebig, and S. Crosson, "A photosensory two-component system regulates bacterial cell attachment.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 104, pp. 18241–18246, 2007.
- [72] T. E. Swartz, T.-S. Tseng, M. A. Frederickson, G. Paris, D. J. Comerici, G. Rajashekara, J.-G. Kim, M. B. Mudgett, G. A. Splitter, R. A. Ugalde, F. A. Goldbaum, W. R. Briggs, and R. A. Bogomolni, "Blue-light-activated histidine kinases: two-component sensors in bacteria.," *Science*, vol. 317, pp. 1090–1093, 2007.
- [73] J. R. Tuckerman, G. Gonzalez, E. H. S. Sousa, X. Wan, J. A. Saito, M. Alam, and M. A. Gilles-Gonzalez, "An oxygen-sensing diguanylate cyclase and phosphodiesterase couple for c-di-GMP control,"

- Biochemistry*, vol. 48, pp. 9764–9774, 2009.
- [74] A. Möglich, R. A. Ayers, and K. Moffat, “Structure and Signaling Mechanism of Per-ARNT-Sim Domains,” *Structure*, vol. 17, pp. 1282–1294, 2009.
- [75] J. Cheung, C. A. Bingman, M. Reyngold, W. A. Hendrickson, and C. D. Waldburger, “Crystal structure of a functional dimer of the PhoQ sensor domain,” *J. Biol. Chem.*, vol. 283, pp. 13762–13770, 2008.
- [76] J. Cheung and W. A. Hendrickson, “Sensor domains of two-component regulatory systems,” *Current Opinion in Microbiology*, vol. 13, pp. 116–123, 2010.
- [77] J. Cheung and W. A. Hendrickson, “Crystal structures of C4-dicarboxylate ligand complexes with sensor domains of histidine kinases DcuS and DctB,” *J. Biol. Chem.*, vol. 283, pp. 30256–30265, 2008.
- [78] S. B. Williams and V. Stewart, “Functional similarities among two-component sensors and methyl-accepting chemotaxis proteins suggest a role for linker region amphipathic helices in transmembrane signal transduction,” *Mol. Microbiol.*, vol. 33, pp. 1093–1102, 1999.
- [79] M. V. A. S. Navarro, P. D. Newell, P. V. Krasteva, D. Chatterjee, D. R. Madden, G. A. O’Toole, and H. Sondermann, “Structural basis for c-di-GMP-mediated inside-out signaling controlling periplasmic proteolysis,” *PLoS Biol.*, vol. 9, 2011.
- [80] H. Park and M. Inouye, “Mutational analysis of the linker region of EnvZ, an osmosensor in *Escherichia coli*,” *J. Bacteriol.*, vol. 179, pp. 4382–4390, 1997.
- [81] M. Hulko, F. Berndt, M. Gruber, J. U. Linder, V. Truffault, A. Schultz, J. Martin, J. E. Schultz, A. N. Lupas, and M. Coles, “The HAMP Domain Structure Implies Helix Rotation in Transmembrane Signaling,” *Cell*, vol. 126, pp. 929–940, 2006.
- [82] M. V. Airola, K. J. Watts, A. M. Bilwes, and B. R. Crane, “Structure of Concatenated HAMP Domains Provides a Mechanism for Signal Transduction,” *Structure*, vol. 18, pp. 436–448, 2010.
- [83] R. Kishii, L. Falzon, T. Yoshida, H. Kobayashi, and M. Inouye, “Structural and functional studies of the HAMP domain of EnvZ, an osmosensing transmembrane histidine kinase in *Escherichia coli*,” *J. Biol. Chem.*, vol. 282, pp. 26401–26408, 2007.

-
- [84] K. E. Swain and J. J. Falke, "Structure of the conserved HAMP domain in an intact, membrane-bound chemoreceptor: A disulfide mapping study," *Biochemistry*, vol. 46, pp. 13684–13695, 2007.
- [85] A. N. Lupas and M. Gruber, "The structure of alpha-helical coiled coils.," *Adv. Protein Chem.*, vol. 70, pp. 37–78, 2005.
- [86] H. U. Ferris, S. Dunin-Horkawicz, N. Hornig, M. Hulko, J. Martin, J. E. Schultz, K. Zeth, A. N. Lupas, and M. Coles, "Mechanism of regulation of receptor histidine kinases," *Structure*, vol. 20, pp. 56–66, 2012.
- [87] H. U. Ferris, S. Dunin-Horkawicz, L. G. Mondéjar, M. Hulko, K. Hantke, J. Martin, J. E. Schultz, K. Zeth, A. N. Lupas, and M. Coles, "The mechanisms of HAMP-mediated signaling in transmembrane receptors," *Structure*, vol. 19, pp. 378–385, 2011.
- [88] A. L. Chang, J. R. Tuckerman, G. Gonzalez, R. Mayer, H. Weinhouse, G. Volman, D. Amikam, M. Benziman, and M. A. Gilles-Gonzalez, "Phosphodiesterase A1, a regulator of cellulose synthesis in *Acetobacter xylinum*, is a heme-based sensor," *Biochemistry*, vol. 40, pp. 3420–3426, 2001.
- [89] Z. Cao, E. Livoti, A. Losi, and W. Gärtner, "A blue light-inducible phosphodiesterase activity in the cyanobacterium *synechococcus elongatus*," *Photochem. Photobiol.*, vol. 86, pp. 606–611, 2010.
- [90] R. Paul, S. Abel, P. Wassmann, A. Beck, H. Heerklotz, and U. Jenal, "Activation of the diguanylate cyclase PleD by phosphorylation-mediated dimerization," *J. Biol. Chem.*, vol. 282, pp. 29170–29177, 2007.
- [91] T. A. Ramelot, A. Yee, J. R. Cort, A. Semesi, C. H. Arrowsmith, and M. A. Kennedy, "NMR structure and binding studies confirm that PA4608 from *Pseudomonas aeruginosa* is a PilZ domain and a c-di-GMP binding protein," *Proteins Struct. Funct. Genet.*, vol. 66, pp. 266–271, 2007.
- [92] J. T. Pratt, R. Tamayo, A. D. Tischler, and A. Camilli, "PilZ domain proteins bind cyclic diguanylate and regulate diverse processes in *Vibrio cholerae*," *J. Biol. Chem.*, vol. 282, pp. 12860–12870, 2007.
- [93] M. Christen, B. Christen, M. G. Allan, M. Folcher, P. Jenö, S. Grzesiek, and U. Jenal, "DgrA is a member of a new family of cyclic diguanosine monophosphate receptors and controls flagellar motor function in *Caulobacter crescentus*," *Proc. Natl. Acad. Sci. U. S. A.*,

- vol. 104, pp. 4112–4117, 2007.
- [94] D. Amikam and M. Y. Galperin, “PilZ domain is part of the bacterial c-di-GMP binding protein,” *Bioinformatics*, vol. 22, pp. 3–6, 2006.
- [95] Y. McCarthy, R. P. Ryan, K. O’Donovan, Y. Q. He, B. Le Jiang, J. X. Feng, J. L. Tang, and J. M. Dow, “The role of PilZ domain proteins in the virulence of *Xanthomonas campestris* pv. *campestris*,” *Mol. Plant Pathol.*, vol. 9, pp. 819–824, 2008.
- [96] A. Duerig, S. Abel, M. Folcher, M. Nicollier, T. Schwede, N. Amiot, B. Giese, and U. Jenal, “Second messenger-mediated spatiotemporal control of protein degradation regulates bacterial cell cycle progression,” *Genes Dev.*, vol. 23, pp. 93–104, 2009.
- [97] P. D. Newell, R. D. Monds, and G. A. O’Toole, “LapD is a bis-(3’,5’)-cyclic dimeric GMP-binding protein that regulates surface attachment by *Pseudomonas fluorescens* Pf0-1.,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 106, pp. 3461–3466, 2009.
- [98] D. Chatterjee, R. B. Cooley, C. D. Boyd, R. A. Mehl, G. A. O’Toole, and H. Sondermann, “Mechanistic insight into the conserved allosteric regulation of periplasmic proteolysis by the signaling molecule cyclic-di-GMP.,” *Elife*, vol. 3, p. e03650, 2014.
- [99] X. Fang, I. Ahmad, A. Blanka, M. Schottkowski, A. Cimmins, M. Y. Galperin, U. Römling, and M. Gomelsky, “GIL, a new c-di-GMP-binding protein domain involved in regulation of cellulose synthesis in enterobacteria,” *Mol. Microbiol.*, vol. 93, pp. 439–452, 2014.
- [100] J. W. Hickman and C. S. Harwood, “Identification of FleQ from *Pseudomonas aeruginosa* as a c-di-GMP-responsive transcription factor,” *Mol. Microbiol.*, vol. 69, pp. 376–389, 2008.
- [101] N. Sudarsan, E. R. Lee, Z. Weinberg, R. H. Moy, J. N. Kim, K. H. Link, and R. R. Breaker, “Riboswitches in eubacteria sense the second messenger cyclic di-GMP.,” *Science*, vol. 321, pp. 411–413, 2008.
- [102] H. Xiao, T. E. Edwards, and A. R. Ferré-D’Amaré, “Structural Basis for Specific, High-Affinity Tetracycline Binding by an In Vitro Evolved Aptamer and Artificial Riboswitch,” *Chem. Biol.*, vol. 15, pp. 1125–1137, 2008.
- [103] J. R. Govan and V. Deretic, “Microbial pathogenesis in cystic fibrosis: mucoid *Pseudomonas aeruginosa* and *Burkholderia cepacia*.,” *Microbiol. Rev.*, vol. 60, pp. 539–574, 1996.

-
- [104] E. E. Smith, D. G. Buckley, Z. Wu, C. Saenphimmachak, L. R. Hoffman, D. A. D'Argenio, S. I. Miller, B. W. Ramsey, D. P. Speert, S. M. Moskowitz, J. L. Burns, R. Kaul, and M. V. Olson, "Genetic adaptation by *Pseudomonas aeruginosa* to the airways of cystic fibrosis patients.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 103, pp. 8487–92, 2006.
- [105] H. K. Huse, T. Kwon, J. E. A. Zlosnik, D. P. Speert, E. M. Marcotte, and M. Whiteley, "Parallel evolution in *Pseudomonas aeruginosa* over 39,000 generations in vivo," *MBio*, vol. 1, 2010.
- [106] V. Burke, J. O. Robinson, C. J. Richardson, and C. S. Bundell, "Longitudinal studies of virulence factors of *Pseudomonas aeruginosa* in cystic fibrosis.," *Pathology*, vol. 23, pp. 145–148, 1991.
- [107] S. Häußler, "Biofilm formation by the small colony variant phenotype of *Pseudomonas aeruginosa*," *Environmental Microbiology*, vol. 6. pp. 546–551, 2004.
- [108] S. Häußler, "Highly adherent small-colony variants of *Pseudomonas aeruginosa* in cystic fibrosis lung infection," *J. Med. Microbiol.*, vol. 52, pp. 295–301, 2003.
- [109] M. Starkey, J. H. Hickman, L. Ma, N. Zhang, S. De Long, A. Hinz, S. Palacios, C. Manoil, M. J. Kirisits, T. D. Starner, D. J. Wozniak, C. S. Harwood, and M. R. Parsek, "*Pseudomonas aeruginosa* Rugose small-colony variants have adaptations that likely promote persistence in the cystic fibrosis lung," *J. Bacteriol.*, vol. 191, pp. 3492–3503, 2009.
- [110] A. Meissner, V. Wild, R. Simm, M. Rohde, C. Erck, F. Bredenbruch, M. Morr, U. Römling, and S. Häußler, "*Pseudomonas aeruginosa* cupA-encoded fimbriae expression is regulated by a GGDEF and EAL domain-dependent modulation of the intracellular level of cyclic diguanylate," *Environ. Microbiol.*, vol. 9, pp. 2475–2485, 2007.
- [111] F. Zähringer, E. Lacanna, U. Jenal, T. Schirmer, and A. Boehm, "Structure and signaling mechanism of a zinc-sensory diguanylate cyclase," *Structure*, vol. 21, pp. 1149–1157, 2013.
- [112] A. A. Baykov, O. A. Evtushenko, and S. M. Avaeva, "A malachite green procedure for orthophosphate determination and its use in alkaline phosphatase-based enzyme immunoassay.," *Anal. Biochem.*, vol. 171, pp. 266–270, 1988.
- [113] "WolframAlpha." [Online]. Available:

<https://www.wolframalpha.com>.

- [114] L. Slabinski, L. Jaroszewski, L. Rychlewski, I. A. Wilson, S. A. Lesley, and A. Godzik, “XtalPred: A web server for prediction of protein crystallizability,” *Bioinformatics*, vol. 23, pp. 3403–3405, 2007.
- [115] “Uniprot database.” [Online]. Available: <http://www.uniprot.org>.
- [116] J. Söding, A. Biegert, and A. N. Lupas, “The HHpred interactive server for protein homology detection and structure prediction,” *Nucleic Acids Res.*, vol. 33, 2005.
- [117] U. Jenal, “Cyclic di-guanosine-monophosphate comes of age: A novel secondary messenger involved in modulating cell surface structures in bacteria?,” *Current Opinion in Microbiology*, vol. 7, pp. 185–191, 2004.
- [118] U. Jenal and J. Malone, “Mechanisms of cyclic-di-GMP signaling in bacteria.,” *Annu. Rev. Genet.*, vol. 40, pp. 385–407, 2006.
- [119] T. Schirmer and U. Jenal, “Structural and mechanistic determinants of c-di-GMP signalling.,” *Nat. Rev. Microbiol.*, vol. 7, pp. 724–735, 2009.
- [120] N. Ausmees, R. Mayer, H. Weinhouse, G. Volman, D. Amikam, M. Benziman, and M. Lindberg, “Genetic data indicate that proteins containing the GGDEF domain possess diguanylate cyclase activity,” *FEMS Microbiol. Lett.*, vol. 204, pp. 163–167, 2001.
- [121] A. D. Tischler and A. Camilli, “Cyclic diguanylate (c-di-GMP) regulates *Vibrio cholerae* biofilm formation,” *Mol. Microbiol.*, vol. 53, pp. 857–869, 2004.
- [122] and M. R. P. Mary Jo Kirisits, Lynne Prost, Melissa Starkey, “Characterization of Colony Morphology Variants Isolated from *Pseudomonas aeruginosa* Biofilms,” *Appl. Environ. Microbiol.*, vol. 71, no. 8, pp. 4809–4821.
- [123] R. D. Finn, A. Bateman, J. Clements, P. Coggill, R. Y. Eberhardt, S. R. Eddy, A. Heger, K. Hetherington, L. Holm, J. Mistry, E. L. L. Sonnhammer, J. Tate, and M. Punta, “Pfam: The protein families database,” *Nucleic Acids Research*, vol. 42, 2014.
- [124] A. G. W. Leslie and H. R. Powell, “Processing diffraction data with MOSFLM,” in *Evolving methods for macromolecular Crystallography*, 2007, pp. 41–51.
- [125] M. D. Winn, C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley,

- P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. W. Leslie, A. McCoy, S. J. McNicholas, G. N. Murshudov, N. S. Pannu, E. A. Potterton, H. R. Powell, R. J. Read, A. Vagin, and K. S. Wilson, "Overview of the CCP4 suite and current developments," *Acta Crystallographica Section D: Biological Crystallography*, vol. 67. pp. 235–242, 2011.
- [126] G. M. Sheldrick, "A short history of SHELX," *Acta Crystallographica Section A: Foundations of Crystallography*, vol. 64. pp. 112–122, 2007.
- [127] G. Langer, S. X. Cohen, V. S. Lamzin, and A. Perrakis, "Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7.," *Nat. Protoc.*, vol. 3, pp. 1171–1179, 2008.
- [128] P. D. Adams, P. V. Afonine, G. Bunkóczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L. W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger, and P. H. Zwart, "PHENIX: A comprehensive Python-based system for macromolecular structure solution," *Acta Crystallogr. Sect. D Biol. Crystallogr.*, vol. 66, pp. 213–221, 2010.
- [129] P. Emsley, B. Lohkamp, W. G. Scott, and K. Cowtan, "Features and development of Coot," *Acta Crystallogr. Sect. D Biol. Crystallogr.*, vol. 66, pp. 486–501, 2010.
- [130] V. B. Chen, W. B. Arendall, J. J. Headd, D. A. Keedy, R. M. Immormino, G. J. Kapral, L. W. Murray, J. S. Richardson, and D. C. Richardson, "MolProbity: All-atom structure validation for macromolecular crystallography," *Acta Crystallogr. Sect. D Biol. Crystallogr.*, vol. 66, pp. 12–21, 2010.
- [131] "DINO: Visualizing Structural Biology (2002) <http://www.dino3d.org>."
- [132] C. T. O. Benfield, D. S. Mansur, L. E. McCoy, B. J. Ferguson, M. W. Bahar, A. P. Oldring, J. M. Grimes, D. I. Stuart, S. C. Graham, and G. L. Smith, "Mapping the IkappaB kinase beta (IKKbeta)-binding interface of the B14 protein, a vaccinia virus inhibitor of IKKbeta-mediated activation of nuclear factor kappaB.," *J. Biol. Chem.*, vol. 286, pp. 20727–20735, 2011.
- [133] L. Y. Geer, M. Domrachev, D. J. Lipman, and S. H. Bryant, "CDART: Protein homology by domain architecture," *Genome Res.*, vol. 12, pp. 1619–1623, 2002.

- [134] M. Kearse, R. Moir, A. Wilson, S. Stones-Havas, M. Cheung, S. Sturrock, S. Buxton, A. Cooper, S. Markowitz, C. Duran, T. Thierer, B. Ashton, P. Meintjes, and A. Drummond, “Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data,” *Bioinformatics*, vol. 28, pp. 1647–1649, 2012.
- [135] E. Krissinel and K. Henrick, “Inference of Macromolecular Assemblies from Crystalline State,” *J. Mol. Biol.*, vol. 372, pp. 774–797, 2007.
- [136] Y. Ye and A. Godzik, “FATCAT: A web server for flexible structure comparison and structure similarity searching,” *Nucleic Acids Res.*, vol. 32, 2004.
- [137] L. Holm and P. Rosenström, “Dali server: Conservation mapping in 3D,” *Nucleic Acids Res.*, vol. 38, 2010.
- [138] A. G. Murzin, S. E. Brenner, T. Hubbard, and C. Chothia, “SCOP: A structural classification of proteins database for the investigation of sequences and structures,” *Journal of Molecular Biology*, vol. 247. pp. 536–540, 1995.
- [139] F. Rao, Y. Qi, H. S. Chong, M. Kotaka, B. Li, J. Li, J. Lescar, K. Tang, and Z.-X. Liang, “The functional role of a conserved loop in EAL domain-based cyclic di-GMP-specific phosphodiesterase.,” *J. Bacteriol.*, vol. 191, no. 15, pp. 4722–4731, Aug. 2009.
- [140] C. Cole, J. D. Barber, and G. J. Barton, “The Jpred 3 secondary structure prediction server.,” *Nucleic Acids Res.*, vol. 36, 2008.
- [141] R. D. Finn, J. Mistry, B. Schuster-Böckler, S. Griffiths-Jones, V. Hollich, T. Lassmann, S. Moxon, M. Marshall, A. Khanna, R. Durbin, S. R. Eddy, E. L. L. Sonnhammer, and A. Bateman, “Pfam: clans, web tools and services.,” *Nucleic Acids Res.*, vol. 34, pp. D247–D251, 2006.
- [142] M. Goujon, H. McWilliam, W. Li, F. Valentin, S. Squizzato, J. Paern, and R. Lopez, “A new bioinformatics analysis tools framework at EMBL-EBI,” *Nucleic Acids Res.*, vol. 38, 2010.
- [143] F. Sievers, A. Wilm, D. Dineen, T. J. Gibson, K. Karplus, W. Li, R. Lopez, H. McWilliam, M. Remmert, J. Söding, J. D. Thompson, and D. G. Higgins, “Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega,” *Molecular Systems Biology*, vol. 7. 2011.

-
- [144] X. Robert and P. Gouet, “Deciphering key features in protein structures with the new ENDscript server,” *Nucleic Acids Res.*, vol. 42, 2014.
- [145] D. W. A. Buchan, F. Minneci, T. C. O. Nugent, K. Bryson, and D. T. Jones, “Scalable web services for the PSIPRED Protein Analysis Workbench,” *Nucleic Acids Res.*, vol. 41, 2013.
- [146] J. Lipfert, L. Columbus, V. B. Chu, S. A. Lesley, and S. Doniach, “Size and shape of detergent micelles determined by small-angle X-ray scattering,” *J. Phys. Chem. B*, vol. 111, pp. 12427–12438, 2007.
- [147] A. Boehm, S. Steiner, F. Zaehring, A. Casanova, F. Hamburger, D. Ritz, W. Keck, M. Ackermann, T. Schirmer, and U. Jenal, “Second messenger signalling governs *Escherichia coli* biofilm induction upon ribosomal stress,” *Mol. Microbiol.*, vol. 72, pp. 1500–1516, 2009.
- [148] A. Krogh, B. Larsson, G. von Heijne, and E. L. Sonnhammer, “Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes,” *J. Mol. Biol.*, vol. 305, pp. 567–580, 2001.
- [149] A. Gattiker, E. Gasteiger, and A. Bairoch, “ScanProsite: a reference implementation of a PROSITE scanning tool,” *Appl. Bioinformatics*, vol. 1, pp. 107–108, 2002.
- [150] M. Brune, J. L. Hunter, J. E. Corrie, and M. R. Webb, “Direct, real-time measurement of rapid inorganic phosphate release using a novel fluorescent probe and its application to actomyosin subfragment 1 ATPase,” *Biochemistry*, vol. 33, pp. 8262–8271, 1994.
- [151] G. Giardina, A. Paiardini, S. Fernicola, S. Franceschini, S. Rinaldo, V. Stelitano, and F. Cutruzzola, “Investigating the allosteric regulation of YfiN from *Pseudomonas aeruginosa*: Clues from the structure of the catalytic domain,” *PLoS One*, vol. 8, 2013.
- [152] K. Jonas, A. N. Edwards, R. Simm, T. Romeo, U. Römling, and Ö. Melefors, “The RNA binding protein CsrA controls cyclic di-GMP metabolism by directly regulating the expression of GGDEF proteins,” *Mol. Microbiol.*, vol. 70, pp. 236–257, 2008.
- [153] E. Brombacher, C. Dorel, A. J. B. Zehnder, and P. Landini, “The curli biosynthesis regulator CsgD co-ordinates the expression of both positive and negative determinants for biofilm formation in *Escherichia coli*,” *Microbiology*, vol. 149, pp. 2847–2857, 2003.

- [154] C. Pesavento, G. Becker, N. Sommerfeldt, A. Possling, N. Tschowri, A. Mehliis, and R. Hengge, “Inverse regulatory coordination of motility and curli-mediated adhesion in *Escherichia coli*,” *Genes Dev.*, vol. 22, pp. 2434–2446, 2008.
- [155] M. Gjermansen, M. Nilsson, L. Yang, and T. Tolker-Nielsen, “Characterization of starvation-induced dispersion in *Pseudomonas putida* biofilms: Genetic elements and molecular mechanisms,” *Mol. Microbiol.*, vol. 75, pp. 815–826, 2010.
- [156] J. S. Parkinson, “Signaling mechanisms of HAMP domains in chemoreceptors and sensor kinases.,” *Annu. Rev. Microbiol.*, vol. 64, pp. 101–122, 2010.
- [157] D. Chatterjee, C. D. Boyd, G. A. O’Toole, and H. Sondermann, “Structural characterization of a conserved, calcium-dependent periplasmic protease from *Legionella pneumophila*,” *J. Bacteriol.*, vol. 194, pp. 4415–4425, 2012.
- [158] E. K. O’Shea, J. D. Klemm, P. S. Kim, and T. Alber, “X-ray structure of the GCN4 leucine zipper, a two-stranded, parallel coiled coil.,” *Science*, vol. 254, pp. 539–544, 1991.
- [159] C. S. Bond, “TopDraw: A sketchpad for protein structure topology cartoons,” *Bioinformatics*, vol. 19, pp. 311–312, 2003.

3.2 In vitro characterization of YfiN, a key regulator of biofilm formation in *Pseudomonas aeruginosa*

3.2.1 Functional characterization of membrane-bound and detergent-extracted YfiN

3.2.1.1 Design and cloning of the YfiN constructs

In the beginning of project, the sequence of YfiN_{GGDEF} was aligned to the two GGDEF domains of the diguanylate cyclases PleD and DgcZ (Figure 3-1) [23], [111]. GGDEF domain boundaries of YfiN were determined by Pfam [141] and predicted to range from residues 249 – 407. The sequences of the GGDEF domains of PleD and DgcZ were aligned starting from $\alpha 0$ to the C-terminal ends of the proteins, omitting the region forming the $\beta 0$ - $\beta 0'$ -hairpin, which is often seen in DGC structures upstream of $\alpha 0$.

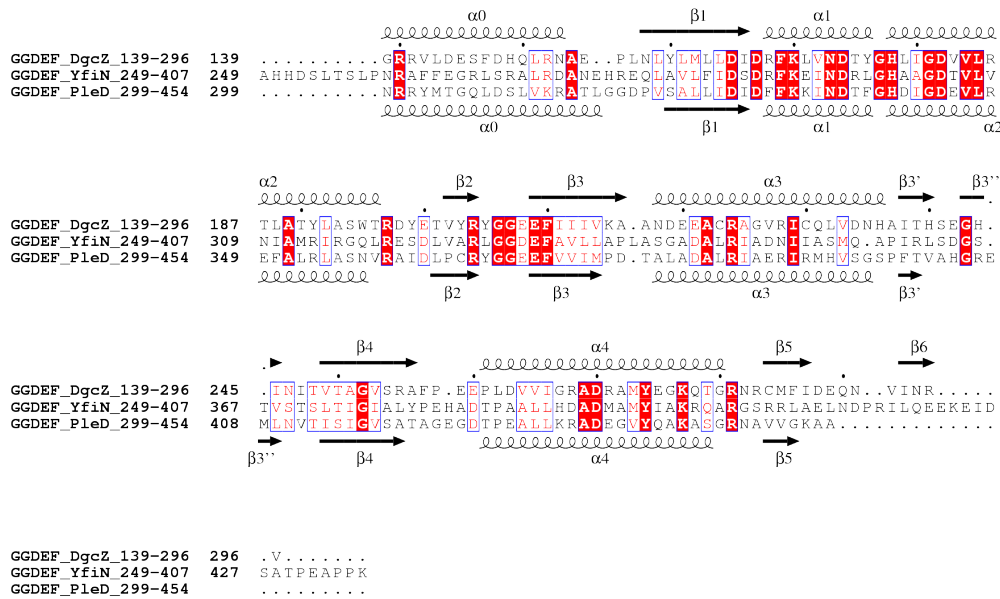


Figure 3-1: Multiple sequence alignment of YfiN_{GGDEF}, DgcZ_{GGDEF} and PleD_{GGDEF} with secondary structural elements indicated. Strictly conserved residues are boxed in white on a red background, and highly conserved residues are boxed in red on a white background. Every 10th residue is indicated with a dot. The α -helices are depicted by a coil, whereas β -strands are shown by an arrow. The secondary structure of PleD and DgcZ are according to [23] and [111], respectively. The alignment was performed using CLUSTAL OMEGA [142] [143]. The figure was generated using ESPript [144].

At the C-terminus, YfiN_{GGDEF} is about 14 - or 22 residues longer than DgcZ or PleD, respectively. However, according to secondary structure prediction [145], the C-terminus is not flexible but predicted with the highest confidence score to be involved in helix formation. Therefore, it was included in the YfiN construct (Figure 3-2).

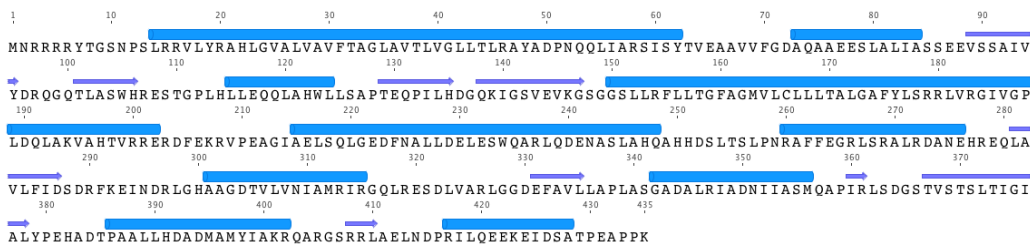


Figure 3-2: Secondary structure prediction for YfiN by psipred analysis [145]. α -helices are colored in blue, whereas β -strands are colored in purple. The figure was generated using Geneious [134].

All YfiN constructs investigated within the framework of this thesis are listed in table Table 3-1.

Name	Feature	N-terminal tag	C-terminal tag
pET28a/YfiN	full-length (1-435)	his ₆	-
		-	his ₆
pET28a/YfiN _{PAS38-159}	periplasmic domain		his ₆
pET28a-SUMO/YfiN _{PAS38-159}	periplasmic domain, fusion	SUMO	his ₆
		SUMO	-
pMal-c2X/YfiN _{PAS38-159}	periplasmic domain, fusion	MBP	his ₆
		MBP	-
pMal-p5X/YfiN _{PAS38-159}	periplasmic domain, fusion	MBP	his ₆
		MBP	-
pET-GSTx/ YfiN _{PAS38-159}	periplasmic domain, fusion	GST	his ₆
		GST	-
pET28a/YfiN _{PAS44-149}	periplasmic domain		his ₆
pET28a-SUMO/ YfiN _{PAS44-149}	periplasmic domain, fusion	SUMO	his ₆
		SUMO	-
pMal-c2X/YfiN _{PAS44-149}	periplasmic domain, fusion	MBP	his ₆
		MBP	-
pMal-p5X/YfiN _{PAS44-149}	periplasmic domain, fusion	MBP	his ₆
		MBP	-
pET28a/YfiN _{PAS44-154}	periplasmic domain		his ₆
pET28a/YfiN _{HAMP-GGDEF178-435}	cytoplasmic domain		his ₆
pET28a/ <i>P. fluorescens</i> YfiN _{HAMP-GGDEF174-420}	cytoplasmic domain, homolog		his ₆
pET28a/ <i>Y. enterocoliticae</i> YfiN _{HAMP-GGDEF188-448}	cytoplasmic domain, homolog		his ₆
pET28a/ <i>S. alaskensis</i> YfiN _{HAMP-GGDEF179-408}	cytoplasmic domain, homolog		his ₆

Table 3-1: YfiN constructs used in the frame of this thesis. The color code is reflecting the different groups of constructs: full-length (yellow), YfiN_{HAMP-GGDEF} constructs from different species (purple) and three YfiN_{PAS} constructs of different length (38-159: orange, 44-149: green and 44-154: blue).

3.2.1.2 Expression of the YfiN constructs

Expression tests of YfiN were performed in either LB - or TB-medium using BL21(DE3) at either 20 °C or 30 °C respectively. Protein expression was induced by addition of 250 μ M IPTG in the case of LB-medium and 50 μ M IPTG in the case of TB-medium. The highest expression levels in membranes were achieved using TB-medium and an expression duration of 5 hours at 30 °C. Placement of the his-tag appeared to be crucial for protein stability, as proteolysis was occurring for the C-terminally his-tagged protein. Two expression gels and their corresponding western blots are shown in Figure 3-3. For protein production, bacteria were cultivated for 5 hours at 30 °C in 2 L Erlenmeyer flasks using TB-medium.

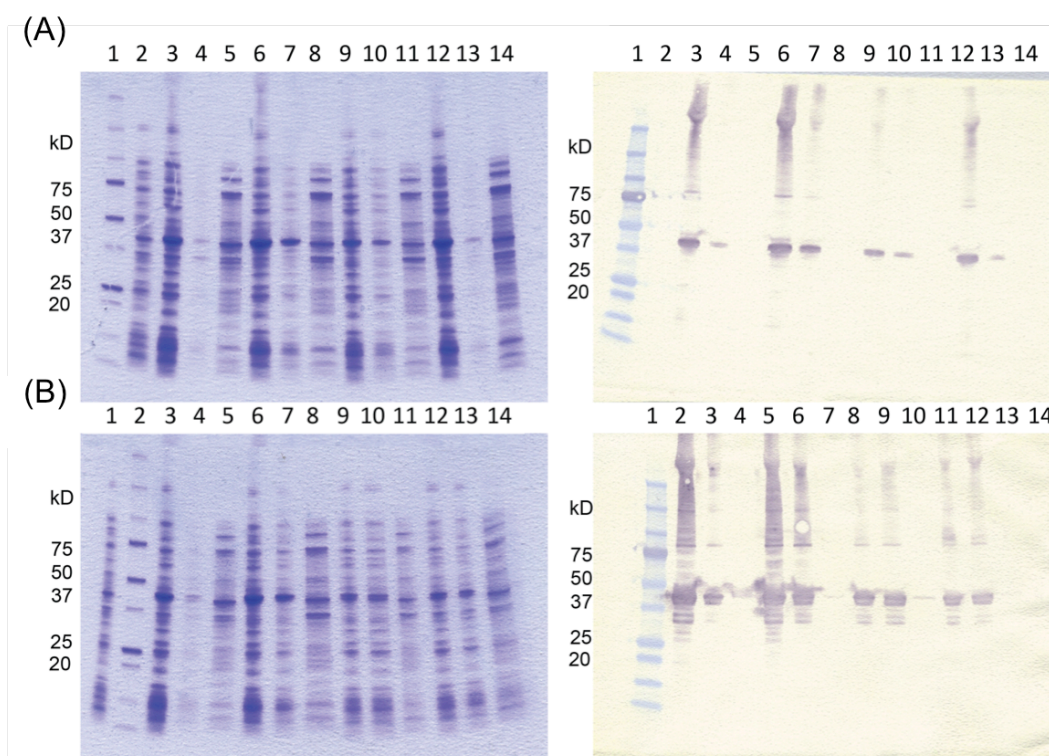


Figure 3-3: SDS-PAGE and western blots for expression tests of either N- or C-terminally his-tagged YfiN. Same amount of cells were loaded in each lane. **(A)** N-terminal his-tag: 1: marker, 2: before induction, 3: cell debris (3h of expression at 30°C), 4: membranes (3h of expression at 30°C), 5: soluble fraction (3h of expression at 30°C), 6: cell debris (5h of expression at 30°C), 7: membranes (5h of expression at 30°C), 8: soluble fraction (5h of expression at 30°C), 9: cell debris (5h of expression at 20°C), 10: membranes (5h of expression at 20°C), 11: soluble fraction (5h of expression at 20°C), 12: cell debris (overnight (o/n) expression at 20°C), 13: membranes (o/n expression at 20°C), 14: soluble fraction (o/n expression at 20°C). **(B)** C-terminal his-tag: the sample order is the same as in (A), except for 1: before induction, 2: marker.

3.2.1.3 Optimization of the preparation protocol of membranes used for a YfiN activity assay

Before proceeding to activity assays using solubilized protein, an assay was established that allowed characterization of enzymatic activity of YfiN in its native membrane environment. Initially, membranes were collected as described in 2.2.2.1 and rinsed only once with reaction buffer (50 mM Tris-HCl, pH 8.0, 200 mM NaCl, 5 mM MgCl₂) prior to the activity assays. They were then incubated with 100 μM GTP and the reaction was allowed to

take place for 30 - and 60 min respectively, where a sample was then taken and analyzed as described in 2.5.1. The resulting chromatograms appeared completely overcrowded with loads of unidentified peaks, making it impossible to draw a conclusion if any GTP consumption or c-di-GMP production was taking place (Figure 3-4).

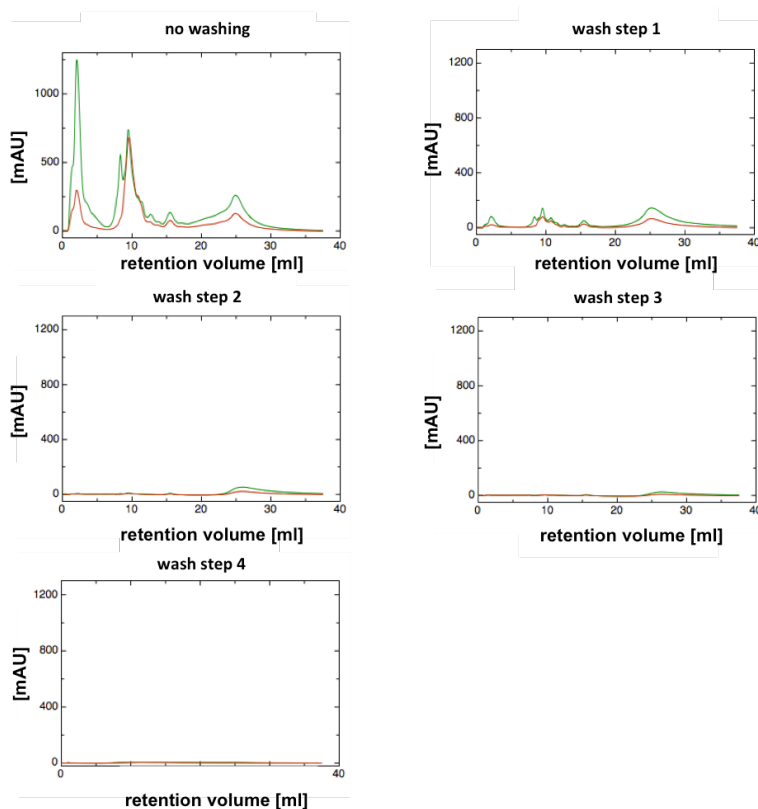


Figure 3-4: Chromatograms of the individual washing steps. The absorption observed at 280 nM is depicted in red, whereas the absorption at 253 nM is colored in green.

Therefore, the preparation of the membranes has to be optimized. Instead of rinsing them only once with reaction buffer in the beginning, they were washed up to four times by completely resuspending the pellet in 4 mL of fresh buffer each time, followed by a centrifugation step to collect the membranes again. This procedure was repeated while the resulting supernatant was analyzed by FPLC. For the activity assay, only membranes

were used that showed no longer any background contamination, suggesting that they were now free from any nucleotides, which were originally present in the cell.

Additionally, reference spectra were generated by running GMP, GDP, GTP and c-di-GMP in different combinations on the ResQ column to establish at which retention volume each nucleotide is eluting and to ensure that all compounds can be separated at sufficient resolution (Figure 3-5). All nucleotides were present at a final concentration of 200 μM .

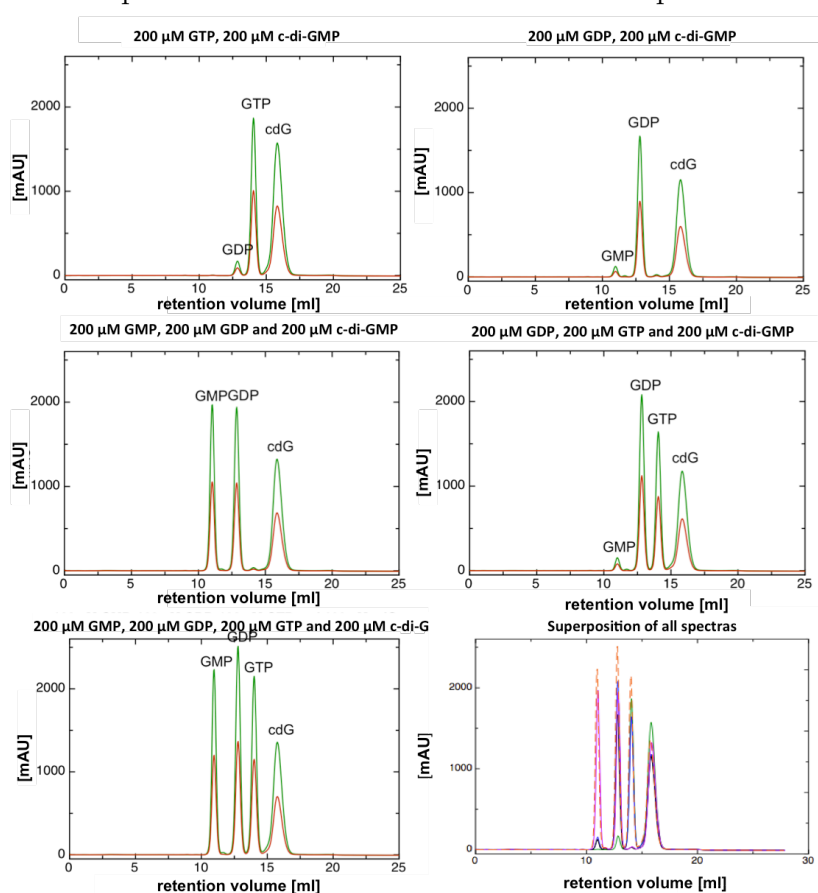


Figure 3-5: Reference spectra of the nucleotides GMP, GDP, GTP and c-di-GMP. The absorption at 253 nm is colored in green, whereas the absorption at 280 nm is depicted in red. The color code for the superposition is as follows: green: sample 1, grey: sample 2, pink: sample 3, blue: sample 4, orange: sample 5.

Results

nucleotide	retention volume [mL]	conductivity [mS/cm]
GMP	11.0	19.6
GDP	12.8	26.5
GTP	14.0	31.0
c-di-GMP	15.8	37.2

Table 3-2: A list of the analyzed nucleotides with the experimentally determined retention volumes and the corresponding conductivity.

