Elucidation of the regulatory mechanisms of the diguanylate cyclases PleD, DgcA and DgcB by structural and biophysical analysis

Inauguraldissertation

zur
Erlangung der Würde eines Doktors der Philosophie
vorgelegt der
Philosophisch-Naturwissenschaftlichen Fakultät
der Universität Basel

von

Paul Wassmann
aus Deutschland

Basel, 2009

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät auf Antrag von

Prof. Dr. Tilman Schirmer
Prof. Dr. Urs Jenal

Basel, den 23.06.2009

Prof. Dr. Eberhard Parlow
Dekan

CONTENTS

Abbreviations

aa	amino acid
AC	adenylate cyclase
ASA	accessible surface area
AUC	analytical ultracentrifugation
c-di-GMP	bis-(3'-5')-cyclic guanosine monophosphate
CD	circular dichroism
DGC	diguanylate cyclase
ε	molar extinction coefficient
GMPCPP	5'-guanylylmethylenebisphosphonate
GTP	guanosine triphosphate
GTPαS	guanosine 5'-O-(1-thiotriphosphate)
HK	histidine kinase
HPt	histidine-containing phosphotransfer protein
I_P	primary inhibition site
I_S	secondary inhibition site
ITC	isothermal titration calorimetry
K_D	dissociation constant
k_{cat}	turnover number, catalytic rate constant
min	minute
MR	molecular replacement
MW	molecular weight
NC	class III nucleotidyl cyclase
NCS	non-crystallographic symmetry
OD	optical density
PDB	protein data bank
PDE	phosphodiesterase
POL	RNA/DNA polymerase
RMSD	root mean square deviation
rpm	revolutions per minute
RU	response units
Rec	response regulator receiver domain

RR	response regulator
s	second
SE AUC	sedimentation equilibrium analytical ultracentrifugation
SEC-MALS	size exclusion chromatography coupled multi angle light scattering
SV AUC	sedimentation velocity analytical ultracentrifugation
SPR	surface plasmon resonance
TCS	two-component system
T_M	melting temperature
V_M	matthews coefficient

Abstract

The ubiquitous bacterial second messenger bis-(3'-5')-cycic di-guanosine monophosphate (c-di-GMP) turns out to be the key regulator of the antagonistic processes: motility of individual cells on the one hand and persistence of a bacterial population in biofilms on the other. The biosynthesis of c-di-GMP by consumption of two molecules GTP is performed by diguanylate cyclases (DGCs). DGCs consist of catalytic GGDEF domains in combination with a plethora of N-terminal, environment sensing regulatory domains.

The previously elucidated crystal structure of the DGC PleD from *C. crescentus* has shown that the catalytic GGDEF domain shares its fold with the well studied adenylate cyclases and DNA polymerases. The ability of the GGDEF domain to bind only one GTP molecule requires formation of a complete joint-active site formed by two GGDEF domains. Therefore, DGCs have to form dimers to be enzymatically active. DGCs widely exploit environment-sensing domains for the regulation of their dimerization. In case of PleD, response regulator receiver domains (Rec1 and Rec2) are used for this process.

Many DGCs are additionally regulated by allosteric product inhibition. The above mentioned crystal structure of PleD revealed binding of intercalated c-di-GMP dimers between a primary inhibition site (I-site), represented mainly by the conserved RxxD motif (GGDEF domain), and arginines of the secondary I-site (Rec2 domain). Two contradicting modes of action were proposed for this regulatory mechanism. (I) Inhibition by c-di-GMP binding to the RxxD motif inducing conformational changes in the active site. (II) Inhibition by c-di-GMP forming cross-links between the GGDEF and the Rec2 domains and preventing hereby the formation of the dimeric active site.

In this study structural, biophysical and biochemical analysis of several DGCs from *C. crescentus* was undertaken, to elucidate the details of the regulatory mechanisms of this class of enzymes.

Analysis of the so-called 'stand-alone' DGCs, which consist of 25-500 amino acids long segments in front of their GGDEF domains, has shown the inability of the GGDEF domains to form dimers autonomously. The 'stand-alone' DGCs utilize their N-terminal segments for dimerization.

To get insights in the environmental cues dependent regulatory mechanism of dimerization, the Rec domains bearing DGC PleD was crystallized in its pseudophosphorylated/activated form. The crystal structure of BeF_3^-•Mg^{2+} activated PleD is the first structure showing a full-length response regulator in its activated state. Comparison with the structure of non-activated PleD resulted in the elucidation of the molecular mechanisms of the dimerization process. Additionally, the formation of a two-fold symmetric, charged pocket at the $(Rec1\text{-}Rec2)_2$ stem interface was observed, which might represent the long-sought 'pole-localization'-signal for PleD.

Besides giving insights in the substrate binding mode of the DGCs, the obtained structures shed light on the catalytic mechanism of DGCs. In combination with biochemical data the structures verified the 'two-metal assisted' catalysis mechanism for the DGCs.

A new, c-di-GMP dimer dependent domain-cross-linking mode was revealed. It is generally applicable to DGCs, involving in the process merely the GGDEF domains. It turned out that the successful inhibition of the DGC PleD relies on the presence of primary- and secondary I-sites, whereas the initial binding of c-di-GMP depends solely on the primary I-site. PleD was shown to need R390 besides the RxxD motif (the whole I-site) to bind c-di-GMP.

Finally, cross-linking of proteins by c-di-GMP, intradimeric like in PleD and intermolecularly like in DgcB, was shown and might represent the main regulatory function of this second messenger.

I INTRODUCTION

Second messengers in bacteria

In order to be able to respond appropriately to environmental cues, many organisms utilize the principle of second messenger dependent signal amplification in their signal transduction pathways. The interaction of the second messengers with their effectors, which are either a part of the target molecules or affect these, is able to promote a multitude of cellular answers on the protein, the DNA or the RNA levels.

Similar to other kingdoms, Bacteria are known to utilize a multitude of nucleotide second messengers (1). Among them are the extensively studied cyclic AMP (cAMP) (2-4) and guanosine tetra/pentaphosphate ((p)ppGpp) (5,6), but also the recently discovered cyclic di-adenosine monophosphate (c-di-AMP) (7,8), the global role of which still has to be shown.

Another member of this group, the bis-(3'-5')-cyclic dimeric guanosine monophosphate (c-di-GMP), has only been recently recognized to play a crucial role in the multicellular behavior of bacteria (9-16).

The ubiquitous bacterial second messenger c-di-GMP

The first reports of the second messenger c-di-GMP date back to 1987, describing it as an allosteric activator of the cellulose synthase in *Gluconacetobacter xylinus* (17).

Structural and biophysical studies on c-di-GMP (18,19) have revealed the high morphological complexity of this second messenger, which forms concentration and divalent-ion dependent intercalated structures. The low cellular concentrations of c-di-GMP (20,21) in the low micro-molar range as well as its structures in complex with proteins (22-26) indicate that probably only the monomeric and the dimeric species are of physiological relevance.

The same scientific group that discovered c-di-GMP was also able to identify the genes, which encode proteins involved in biosynthesis and breakdown of this second messenger (27). The catalysis of c-di-GMP by diguanylate cyclases (DGCs) consuming two molecules of GTP was assigned to the GGDEF domain, whereas degradation is performed by phosphodiesterases (PDEs) containing either the EAL or the HD-GYP domain. All three domains are actually named according to the consensus sequences in their active sites.

The pivotal role of c-di-GMP as a global second messenger has not been realized until the onset of the genome sequencing times that showed broad phylogenetic distribution of DNA sequences encoding GGDEF, EAL and HD-GYP domain proteins (28). The complete absence of proteins bearing these domains in the kingdoms of Eukaryotes and Archaea and the ubiquitous presence in the eubacterial kingdom as well as the enormous magnitude of these proteins in almost each bacterial species woke a profound interest in c-di-GMP and in the related signal transduction pathway. Since then, giant scientific strides were made in recognition of the c-di-GMP's role in the regulation of cellular processes concerning the community behavior of bacteria.

The role of c-di-GMP in bacterial pathogenesis: regulation of motility, biofilm formation and virulence gene expression

Besides regulation of such processes as cell cycle control (29) or antibiotic production (30), c-di-GMP seems to control mainly the transition between the motile and the sessile 'lifestyles' in bacteria (9). Both bacterial states, the planctonic and the sessile (biofilm), are involved in the pathogenicity of these microorganisms. Motility is known to be a crucial factor in the early stage of the host infection (31,32). Biofilms, which are represented by complex microorganism communities embedded in an organic matrix, pose a high danger in the healthcare and the industrial sectors being highly resistant to antibiotics (33,34). The regulation of the cellular level of c-di-GMP by DGCs and PDEs enables the bacteria, harboring such pathogens as *Vibrio cholerae, Yersinia pestis, Salmonella* Typhimurium and *Pseudomonas aeruginosa,* to impede the inhibition of motility decreasing the concentration of c-di-GMP (35-37). On the other hand, raising the cellular concentration enables them to induce formation of the biofilm (38-40).

Additionally, c-di-GMP was shown to be able to manipulate the expression of virulence factors (41,42). In order to understand the involvement of c-di-GMP in these pathogenic processes and to use this knowledge in the clinical intervention against the presented bacterial traits, the complex interaction between c-di-GMP and proteins involved in its synthesis, breakdown and transmission (effectors) has to be elucidated.

Controlled biosynthesis and degradation of c-di-GMP

The identification of the enzymatic function of the GGDEF domain (previously named DUF1 = domain of unknown function) to produce one c-di-GMP and two pyrophosphate molecules using up two molecules GTP attracted the general scientific interest. The consensus sequence (GGDEF) was shown by mutagenesis to be a part of the active site (43,44).

The overall fold of the GGDEF domains was first predicted by bioinformatics (45) and later on shown by X-ray crystallography (22) to be shared by adenylate cyclases (ACs) and DNA/RNA polymerases (POLs). The structure of the DGC PleD from *C. crescentus* has additionally shown that the GGDEF domains are restricted in binding a single GTP molecule. Therefore, formation of DGC dimers sharing a common active site formed at the interface of two GGDEF domains is the inevitable imperative for the enzymatic activity of DGCs.

The detection of allosterically bound c-di-GMP dimers in the crystal structure of PleD and the subsequent biochemical analysis revealed the regulation of DGCs by non-competitive product inhibition (22). The overall function of this regulatory process might be relevant for the concentration limitation of the enzymatic product and prevention of GTP depletion. Two distinct theories were postulated explaining the inhibitory mechanism. The first, the so-called “inhibition by domain immobilization”, claims a c-di-GMP dependent cross-linking of the catalytic GGDEF domain with the regulatory domains of the protein (22). In the structure of PleD the c-di-GMP dimers are found in interaction with R359xxD362 + R390 (= primary I-site) of the GGDEF domain and R148 + R178 of the regulatory domains (= secondary I-site). Another theory proposes a classical allosteric mechanism, by which the binding of c-di-GMP to the highly conserved R359xxD362 (RxxD motif) results in conformational changes in the active site (46).

The EAL domain PDEs, which were shown to be active as monomers (24), linearize c-di-GMP to pGpG. Whether pGpG is acting as a second messenger is not known yet. The process of degradation was shown to be dependent on divalent cations. Mg^{2+} or Mn^{2+} are essential for the catalysis, whereas presence of Ca^{2+} or Zn^{2+} lead to inhibition (47-49).

The second protein family of c-di-GMP degrading PDEs are the metal-dependent HD-GYP domain proteins, which are not related to the EAL domain proteins. These

were shown to linearize c-di-GMP to pGpG and to be able to degrade the later ones to GMP molecules.

One interesting class of the c-di-GMP processing proteins are the so-called "tandem-proteins" composed of GGDEF and one of the PDE domains, EAL or HD-GYP. For the GGDEF-EAL domain protein PdeA from *C. crescentus* the PDE activity was shown, which can be modulated by GTP binding to the degenerated active site of the GGDEF domain (47). This speaks for GGDEF-EAL interdomain communication. In contrast to PdeA most "tandem-proteins" show no degeneration of the active sites, in neither of both domains. Insights in the regulatory mechanisms of these proteins will be the next step in the understanding of the c-di-GMP signal transduction pathway.

One of the biggest riddles posed by the c-di-GMP signaling is the enormous number of GGDEF-, EAL- and HD-GYP domain proteins encoded by the genomes of bacteria. A single species can have up to 100 of such proteins, e.g. *V. vulnificus* (50). Such a high diversity of proteins having the same tasks, namely biosynthesis or degradation of c-di-GMP, stands for a highly sensitive signaling pathway, but also for the requirement of strict regulation. Most of the DGCs and PDEs utilize regulatory input domains at their N-terminus to do so. The high variability of these input modules, e.g. transmembrane domains, haem- or flavin associated PAS domains (51), light sensitive BLUF domains (52,53), small molecule binding GAF domains (54), response regulator receiver domains (Rec) of the 'two-component' system (22,23) and so on, seems to involve c-di-GMP signaling in sensing of a high number of environmental and cellular traits (50).

Additionally, DGCs and PDEs that harbor Rec domains fusing the 'two-component' and the c-di-GMP signaling pathways indicate signal input in the c-di-GMP pools from 'quorum sensing' (1).

Downstream effectors of c-di-GMP

For a long time the molecular output activities of c-di-GMP have stayed hidden, but are now starting to emerge. One of the c-di-GMP binding effectors is known since the discovery of c-di-GMP, namely the cellulose synthase from *G. xylinus* (17). Inspired by this fact, computational analysis returned with a prediction of the so-called PilZ domain as a putative c-di-GMP effector (55). The genes encoding this domain are broadly distributed in bacteria. The domain is found in proteins in combination with several other domains as well as with GGDEF- and EAL domains. Binding of c-di-

GMP by PilZ domains was later on shown by biochemistry and structural biology (25,26,56).

Prediction of another c-di-GMP effector was made on the basis of GGDEF domains possessing a degenerated active site, but an intact I-site (14). One of such proteins, PopA from *C. crescentus*, has recently been shown to target proteins for degradation in presence of c-di-GMP (29). The I-site motif of the DGCs seems also to be used independently of the GGDEF domain context. PelD of *P.aeruginosa* is activated by c-di-GMP binding to a motif that is analog to the I-site motif of PleD (57).

The targets of c-di-GMP are not exclusively proteins. This was recently shown by the interaction of the second messenger with riboswitches (58), targeting the translation of downstream effectors.

Additionally, c-di-GMP might exert its second messenger effect by binding to or by forming protein complexes according to its cross-linking function in the DGC PleD (22). Further experiments have to show the biological relevance of this theory as well as to extend the characterization of the identified targets on molecular, biochemical and structural levels.

Aims of this work

The analysis of the DGCs shed light not only on the c-di-GMP biosynthesis, but also on its regulatory mechanisms as well as on c-di-GMP effectors (see above). On the other hand, several questions stayed unanswered.

The regulation of DGCs by their N-terminal input domains was predicted to depend on environmental cues and to result in dimerization of these DGCs. The best understood model DGC is PleD consisting of two N-terminal Rec domains, a phosphorylatable Rec1 and a non-phosphorylatable Rec2, as input domains. The phosphorylation of Rec1 by its cognate histidine-kinases PleC and DivJ results in dimerization of PleD (43,59,60) and subsequent sequestration to the stalked pole of the microorganism. Since PleD was crystallized in the non-phosphorylated state, it would be beneficial to gain the structure of the activated protein to understand the molecular mechanism of the dimerization process as well as the function of the non-phosphorylatable Rec2.

The structure of the DGC PleD was solved in complex with its product c-di-GMP bound not only to the allosteric sites, but also to the active site (22). The scientists proposed a catalytic mechanism, where E371 of the GGDEF sequence motif plays the

role of the general base in the process of the phosphodiester-bond formation. This mechanism is distinct from the one utilized by the structurally related ACs and POLs, which utilize the 'two-metal assisted' catalysis. To test the postulated mechanism as well as to understand how DGCs manage to perform both phosphodiester-bond-formations simultaneously to create the macrocycle of c-di-GMP, the DGC PleD was tried to be trapped crystallographically in the productive state.

The identification of the allosteric binding site of c-di-GMP in DGCs was a crucial step in the identification of c-di-GMP effectors (see above), but it is still unclear which of the two proposed allosteric mechanisms is utilized by the DGCs and what are the molecular and kinetic properties of this process.

Although the GGDEF domain proteins were shown to be active as dimers, it is not clear whether these domains are able to dimerize autonomously. Several DGCs are identified to consist of an unassigned N-terminal segment ranging from 20 to 500 amino acids besides the C-terminal GGDEF domains. Are such 'stand-alone' DGCs enzymatically active? And if yes, how do they dimerize and get input from their environment?

I have tried to answer these questions during my PhD work using structural, biophysical and biochemical techniques presenting the results in the following sections. Chapter II.1 and II.2 are dealing with the structures of activated PleD concentrating on the mechanisms of the Rec domain dependent dimerization of PleD, the DGC catalysis and the allosteric product inhibition. Chapter II.3 presents the analysis of the allosteric product inhibition mechanism by the means of biochemical and biophysical techniques. Chapter II.4 describes the characterization of the 'stand-alone' DGCs, DgcA and DgcB from *C. crescentus*.

II RESULTS

Structure
Article

Structure of BeF_3^--Modified Response Regulator PleD: Implications for Diguanylate Cyclase Activation, Catalysis, and Feedback Inhibition

Paul Wassmann,[1] Carmen Chan,[1,3] Ralf Paul,[2] Andreas Beck,[1] Heiko Heerklotz,[1,4] Urs Jenal,[2] and Tilman Schirmer[1,*]
[1] Core Program of Structural Biology and Biophysics
[2] Focal Area of Growth and Development
Biozentrum, University of Basel, Klingelbergstrasse 70, CH-4056 Basel, Switzerland
[3] Present address: Lonza Biopharmaceuticals, Lonza Ltd., CH-3930 Visp, Switzerland.
[4] Present address: Leslie Dan Faculty of Pharmacy, University of Toronto, 144 College Street, Toronto, Ontario M5S 3M2, Canada.
*Correspondence: tilman.schirmer@unibas.ch
DOI 10.1016/j.str.2007.06.016

SUMMARY

Cyclic di-guanosine monophosphate (c-di-GMP) is a ubiquitous bacterial second messenger involved in the regulation of cell surface-associated traits and persistence. We have determined the crystal structure of PleD from *Caulobacter crescentus*, a response regulator with a diguanylate cyclase (DGC) domain, in its activated form. The BeF_3^- modification of its receiver domain causes rearrangement with respect to an adaptor domain, which, in turn, promotes dimer formation, allowing for the efficient encounter of two symmetric catalytic domains. The substrate analog GTPαS and two putative cations are bound to the active sites in a manner similar to adenylate cyclases, suggesting an analogous two-metal catalytic mechanism. An allosteric c-di-GMP-binding mode that crosslinks DGC and an adaptor domain had been identified before. Here, a second mode is observed that crosslinks the DGC domains within a PleD dimer. Both modes cause noncompetitive product inhibition by domain immobilization.

INTRODUCTION

The central role of bis-(3′ → 5′)-cyclic di-guanosine monophosphate (c-di-GMP) as a signaling molecule has been realized only upon the recent recognition of the omnipresence of genes coding for diguanylate cyclase domains (DGC or GGDEF domains) in bacterial genomes. C-di-GMP regulates cell surface-associated traits and community behavior such as biofilm formation in most eubacteria (Jenal and Malone, 2006), and its relevance to the virulence of pathogenic bacteria has been demonstrated (Tischler and Camilli, 2004). In particular, the dinucleotide has been proposed to orchestrate the switch between acute and persistent phases of infection (Malone et al., 2007). C-di-GMP is synthesized out of two molecules of GTP and is degraded into the linear dinucleotide pGpG by the opposing activities of DGCs and c-di-GMP-specific phosphodiesterases. Both enzymes occur in combinations with various other, mostly sensory or regulatory, domains. It is believed that in this way environmental or internal stimuli can control the production of c-di-GMP, which, in turn, will affect downstream targets (Jenal and Malone, 2006). One of these c-di-GMP effector domains (PilZ) has recently been identified (Amikam and Galperin, 2006; Christen et al., 2007).

About 10% of the known DGCs are response regulators (RRs) as part of two-component systems (Jenal and Malone, 2006). RRs are activated by cognate histidine kinases that phosphorylate a conserved aspartate of the receiver domain (Rec) (Stock et al., 2000). Various Rec domains have been studied in great detail, and the structural changes upon activation have been described (for a review, see Robinson et al., 2000).

Due to the instability of the aspartyl phosphoanhydride, these studies have been performed with phosphoryl analogs. Notably, modification of CheY by beryllium fluoride (BeF_3^-) (Lee et al., 2001) resulted in structural changes fully equivalent to those obtained by phosphorylation (Birck et al., 1999; Lewis et al., 1999). Upon activation, a Thr/Ser side chain is pulled toward the modified aspartate, and a Phe/Tyr side chain follows to fill the gap and changes from a semiexposed to a buried position. Concomitantly, mainly the β4-α4 loop of the Rec domain is changing its conformation. Much less is known about how these conformational changes elicited in the Rec domain are signaled downstream either to target proteins or to the output domain of full-length RRs. The structures of full-length CheB (Djordjevic et al., 1998), NarL (Maris et al., 2002), PrrA (Nowak et al., 2006), and members of the OmpR/PhoB subfamily (DrrB, DrrD) (Robinson et al., 2003) have been determined, but only in the nonactivated state. Distinct mechanisms of activation have been proposed, such as relief of active site obstruction (e.g., CheB) (Djordjevic et al., 1998) (PrrA) (Nowak et al., 2006), dimer or oligomer formation (e.g., PhoB) (Bachhawat et al., 2005), or both (NarL) (Maris et al., 2002).

PleD from *Caulobacter crescentus* is an unorthodox RR consisting of a Rec domain (D1) with the phosphorylatable aspartate, a Rec-like adaptor domain (D2), and the enzymatic DGC domain, also called the GGDEF domain according to a conserved sequence motif. The protein is intimately involved in the transition of the *Caulobacter* cells from the motile to the sessile form (Aldridge et al., 2003). To gain insight into the molecular mechanisms of catalysis and regulation exerted by DGCs, we have previously determined the crystal structure of PleD (Chan et al., 2004) (Figures 2A and 2C). The fold of DGC (PFAM00990; http://www.sanger.ac.uk/Software/Pfam/) closely resembles that of class III nucleotidyl cyclases, which include bacterial and eukaryotic adenylyl cyclases (ACs) (Chan et al., 2004). A recent review with a detailed comparison of DGC and AC has been published recently (Sinha and Sprang, 2006).

Enzymatic studies showed that pseudo phosphorylation of PleD by BeF_3^- results in a 35-fold increase of specific activity (R.P. et al., unpublished data). The nonactivated structure of PleD (Chan et al., 2004) suggests that this activation may occur via dimerization of the D1/D2 stem domains, which brings two DGC domains into proximity as a prerequisite for the condensation reaction to occur ("activation by dimerization"). At the same time, the structure showed that the c-di-GMP product crosslinks the DGC with the adaptor domain, suggesting, together with biochemical data, that the product allosterically inhibits the enzyme by hampering the productive encounter of the two DGC domains of the dimer ("inhibition by immobilization"). Later, the central role in feedback inhibition of the allosteric-binding site on the DGC domain (with a characteristic RxxD sequence motif) was demonstrated by mutagenesis, genetic screens, and sequence comparison, and an alternative allosteric model for the inhibition mechanism was proposed (Christen et al., 2006).

Here, we present the crystal structure of PleD after modification of the active aspartate with the phosphoryl analog BeF_3^- and compare it with the nonactivated structure (Chan et al., 2004). The modification induces a change in the relative arrangement of the two Rec domains within the monomer, resulting in the stabilization of the dimer that is the catalytically competent form (R.P. et al., unpublished data). Details about the binding mode of the substrate analog GTPαS/Mg^{2+} are elucidated, and possible catalytic mechanisms are discussed. Unexpectedly, a new, to our knowledge, and possibly general mode of noncompetitive product inhibition for DGCs in which c-di-GMP immobilizes the two DGC domains of the PleD dimer in a nonproductive arrangement is revealed.