3.2.1.4 Quantification of DGC activity using membrane-bound YfiN

3.2.1.4.1 Non-competitive inhibition

Non-competitive inhibition characterizes a system where the inhibitor and the substrate may be bound to the enzyme at any given time. When both the substrate and the inhibitor are bound, the enzyme-substrate-inhibitor complex cannot form product and can only be converted back to the enzyme-substrate complex or the enzyme-inhibitor complex.

A noncompetitive inhibitor binds to a different (allosteric) site that is not the active site and thereby induces changes in the structure of the enzyme. The enzyme is thereby prevented from forming product, leading to a decrease in the rate of the chemical reaction of enzyme and substrate, which cannot be changed by increasing concentration of substrate. The binding of an uncompetitive inhibitor thus decreases V_{\max} by a factor α but has no change on the K_m of the chemical reaction.

$$\alpha = 1 + \frac{[I]}{K_i} \quad \text{and} \quad \alpha' = 1 + \frac{[I]}{K'_i}$$

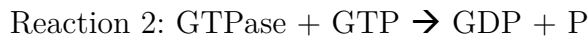
Equation 7: K_i is the dissociation constant for binding of the inhibitor to the enzyme whereas K'_i describes binding of the inhibitor to the enzyme-substrate complex.

		Type of inhibition	Km apparent	Vmax apparent
$K_i = K_i'$	$\alpha = \alpha'$	non-competitive	K_m	V_{max}/α

Table 3-3

3.2.1.4.2 The function used to quantify DGC activity using membrane-bound YfiN

In the system present there are 2 enzymatic reactions that are taking place simultaneously involving the same substrate (GTP) but not the same enzymes (DGC and GTPase).



In the case of DGCs (enzyme 1) non-competitive inhibition by the product c-di-GMP and - if added to the experiments - the repressor YfiR was observed (Figure 3-6, Figure 3-7). According to Table 3-3 illustrating the effects of non-competitive inhibition on Km and Vmax, the apparent kcat ($kcat1'$) can be expressed as a function of the effective kcat ($kcat1$), the product c-di-GMP concentration (P10) and the dissociation constants of the inhibitors c-di-GMP (K_i) and YfiR (K_r):

$$kcat1' = \frac{kcat1}{\left(1 + \frac{[P10]}{K_i}\right) \times \left(1 + \frac{[R]}{K_r}\right)}$$

Based on Equation 5 (2.5.4.2), the reaction rate of c-di-GMP production can be expressed as follows:

$$\text{reaction 1} = \frac{[E1] \times [S1] \times kcat1'}{(Km1 + [S1])}$$

For the GTPases (enzyme 2) a time-dependent inactivation was observed that was accounted for by the introduction of an additional parameter “ $t_{1/2}$ ” to the formula.

$$kcat2' = kcat2 \times (e^{\frac{-t}{t_{1/2}}})$$

Analogous to reaction 1, reaction 2 can therefore be expressed as follows:

$$\text{reaction 2} = \frac{[E2] \times [S1] \times kcat2}{(Km2 + [S1])} \times e^{\frac{-t}{t_{half}}}$$

The total reaction rate $\frac{dP}{dt}$ is therefore a sum of both reactions that are taking place in the system simultaneously and can be expressed as follows:

$$\frac{dP}{dt} = \left(kcat1' \times \frac{[E1] \times [S1]}{Km1 + [S1]} \right) + \left(kcat2' \times \frac{[E2] \times [S1]}{Km2 + [S1]} \right)$$

Equation 8

3.2.1.4.3 Experimental setup used to quantify DGC activity

Membranes containing YfiN were prepared as described in section 2.4 and used for a series of experiments using different conditions. In a first experiment (I), the membranes were incubated with reaction buffer (50 mM Tris-HCl, pH 8.0, 200 mM NaCl, 5 mM MgCl₂) containing 2 mM GTP to screen for DGC activity. In a second experiment (II), the reaction buffer was supplemented with 155 μM c-di-GMP in addition to 2 mM GTP to verify if potential DGC activity could be abolished by non-competitive product inhibition. In a third experiment (III), 148 μM purified YfiR was added to the membranes in buffer containing 2.2 mM GTP to test if YfiR is inhibiting YfiN also in vitro. Two negative controls were included as well, incubating either GTP or c-di-GMP in the reaction buffer over the same amount of time as the experiments. Samples were then taken and analyzed at the start and at the end of the experiment, which revealed complete stability of the nucleotides over the time-course investigated (data not shown). When incubated with membranes, there is a clear indication that GTP got turned over to either GDP or c-di-GMP and indirectly to GMP over time (Figure 3-6 (A), (C), (E)). The appearance of GDP as a side-product indicates that other enzymes, notably GTPases, are present in the membranes and show catalytic activity. However, this activity is clearly slowing down within the first 30 min of the experiments, most probably due to inactivation of the proteins in question. This observation was accounted for by introducing an additional parameter, “ $t_{1/2}$ ”, to the formula used to fit the data (3.2.1.4.2, Equation 8), which described the half-life of the GTPases.

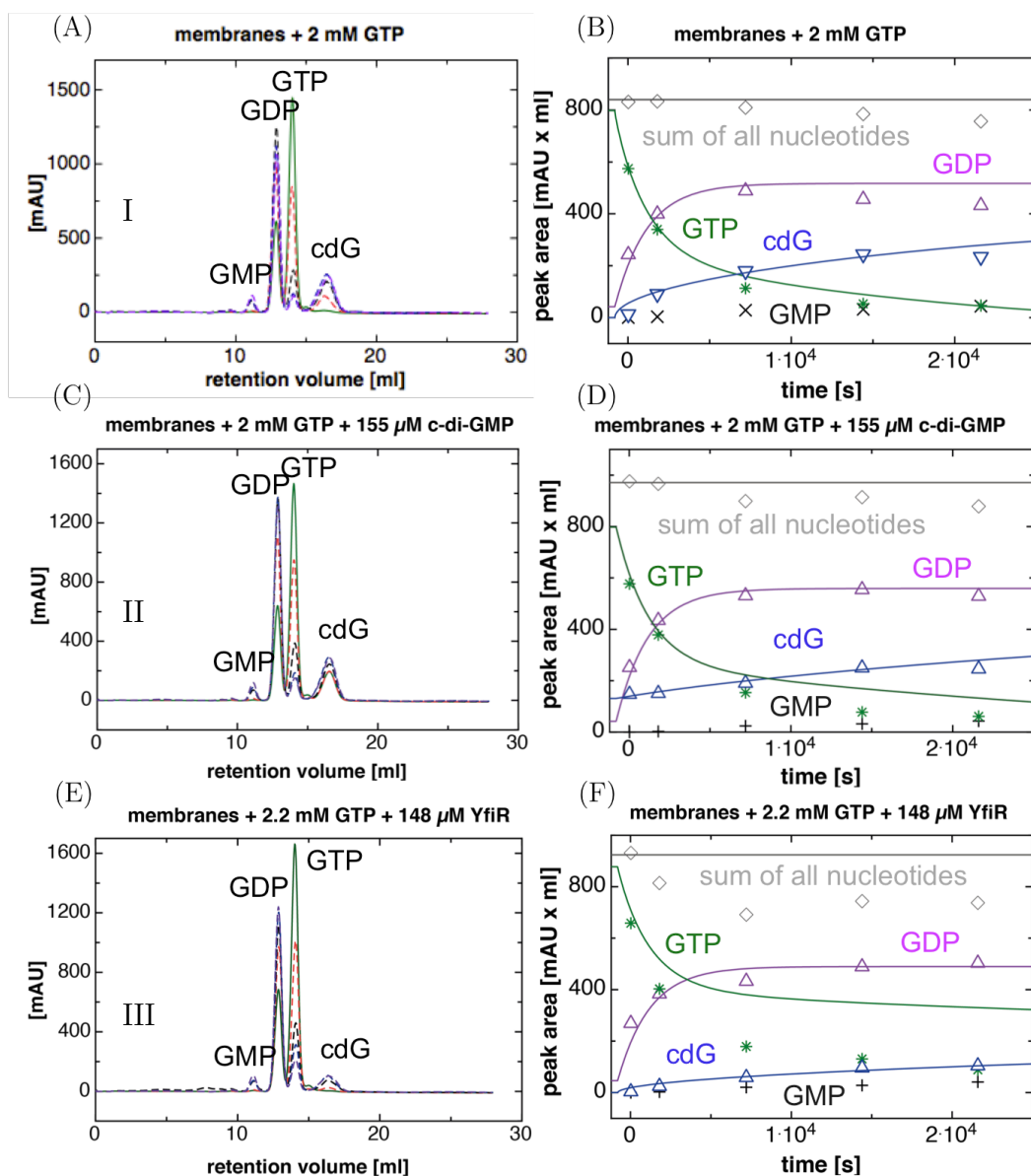


Figure 3-6: Activity test using membranes containing YfiN + 2 mM GTP (first row, experiment I), 2 mM GTP + 155 μ M c-di-GMP (second row, experiment II) and 2.2 mM GTP + 148 μ M YfiR (third row, experiment III). (A)(C)(E) The obtained chromatograms showing nucleotide turnover after different time points: green (20 sec), orange (30 min), black (120 min), blue (240 min) and purple (360 min). The different species are labeled accordingly. (B)(D)(F) The peak areas of the data points were plotted against the time and fitted to a Michaelis-Menten model (Equation 5). GMP production was not included in the model as only minor amounts were generated and the reaction is not interfering with the GTPase and diguanylate cyclase activity. Data points are colored in black (GMP), pink (GDP), green (GTP) and blue (c-di-GMP). The experimentally obtained peak areas of GMP, GDP, GTP and c-di-GMP were summed up (grey diamonds) and the theoretical sum of the nucleotides is marked with a grey bar.

All three data sets were evaluated in the following way. First, the kinetic constant, k_{cat2} , which stands for the turnover rate of the GTPases, was fitted for each experiment individually (Figure 3-6, (B), (D), (F), pink line). For experiment I, only the first three data points were used for the fitting as a decrease of absorption was observed, which cannot be explained by the applied model. For the same reason the last data point was omitted for fitting in experiment II.

In a second step, the k_{cat1} , which describes the turnover rate of the DGC(s), was fitted for each experiment individually and the resulting curve was plotted in blue (Figure 3-6, (B), (D), (F), blue line). As it was observed before for the k_{cat2} value, the last two data points of experiments I and II were decreasing and hence they were omitted for the fitting.

As the decrease in GTP concentration can be exclusively attributed to the synergistic activities of DGCs and GTPases, the previously individually fitted k_{cat} values (k_{cat1} , k_{cat2}) were then used as input values for each experiment and the decrease in substrate concentration was chosen as an output. This yielded a progression curve that was reflecting the decrease in substrate concentration as a function of the two k_{cat} values (Figure 3-6, (B), (D), (F), green line).

In addition to GDP, GMP was observed as a side-product, however, only very small amounts were detected. The GMP-producing activity was found to be interfering with neither DGCs nor GTPases, as both enzymes were still in full substrate saturation, and therefore the activity was neglected for the fitting procedure.

It was observed that a small amount of nucleotides was lost during the course of all three experiments, often starting around prolonged incubation times (Figure 3-6, (B), (D), (F), grey data points). For experiments I and II,

the sum of all nucleotides at the end of the experiment was still comparable to the amount present in the beginning of the experiment. Nucleotide loss was especially pronounced in experiment III, possibly explaining the rather large discrepancy between the theoretical GTP progression curve and the actual amount of GTP present in the sample (Figure 3-6, green data points and green line). The experiments should be repeated in order to establish if this observations

In a final step, global fitting of the k_{cat1} and the K_r , which is the dissociation constant of the YfiR-YfiN binding, was performed based on DGC activity data from all three experiments and the resulting curves were plotted in Figure 3-7. An overview of all input and output parameters used for the fitting is presented in Table 3-4.

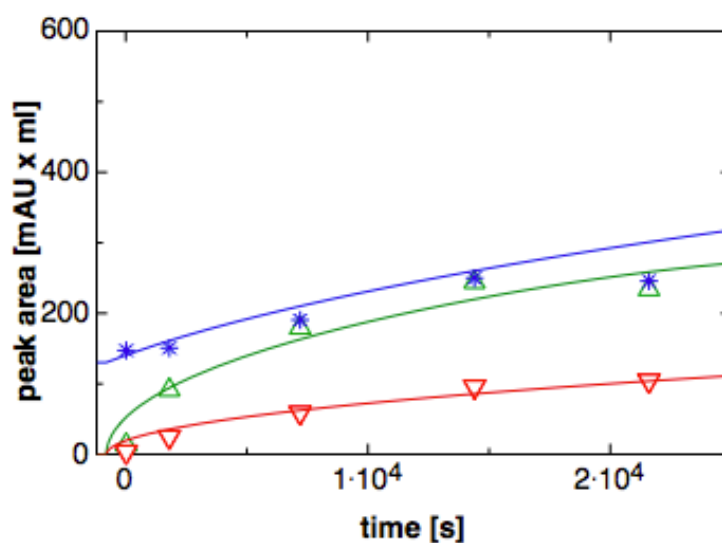


Figure 3-7: DGC activity, which was observed in the presence of GTP only (experiments I, green data points), GTP and c-di-GMP (experiment II, blue data points) as well as GTP and YfiR (experiment III, red data points).

	GTP	GTP + c-di-GMP	GTP + YfiR	Global fit of data
Unrefined parameters				
[E1] [M] (DGC)	1.00E-09	=	=	
[E2] [M] (GTPase)	1.00E-09	=	=	
[S] [M]	1.90E-03	=	2.09E-03	
[P10] [M] (c-di-GMP)	-	1.55E-04	-	
[P20] [M] (GDP)	1.00E-04	=	1.10E-04	
[R] [M] (YfiR)	-	-	1.48E-04	
Km1 [M]	4.00E-05	=	=	
Km2 [M]	4.00E-05	=	=	
Ki1 [M]	1.00E-06	=	=	
t _{1/2} [s]	2000	=	=	
t_start [s]	-800	=	=	
sc0 (GTP)	4.20E+05	=	=	
sc1 (c-di-GMP)	8.40E+05	=	=	
sc2 (GDP)	4.20E+05	=	=	
Refined parameters				
kcat1 [rel. units] (DGC)	5676.5	4205.9	4533.7	5095.8
kcat2 [rel. units] (GTPase)	394.4	427.4	363.9	
Kr [M]	-	-	2.83E-05	2.53E-05
Ratio kcat1/kcat2	14.4	9.8	12.5	

Table 3-4: Summary of all input and output parameters described in the Michaelis-Menten model that was used for fitting of the data. Columns 2-4 report values that were obtained by individual fitting of the data, whereas in column 5 globally fitted values are stated. Numbers 1 and 2 in conjunction with a term indicate that the value is either associated with the DGC activity (1) or the GTPase activity (2). The symbol “=” indicates that the number marked in one of the columns was the same for all three experiments. **[E1]** and **[E2]** represent the unknown concentrations of the DGC and GTPases, which were fixed at 1 nM each. The substrate concentration is in-fact made up of two contributions, as even a freshly prepared GTP **[S]** stock shows a GDP contamination **[P20]**, which amounts to roughly 5 % of the total concentration ($[S]+[P20]$). **[P10]** and **[R]** refer to the concentrations at which c-di-GMP and YfiR were added to the experimental set-up. **Km1** and **Ki1** are associated with the diguanylate cyclase(s), whereas **Km2** refers to the GTPases. As the GTPase activity decreased steadily over time, most probably caused by inactivation of the proteins, the parameter “t_{1/2}” was introduced to account for this observation. “t_start” is referring to the time delay between setting up the experiment and the heat inactivation of the protein for the first time point sample. The three scaling factors are relating the determined peak area to the corresponding concentrations of each nucleotide. The **kcat1** and **kcat2** values referring to the turnover rate (relative units) of either DGC or GTPase, were refined locally for each dataset and the resulting ratio is indicated in the last row. In addition, kcat1 and Kr were fitted globally using the DGC activity data from all three experiments. The scaling factors (**sc0,sc1,sc2**) are required to convert the obtained peak areas into concentrations of the nucleotides.

3.2.1.5 Solubilization and purification of YfiN

The first step of the purification of YfiN was its extraction from the inner membrane. Therefore, the total membrane fraction was collected and incubated in three different experiments with the same buffer supplemented with various detergent concentration for solubilization/purification: 21 mM/3 mM DM, 20 mM DDM/0.3 mM DDM and 51 mM/35 mM β -OG. The solubilization process was allowed to take place for one hour with steady stirring applied. After a centrifugation step to separate the solubilized fraction from the residual insoluble material, the supernatant was then loaded onto a 5 mL HisTrap column for a subsequent Ni-affinity chromatography. Already at this stage it became evident that barely no protein (48.8 kD) was extracted when β -OG was used as detergent (Figure 3-8) and therefore, no further SEC was pursued and the experiment was terminated.

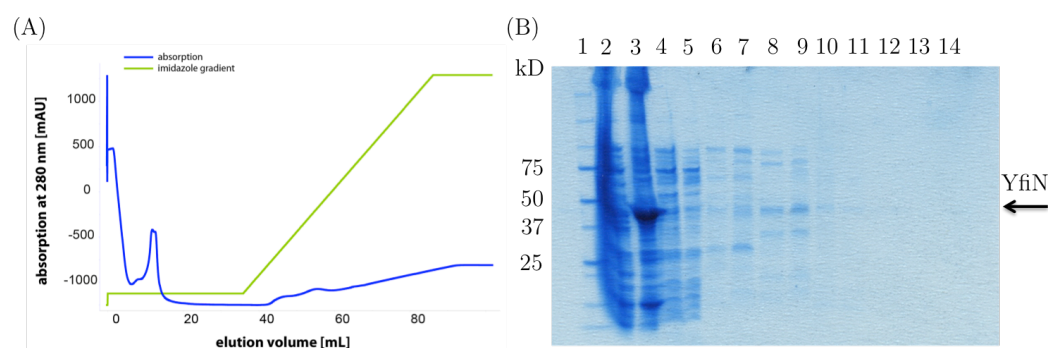


Figure 3-8: IMAC purification step of YfiN solubilized by β -OG. **(A)** Chromatogram of YfiN purification by IMAC using a 5 mL HisTrap column. **(B)** SDS-PAGE illustrating each step of YfiN purification: 1: marker, 2: whole cell, 3: non-solubilized membrane fraction, 4: load IMAC, 5: flow-through IMAC, 6: IMAC 45 mL-50 mL, 7: IMAC 50 mL-55 mL, 8: IMAC 55 mL-60 mL, 9: IMAC 60 mL-65 mL, 10: IMAC 65 mL-70 mL, 11: IMAC 70 mL-75 mL, 12: IMAC 75 mL-80 mL, 13: IMAC 80 mL-85 mL, 14: IMAC 85 mL-90 mL.

The solubilization capacity of DM (Figure 3-9) was found to be superior to the one from DDM (Figure 3-11) but nonetheless a size-exclusion chromatography was performed for both of them to screen for the quality of the protein. A huge aggregation peak was observed for both detergents after SEC, with only very little protein not eluting in the void volume in the case of DM (Figure 3-10). Due to time limitations, it was decided to proceed with DM as detergent.

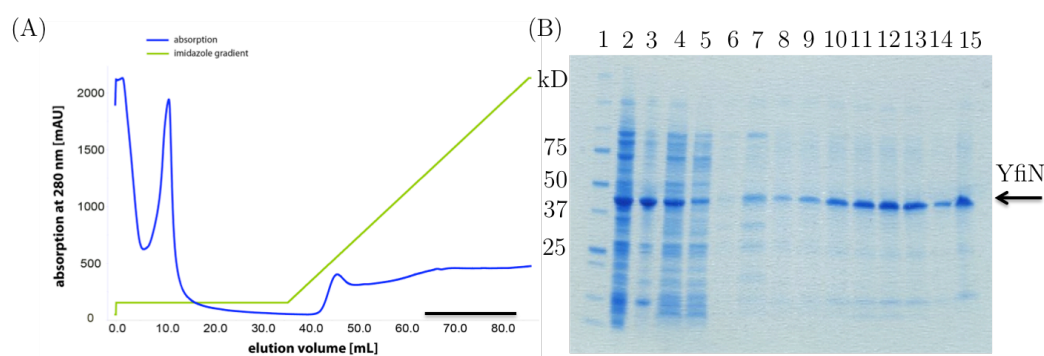


Figure 3-9: IMAC purification step of YfiN solubilized by DM. **(A)** Chromatogram of YfiN purification by IMAC using 5 mL HisTrap column. The major part of YfiN starts eluting at an imidazole concentration of 300 mM. Fractions pooled and used for subsequent purification are marked by a line. **(B)** SDS-PAGE illustrating each step of YfiN purification: 1: marker, 2: whole cell, 3: non-solubilized membrane fraction, 4: load IMAC, 5: flow-through IMAC, 6: IMAC 45 mL-50 mL, 7: IMAC 50 mL-55 mL, 8: IMAC 55 mL-60 mL, 9: IMAC 60 mL-65 mL, 10: IMAC 65 mL-70 mL, 11: IMAC 70 mL-75 mL, 12: IMAC 75 mL-80 mL, 13: IMAC 80 mL-85 mL, 14: IMAC 85 mL-90 mL, 15: 85 mL-90 mL.

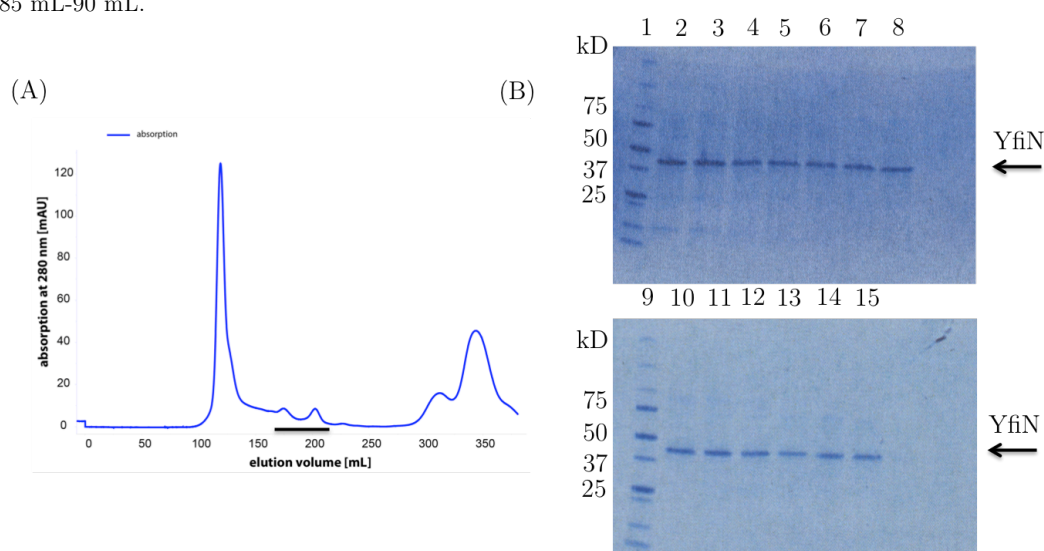


Figure 3-10: SEC purification step of YfiN solubilized by DM. **(A)** Chromatogram of YfiN purification by gel filtration chromatography using a HiLoad- 26/60-Superdex-200 prep grade column. **(B)** SDS-PAGE illustrating each step of YfiN purification: 1: marker, 2: SEC 110 mL-114 mL, 3: SEC 114 mL-118 mL, 4: SEC 118 mL-122 mL, 5: SEC 122 mL-126 mL, 6: SEC 126 mL-130 mL, 7: SEC 162 mL-166 mL, 8: SEC 166 mL-170 mL, 9: marker, 10: SEC 162 mL-166 mL, 11: 166 mL-170 mL, 12: SEC 170 mL-174 mL, 13: SEC 186 mL-190 mL, 14: SEC 190 mL-194 mL, 15: SEC 194 mL-198 mL. Fractions pooled and used for subsequent activity assays are marked by a line.

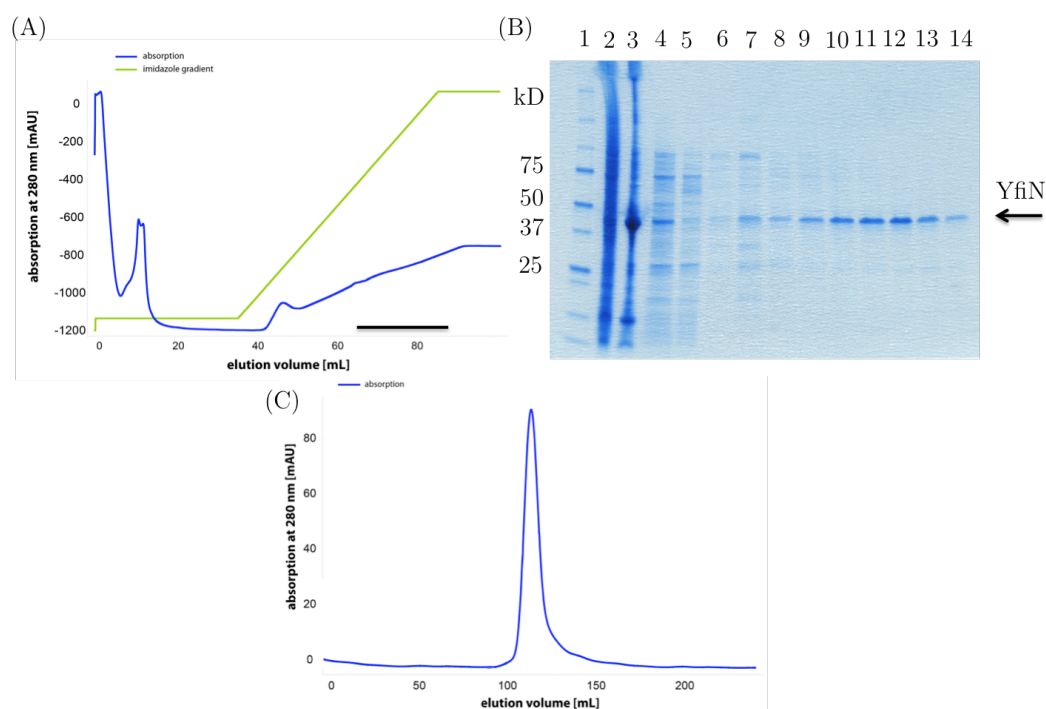


Figure 3-11: IMAC and SEC purification steps of YfiN solubilized by DDM. **(A)** Chromatogram of YfiN purification by IMAC using 5 mL HisTrap column. The major part of YfiN starts eluting at an imidazole concentration of 300 mM. Fractions pooled and used for subsequent purification are marked by a line. **(B)** SDS-PAGE illustrating each step of YfiN purification: 1: marker, 2: whole cell, 3: non-solubilized membrane fraction, 4: load IMAC, 5: flow-through IMAC, 6: IMAC 45 mL-50 mL, 7: IMAC 50 mL-55 mL, 8: IMAC 55 mL-60 mL, 9: IMAC 60 mL-65 mL, 10: IMAC 65 mL-70 mL, 11: IMAC 70 mL-75 mL, 12: IMAC 75 mL-80 mL, 13: IMAC 80 mL-85 mL, 14: IMAC 85 mL-90 mL. **(C)** Chromatogram of YfiN purification by gel filtration chromatography using a HiLoad- 26/60-Superdex-200 prep grade column revealing only aggregated protein.

The SEC for DM-solubilized YfiN was run at the very low flowspeed of 0.2 mL/min to separate the protein from a nucleotide that has been found bound (Figure 3-12).

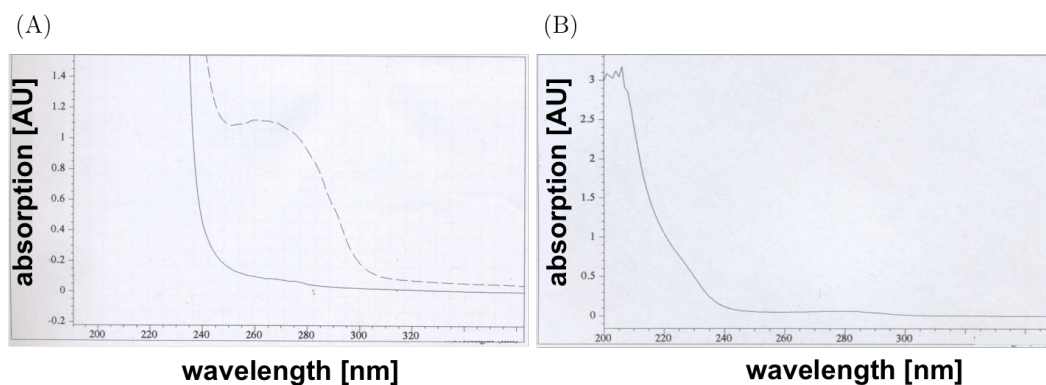


Figure 3-12: The obtained absorption spectra of YfiN before **(A)** and after **(B)** SEC illustrating nucleotide binding.

YfiN has been shown to co-elute with *c*-di-GMP before by Malone et al. [12]. 500 μ l of the aggregation peak (Figure 3-10(A)) were removed, cooked at 75 °C for 10 min and the precipitate was collected by centrifugation. The resulting supernatant was diluted 1:10 with running buffer (5 mM NH_4HCO_3 , pH 8.0), loaded onto a 5 mL ResQ column and analyzed by FPLC. *C*-di-GMP was included as a reference (Figure 3-13).

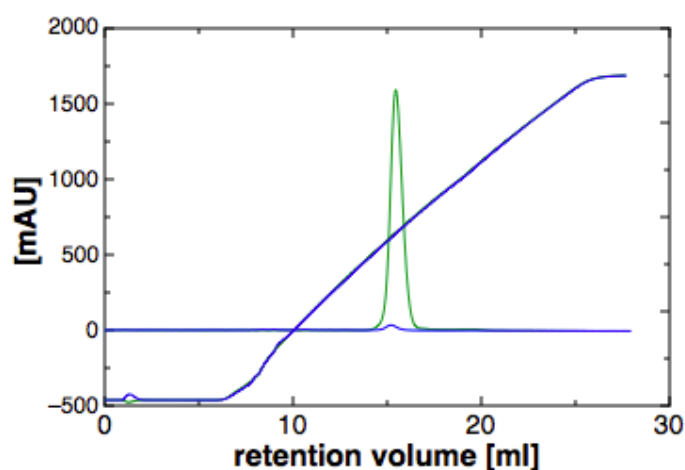


Figure 3-13: Superposition of the absorption spectra of *c*-di-GMP (green) and the supernatant of YfiN (blue), recorded at 253 nm. Both nucleotides were loaded onto a 5 mL ResQ column and eluted by a gradient of 0.005 M to 1 M NH_4HCO_3 .

The bound nucleotide was identified as *c*-di-GMP as it eluted at the same position as the *c*-di-GMP reference (Figure 3-13). The observed 260/280 ratio before SEC amounted to 1.15 (Figure 3-12, (A)), corresponding to a YfiN:*c*-di-GMP ratio of 1:1.5. After SEC a molar ratio of 1:0.3 was obtained, corresponding to almost nucleotide-free protein.

SEC fractions, which contained YfiN eluting from 162 mL-174 mL and 186 mL-198 mL (Figure 3-10), corresponding to the small peaks 2 and 3, were

concentrated to 1 mg/mL each and analyzed by dynamic light scattering (DLS).

	Hydrodynamic radius \pm SD [nm]	Estimated MW \pm SD [kD]	% Polydispersity
Peak 2	4.36 \pm 1.92	105.6 \pm 51.5	50.4
Peak 3	3.77 \pm 1.38	74.9 \pm 29.9	42.6

Table 3-5: The hydrodynamic radius as well as the corresponding molecular weight of YfiN eluting in peak 2 and 3 was determined by DLS.

The size of a YfiN monomer amounts to 48.8 kD, whereas a DM micelle has the size of a globular protein of roughly 60-70 kD [146]. As the observed standard deviations are quite high, it is difficult to get a reliable estimation of the oligomeric state of YfiN in solution (Table 3-5). Peak 3 could correspond to a monomer, whereas peak 2 most probably corresponds to a slightly different, e.g. more elongated monomeric state. The data is discussed in more detail in 3.5.2. Unfortunately, an attempt to get more reliable information about the oligomeric state of YfiN by multiple-angle laser light-scattering (MALLS) failed due to the necessary strong concentration of the protein, resulting in too high detergent concentrations present in the sample (data not shown). It was decided to go ahead with the protein from the two peaks, as even when the oligomeric state was not resolved, the protein clearly was not aggregated and therefore presumably suited for an activity assay.

3.2.1.6 Enzymatic characterization of detergent-solubilized YfiN using the malachite-green assay

3.2.1.7 Establishing the assay in a 96-well plate format using the well characterized DGC DgcZ

Before carrying out the assay with YfiN, DgcZ was used as a reference protein to verify if the previously obtained k_{cat} values could be reproduced if the assay was performed in a 96-well format [111], [147]. Therefore, 200 nM zinc-free DgcZ was incubated with either 1 μ M or 10 μ M GTP and samples were analyzed after 0 min, 5 min, 10 min, 20 min and 30 min as described in section 2.5.3. A negative control was included in both experiments, where GTP was incubated with the reaction buffer over the time span of the experiment without the addition of protein (Figure 3-14).

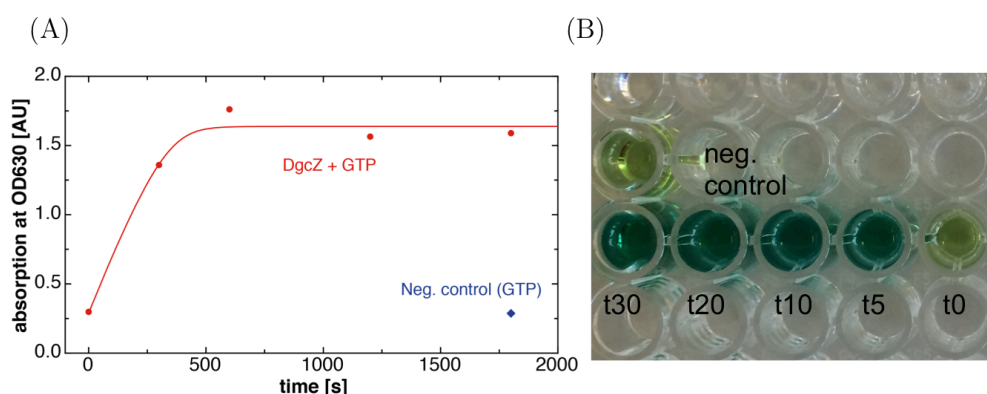


Figure 3-14: Characterizing the activity of DgcZ using the malachite-green assay in a 96-well plate format. (A) The obtained absorption values were plotted against the time and fitted to yield the k_{cat} value. of 0.19 s^{-1} . (B) The formation of the [malachite green-P_i] complex is leading to the characteristic color change from yellow to green, therefore being a direct measure for the amount of P_i in the reaction mixture.

The obtained increase in absorption values could be directly linked to the amount of generated inorganic phosphate using a scaling factor

(5.3299E+04) previously determined by a calibration curve with a correlation coefficient of 99.88 %. The linear range of the assay was determined to lie between 0 μM and 21 μM Pi, corresponding to an absorption limit of $\sim 0\text{-}1.5$ [AU]. The resulting k_{cat} was determined to be 0.19 s^{-1} , which is in agreement with the previously determined k_{cat} of 0.6 s^{-1} [111]. The malachite green assay has therefore been successfully established in the 96-well plate format for the characterization of DGC activity.

3.2.1.8 Using the established protocol to characterize the activity of detergent-solubilized YfiN

Before presenting the performed experiments, it has to be mentioned that the presumable DGC activity observed in one of the following experiments was actually due to a GTPase contamination in the YfiN sample. The nature of the assay does not allow discrimination between GTPase activity and DGC activity if pyrophosphatase (PPase) is present in the sample, as both lead to the production of inorganic phosphate, which is the readout of the assay. As the assay has been used before by others to characterize DGC (+PPase) as well as PDE (+alkaline phosphatase, AP) activity and none actually routinely included a negative control omitting the PPase/AP, it was not evident in the beginning to perform said negative control. However, after having observed that a YfiR sample incubated with GTP showed very low enzymatic activity (Figure 3-18), most likely due to a GTPase contamination, an experiment was conducted, where YfiN was incubated with GTP in presence and absence of PPase (Figure 3-15, Table 3-6).

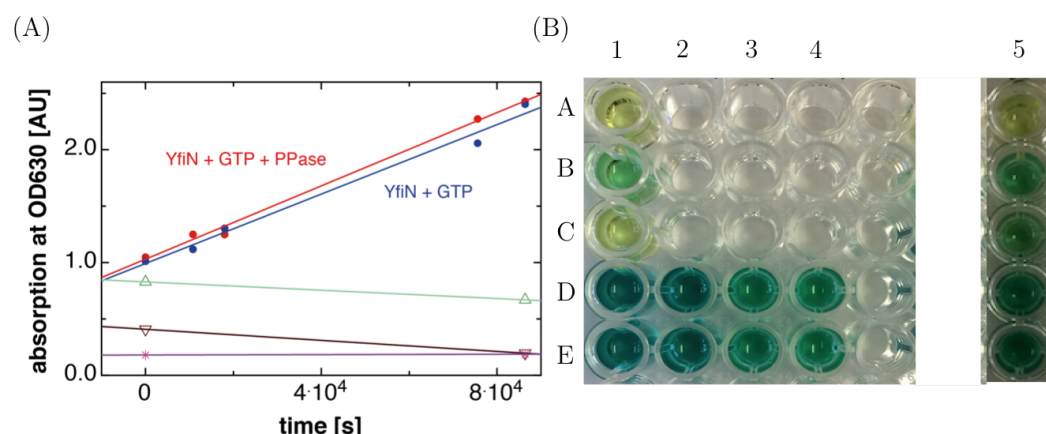


Figure 3-15: Activity assay in presence and absence of PPase. (A) The measured absorption values were plotted against the time and linear fit of the data was performed. For both experiments (YfiN + GTP + PPase) and (YfiN + GTP - PPase), only the first three data points were used for the fitting, in order to compare only the initial slopes of the curves and to exclude data that might not be lying in the linear range of the assay anymore. The resulting fits are indicated by solid lines and are colored according to the experiment: YfiN + GTP + PPase (red), YfiN + GTP - PPase (blue), YfiN + buffer (green), GTP + buffer (brown) and buffer alone (pink). (B) Time-course of the activity showing the formation of the molybdophosphoric acid complex, which is directly related to the free inorganic phosphate concentration. Lanes 1-5 correspond to the following time-points: 24 hours, 20 hours, 6 hours, 4 hours and 0 hours. Row A shows the buffer only control, row B the YfiN + buffer control, row C the GTP + buffer control, row D the YfiN + GTP + PPase time-course and row E the YfiN + GTP time-course.

	100 μ M GTP	v [μ M/s]	v_{corr} [μ M/s]
2.3 μ M YfiN + PPase	-	3.8E-5	
	+	3.6E-4	4.5E-4
2.3 μ M YfiN - PPase	+	3.1E-4	4.0E-4
no protein	-	0.0	
	+	4.7E-5	

Table 3-6: Summary of the data fitted by linear regression to obtain the initial velocity v of the protein. v_{corr} stands for the corrected initial velocity where the value of the negative control was subtracted to take potential GTP hydrolysis as well as other effects into account.

The experiment showed that in presence and absence of PPase, the same initial slope was obtained, confirming that the observed activity is not caused by a DGC but by a GTP-hydrolyzing enzyme. Incubation of YfiN and GTP without PPase should not lead to any change in signal as it exclusively produces pyrophosphate, which is not detected by the assay.

In the following, the performed experiments are nonetheless described to demonstrate how characterization of the presumable YfiN activity was tackled.

For the first activity test, YfiN was used directly after eluting from SEC without further concentrating the protein. The overall concentration was thus very low, amounting to less than 100 nM protein. The three different SEC peak fractions containing YfiN were kept separate and each of them was screened for potential activity. GTP was present at a concentration of 20 μ M for all experiments and a negative control, containing only GTP and buffer was included as well.

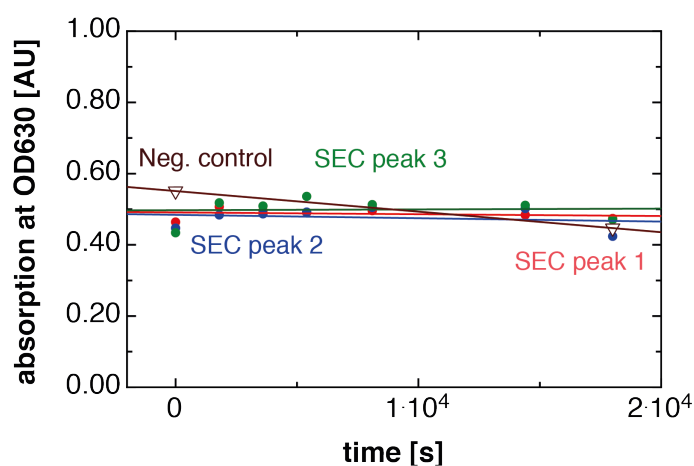


Figure 3-16: Malachite green assay using unconcentrated YfiN directly eluted from SEC. The measured absorption values were plotted against the time and linear fit of the data was performed.

	20 μM GTP	v [$\mu\text{M}/\text{s}$]	v_{corr} ($\mu\text{M}/\text{s}$)
SEC peak 1	+	-9.5E-06	1.0E-04
SEC peak 2	+	-1.8E-05	9.2E-05
SEC peak 3	+	4.7E-06	1.15E-04
no protein	+	-1.1E-04	-

Table 3-7: Summary of the data fitted by linear regression to obtain the initial velocity v of the protein. v_{corr} stands for the corrected initial velocity where the value of the negative control was subtracted to take potential GTP hydrolysis into account.

The starting point of the negative control was clearly an outlier, most probably due to a pipetting error, as the value was even higher than in the experiments containing GTP and protein (Figure 3-16). Overall, no clear trend could be detected for any of the peaks. Instead, the data seemed to rather fluctuate randomly, most probably due to the very low protein concentration, leading to little or no change in signal (Table 3-7). For the next experiments it was thus decided to concentrate the protein after SEC.

In the next experiment, YfiN was present at a final concentration of 2.5 μM and thus had to be strongly concentrated due to repeatedly low protein yields, leading to a simultaneous increase of DM present in the sample. In a second experimental set-up, YfiR was added to YfiN in a final concentration of 7.5 μM , to check if inhibition of the potential DGC activity is taking place. Based on these concentrations, a simulation was made assuming different K_d -values for the YfiR([A0])-YfiN(B0) interaction, what concentration of the complex [AB] is obtained using the equation derived in 2.5.4.1. A complex concentration of 2.5E-06 M therefore corresponds to the

maximum, as it is limited by the lowest concentration of the individual components forming the complex, which was in this case YfiN.

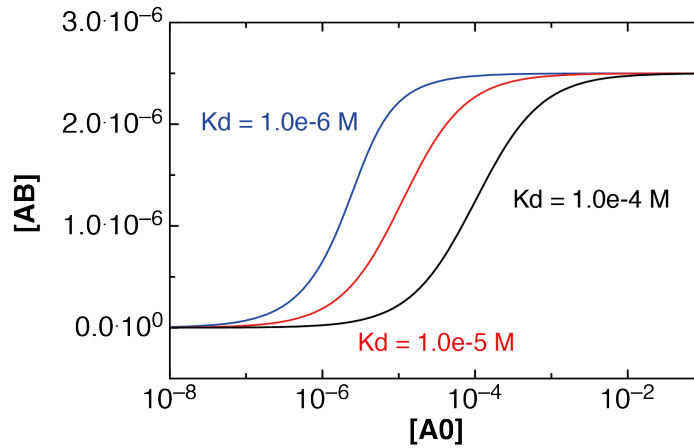


Figure 3-17: Theoretical concentrations of the complex [AB] in relation to varying concentrations of [A0], assuming different Kd-values values for the YfiR-YfiN interaction. [B0] amounted to 2.5E-06 M, thereby posing the upper limit for the expected amount of complex.

[A0] [M]	[B0] [M]	Kd [M]	[AB] [M]
7.5E-6	2.5E-6	1.0E-6	2.1E-6
7.5E-6	2.5E-6	1.0E-5	9.9E-7
7.5E-6	2.5E-6	5.0E-5	3.1E-7
7.5E-6	2.5E-6	1.0E-4	1.7E-7

Table 3-8: Complex concentrations [AB] that are expected assuming different Kd values for the YfiR-YfiN interaction and [A0] = 7.5E-6 M and [B0] = 2.5E-6 M.

According to the simulation, only a Kd of 1.0E-06 M would lead to an almost fully complexed YfiN at the concentrations used in the experiment, whereas for the other Kds, the amount of the expected complex is significantly reduced (Table 3-8).

In addition to the two experiments containing either YfiN alone or YfiN in presence of YfiR, a third experiment was set up, where 25 μ M c-di-GMP

was added to the protein to check if non-competitive product inhibition is taking place. In all three experiments, GTP was present at a concentration of 20 μM . Negative controls included incubation of either the protein alone or GTP alone in buffer and incubation of YfiR in presence of GTP and buffer.

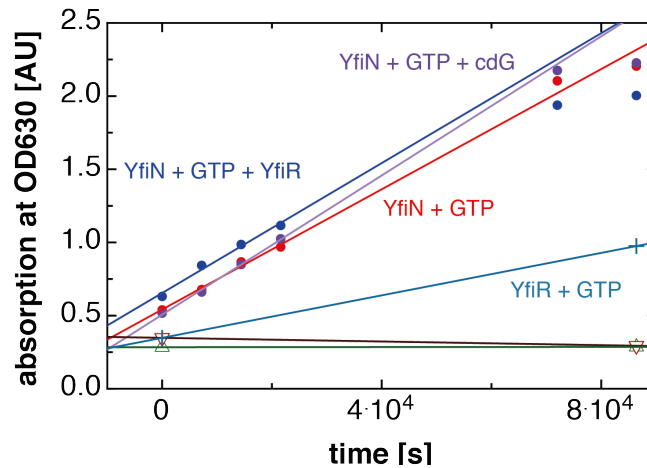


Figure 3-18: The measured absorption values were plotted against the time and linear fit of the data was performed. For all three experiments (YfiN + GTP, YfiN + GTP + YfiR as well as YfiN + GTP + c-di-GMP), only the first four data points were used for the fitting, in order to compare only the initial slopes of the curves and to exclude data that might not be lying in the linear range of the assay. The resulting fits are indicated by solid lines and are colored according to the experiment: YfiN + GTP (red), YfiN + GTP + YfiR (blue), YfiN + GTP + c-di-GMP (purple), YfiR + GTP (blue-green), YfiN + buffer (green) and GTP + buffer (brown).

	20 μM GTP	7.5 μM YfiR	v [$\mu\text{M}/\text{s}$]	v _{corr} [$\mu\text{M}/\text{s}$]
2.5 μM YfiN	-	-	4.3E-07	-
	+	-	3.7E-04	3.8e-04
	+	+	4.2E-04	4.3e-04
2.5 μM YfiN + 25 μM cdG	+	-	4.4E-04	4.5e-04
no protein	+	-	-1.2E-05	-
	+	+	1.4E-04	1.5e-04

Table 3-9: Summary of the data fitted by linear regression to obtain the initial velocity v of the protein in absence of any inhibiting agent or in presence of either YfiR or c-di-GMP. v_{corr} stands for the corrected initial velocity where the value of the corresponding negative control(s) was/were subtracted to take into account potential GTP hydrolysis or other effects. The cells containing the final initial velocities exhibited by YfiN in all conditions investigated are colored according to the same color scheme as in Figure 3-18 to facilitate comparison. Accordingly, the catalytic activity observed in the YfiR sample was colored in green.

As indicated by the data (Table 3-9), there was no significant difference in initial slopes detectable. This can be explained by the fact that the observed signal change was not due to the activity of YfiN but the GTPase contamination present in the sample, which is neither regulated by YfiR nor product-inhibited by c-di-GMP. YfiR most likely was contaminated with a GTPase as well as it showed low enzymatic activity, which was in the same range as the YfiN sample. Both activities appear very low, however considering the fact that YfiN as well as YfiR are 99 % pure according to an SDS-PAGE (Figure 3-19) and thus only very minor amounts of GTPases are present, their k_{cat} -values are actually considerably high.

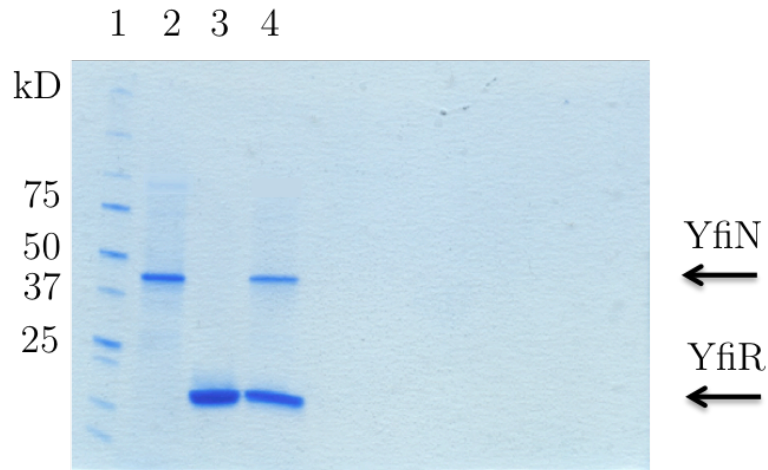


Figure 3-19: An SDS-PAGE illustrating the purity of the YfiN and YfiR stock solution as well as the complex mixture. 1) marker, 2) YfiN stock solution (4.5 μ M), 3) YfiR stock solution (15 μ M), 4) YfiN-YfiR complex sample.

3.3 Individual domains of YfiN

3.3.1 The periplasmic PAS domain

3.3.1.1 Design and cloning of the YfiN_{PAS} constructs

In the beginning of this project, the YfiN sequence was subjected to transmembrane-prediction analysis using TMHMM to determine the boundaries of the periplasmic PAS domain [148]. The first transmembrane-helix was predicted to range from residues 21-43 and the second one from residues 155-177. The sequence corresponding to the PAS domain was then subjected to HHPred analysis, to screen for already existing structures of structurally closely related proteins. This search resulted in four templates with an E-score higher than 10^{-7} (Table 3-10, E-value 2011), with the highest score being 10^{-9} .

PDB identifier	template protein	Sequence identity [%]	E-value, 2011	E-value, 2014
1p0z	CitA	21	1.7E-09	2.8E-05
3b42	GSU0935 MCP	17	2.9E-07	9.1E-07
3by8	DcuS	13	1.0E-08	2.1E-03
3b47	GSU0582 MCP	14	4.4E-07	3.2E-06
3pju	LapD	15	1.6E-04	6.4E-07

Table 3-10: A list of the four proteins identified by HHPred in 2011 that showed an E-value which was larger than 10^{-4} . The corresponding PDB identifiers and the sequence identities between YfiN and the template proteins are indicated. In 2014, the same procedure was repeated, yielding different E-values and an additional template.

In 2014, the same procedure was repeated, thereby revealing that the E-values drastically changed for most proteins. As the PAS domain of LapD ranked now amongst the best, it was included as well for the design of an additional construct. The HHPred alignment of YfiN against the five template proteins revealed that all except LapD proposed the same starting residue at the N-terminal end, which was T44 (data not shown). In the case of LapD, the alignment suggested a slightly longer construct, starting at residue T38. For the design of the C-terminal end, again slightly different construct lengths were obtained. In CitA and DcuS, the structures revealed a well-ordered C-terminal end, which was found to be forming $\alpha 6$ of the classical periplasmic PAS fold (Figure 1-3). Based on these two structures, a construct was designed, which was encompassing the whole predicted PAS domain of YfiN, ranging from residues 44-154 (Table 3-11).

The structures of 3b42 and 3b47 revealed a C-terminal end, which was involved in β -strand formation, followed by an unstructured linker of roughly 10 residues. The end of the structured region of the proteins aligned to residue G149 in YfiN and therefore a second, slightly shorter construct was generated accordingly, encompassing residues 44-149 (Table 3-11).

In agreement with CitA and DcuS, the C-terminal end of LapD was found to be well-structured and engaged in helix formation. However, the last ordered residue of the construct was matched to residue G159 in YfiN in the HHPred alignment, thus suggesting a slightly longer construct than CitA and DcuS, encompassing residues 38-159 (Table 3-11).

YfiN_{PAS} construct length	Template proteins	PDB identifier
38-159	LapD	3pju
44-149	GSU0935	3b42
	GSU0582	3b47
44-154	CitA	1p0z
	DcuS	3by8

Table 3-11: Overview of the YfiN_{PAS} constructs investigated within the framework of this thesis and the underlying template proteins used for their design.

3.3.1.2 Expression of the YfiN_{PAS} constructs

All three C-terminally his-tagged YfiN_{PAS} constructs (38-159, 44-149, 44-154) exclusively got expressed in inclusion bodies, with only very little soluble expression for YfiN_{PAS44-149} (data not shown). Neither expression of the constructs at 10 °C overnight in ArcticExpress™ nor the usage of the BL21-AI™ strain to control potential leaky expression resulted in increased yields of soluble protein. Therefore, it was decided to fuse YfiN_{PAS38-159} and YfiN_{PAS44-149} to either a SUMO-(12 kD), GST- (26 kD) or MBP-tag (42 kD). Expression tests of the N-terminally fusion-tagged YfiN_{PAS} constructs were performed in either LB- or TB-medium using the BL21(DE3) strain at either

Results

20 °C or 30 °C . Protein expression was induced by the addition of 250 μ M or 50 μ M IPTG in the case of LB-medium or TB-medium, respectively. Overall, expression levels did not change in relation to the presence or absence of a C-terminal his-tag, therefore only the his-tagged constructs will be considered in the following.

MBP-tagged YfiN_{PAS38-159} -and YfiN_{PAS44-149}, led to the highest expression levels using TB-medium and an expression duration of 4 hours at 30 °C for YfiN_{PAS38-159} and 20 °C overnight for YfiN_{PAS44-149}. Exemplarily, two expression gels are shown in Figure 3-20. For protein production, bacteria were cultivated in a 2 L scale with the conditions established in the expression tests.

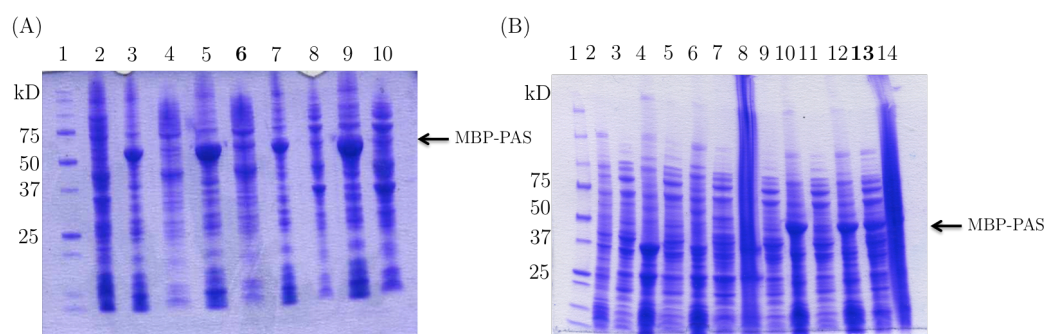


Figure 3-20: SDS-PAGE of the expression tests of YfiN_{PAS38-159} and YfiN_{PAS44-149}. The condition chosen for expression is marked in bold. The same amount of cells was loaded in each lane. **(A)** YfiN_{PAS38-159} : 1: marker, 2: before induction, 3: insoluble fraction (2h of expression at 30 °C), 4: soluble fraction (2h of expression at 30 °C), 5: insoluble fraction (4h of expression at 30 °C), 6: soluble fraction (4h of expression at 30 °C), 7: insoluble fraction (5h of expression at 20 °C), 8: soluble fraction (5h of expression at 20 °C), 9: insoluble fraction (overnight expression at 20 °C), 10: soluble fraction (overnight expression at 20 °C). **(B)** YfiN_{PAS44-149}: lanes 3-8 correspond to the GST-tagged construct, whereas lanes 9-14 correspond to the MBP-tagged construct. 1: marker, 2: before induction, 3: soluble fraction (3h of expression at 30 °C), 4: insoluble fraction (3h of expression at 30 °C) 5: soluble fraction (5h of expression at 20 °C), 6: insoluble fraction (5h of expression at 30 °C), 7: soluble fraction (overnight expression at 20 °C), 8: insoluble fraction (overnight expression at 20 °C), lanes 9-14: same as lanes 3-8 just for the MBP-tagged construct.

3.3.1.3 Purification of the YfiN_{PAS} constructs

C-terminally his-tagged YfiN_{PAS44-149} was purified as described in 2.2.2.2. The SEC revealed one peak eluting at 40 mL, corresponding to aggregated protein (data not shown). In addition, SDS-PAGE analysis revealed the presence of numerous contaminants, eluting together with in the aggregated YfiN_{PAS44-149} during SEC. As only very minor amounts of the protein of interest were obtained throughout the experiment, the focus was shifted towards refolding the protein expressed in inclusion bodies.

Inclusion bodies were prepared and solubilized as described in 2.2.2.3. As expected, large amounts of refolded protein were obtained after IMAC, however the subsequent SEC revealed the presence of exclusively aggregated YfiN_{PAS44-149} once again (data not shown). As all the C-terminally his-tagged PAS constructs were either not expressed in a soluble form or showed minor soluble expression but aggregation-prone behavior, the focus was shifted towards the fusion-tagged constructs at this stage.

Both MBP-tagged YfiN_{PAS38-159} and YfiN_{PAS44-149} constructs were purified as described in 2.2.2.2. In the following, only results from YfiN_{PAS38-159} are shown as both constructs behaved in the same way. A chromatogram of a typical run is shown in Figure 3-21(A) and the corresponding SDS-PAGE is shown in Figure 3-21(B).

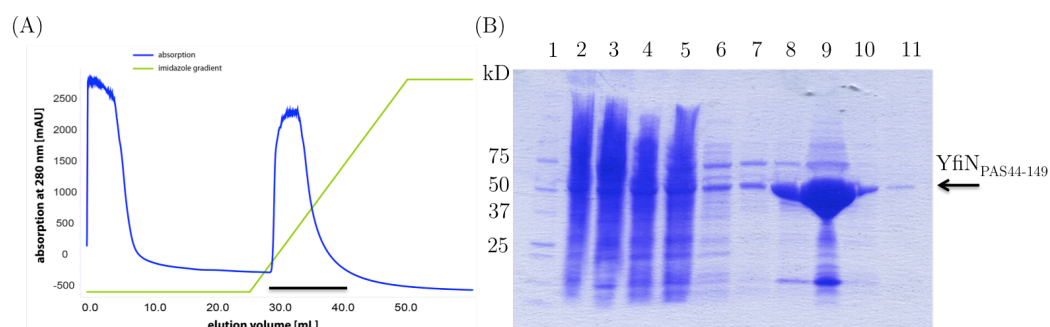


Figure 3-21: IMAC purification step of MBP-tagged YfiN_{PAS38-159}. **(A)** Chromatogram of YfiN_{PAS38-159} purification by IMAC using 5 mL MBPTrap column. The maltose gradient is ranging from 0 mM to 10 mM in 5 CV. Fractions pooled and used for subsequent purification are marked by a line. **(B)** SDS-PAGE illustrating each step of the purification: 1: marker, 2: whole cell, 3: load MBPTrap, 4: flow-through MBPTrap, 5: IMAC 0 mL-5 mL, 6: IMAC 5 mL-10 mL, 7: IMAC 10 mL-15 mL, 8: IMAC 25 mL-30 mL, 9: IMAC 30 mL-35 mL, 10: IMAC 35 mL-40 mL.

MBP-tagged YfiN_{PAS38-159} started eluting at a maltose concentration of about 1.5 mM, which resulted in a high yield of almost pure fusion-protein. The SDS-PAGE showed an additional band running slightly below 15 kD, which indicated the presence of the C-terminally his-tagged YfiN_{PAS38-159} lacking the MBP-tag. After Ni-chromatography, MBP-tagged YfiN_{PAS38-159} was subjected to gel filtration as a final purification step.

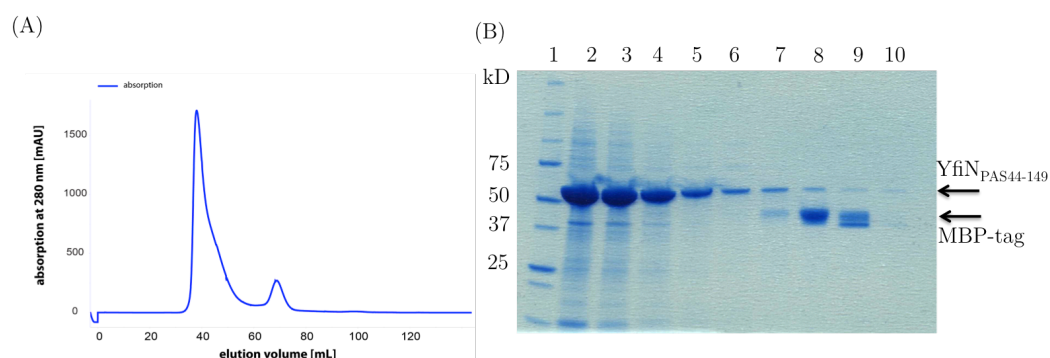


Figure 3-22: SEC purification step of MBP-tagged YfiN_{PAS38-159} using a S75 16/60 column. **(A)** Chromatogram of MBP-tagged YfiN_{PAS38-159} purification by gel filtration chromatography using a HiLoad- 16/60-Superdex-75 prep grade column. **(B)** SDS-PAGE illustrating each step of the purification: 1:marker, 2: SEC 35 mL-40 mL, 3: SEC 40 mL-45 mL, 4: SEC 45 mL-50 mL, 5: SEC 50 mL-55 mL, 6: SEC 55 mL-60 mL, 7: SEC 60 mL-65 mL, 8: SEC 65 mL-70 mL, 9: SEC 70 mL-75 mL, 10: SEC 75 mL-80 mL.

Most of the MBP-tagged YfiN_{PAS38-159} was eluting in the form of aggregates at 40 mL, resulting in a severely trailing peak (Figure 3-22). A second peak at 70 mL was observed containing the free MBP-tag, most probably stemming from sample degradation. In an attempt to stabilize the fusion-protein, two additional purifications were performed with buffers that were either supplemented with 0.2 % Tween-20 or 50 mM arginine and glutamate (data not shown). However, none of these measures led to an optimization of protein behavior and the approach, focusing on the isolated PAS domain was thus terminated at this stage. Table 3-12 is giving a short overview of all the results obtained for the YfiN_{PAS} constructs investigated within the frame of this thesis.

YfiN _{PAS} construct	Tag(s)	Size of fusion-tag [kDa]	Status
38-159	C-his	0.9	No soluble expression in BL21(DE3) at different expression temperatures.
	SUMO, +/- C-his	12.0	No soluble expression in LB and TB medium at the temperatures tested (20 °C, 30 °C)
	GST, +/- C-his	26.0	Little soluble expression in TB medium, however the protein did not bind to a GST-trap in a first experiment
	MBP, cyto. and peri., +/- C-his	42.5	Abundant soluble expression in BL21(DE3) using TB-medium but SEC revealed a severely trailing aggregation peak (30ml) on a S75 16/60.
44-149	C-his	0.9	No soluble expression in ArcticExpress™ or BL21-AI™. Little soluble expression in BL21(DE3), but SEC revealed loads of impurities and aggregation-prone protein. On-column refolding of inclusion bodies led to aggregated protein.
	SUMO, +/- C-his	12	No soluble expression in LB and TB medium at all temperatures tested (20 °C, 30 °C and 37 °C)
	GST, +/- C-his	26	No soluble expression in LB and TB medium at the temperatures tested (20 °C, 30 °C)
	MBP, cyto. and peri., +/- C-his	42.5	Abundant soluble expression in BL21(DE3) using TB-medium, however severe aggregation revealed by SEC. Addition of either detergents (0.2 % Tween-20) or 50 mM Arg/Glu to the buffers did not improve the quality of the protein.
44-154	C-his	0.9	No soluble expression in ArcticExpress™, BL21-AI™ and BL21(DE3)

Table 3-12: A summary of all the results that were obtained for the YfiN_{PAS} constructs investigated within the framework of this thesis. The color code is the same as for Table 3-1.

3.3.2 The cytoplasmic HAMP and GGDEF domains

3.3.2.1 Design and cloning of the YfiN_{HAMP-GGDEF} constructs

In *P. aeruginosa*, the second transmembrane helix from YfiN is predicted to range from 155-177 as determined before in section 3.3.1.1. The cytoplasmic part is therefore ranging from residues 178-435. This sequence was then analyzed by Pfam, SMART and PROSITE [149] to identify the boundaries of the cytoplasmic HAMP domain. Both SMART and PROSITE predicted the HAMP domain to range from residues 183-236, whereas Pfam predicted 164-236, including a part of the transmembrane region. The cytoplasmic construct was therefore designed to start at residue 178 to add five additional amino acids between the start of the construct and the HAMP domain in order to raise the chances for proper folding.

The C-terminal end was designed in analogy to the full-length protein (3.2.1.1), resulting in a *P. aeruginosa* YfiN_{HAMP-GGDEF} construct that encompassed the whole cytoplasmic region from residues 178 to 435. Based on this construct, a search was launched as described in 2.6.1 to identify YfiN homologs that were predicted to crystallize more readily than the one from *P. aeruginosa*. Several additional criteria had to be met by the potential candidates, such as the presence of a PAS, HAMP and a GGDEF domain as well as an intact operon coding for all three Yfi proteins. Ultimately, three homologs were chosen, meeting all of the above stated criteria: YfiN from *Pseudomonas fluorescens*, sharing 66 % sequence identity with YfiN from *P. aeruginosa*, YfiN from *Yersinia enterocoliticae* with a sequence identity of 58 % and YfiN from *Sphingopyxis alaskensis* with a sequence identity of 48 % (Figure 3-23, Table 3-13).

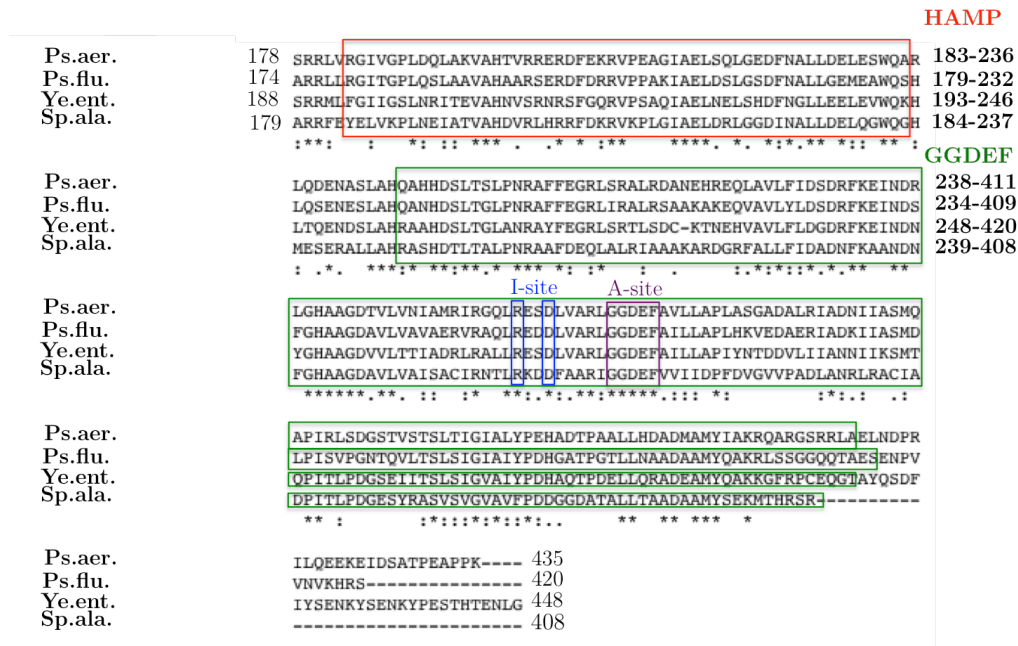


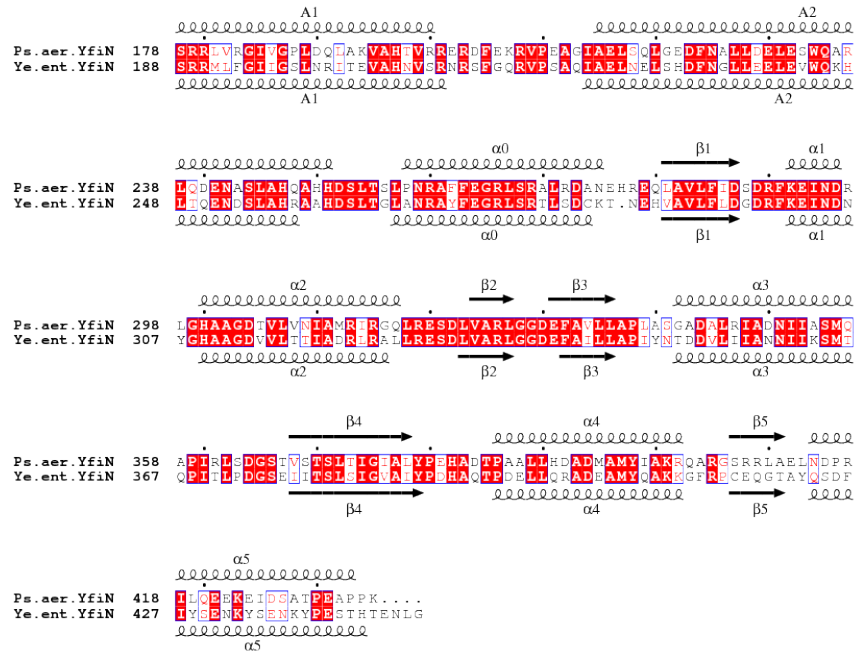
Figure 3-23: Multiple sequence alignment of the cytoplasmic part of YfiN homologs from four different species. Ps.aer. stands for *P. aeruginosa*, Ps.flu. for *P. fluorescens*, Ye.ent. for *Yersinia enterocolitica* and Sp.ala. for *Sphingopyxis alaskensis*. The numbers of the starting and terminal residues are indicated. The HAMP- and GGDEF domain residues are marked with a red box and green box respectively as predicted by SMART and PROSITE. Functionally important residues as the I-site (blue box) or the A-site (purple box) are indicated separately. The multiple sequence alignment was generated by CLUSTAL OMEGA [142] [143].

Species	Length of cytoplasmic part	Molecular weight [kD]	Sequence identity to Ps.aer. YfiN [%]	XtalPred crystallization class
<i>P. aeruginosa</i>	178-435	29.7	100	3
<i>P. fluorescens</i>	174-420	27.8	66	2
<i>Y. enterocoliticae</i>	188-448	30.4	58	2
<i>S. alaskensis</i>	179-408	26.2	48	1

Table 3-13: Table summarizing the properties of all four YfiN homologs.

Constructs for the three homologs were designed in the same way as for *Ps.aer.*, as all of them were predicted to have a well-structured C-terminal end, although of slightly different length (Figure 3-24).

(A)



(B)

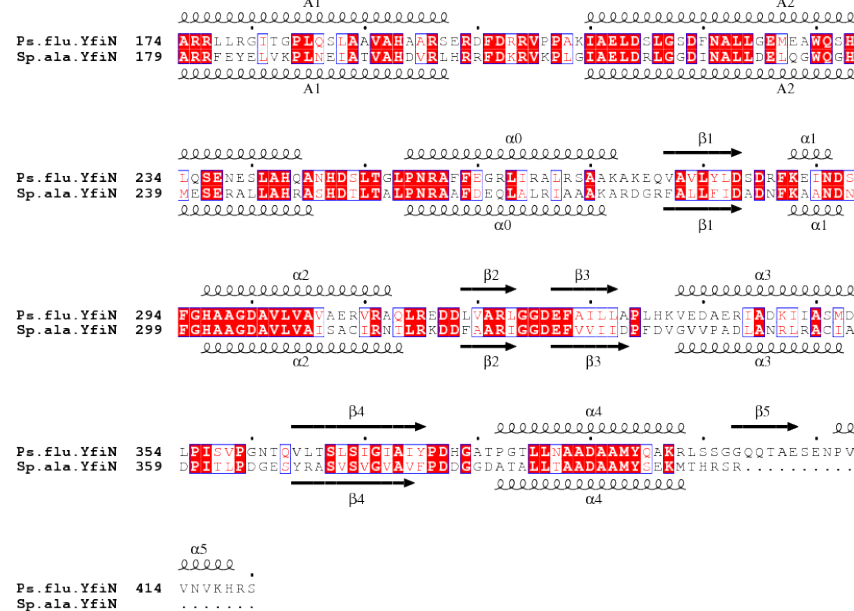


Figure 3-24: Pairwise multiple sequence alignment of YfiN_{HAMP-GGDEF} from *Ps.aer.* and *Ye.ent.* (A) as well as *Ps.flu.* and *Sp.ala.* (B)... Secondary structural elements were predicted by Jpred [140] and are either shown on the top or the bottom of the sequences. Capital letters refer to the HAMP nomenclature and Greek letters refer to the GGDEF nomenclature. Strictly conserved residues are boxed in white on a red background, and highly conserved residues are boxed in red on a white background. Every tenth residue is indicated with a dot (.). The α -helix is depicted by a coil, whereas β -strands are depicted as an arrow. The alignment was performed using CLUSTAL OMEGA [142] [143]. The figure was generated using ESPrpt [144].

3.3.2.2 Expression of the YfiN_{HAMP-GGDEF} constructs

Expression tests of the four YfiN_{HAMP-GGDEF} were performed either in BL21(DE3) or BL21 StarTM (DE3) at three different temperatures of 20 °C, 30 °C and 37°C. Protein expression was induced with 250 µM IPTG. In most of the tested conditions, the constructs were expressed in abundant amounts in the soluble fraction. The best expression levels were achieved in BL21(DE3) at 30 °C using 5 h of induction. Exemplarily an expression gel for each construct is shown in Figure 3-25. For protein production, bacteria were cultivated on a 2 L scale with the conditions established in the expression tests.

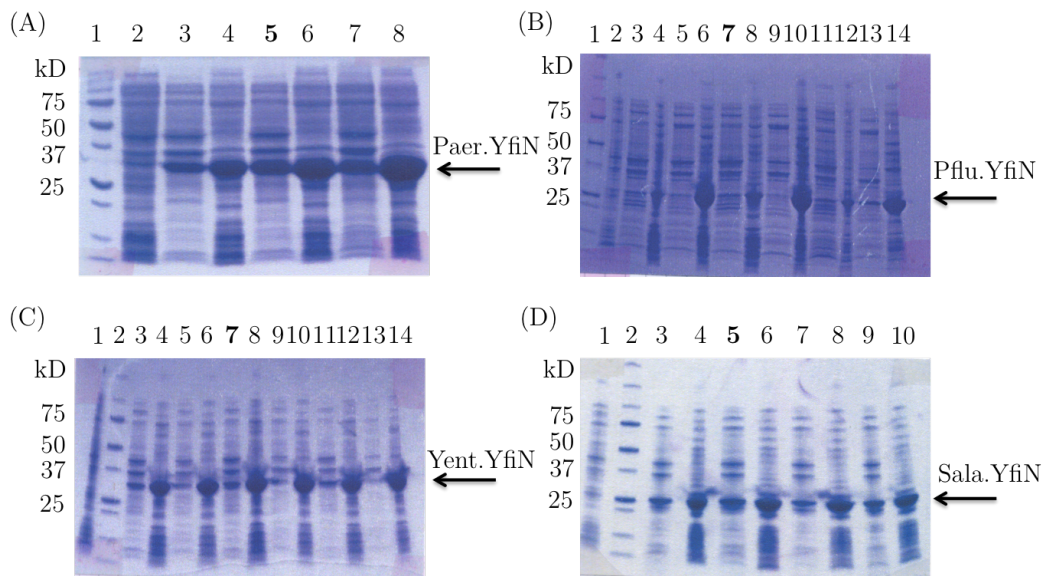


Figure 3-25: SDS-PAGE of the expression tests of YfiN_{HAMP-GGDEF} from different species. The condition chosen for expression is marked in bold. The same amount of cells was loaded in each lane. **(A)** Ps.aer. YfiN_{HAMP-GGDEF}: 1: marker, 2: before ind., 3: soluble fraction (3h of expression at 30 °C), 4: insoluble fraction (3h of expression at 30 °C), 5: soluble fraction (5h of expression at 30 °C), 6: insoluble fraction (5h of expression at 30 °C), 7: soluble fraction (7h of expression at 30 °C), 8: insoluble fraction (7h of expression at 30 °C) **(B)** Ps.flu. YfiN_{HAMP-GGDEF}: 1: marker, 2: before induction, 3: soluble fraction (3h of expression at 30 °C), 4: insoluble fraction (3h of expression at 30 °C), 5: soluble fraction (3h of expression at 37 °C), 6: insoluble fraction (3h of expression at 37 °C), 7: soluble fraction (5h of expression at 30 °C), 8: insoluble fraction (5h of expression at 30 °C), 9: soluble fraction (5h of expression at 37 °C), 10: insoluble fraction (5h of expression at 37 °C), 11: soluble fraction (7h of expression at 30 °C), 12: insoluble fraction (7h of expression at 30 °C), 13: soluble fraction (7h of expression at 37 °C), 14: insoluble fraction (7h of expression at 37 °C). **(C)** Ye.ent. YfiN_{HAMP-GGDEF}: the sample order is the same as in (B), except for 1: before induction, 2: marker. **(D)** Sp.ala. YfiN_{HAMP-GGDEF}: 1: before induction, 2: marker, 3: soluble fraction (3h of expression at 30 °C), 4: insoluble fraction (3h of expression at 30 °C), 5: soluble fraction (5h of expression at 30 °C), 6: insoluble fraction (5h of expression at 30 °C), 7: soluble fraction (3h of expression at 37 °C), 8: insoluble fraction (3h of expression at 37 °C), 9: soluble fraction (5h of expression at 37 °C), 10: insoluble fraction (5h of expression at 37 °C).

3.3.2.3 Purification of the YfiN_{HAMP-GGDEF} constructs

Ps.aer. YfiN_{HAMP-GGDEF} was purified to homogeneity using a two-step purification procedure, consisting of Ni-affinity and size-exclusion chromatography. After cell lysis with French press, Ni-affinity chromatography using a HisTrap column was performed. A chromatogram of a typical run is shown in Figure 3-26(A) and the corresponding SDS-PAGE is shown in Figure 3-26(B). As imidazole concentrations higher than 270 mM led to almost exclusively aggregated protein (data not shown) the IMAC protocol was modified, applying a less steep gradient, which was increasing over 15 CV instead of 10 CV. Furthermore, all buffers were supplemented with 5 % glycerol to provide a stabilizing effect. Owing to the optimized protocol, Ps.aer. YfiN_{HAMP-GGDEF} was eluting almost completely pure from the HisTrap at imidazole concentrations of about 150 mM-230 mM (Figure 3-26(A)). After Ni-chromatography, the protein was subjected to gel filtration as a final purification step. Ps.aer. YfiN_{HAMP-GGDEF} ran as one single peak at an elution volume of 70 mL (Figure 3-26(C)). According to a calibration curve for the present column, this corresponds to a molecular mass of 40 kDa. The mass of one Ps.aer. YfiN_{HAMP-GGDEF} chain is 29.7 kD and considering the elongated shape of the protein, this elution volume most likely corresponds to a monomer. The final yield was about 2.5 mg/L culture. At the end of the SEC run, a second peak was observed with an inversed 254/280 nm ratio, clearly indicating the presence of a nucleotide that was bound to YfiN. The same has been observed before for the full-length protein in 3.2.1.5 (Figure 3-13), where the co-eluting nucleotide was identified as c-di-GMP.

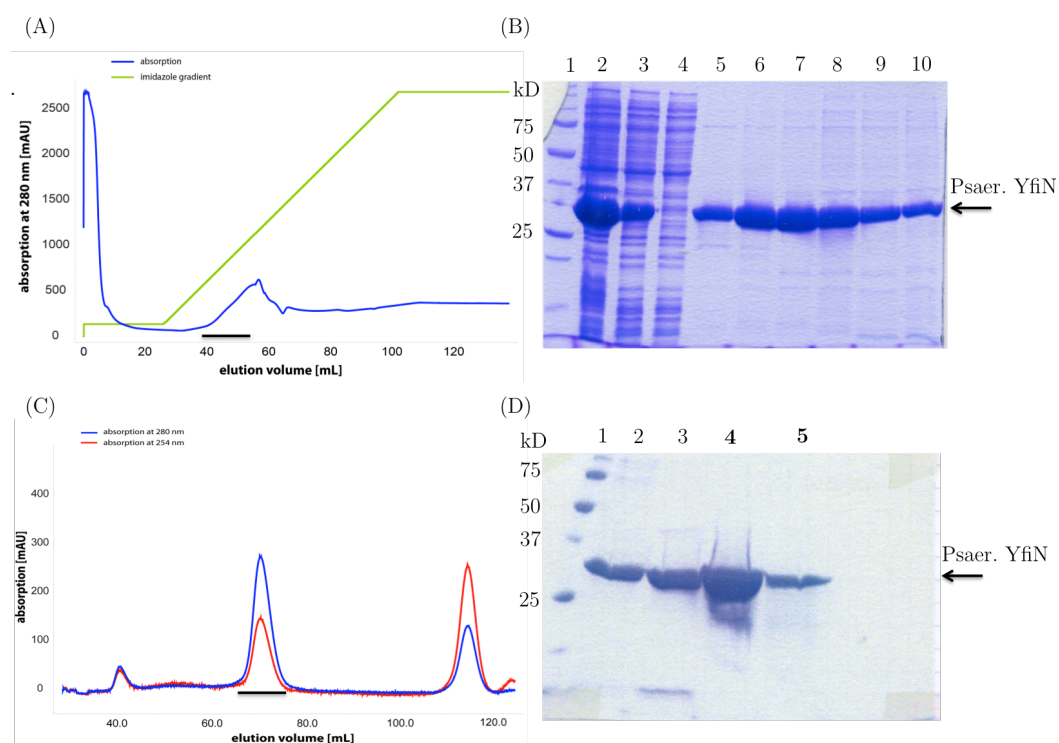


Figure 3-26: IMAC and SEC purification steps of Ps.aer. YfiN_{HAMP-GGDEF}. **(A)** Chromatogram of YfiN purification by IMAC using a 5 mL HisTrap column. Fractions pooled and used for subsequent purification are marked by a line. **(B)** SDS-PAGE illustrating each step of YfiN purification: 1: marker, 2: whole cell, 3: load IMAC, 4: flow-through IMAC, 5: IMAC 40 mL-45 mL, 6: IMAC 45 mL-50 mL, 7: IMAC 50 mL-55 mL, 8: IMAC 55 mL-60 mL, 9: IMAC 60 mL-65 mL, 10: IMAC 65 mL-70 mL. **(C)** Chromatogram of YfiN purification by gel filtration chromatography using a HiLoad- 16/60-Superdex-75 prep grade column. Pooled YfiN_{HAMP-GGDEF} fractions are marked by a line. **(D)** SDS-PAGE illustrating each step of Ps.aer. YfiN_{HAMP-GGDEF} purification. Relevant lanes are marked in bold. 1: marker, lane 2 and 3: YfiN_{HAMP-GGDEF} samples from a different preparation, 4: SEC 63 mL-68 mL, 5: SEC 68 mL-73 mL.