RESULTS

Dimerization in Solution

Size-exclusion chromatography gave evidence that nonactivated PleD partly dimerizes in a concentration-dependent manner at high protein concentrations (Figure 1A). Most relevant, dimerization was enhanced by the addition of BeF_3^- and divalent cations (Mg^{2+} or Mn^{2+}) (Figure 1B), and the manganese cation was shown to be more efficient. In parallel, BeF_3^--induced dimerization has also been shown via chemical crosslinking, and it was shown that the dimeric fraction entirely contains the catalytic activity (R.P. et al., unpublished data).

To thermodynamically quantify the dimerization affinity, isothermal titration calorimetry (ITC) was performed. Figure 1C shows the heat peaks measured after injections of predominantly dimerized PleD into matching buffer; in the beginning of the titration, the dilution of the titrant is largest and causes the dissociation of the dimers. Since dimerization is exothermic in this case, heat is consumed for dissociation after each injection. With increasing concentration in the cell, more and more dimers persist, and the heat of dissociation decreases gradually. Figure 1D illustrates two pairs of such data sets obtained by normalization and baseline (Q_{dil}) correction of ITC curves with activated and nonactivated protein, respectively, on a logarithmic scale. Q_{dil} ranged from ~1 kcal/mol for the nonactivated PleD mutant to −6 and −3 kcal/mol for the activated protein, and they include dilution effects and slight pH or temperature mismatches of titrant and cell content. At first glance, it is seen that the activated protein dimerizes at a much lower concentration. Upon activation with BeF_3^-, the fitted dissociation constant, K_D, of dimerization decreased from about 100 μM to < 10 μM. Hereby, the latter value represents an upper limit, because, for this kind of experiment, the instrument was close to the detection limit. The enthalpy of dimerization is −10 kcal/mol for the nonactivated protein and is apparently somewhat less exothermic for the activated protein, suggesting that activation eliminates an entropic hindrance of dimerization.

Structural Changes Induced by Activation

To obtain the structure of activated PleD, crystallization was attempted with BeF_3^--modified protein. Crystals were obtained upon addition of c-di-GMP, which had been shown earlier to rigidify the multidomain protein by crosslinking an allosteric site of the DGC domain with a neighboring domain (Chan et al., 2004) and the substrate analog GTPαS/Mg^{2+} (Table 1). The pseudo phosphorylation at Asp53 in domain D1 resulted in a train of structural changes (Figures 2 and 3), ultimately leading to a dramatic tightening of the dimer interface of the $(D1/D2)_2$ stem compared to the nonactivated structure (Figure 2). Furthermore, but probably not as a consequence of activation, the arrangement of the DGC domains is drastically different from the nonactivated structure.

Asp53 of Rec domain D1 appears fully modified (Figures 3A and 3B) and closely resembles phospho-aspartate (Lewis et al., 1999), with a Be-O distance of 1.58 Å (restrained in refinement to 1.55 Å) and a C_γ-O_δ-Be-F_1 dihedral *cis* conformation. The moiety forms four H bonds with the binding pocket and contributes to the coordination of the adjacent Mg^{2+} ion. The modification results in a restructuring of the β4-α4 loop, with Thr83 moving by more than 3 Å relative to the position in the nonactivated state to form an H bond with the BeF_3^- moiety (Figures 3A and 3B). The vacated space, in turn, is claimed by Phe102,

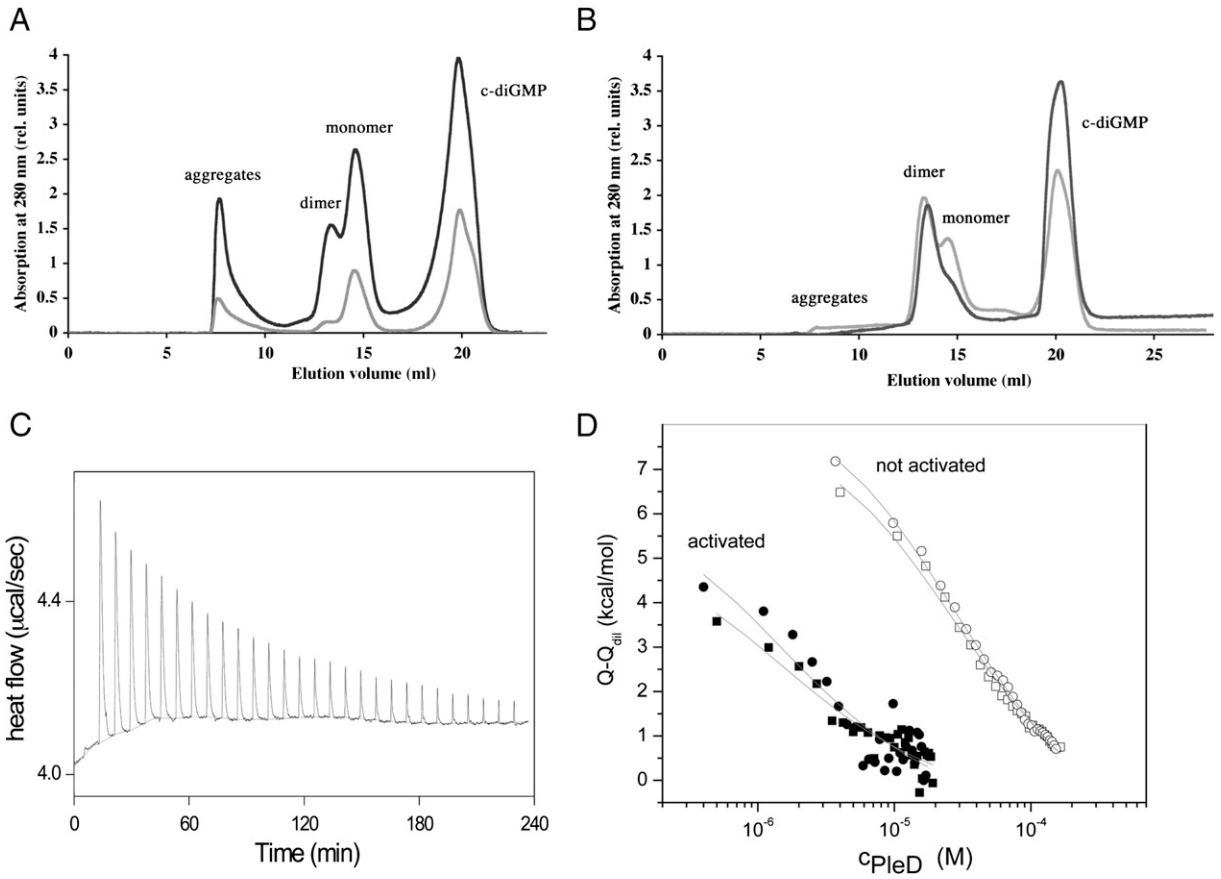


Figure 1. Size-Exclusion Chromatography Elution Profiles and Isothermal Titration Calorimetry of Purified PleD Mutant R313A

(A–D) This mutant with unmodified dimerization domains had been chosen for technical reasons (availability of material). (A) Elution profiles of nonactivated PleD at a protein concentration of 33 μM (gray) and 66 μM (black). (B) Elution profiles of BeF_3^--modified PleD (66 μM) in presence of 10 mM $MgCl_2$ (gray) and 1 mM $MnCl_2$ (black). Note that c-di-GMP is released and separated from the protein during the runs. (C) The primary ITC data for a dilution experiment (initial PleD concentration = 0.86 mM) with nonactivated PleD. (D) The integrated two ITC data sets for the activated (solid symbols) and nonactivated (open symbols) PleD mutant, after subtraction of the baseline heats, Q_{dil}, obtained by model-based extrapolation of the curves to high concentrations. The curves illustrate the best fits with shared K_D values of 100 μM (nonactivated) and 10 μM (activated). Due to the noisy data, the latter value represents an upper limit.

which changes its side chain conformation from gauche+ to *trans* to become buried; part of the β5 main chain (residues 102–104) moves up to 2 Å.

Domain D1 interacts closely and in a pseudo two-fold symmetric manner with the adaptor domain D2 across the α4-β5-α5 faces of both domains. Activation results in a domain rotation/shift of 14°/6 Å of D2 relative to D1 and a massive repacking of the domain interface (Figure 3A). In Figures 3C and 3D, the D1/D2 interface is displayed in its two states. While most of the contacting D1 residues stay in touch with the D2 domain, Arg91, Val110, and Ile251 contribute to the interface only in the nonactivated state, and residues Arg88, Ile92, Leu95, Val241, Gln259, and Ala263 contribute only in the activated state. Amazingly, the three interdomain salt bridges (Asp108-Arg237, Arg115-Asp250, Asp101-Arg264) are not disrupted during the transition, although the partners move up to 7 Å with respect to each other. However, upon activation, Asp257 swaps its ion-pair partner Arg91 for Arg88. Upon activation, the hydrophobic contact patch around Met111, Ala114, Met240, and Leu244 becomes repacked, and an additional apolar contact is formed that involves Ile92, Leu95, and Ala263 (Figures 3C and 3D).

On the quaternary level, in both crystal structures, the Rec and the adaptor domains form a dyad symmetric $(D1/D2)_2$ "stem" with the equivalent domains of a second monomer (Figure 2). Since both domains contribute to the interchain interface, it is obvious that the relative repositioning of the domains within each chain also has consequences on the quaternary structure. The weak dimer interface observed in the nonactivated structure (Chan et al., 2004) is greatly tightened upon activation (Figures 2C and 2D), and the buried accessible surface areas per monomer (ΔASA) increased from 900 $Å^2$ to 1436 $Å^2$.

In both states, the interchain contacts between D1 and D2 are formed in an isologous, i.e., two-fold symmetric manner. In the nonactivated structure, there is only a small contact patch around Tyr26 resulting in a discontinuous

Table 1. Data Collection and Refinement Statistics of Activated PleD in Complex with c-di-GMP and GTPαS

Data Collection	
Space group	$P2_12_12$
Cell dimensions	
a, b, c (Å)	128.9, 132.6, 88.4
Resolution (Å)	30–2.71 (2.85–2.71)
R_{merge} (%)	9.8 (43.6)
I/σ(I)	10.2 (1.9)
Completeness (%)	95.0 (78.0)
Redundancy	2.9 (2.4)
Refinement	
Number of reflections	39,043
R_{work}/R_{free}	21.7/25.4
Number of atoms	
Protein	7,012
Ligands	272
Water	14
B factors (Å^2)	
Protein	41.0
Ligands	43.7
Water	29.2
Rmsds	
Bond lengths (Å)	0.011
Bond angles (°)	1.5

Data in parentheses belong to the outer resolution shell.

interface (Figure 4A). While this contact is maintained in the activated structure, a multitude of additional interactions are formed involving the α1, β2, β3, α5 face of D1 and the α3, α3-β3, α5 face of D2 (Figure 4B). Three salt bridges are formed between the subunits, and there is a sulfate ion on the two-fold symmetry axis at the center of a basic cluster formed by Arg117 and Arg121 with their symmetry mates (Figure 4B). In addition, there are a few polar interactions and a small hydrophobic cluster formed by Leu124 and Val125 and their symmetry mates.

Upon activation, the α5 helices of the stem domains become extended by 4–5 residues at the C terminus. It is these helices that experience the largest change in relative interchain distance, moving from a closest distance of about 10 Å to direct van der Waals contact (Figures 4A and 4B). Together with their symmetry mates, the C-terminal thirds of the four α5 helices form a parallel α helix bundle in the dimer (Figure 2B).

Active Site

To gain further insight into the catalytic action of DGCs, the substrate analog GTPαS and Mg^{2+} had been added to the crystallization set-ups. In the activated structure, both active sites of the DGC domains, which face away from each other (see below), are occupied by the ligands. Figure 5 shows that the β- and γ-phosphates are tightly bound and form two H bonds to main chain amides of the short P loop between strand β1 and helix α1 and ionic interactions to Lys442, Arg446, and a putative Mg^{2+} ion (named metal B in analogy to the situation in adenylate cyclases [Tesmer et al., 1999]). The cation is additionally coordinated by main chain carbonyl 328 of the P loop, both carboxylate oxygens of Asp327, and Glu370, OE1. In one subunit of the dimer (subunit A), a second putative Mg^{2+} ion (metal A) is bound from another side to the same two side chains, interacting with atoms Asp327, OD2 and Glu370, OE2, respectively. Consistent with the structure, mutation of Asp327 to Ala resulted in a complete loss of enzymatic activity, as has been seen previously for mutant E370Q (Christen et al., 2006). The elongated density of the omit map (Figure 5) and the elevated B factor of the ion (52 Å^2) possibly indicate alternative positions for this ion. Neither the α-thiophosphate nor the ribose moiety are bound to the protein by specific interactions, resulting in comparatively high B values (~60 Å^2). The guanine moiety, in contrast, is well adopted, forming H bonds with Asn335 and Asp344 of the guanine-binding pocket, also already identified in the nonactivated protein-product complex (Chan et al., 2004).