Surprisingly, all four YfiN_{HAMP-GGDEF} constructs appeared to be prone to varying degrees of proteolysis during the process of purification. In some cases, the large amounts of protein on the SDS-PAGE impeded proper differentiation between one and several bands that were running closely together (Figure 3-26(D), lane 4). In the following, examples are shown for P.flu. (Figure 3-27) and Y.ent. YfiN_{HAMP-GGDEF} (Figure 3-28), purified in the same way as described for Ps.aer, illustrating the appearance of either a double band or even a triple band after SEC. Sp.ala. YfiN_{HAMP-GGDEF} was subjected to proteolysis as well (data not shown).

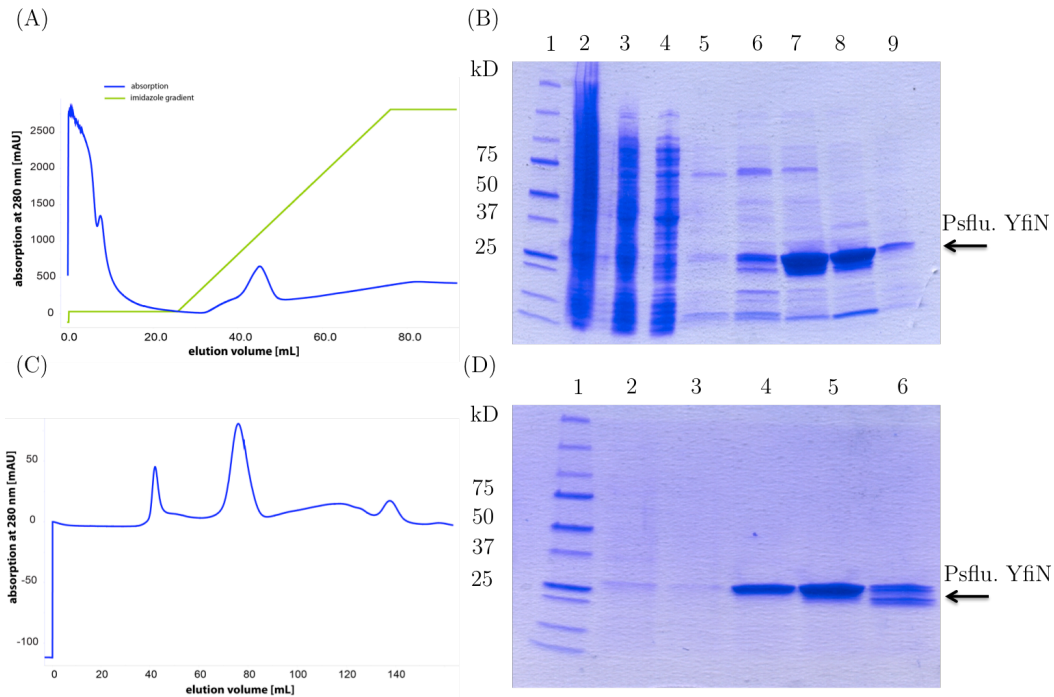


Figure 3-27: IMAC and SEC purification steps of Ps.flu. YfiN_{HAMP-GGDEF}. **(A)** Chromatogram of YfiN purification by IMAC using a 5 mL HisTrap column. Fractions pooled and used for subsequent purification are marked by a line. **(B)** SDS-PAGE illustrating each step of YfiN purification: 1: marker, 2: whole cell, 3: load IMAC, 4: flow-through IMAC, 5: IMAC 30 mL-35 mL, 6: IMAC 35 mL-40 mL, 7: IMAC 40 mL-45 mL, 8: IMAC 45 mL-50 mL, 9: IMAC 50 mL-55 mL. **(C)** Chromatogram of YfiN purification by gel filtration chromatography using a HiLoad- 16/60-Superdex-75 prep grade column. **(D)** SDS-PAGE illustrating each step of Ps.flu. YfiN_{HAMP-GGDEF} purification: 1: marker, 2: SEC 40 mL-45 mL, 3: SEC 45 mL-50 mL, 4: SEC 70 mL-75 mL, 5: SEC 75 mL-80 mL, 6: SEC 80 mL-85 mL.

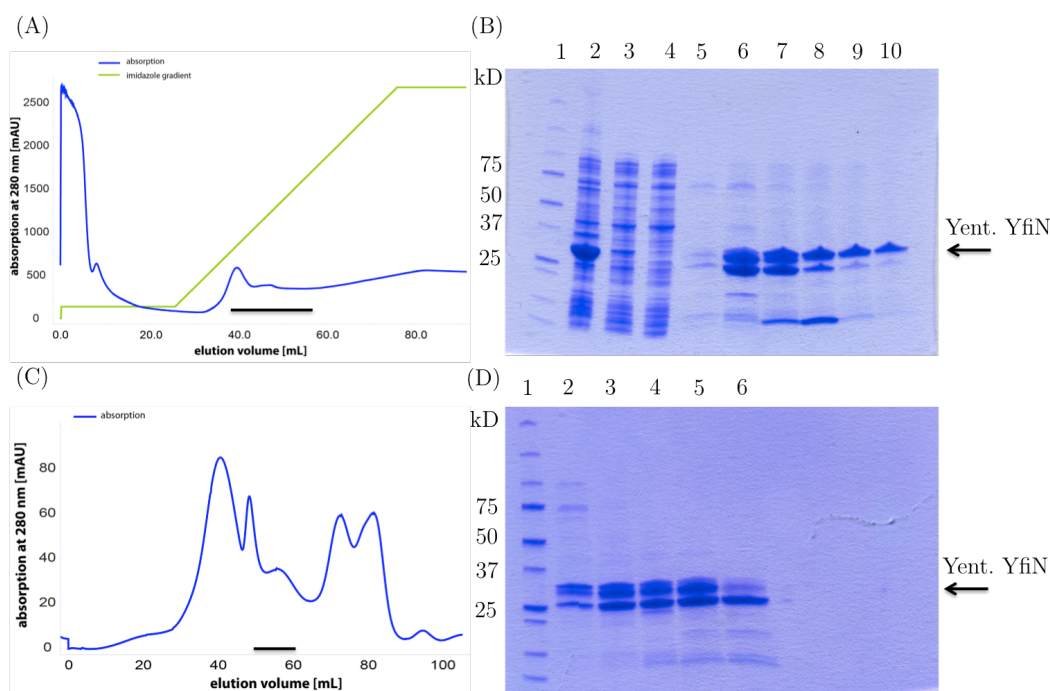


Figure 3-28: IMAC and SEC purification steps of Ye.ent. YfiN_{HAMP-GGDEF}. (A) Chromatogram of YfiN purification by IMAC using a 5 mL HisTrap column. Fractions pooled and used for subsequent purification are marked by a line. (B) SDS-PAGE illustrating each step of YfiN purification: 1: marker, 2: whole cell, 3: load IMAC, 4: flow-through IMAC, 5: IMAC 30 mL-35 mL, 6: IMAC 35 mL-40 mL, 7: IMAC 40 mL-45 mL, 8: IMAC 45 mL-50 mL, 9: IMAC 50 mL-55 mL, 10: SEC 55 mL-60 mL. (C) Chromatogram of YfiN purification by gel filtration chromatography using a HiLoad- 16/60-Superdex-75 prep grade column. Fractions used for crystallization containing non-aggregated species that did not show proteolysis are marked by a line (data not shown). (D) SDS-PAGE illustrating each step of Ps.flu. YfiN_{HAMP-GGDEF} purification: 1: marker, 2: SEC 40 mL-45 mL, 3: SEC 45 mL-50 mL, 4: SEC 70 mL-75 mL, 5: SEC 75 mL-80 mL, 6: SEC 80 mL-85 mL.

In order to identify the degradation products, the three Ye.ent. YfiN_{HAMP-GGDEF} bands detected on a SDS-PAGE (Figure 3-28) were cut out and sent for tryptic digest and mass spectrometry analysis.

3.3.2.4 Characterization of the degradation products by tryptic digest/MS

For all three samples, the generated peptides were found distributed throughout the protein sequence, indicating that the whole protein was flying in the mass spectrometer (Figure 3-29). However, the frequencies at

which some of some of the observed peptides occurred quite drastically changed depending on the investigated sample. For simplicity reasons, the three samples are referred to as follows: sample 1= upper band of the observed double band, sample 2= lower band of the observed double band, and sample 3= the lowest band appearing on the SDS-PAGE (Figure 3-28).

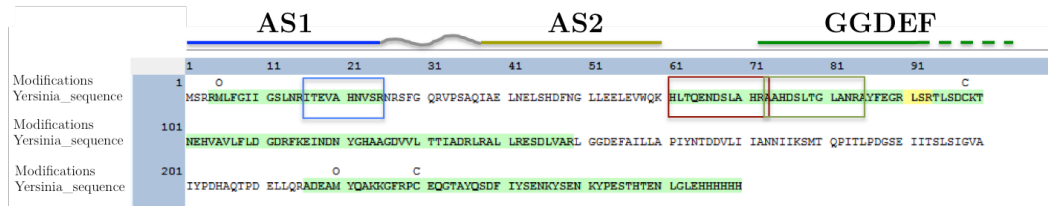


Figure 3-29: Sequence of the cytoplasmic part of *Ye.ent.* Yfi^{NHAMP}-GGDEF. The parts of the protein that were observed in the mass spectrometry analysis are underlined in green. The peptides relevant for the identification of the degradation products are boxed in blue, red and green. The two HAMP helices, the interconnecting linker and the beginning of the GGDEF domain are represented schematically.

Position	Sequence	observed area in sample 1	observed area in sample 2	observed area in sample 3	Location of cleavage site
14-25	ITEVAHNVSR	9.02e+08	1.38e+09	8.32e+08	full-length cytoplasmic construct
61-72	HLTQENDSLAHR	1.67e+07	4.56e+09	4.62e+08	within the HAMP domain
72-84	AAHDSLTLGLANR	5.85e+06	3.73e+07	6.47e+08	between the HAMP and GGDEF domain

Table 3-14: Table summarizing the position and sequence of the relevant peptides as well as their observed frequencies depending on the sample.

Mass spectrometry analysis revealed that the protein got cleaved at two different cleavage sites, one located within the HAMP domain and one

situated between the HAMP and the GGDEF domain. These findings are discussed in more detail in section 3.5.4.2. Nonetheless, the cleaved products could usually be separated from the full-length cytoplasmic YfiN constructs and crystallization was pursued with *Ps.aer.* and *Ye.ent.* YfiN_{HAMP-GGDEF}.

3.3.2.5 Initial crystallization attempts of *Ps.aer.* and *Ye.ent.* YfiN_{HAMP-GGDEF}

For the crystallization trials of *Ps.aer.* and *Ye.ent.* YfiN_{HAMP-GGDEF}, commercial crystallization screens were employed at different protein concentrations in 96-well sitting drop set-ups. All trials are listed in Table 3-15. As the first attempts to obtain crystals were not successful, both constructs were transferred to SEC-buffer without the addition of glycerol. However, no crystals were growing under such conditions either. Instead, severe oiling out and precipitation was observed in many of the conditions, most probably due to the heterogeneity of the proteins in question. The attempt to trap *Ps.aer.* YfiN_{HAMP-GGDEF} in an inhibited state by the addition of c-di-GMP in a 1:3 molar ratio resulted in protein precipitation. Furthermore, it has never been investigated if the full-length cytoplasmic constructs eluted from SEC might still degrade in a time-dependent fashion, further impeding crystallization. As initial crystallization attempts failed, it was decided to do a functional characterization of the *Ps.aer.* YfiN_{HAMP-GGDEF} construct.

construct	screen	protein concentration [mg/mL]
Ye.ent. YfiN _{HAMP-GGDEF}	PACT Premier	6.7
	Index	6.7
		13.1
	PEGR _x	6.7
		13.1
	PEG-Ion	13.1
Structure Screen I+II	6.7	
Ps.aer. YfiN _{HAMP-GGDEF}	PACT premier	1
		7.6
		15.7
	JCSG	1
		3.1
		7.6
	PEGR _x	7.6
	Index	7.6
	CSS I	7.6
	Structural Screen I + II	7.6

Table 3-15: Crystallization trials for Ps.aer. and Ye.ent. YfiN_{HAMP-GGDEF}.

3.3.2.6 Functional characterization of Ps.aer. YfiN_{HAMP-GGDEF} by the phosphate sensor assay

3.3.2.6.1 Mechanistic principle of the phosphate sensor assay

The phosphate sensor is a simple assay system used to directly measure the amount of inorganic phosphate generated in an enzymatic reaction. It is based on a phosphate-binding protein from *E. coli*, chemically modified with the MDCC fluorophore [150], which upon binding of inorganic phosphate increases its fluorescence between 6-8 times. This simple direct measurement of the release of phosphate is not dependent on a specific substrate or enzyme, making it suitable for almost any target of interest. In order to

study DGC activity, the assay can be coupled with a second enzyme pyrophosphatase, which will convert the pyrophosphate, generated during the process of c-di-GMP production, to inorganic phosphate (Figure 3-30). This slight modification of the assay allows on-line quantification of DGC activity, which is less laborious than methods depending on time point measurements (2.5.1; 2.5.3).

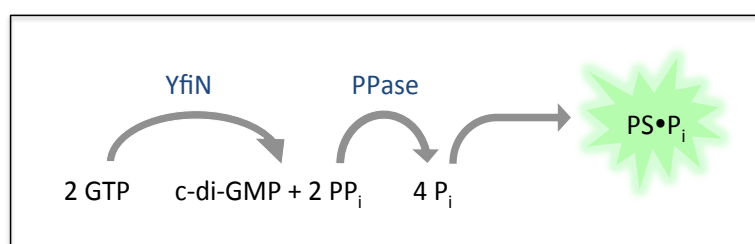


Figure 3-30: Mechanistic principle of phosphate detection by the phosphate sensor assay as a method to quantify DGC activity

3.3.2.6.2 Establishing the assay for DGCs using the well characterized DGC DgcZ as a reference protein

The phosphate sensor has been previously established for the characterization of the PDE YahA. In order to establish the assay for the characterization of DGC activity, the well-characterized DGC DgcZ was used as a reference protein. 200 nM zinc-free DgcZ was incubated with either 5 μM , 10 μM , 20 μM , 50 μM or 100 μM GTP and phosphate production was monitored in real-time as an indirect output for DGC activity. Global fitting of all datasets to a Michaelis-Menten kinetic model yielded a k_{cat} -value of $6.2\text{E-}02 \text{ s}^{-1}$ (data not shown), which is in perfect agreement with the published k_{cat} value of 0.6 s^{-1} [111]. The assay was

therefore applied in a next step for the characterization of Ps.aer. YfiN_{HAMP-GGDEF}.

3.3.2.6.3 Enzymatic characterization of Ps.aer. YfiN_{HAMP-GGDEF}

5 μM Ps.aer. YfiN_{HAMP-GGDEF} was incubated with 50 μM GTP and the increase in fluorescence was monitored for 20 min as described in 2.5.2. The obtained data points were then fit to the Michaelis-Menten model, described in **Error! Reference source not found.** to obtain a k_{cat} -value of $3.4\text{E-}04$ s^{-1} . For a detailed discussion of the results, refer to chapter 3.5.4.2.

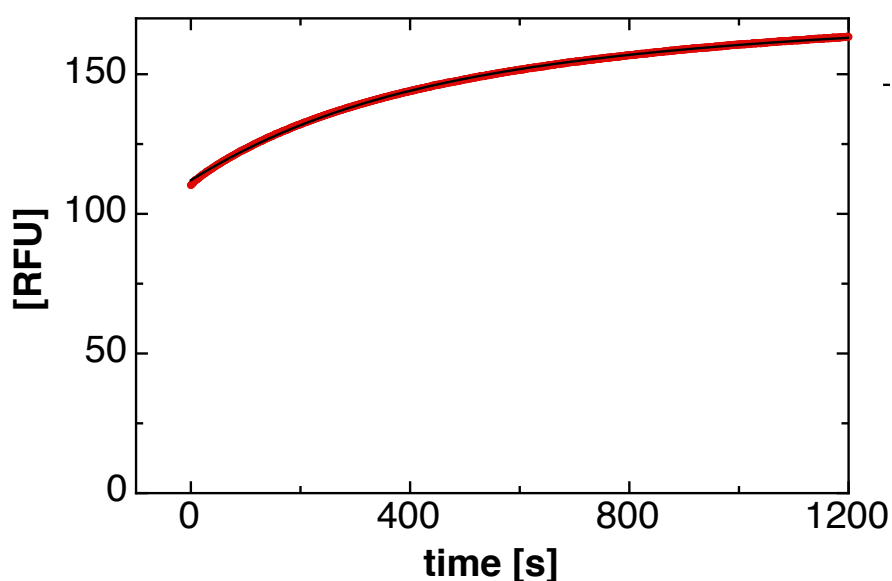


Figure 3-31: The measured increase in fluorescence was plotted against the time and the data was fit to a Michaelis-Menten model to yield a k_{cat} of $3.2\text{e-}4$.

3.4 Bioinformatical analysis of YfiN: a model of the product-inhibited and the active state of YfiN

As described in 1.3, most DGCs are subjected to non-competitive product inhibition, involving c-di-GMP mediated cross-linking between either the two DGC domains or the DGC domain and an accessory domain [23] [24]. Prerequisite for this binding mode is the presence of an intact primary I-site (RxxD) on the DGC domain, and a secondary I-site either on the DGC or the accessory domain. The crystal structure of the YfiN_{CGDEF} domain, which was solved by Giardina et al. [151] revealed a degenerate secondary I-site, where two of the three arginines required for c-di-GMP binding were missing. Therefore, a bioinformatical approach was set-up with the goal of identifying highly conserved Arg residues in the HAMP domain of YfiN that could potentially serve as a secondary I-site. Consequently, a typical sequence logo of HAMP domains found in YfiN orthologs was generated as described in 2.6.2 and compared to the sequence logo from HAMP domains found in methyl-accepting chemotaxis protein that do not show c-di-GMP mediated regulation. For both generated logos, the YfiN sequence was portrayed as well to facilitate comparison.

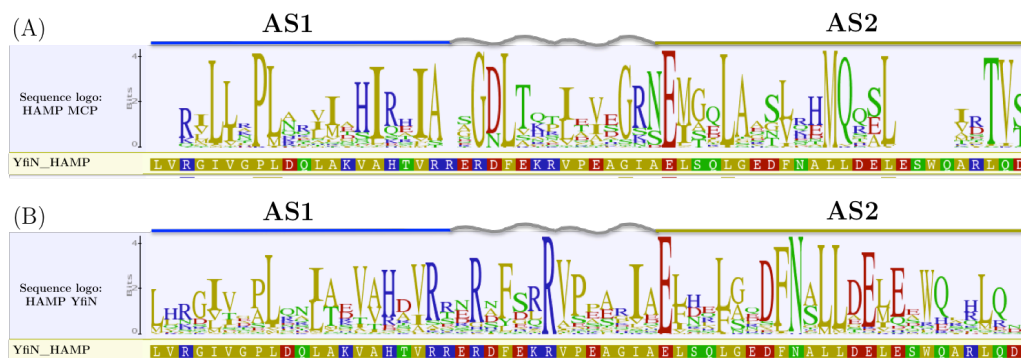


Figure 3-32: Sequence logos of HAMP domains either found in MCPs (A) or in YfiN orthologs (B). The secondary structure elements are represented schematically above the logo. The amino acids are colored according to polarity. The image was generated using Geneious [134].

Comparison of the sequence logos reveals two highly conserved and one completely conserved arginine, which are present in HAMP domains found in YfiN orthologs but not in HAMP domains of MCPs (Figure 3-32). One of them is located at the end of AS1 of the HAMP domain (R200), whereas the following two ones are located in the flexible linker (R203 and R208). As a next step, a homology model of the cytoplasmic part of YfiN was generated as described in 2.6.3 to verify if a 3D arrangement of the HAMP and the GGDEF domain permits these arginines to serve as a potential secondary I-site.

The homology model of YfiN_{HAMP-GGDEF} is showing c-di-GMP mediated cross-linking of the HAMP and the GGDEF domain (Figure 3-33), resembling the product-inhibited state of PleD, where cross-linking is observed between the GGDEF and its associated REC domain [23]. A linker of 6 amino acids, connecting the HAMP and the GGDEF domain could not be modeled and was omitted from the model. Therefore, the effective placement and orientation of the GGDEF domain in respect to the HAMP domain might be slightly different. C-di-GMP binding is mediated by the primary I-site, consisting of the characteristic R(319)xxD(322) motif and the putative secondary I-site, consisting of R208 (Figure 3-33(B)). The second Arg (R203) located in the linker region seems to be too far away to be actively involved in product binding.

As a feasible model of the inactivated state of YfiN had been proposed, the focus was shifted towards generating a model of the active state of YfiN to potentially get deeper insights into the domain movements required to link the two states. Therefore, a model of the active state of DgcZ (Tilman Schirmer, pers. comm.) was used as a template for potentially correct orientation and placement of the GGDEF domains.

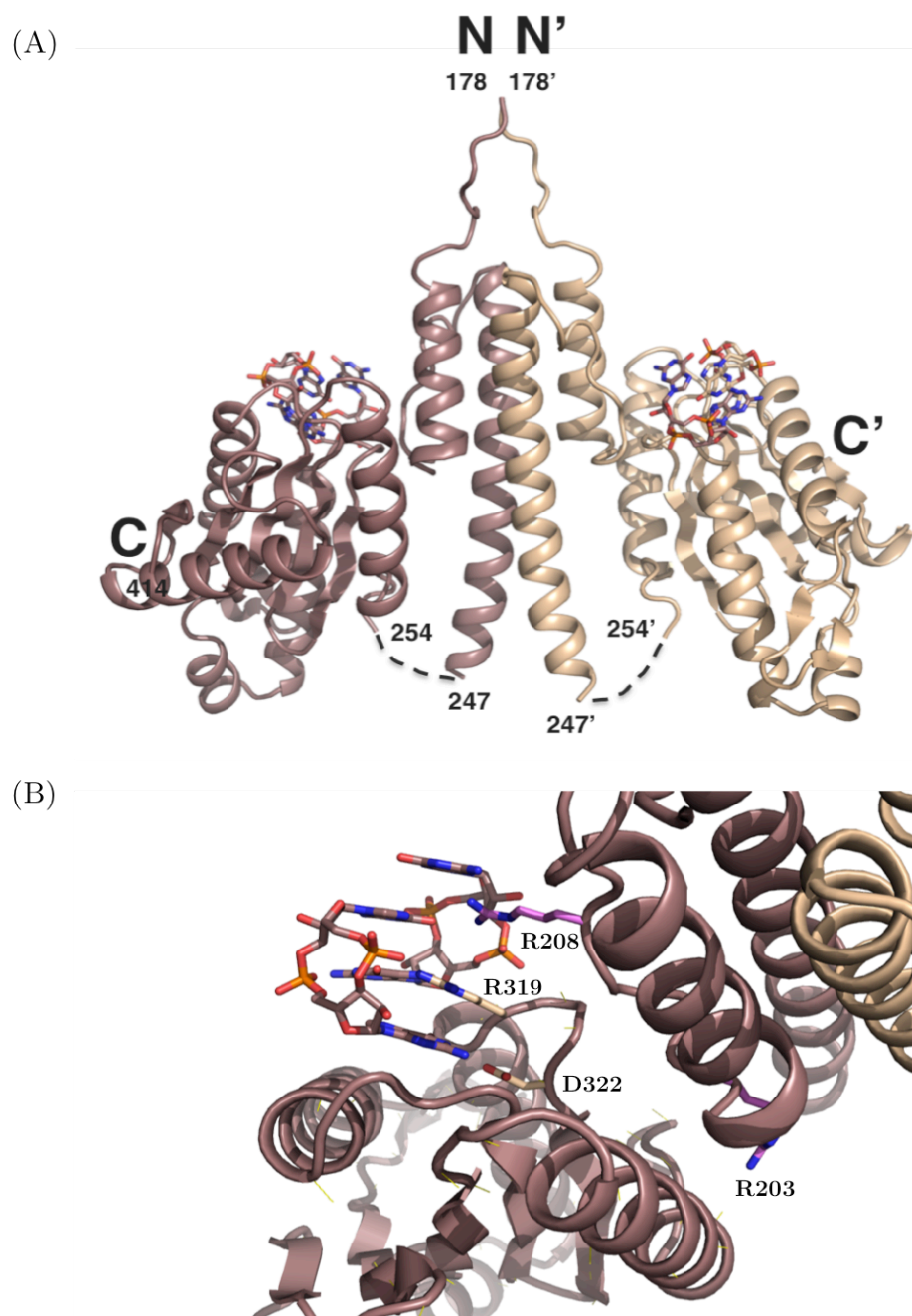


Figure 3-33: The proposed model of the product-inhibited state of YfiN. The homology model of the HAMP domain was generated based on EnvZ by HHPred [116]. The structure of the GGDEF domain of YfiN (PDB entry 4iob) was superposed onto the GGDEF domain of PleD in the product-inhibited state (PDB entry 1w25). Inhibition is mediated by *c*-di-GMP, which is cross-linking the GGDEF and the HAMP domain, in analogy to the product-inhibited state of PleD [24]. The two subunits are distinguished by colors and chain ends are labeled with their corresponding residue numbers. **(A)** Two YfiN monomers come together to form a dimeric assembly, where the N-termini are facing towards the inner membrane. There is a linker of 6 residues, connecting the HAMP to the GGDEF domain, which could not be modeled. **(B)** Close-up view of the *c*-di-GMP dimer, which is cross-linking the HAMP and the GGDEF domain. The I-site residues of YfiN, R319 and D322, are shown as sticks and colored in beige. The two-conserved Arg residues, which are situated in the linker region are shown as sticks and colored in purple.

In a first step, the N-terminal all-helical CZB domain of DgcZ was aligned to the HAMP-helices of YfiN and subsequently the structure of the YfiN GGDEF domain was superposed onto the two DgcZ GGDEF domains.

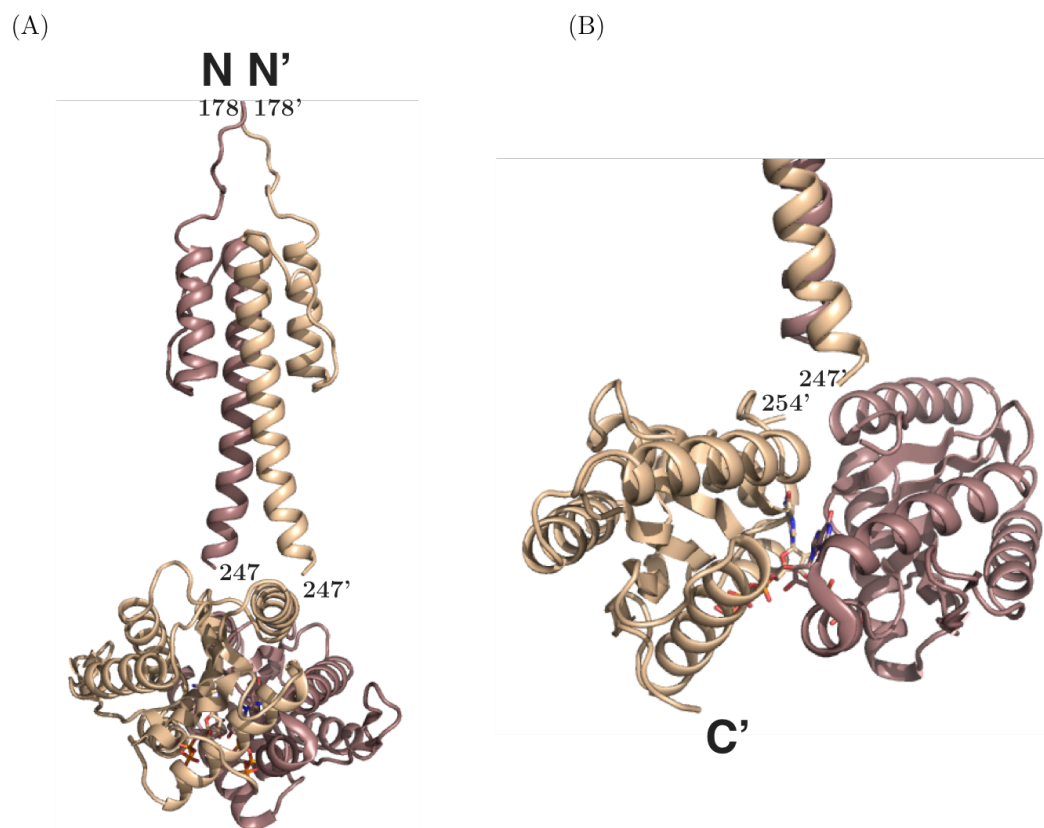


Figure 3-34: The proposed model of hypothetical active state of YfiN. The coloration scheme is the same as for Figure 3-33. (A) Overview of the whole protein in the proposed active conformation. (B) Zoomed-in view into the active site revealing two GTP molecules, which are aligned in an antiparallel fashion for catalysis.

The model of the active conformation is revealing quite a drastic movement of the GGDEF domain in comparison to the inactive state (Figure 3-33, Figure 3-34). As a model of both states was available, a morphing was performed to illustrate the conformational rearrangements required to switch from one state to the other (Figure 3-35).



Figure 3-35: Scan the QR code to get redirected to the morphing video.

3.5 Discussion

The focus of this thesis was the investigation of the YfiR-YfiN interaction and the *in vitro* characterization of the individual proteins. In order to quantify the DGC activity of YfiN, two different approaches were chosen: the first one was based on measuring the activity of membrane-bound YfiN and the second one involved solubilization and purification of YfiN prior to the activity assay. In the following, the results of each subchapter are discussed separately.

3.5.1 Activity assay of membrane-bound YfiN

Measuring the activity of a YfiN sample still embedded in the membranes involved extensive washing of the latter in order to get rid of numerous nucleotides that had been bound before. In a next step, the purified membranes were then incubated either with GTP alone or in combination with YfiR or *c*-di-GMP to check if repression or product inhibition occurred. The appearance of GDP as a side-product indicated that GTPases were

present in the membranes and showed catalytic activity. This might be a proof for the integrity of the membrane fraction, suggesting that the washing protocol was not too harsh and retained a quasi-native environment. The observed GTPase activity was constant throughout all three experiments (Table 3-4) and did not lead to substrate depletion as the enzymes got inactivated over time (Figure 3-6). The presence of GTPases does therefore not pose a limitation to the activity assay as long as substrate concentrations equal or higher than 2 mM are employed.

The second observation was that the membrane sample clearly showed time-dependent production of c-di-GMP in all three experiments, however at different rates (Figure 3-6, Table 3-4). Comparing the progression curve of a sample incubated with GTP to the one from a sample incubated with GTP and YfiR, clearly revealed YfiR-mediated repression of the observed DGC activity with an affinity constant K_r of 25 μ M (Figure 3-7). This result was in fact indicative of two findings at the same time: first of all, clear evidence was obtained that YfiR-mediated repression is indeed also taking place in vitro and second, it proved that the observed DGC activity was to a large extent YfiN-specific. In theory, several DGCs are present in *E. coli* that are localized at the inner-membrane including YfiN (*E. coli*), YcdT [152], YaiC [153], YegE [154], YeaI [154] and YedQ [154] for all of which in vivo activity have been shown. However, considering the vast excess of overexpressed YfiN from *P. aeruginosa* in the membranes, the potential activity of these enzymes can most probably be neglected. Additionally, for YedQ and YegE DGCs only basal expression levels have been shown in growing cells, which are stimulated by the factor sigma(S) upon the transition from the post exponential to the stationary phase [154].

However, a control should be included in the next experimental series, featuring cells that were transformed with either an empty plasmid or ideally a plasmid coding for a catalytically inactive YfiN, having a mutated GGDEF motif [44]. This would in turn give indications if a potential background c-di-GMP producing activity is observed, catalyzed by other membrane-bound DGCs in *E. coli*.

The determined YfiR affinity constant K_r of 25 μM is lying in a physiological range for protein-protein interactions, especially taking into account that YfiR most likely is binding as a dimer to YfiN, which increases the local concentration of YfiR molecules. However, the K_r value has been determined using only one YfiR concentration so far and the experiment should be repeated at additional concentrations especially in the range of the K_r to get a more accurate determination of the value.

The YfiN-YfiR interaction shares some striking similarities with the LapD-LapG system from *Pseudomonas fluorescens* and *Pseudomonas putida* [3], [79], [97], [155]. In *P. fluorescens*, the availability of the essential nutrient inorganic phosphate is an environmental signal that controls biofilm formation through a c-di-GMP signaling pathway. In presence of low amounts of inorganic phosphate, c-di-GMP levels are depleted in the cell and these changes are sensed by the transmembrane c-di-GMP receptor protein LapD. It shows a similar domain organization as YfiN, containing cytoplasmic degenerate GGDEF and EAL domains that lack catalytic activity but the phosphodiesterase domain retained the capacity of c-di-GMP binding. When c-di-GMP levels are low, LapD is kept in an “off” state that allows LapG, a periplasmic protease, to interact with LapA and cleave the N-terminal domain of this adhesion protein, releasing LapA from the cell

surface and thereby leading to biofilm detachment. Under abundant phosphate conditions, c-di-GMP is binding to the EAL domain of LapD in the cytoplasm, whereupon the signal is then relayed to the periplasmic PAS output domain through an inside-out signaling mechanism that utilizes a juxtamembrane HAMP domain. This in turn leads to binding and sequestering of LapG in the periplasm, promoting cell adhesion via maintenance of LapA on the cell surface [97].

In analogy to LapD, YfiN also appears to switch between discrete inactive and active functional states depending on the presence or absence of the inhibitor YfiR. This seems to be a common feature of many well-characterized transmembrane signaling proteins, such as histidine kinases and methyl-accepting chemotaxis proteins, where ligand binding is causing a conformational change that is propagated throughout the whole protein [156]. However, YfiN activity is controlled by YfiR in an outside-inside signaling mechanism in contrast to the inside-outside mechanism observed for LapD [14]. Release of repression leads to a conformational shift in YfiN, which is relayed through the PAS and transmembrane domains to the cytoplasmic HAMP domain, leading to an active conformation. This is then in turn transmitted to the effector GGDEF domains, allowing them to form a catalytically competent homodimer required for c-di-GMP production.

In contrast to the proposed 2:2 stoichiometry for the YfiR-YfiN complex, the crystal structure of the LapD-LapG complex revealed a 2:1 ratio [98]. However, unlike YfiR, LapG has not shown any signs of homodimerization in studies so far, which could therefore explain the observed 2:1 ratio [98], [157]. The binding mode of LapD-LapG shows distinct similarities to the proposed way of interaction between YfiR-YfiN, both involving hydrophobic binding sites. This involves a completely conserved tryptophan residue

W126 (*Legionella pneumophila*, corresponding to W125 in *P. fluorescens*), which serves as the major anchor point on LapD as it inserts into a hydrophobic pocket at the bottom of LapG (Figure 3-36) [98].

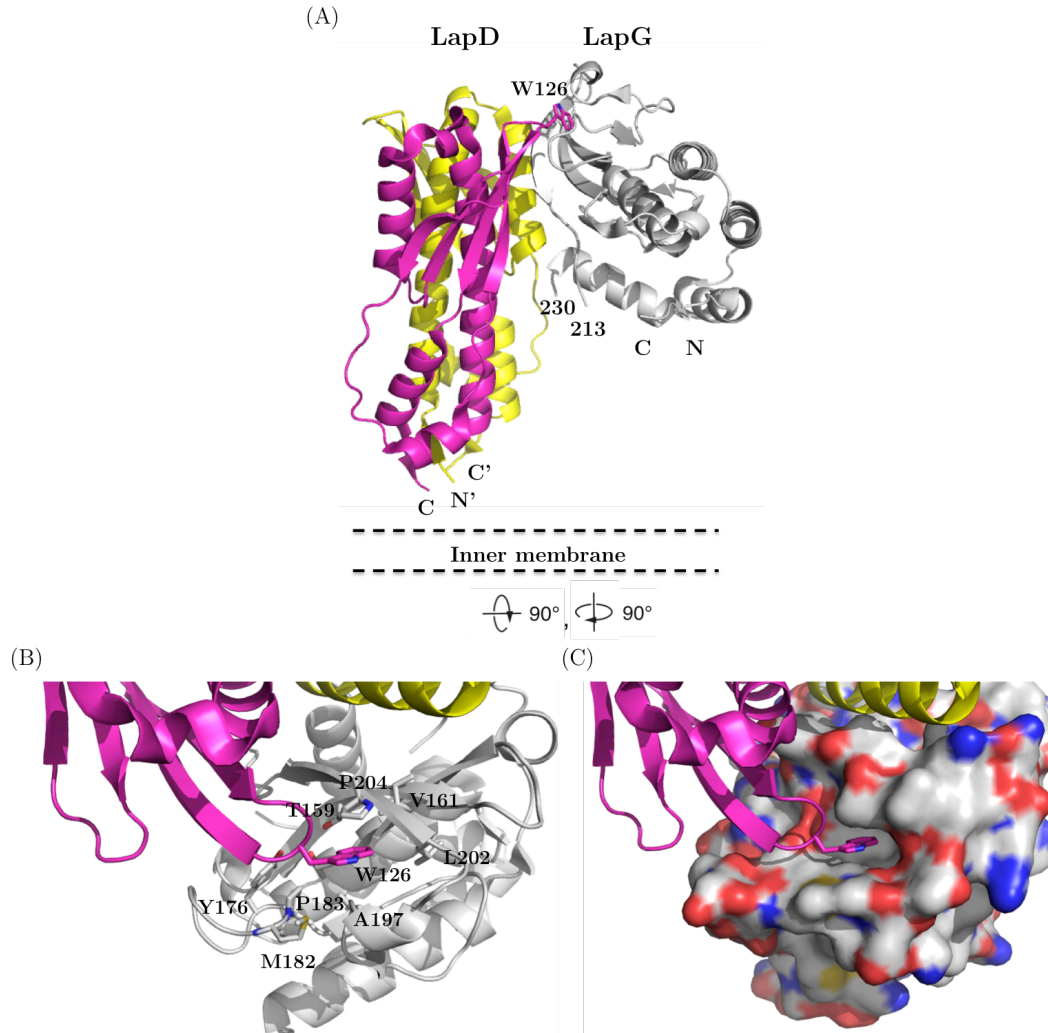


Figure 3-36: Crystal structure of a *P. fluorescens* LapG and *L. pneumophila* LapD_{PAS} complex, based on PDB entry 4U65. (A) Side-view of the complex structure. The LapG-LapD_{PAS} complex is shown as a ribbon representation, with the two protomer chains of LapD_{PAS} colored in pink and yellow and LapG colored in grey. The highly conserved tryptophan residue (W126) is shown as sticks for one of the subunits. Two dashed lines indicate the relative position of the inner membrane to the complex. (B) The hydrophobic binding pocket for W126 from LapD_{PAS} in LapG is shown as a close-up view. Residues forming the pocket are shown as sticks and labeled accordingly. A 90° rotation around x, followed by a 90° rotation around y was applied to LapD from the complex structure to view the interface. (C) Surface representation of (B), illustrating the hydrophobic nature of the binding pocket.

So far, the activity assay using membrane-bound YfiN has revealed time-dependent c-di-GMP production, which could be reduced by the addition of YfiR to the experimental setup.

Furthermore, experiment II, where c-di-GMP was added to the membranes in addition to GTP revealed that YfiN indeed was subjected to non-competitive product inhibition (Figure 3-7). This finding is further supported by the personal observations made during the purification process, where YfiN was co-eluting with c-di-GMP, clearly proving that binding is taking place *in vivo* (Figure 3-12, Figure 3-13). Moreover, this result is additionally confirmed by an independent study from the Jenal group, where a YfiN_{HAMP-GGDEF} construct was clearly shown to bind to c-di-GMP [12]. However, this finding stood in clear contrast to another study, where the structure of the GGDEF domain of YfiN revealed that the presumable secondary I-site present on the GGDEF itself, which is required for domain cross-linking, was degenerate (Figure 3-37) [151]. The presence of an intact primary and secondary I-site is a prerequisite for the c-di-GMP mediated cross-linking of GGDEF domains, which results in the spatial separation of the two active sites and thereby inactivation of the protein. However, the secondary I-site can also be localized on an accessory domain, thereby leading to intramolecular cross-linking of the GGDEF domain to the accessory domain upon c-di-GMP binding [23].

In the aforementioned study, no binding was observed of YfiN_{HAMP-GGDEF} to c-di-GMP in an ITC experiment and it was therefore concluded that YfiN is not subjected to product inhibition due to lack of a secondary I-site on the GGDEF as well as on the HAMP domain.

Having a closer look at the AUC data, which was performed on the YfiN_{HAMP-GGDEF} construct revealed, however, that as it was already observed

in the framework of this thesis, the construct used in said study got proteolytically cleaved, corresponding to the size of the isolated GGDEF. This was also the reason why the HAMP domain was not present in the crystal structure. The conclusion drawn in the study was that YfiN_{HAMP-GGDEF} was eluting as a monomer from SEC but no references was made to the occurring proteolysis. It remains thus unclear, if the authors were aware of this issue or not.

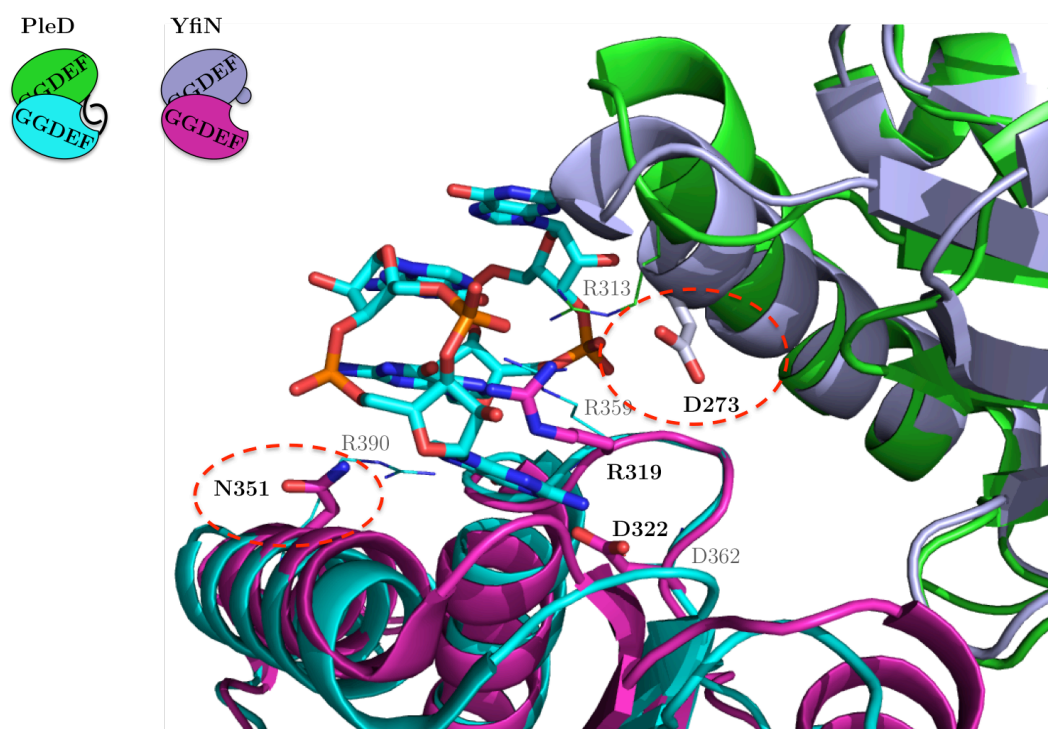


Figure 3-37: Comparison of the Ip and Is sites of YfiN (PDB entry 4iob) and PleD (PDB entry: 2v0n). The two subunits of a hypothetical product-inhibited YfiN, which were superposed onto the PleD dimer, are shown in magenta and purple, while PleD is shown in green and blue. The coordinates for c-di-GMP were taken from PleD and shown in sticks. The residues involved in binding are shown in stick as well and are labeled in bold and black for YfiN and in grey for PleD. YfiN lacks two of the three arginines required for c-di-GMP binding, which are replaced by D273 and N351.

Assuming that the secondary I-site of YfiN is in fact localized on the HAMP domain and taking into consideration that the construct got proteolytically cleaved, this could thus pose a potential explanation for the fact that no c-di-GMP binding was observed. This control mechanism is in analogy to

PleD, where besides a *c*-di-GMP mediated GGDEF-GGDEF domain cross-link, a second mode of domain immobilization is found, involving *c*-di-GMP mediated cross-linking of the GGDEF domain to the accessory REC domain (Figure 3-38) [23] [24].

PleD

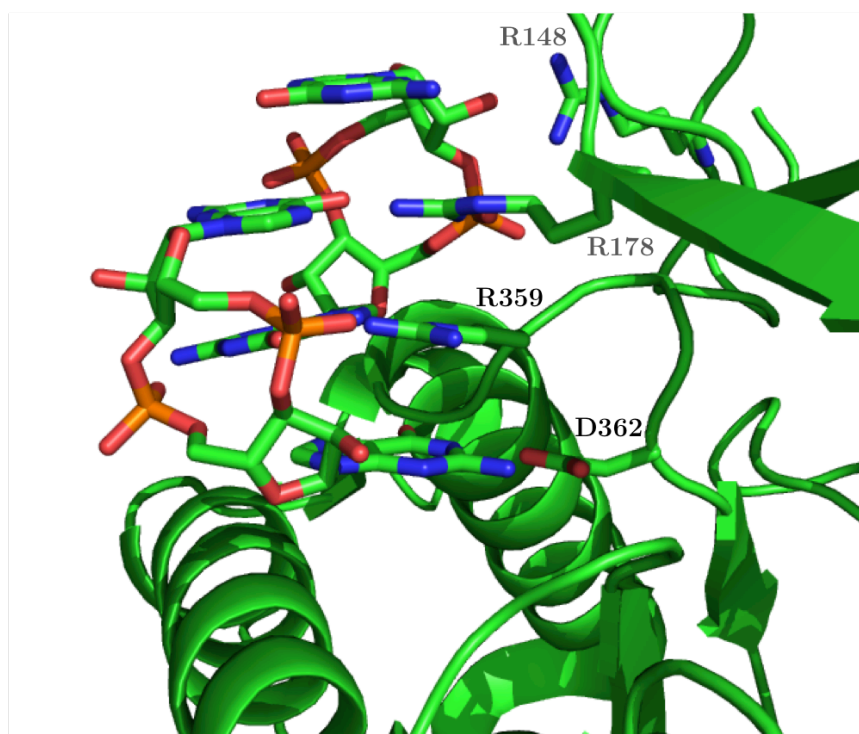


Figure 3-38: The non-competitive product inhibited state of PleD is involving a *c*-di-GMP mediated cross-link between the REC and the GGDEF domains. The *c*-di-GMP dimer is shown as sticks. The residues involved in binding are shown as sticks and labeled in bold and black for the Ip-site and in bold and grey for the Is-site.

In order to identify highly conserved Arg residues in the HAMP domain of YfiN, which could potentially serve as a secondary I-site, a sequence logo of HAMP domains found in YfiN orthologs was compared to the sequence logo from HAMP domains found in methyl-accepting chemotaxis protein that do not show *c*-di-GMP mediated regulation.

One highly (R203) and one completely conserved Arg (R208) were identified in the HAMP domains of YfiN orthologs, which were not present in the HAMP domains of MCPs at the equivalent position (Figure 3-32). A homology model of the product-inhibited state of YfiN, which was generated based on PleD, revealed that R208 is indeed situated close enough to coordinate the N7 position and the oxygen of the phosphate group of the c-di-GMP dimer, comparable to R178 in PleD (Figure 3-33 (B), Figure 3-38). There seems to be no direct involvement of Arg203 in c-di-GMP binding but it might be conserved for structural reasons such as stabilizing the helix bundle. From the model, it is not evident if there are more residues involved in the binding and coordination of c-di-GMP besides the primary I-site, consisting of the characteristic RxxD motif, and the putative secondary I-site R208. Taken all these findings together, this concludes that YfiN is clearly subjected to non-competitive product inhibition, presumably by c-di-GMP mediated cross-linking and that product binding is accomplished by the primary I-site and a secondary I-site, which is located on the HAMP domain. However, the role of Arg208 as the potential secondary I-site remains to be verified by mutational studies *in vitro* or *in vivo*.

In summary, the activity assay using membrane-bound YfiN yielded astonishingly clean and easy to interpret chromatograms due to the improved washing protocol. This might therefore pose a valid approach for membrane proteins in general to probe for activity before solubilization and investigate the qualitative effect of inhibitors or interaction partners on the activity. For quantitative statements, a clear estimation of the protein concentration is required, which could be obtained by Western Blot analysis.

In the future, the protocol should be slightly adapted, using shorter experimental durations, as after 2 hours nucleotide loss started occurring. This could be caused by nucleotide binding to the membranes, where they are withdrawn from the reaction mixture upon heat-deactivation and precipitation (Figure 3-6).

3.5.2 Purification of detergent-solubilized YfiN

YfiN was expressed in abundant amounts in the membranes at the established condition (Figure 3-3(A), lane 7). Placement of the his-tag at the N-terminus turned out to be very important for protein stability, as a C-terminally tagged construct appeared completely degraded. In order to extract the protein from the membranes and keep it in a solubilized state, three different detergents were tested in regards to their extraction capacities. DM thereby showed the highest extraction efficiencies in comparison to DDM, which still extracted some amounts of protein (Figure 3-11) and β -OG, which did not lead to any considerable protein extraction (Figure 3-8) and was therefore omitted from further tests. SEC revealed completely aggregated protein for DDM (Figure 3-11) and a prominent aggregation peak for DM (Figure 3-10) but for the latter two peaks were obtained corresponding to smaller molecular weights, however the peak intensities amounted to only a few mAU units. Due to time limitations, it was decided to proceed with DM as detergent, as not so much protein was required for the subsequent activity assays. However, what usually should be done at this stage is to go for extensive screening of detergents and

additives, which stabilize the target protein to maximize the yield of working material for downstream applications.

For all purifications, c-di-GMP binding to YfiN was observed in a 1.5: 1 molar ratio on average. Therefore, the subsequent SEC was usually run at very low flow speeds to get rid of the majority of bound nucleotide (Figure 3-12).

In order to get information about the oligomeric state of YfiN in solution, DLS analysis was performed for both SEC peaks individually. As the observed standard deviations were quite high, it was difficult to get a reliable estimation of the oligomeric state of YfiN in solution (Table 3-5). Both peaks correspond most probably to a monomeric YfiN but in different conformations. It is however not straightforward to deduce the nature of these two conformations. Due to the very low absorption values obtained, usually lying in the range of $10E-02$ - $10E-03$, both peak fractions had to be pooled together and concentrated in order to get reliable measurements. It can therefore not be excluded that, for instance, the two peaks show different 260/280 ratios. One of the hypotheses could be that as residual c-di-GMP binding is observed even after SEC at a protein-nucleotide molar ratio of 1: 0.3 on average, one of the conformations could correspond to a c-di-GMP bound and the other one to a c-di-GMP-free form. The product-inhibited state forms a much more globular shape as the product-free form, which is adopting a more elongated shape (Figure 3-33, Figure 3-34). The former could therefore correspond to peak 3, whereas the latter is eluting earlier from the SEC and therefore would be contained in peak 2. However, a protein: nucleotide ratio of 0.3 is equivalent to every 6. protein being fully loaded with c-di-GMP and hence the same difference in peak area should be obtained for the two peaks, which is clearly not the case for YfiN (Figure 3-

10). The identities of the two conformations remain therefore speculative. Nonetheless, it is evident that if buffer conditions are identified, which lead to higher protein yields, more accurate determination of the molecular weight is possible by MALLS, as strong concentration of the protein is no longer required.

3.5.3 Activity assay of detergent-solubilized YfiN

A protocol was successfully set-up for the characterization of DGC activity in 96-well format using the malachite-green assay. The k_{cat} values obtained for the reference protein DgcZ (3.2.1.7), were in perfect agreement with the previously published ones [111]. However, when the presumable activity of detergent-solubilized YfiN was measured in a next step it turned out to be caused by a GTPase contamination present in the YfiN sample (Figure 3-15). This finding was revealed by a control where YfiN and GTP were incubated in presence and absence of PPase. In the absence of PPase, no signal should be obtained for a DGC as only inorganic phosphate is detected by the assay. However, the experiment revealed no significant differences in the initial velocities of YfiN in presence and absence of PPase (Table 3-6), conforming that the observed activity was indeed due to a GTPase contamination.

In addition, YfiR also showed activity when incubated with GTP, which was in the same range as YfiN (Figure 3-18). This activity is most probably caused by a GTPase contamination as well (Figure 3-18) as no physiologically relevant GTPase activity is expected of a periplasmic protein due to the lack of nucleotides in the periplasm.

Taken all these findings together, there was no detectable YfiN activity taking place during the experiment. The most likely reason for this is due to YfiN inactivation as a result of the necessary strong concentration process, which at the same time led to presumably very high DM concentrations of 30 mM – 60 mM in the sample depending on the experiment. As stated before, more time should therefore be invested on the identification of buffers, detergents and additives, which increase the amount of extracted and well-behaving protein.

Furthermore, the fact that the malachite-green assay does not provide a direct readout for DGC activity but instead needs to be coupled to a second enzyme posed a certain limitation in this study. An FPLC-based method would have directly visualized that the investigated activity was in fact generating GDP instead of c-di-GMP, giving a tighter control on the nucleotide turnover, which is taking place during the experiment. However, the malachite-green assay constitutes a less laborious and easy-to-use alternative if the investigated enzymes show an activity of certain significance such as DgcZ, which gave rise to a signal that was at least two orders of magnitudes higher than the contamination.

3.5.4 Preliminary results of the individual domains of YfiN

3.5.4.1 The periplasmic PAS domain

Throughout all experiments, two of the three investigated PAS constructs showed no soluble expression at all and only for YfiN_{PAS44-149} very low amounts of soluble protein were obtained. However, the following SEC revealed only aggregated protein, which was co-eluting with numerous

contaminants. When the same construct was fused to a MBP, GST or SUMO-tag, abundant soluble expression was obtained for the MBP-PAS₄₄₋₁₄₉ fusion, whereas only insoluble expression was observed for the other two fusion constructs. However, also the MBP-fusion construct eluted exclusively as aggregates from the SEC, again indicating folding problems of the PAS domain. One explanation for the observed aggregation-prone and unstable behavior of the PAS domain could be that it needs the presence of the downstream transmembrane and HAMP domains for dimerization and thus stabilization of the protein. The dimerization interface of the YfiN_{PAS} domain is overall well conserved, with the potential binding site motif AAVVF showing almost complete conservation (3.1), whereas the rest of the domain is rather variable. The presence of this hydrophobic patch is distinct for YfiN, as none of the other PAS domains investigated so far have been shown to operate via a hydrophobic binding site, which is situated on the surface of the protein. Even the similarly operating LapD-LapG interaction is mediated via a hydrophobic pocket located on LapG, which is shielded from the aqueous environment [98]. Taking into account that YfiN presumably switches between a discrete inactive and active state, depending on the presence or absence of bound YfiR, it is likely that the hydrophobic patch is either engaging in YfiR binding or is shielded from the aqueous environment by the two PAS domains coming closer together.

Future experiments should therefore focus on stabilizing either the YfiR-bound state or generating a leucine-zipper PAS construct, which would lead to the stabilization of a PAS dimer. Leucine zippers have been shown to form parallel coiled-coils and, hence, serves as a strong dimerization modules [158]. A similar approach was used for investigation of the DGC WspR [26].

3.5.4.2 The cytoplasmic HAMP and GGDEF domains

The four YfiN_{HAMP-GGDEF} constructs from different species showed overall the same kind of behavior. Expression tests revealed that the cytoplasmic constructs got expressed in soluble abundant amounts (Figure 3-25). *Ps.aer.*, *Ps.flu.* and *Ye.ent.* YfiN_{HAMP-GGDEF} were purified to homogeneity by a two-step purification process including IMAC and SEC. Slight adjustments had to be made to the IMAC protocol, applying a less steep imidazole gradient, in order to reduce aggregation of the proteins. SDS-PAGE analysis during the purification process showed that all constructs were subjected to proteolysis (Figure 3-27). Tryptic digest in combination with mass spectrometry analysis revealed that *Ye.ent.* YfiN got cleaved at two positions within the protein, once within the HAMP domain and once between the HAMP and the GGDEF domain (Figure 3-29, Table 3-14). This could indicate that these parts of the sequence are not well structured and thus prone for proteolytic degradation. The linker between the two HAMP helices could therefore be a likely target for cleavage to occur. Nonetheless, the cleaved products could usually be separated from the full-length cytoplasmic YfiN constructs and initial crystallization was pursued with *Ps.aer.* and *Ye.ent.* YfiN_{HAMP-GGDEF}. However, in many drops only precipitation and severe oiling out were observed, most probably due to the heterogeneity of the proteins as a result of nucleotide binding and proteolysis. It should be investigated if the protein is stable after SEC or if time-dependent proteolysis is occurring, which would in turn hinder crystallization. As for the full-length YfiN, the cytoplasmic constructs could also be run on a S200 26/60 column at a low flow speed, in order to separate the protein from the bound nucleotide (Figure 3-12). Further

crystallization trials could be performed in the presence of a non-hydrolyzable GTP analogue, such as GTP α S, in order to trap the protein in one state.

As initial crystallization attempts failed, it was decided to do a functional characterization of the Ps.aer. YfiN_{HAMP-GGDEF} construct. The phosphate sensor assay was used to determine the k_{cat} of YfiN_{HAMP-GGDEF}, which amounted to $3.4E-04\text{ s}^{-1}$. This might seem very low, however the value is in the same order of magnitude as the k_{cat} of non-activated PleD, which amounted to $9.0E-04\text{ s}^{-1}$ [24]. DGCs activity requires the formation of a dimer, which depends on the diffusion as well as the proper orientation of the two GGDEF domains, where each domain contributes one GTP substrate molecule to the reaction. However, it is likely that nature optimized the macroscopic electrostatic interactions involved in dimer formation in order to facilitate the proper orientation of the two domains with respect to each other. Therefore, the process of dimer formation predominantly depends on diffusion, which can in turn be increased due to law of mass action by raising the concentration of the proteins. A monomeric GGDEF domain is thus still expected to show some basal activity, which is indeed what is observed for the monomeric Ps.aer. YfiN_{HAMP-GGDEF}.

4 Conclusion and Outlook

To better understand the formation of *P. aeruginosa* small-colony variants (SCVs) during long-term lung infections, knowledge about the interaction mechanisms of the involved YfiBNR proteins is essential. In vivo and in vitro studies partly characterized these interactions, however, according to my knowledge, no structural information was available to date. Furthermore, in vitro functional studies of the effector of the system YfiN and its repressor YfiR were still missing.

Within the framework of this thesis, YfiN and YfiR have therefore been analyzed by structural, functional and biochemical means. Together with the results from the previous studies, this facilitates a better insight into the mechanism of how YfiN activity is being regulated.

The three-dimensional structure of YfiR was solved using X-ray Crystallography, which revealed that it is adopting a novel fold. Moreover, the fact that a dimeric YfiR was found in the crystal structure, raised questions about the physiological relevance of this dimer. Probing the oligomeric state of YfiR by SEC-MALLS confirmed that it indeed dimerizes in solution with a dissociation constant (Kd) in the nanomolar range. Furthermore, a mutation in the putative dimerization interface was shown to significantly weaken the dimerization by two orders of magnitude compared to the wildtype. This indicates that the same kind of dimer is adopted in solution as in the crystal structure.

In order to characterize the YfiR-YfiN interaction in vitro, a functional assay was established using membrane-bound YfiN. Time-dependent c-di-GMP production was observed, which confirmed that YfiN showed catalytic

activity *in vitro*. Upon addition of YfiR, significant repression of YfiN activity was demonstrated, with a Kd in the lower micromolar range. Considering that the Kd of YfiR dimerization is two orders of magnitude lower, it is therefore proposed that YfiR is binding to YfiN in a dimeric state.

Furthermore, membrane-bound YfiN appeared to be subjected to non-competitive feedback inhibition, which was in agreement with the observations that all YfiN constructs investigated during this thesis co-eluted with *c*-di-GMP. This clearly indicates the presence of an intact primary I-site on the GGDEF domain and an intact secondary I-site localized either on the HAMP or the GGDEF domain. Further insights were gained by the crystal structure of YfiN_{GGDEF}, solved by Giardina et al. [151], which revealed that the secondary I-site on the GGDEF domain is degenerate and thus incapable of *c*-di-GMP binding. This in turn prompted the generation of a homology model of the product-inhibited state of YfiN, involving the primary I-site on the GGDEF domain and one highly conserved arginine R208 on the HAMP domain, which could serve as the potential secondary I-site. To identify the domain rearrangements required to switch between inactive and active state, a second homology model was generated of the presumable activated state of YfiN. The resulting overlay of both states revealed a rather massive rearrangement of the GGDEF domain, of which the physiological occurrence could, however, still be envisaged.

Preliminary attempts to investigate the activity of detergent-solubilized YfiN not been successful so far due to low protein yields.

In conclusion, Figure 4-1 shows an updated version of the proposed mechanism by which YfiN activity is regulated in the cell, incorporating all relevant findings of this study.

In the future, the established activity assay using membrane-bound YfiN should be applied for further investigations. Among others, the performed experiments should be repeated to verify the reproducibility of the results. In addition, a more accurate determination of the dissociation constant for the YfiR-YfiN binding should be obtained by repeating the same experiment at different YfiR concentrations. The assay could then be used to further study the interactions between the currently available YfiR and YfiN gain-of-function mutants and between YfiN and the YfiR dimerization interface mutant.

For the solubilization of YfiN, more elaborate efforts are required to screen for buffers, additives and detergents, which increase the amount of solubilized protein for downstream applications. Once a purification protocol has been established, the application of nanodiscs should be pursued for further functional assays *in vitro*.

Ultimately, the role of R208 as the putative secondary I-site remains to be investigated by *in vitro* or *in vivo* experiments.

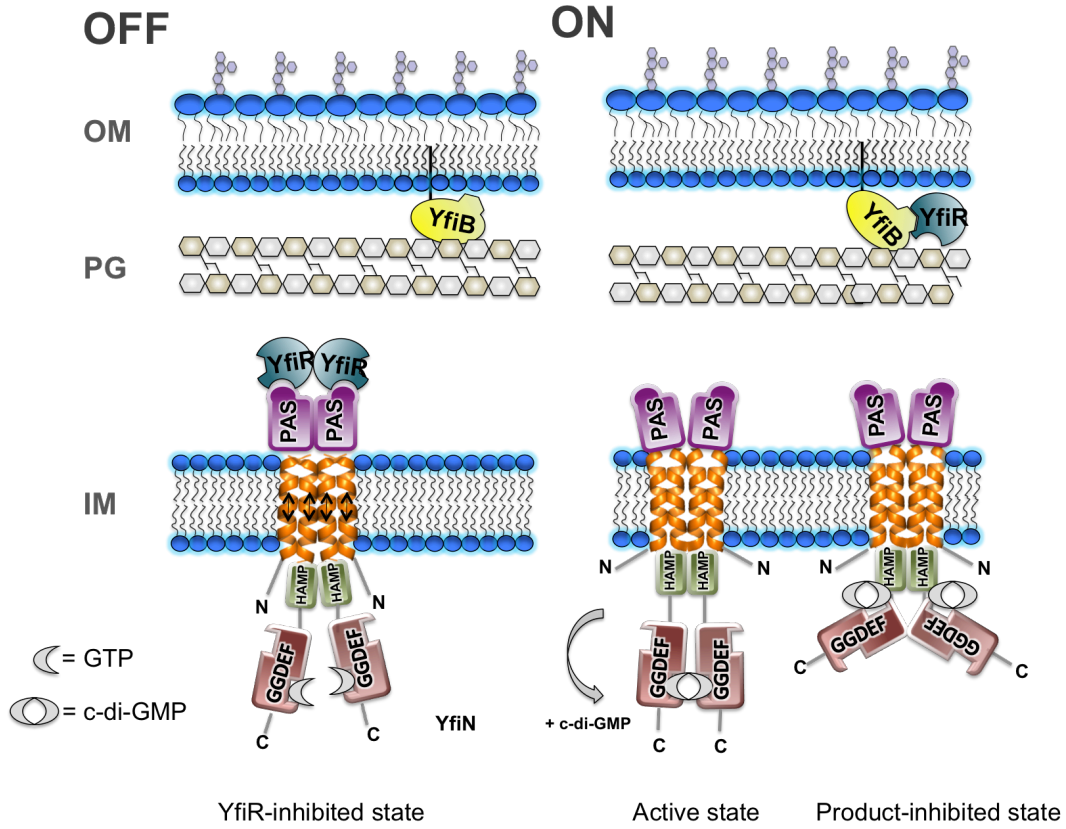


Figure 4-1: Overview of the YfiB/NR system in the off state (left) and in the on state (right). All three proteins as well as GTP and c-di-GMP are depicted in a schematic. Three different conformational states are proposed, a YfiR-inhibited, an active and a transient product inhibited state to limit excess substrate consumption.

5 Contrasting Mechanism of Inhibition of SHV5 and KPC2 β -Lactamases

5.1 Research article II (Kauer and Zähringer *et al.*, in preparation)

Contrasting Mechanisms of Inhibition of SHV-5 and KPC-2 Class A β -Lactamases by Sulbactam, (3R)-3-Acetamido-2-hydroxy-3,4-dihydro-1,2-benzoxaborinine-8-carboxylic Acid and Avibactam

Stefanie Kauer^{†1}, Franziska Zähringer^{†1}, Patrick Caspers^{2†}, Clothilde Dantier², Eric Desarbre², Stefan Reinelt², Tilman Schirmer¹, Malcolm G. P. Page^{2}*

1. Structural Biology and Biophysics, Biozentrum, University of Basel, Klingelbergstrasse 50/70 CH - 4056 Basel, Switzerland. 2 Basilea Pharmaceutica International Ltd, Grenzacherstrasse 487, CH-4058 Basel, Switzerland.

‡These authors contributed equally.

*Corresponding author.

Manuscript in preparation

Statement of my own contributions

I contributed to the manuscript by co-crystallizing SHV-5 in presence of (3R)-3-Acetamido-2-hydroxy-3,4-dihydro-1,2-benzoxaborinine-8-carboxylic Acid and solving the crystal structure of the complex. I wrote the methods sections “Crystal Growth and Structure Determination of Purified SHV-5 “ and “Data collection and Processing“. I furthermore wrote the results section “Crystal structure of the covalent AHB-SHV-5 complex“. Figure 5 and figure 7 have been prepared by me, as well as the original drafts to figures 6 and 8. I also prepared tables 1 and 3.

Contrasting Mechanisms of Inhibition of SHV-5
and KPC-2 Class A β -Lactamases by Sulbactam,
(3R)-3-Acetamido-2-hydroxy-3,4-dihydro-1,2-
benzoxaborinine-8-carboxylic Acid and Avibactam

Stefanie Kauer^{‡1}, Franziska Zähringer^{‡1}, Patrick Caspers^{2†}, Clothilde Dantier², Eric Desarbre², Stefan Reinelt², Tilman Schirmer¹, Malcolm G. P. Page^{2}*

1. Structural Biology and Biophysics, Biozentrum, University of Basel, Klingelbergstrasse 50/70 CH - 4056 Basel, Switzerland. 2 Basilea Pharmaceutica International Ltd, Grenzacherstrasse 487, CH-4058 Basel, Switzerland.

[‡]These authors contributed equally.

*Corresponding author.

ABSTRACT: This study examines the reaction of the Ambler class A extended-spectrum β -lactamase, SHV-5, and the Ambler class A carbapenemase, KPC-2, with a mechanism-based inhibitor (sulbactam), a potential transition-state analogue (3-acetamido-2-hydroxy-3,4-dihydro-1,2-

benzoxaborinine-8-carboxylic acid) and the non- β -lactam, avibactam. Branched pathways that result in stabilized intermediates appear to make a significant contribution to the inhibition of SHV-5 and the crystal structure of one such inhibitor complex formed with (3-acetamido-2-hydroxy-3,4-dihydro-1,2-benzoxaborinine-8-carboxylic acid) is described. The inhibitor moiety forms a closed structure with boron in tetrahedral configuration, potentially analogous to the transition state for acyl-enzyme formation, rather than a previously reported open structure previously reported for the complex formed with TEM-1 (Ness, S., Martin, R., Kindler, A. M., Paetzel, M., Gold, M., Jensen, S.E., Jones, J.B., Strynadka, N.C. (2000). *Biochemistry* 39, 5312-5321). In contrast, although there is evidence for a branched pathway in the reaction between sulbactam and KPC-2, this does not result in stable intermediates. Avibactam reacts more slowly than either of the other types of inhibitor and follows a two-step reaction to form a stable carbamoyl ester with the active site serine. The crystal structure shows that the inhibitor moiety adopts a subtly different conformation in the active site that could account for its degradation pathway.

β -Lactam antibiotics have been a cornerstone of the treatment of infections caused by Gram-negative bacteria for the last several decades.¹ The β -lactams inhibit penicillin-binding proteins, preventing an essential transpeptidation reaction and thereby triggering a series of, as yet

incompletely understood, reactions that lead to cell lysis and death.^{2,3} The Gram-negative bacteria have acquired resistance to β -lactam antibiotics through mutations that increase the filtration effect of their outer membrane, or enhance the rate at which the antibiotic is extruded from the cell and, above all, through the expression of β -lactamases that hydrolyze the β -lactam ring, rendering the antibiotic inactive.^{4,5} Four classes of β -lactamases, A, B, C and D, are commonly recognized, based on their amino acid sequences.⁶⁻⁸ The class A β -lactamases are the most prevalent in human pathogens and have evolved into a number of lineages with modified kinetic properties that extend their substrate specificity profile to include advanced generation cephalosporins (extended-spectrum β -lactamases, or ESBLs) or carbapenemases that can hydrolyze carbapenem antibiotics that normally inhibit these enzymes.^{9,10} Members of the SHV lineage of class A β -lactamases are commonly found in *Klebsiella pneumoniae* and *Escherichia coli*. SHV-5 acquires an ESBL phenotype compared to SHV-1 through substitutions at positions Gly-238, to Ser, and Glu-240, to Lys, that extend its substrate specificity to include aminothiazolyloxyimino-cephalosporins such as cefotaxime, ceftriaxone and ceftazidime.¹¹ The changes in kinetic properties, which are associated with branched pathways driven by enzyme conformational changes dependent on the reaction with substrates, has the unexpected consequence of rendering the mutant enzymes more susceptible to mechanism-based inhibitors such as the clinically used β -lactamase inhibitors clavulanic acid, sulbactam and tazobactam.^{12,13} Nevertheless, because of their abundance, diversity and potential high-level expression, the ESBL enzymes of this type remain a concern and there is considerable interest in the development of new, more potent inhibitors to combine with

β -lactamase-labile antibiotics. More challenging, in 2013, the Centers for Disease Control and Prevention (CDC) in the United States announced that carbapenem-resistant Enterobacteriaceae (CRE) are one of the highest priority target pathogens for the identification and development of novel antibacterials.¹⁴ Klebsiellae expressing the class A carbapenemase KPC-2¹⁵ are among the most prevalent CRE in many parts of the world including the United States, Europe and Israel.¹⁶ Although designated a carbapenemase, KPC-2 is able hydrolyze β -lactams from all classes¹⁷ with only a few agents, such as BAL30072, being hydrolyzed slowly enough to retain antimicrobial activity in the presence of KPC-2.¹⁸ Further, KPC-2 is not susceptible to the commercially available β -lactamase inhibitors and only the newly approved avibactam is able to afford significant protection to labile β -lactam antibiotics.^{19,20,24} Boronic acid-based inhibitors have been studied by several laboratories as potential inhibitors of β -lactamases including KPC-2.^{21,22} ²³ The inhibitor (3R)-3-acetamido-2-hydroxy-3,4-dihydro-1,2-benzoxaborinine-8-carboxylic acid (AHB) was described as the product of structure-based design and is one of the most potent boronic acid-based inhibitors of class A β -lactamases.²²

Sulbactam (Fig 1A) is a mechanism-based inhibitor that reacts with the β -lactamases in the same way as simple substrates, such as penicillins, do forming an acyl-enzyme intermediate (cf Scheme 1A) but then undergoes a complicated series of chemical rearrangements that lead to enamine intermediates (Scheme 2) in which the acyl ester is deactivated.^{12,13} The boronic acid-based AHB (Fig. 1B) was postulated to be a transition-state analogue forming a tetrahedral boronyl ester that resembles the hypothetical tetrahedral intermediate for the hydrolysis of the acyl-enzyme

formed with β -lactams (Scheme 1B).²² Avibactam, (1R,2S,5R)-2-carbamoyl-5-methyl-7-oxo-1,6-diazabicyclo[3.2.1]octan-6-yl hydrogen sulfate (Fig. 1C), represents a new class of β -lactamase inhibitors that are able to modify the active site serine to form a carbamoyl ester (Scheme 1B), which is significantly less susceptible to hydrolysis than acyl esters.²⁰ It has been established that regeneration of the free enzyme from the carbamoyl intermediate does not proceed through direct hydrolysis, but may involve either neighbouring group attack of N6 on the carbamoyl ester to release the starting compound, or fragmentation of the inhibitor moiety (Scheme 1C).²⁴

While studying the kinetics of inhibition of ESBLs by model β -lactamase inhibitors including sulbactam, AHB and avibactam, we observed some interesting features in the reaction of SHV-5 and KPC-2 and we describe here the inhibition kinetics and crystal structure of complexes formed between these two inhibitors and the enzymes.

EXPERIMENTAL PROCEDURES

Bacterial Strains, Cloning and Expression. Plasmids carrying the β -lactamase genes were extracted using standard methods²⁵ from the host organisms: *Klebsiella pneumoniae* (pYW1) for KPC-2, *Escherichia coli* HB101 for SHV-1 and *Escherichia coli* MPA-5 for SHV-5, all isolates were from the Basilea strain collection. The PCR products were purified, digested with appropriate restriction enzymes and cloned into the correspondingly digested *E. coli* inducible expression vector pPC56. The

ligation reactions were transformed into *E. coli* M15 containing the compatible plasmid pREP4, coding for the lacI repressor and kanamycin resistance.²⁶ (Stueber et al). Transformants were selected on LB agar plates containing chloramphenicol (15mg/L) and kanamycin (25 mg/L). Individual clones were tested for the presence of an insert of the expected size. The DNA of the inserts were completely sequenced using vector-based flanking primers and a set of internal sequencing primers. The strains expressing the appropriate enzyme were grown overnight in Luria broth (LB) containing 25 μ g/ml kanamycin and 15 μ g/ml chloramphenicol, 20% by volume of glycerol was added and the suspension stored at -80°C for further use. For growth, 5 ml of LB + 25 μ g/ml kanamycin + 15 μ g/ml chloramphenicol was inoculated with bacteria from the -80°C stock and cultured overnight at at 37°C with gentle shaking. 2.5 mL of this overnight culture was introduced into a conical flask containing 100 ml LB + 25 μ g/ml kanamycin + 15 μ g/ml chloramphenicol and culture continued overnight at 37°C. This culture was used to inoculate 6 x 1L of Terrific broth + 25 μ g/ml kanamycin + 15 μ g/ml chloramphenicol (10 ml per flask) and the cultures were grown to OD 600nm between 0.7 and 0.8 when isopropylthio- β -D-galactoside was added at a final concentration of 0.1 mM to induce β -lactamase expression. The cultures were incubated for 2h at 30°C with orbital shaking at 100 r.p.m. The bacteria were collected by centrifugation and the cell pellets were frozen at -80°C for further use.

β -Lactamase purification. The cell pellets were thawed on ice and the cells re-suspended in 20mM Tris-HCl pH 7.5, 10mM MgCl₂ containing 10 μ g/mL DNase I. The cell suspension was passed twice through a pre-cooled (4°C) French Press cell (SLM AMINCO Instruments) at 1250 PSI and then

centrifuged at 35000g, 1 hour, 4°C. Activity in the supernatant and the pellet was determined with nitrocefin as reporter substrate and protein quantification using Lowry method. The supernatant was dialysed against two changes of 2L of 20mM Tris-HCl pH7.5 at 4°C. Ammonium acetate (5M, pH 5.0) was added drop-wise, while stirring, on ice to give 0.1M final concentration and the cloudy suspension was centrifuged at 46000g for 10min. at 4°C. The supernatant was loaded onto a cm x cm SP sepharose FF equilibrated with 0.1M ammonium acetate pH 5.0 at 1.5mL/min. The column was washed overnight with 0.1M ammonium acetate pH 5.0. The protein was elution using a linear gradient of 1M ammonium acetate pH8.5. Fractions were collected and tested for β -lactamase activity and protein concentration. For purification of SHV-5 β -lactamase, the eluate from the SP sepharose FF was dialysed against 20 mM Tris-HCl pH 7.4, concentrated using 10kDa Amicon concentrators at 3000g at 4°C in a swing- rotor centrifuge and applied to a Superdex 200 Hi Load Prep grade column equilibrated with 20 mM Tris-HCl pH 7.4. The protein was eluted in an isocratic run in the same buffer. For the purification of KPC-2 β -lactamase, the active fractions eluted from the SP sepharose FF column were pooled and dialysed against two changes of 2L of 20mM Tris-HCl pH7.5 at 4°C. The dialysate was concentrated before loading on to a Superdex 200 Hi Load Prep grade column (GE Code No.: 17-1071-01) equilibrated with 50mM sodium phosphate buffer pH7.5 at 1.5mL/min and the protein was eluted in an isocratic run in the same buffer.

Reagents. Sulbactam, (3R)-3-acetamido-2-hydroxy-3,4-dihydro-1,2-benzoxaborinine-8-carboxylic acid, avibactam and nitrocefin were

synthesized in the laboratories of Basilea Pharmaceutica International Ltd according to published procedures.

Enzyme Inhibition Kinetics The interactions of the inhibitors with the two β -lactamases were studied using nitrocefin as reporter substrate. Preliminary apparent affinities (IC_{50}) were determined by incubating the enzymes with a range of concentrations of the inhibitor for 5 min. and then diluting with 100 μ L 100 μ M nitrocefin solution. The observed rates of modification of the enzyme by inhibitor were determined by incubating a 2 μ L reaction mixture containing the enzymes and the inhibitor for times between 1s and 180 min. before diluting with 100 μ L 100 μ M nitrocefin solution. Reactions requiring incubations less than 30s were performed using a SF61 multimixing stopped flow spectrophotometer (HI-Tech Scientific, Salisbury, UK). The recovery rates of free enzyme from inhibition were determined by incubating a 2 μ L reaction mixture containing the enzymes and the inhibitor to allow complete modification before diluting with 100 μ L 100 μ M nitrocefin solution. The reaction of the two enzymes with sulbactam was followed at 233nm and 270nm using a SF61 multimixing stopped flow spectrophotometer (Hi-Tech Scientific, Salisbury, UK). Enzyme concentrations were 50nM, for multiple turnovers, to 5 μ M for single turnover. Sulbactam concentrations were from 4 to 500 μ M. Curve fitting was performed using Grafit (Erithacus Software).

Crystal Growth and Structure Determination Purified SHV-5 samples [20 mM Tris-HCL, pH 8.0] were concentrated to between 22 mg/mL and 26.5 mg/mL, AHB dissolved in 100 % DMSO was added in a 1:4 molar excess to the protein and the solution was incubated for 1 hour. The drops were prepared by mixing 0.2 μ L SHV-5 - inhibitor solution with 0.2 μ L reservoir

solution from various commercial screens, resulting in a final DMSO concentration of 3 % and a ligand concentration of 4 mM respectively in the crystallization drops, and the drops were equilibrated against 50 μ L reservoir buffer. Thin, elongated, plate-like crystals grew within a few hours to one day to a sufficient size of 3 mm in length in 1.8 M Ammonium sulfate, 0.1 M BIS-TRIS pH 6.5, 2% PEG monomethyl ether 530]. For cryoprotection, 1 μ L of perfluoropolyether (Hampton Research) was added on top of the crystallization drop from a condition containing 1.8 M ammonium sulfate, 0.1 M Bis-Tris pH 6.5, 2% PEG monomethyl ether 530, and subsequently crystals were fished and flash frozen in liquid nitrogen.