Allosteric Product Binding

Compared to the previously determined nonactivated structure of PleD, the DGC domains adopt a drastically different position with respect to the stem domains. While, in the previous structure, the weak D2/DGC interface is strengthened by a bound c-di-GMP dimer that effectively crosslinks the domains (Figure 2A), in the present structure, this interface is completely disrupted and the DGC domains are swung out to form a two-fold symmetric c-di-GMP-crosslinked dimer (Figure 6). The difference in the relative orientation of the DGC domain with respect to the stem is probably not a direct consequence of Rec modification, since, assuming a flexible D2-DGC linker, a change of the crosslinking mode appears to be structurally feasible in either state. Thus, both domain organizations would occur in thermodynamic equilibrium in solution.

C-di-GMP forms dimers with the four guanyl bases stacked and intercalated as in small-molecule crystals (Egli et al., 1990; Liaw et al., 1990). Figure 6B shows that this form is found bound to the allosteric site. Two guanyl bases interact with DGC residues Arg359, Asp362, and Arg390 (primary I site, I_p), whereas the neighboring third base is bonded to Arg313 of the adjacent DGC domain (secondary I site, $I_{S,DGC}$). Due to symmetry, there are two isologous crosslinks within the DGC dimer (Figure 6). The dimer interface (with a ΔASA of 517 Å^2 that increases to 1044 Å^2 upon ligand binding) is exclusively hydrophobic and involves Ala360, Ile361, Pro377, and the hydrophobic part of the Ser309 side chain. Noteworthy, in the nonactivated structure, a c-di-GMP dimer is bound to I_p in exactly the same way, but it crosslinks with the third base and a phosphate moiety to Arg148 and Arg178 of the D2 domain ($I_{S,D2}$).

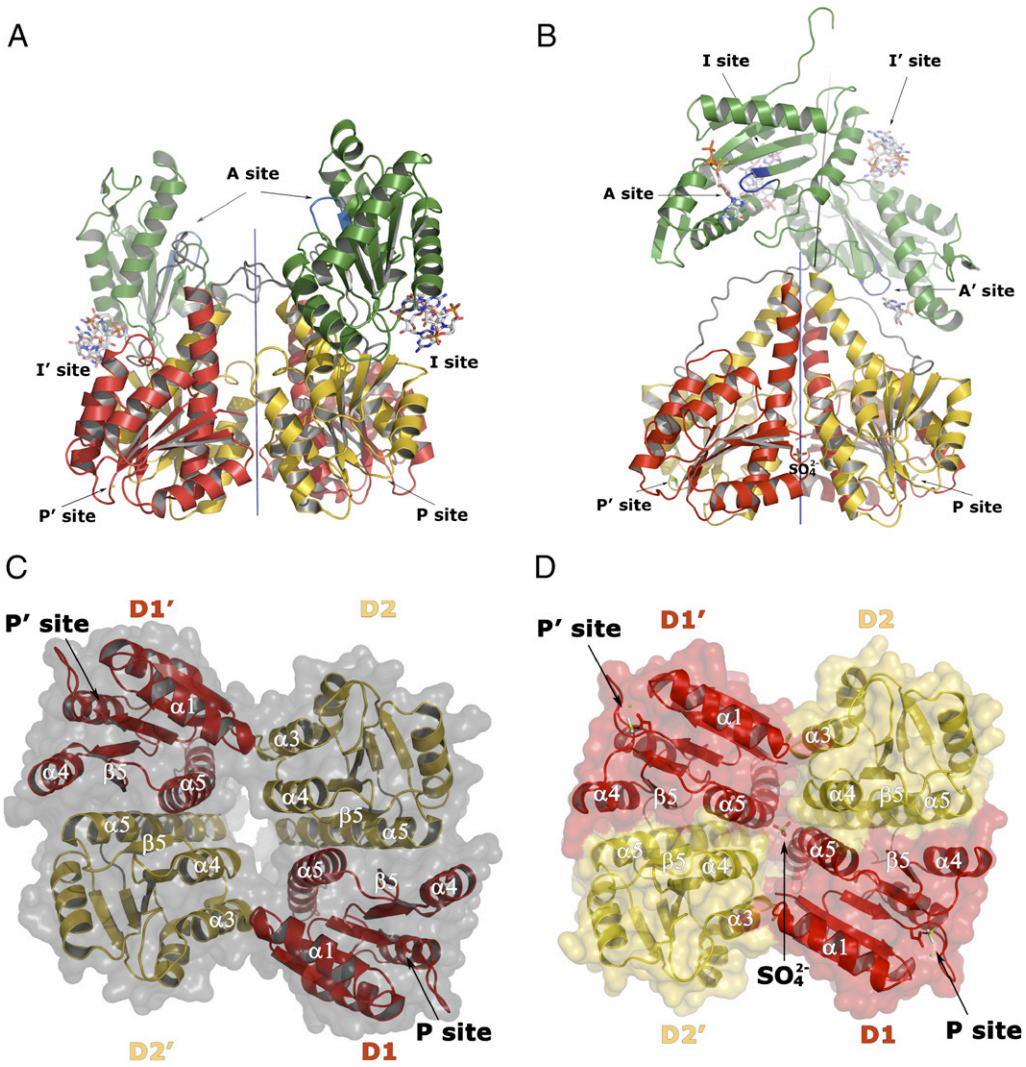


Figure 2. Ribbon Diagrams of the Dimeric Crystal Structures of Nonactivated PleD and BeF_3^--Activated PleD

(A–D) In (A) and (B) the view is perpendicular to the two-fold axis of the stem. In (C) and (D), the view is rotated by 90° around a horizontal axis with respect to the top panels, showing the bottom view of the $(D1/D2)_2$ stem with the DGC domains in the rear clipped off for clarity. The domains are colored in red (Rec domain D1), yellow (adaptor domain D2), and green (enzymatic domain DGC, with the GGEEF signature motif highlighted in blue). The disordered parts of the interdomain linkers are shown in gray. Labels with a prime indicate symmetry-related elements. The two-fold symmetry axes are shown as thin, straight lines. (A and C) Nonactivated PleD (Chan et al., 2004) is associated to a loose dimer. The active sites (A sites) are occupied by c-di-GMP, which crosslinks to an adjacent dimer (not shown). Intercalated $(c\text{-di-GMP})_2$ dimers are bound to allosteric inhibition sites I and I′. Each inhibition site is comprised of a primary inhibition site on DGC (I_p; Arg359, Asp362, Arg390) and a secondary site on the adaptor domain ($I_{s,D2}$;R148, R178). (B and D) In the activated structure, the phosphorylation site (P site) is modified by BeF_3^- and Mg^{2+}, and the active site (A site) harbors GTPαS/Mg^{2+}. $(C\text{-di-GMP})_2$ dimers are bound to the dyad-related sites I and I′. Each site is comprised of the primary I_p site, as in the nonactivated structure, and a secondary I site of the symmetry-related DGC ($I_{s,DGC}$; R313; also see Figure 6). The two A sites face in opposite directions, rendering the enzyme catalytically incompetent.

Feedback Inhibition Probed by Mutagenesis

PleD shows noncompetitive product inhibition with a K_i of about 0.5 μM (Chan et al., 2004) (Table 2). Crosslinking of the DGC domain to the D2 domain, as seen in the crystal structure of nonactivated PleD, has been proposed to be the mechanism of product inhibition in PleD, since this would prevent the productive encounter of the two GTP-loaded enzyme domains (Chan et al., 2004). To test the "inhibition by domain immobilization" hypothesis, pertinent PleD mutants have been analyzed recently (Christen et al., 2006). Mutations of primary I site (I_p) residues were, apart from R390A, largely deleterious for activity and, thus, inconclusive. Mutation R390A, however, showed a considerably increased K_i. Truncation of the secondary I site residues on D2 ($I_{s,D2}$; mutant R148A/R178A), on the other hand, did not affect feedback inhibition.

Here, we applied a photometric pyrophosphate assay to reinvestigate this $I_{s,D2}$ mutant and to probe the role of the secondary $I_{s,DGC}$ site (Arg313). As can be seen from Table 2, mutating the $I_{s,D2}$ or the $I_{s,DGC}$ site alone had a

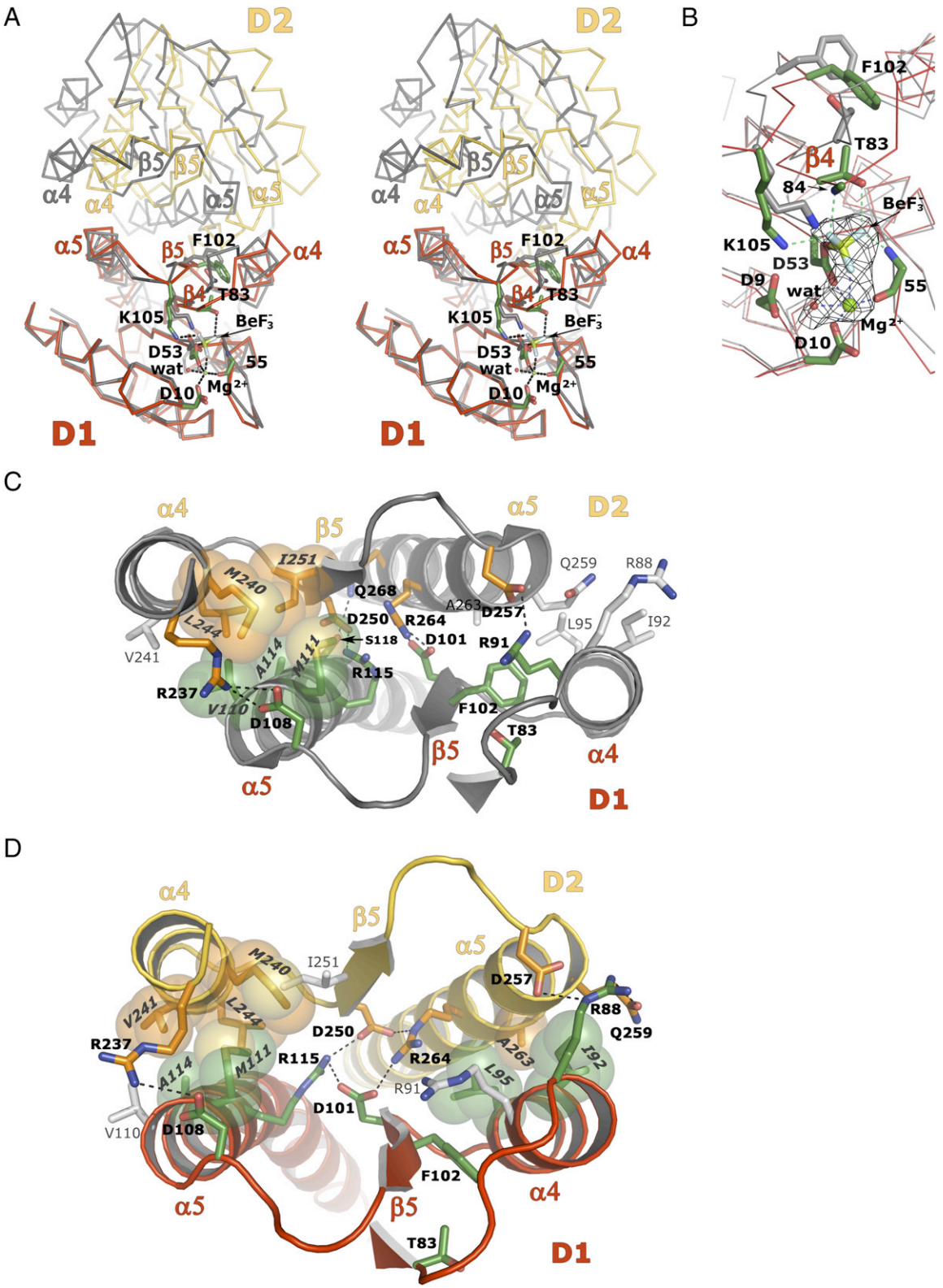


Figure 3. Structural Changes Invoked by BeF_3^- Modification of Asp53

(A) Comparison of the Cα traces of the D1/D2 PleD stem in its activated (D1, red; D2, yellow) and nonactivated (gray) conformation after superposition of the D1 domains. Residues of the acidic pocket as well as Thr83 and Phe102 are shown in full.