Purified KPC-2 samples were concentrated to 15 - 18 mg/mL. Vapor diffusion using the crystallization conditions described by Ke et al.²⁷ were tested first (Table 1). The drops were prepared by mixing 1 μ L KPC-2 (17 mg/mL, in 20 mM Tris pH 7.6) with 0.5 μ L reservoir solution and were equilibrated against 500 μ L reservoir buffer. Most conditions yielded crystals but the crystallization drops contained many nuclei from which long thin rods of crystals up to a size of 1000 x 40 x 40 μm^3 emanated. To reduce the number of crystallization nuclei, the precipitant concentration, the pH value, the protein concentration and the crystallization temperature were varied. Decreasing temperature, PEG 6'000 or protein concentration did not result in bigger crystals. The crystals grew more slowly, but there were still a lot of nuclei and they never reached the size reported by Ke et al.²⁷. Single crystals of sufficient size were obtained by adding 10 mM spermine to the original crystallization condition. The crystals grew to a final size of 400 x 40 x 40 μm^3 within 5 days. In order to get KPC-2 inhibitor complex crystals, native KPC-2

crystals were soaked in a 5 μ L soaking solution, containing 100 mM bicine pH 9, 16% PEG 6'000, 10 mM spermine and 2 mM inhibitor for 4 h. For cryoprotection, the soaked crystals were successively placed in soaking solution with 5, 10, 15% (v/v) added glycerol with soaking times of 20 sec per step. Subsequently crystals were flash frozen in liquid nitrogen.

Data collection and processing

Data collection was performed at the beamline PXIII X06DA at the Swiss Light Source (Villigen, Switzerland). A complete dataset was collected to a resolution of 2.4 Å. For processing and scaling the programs MOSFLM/SCALA^{28 [collab]} were used. The parameters of the data collection and the statistics of the collected data set are summarized in Table 1.

RESULTS

Inhibition of SHV-5. SHV5 β -lactamase was strongly inhibited by sulbactam. At low concentrations, there was an initial rapid loss of approximately 20% of the initial activity (Fig. 2A), accompanied by a burst of increased absorbance (Fig 3A) at peak wavelengths of 233 nm, which accompanies acylation, and 278 nm and a shoulder around 300 nm, which accompany imine and enamine intermediate formation (Scheme 2). The initial rate of modification increased linearly with sulbactam concentration over the range studied (Fig. 2B), with an apparent second order rate constant $0.5 \pm 0.08 \times 10^6 \text{ M}^{-1} \cdot \text{s}^{-1}$. Further loss of activity occurred at a lower rate that showed saturation kinetics (Fig. 2C) and the inhibition eventually reached a limiting fractional residual activity, $F = 0.15$. This is expected in

a reaction where the rate of breakdown of the acyl intermediate is slower than the rate of its formation (eqn 1)²⁹

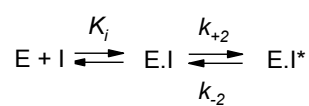


$$\text{where } k'_1 = k_{ac} [I]/(K_I + [I]) \quad \text{Eqn 2}$$

Fitting the observed rate of modification to Eqn 2 gives a limiting rate, $k_{ac} = 0.45 \pm 0.03 \text{ s}^{-1}$ and the apparent affinity, $K_I = 16 \text{ }\mu\text{M}$. That the residual activity does not go to zero but reaches a limiting value is attributable breakdown of the intermediate EI in Eqn 1 to regenerate free enzyme at a rate k_{rec} . The limiting fractional residual activity (F_∞) is given by $F_\infty = k_{rec}/k_{ac}$,³⁰ from which k_{rec} can be calculated to be 0.068 s^{-1} . The increase in A_{278} continued at a decreased rate during the secondary phase of inhibition, rising to a plateau value at about the same concentration as the limiting level of inhibition (Fig. 3). The change in A_{233} approached a steady-state rate of hydrolysis through a hyperstoichiometric burst, consistent with the branched pathway (Scheme 3A) expected for the reaction of sulbactam and class A β -lactamases.^{12, 31} The absorbance at 300 nm passed through a maximum early in the approach to the steady-state, suggesting the formation of an intermediate (Y in Scheme 3A) in the initial phase of reaction that decays in favor of a more stable complex (Z in Scheme 3A), through which most of the reaction flows in the steady-state. SHV-5 has been shown to accumulate the trans-enamine intermediate in an enzyme complex that has a conformation with decreased β -sheet structure during the steady-state phase of the reaction of sulbactam.¹³ After single turnovers, or when the sulbactam is depleted in multi-turnover experiments, the activity of the enzyme recovered at approximately the same rate as the

decay in A_{278} (Fig. 4). The rate of either recovery of activity or decay in A_{278} after single turnover was biphasic with an initial concentration-dependent phase that decreased from 0.12 s^{-1} at $0.25 \text{ }\mu\text{M}$ sulbactam to $0.05 \pm 0.01 \text{ s}^{-1}$ at $2 \text{ }\mu\text{M}$ and a slower concentration-independent phase with a rate of $0.015 \pm 0.004 \text{ s}^{-1}$. The concentration-dependent phase of recovery is dominated by the redistribution between intermediates X and Y (k_{+5} , upper limit 0.068 s^{-1}), whereas the concentration independent phase most likely approaches the decay of Z to X (k_4).³² After multiple turnovers, the rate of recovery was also biphasic, independent of the initial sulbactam concentration, with rates of 0.012 s^{-1} (amplitude approximately 10% of total activity) and 0.003 s^{-1} (amplitude approximately 90% of total activity). These two phases of recovery can be most simply identified with decay of intermediate Z through Y ($k_4, k_{+5} = 0.012 \text{ s}^{-1}$) and the slower decay of Z directly to regenerate free enzyme ($k_{+6} = 0.003 \text{ s}^{-1}$).

The boronic acid inhibitor, AHB, modified SHV-5 with an initial rapid loss of approximately 70% of the initial activity. The initial rate of modification increased linearly with AHB concentration over the range studied (Fig. 2B), giving an apparent second order rate constant of $2.1 \pm 0.1 \times 10^6 \text{ M}^{-1} \cdot \text{s}^{-1}$. Further loss of activity occurred at a lower rate that showed saturation kinetics (Fig. 2C). Boronic acids are generally considered to react with β -lactamases in a two-step mechanism in which a rapid equilibrium (non-covalent binding step) precedes a rate-determining step where covalent attachment occurs²¹



Eqn 3

The overall dissociation constant K_i^* is related to the dissociation constant of the first step by:

$$K_i^* = K_i \cdot \frac{k_{-2}}{(k_{+2} + k_{-2})} \quad \text{Eqn 4}$$

and the observed rate of modification is given by $k_{\text{obs}} = k_{-2} + (k_{+2} \cdot [I]/K_i)$.

The second order rate constant determined above = $k_{+2} / K_i = 2.1 \pm 0.1 \times 10^6 \text{ M}^{-1} \cdot \text{s}^{-1}$.

Determination of the AHB IC_{50} with 5 min incubation of enzyme and inhibitor gave biphasic dose response curves with 70 % of activity inhibited with high apparent affinity (20 nM) and the 30% of residual activity being inhibited with lower apparent affinity (9.5 μM) (Table 2). The recovery of free enzyme once the enzyme-inhibitor complex was diluted in the presence of excess nitrocefin over AHB was also biphasic, independent of the initial AHB concentration, with rates of 0.05 s^{-1} (amplitude approximately 30% of total activity) and 0.0009 s^{-1} (amplitude approximately 70% of total activity). The interaction of AHB with SHV-5 is thus not consistent with the simple two-step reaction represented in Eqn 3, but it is consistent with a modified branched pathway, analogous to that observed with sulbactam, where the isomerized enzyme-inhibitor complex Z, yields the starting molecule AHB and free enzyme (Scheme 3B).

SHV-5 was relatively slowly inhibited by avibactam. The reaction kinetics conformed with the two-step binding model proposed for the ESBL CTX-M 15. ^{24(Ehmann)} There was an initial loss of approximately 90% of the initial activity with a rate of modification that increased linearly with avibactam concentration over the range studied (Fig. 2B), giving an apparent second

order rate constant of $1.5 \pm 0.07 \times 10^5 \text{ M}^{-1}\cdot\text{s}^{-1}$. Recovery of the enzyme activity after dilution into nitrocefin was slow and apparently monophasic with a recovery rate of $2 \times 10^{-4} \text{ s}^{-1}$.

Crystal structure of a covalent AHB-SHV-5 complex. The crystals belonged to space group P2₁2₁2 with 2 molecules in the asymmetric unit and a solvent content of 43 %. The measured intensities were rather weak owing to the thinness of the crystals, which resulted in a relatively high overall R_{meas} (17.2 %). Analysis of the data by HKLVIEW²⁸ revealed that pseudo B-centering was present at low resolution. Indeed, a pseudo-origin peak was found in the Patterson map at 0.5, 0.0, 0.5, displaying half the height of the origin peak. The final model is comprised of all 264 residues in both subunits, 50 water molecules and two inhibitor molecules. The two molecules in the asymmetric unit of the complex structure were considered identical because the rmsd of the corresponding C α position was 0.35 Å. In the following the structure of subunit A is described. The SHV-5-inhibitor structure was solved by molecular replacement using PHASER³³ with chain A of SHV-1 D104K mutant (PDB entry code 2G2W,³⁴) as the initial search model. A unique solution with two molecules in the asymmetric unit was identified. The two subunits in the asymmetric unit form a dimer with 2-fold symmetry. Consistent with the pseudo B-centering subunit A and a symmetry mate of B are almost identically oriented and are related by a translation vector of (0.5, 0.0, 0.5). The model was refined with REFMAC (5.6.0117).²⁸ NCS restraints were applied for the two chains, and TLS refinement was carried out treating each subunit as an individual group. After refinement of the protein to an R/R_{free} of 28 % / 32 %, clear difference density for the ligand remained close to the active site Ser70. The inhibitor

was modeled into the electron density in the closed conformation with COOT.³¹ Water molecules were added manually and checked with COOT. The quality of the model was analyzed using MOLPROBITY³² and PROCHECK³³. All figures were prepared with PyMOL. Difference density for the inhibitor bound in the active site of the enzyme (Fig. 5) was already apparent in the first electron density map. Unexpectedly, the 2Fo-Fc map clearly suggested that the inhibitor was present in the closed conformation (Fig. 6) instead of the open conformation reported by Ness et al.²¹ No structure factors were deposited for that structure (PDB entry code 1ERQ), therefore an omit map could not be calculated to compare the ligand difference density obtained for both structures. There is clear evidence for the covalent link of the boron atom to the O γ side chain oxygen of Ser70 and to the exocyclic oxygen (O5) of the inhibitor, resulting in a tetrahedral stereochemistry of the boronate center. The inhibitor is stabilized via H-bonds with residues Ser70, Ser130, Asn132, Thr235, Ala237 and Arg243 (Fig. 6). The distances are listed in Table 3. The average B-factor for the inhibitor in molecule A is 19.1 Å² and 17.9 Å² for molecule B respectively. In the active site (Fig. 6), a water molecule is found in the same position as the deacylation water observed in many β -lactamase structures before.^{38,39} It is interacting with Ser70, Glu166 and Asn170 via hydrogen bonds and has a B-factor of 10.9 Å².

Inhibition of KPC-2 Compared to SHV-5, KPC-2 was slowly and inefficiently modified by sulbactam: stable acylation, as seen with SHV-5, could not be demonstrated and sulbactam behaved as a competing substrate towards nitrocefin. The enamine intermediates with absorption around 280 nm were only observed at an early stage of the reaction, prior to

establishment of the steady-state hydrolysis (Fig. 3B). Their abundance decayed as the steady-state was established and they were only present at a very low level during the steady-state, consistent with the lack of stable acylation. It appears that KPC-2 establishes a branched pathway (Scheme 3) in which direct hydrolysis of the acyl-enzyme intermediate (equivalent to intermediate Y in the reaction with SHV-5) is favoured over the chemical rearrangement leading to enamine intermediates and a stabilizing enzyme conformation (intermediate Z in the reaction with SHV-5).

Modification of KPC-2 by both AHB and avibactam conformed with the two-step reaction previously proposed for avibactam and KPC-2.²⁴ The rate of modification was slow compared to SHV-5 and the monophasic dose response curves gave lower apparent affinities (IC_{50}) around 500 and 200 nM for AHB and avibactam, respectively (Table 2). The recovery of free enzyme after dilution of the modified enzyme into excess nitrocefin was slow for both types of inhibitor. In both cases, the effective dissociation constant (K_d , = recovery rate/modification rate)²⁴ was similar to the IC_{50} determined by competition (0.51 vs 0.50 and 0.2 vs 0.17 for K_d and IC_{50} for AHB and avibactam, respectively).

Crystal structure of the avibactam-KPC-2 complex A complete dataset was collected to a resolution of 2.2 Å. For processing and scaling the programs MOSFLM/SCALA²⁸ were used. The parameters of the data collection and the statistics of the collected data set are summarized in Table 1. The crystals belonged to space group P3₁ with 3 molecules in the asymmetric unit and a solvent content of 50.5%. Unfortunately, the data set turned out to be twinned, as indicated by the Meroherdal Crystal

Twinning Server.⁴⁰ The twinning operation was (h, -h-k, -l) and the twinning fraction 0.41. The dataset was detwinned using DETWIN²⁸ and the structure was determined with the detwinned data by molecular replacement using MOLREP⁴¹ using chain A of the KPC-2 structure (PDB entry 2OV5) as the initial search model. MOLREP found a unique solution with three molecules in the asymmetric unit. The model was refined with REFMAC5.²⁸ NCS restraints were applied between the three chains, and TLS refinement was carried out using each chain as an individual group. The model of the inhibitor was generated with PRODRG⁴² and modeled into electron density with COOT.³⁵ Water molecules were added with ARP⁴³ and manually checked with COOT. The quality of the model was analyzed using PROCHECK³⁷. All figures were prepared with PyMol. Considering that the data had to be detwinned, the model has good quality indicators (Table 4) and 99.5% of the residues had Φ and Ψ angles in the allowed regions of the Ramachandran plot. The final model is comprised of all 261 residues in all three chains, 93 water molecules and three inhibitor molecules. The complexed protein structure is virtually identical to the native structure (rmsd C^α of 0.06 Å). The three molecules in the asymmetric unit of the complex structure were considered identical, because the rms C^α differences of the pairwise comparisons were below 0.06 Å. In the following, we describe the structure of molecule A. The inhibitor appears covalently linked by an ester bond to Ser70 and extends into the active side cleft (Figure 7). KPC-2 molecule A showed the best density. The omit electron density is shown in Figure 8. The average B factor for the inhibitor in KPC-2 molecule A is 36.2 Å², for molecules B and C the average B factors are somewhat higher (43.6 Å²

and 52.8 \AA^2). The inhibitor is stabilized via H-bonds with residues Ser70, Ser130, Asn132, Thr235 and Thr237 and hydrophobic contact with the residue Trp105 (Figure 6). The distances are listed in Table 3. A water molecule is found in the active site in a position to potentially deacylate the enzyme (Figure 8). It is interacting with Glu166 and Asn170 and has a B factor of 37.4 \AA^2 .

DISCUSSION

The kinetic analysis suggests that branched pathways involving enzyme conformation change as well as chemical modification of reactive intermediates, which were expected for the reaction between sulbactam and SHV-5, are may underlie the reactions of both this ESBL and the carbapenemase KPC-2 with other, non- β -lactam, classes of inhibitor. The structure of the KPC enzymes appears have evolved to disfavor the formation of chemically stabilized intermediates, such as the enamines formed with clavulanic acid, sulbactam and tazobactam, as well as the Δ^1 -tautomer of the pyrroline double bond in the acyl-enzyme species that occurs in the reaction with carbapenems.¹²

The structure of the AHB/SHV-5 complex was compared to structures of the closely-related class-A β -lactamases SHV-1, SHV-2 and TEM-1.^{22,39,44,45} As expected from the high sequence similarity of these proteins, a good superposition of the overall-structures is observed (RMSD $1.0\text{-}1.2\text{\AA}$). Interestingly, the conformation of the H10-helix of SHV-5 (residues 218-230) deviates significantly from the ones observed in the structure of SHV-1 and

SHV-2 but is identical to the one in the structure of AHB-TEM-1 complex. Since this deviation has been observed previously³⁴ and has been linked to the presence of a detergent molecule, this observation is not discussed in detail here. Relative to the wild-type enzyme SHV-1, SHV-5 differs by the mutations Gly238Ser and Glu239Lys, which are located in a short loop which connects B3 and B4 β -sheets. Very similar conformations are observed for this loop in both monomers but only weak electron density is observed for residues Lys239 and Arg240 in monomer A, which reflects the conformational plasticity of this loop. However, the loop conformation of SHV-5 deviates significantly to the ones observed in X-ray structures of SHV-1.^{44,45} Residues 240–241 have flipped, resulting in a drastic shift of the corresponding C α -positions (3.8 Å for of Lys239 monomer B). The finding correlates well with the structure of SHV-2, for which an altered conformation of this loop is also described, albeit to a lesser extent.³⁹ As in case of the structure of SHV-2, no significant alteration of the catalytic residues in the active site is observed.

Interestingly, the electron density suggests binding of the inhibitor in the form of a bicyclic boronate ester, which is in contrast to its TEM-1 binding mode²² where the boronic acid state is observed. Despite this difference, the key interactions of the inhibitor observed in the SHV-5 structure are highly similar to the ones observed for TEM-1. In brief, strong hydrogen-bond interactions are formed to the the main chain nitrogens of Ser70 and Ala237, the side chains of Arg244 and Lys234 and to the backbone-carbonyl, the main chain carbonyl at Ala237 and the side chain amide nitrogen of Asn132. Also similar to TEM-1, favorable van der Waals interactions between the aromatic side chain of Tyr105 and the aromatic ring of the

inhibitor is observed. However, since an alternative conformation is observed for Tyr105 in which the side-chain is rotated by almost 120 degrees, this interaction might be less relevant compared to TEM-1.

A high similarity of the overall-structures was observed upon superposition of the structure of the KPC-2/avibactam-complex with the structures reported by Ke et al.,^{27,47} in which either a bicine molecule (pdb-code 2OV5) or a 3-nitrophenyl-boronic acid (3-NBA, pdb-code 3RXX) are present in the active site. A good similarity is also observed for the active sites, including the catalytic residues and the di-sulfide bridge formed by Cys69 and Cys238. However, significant differences could be observed for the orientation of the loop at the end of the β 3-strand. Whereas the structures of the avibactam and the bicine-complex superimpose well, the loop adopts a different conformation in case of the 3-NBA complex resulting in the shift of the carbonyl-oxygen of Gly239 by 2.9Å. Furthermore, a significant deviation of the 3-NBA-complex-structure is observed on the opposite side of the pocket. Here the loop between Val103 and Ser106 is shifted, which also affects the side-chain conformation of Trp105. This strong conformational plasticity is reflected by a distance of 6.5 Å between the positions of the aromatic nitrogen of Trp105 of the avibactam and 3-TBA-complex. Also a high similarity of the KPC-2/avibactam-complex is observed to the structure KPC-2 in the apo-state (pdb-code 3DW0),⁴⁸ which closely resembles the structure of the 3-NBA-complex, in particular regarding the orientation of Trp105. However, a significant deviation is observed for the side-chain orientation of catalytic residue Ser130. Here, the side-chain is directed towards a water molecule located between Thr235 and Val127, whereas in the other structures, the side-chain is pointing towards

Ser90. The importance of the de-acylation water molecule, which is located between the residues Ser70 and Asn170, is reflected by the only minor variation of its position observed in the different structures.

The binding mode of avibactam to KPC-2 was compared to the ones observed for complexes with the class-A β -lactamase CTX-M 15.^{49, 50} As expected from the high similarity of the active sites of CTX-M 15 and KPC-2, the inhibitors superimpose well in the respective structures, with an average distance of approx. 0.5 Å for atoms of the piperidine ring. The amide-groups are located in an almost identical position and the piperidine rings share a similar conformation. However, a notable difference is observed for the orientation of the sulfoxyamino-groups. Whereas the sulfur-atoms are located in close proximity, the nitrogen and oxygen atoms of the linker adopt different positions. In case of the KPC-2/avibactam complex, the oxygen is in a position beneath the piperidine ring whereas for CTX-M 15, the N-O bond is pointing away from the piperidine. As a consequence, protons of the nitrogen are modeled in a more distant position from the catalytic serine residues in case of KPC-2 (2.4 and 3.4 Å to Ser130 and Ser70 compared to 1.9 and 2.5 Å in case of CTX-M 15). The difference in conformation of the N-O bond and consequent different hydrogen-bonding may decrease the propensity for neighbouring group attack by N6 to regenerate the starting inhibitor molecule and free KPC-2 and therefore account for the unique degradation pathway observed for KPC-2.²⁴

FIGURES

Fig. 1 Structures of (A) sulbactam, (B) (3R)-3-acetamido-2-hydroxy-3,4-dihydro-1,2-benzoxaborinine-8-carboxylic acid (AHB) and (C) avibactam.

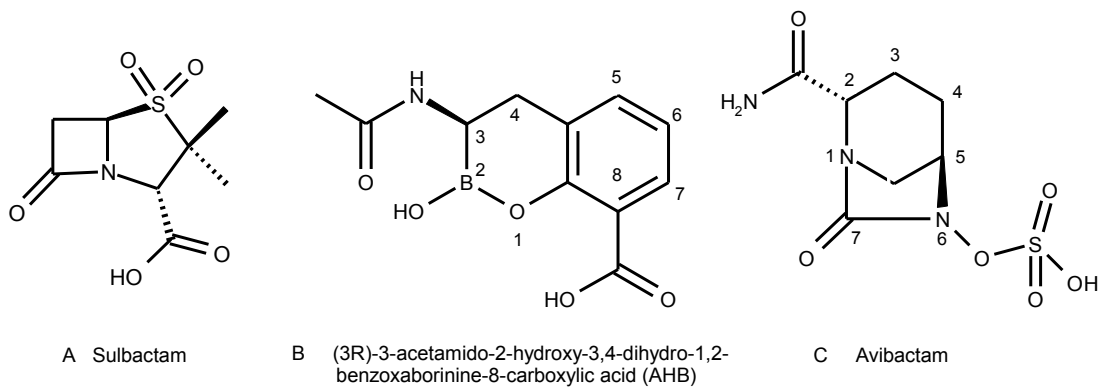


Figure 2. Inhibition of SHV-5: (A) time course for the initial modification by sulbactam at 1 μ M (solid circles), 2 μ M (solid triangles), 5 μ M (solid squares), 10 μ M (solid diamonds), 20 μ M (open circles), 40 μ M (open triangles) and 80 μ M (open squares); (B) initial second order rates of modification by sulbactam (solid circles), AHB (solid squares) and avibactam (solid triangles); (C) limiting first order rates of modification by sulbactam (circles) and AHB (solid squares)

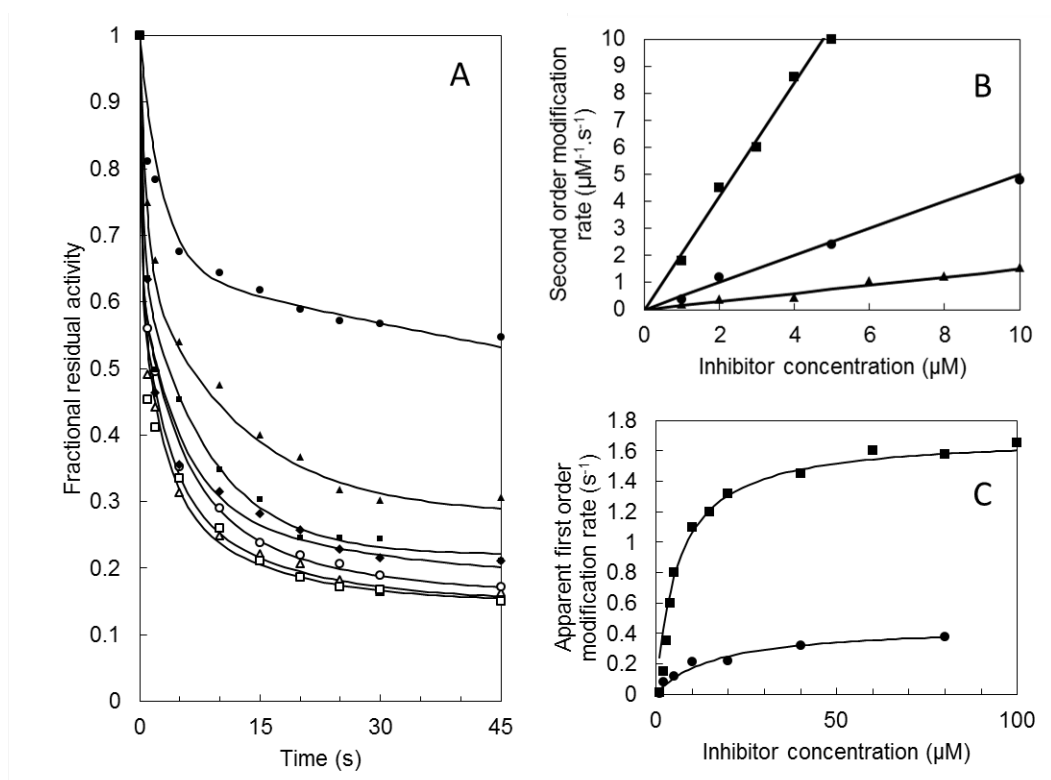


Figure 3: Time course of the reaction between sulbactam (500 μ M) and either (A) SHV-5 at 1 μ M or (B) KPC-2 at 5 μ M, followed at 233 nm (purple lines), 278 nm (blue lines) and 300 nm (orange line).

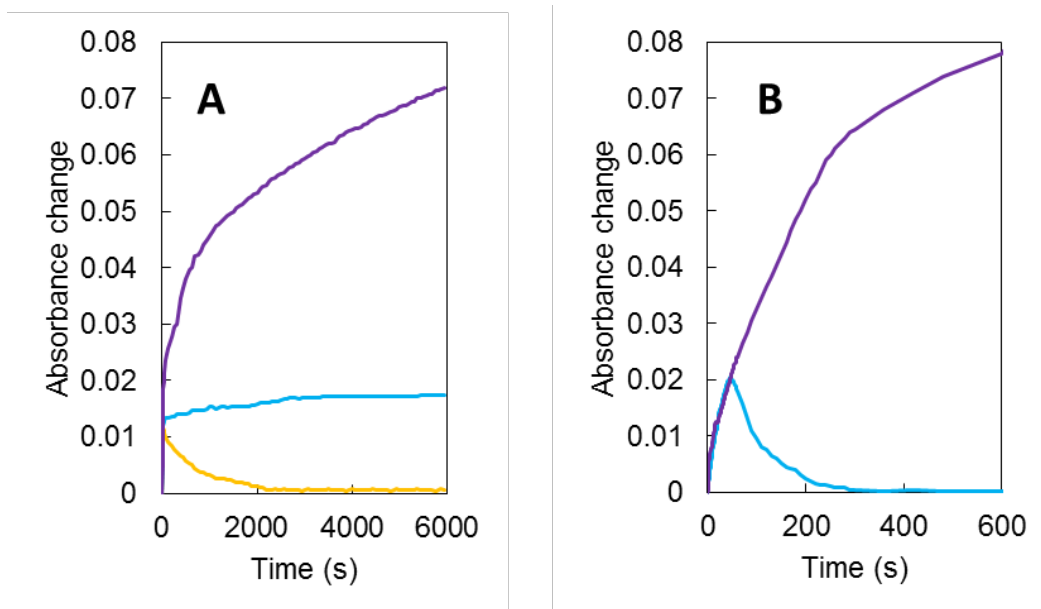


Figure 4: (A) Time course of the reaction between sulbactam and SHV-5 (1 μ M) at 2 μ M (blue line and triangles) and 20 μ M (orange line and triangles). The solid lines show the absorbance change at 278nm, normalized with respect to the mean steady-state absorbance, and the triangles show the fractional residual activity towards nitrocefin hydrolysis. (B) Semi-logarithmic replots of the data shown in A, for sulbactam 2 μ M (blue line and circles). and 20 μ M (orange line and circles: time points recalculated to show the time elapsed after the end of the steady-state) with the residual activity transformed to fractional inhibition. Also shown are the fractional inhibition of SHV-5 (squares) and KPC-2 (triangles) by (3R)-3-acetamido-2-hydroxy-3,4-dihydro-1,2-benzoxaborinine-8-carboxylic acid after multiple turnovers.

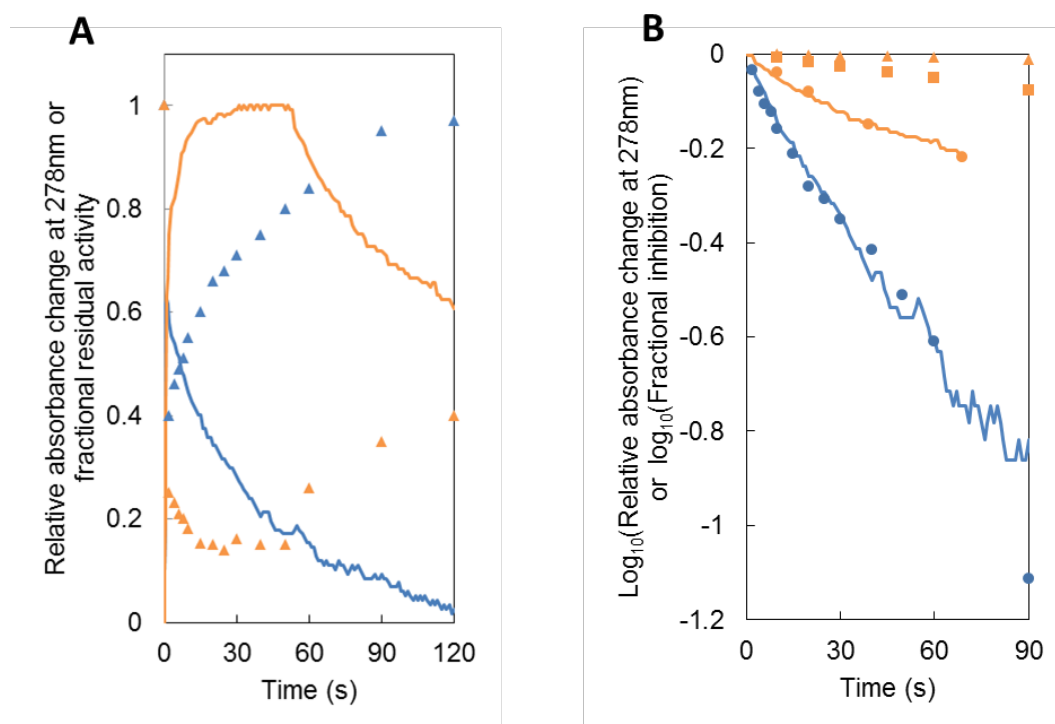


Figure 5: Structure of SHV-5 β -lactamase in complex with the inhibitor (3R)-3-acetamido-2-hydroxy-3,4-dihydro-1,2-benzoxaborinine-8-carboxylic acid. (A) Cartoon diagram of SHV-5; the inhibitor moiety is covalently bound to Ser70 (shown in sticks). (B) Surface representation of SHV-5; the inhibitor moiety is located in the catalytic cleft.

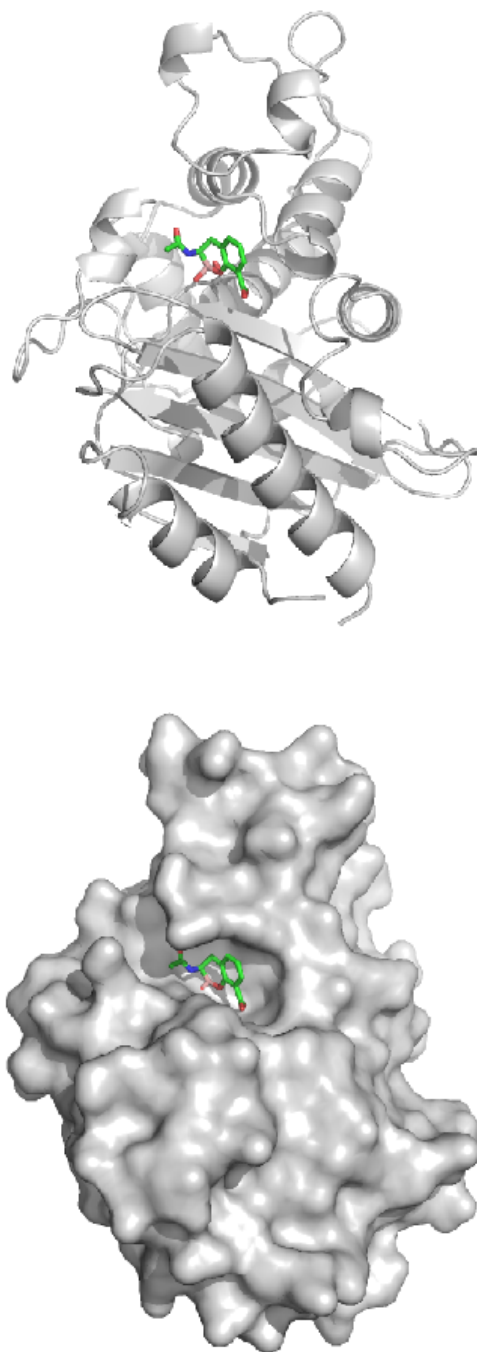


Figure 6: Two stereo views of the active site of the SHV-5 β -lactamase with bound AHB. The interaction residues and the inhibitor are shown in full. The H-bonds are shown as dashed lines. The omit map for the ligand is contoured at 2σ . The F_o-F_c omit map, contoured at 2σ , clearly supports the scenario in which the ligand binds to Ser70 in the closed conformation.

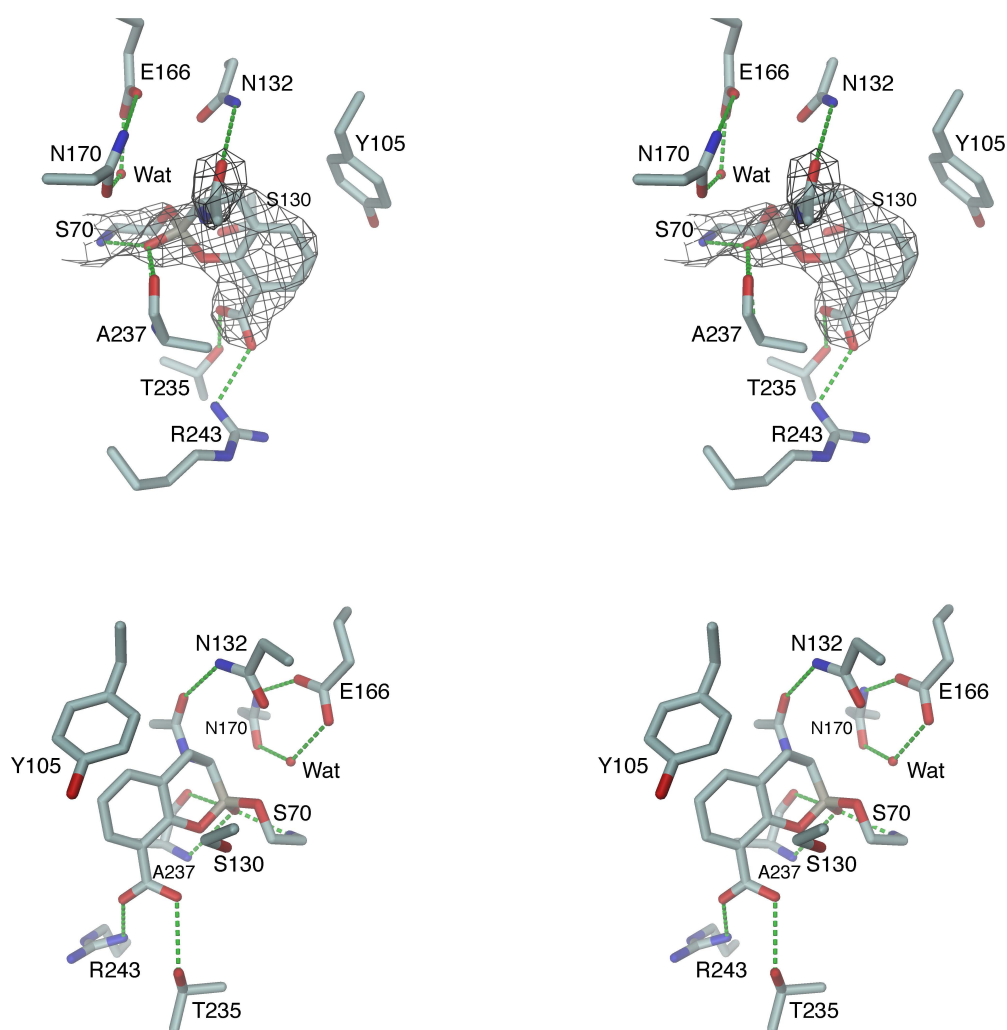


Figure 7. Structure of the KPC-2 β -lactamase in complex with the inhibitor avibactam. (A) Ribbon diagram of the KPC-2; the inhibitor moiety is covalently bound to Ser70 (shown in sticks). (B) Surface representation of the KPC-2; the inhibitor moiety is located in the catalytic cleft.

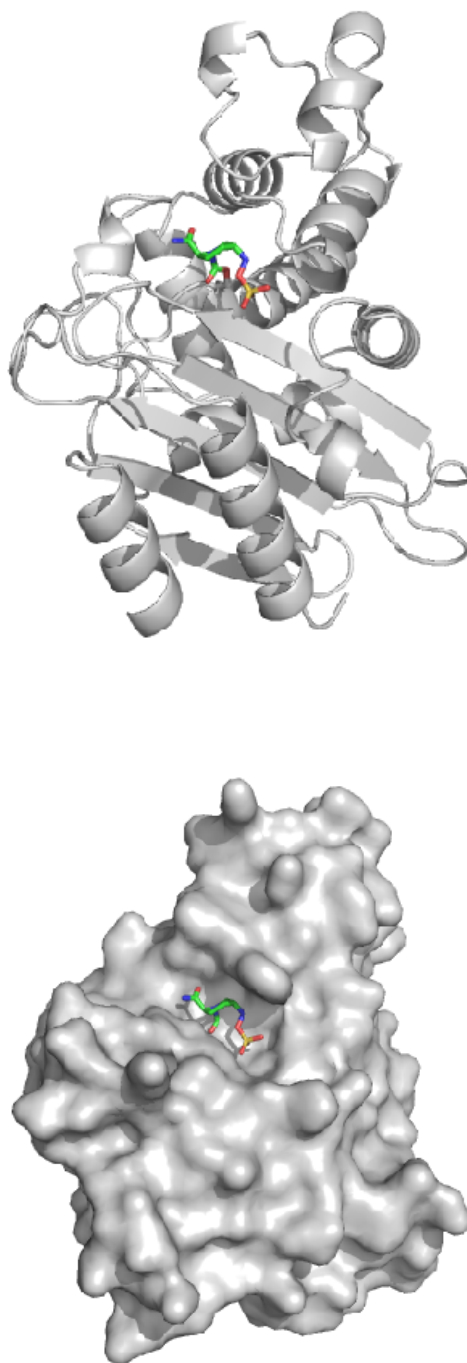
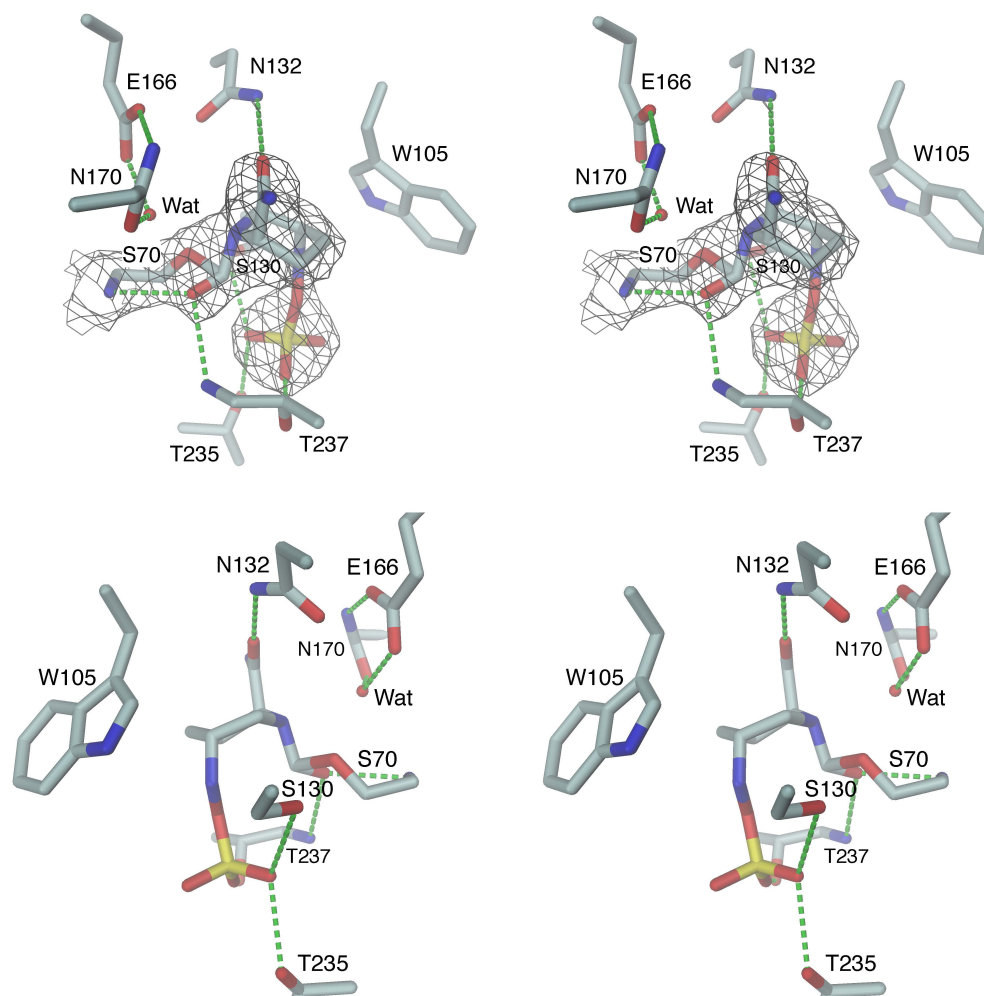
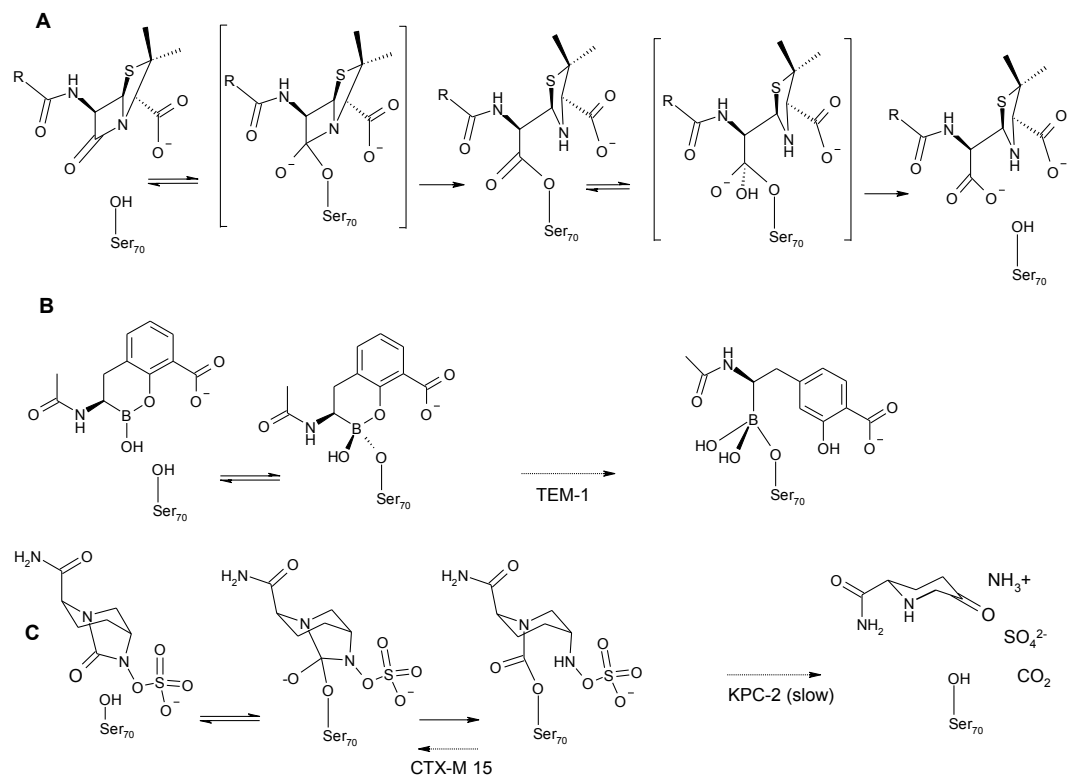


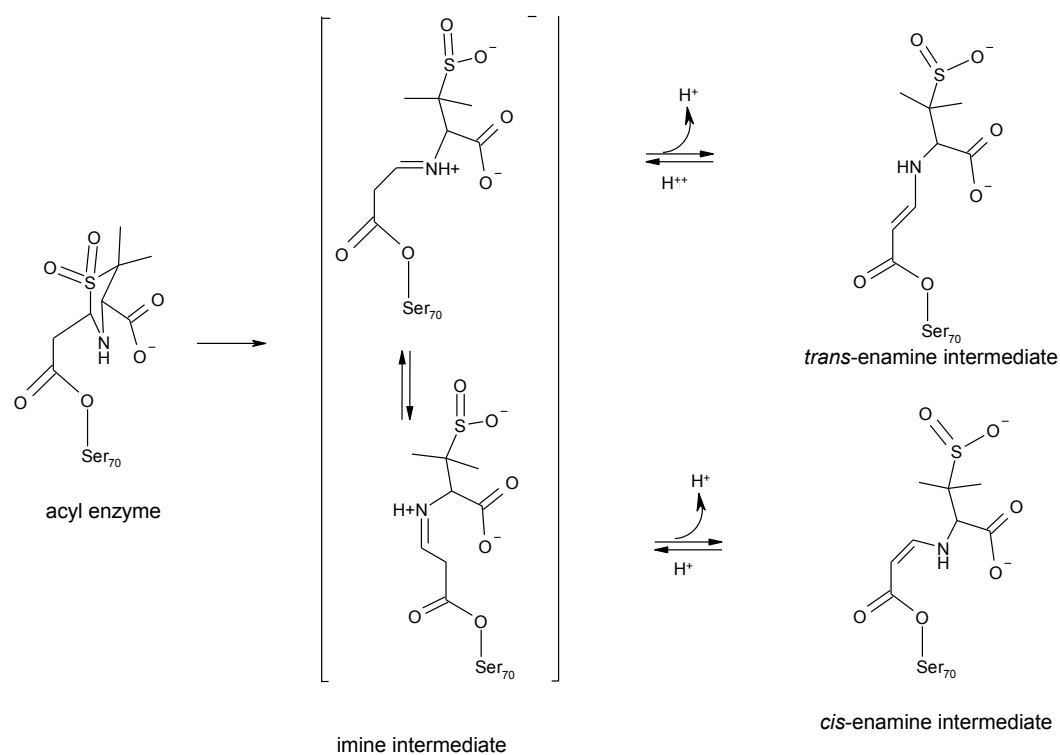
Figure 8. Two stereo views of the active site of the KPC-2 β -lactamase with bound avibactam. The interaction residues and the inhibitor are shown in full. The H-bonds are shown as dashed lines. The omit map for the ligand is contoured at 2σ .



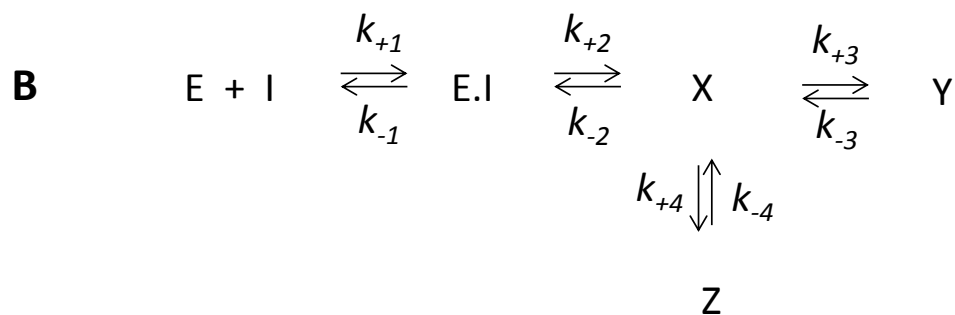
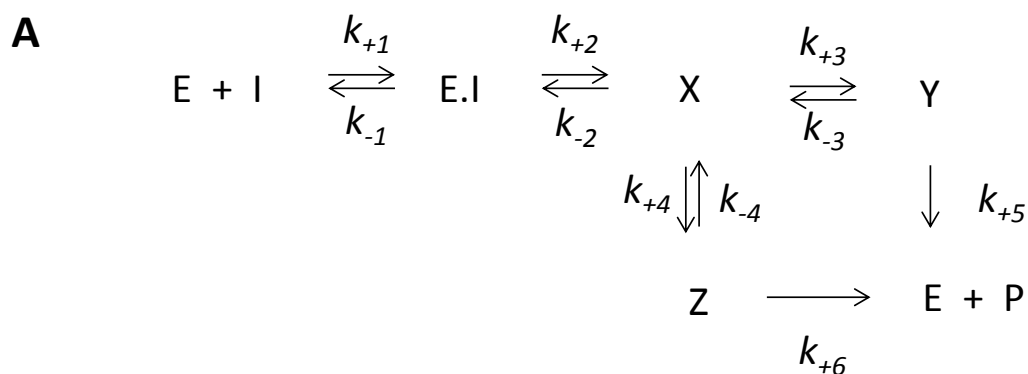
Scheme 1. Schematic representation of the intermediates possible during reaction of a β -lactamase with (A) a penicillin, (B) 3-acetamido-2-hydroxy-3,4-dihydro-1,2-benzoxaborinine-8-carboxylic acid and (C) avibactam



Scheme 2. Schematic representation of the intermediates possible during rearrangement of the acyl-enzyme intermediate formed with sulbactam



Scheme 3. Branched pathways involving enzyme conformation change, modified from Frere (1981) and Waley (1991), showing (A) a reaction pathway proceeding to a product (e.g. sulbactam and SHV-5) and (B) and inhibition pathway regenerating the inhibitor and free enzyme (e.g. AHB and SHV-5).



Contrasting Mechanism of Inhibition of SHV5 and KPC2 β -Lactamases

Table 1: Data collection and refinement statistics. Statistics for the highest-resolution shell are shown in parentheses.

Parameter	SHV5	KPC-2
Wavelength [\AA]	1	1
Resolution range [\AA]	27 - 2.4 (2.486 - 2.4)	30 - 2.2
Space group	P 21 21 2	P31
Unit cell [\AA , $^\circ$]	69.7 106.8 62.6 90 90 90	116.29 116.29 52.79 120 120 120
Total reflections	41629 (6276)	95396 (13326)
Unique reflections	17676 (2616)	40113 (5809)
Multiplicity	2.6 (2.4)	2.4 (2.3)
Completeness [%]	93.0 (97.6)	99.0 (98.1)
Mean I/sigma(I)	4.6 (2.3)	15.7 (2.4)
R-merge [%]	12.4 (35.2)	7.1 (30.2)
R-meas [%]	17.2 (48.8)	9.1 (39.1)
R-work [%]	22.2	23.6
R-free [%]	28.8	27.5
Numbers of non-hydrogen atoms		
protein	4111	11652
ligands	36	51
RMS(bonds) [\AA^2]	0.021	0.010
RMS(angles) [$^\circ$]	1.85	1.174
Ramachandran favored [%]	96	94.3
Ramachandran allowed [%]	3.8	5.1
Average B-factors [\AA^2]		
protein	17.07	40.4
ligands	18.52	44.2
solvent	12.43	33.6

Table 2. Observed Kinetic Parameters Describing Inhibition of SHV-5 and KPC-2

Contrasting Mechanism of Inhibition of SHV5 and KPC2 β -Lactamases

Parameter	SHV-5			KPC-2			
	SUL	AHB	AVI	SUL	AHB	AVI	
Apparent affinity (μM)	IC ₅₀		0.028	0.019			
	(Residual activity, %)	0.6	(72)	(95)	>100	0.51	0.17
			9.5	2.8		(100)	(100)
			(28)	(5)			
K _i , initial phase of modification	16	6	ND	ND	ND	ND	
K _d , effective dissociation constant	ND	ND	0.023	ND	0.50	0.20	
Modification rate	Apparent 2 nd order rate constant, initial phase, M ⁻¹ .s ⁻¹	5.1 x 10 ⁵	2.1 x10 ⁶	1.5 x 10 ⁵	ND	7.5 x 10 ⁴	1.4 x 10 ⁴
	Limiting rate of modification, k _{ac} , s ⁻¹	0.45	1.7	ND	ND	ND	ND
	Calculated rate from limiting fractional activity, k _{rec}	0.068	ND	ND	ND	ND	ND
Recovery rate (s ⁻¹)	Recovery rate · Single turnover (Residual activity, %)	0.05 (60%) 0.015 (40%)	ND	ND	ND	ND	ND
	Recovery rate · Multiple turnover or equilibrium mixture for AHB and avibactam. (Residual activity, %)	0.012 (10%) 0.003 (90%)	0.05 (30%) 0.0009 (70%)	2 x 10 ⁻⁴ (100)		3.8 x 10 ⁻⁴ (100)	0.003 (100)

Table 3: Hydrogen Bonding Distances in the Inhibitor/Enzyme Complexes

Location	Inhibitor Atom	Enzyme atom	Deacylation water	Distance (Å)	
				SHV-5	TEM-1 ^a
	OB1	Ser70N	-	2.40	(Ser70 N) 2.80
	OB1	Ser70OG	-	2.51	(Ser70 OG) 2.31
	O5	Ser130 OG	-	3.06	(Ser130 OG) 2.73
	O2	Asn132 ND2	-	3.10	(Asn132 ND2) 2.90
	O3	Thr235 OG1	-	2.88	(Ser235 OG1) 2.80
	N1	Ala237 O	-	3.01	(Ala237 O) 3.20
Boronate	OB1	Ala 237 N	-	3.15	(Ala237 N) 3.07
	O4	Arg243 NH1	-	2.85	(Arg243 NH1) 2.80
Deacylation		Glu166 OE2	H ₂ O 37	2.48	2.50
water		Asn170 OD1	H ₂ O 37	2.78	2.70
				KPC-2	CTXM-15 ^b
	O7	Ser70 N	-	2.87	(Ser70 N) 2.71
	O3	Ser130 OG	-	3.32	(Ser130 OG) 3.08
Avibactam	O14	Asn132 ND2	-	3.21	(Asn132 ND2) 2.85
	O3	Thr235 OG1	-	2.77	(Thr235 OG1) 2.62
	O4	Thr237 OG1	-	2.42	(Ser237 OG) 3.29
	O7	Thr237 N	-	2.95	(Ser237 N) 2.90
Deacylation		Glu166 OE2	H ₂ O 35	2.48	(Glu166 OE2) 2.60
water		Asn170 OD1	H ₂ O 35	2.32	(Asn170 OD1) 2.71
			H ₂ O 35	2.89	

^aNess et al. (2000) *Biochemistry* 39, 5312-5321. ^bLahiri et al. (2013) *Antimicrob. Agents Chemother.* 57, 2496-2505.

AUTHOR INFORMATION

Corresponding Author

*Corresponding author. Mailing address: Basilea Pharmaceutica International Ltd, Grenzacherstrasse 487, CH-4058 Basel, Switzerland. Phone +41616061399. FAX +41616061112

E-mail:Malcolm.page@basilea.com

Present Addresses

†Present Address: Actelion Pharmaceuticals Ltd, Gewerbestrasse 16, CH-4123 Allschwil, Switzerland

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript. ‡These authors contributed equally.

ACKNOWLEDGMENT

We thank Malgosia Kania for her excellent technical assistance during part of this work.

ABBREVIATIONS

REFERENCES

1. Page, M. G. P. (2012a) Beta-Lactam Antibiotics. In: Dougherty, T. J., and Pucci, M. J. (Eds) Antibiotic Discovery and Development. Springer Science+Business Media, LLC, New York, Dordrecht, Heidelberg, London. pp.79-117.
2. Frère, J.-M., and Page, M. G. P. (2014) Penicillin-binding proteins: evergreen drug targets. *Current Opinion in Pharmacology* 18: 112-119.
3. Cho, H., Uehara, T., and Bernhardt, T. G. (2014) Beta-lactam antibiotics induce a lethal malfunctioning of the bacterial cell wall synthesis machinery. *Cell* 159: 1300-1311.
4. Page, M. G. P. (2012b) The role of the outer membrane of Gram-negative bacteria in antibiotic resistance: Ajax' shield of Achilles' heel? In: Coates, A.R.M. (Ed.) Antibiotic Resistance, Handbook of Experimental Pharmacology 211. Springer Verlag Berlin Heidelberg. pp. 67-86.
5. Bush, K. (2012) Evolution of β -lactamases: past, present and future. In: Dougherty, T. J., and Pucci, M. J. (Eds) Antibiotic Discovery and Development. Springer Science+Business Media, LLC, New York, Dordrecht, Heidelberg, London. pp.427-453.
6. Ambler, R. P. (1980) The structure of β -lactamases. *Philos. Trans. R. Soc. Lond. [Biol.]* 289, 321-331.
7. Huovinen P., Huovinen S, Jacoby G. A. (1988) Sequence of PSE-2 beta-lactamase. *Antimicrob. Agents Chemother.* 32. 134-136.
8. Jaurin, B., Grundstrom, T. (1981) ampC cephalosporinase of *Escherichia coli* K-12 has a different evolutionary origin from that of

- β -lactamases of the penicillinase type. Proc. Natl Acad. Sci USA 78:4897-4901.
9. Bush, K., and Jacoby G. A. (2010) Updated Functional Classification of β -Lactamases. Antimicrob Agents Chemother. 54, 969–976.
 10. Jacoby G. A., Bush K. (2015) β -Lactamase Classification and Amino Acid Sequences for TEM, SHV and OXA Extended-Spectrum and Inhibitor Resistant Enzymes. [Online: consulted].
 11. Billot-Klein, D., L. Gutmann, and E. Collatz. 1990. Nucleotide sequence of the SHV-5 β -lactamase gene of a Klebsiella pneumoniae plasmid. Antimicrob. Agents Chemother. 34:2439-2441.
 12. Page, M.G.P. (2008) Extended-spectrum β -lactamases: structure and kinetic mechanism. CMI 14(Suppl.1) 63-74
 13. Kalp, M., Bethel, C. R., Bonomo, R. A. & Carey P. R. (2009) Why the Extended-Spectrum β -Lactamases (ESBLs) SHV-2 and SHV-5 are “Hyper-Susceptible” to Mechanism-Based Inhibitors. Biochemistry 48, 9912–9920.
 14. United States Department of Health and Human Services and Centers for Disease Control and Prevention. Antibiotic resistance threats in the United States, 2013, Atlanta, GA. <http://www.cdc.gov/drugresistance/threat-report-2013/index.html>.
 15. Yigit, H., Queenan, A. M., Rasheed, J. K., Biddle, J. W., Domenech-Sanchez, A., Alberti, S., Bush, K., and Tenover, F. C. (2003) Carbapenem-resistant strain of Klebsiella oxytoca harboring carbapenem-hydrolyzing β -lactamase KPC-2, Antimicrob Agents Chemother 47, 3881-3889.
 16. da Silva R.M. , Traebert J., Galato D. (2012) Klebsiella pneumoniae

- carbapenemase (KPC)-producing *Klebsiella pneumoniae*: a review of epidemiological and clinical aspects. *Expert Opin Biol Ther.* 12: 663-671.
17. Walther-Rasmussen, J. and Høiby, N. (2007) Class A carbapenemases. *J. Antimicrob. Chemother.* 60, 470-482
18. Hofer, B., Dantier, C., Gebhardt, K., Desarbres, E., Schmitt-Hoffmann, A., and Page, M.G.P. (2013) Combined effects of the siderophore monosulfactam BAL30072 and carbapenems on multidrug-resistant Gram-negative bacilli. *J. Antimicrob Chemother.* 68:1120-1129.
19. Papp-Wallace, K. M., Bethel, C. R., Distler, A. M., Kasuboski, C., Taracila, M., and Bonomo, R. A. (2010) Inhibitor resistance in the KPC-2 β -lactamase, a preeminent property of this class A β -lactamase, *Antimicrob Agents Chemother* 54, 890-897
20. Ehmman, D. E., Jahic, H., Ross, P. L., Gu, R. F., Hu, J., Kern, G., Walkup, G. K., and Fisher, S. L. (2012) Avibactam is a covalent, reversible, non- β -lactam β -lactamase inhibitor, *Proc Natl Acad Sci U S A* 109, 11663-11668
21. Crompton, I E , Cuthbert, B K, Lowe, G and Waley S G (1988) Beta-lactamase inhibitors. The inhibition of serine beta-lactamases by specific boronic acids. *Biochem J.* 251 453–459.
22. Ness, S., Martin, R., Kindler, A. M., Paetzel, M., Gold, M., Jensen, S.E., Jones, J.B., Strynadka, N.C. (2000). Structure-based design guides the improved efficacy of deacylation transition state analogue inhibitors of TEM-1 beta-lactamase. *Biochemistry* 39, 5312-5321.
23. Hecker, S., Reddy, K.R., Totrov, M., Hirst, G.C., Lomovskaya, O.,

- Griffith D. C., King, P., Tsivkovski R., Sun D., Sabet, M., Tarazi, Z., Clifton, M. C., Atkins, K., Raymond, A., Potts, K. T., Abendroth, J., Boyer, S. H., Loutit, J. S., Morgan, E. E., Durso, S., Dudley, M. N.. (2015). Discovery of a cyclic boronic acid β -lactamase inhibitor (RPX7009) with utility vs class A serine carbapenemases. *J. Med. Chem.* DOI:10.1021/acs.jmedchem.5b00127.
24. Ehmman, D. E., Jahic, H., Ross, P. L., Gu, R. F., Hu, J., Durand-Reville, T. F., Lahiri, S., Thresher, J., Livchak, S., Gao, N., Palmer, T., Walkup, G. K., and Fisher, S. L. (2013) Kinetics of avibactam inhibition against Class A, C, and D β -lactamases, *J Biol Chem* 288, 27960-27971
25. Sambrook, J. and Russell, D. W. (2001) *Molecular cloning, a laboratory manual*, 3rd edn, Cold Spring Harbor Laboratory Press, Cold Spring Harbour, New York
26. Stüber, D., Matile, H., and Garrotta, G. (1990) In *Immunological Methods* (Lefkovits, I., and Pernis, B., eds) Academic Press, Orlando, pp. 121-152.
27. Ke, W., Bethel, C. R., Thomson, J. M., Bonomo, R. A. & van den Akker, F. (2007). Crystal structure of KPC-2: Insights into carbapenemase activity in class A β -lactamases. *Biochemistry* 46, 5732 – 5740
28. Collaborative Computational Project, Number 4 (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* 50, 760–763.

29. Fersht, A. (1999) Structure and mechanism in protein science. A guide to enzyme catalysis and protein folding. New York: Freeman. 631 pp.
30. Christensen, H., Martin, M. T., and Waley, S. G. (1990) Beta-lactamases as fully efficient enzymes. Determination of all the rate constants in the acyl-enzyme mechanism. *Biochem J.* 266: 853–861.
31. Li, R., Wang, Y.T., Chen, C. L. (2013) Why tazobactam and sulbactam have different intermediates population with SHV-1 β -lactamase: a molecular dynamics study. *J Mol Model.* 19:2519-24.
32. Waley, S. G. (1991) Kinetics of substrate-induced inactivation. *Biochem. J.* 279: 87-94
33. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. A., Storoni, L. C., Read, R. J. (2007). Phaser Crystallographic Software. *J. Appl. Cryst.* 40, 658-674
34. Reynolds, K.A., Thomson, J.M., Corbett, K.D., Bethel, C.R., Berger, J.M., Kirsch, J.F., Bonomo, R.A., Handel, T.M. (2006) Structural and computational characterization of the SHV-1 beta-lactamase-beta-lactamase inhibitor protein interface. *J.Biol.Chem.*281: 26745-26753
35. Emsley, P. & Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallographica D* 60, 2126–2132.
36. Chen et al. (2010). MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallographica D* 66, 12-21
37. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* 26, 283-291.

38. Minasov, G., Wang, X., Shoichet, B.K. (2002). An Ultrahigh Resolution Structure of TEM-1 β -Lactamase suggests a Role for Glu166 as the General Base in Acylation. *J. Am. Chem. Soc.* 124, 5333-5340
39. Nukaga, M., Mayama, K., Hujer, A.M., Bonomo, R.A., Knox, J.R. (2003). Ultrahigh Resolution Structure of a Class A β -Lactamase: On the Mechanism and Specificity of the Extended-spectrum SHV-2 Enzyme. *Journal of Molecular Biology.* 328, 289-301.
40. Yeates, T. O. (1997). Detecting and overcoming crystal twinning. *Methods Enzymol.* 276, 344-358.
41. Vagin, A. & Teplyakov, A. (1997). MOLREP: an automated program for molecular replacement. *J. Appl. Cryst.* 30, 1022-1025.
42. Schüttelkopf, A. W. & van Aalten, D. M. F. (2004). PRODRG: a tool for high throughput crystallography of protein-ligand complexes. *Acta Crystallogr. D* 60, 1355–1363.
43. Perrakis, A., Morris, R. M. & Lamzin, V. S. (1999). Automated protein model building combined with iterative structure refinement. *Nature Struct. Biol.* 6, 458–463.
44. Padayatti, P.S., Sheri, A., Totir, M.A., Helfand, M.S., Carey, M.P., Anderson, V.E., Carey, P.R., Bethel, C.R., Bonomo, R.A., Buynak, J.D., van den Akker, F.
Rational design of a beta-lactamase inhibitor achieved via stabilization of the trans-enamine intermediate: 1.28 Å crystal structure of wt SHV-1 complex with a penam sulfone. *J.Am.Chem.Soc.*128: 13235-13242 (2006), pdb-code 2H5S

45. Ke, W., Sampson, J.M., Ori, C., Prati, F., Drawz, S.M., Bethel, C.R., Bonomo, R.A., van den Akker, F. (2011) Novel insights into the mode of inhibition of class A SHV-1 beta-lactamases revealed by boronic acid transition state inhibitors. *Antimicrob. Agents Chemother.* 55: 174-183 pdb-code 3MKE
46. Reynolds KA1, Thomson JM, Corbett KD, Bethel CR, Berger JM, Kirsch JF, Bonomo RA, Handel TM., Structural and computational characterization of the SHV-1 beta-lactamase-beta-lactamase inhibitor protein interface. *J Biol Chem.* 2006 Sep 8;281(36):26745-53.
47. Ke, W., Bethel, C.R., Papp-Wallace, K.M., Pagadala, S.R., Nottingham, M., Fernandez, D., Buynak, J.D., Bonomo, R.A., van den Akker, F. (2012). Crystal structures of KPC-2 β -lactamase in complex with 3-nitrophenyl boronic acid and the penam sulfone PSR-3-226. *Antimicrob. Agents Chemother.* 56: 2713-2718.
48. Petrella, S., Ziental-Gelus, N., Mayer, C., Renard, M., Jarlier, V., Sougakoff, W. (2008) Genetic and structural insights into the dissemination potential of the extremely broad-spectrum class A beta-lactamase KPC-2 identified in an *Escherichia coli* strain and an *Enterobacter cloacae* strain isolated from the same patient in France. *Antimicrob. Agents Chemother.* 52: 3725-3736
49. Lahiri SD, Mangani S, Durand-Reville T, Benvenuti M, De Luca F, Sanyal G, Docquier JD. (2013). Structural insight into potent broad-spectrum inhibition with reversible recyclization mechanism: avibactam in complex with CTX-M-15 and *Pseudomonas aeruginosa* AmpC β -lactamases. *Antimicrob Agents Chemother* 57:2496-2505.
50. King, D.T., King, A.M., Lal, S.M., Wright, G.D., Strynadka,

N.C.J. (2015) Molecular Mechanism of Avibactam Mediated Beta-Lactamase Inhibition ACS.Infect.Dis, in press

6 References

- [1] U. Römling, M. Gomelsky, and M. Y. Galperin, “C-di-GMP: The dawning of a novel bacterial signalling system,” *Molecular Microbiology*, vol. 57, pp. 629–639, 2005.
- [2] R. D. Monds, P. D. Newell, R. H. Gross, and G. A. O’Toole, “Phosphate-dependent modulation of c-di-GMP levels regulates *Pseudomonas fluorescens* Pf0-1 biofilm formation by controlling secretion of the adhesin LapA,” *Mol. Microbiol.*, vol. 63, pp. 656–679, 2007.
- [3] P. D. Newell, C. D. Boyd, H. Sondermann, and G. A. O’Toole, “A c-di-GMP effector system controls cell adhesion by inside-out signaling and surface protein cleavage,” *PLoS Biol.*, vol. 9, 2011.
- [4] O. Kirillina, J. D. Fetherston, A. G. Bobrov, J. Abney, and R. D. Perry, “HmsP, a putative phosphodiesterase, and HmsT, a putative diguanylate cyclase, control Hms-dependent biofilm formation in *Yersinia pestis*,” *Mol. Microbiol.*, vol. 54, pp. 75–88, 2004.
- [5] M. Merighi, V. T. Lee, M. Hyodo, Y. Hayakawa, and S. Lory, “The second messenger bis-(3’-5’)-cyclic-GMP and its PilZ domain-containing receptor Alg44 are required for alginate biosynthesis in *Pseudomonas aeruginosa*,” *Mol. Microbiol.*, vol. 65, pp. 876–895, 2007.
- [6] V. T. Lee, J. M. Matewish, J. L. Kessler, M. Hyodo, Y. Hayakawa, and S. Lory, “A cyclic-di-GMP receptor required for bacterial exopolysaccharide production,” *Mol. Microbiol.*, vol. 65, pp. 1474–1484, 2007.
- [7] D. A. D’Argenio, M. W. Calfee, P. B. Rainey, and E. C. Pesci, “Autolysis and autoaggregation in *Pseudomonas aeruginosa* colony morphology mutants,” *J. Bacteriol.*, vol. 184, pp. 6481–6489, 2002.
- [8] R. Simm, M. Morr, A. Kader, M. Nimtz, and U. Römling, “GGDEF and EAL domains inversely regulate cyclic di-GMP levels and transition from sessility to motility,” *Mol. Microbiol.*, vol. 53, pp. 1123–1134, 2004.
- [9] A. Nakhamchik, C. Wilde, and D. A. Rowe-Magnus, “Cyclic-di-GMP regulates extracellular polysaccharide production, biofilm formation, and rugose colony development by *Vibrio vulnificus*,” *Appl. Environ. Microbiol.*, vol. 74, pp. 4199–4209, 2008.

- [10] H. Kulasakara, V. Lee, A. Brencic, N. Liberati, J. Urbach, S. Miyata, D. G. Lee, A. N. Neely, M. Hyodo, Y. Hayakawa, F. M. Ausubel, and S. Lory, "Analysis of *Pseudomonas aeruginosa* diguanylate cyclases and phosphodiesterases reveals a role for bis-(3'-5')-cyclic-GMP in virulence.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 103, pp. 2839–2844, 2006.
- [11] A. G. Bobrov, O. Kirillina, D. A. Ryjenkov, C. M. Waters, P. A. Price, J. D. Fetherston, D. Mack, W. E. Goldman, M. Gomelsky, and R. D. Perry, "Systematic analysis of cyclic di-GMP signalling enzymes and their role in biofilm formation and virulence in *Yersinia pestis*," *Mol. Microbiol.*, vol. 79, pp. 533–551, 2011.
- [12] J. G. Malone, T. Jaeger, C. Spangler, D. Ritz, A. Spang, C. Arrieumerlou, V. Kaever, R. Landmann, and U. Jenal, "YfiBNR mediates cyclic di-GMP dependent small colony variant formation and persistence in *Pseudomonas aeruginosa*," *PLoS Pathog.*, vol. 6, 2010.
- [13] S. Haussler, B. Tummeler, H. Weissbrodt, M. Rohde, and I. Steinmetz, "Small-colony variants of *Pseudomonas aeruginosa* in cystic fibrosis," *Clin Infect Dis*, vol. 29, pp. 621–625, 1999.
- [14] J. G. Malone, T. Jaeger, P. Manfredi, A. Dötsch, A. Blanka, R. Bos, G. R. Cornelis, S. Häussler, and U. Jenal, "The YfiBNR signal transduction mechanism reveals novel targets for the evolution of persistent *Pseudomonas aeruginosa* in cystic fibrosis airways," *PLoS Pathog.*, vol. 8, 2012.
- [15] J. G. Leid, C. J. Willson, M. E. Shirtliff, D. J. Hassett, M. R. Parsek, and A. K. Jeffers, "The exopolysaccharide alginate protects *Pseudomonas aeruginosa* biofilm bacteria from IFN-gamma-mediated macrophage killing.," *J. Immunol.*, vol. 175, pp. 7512–7518, 2005.
- [16] A. L. Spoering and K. Lewis, "Biofilms and planktonic cells of *Pseudomonas aeruginosa* have similar resistance to killing by antimicrobials," *J. Bacteriol.*, vol. 183, pp. 6746–6751, 2001.
- [17] S. Häussler, C. Lehmann, C. Breselge, M. Rohde, M. Classen, B. Tummeler, P. Vandamme, and I. Steinmetz, "Fatal outcome of lung transplantation in cystic fibrosis patients due to small-colony variants of the *Burkholderia cepacia* complex.," *Eur. J. Clin. Microbiol. Infect. Dis.*, vol. 22, pp. 249–253, 2003.
- [18] M. J. Blommers, C. A. Haasnoot, J. A. Walters, G. A. van der Marel,

- J. H. van Boom, and C. W. Hilbers, "Solution structure of the 3'-5' cyclic dinucleotide d(pApA). A combined NMR, UV melting, and molecular mechanics study.," *Biochemistry*, vol. 27, pp. 8361–8369, 1988.
- [19] C. A. Frederick, M. Coll, G. A. van der Marel, J. H. van Boom, and A. H. Wang, "Molecular structure of cyclic deoxydiadenylic acid at atomic resolution.," *Biochemistry*, vol. 27, pp. 8350–8361, 1988.
- [20] M. Egli, R. V Gessner, L. D. Williams, G. J. Quigley, G. A. van der Marel, J. H. van Boom, A. Rich, and C. A. Frederick, "Atomic-resolution structure of the cellulose synthase regulator cyclic diguanylic acid.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 87, pp. 3235–3239, 1990.
- [21] Y. Guan, Y. G. Gao, Y. C. Liaw, H. Robinson, and A. H. J. Wang, "Molecular-Structure of Cyclic Diguanylic Acid at 1 Angstrom Resolution of 2 Crystal Forms - Self-Association, Interactions with Metal Ion/Planar Dyes and Modeling Studies," *J. Biomol. Struct. Dyn.*, vol. 11, pp. 253–276, 1993.
- [22] Y. C. Liaw, Y. G. Gao, H. Robinson, G. M. Sheldrick, L. A. Sliedregt, G. A. van der Marel, J. H. van Boom, and A. H. Wang, "Cyclic diguanylic acid behaves as a host molecule for planar intercalators.," *FEBS Lett.*, vol. 264, pp. 223–227, 1990.
- [23] C. Chan, R. Paul, D. Samoray, N. C. Amiot, B. Giese, U. Jenal, and T. Schirmer, "Structural basis of activity and allosteric control of diguanylate cyclase.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 101, pp. 17084–17089, 2004.
- [24] P. Wassmann, C. Chan, R. Paul, A. Beck, H. Heerklotz, U. Jenal, and T. Schirmer, "Structure of BeF₃-Modified Response Regulator PleD: Implications for Diguanylate Cyclase Activation, Catalysis, and Feedback Inhibition," *Structure*, vol. 15, pp. 915–927, 2007.
- [25] N. De, M. Pirruccello, P. V. Krasteva, N. Bae, R. V. Raghavan, and H. Sondermann, "Phosphorylation-independent regulation of the diguanylate cyclase WspR," *PLoS Biol.*, vol. 6, pp. 0601–0617, 2008.
- [26] N. De, M. V. A. S. Navarro, R. V. Raghavan, and H. Sondermann, "Determinants for the Activation and Autoinhibition of the Diguanylate Cyclase Response Regulator WspR," *J. Mol. Biol.*, vol. 393, pp. 619–633, 2009.
- [27] D. A. Ryjenkov, R. Simm, U. Römling, and M. Gomelsky, "The PilZ

- domain is a receptor for the second messenger c-di-GMP: The PilZ domain protein YcgR controls motility in enterobacteria,” *J. Biol. Chem.*, vol. 281, pp. 30310–30314, 2006.
- [28] J. Ko, K. S. Ryu, H. Kim, J. S. Shin, J. O. Lee, C. Cheong, and B. S. Choi, “Structure of PP4397 reveals the molecular basis for different c-di-GMP binding modes by pilz domain proteins,” *J. Mol. Biol.*, vol. 398, pp. 97–110, 2010.
- [29] J. Habazettl, M. G. Allan, U. Jenal, and S. Grzesiek, “Solution structure of the PilZ domain protein PA4608 complex with cyclic di-GMP identifies charge clustering as molecular readout,” *J. Biol. Chem.*, vol. 286, pp. 14304–14314, 2011.
- [30] J. S. Shin, K. S. Ryu, J. Ko, A. Lee, and B. S. Choi, “Structural characterization reveals that a PilZ domain protein undergoes substantial conformational change upon binding to cyclic dimeric guanosine monophosphate,” *Protein Sci.*, vol. 20, pp. 270–277, 2011.
- [31] P. V Krasteva, J. C. N. Fong, N. J. Shikuma, S. Beyhan, M. V. A. S. Navarro, F. H. Yildiz, and H. Sondermann, “*Vibrio cholerae* VpsT regulates matrix production and motility by directly sensing cyclic di-GMP.,” *Science*, vol. 327, pp. 866–868, 2010.
- [32] K. D. Smith, S. V Lipchock, T. D. Ames, J. Wang, R. R. Breaker, and S. A. Strobel, “Structural basis of ligand binding by a c-di-GMP riboswitch.,” *Nat. Struct. Mol. Biol.*, vol. 16, pp. 1218–1223, 2009.
- [33] C.-Y. Yang, K.-H. Chin, M. L.-C. Chuah, Z.-X. Liang, A. H.-J. Wang, and S.-H. Chou, “The structure and inhibition of a GGDEF diguanylate cyclase complexed with (c-di-GMP)₂ at the active site.,” *Acta Crystallogr. D. Biol. Crystallogr.*, vol. 67, no. Pt 12, pp. 997–1008, Dec. 2011.
- [34] G. Minasov, S. Padavattan, L. Shuvalova, J. S. Brunzelle, D. J. Miller, A. Baslé, C. Massa, F. R. Collart, T. Schirmer, and W. F. Anderson, “Crystal structures of Ykul and its complex with second messenger cyclic Di-GMP suggest catalytic mechanism of phosphodiester bond cleavage by EAL domains,” *J. Biol. Chem.*, vol. 284, pp. 13174–13184, 2009.
- [35] T. R. Barends, E. Hartmann, J. J. Griese, T. Beitlich, N. V Kirienko, D. A. Ryjenkov, J. Reinstein, R. L. Shoeman, M. Gomelsky, and I. Schlichting, “Structure and mechanism of a bacterial light-regulated cyclic nucleotide phosphodiesterase,” *Nature*, vol. 459, pp. 1015–1018,

- 2009.
- [36] A. Sundriyal, C. Massa, D. Samoray, F. Zehender, T. Sharpe, U. Jenal, and T. Schirmer, “Inherent regulation of EAL domain-catalyzed hydrolysis of second messenger cyclic di-GMP,” *J. Biol. Chem.*, vol. 289, pp. 6978–6990, 2014.
- [37] J. Benach, S. S. Swaminathan, R. Tamayo, S. K. Handelman, E. Folta-Stogniew, J. E. Ramos, F. Forouhar, H. Neely, J. Seetharaman, A. Camilli, and J. F. Hunt, “The structural basis of cyclic diguanylate signal transduction by PilZ domains.,” *EMBO J.*, vol. 26, pp. 5153–5166, 2007.
- [38] D. L. Burdette, K. M. Monroe, K. Sotelo-Troha, J. S. Iwig, B. Eckert, M. Hyodo, Y. Hayakawa, and R. E. Vance, “STING is a direct innate immune sensor of cyclic di-GMP,” *Nature*, vol. 478, pp. 515–518, 2011.
- [39] N. Tschowri, M. A. Schumacher, S. Schlimpert, N. B. Chinnam, K. C. Findlay, R. G. Brennan, and M. J. Buttner, “Tetrameric c-di-GMP mediates effective transcription factor dimerization to control *Streptomyces* development.,” *Cell*, vol. 158, no. 5, pp. 1136–1147, Aug. 2014.
- [40] Z. Zhang, B. L. Gaffney, and R. A. Jones, “c-di-GMP displays A monovalent metal ion-dependent polymorphism,” *J. Am. Chem. Soc.*, vol. 126, pp. 16700–16701, 2004.
- [41] Z. Zhang, S. Kim, B. L. Gaffney, and R. A. Jones, “Polymorphism of the signaling molecule c-di-GMP,” *J. Am. Chem. Soc.*, vol. 128, pp. 7015–7024, 2006.
- [42] M. Gentner, M. G. Allan, F. Zaehring, T. Schirmer, and S. Grzesiek, “Oligomer formation of the bacterial second messenger c-di-GMP: Reaction rates and equilibrium constants indicate a monomeric state at physiological concentrations,” *J. Am. Chem. Soc.*, vol. 134, pp. 1019–1029, 2012.
- [43] R. Paul, S. Weiser, N. C. Amiot, C. Chan, T. Schirmer, B. Giese, and U. Jenal, “Cell cycle-dependent dynamic localization of a bacterial response regulator with a novel di-guanylate cyclase output domain,” *Genes Dev.*, vol. 18, pp. 715–727, 2004.
- [44] J. G. Malone, R. Williams, M. Christen, U. Jenal, a. J. Spiers, and P. B. Rainey, “The structure-function relationship of WspR, a *Pseudomonas fluorescens* response regulator with a GGDEF output