(B) Blow-up of the superimposed D1 domains around the activation site; residues involved in pseudo phosphorylation are shown in full. The omit map for Asp53-BeF_3^- and the associated Mg^{2+} ion is depicted at 3α.

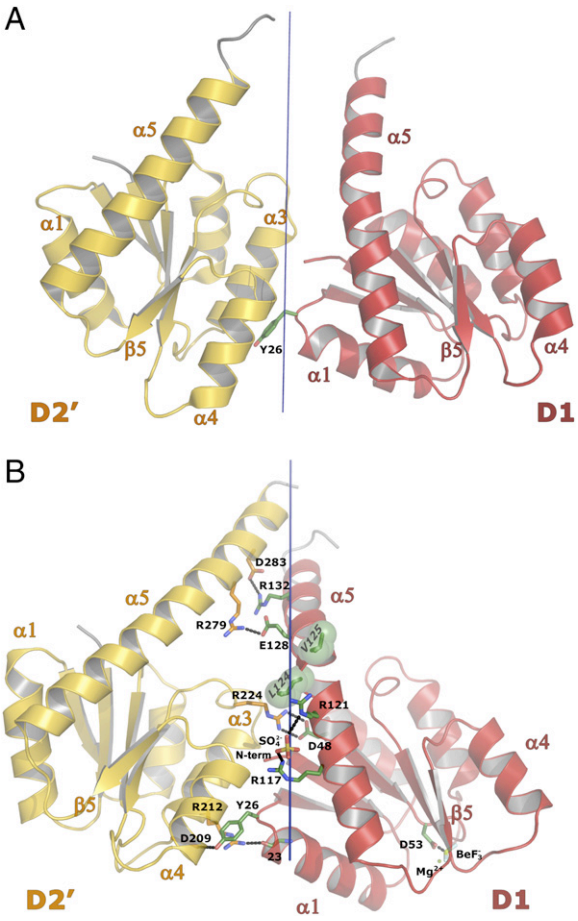


Figure 4. Intermonomer D1/D2′ Contact before and after Activation

(A and B) In the dimeric stem, all shown contacts occur twice due to the two-fold symmetry. Interface residues are shown in full. (A) In nonactivated PleD, the interaction is restricted to a small contact patch around Tyr26. In (B) activated PleD, there is a multitude of polar and ionic interactions and an apolar contact between Leu124, Val125 (green spheres,) and the corresponding residues of the symmetry-related D1′ (not shown). A putative sulfate ion is found on the symmetry axis crosslinking Ag117 and Arg121 with their symmetry mates.

moderate effect on the inhibition constants, with a 10-fold and 2-fold increase, respectively. Our values deviate slightly from those measured by Christen et al. (2006) for wild-type and R148A/R178A PleD, possibly because their data had been measured by a different technique (thin-layer chromatography). Furthermore, the published K_i values were actually IC_{50} values. Truncating both secondary I sites in the triple mutant R148A/R178A/R313A increased the K_i by more than 60-fold. Interestingly, it was not possible to inhibit this mutant completely, with the inhibited state showing a residual k_{cat}' rate constant of about 5% of k_{cat} (Table 2; Figure S2D, see the Supplemental Data available with this article online). All mutants, but in particular those involving $I_{s,D2}$ residues, showed considerably elevated activity (Table 2). This may be due to a decreased dimerization K_D of the nonactivated mutant proteins.

The K_D of c-di-GMP binding to the triple mutant is about 4 μM (ITC measurement, data not shown) and, thus, is increased by only a factor of about 10 with respect to PleD wild-type (R.P. et al., unpublished data). This affinity most probably is due to c-di-GMP binding to the uncompromised I_p site. In summary, the data indicate that the two immobilization modes operate redundantly, i.e., the integrity of only one of the I_s sites is required for noncompetitive product inhibition, presumably via domain crosslinking.

DISCUSSION

DGCs play a central role in bacterial c-di-GMP signaling. Therefore, their action, synthesis of the secondary messenger, must be tightly controlled (Christen et al., 2006). Comparison of the structures of nonactivated and activated PleD allows us to propose the mechanisms of activation by phosphorylation and, unexpectedly, suggests two redundant modes of feedback inhibition. For a schematic overview of the various structural states of PleD, see Figure 7. Furthermore, the binding mode of the substrate analog GTPαS gives insight into the catalytic mechanism of the cyclization reaction.

Activation

It has recently become apparent that BeF3-modified Rec domains of RR transcription factors of the OmpR/PhoB family form two-fold symmetric homodimers (Bachhawat et al., 2005; Toro-Roman et al., 2005a, 2005b). Since their DNA-binding domains bind cooperatively to their target sequences, it was suggested that protein dimerization is the mechanism of activation. In some cases, the active dimer structure was also seen in crystals of the native protein (Toro-Roman et al., 2005a, 2005b), which may, however, be of no physiological relevance considering the high protein concentration needed for crystallization.

In PleD, the signal generated by pseudo phosphorylation of Asp53 is transduced to the α4-β5-α5 face, where it promotes the D1/D2 rearrangement, which, in turn, facilitates dimerization. As in other Rec domains, the conserved Phe/Tyr at the center of this interface follows the movement of a Thr/Ser and changes its rotameric state. The movement of Phe102 is compensated by a slight twist of the C terminus of helix α4, allowing residues Ile92 and Leu95 to make apolar interdomain contacts with Ala263. Together with other rather subtle differences at the α4-β5-α5 face of D1 a substantial domain rearrangement is induced (Figure 3A). It is unlikely that these rearrangements are caused by crystal packing, since, in the loosely

(C and D) Close-up view of the D1/D2 interface in the (C) nonactivated and (D) activated state showing the quasi-two-fold symmetric interface. Both domains contribute their α4-β5-α5 face. The viewing direction is along the quasi-dyad. All residues of the D1/D2 interface as well as Thr83 and Phe102 are shown in stick and sphere representation for polar and apolar residues, respectively. Residues, which do not interact in one state, but do so in the other, are shown in white.

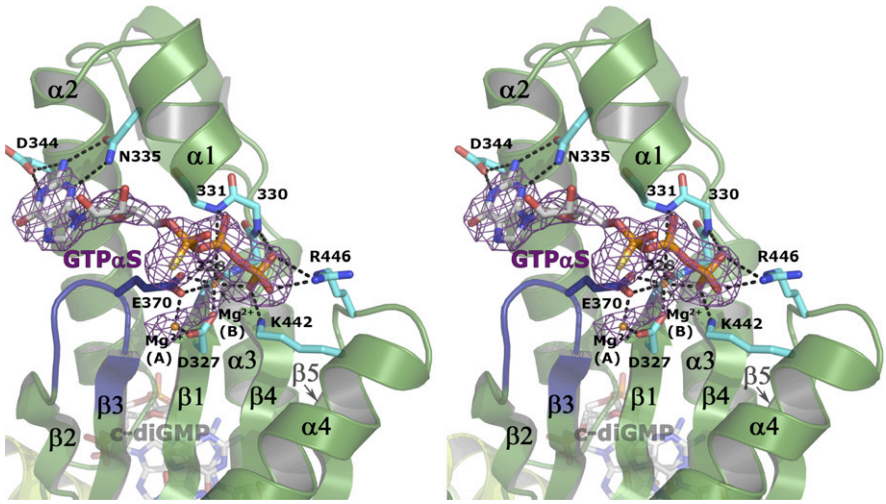


Figure 5. Substrate Analog GTPαS and Mg^{2+} Bound to the Active Site of PleD

The omit map for the ligands is contoured at 3σ. The DGC domain is shown in ribbon representation; the GGEEF signature hairpin is shown in dark blue, and all interacting residues and the P loop main chain (residues 328–331) are shown in full. $(c\text{-di-GMP})_2$ bound to the I_p site of the DGC domain can be seen in the rear.

packed crystals (V_M = 3.8 Å^3/Da), the crystal contact areas are more than 2-fold smaller than the dimer interface.

Interestingly, the two PleD Rec domains of a given chain interact quasi-two-fold symmetrically and in a very similar way as what is observed for the homodimerization of Rec domains of the OmpR/PhoB subfamily. While the α4-β5-α5 face is the interaction surface in both cases, the similarity even extends to the residue level. In PleD, Asp101/Arg115 and Asp250/Arg264 form two quasi-isologous interdomain salt bridges at the center of the interface (Figures 3C and 3D), whereas homologous residues form isologous interchain salt bridges in the Rec domains of the OmpR/PhoB family (e.g., Asp99/Arg113 for ArcA) (Toro-Roman et al., 2005a).

DGCs catalyze the condensation of two identical substrates (GTP) to form the two-fold symmetric c-di-GMP product. Since the DGC domain binds only one substrate, two such domains are needed for catalysis, as depicted in the mechanistic model of Figure 7 (state 4). Control of dimerization allows for simple and efficient regulation of the catalytic activity, as the reaction type for the encounter of substrate-loaded DGC domains is changed from bi- to unimolecular. At the high protein concentration of 20 μM used in the enzymatic assays (Table 2), dimerization of nonactivated PleD with a K_D of about 100 μM is not negligible, explaining the observed constitutive activity. The physiological PleD concentration is not known and may vary within the cell. But, clearly, activation by reduction of the dissociation constant, K_D, will be most effective at protein concentrations well below the K_D of nonactivated PleD. Indeed, it has been shown that, at a lower protein concentration of 5 μM, the catalytic rate is enhanced by a factor of 35 upon BeF_3^- modification, and that it is the dimeric species that carries the activity (R.P. et al., unpublished data).

For WspR from *Pseudomonaas fluorescens*, which is a DGC with a single N-terminal Rec domain, a more indirect dimerization mechanism, with phosphorylation relieving the dimer interface obstruction caused by the effector domain, has been proposed (Malone et al., 2007). For RRp1 from *Borrelia burgdorferi*, which has the same Rec-DGC domain organization, phosphorylation-dependent cyclase activity has been demonstrated (Ryjenkov et al., 2005). Thus, in these cases, a second Rec-like domain such as the adaptor domain of PleD appears to be dispensible for dimerization, and it may serve an additional regulatory function in RRs with a Rec-Rec-DGC domain composition (see below). How DGCs without Rec domains are activated remains to be investigated. Several of them may be constitutive dimers such as DgcA (CC3285) from *C. crescentus* with its putative N-terminal coiled-coil domain (P.W., unpublished data).

Catalysis

The active site of the DGC domain of PleD is identified by the bound substrate analog GTPαS. It consists of well-defined subsites for the β- and γ-phosphates and for the guanine base (Figure 5), which explains the observed substrate specificity (Chan et al., 2004; Paul et al., 2004). The structural similarity of the DGC domain to adenylate cyclase (AC) and DNA polymerase (POL) has been pointed out (Chan et al., 2004). From this work, it is apparent that the similarity also extends to the mode of substrate binding as far as the position of the terminal phosphates close to the P loop, the presence of metal B, and its coordination by the phosphates and two invariant carboxylates (D327 and E370 in PleD) are concerned. Metal site A is also occupied in one of the subunits (Figure 5), but the α-phosphate is not in coordinating distance as in the bacterial AC CyaC (Steegborn et al., 2005) and in POL

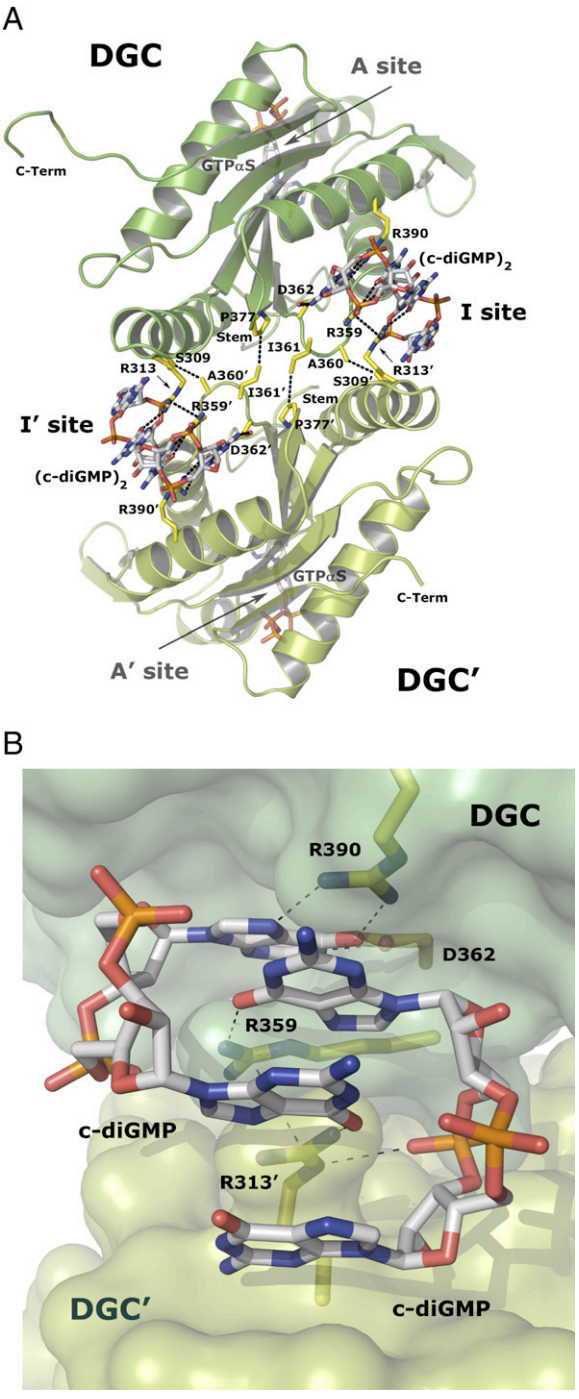


Figure 6. Crosslinking of the DGC Domains by c-di-GMP

(A) Ribbon diagram of the DGC dimer along its symmetry axis; both $(c\text{-di-GMP})_2$ ligands and interacting residues are shown in full. The $(c\text{-di-GMP})_2$ ligands are bound to the I site (I_P site and the $I_{s,DGC}$ site of the adjacent subunit).

(B) Close-up view of the intercalated $(c\text{-di-GMP})_2$ ligand, which crosslinks the two DGC domains (dark- and light-green surface representation) by binding to the I_p (Arg359, Asp362, Arg390) and the $I_{s,DGC}$ (R313′) site of the adjacent domain.

(Doublie et al., 1998). This may be caused by the thiol modification of the ligand or the absence of the second substrate. Modeling shows that moderate torsional adjustments can bring the α-phosphate in coordinating distance to metal A, with an orientation ready for in-line attack onto the P_α-P_β diester bond. This can occur without compromising binding of the terminal phosphates or the guanine.

The role of the conserved Glu371 in DGCs (Figure S1) has not been clarified, but it may coordinate metal A transiently and/or serve as a proton acceptor for the incoming 3′-hydroxyl of the other subunit. Indeed, mutagenesis has shown that Glu371 is as indispensable for catalysis as are the magnesium ligands Glu370 (Christen et al., 2006) and Asp327 (data not shown). Another conserved residue in the active site is Lys332, whose side chain amino group can easily adopt a position from which it could stabilize the charge of the pentavalent phosphoryl transition state and the pyrophosphate leaving group (Figure S1). The same role has been postulated for the nonhomologous Arg1150 in CyaC (Steegborn et al., 2005).

To form the catalytically competent enzyme-substrate complex (Michaelis-Menten complex), two substrate-loaded DGC domains have to line-up in antiparallel orientation, such that the 3′-hydroxyl groups of the bound substrates are brought in close proximity to the α-phosphate of the other GTP molecule and such that they are properly positioned to perform an intermolecular in-line nucleophilic attack onto P_α from the side opposing the susceptible P_α-P_β diester bond. Modeling shows that this can be achieved without clashes in a dimeric arrangement, in which helix α4 of one domain contacts the small β sheet (β0, β0′) of the symmetry-related domain (Figure S1). The interface carries a number of ionic side chains that probably are involved in intermolecular salt bridges. The model of the $(DGC)_2$-c-di-GMP product complex obtained after only minor rearrangements of the reactive groups (Figure S1B) is different from that observed experimentally in nonactivated PleD (Chan et al., 2004), where c-di-GMP crosslinks A sites of adjacent dimers. This is due to a difference in the position of the ribose and α-phosphate moieties relative to the protein, whereas the guanine base is bound in the same way. The orientation of the product in the active site of nonactivated PleD is most probably enforced by the artifactual, ligand-mediated dimer-dimer association in the crystal.

The close resemblance of the constellation of reactive groups and metal ions in the modeled PleD Michaelis-Menten complex with that in AC and POL suggests that the same two-metal-assisted mechanism for phosphodiester formation as suggested for ACs (Tesmer et al., 1999) and for POLs (Steitz, 1999) is operational in DGCs. More kinetic and mutagenesis experiments have to be performed to reveal the precise catalytic mechanism of DGCs.

Feedback Inhibition

DGCs exhibit exquisite noncompetitive product inhibition, as demonstrated biochemically for PleD (Chan et al., 2004; Paul et al., 2004) and for DgcA from *C. crescentus*

Table 2. Kinetic Data of Nonactivated Wild-Type and Mutant PleD

	Wild-type PleD	PleD R148A/R178A	PleD R313A	PleD R148A/R178A/R313A
IC_{50} (μM)	7 ± 2	17 ± 4	9 ± 2	30 ± 5
K_i (μM)	0.5[a]	5	0.9	33
K_s (μM)	26	20[a]	21	25
k_{cat} (μM c-di-GMP/[s · μM PleD])	$9.0 \cdot 10^{-4}$	$1.5 \cdot 10^{-2}$	$4.0 \cdot 10^{-3}$	$1.6 \cdot 10^{-2}$
k_{cat}' (μM c-di-GMP/[s · μM PleD])	0[a]	0[a]	0[a]	$8.1 \cdot 10^{-4}$

IC_{50} values represent experimental values with estimated standard deviations based on two measurements at GTP concentrations of 100 μM and 500 μM. The concentration of the protein was 20 μM. The other kinetic parameters have been obtained by fitting numerical simulations to the time course of c-di-GMP concentration at various inhibitory concentrations (Figure S2). In the simple kinetic scheme, enzyme and substrate as well as enzyme and product are in thermodynamic equilibrium, with dissociation constants of K_s and K_i, respectively, with the product binding noncompetitively to an allosteric site. Also, product and substrate binding are assumed to be independent of each other. k_{cat} and k_{cat}' are the catalytic rate constants of substrate turnover for noninhibited and inhibited protein, respectively. The fitted values for the on-rate constants of substrate and product ranged between 100 s^{-1} and 350 s^{-1}. It should be noted that the fits for PleD wild-type and mutant R148A/R178A were poor (Figures S2A and S2B), and, hence, the corresponding parameters not very well determined. This may be due to the simplicity of the kinetic model.
[a] Parameter was fixed in simulation.

(Christen et al., 2006), and, based on conserved I_p sites, this is expected to be the case for the majority of DGC domains (Christen et al., 2006). Feedback inhibition prevents futile GTP consumption and probably defines an upper ceiling for the concentration of the second messenger c-di-GMP in the bacterial cell. Christen et al. (2006) showed that the allosteric I_p site is crucial for binding and product inhibition and suggested that communication between the

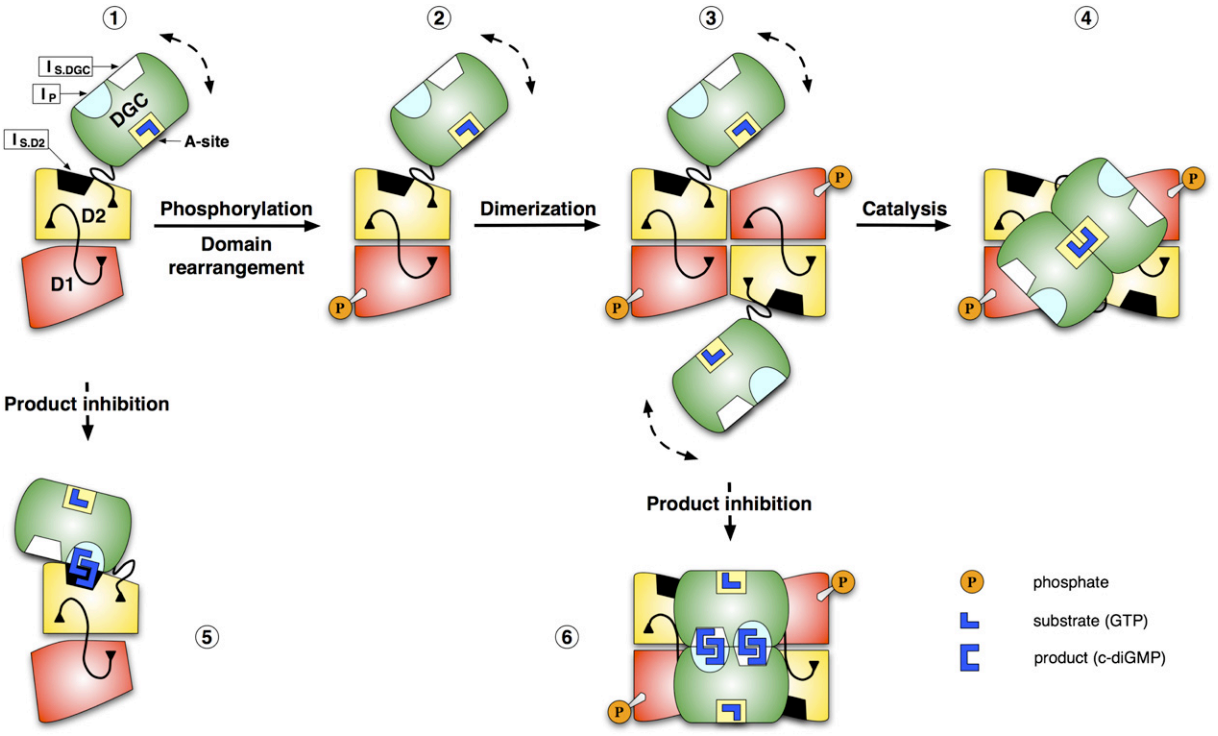


Figure 7. Mechanistic Model of PleD Regulation
The model is adapted from Chan et al. (2004). The DGC domain (green) is connected via a flexible linker to the stem (receiver domain D1 [red] and adaptor domain D2 [yellow]) and is supposed to be mobile relative to it. (Upper row) Activation. Phosphorylation of domain D1 leads to a rearrangement of the stem domains, which, in turn, allows for formation of a tight dimeric stem (3). The dimeric arrangement is a prerequisite for an efficient and productive encounter of the two substrate-loaded DGC domains to form the c-di-GMP product (4). (Lower row) Product inhibition. Dimeric product molecules, $(c\text{-di-GMP})_2$, can crosslink the primary inhibition site on DGC, I_p, with a secondary binding site either on D2, $I_{s,D2}$ (5) or on the adjacent DGC domain, $I_{s,DGC}$ (6). The former structure has been observed experimentally with nonactivated PleD (Chan et al., 2004), the latter structure is presented in this report. In both cases, the DGC domains become immobilized, and the active sites are hampered from a productive encounter. Note that a possible direct communication between I_p and A sites (Christen et al., 2006) is not depicted.

I_p site and the active site of a DGC domain could cause the observed noncompetitive product inhibition. However, in the nonactivated structure (Chan et al., 2004), the enzymatic product in the form of two intercalated c-di-GMP molecules was found to crosslink the I_p with the $I_{s,D2}$ site on the adjacent D2 domain (Figure 7, state 5). Therefore, an ''inhibition by domain immobilization'' mechanism, in which domain crosslinking by c-di-GMP would prevent productive encounter of the two enzymatic domains of the dimer, was proposed. Indeed, $(c\text{-di-GMP})_2$ can be regarded as a multivalent ligand that offers, among other features, the O6-N7 edges of its four guanyl bases for specific pairing with arginine guanidyl groups, as observed in DNA-protein interactions. In nonactivated PleD, the ligand crosslinks the well-defined and well-conserved (Christen et al., 2006) I_p site on the $\alpha 2$-$\beta 2$ corner of DGC (formed by Arg359, Asp362, Arg392) with arginines 148 and 178 of D2 ($I_{s,D2}$).