- domain,” *Microbiology*, vol. 153, pp. 980–994, 2007.
- [45] J. W. Hickman, D. F. Tifrea, and C. S. Harwood, “A chemosensory system that regulates biofilm formation through modulation of cyclic diguanylate levels,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 102, pp. 14422–14427, 2005.
- [46] D. A. Ryjenkov, M. Tarutina, O. V. Moskvina, and M. Gomelsky, “Cyclic diguanylate is a ubiquitous signaling molecule in bacteria: Insights into biochemistry of the GGDEF protein domain,” *J. Bacteriol.*, vol. 187, pp. 1792–1798, 2005.
- [47] B. Christen, M. Christen, R. Paul, F. Schmid, M. Folcher, P. Jenoe, M. Meuwly, and U. Jenal, “Allosteric control of cyclic di-GMP signaling,” *J. Biol. Chem.*, vol. 281, pp. 32015–32024, 2006.
- [48] M. Christen, B. Christen, M. Folcher, A. Schauerte, and U. Jenal, “Identification and characterization of a cyclic di-GMP-specific phosphodiesterase and its allosteric control by GTP,” *J. Biol. Chem.*, vol. 280, pp. 30829–30837, 2005.
- [49] R. Tamayo, A. D. Tischler, and A. Camilli, “The EAL domain protein VieA is a cyclic diguanylate phosphodiesterase,” *J. Biol. Chem.*, vol. 280, pp. 33324–33330, 2005.
- [50] A. J. Schmidt, D. A. Ryjenkov, and M. Gomelsky, “The ubiquitous protein domain EAL is a cyclic diguanylate-specific phosphodiesterase: Enzymatically active and inactive EAL domains,” *J. Bacteriol.*, vol. 187, pp. 4774–4781, 2005.
- [51] F. Rao, Y. Yang, Y. Qi, and Z. X. Liang, “Catalytic mechanism of cyclic di-GMP-specific phosphodiesterase: A study of the EAL domain-containing RocR from *Pseudomonas aeruginosa*,” *J. Bacteriol.*, vol. 190, pp. 3622–3631, 2008.
- [52] M. Y. Galperin, D. A. Natale, L. Aravind, and E. V. Koonin, “A specialized version of the HD hydrolase domain implicated in signal transduction,” *Journal of molecular microbiology and biotechnology*, vol. 1, pp. 303–305, 1999.
- [53] A. S. N. Seshasayee, G. M. Fraser, and N. M. Luscombe, “Comparative genomics of cyclic-di-GMP signalling in bacteria: Post-translational regulation and catalytic activity,” *Nucleic Acids Res.*, vol. 38, pp. 5970–5981, 2010.
- [54] B. R. Boles and L. L. McCarter, “*Vibrio parahaemolyticus* scrABC, a

- novel operon affecting swarming and capsular polysaccharide regulation,” *J. Bacteriol.*, vol. 184, pp. 5946–5954, 2002.
- [55] M. Tarutina, D. A. Ryjenkov, and M. Gomelsky, “An unorthodox bacteriophytochrome from *Rhodobacter sphaeroides* involved in turnover of the second messenger c-di-GMP,” *J. Biol. Chem.*, vol. 281, pp. 34751–34758, 2006.
- [56] R. B. R. Ferreira, L. C. M. Antunes, E. P. Greenberg, and L. L. McCarter, “*Vibrio parahaemolyticus* ScrC modulates cyclic dimeric GMP regulation of gene expression relevant to growth on surfaces,” *J. Bacteriol.*, vol. 190, pp. 851–860, 2008.
- [57] H. Weber, C. Pesavento, A. Possling, G. Tischendorf, and R. Hengge, “Cyclic-di-GMP-mediated signalling within the sigma network of *Escherichia coli*,” *Mol. Microbiol.*, vol. 62, pp. 1014–1034, 2006.
- [58] R. Tamayo, S. Schild, J. T. Pratt, and A. Camilli, “Role of cyclic Di-GMP during El Tor biotype *Vibrio cholerae* infection: Characterization of the in vivo-induced cyclic Di-GMP phosphodiesterase CdpA,” *Infect. Immun.*, vol. 76, pp. 1617–1627, 2008.
- [59] B. I. Kazmierczak, M. B. Lebron, and T. S. Murray, “Analysis of FimX, a phosphodiesterase that governs twitching motility in *Pseudomonas aeruginosa*,” *Mol. Microbiol.*, vol. 60, pp. 1026–1043, 2006.
- [60] M. V. A. S. Navarro, N. De, N. Bae, Q. Wang, and H. Sondermann, “Structural Analysis of the GGDEF-EAL Domain-Containing c-di-GMP Receptor FimX,” *Structure*, vol. 17, pp. 1104–1116, 2009.
- [61] Y. Qi, M. L. C. Chuah, X. Dong, K. Xie, Z. Luo, K. Tang, and Z. X. Liang, “Binding of cyclic diguanylate in the non-catalytic EAL domain of FimX induces a long-range conformational change,” *J. Biol. Chem.*, vol. 286, pp. 2910–2917, 2011.
- [62] M. Y. Galperin, A. N. Nikolskaya, and E. V. Koonin, “Novel domains of the prokaryotic two-component signal transduction systems,” *FEMS Microbiology Letters*, vol. 203, pp. 11–21, 2001.
- [63] M. Y. Galperin, “Bacterial signal transduction network in a genomic perspective,” *Environmental Microbiology*, vol. 6, pp. 552–567, 2004.
- [64] J. Lee, D. R. Tomchick, C. A. Brautigam, M. Machius, R. Kort, K. J. Hellingwerf, and K. H. Gardner, “Changes at the KinA PAS-A

- dimerization interface influence histidine kinase function,” *Biochemistry*, vol. 47, pp. 4051–4064, 2008.
- [65] X. Ma, N. Sayed, P. Baskaran, A. Beuve, and F. Van Den Akker, “PAS-mediated dimerization of soluble guanylyl cyclase revealed by signal transduction histidine kinase domain crystal structure,” *J. Biol. Chem.*, vol. 283, pp. 1167–1178, 2008.
- [66] Y. Oka, T. Matsushita, N. Mochizuki, P. H. Quail, and A. Nagatani, “Mutant screen distinguishes between residues necessary for light-signal perception and signal transfer by phytochrome B,” *PLoS Genet.*, vol. 4, 2008.
- [67] B. L. Taylor and I. B. Zhulin, “PAS domains: internal sensors of oxygen, redox potential, and light.,” *Microbiol. Mol. Biol. Rev.*, vol. 63, pp. 479–506, 1999.
- [68] S. Reinelt, E. Hofmann, T. Gerharz, M. Bott, and D. R. Madden, “The structure of the periplasmic ligand-binding domain of the sensor kinase CitA reveals the first extracellular pas domain,” *J. Biol. Chem.*, vol. 278, pp. 39189–39196, 2003.
- [69] Y. Qi, F. Rao, Z. Luo, and Z. X. Liang, “A flavin cofactor-binding PAS domain regulates c-di-GMP synthesis in AxDGC2 from *Acetobacter xylinum*,” *Biochemistry*, vol. 48, pp. 10275–10285, 2009.
- [70] E. B. Purcell, C. A. McDonald, B. A. Palfey, and S. Crosson, “An analysis of the solution structure and signaling mechanism of LovK, a sensor histidine kinase integrating light and redox signals,” *Biochemistry*, vol. 49, pp. 6761–6770, 2010.
- [71] E. B. Purcell, D. Siegal-Gaskins, D. C. Rawling, A. Fiebig, and S. Crosson, “A photosensory two-component system regulates bacterial cell attachment.,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 104, pp. 18241–18246, 2007.
- [72] T. E. Swartz, T.-S. Tseng, M. A. Frederickson, G. Paris, D. J. Comerci, G. Rajashekara, J.-G. Kim, M. B. Mudgett, G. A. Splitter, R. A. Ugalde, F. A. Goldbaum, W. R. Briggs, and R. A. Bogomolni, “Blue-light-activated histidine kinases: two-component sensors in bacteria.,” *Science*, vol. 317, pp. 1090–1093, 2007.
- [73] J. R. Tuckerman, G. Gonzalez, E. H. S. Sousa, X. Wan, J. A. Saito, M. Alam, and M. A. Gilles-Gonzalez, “An oxygen-sensing diguanylate cyclase and phosphodiesterase couple for c-di-GMP control,” *Biochemistry*, vol. 48, pp. 9764–9774, 2009.

-
- [74] A. Möglich, R. A. Ayers, and K. Moffat, “Structure and Signaling Mechanism of Per-ARNT-Sim Domains,” *Structure*, vol. 17. pp. 1282–1294, 2009.
- [75] J. Cheung, C. A. Bingman, M. Reyngold, W. A. Hendrickson, and C. D. Waldburger, “Crystal structure of a functional dimer of the PhoQ sensor domain,” *J. Biol. Chem.*, vol. 283, pp. 13762–13770, 2008.
- [76] J. Cheung and W. A. Hendrickson, “Sensor domains of two-component regulatory systems,” *Current Opinion in Microbiology*, vol. 13. pp. 116–123, 2010.
- [77] J. Cheung and W. A. Hendrickson, “Crystal structures of C4-dicarboxylate ligand complexes with sensor domains of histidine kinases DcuS and DctB,” *J. Biol. Chem.*, vol. 283, pp. 30256–30265, 2008.
- [78] S. B. Williams and V. Stewart, “Functional similarities among two-component sensors and methyl-accepting chemotaxis proteins suggest a role for linker region amphipathic helices in transmembrane signal transduction,” *Mol. Microbiol.*, vol. 33, pp. 1093–1102, 1999.
- [79] M. V. A. S. Navarro, P. D. Newell, P. V. Krasteva, D. Chatterjee, D. R. Madden, G. A. O’Toole, and H. Sondermann, “Structural basis for c-di-GMP-mediated inside-out signaling controlling periplasmic proteolysis,” *PLoS Biol.*, vol. 9, 2011.
- [80] H. Park and M. Inouye, “Mutational analysis of the linker region of EnvZ, an osmosensor in *Escherichia coli*,” *J. Bacteriol.*, vol. 179, pp. 4382–4390, 1997.
- [81] M. Hulko, F. Berndt, M. Gruber, J. U. Linder, V. Truffault, A. Schultz, J. Martin, J. E. Schultz, A. N. Lupas, and M. Coles, “The HAMP Domain Structure Implies Helix Rotation in Transmembrane Signaling,” *Cell*, vol. 126, pp. 929–940, 2006.
- [82] M. V. Airola, K. J. Watts, A. M. Bilwes, and B. R. Crane, “Structure of Concatenated HAMP Domains Provides a Mechanism for Signal Transduction,” *Structure*, vol. 18, pp. 436–448, 2010.
- [83] R. Kishii, L. Falzon, T. Yoshida, H. Kobayashi, and M. Inouye, “Structural and functional studies of the HAMP domain of EnvZ, an osmosensing transmembrane histidine kinase in *Escherichia coli*,” *J. Biol. Chem.*, vol. 282, pp. 26401–26408, 2007.
- [84] K. E. Swain and J. J. Falke, “Structure of the conserved HAMP

- domain in an intact, membrane-bound chemoreceptor: A disulfide mapping study,” *Biochemistry*, vol. 46, pp. 13684–13695, 2007.
- [85] A. N. Lupas and M. Gruber, “The structure of alpha-helical coiled coils,” *Adv. Protein Chem.*, vol. 70, pp. 37–78, 2005.
- [86] H. U. Ferris, S. Dunin-Horkawicz, N. Hornig, M. Hulko, J. Martin, J. E. Schultz, K. Zeth, A. N. Lupas, and M. Coles, “Mechanism of regulation of receptor histidine kinases,” *Structure*, vol. 20, pp. 56–66, 2012.
- [87] H. U. Ferris, S. Dunin-Horkawicz, L. G. Mondéjar, M. Hulko, K. Hantke, J. Martin, J. E. Schultz, K. Zeth, A. N. Lupas, and M. Coles, “The mechanisms of HAMP-mediated signaling in transmembrane receptors,” *Structure*, vol. 19, pp. 378–385, 2011.
- [88] A. L. Chang, J. R. Tuckerman, G. Gonzalez, R. Mayer, H. Weinhouse, G. Volman, D. Amikam, M. Benziman, and M. A. Gilles-Gonzalez, “Phosphodiesterase A1, a regulator of cellulose synthesis in *Acetobacter xylinum*, is a heme-based sensor,” *Biochemistry*, vol. 40, pp. 3420–3426, 2001.
- [89] Z. Cao, E. Livoti, A. Losi, and W. Gärtner, “A blue light-inducible phosphodiesterase activity in the cyanobacterium *synechococcus elongatus*,” *Photochem. Photobiol.*, vol. 86, pp. 606–611, 2010.
- [90] R. Paul, S. Abel, P. Wassmann, A. Beck, H. Heerklotz, and U. Jenal, “Activation of the diguanylate cyclase PleD by phosphorylation-mediated dimerization,” *J. Biol. Chem.*, vol. 282, pp. 29170–29177, 2007.
- [91] T. A. Ramelot, A. Yee, J. R. Cort, A. Semesi, C. H. Arrowsmith, and M. A. Kennedy, “NMR structure and binding studies confirm that PA4608 from *Pseudomonas aeruginosa* is a PilZ domain and a c-di-GMP binding protein,” *Proteins Struct. Funct. Genet.*, vol. 66, pp. 266–271, 2007.
- [92] J. T. Pratt, R. Tamayo, A. D. Tischler, and A. Camilli, “PilZ domain proteins bind cyclic diguanylate and regulate diverse processes in *Vibrio cholerae*,” *J. Biol. Chem.*, vol. 282, pp. 12860–12870, 2007.
- [93] M. Christen, B. Christen, M. G. Allan, M. Folcher, P. Jenö, S. Grzesiek, and U. Jenal, “DgrA is a member of a new family of cyclic diguanosine monophosphate receptors and controls flagellar motor function in *Caulobacter crescentus*,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 104, pp. 4112–4117, 2007.

-
- [94] D. Amikam and M. Y. Galperin, "PilZ domain is part of the bacterial c-di-GMP binding protein," *Bioinformatics*, vol. 22, pp. 3–6, 2006.
- [95] Y. McCarthy, R. P. Ryan, K. O'Donovan, Y. Q. He, B. Le Jiang, J. X. Feng, J. L. Tang, and J. M. Dow, "The role of PilZ domain proteins in the virulence of *Xanthomonas campestris* pv. *campestris*," *Mol. Plant Pathol.*, vol. 9, pp. 819–824, 2008.
- [96] A. Duerig, S. Abel, M. Folcher, M. Nicollier, T. Schwede, N. Amiot, B. Giese, and U. Jenal, "Second messenger-mediated spatiotemporal control of protein degradation regulates bacterial cell cycle progression," *Genes Dev.*, vol. 23, pp. 93–104, 2009.
- [97] P. D. Newell, R. D. Monds, and G. A. O'Toole, "LapD is a bis-(3',5')-cyclic dimeric GMP-binding protein that regulates surface attachment by *Pseudomonas fluorescens* Pf0-1.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 106, pp. 3461–3466, 2009.
- [98] D. Chatterjee, R. B. Cooley, C. D. Boyd, R. A. Mehl, G. A. O'Toole, and H. Sondermann, "Mechanistic insight into the conserved allosteric regulation of periplasmic proteolysis by the signaling molecule cyclic-di-GMP.," *Elife*, vol. 3, p. e03650, 2014.
- [99] X. Fang, I. Ahmad, A. Blanka, M. Schottkowski, A. Cimmins, M. Y. Galperin, U. Römling, and M. Gomelsky, "GIL, a new c-di-GMP-binding protein domain involved in regulation of cellulose synthesis in enterobacteria," *Mol. Microbiol.*, vol. 93, pp. 439–452, 2014.
- [100] J. W. Hickman and C. S. Harwood, "Identification of FleQ from *Pseudomonas aeruginosa* as a c-di-GMP-responsive transcription factor," *Mol. Microbiol.*, vol. 69, pp. 376–389, 2008.
- [101] N. Sudarsan, E. R. Lee, Z. Weinberg, R. H. Moy, J. N. Kim, K. H. Link, and R. R. Breaker, "Riboswitches in eubacteria sense the second messenger cyclic di-GMP.," *Science*, vol. 321, pp. 411–413, 2008.
- [102] H. Xiao, T. E. Edwards, and A. R. Ferré-D'Amaré, "Structural Basis for Specific, High-Affinity Tetracycline Binding by an In Vitro Evolved Aptamer and Artificial Riboswitch," *Chem. Biol.*, vol. 15, pp. 1125–1137, 2008.
- [103] J. R. Govan and V. Deretic, "Microbial pathogenesis in cystic fibrosis: mucoid *Pseudomonas aeruginosa* and *Burkholderia cepacia*," *Microbiol. Rev.*, vol. 60, pp. 539–574, 1996.
- [104] E. E. Smith, D. G. Buckley, Z. Wu, C. Saenphimmachak, L. R.

- Hoffman, D. A. D'Argenio, S. I. Miller, B. W. Ramsey, D. P. Speert, S. M. Moskowitz, J. L. Burns, R. Kaul, and M. V. Olson, "Genetic adaptation by *Pseudomonas aeruginosa* to the airways of cystic fibrosis patients.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 103, pp. 8487–92, 2006.
- [105] H. K. Huse, T. Kwon, J. E. A. Zlosnik, D. P. Speert, E. M. Marcotte, and M. Whiteley, "Parallel evolution in *Pseudomonas aeruginosa* over 39,000 generations in vivo," *MBio*, vol. 1, 2010.
- [106] V. Burke, J. O. Robinson, C. J. Richardson, and C. S. Bundell, "Longitudinal studies of virulence factors of *Pseudomonas aeruginosa* in cystic fibrosis.," *Pathology*, vol. 23, pp. 145–148, 1991.
- [107] S. Häußler, "Biofilm formation by the small colony variant phenotype of *Pseudomonas aeruginosa*," *Environmental Microbiology*, vol. 6, pp. 546–551, 2004.
- [108] S. Häußler, "Highly adherent small-colony variants of *Pseudomonas aeruginosa* in cystic fibrosis lung infection," *J. Med. Microbiol.*, vol. 52, pp. 295–301, 2003.
- [109] M. Starkey, J. H. Hickman, L. Ma, N. Zhang, S. De Long, A. Hinz, S. Palacios, C. Manoil, M. J. Kirisits, T. D. Starner, D. J. Wozniak, C. S. Harwood, and M. R. Parsek, "*Pseudomonas aeruginosa* Rugose small-colony variants have adaptations that likely promote persistence in the cystic fibrosis lung," *J. Bacteriol.*, vol. 191, pp. 3492–3503, 2009.
- [110] A. Meissner, V. Wild, R. Simm, M. Rohde, C. Erck, F. Bredenbruch, M. Morr, U. Römling, and S. Häußler, "*Pseudomonas aeruginosa* cupA-encoded fimbriae expression is regulated by a GGDEF and EAL domain-dependent modulation of the intracellular level of cyclic diguanylate," *Environ. Microbiol.*, vol. 9, pp. 2475–2485, 2007.
- [111] F. Zähringer, E. Lacanna, U. Jenal, T. Schirmer, and A. Boehm, "Structure and signaling mechanism of a zinc-sensory diguanylate cyclase," *Structure*, vol. 21, pp. 1149–1157, 2013.
- [112] A. A. Baykov, O. A. Evtushenko, and S. M. Avaeva, "A malachite green procedure for orthophosphate determination and its use in alkaline phosphatase-based enzyme immunoassay.," *Anal. Biochem.*, vol. 171, pp. 266–270, 1988.
- [113] "WolframAlpha." [Online]. Available: <https://www.wolframalpha.com>.

-
- [114] L. Slabinski, L. Jaroszewski, L. Rychlewski, I. A. Wilson, S. A. Lesley, and A. Godzik, “XtalPred: A web server for prediction of protein crystallizability,” *Bioinformatics*, vol. 23, pp. 3403–3405, 2007.
- [115] “Uniprot database.” [Online]. Available: <http://www.uniprot.org>.
- [116] J. Söding, A. Biegert, and A. N. Lupas, “The HHpred interactive server for protein homology detection and structure prediction,” *Nucleic Acids Res.*, vol. 33, 2005.
- [117] U. Jenal, “Cyclic di-guanosine-monophosphate comes of age: A novel secondary messenger involved in modulating cell surface structures in bacteria?,” *Current Opinion in Microbiology*, vol. 7, pp. 185–191, 2004.
- [118] U. Jenal and J. Malone, “Mechanisms of cyclic-di-GMP signaling in bacteria.,” *Annu. Rev. Genet.*, vol. 40, pp. 385–407, 2006.
- [119] T. Schirmer and U. Jenal, “Structural and mechanistic determinants of c-di-GMP signalling.,” *Nat. Rev. Microbiol.*, vol. 7, pp. 724–735, 2009.
- [120] N. Ausmees, R. Mayer, H. Weinhouse, G. Volman, D. Amikam, M. Benziman, and M. Lindberg, “Genetic data indicate that proteins containing the GGDEF domain possess diguanylate cyclase activity,” *FEMS Microbiol. Lett.*, vol. 204, pp. 163–167, 2001.
- [121] A. D. Tischler and A. Camilli, “Cyclic diguanylate (c-di-GMP) regulates *Vibrio cholerae* biofilm formation,” *Mol. Microbiol.*, vol. 53, pp. 857–869, 2004.
- [122] and M. R. P. Mary Jo Kirisits, Lynne Prost, Melissa Starkey, “Characterization of Colony Morphology Variants Isolated from *Pseudomonas aeruginosa* Biofilms,” *Appl. Environ. Microbiol.*, vol. 71, no. 8, pp. 4809–4821.
- [123] R. D. Finn, A. Bateman, J. Clements, P. Coggill, R. Y. Eberhardt, S. R. Eddy, A. Heger, K. Hetherington, L. Holm, J. Mistry, E. L. L. Sonnhammer, J. Tate, and M. Punta, “Pfam: The protein families database,” *Nucleic Acids Research*, vol. 42, 2014.
- [124] A. G. W. Leslie and H. R. Powell, “Processing diffraction data with MOSFLM,” in *Evolving methods for macromolecular Crystallography*, 2007, pp. 41–51.
- [125] M. D. Winn, C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. W. Leslie, A.

- McCoy, S. J. McNicholas, G. N. Murshudov, N. S. Pannu, E. A. Potterton, H. R. Powell, R. J. Read, A. Vagin, and K. S. Wilson, "Overview of the CCP4 suite and current developments," *Acta Crystallographica Section D: Biological Crystallography*, vol. 67. pp. 235–242, 2011.
- [126] G. M. Sheldrick, "A short history of SHELX," *Acta Crystallographica Section A: Foundations of Crystallography*, vol. 64. pp. 112–122, 2007.
- [127] G. Langer, S. X. Cohen, V. S. Lamzin, and A. Perrakis, "Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7.," *Nat. Protoc.*, vol. 3, pp. 1171–1179, 2008.
- [128] P. D. Adams, P. V. Afonine, G. Bunkóczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L. W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger, and P. H. Zwart, "PHENIX: A comprehensive Python-based system for macromolecular structure solution," *Acta Crystallogr. Sect. D Biol. Crystallogr.*, vol. 66, pp. 213–221, 2010.
- [129] P. Emsley, B. Lohkamp, W. G. Scott, and K. Cowtan, "Features and development of Coot," *Acta Crystallogr. Sect. D Biol. Crystallogr.*, vol. 66, pp. 486–501, 2010.
- [130] V. B. Chen, W. B. Arendall, J. J. Headd, D. A. Keedy, R. M. Immormino, G. J. Kapral, L. W. Murray, J. S. Richardson, and D. C. Richardson, "MolProbity: All-atom structure validation for macromolecular crystallography," *Acta Crystallogr. Sect. D Biol. Crystallogr.*, vol. 66, pp. 12–21, 2010.
- [131] "DINO: Visualizing Structural Biology (2002) <http://www.dino3d.org>."
- [132] C. T. O. Benfield, D. S. Mansur, L. E. McCoy, B. J. Ferguson, M. W. Bahar, A. P. Oldring, J. M. Grimes, D. I. Stuart, S. C. Graham, and G. L. Smith, "Mapping the IkappaB kinase beta (IKKbeta)-binding interface of the B14 protein, a vaccinia virus inhibitor of IKKbeta-mediated activation of nuclear factor kappaB.," *J. Biol. Chem.*, vol. 286, pp. 20727–20735, 2011.
- [133] L. Y. Geer, M. Domrachev, D. J. Lipman, and S. H. Bryant, "CDART: Protein homology by domain architecture," *Genome Res.*, vol. 12, pp. 1619–1623, 2002.
- [134] M. Kearse, R. Moir, A. Wilson, S. Stones-Havas, M. Cheung, S.

- Sturrock, S. Buxton, A. Cooper, S. Markowitz, C. Duran, T. Thierer, B. Ashton, P. Meintjes, and A. Drummond, "Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data," *Bioinformatics*, vol. 28, pp. 1647–1649, 2012.
- [135] E. Krissinel and K. Henrick, "Inference of Macromolecular Assemblies from Crystalline State," *J. Mol. Biol.*, vol. 372, pp. 774–797, 2007.
- [136] Y. Ye and A. Godzik, "FATCAT: A web server for flexible structure comparison and structure similarity searching," *Nucleic Acids Res.*, vol. 32, 2004.
- [137] L. Holm and P. Rosenström, "Dali server: Conservation mapping in 3D," *Nucleic Acids Res.*, vol. 38, 2010.
- [138] A. G. Murzin, S. E. Brenner, T. Hubbard, and C. Chothia, "SCOP: A structural classification of proteins database for the investigation of sequences and structures," *Journal of Molecular Biology*, vol. 247. pp. 536–540, 1995.
- [139] F. Rao, Y. Qi, H. S. Chong, M. Kotaka, B. Li, J. Li, J. Lescar, K. Tang, and Z.-X. Liang, "The functional role of a conserved loop in EAL domain-based cyclic di-GMP-specific phosphodiesterase.," *J. Bacteriol.*, vol. 191, no. 15, pp. 4722–4731, Aug. 2009.
- [140] C. Cole, J. D. Barber, and G. J. Barton, "The Jpred 3 secondary structure prediction server.," *Nucleic Acids Res.*, vol. 36, 2008.
- [141] R. D. Finn, J. Mistry, B. Schuster-Böckler, S. Griffiths-Jones, V. Hollich, T. Lassmann, S. Moxon, M. Marshall, A. Khanna, R. Durbin, S. R. Eddy, E. L. L. Sonnhammer, and A. Bateman, "Pfam: clans, web tools and services.," *Nucleic Acids Res.*, vol. 34, pp. D247–D251, 2006.
- [142] M. Goujon, H. McWilliam, W. Li, F. Valentin, S. Squizzato, J. Paern, and R. Lopez, "A new bioinformatics analysis tools framework at EMBL-EBI," *Nucleic Acids Res.*, vol. 38, 2010.
- [143] F. Sievers, A. Wilm, D. Dineen, T. J. Gibson, K. Karplus, W. Li, R. Lopez, H. McWilliam, M. Remmert, J. Söding, J. D. Thompson, and D. G. Higgins, "Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega," *Molecular Systems Biology*, vol. 7. 2011.
- [144] X. Robert and P. Gouet, "Deciphering key features in protein

- structures with the new ENDscript server,” *Nucleic Acids Res.*, vol. 42, 2014.
- [145] D. W. A. Buchan, F. Minneci, T. C. O. Nugent, K. Bryson, and D. T. Jones, “Scalable web services for the PSIPRED Protein Analysis Workbench.,” *Nucleic Acids Res.*, vol. 41, 2013.
- [146] J. Lipfert, L. Columbus, V. B. Chu, S. A. Lesley, and S. Doniach, “Size and shape of detergent micelles determined by small-angle X-ray scattering,” *J. Phys. Chem. B*, vol. 111, pp. 12427–12438, 2007.
- [147] A. Boehm, S. Steiner, F. Zaehring, A. Casanova, F. Hamburger, D. Ritz, W. Keck, M. Ackermann, T. Schirmer, and U. Jenal, “Second messenger signalling governs *Escherichia coli* biofilm induction upon ribosomal stress,” *Mol. Microbiol.*, vol. 72, pp. 1500–1516, 2009.
- [148] A. Krogh, B. Larsson, G. von Heijne, and E. L. Sonnhammer, “Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes.,” *J. Mol. Biol.*, vol. 305, pp. 567–580, 2001.
- [149] A. Gattiker, E. Gasteiger, and A. Bairoch, “ScanProsite: a reference implementation of a PROSITE scanning tool.,” *Appl. Bioinformatics*, vol. 1, pp. 107–108, 2002.
- [150] M. Brune, J. L. Hunter, J. E. Corrie, and M. R. Webb, “Direct, real-time measurement of rapid inorganic phosphate release using a novel fluorescent probe and its application to actomyosin subfragment 1 ATPase.,” *Biochemistry*, vol. 33, pp. 8262–8271, 1994.
- [151] G. Giardina, A. Paiardini, S. Fericola, S. Franceschini, S. Rinaldo, V. Stelitano, and F. Cutruzzola, “Investigating the allosteric regulation of YfiN from *Pseudomonas aeruginosa*: Clues from the structure of the catalytic domain,” *PLoS One*, vol. 8, 2013.
- [152] K. Jonas, A. N. Edwards, R. Simm, T. Romeo, U. Römling, and Ö. Melefors, “The RNA binding protein CsrA controls cyclic di-GMP metabolism by directly regulating the expression of GGDEF proteins,” *Mol. Microbiol.*, vol. 70, pp. 236–257, 2008.
- [153] E. Brombacher, C. Dorel, A. J. B. Zehnder, and P. Landini, “The curli biosynthesis regulator CsgD co-ordinates the expression of both positive and negative determinants for biofilm formation in *Escherichia coli*,” *Microbiology*, vol. 149, pp. 2847–2857, 2003.
- [154] C. Pesavento, G. Becker, N. Sommerfeldt, A. Possling, N. Tschowri,

-
- A. Mehlis, and R. Hengge, “Inverse regulatory coordination of motility and curli-mediated adhesion in *Escherichia coli*,” *Genes Dev.*, vol. 22, pp. 2434–2446, 2008.
- [155] M. Gjermansen, M. Nilsson, L. Yang, and T. Tolker-Nielsen, “Characterization of starvation-induced dispersion in *Pseudomonas putida* biofilms: Genetic elements and molecular mechanisms,” *Mol. Microbiol.*, vol. 75, pp. 815–826, 2010.
- [156] J. S. Parkinson, “Signaling mechanisms of HAMP domains in chemoreceptors and sensor kinases,” *Annu. Rev. Microbiol.*, vol. 64, pp. 101–122, 2010.
- [157] D. Chatterjee, C. D. Boyd, G. A. O’Toole, and H. Sondermann, “Structural characterization of a conserved, calcium-dependent periplasmic protease from *Legionella pneumophila*,” *J. Bacteriol.*, vol. 194, pp. 4415–4425, 2012.
- [158] E. K. O’Shea, J. D. Klemm, P. S. Kim, and T. Alber, “X-ray structure of the GCN4 leucine zipper, a two-stranded, parallel coiled coil,” *Science*, vol. 254, pp. 539–544, 1991.
- [159] C. S. Bond, “TopDraw: A sketchpad for protein structure topology cartoons,” *Bioinformatics*, vol. 19, pp. 311–312, 2003.

7 Acknowledgments

First and foremost, I would like to express my special gratitude to my supervisor, Prof. Tilman Schirmer for giving me the opportunity and the freedom to work on such a demanding and at the same time exciting project. Under his supervision, I learnt how to work independently and think critically. He taught me to not jump to conclusions and to critically investigate the problems at hand - a skill, which is not only highly useful in science but in life overall.

Next, I would like to thank Urs Jenal and Timm Maier for all the scientifically stimulating discussions during joint seminars and committee meetings that helped to guide the project into the right directions. Furthermore, I would like to thank Timm Maier for expert crystallographic advice and for always having an open ear for me during several occasions regarding scientific but also personal matters.

A big **thanks, danke, merci, ribuan terimakasih, gracias, obrigada, shukria, hvala, Спасибо, धन्यवाद** goes out to Anna, Chee, Diana, David, Evy, Frédéric, Habib, Hugo, Leo, Marija and Marlise for always being there for me during these four years of PhD. You have become very dear to me and I would like to thank you for the various joyful events that we have done together, be it pool parties at our place, a weekend in Graubünden, the many crazy LoL nights, party nights, “beering” and so on. A special thanks thereby goes to Habib, for you have occupied a very special place in my

heart and I'd like to thank you for all the things that we have done together. You truly are the best friend I had at the Biozentrum!

Last but not least I would like to thank the whole Lim gang, Bref, Janne, Kai Marko, Lejko and Rod for all the good times we've spent together!

I am truly grateful to Bref, Chee and Habib for proofreading my thesis. Thanks guys!! Another thanks goes to the past and present members of the Schirmer lab Amit Sundriyal, Badri Nath Dubey, Claudia Massa, Dietrich Samoray, Franziska Zähringer, Frédéric Stanger, Ivan Plaza Menacho, Johanan Rueher, Nicole Schwob, Nikolaus Schmitz, Susanne Brükner, Tillmann Heinisch and Valentin Köhler for their helpfulness in the lab and for their funny discussions during the lunch breaks.

Liseli, jetz bisch o no ganz persönlich i mire Thesis verweigst ☺ Merci viu mou für di 4 Jahr, wo du immer für mi do bisch gsi, wenni mou wieder Sörgeli ha gha wägem PhD oder anderwiitig. Du hesch immer es offnigs Ohr für mi gha und hesch mi immer wieder uf dBei gschteht. Merci viu viu Mou für das!

En dehors du Bio, j'aimerais remercier les deux Malkis Karimba et Celita pour tous leurs efforts, surtout de ce dernier temps. Merci pour m'avoir préparé à manger, pour pleins de bisous, pour toujours m'écouter quand j'avais des doutes, pour m'avoir épaulée et soutenue... bref, merci pour toujours avoir été là pour moi. Je vous aime!! ☺ On a déjà passé une super belle année ensemble et la vie fait que commencer!

Acknowledgments

Karim, tu es l'homme de ma vie et je me réjouis tellement que tout sera bientôt terminé et après la vraie vie va enfin commencer.... "Oui, je le veux" ☺ Je t'aime mon coeur!

Célia, tu m'as apportée déjà beaucoup de joie dans si peu de temps! Je me réjouis de faire encore beaucoup de choses avec toi en faisant des swaps, profiter de l'été, faire du shopping..... ☺

I would also like to thank all my friends outside of the biozentrum, for keeping up with me and my busy schedule.... I know that I have not always been a very good friend in the past four years due to the limited time I had outside of the lab, this is why I'd like to apologize to you! You are the best friends, I've ever had and no one will ever replace you girls... So thank you Caro, Dagmar, Isä, Mesä, Mirä and Misch!

E Ehreplatz gilt mine Eltere: I möcht mi bi öich beidne vo Härze bedanke, für alles was dir I de letschte 27 Jahr für mi gmacht heit. Es isch so wunderschön zwüsse, dass me immer zwöi Mönsche het im Läbe, wo me sich cha druf verloh. Mir heis so guet zäme und I fröie mi, mit öich no oft schöni Stunde hie im Elsass verbringe zdürfe ☺ Ha öich gärn!

CURRICULUM VITAE

PERSONALIEN

Name	Malki	
Vorname	Stefanie	
Adresse	5 rue du Maréchal Juin F - 68300 St.Louis	
Telefon	0033 69 44 45 34	
Mobile	0041 79 422 75 75	
Email	stefanie.kauer@unibas.ch	
Geburtsdatum	13. Juli 1987	
Heimatort	Dürrenroth, Kanton Bern	
Zivilstand	verheiratet (lediger Name: Kauer) Mutter zweier Töchter, geb. 23.11.2015 und 15.02.2017	
Linkedin	https://ch.linkedin.com/in/stefanie-malki-90aa2b8a	

BERUFLICHE TÄTIGKEITEN

04/2016 – aktuell	Recruitment Consultant swisselect
05/2015- 03/2016	Administrative Assistant/Lab Technician von Prof. Dr. Roderick Lim Biozentrum, Universität Basel <ul style="list-style-type: none"> • Verantwortlich für die reibungslose Abwicklung administrativer und persönlichen Anliegen einer Gruppe von 14 Personen aus zehn Nationen. • Optimierung der Kommunikationswege sowie Innehaben einer Vermittlerrolle zwischen Vorgesetztem und Mitarbeitenden.
2011- 04/2015	PhD in Structural Biology and Biophysics Biozentrum, Universität Basel <ul style="list-style-type: none"> • Selbständiges Arbeiten an einem Projekt von Beginn bis Ende. Dabei lernte ich, mit Niederlagen umzugehen und sie konstruktiv zu nutzen, denn in der Forschung wird man zwangsläufig damit konfrontiert. Durch den Abschluss meines Doktorats

lernte ich, nach einem Fehlschlag niemals aufzugeben.

- Koordination einer Seminarreihe mit nationalen und internationalen Referenten. Ich war verantwortlich für die Kontaktaufnahme zwischen den Referenten, erstellte ein Programm für sie und organisierte deren Aufenthalt von Ankunft bis Abreise. Ich habe es dabei sehr genossen, mit unterschiedlichen Persönlichkeiten in Kontakt zu treten und empfand die verschiedenen Biografien als bereichernd und sehr interessant. Ich konnte mich problemlos der jeweiligen Situation anpassen. Die Liste meiner Gäste umfasste Postdocs, Professoren und sogar Vertreter des Adels.
- Teilnahme an diversen Konferenzen, um wissenschaftliche Resultate in einem Posterformat oder einem Talk zu präsentieren. Diese Erfahrungen haben mir zu einem selbstsicheren Auftreten und einem routinierten, kompetenten Umgang mit Präsentationsaufgaben verholfen.
- Coaching von mehreren Studentengruppen bezüglich Team-work, sorgfältiger Datenanalyse sowie Datenpräsentation.

SCHULEN UND AUSBILDUNG

2011-April 2015

PhD in Structural Biology and Biophysics

Biozentrum, Universität Basel

Titel der Arbeit: „Structural and functional investigations of the Pseudomonas aeruginosa YfiBNR-system“ (Note 5.24, cum laude)

2009-2011

MSc Biology

Universität Bern

Master thesis in der Gruppe von PD Dr. Achim Stocker (Note 6, summa cum laude)

2006 – 2009

BSc Biology

Universität Bern

Bachelor of Science in Cell Biology (grade 4.73)

Bachelor thesis in der Gruppe von Prof. Dr. Ulrich Baumann

(Note 6, summa cum laude)

2003 – 2006

Gymnasium Oberaargau

Langenthal

Maturaabschluss Juni 2006 (Schwerpunktfach Englisch)

SPRACHEN

DEUTSCH

Erstsprache

ENGLISCH

ausgezeichnet in Wort und Schrift (Kompetenzstufe C2)

FRANZÖSISCH

sehr gut in Wort und Schrift (Kompetenzstufe C2)

REFERENZEN

Referenzpersonen gebe ich Ihnen gerne anlässlich eines Vorstellungsgespräches bekannt.

Über meine Person

SELBSTKOMPETENZ

- Wenn ich ein Ziel vor Augen habe, verfolge ich dieses engagiert mit grossem Durchsetzungsvermögen und gebe nicht auf. Diese Kompetenz habe ich mir während meines Studiums im Bereich der Forschung angeeignet: Rückschläge gehören zur Tagesordnung; ich lernte, sie wegzustecken und die Motivation aufrecht zu erhalten. Meine Frustrationstoleranz ist dadurch gestiegen.
- Ich kann sehr effizient und ausdauernd an einer Sache dran bleiben. So bereitet es mir keine Mühe, wenn ich zwischendurch auch einmal am Wochenende oder am Abend länger arbeiten muss.
- Ich habe eine sehr schnelle Auffassungsgabe. In einer Diskussion nehme ich die verschiedenen Meinungen und Aussagen umfassend auf, kann sie kombinieren, zusammenfassen und innert Kürze auswerten.
- Ich kann gut organisieren. Ich habe während meines Doktorates die Aufenthalte diverser nationaler und internationaler Sprecher organisiert. Dazu habe ich diese Personen auch persönlich während ihres Aufenthaltes betreut.

SOZIALKOMPETENZ

- Ich bin eine ausgesprochen fröhliche Person: Grossen Wert lege ich auf einen respektvollen und wertschätzenden Umgang mit meinen Mitmenschen. Meine Umgangsformen sind tadellos.
- Ich kann gut auf unterschiedlichste Menschen zugehen, mich auf sie einlassen und adäquat kommunizieren/agieren.
- Ich arbeite sehr gerne in einem Team. Dies ist etwas, was mir in der Forschung fehlte. Gegenseitige Unterstützung, Austausch und eine offene Feedbackkultur sind mir wichtig.
- Es bereitet mir keine Mühe, mich in einer bestehenden Gruppe einzufügen: Ich kann mich, wenn nötig, unterordnen, aber auch eine Leitungsposition einnehmen, um gemeinsam und koordiniert mit meiner Gruppe ein Ziel zu erreichen.

FACHKOMPETENZ

- Ich spreche und schreibe fließend Deutsch, Englisch sowie Französisch. Dabei lege ich grossen Wert auf Schreibrichtigkeit.
-

METHODENKOMPETENZ

- Ich besitze eine schnelle Auffassungsgabe. Ich schrecke nicht davor zurück, neue Hürden und Herausforderungen in Angriff zu nehmen; ich meistere diese grundsätzlich auf eine effiziente Art und Weise.
- Durch meinen wissenschaftlichen Hintergrund verfüge ich über ein ausgeprägtes analytisches Denken.