In the present structure, a similar mechanism can be deduced. In this mechanism, the dimeric ligand is again bound to the I_p site, but its other valency crosslinks to a hitherto unrecognized, to our knowledge, secondary site on the adjacent DGC domain of the dimer ($I_{s,DGC}$). As a result, the DGC domains interact in a nonproductive manner with the two A sites, which point away from each other (Figure 2B). Our functional analysis of pertinent mutants (Table 2) suggests that domain immobilization might indeed contribute to feedback regulation, and that the two inhibition modes operate redundantly (Figure 7, states 5 and 6). The advantage of redundancy in this context is not clear, but we note that efficient DGC-DGC crosslinking obviously can operate only after dimerization, while D2-DGC crosslinking occurs within the nonactivated monomer. Why the triple mutant is not fully active upon c-di-GMP binding, but rather switches to a slower rate (Table 2), remains to be investigated. This may be due to the residual binding capacity of the mutated I_s sites, the existence of additional I_s sites with low affinity, or a change of active site structure and/or dynamics upon allosteric product binding to the I_p site, as suggested by molecular simulations (Christen et al., 2006).

The relevance of feedback inhibition has been discussed before upon recognizing that the $R_{359}xxD_{362}$ motif of the I_p site is largely conserved (Christen et al., 2006). It is present in 59% of all GGDEF domain sequences of the current PFAM release (5200 sequences) that have an intact GGD(E)EF motif and are, therefore, putative DGCs. A large part of these seem to operate via c-di-GMP crosslinking of I_p with an $I_{s,DGC}$ site, as observed in the present PleD structure. This can be inferred from the fact that 64% of the DGC sequences with the RxxD motif show an arginine at position 313 of helix $\alpha 0$ or at a position one turn of the helix farther up (positions 316 and 317), from where the long side chain probably could also reach the ligand. Covariation calculations show a significant correlation of 0.45 (0.32, 0.61) for simultaneously finding the RxxD motif and an arginine at position 313 (316, 317). Of the subset of 19 DGCs for which guanylate cyclase activity has been shown experimentally (see Figure S4 in Christen et al. [2006]), 14 display the RxxD motif of the I_p site, and all but one display an arginine residue at one of the secondary positions.

This implies that catalytic domain immobilization by DGC-DGC crosslinking is a rather common feedback inhibition mechanism of DGCs and does not rely on secondary inhibition sites on other domains. This autonomy of the DGC domain has probably been an advantage in evolution to conserve feedback inhibition. Interestingly, feedback inhibition is also observed in the prototype of a ''single-domain'' DGC, DgcA from *C. crescentus*, whose sequence exhibits both I_p and $I_{s,DGC}$ sites as well as a small N-terminal extension of possibly coiled-coil structure for dimerization (Christen et al., 2006).

The multivalency of the second messenger c-di-GMP appears perfectly suited for signal transfer. Here, we showed evidence that it can be utilized for domain crosslinking, which elicits large structural changes and which may well also work for downstream signaling processes. Interestingly, c-di-GMP binds to the N terminus of the PilZ effector domain (Christen et al., 2007), i.e., to a strategic position where it could control domain arrangement in multidomain targets carrying this domain. The identification of additional c-di-GMP effector proteins will reveal to what extent this notion can be generalized.

EXPERIMENTAL PROCEDURES

Mutagenesis, Expression, and Purification

PleD mutations, D327A, R313A and R148A/R178A/R313A, were performed by using the QuickChange II Site-Directed Mutagenesis Kit (Stratagene) based on pRUN plasmids (derived from the pBR322 vector) containing the C-terminally His_6-tagged wild-type gene (Chan et al., 2004) or the R148A/R178A double mutant (Christen et al., 2006). C-terminally His_6-tagged full-length PleD and mutants were expressed in the *E. coli* BL21(DE3)pLysS strain, induced by 0.5 mM IPTG (3–4 hr, 30°C). Cells were washed twice with 20 mM Tris-HCl (pH 8.0), 500 mM NaCl buffer and were stored at −80°C. The proteins were purified by affinity chromatography with 5 ml HisTrap columns (GE Healthcare) and ~200 mM imidazole for elution. The pooled protein fractions were concentrated and further purified by size-exclusion (SEC) chromatography by using the Superdex 200 HR 26/60 column (GE Healthcare) and 20 mM Tris-HCL (pH 8.0), 100 mM NaCl as running buffer (SEC buffer). For this step, the protein concentration was kept low (5–10 mg/ml) to ensure dissociation and removal of bound c-diGMP. The protein was concentrated to 10 mg/ml, aliquoted, and stored at −20°C for further use.

Activation of PleD by BeF_3^- Modification

Activation of PleD was accomplished by supplementing the SEC buffer with $BeCl_2$, NaF, and either $MgCl_2$ or $MnCl_2$ (A buffer). Final concentrations were 200 μM for PleD, 1 mM for $BeCl_2$, 10 mM for NaF, 10 mM for $MgCl_2$, and 1 mM for $MnCl_2$. The protein was incubated for at least 30 min before experiments were conducted.

C-di-GMP

Pure samples of c-di-GMP were obtained from N. Amiot, Department of Chemistry, University of Basel. It was chemically synthesized according to procedures by Amiot et al. (2006).

Enzymatic Assays

DGC activity of wild-type and mutant PleD was analyzed by monitoring the production of pyrophosphate by using a pyrophosphatase-coupled spectrophotometric assay. Details are given elsewhere (Chan

et al., 2004). The kinetic data were fitted to numerical simulations of a scheme of noncompetitive product inhibition by using the program Berkeley Madonna.

Analytical Size-Exclusion Chromatography

Dimerization of PleD was monitored by size-exclusion chromatography (SEC) by using a Superdex 200 HR 10/30 column (GE Healthcare). SEC and A buffers were used as running buffers for nonactivated and activated protein samples, respectively. Runs were performed on an ÄKTA*purifier* (GE Healthcare) system at a flow rate of 0.7 ml/min. The concentrations of the eluted peak fractions were measured by an online refractometer (Optilab rEX, Wyatt Technology).

Isothermal Titration Calorimetry

Samples of the PleD mutant protein R313A were first diluted 15-fold either in SEC or A buffer (containing Mn^{2+} as divalent cation) and then concentrated by using a 15 ml Amicon Ultra MWCO 10 kDa (Millipore) concentrator. Protein samples and buffers were degassed (ThermoVac, MicroCal) at 2° below the temperature used in the experiments before loading into the VP-isothermal titration calorimetry (ITC) (Microcal, Northampton, MA). Details of the ITC measurement procedure and analysis of self-association have been described previously (McPhail and Cooper, 1997; Velazquez-Campoy et al., 2004). In short, 10 μl samples of nonactivated (two experiments at concentrations 929 μM and 861 μM) and activated (108 μM and 99 μM) PleD were injected into the cell filled with buffer (1.4 ml volume) at 6 min intervals. The measurements of nonactivated and activated protein were performed at 15°C and at 10°C, respectively. Both measurements of each state were fit together to obtain global ΔH and K_D values. A constant background heat, produced by the dilution of the titrant, residual differences in the buffer composition in syringe and cell, and technical effects was eliminated by allowing for a constant heat contribution (Q_{dil}) as another fit parameter upon data evaluation.

Crystallography

Crystallization of activated PleD was performed at room temperature by the hanging-drop vapor-diffusion technique. Before crystallization, c-di-GMP and Rp-GTP-α-S (BioLog) were added with final concentrations of 0.2 mM and 1 mM, respectively, to BeF_3^-/Mg^{2+}-modified PleD samples (final concentration 100 μM). A (1 μl + 1 μl) mixture of protein and well solution (0.1 M HEPES [pH 8.0], 0.73 M Na_2SO_4) gave rise to crystals of needle shape. Diffraction data were collected from a single crystal at the Swiss Light Source, Paul-Scherrer-Institute, Villigen, Switzerland, and were processed with MOSFLM/SCALA (CCP4, 1994). Orientations and positions of individual domains were determined by molecular replacement with the structure of nonactivated PleD (PDB code 1W25) by using PHASER (McCoy, 2007). The model was built by using COOT (Emsley and Cowtan, 2004) and was refined with REFMAC (CCP4, 1994). Two-fold NCS restraints were imposed.

Supplemental Data

Supplemental data include a figure of the modeled structure of the competent PleD-substrate complex, a figure of the PleD-product complex and a figure with the kinetic data of wild-type and mutant PleD and are available at http://www.structure.org/cgi/content/full/15/8/915/DC1/.

ACKNOWLEDGMENTS

Dietrich Samoray and Arnaud Baslé are greatly acknowledged for expert technical support. We thank the staff of the Swiss Light Source (SLS, PSI, Villigen, Switzerland) for assistance with data collection and N. Amiot, Department of Chemistry, University of Basel, for the gift of chemically synthesized c-di-GMP. We thank Dietrich Samoray, Dagmar Klostermeier, and Zora Markovic-Housley for critical reading of the manuscript. The work was supported by grant 3100A0-105587 of the Swiss National Science Foundation (to T.S.).

Received: January 27, 2007
Revised: June 25, 2007
Accepted: June 25, 2007
Published: August 14, 2007

REFERENCES

Aldridge, P., Paul, R., Goymer, P., Rainey, P., and Jenal, U. (2003). Role of the GGDEF regulator PleD in polar development of *Caulobacter crescentus*. Mol. Microbiol. *47*, 1695–1708.

Amikam, D., and Galperin, M.Y. (2006). PilZ domain is part of the bacterial c-di-GMP binding protein. Bioinformatics *22*, 3–6.

Amiot, N., Heintz, K., and Giese, B. (2006). New approach for the synthesis of c-di-GMP and its analogues. Synthesis *24*, 4290–4296.

Bachhawat, P., Swapna, G.V., Montelione, G.T., and Stock, A.M. (2005). Mechanism of activation for transcription factor PhoB suggested by different modes of dimerization in the inactive and active states. Structure *13*, 1353–1363.

Birck, C., Mourey, L., Gouet, P., Fabry, B., Schumacher, J., Rousseau, P., Kahn, D., and Samama, J.P. (1999). Conformational changes induced by phosphorylation of the FixJ receiver domain. Struct. Fold. Des. *7*, 1505–1515.

CCP4 (Collaborative Computational Project, Number 4) (1994). The CCP4 suite: programs for protein crystallography. Acta Crystallogr. D Biol. Crystallogr. *50*, 760–763.

Chan, C., Paul, R., Samoray, D., Amiot, N.C., Giese, B., Jenal, U., and Schirmer, T. (2004). Structural basis of activity and allosteric control of diguanylate cyclase. Proc. Natl. Acad. Sci. USA *101*, 17084–17089.

Christen, B., Christen, M., Paul, R., Schmid, F., Folcher, M., Jenoe, P., Meuwly, M., and Jenal, U. (2006). Allosteric control of cyclic di-GMP signaling. J. Biol. Chem. *281*, 32015–32024.

Christen, M., Christen, B., Allan, M.G., Folcher, M., Jeno, P., Grzesiek, S., and Jenal, U. (2007). 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. USA *104*, 4112–4117.

Djordjevic, S., Goudreau, P.N., Xu, Q., Stock, A.M., and West, A.H. (1998). Structural basis for methylesterase CheB regulation by a phosphorylation-activated domain. Proc. Natl. Acad. Sci. USA *95*, 1381–1386.

Doublie, S., Tabor, S., Long, A.M., Richardson, C.C., and Ellenberger, T. (1998). Crystal structure of a bacteriophage T7 DNA replication complex at 2.2 Å resolution. Nature *391*, 251–258.

Egli, M., Gessner, R.V., Williams, L.D., Quigley, G.J., van der Marel, G.A., van Boom, J.H., Rich, A., and Frederick, C.A. (1990). Atomic-resolution structure of the cellulose synthase regulator cyclic diguanylic acid. Proc. Natl. Acad. Sci. USA *87*, 3235–3239.

Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. *60*, 2126–2132.

Jenal, U., and Malone, J. (2006). Mechanisms of cyclic-di-GMP signaling in bacteria. Annu. Rev. Genet. *40*, 385–407.

Lee, S.Y., Cho, H.S., Pelton, J.G., Yan, D., Berry, E.A., and Wemmer, D.E. (2001). Crystal structure of activated CheY. Comparison with other activated receiver domains. J. Biol. Chem. *276*, 16425–16431.

Lewis, R.J., Brannigan, J.A., Muchova, K., Barak, I., and Wilkinson, A.J. (1999). Phosphorylated aspartate in the structure of a response regulator protein. J. Mol. Biol. *294*, 9–15.

Liaw, Y.C., Gao, Y.G., Robinson, H., Sheldrick, G.M., Sliedregt, L.A., van der Marel, G.A., van Boom, J.H., and Wang, A.H. (1990). Cyclic diguanylic acid behaves as a host molecule for planar intercalators. FEBS Lett. *264*, 223–227.

Malone, J.G., Williams, R., Christen, M., Jenal, U., Spiers, A.J., and Rainey, P.B. (2007). The structure-function relationship of WspR,

a *Pseudomonas fluorescens* response regulator with a GGDEF output domain. Microbiology *153*, 980–994.

Maris, A.E., Sawaya, M.R., Kaczor-Grzeskowiak, M., Jarvis, M.R., Bearson, S.M., Kopka, M.L., Schroder, I., Gunsalus, R.P., and Dickerson, R.E. (2002). Dimerization allows DNA target site recognition by the NarL response regulator. Nat. Struct. Biol. *9*, 771–778.

McCoy, A.J. (2007). Solving structures of protein complexes by molecular replacement with Phaser. Acta Crystallogr. D Biol. Crystallogr. *63*, 32–41.

McPhail, D., and Cooper, A. (1997). Thermodynamics and kinetics of dissociation of ligand-induced dimers of vancomycin antibiotics. J. Chem. Soc., Faraday Trans. *93*, 2283–2289.

Nowak, E., Panjikar, S., Konarev, P., Svergun, D.I., and Tucker, P.A. (2006). The structural basis of signal transduction for the response regulator PrrA from *Mycobacterium tuberculosis*. J. Biol. Chem. *281*, 9659–9666.

Paul, R., Weiser, S., Amiot, N.C., Chan, C., Schirmer, T., Giese, B., and Jenal, U. (2004). Cell cycle-dependent dynamic localization of a bacterial response regulator with a novel di-guanylate cyclase output domain. Genes Dev. *18*, 715–727.

Robinson, V.L., Buckler, D.R., and Stock, A.M. (2000). A tale of two components: a novel kinase and a regulatory switch. Nat. Struct. Biol. *7*, 626–633.

Robinson, V.L., Wu, T., and Stock, A.M. (2003). Structural analysis of the domain interface in DrrB, a response regulator of the OmpR/PhoB subfamily. J. Bacteriol. *185*, 4186–4194.

Ryjenkov, D.A., Tarutina, M., Moskvin, O.V., and Gomelsky, M. (2005). Cyclic diguanylate is a ubiquitous signaling molecule in bacteria: insights into biochemistry of the GGDEF protein domain. J. Bacteriol. *187*, 1792–1798.

Sinha, S.C., and Sprang, S.R. (2006). Structures, mechanism, regulation and evolution of class III nucleotidyl cyclases. Rev. Physiol. Biochem. Pharmacol. *157*, 105–140.

Steegborn, C., Litvin, T.N., Levin, L.R., Buck, J., and Wu, H. (2005). Bicarbonate activation of adenylyl cyclase via promotion of catalytic active site closure and metal recruitment. Nat. Struct. Mol. Biol. *12*, 32–37.

Steitz, T.A. (1999). DNA polymerases: structural diversity and common mechanisms. J. Biol. Chem. *274*, 17395–17398.

Stock, A.M., Robinson, V.L., and Goudreau, P.N. (2000). Two-component signal transduction. Annu. Rev. Biochem. *69*, 183–215.

Tesmer, J.J., Sunahara, R.K., Johnson, R.A., Gosselin, G., Gilman, A.G., and Sprang, S.R. (1999). Two-metal-Ion catalysis in adenylyl cyclase. Science *285*, 756–760.

Tischler, A.D., and Camilli, A. (2004). Cyclic diguanylate (c-di-GMP) regulates *Vibrio cholerae* biofilm formation. Mol. Microbiol. *53*, 857–869.

Toro-Roman, A., Mack, T.R., and Stock, A.M. (2005a). Structural analysis and solution studies of the activated regulatory domain of the response regulator ArcA: a symmetric dimer mediated by the α4-β5-α5 face. J. Mol. Biol. *349*, 11–26.

Toro-Roman, A., Wu, T., and Stock, A.M. (2005b). A common dimerization interface in bacterial response regulators KdpE and TorR. Protein Sci. *14*, 3077–3088.

Velazquez-Campoy, A., Leavitt, S.A., and Freire, E. (2004). Characterization of protein-protein interactions by isothermal titration calorimetry. Methods Mol. Biol. *261*, 35–54.

Accession Numbers

The coordinates and structure factors of activated PleD in complex with c-diGMP and Rp-GTP-α-S have been deposited in the Protein Data Bank under accession code 2v0n.

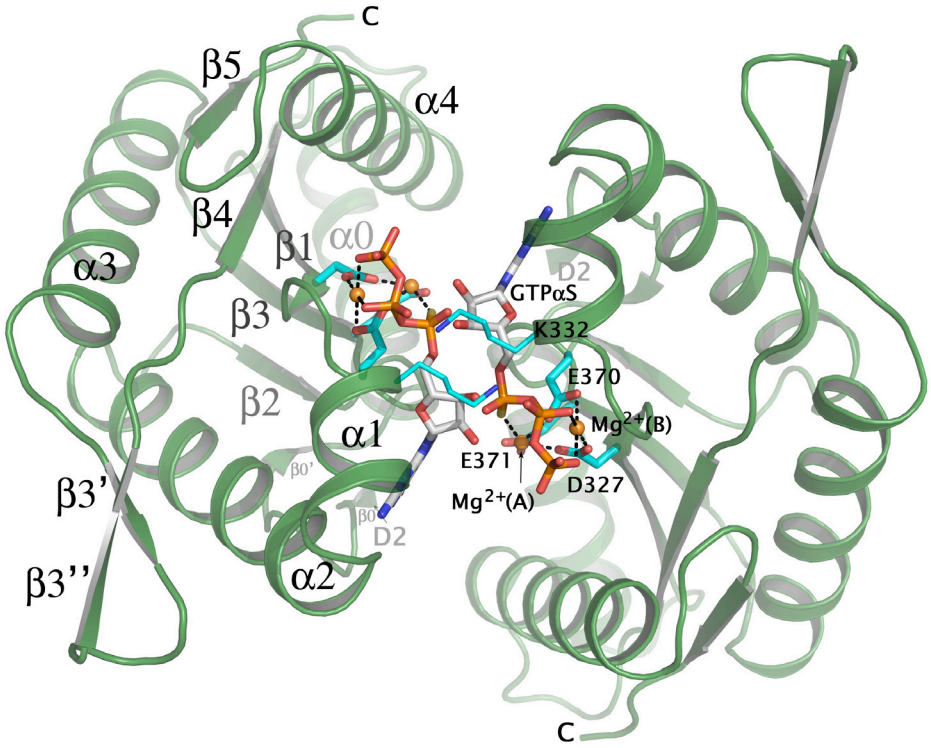
Structure *15*

Supplemental data

Structure of BeF_3^--Modified Response Regulator PleD: Implications for Diguanylate Cyclase Activation, Catalysis, and Feedback Inhibition

Paul Wassmann, Carmen Chan, Ralf Paul, Andreas Beck, Heiko Heerklotz, Urs Jenal, and Tilman Schirmer

A


C
β5
α4
β4
α0
β1
α3
β3
D2
GTPαS
K332
E370
β2
α1
Mg2+(B)
E371
D327
β3'
β0'
Mg2+(A)
D2
β3''
α2
C

B

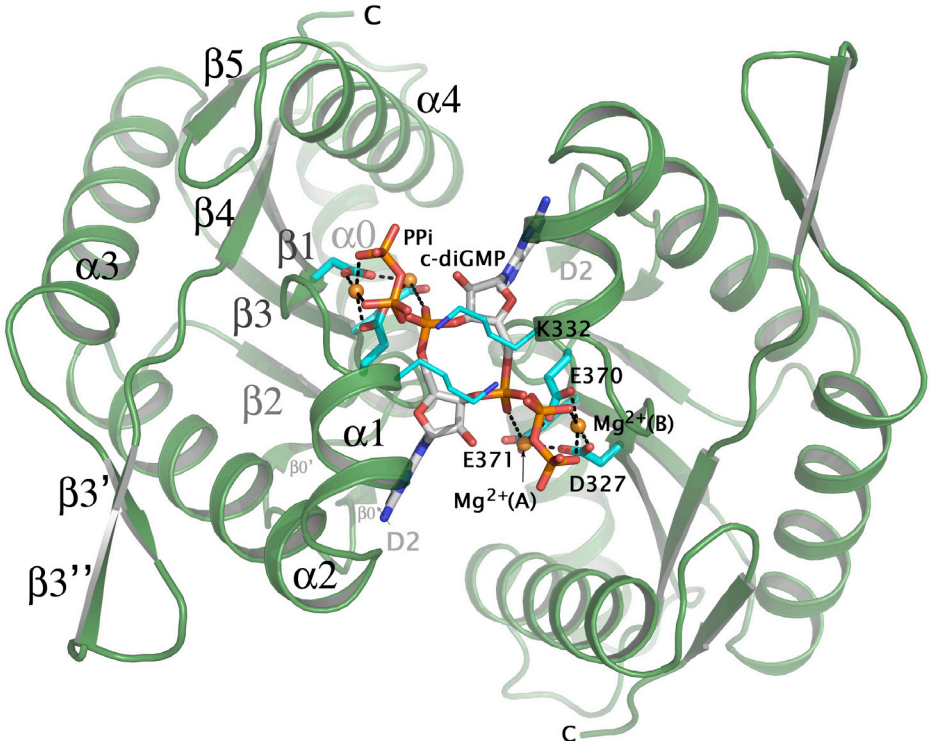
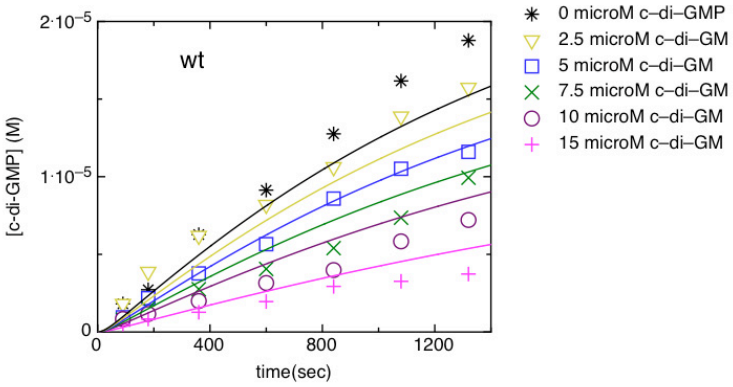

C
β5
α4
β4
α0
PPi
β1
c-diGMP
α3
D2
β3
K332
E370
β2
Mg2+(B)
α1
E371
D327
β3'
β0'
Mg2+(A)
D2
β3''
α2
C

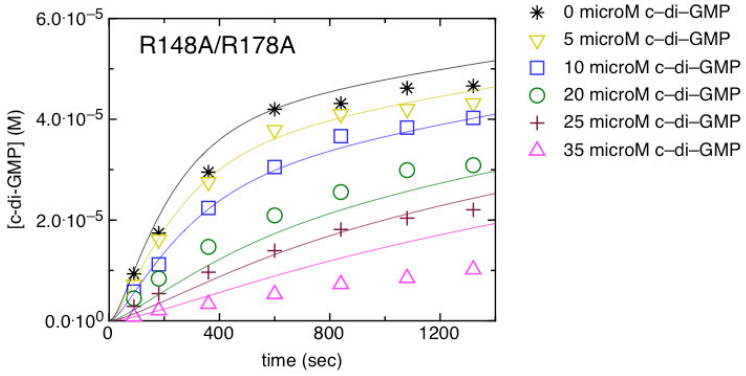
Figure S1. Model of the Catalytically Competent Dimer Constellation of Diguanylate Cyclase

Ribbon diagram with active site residues and ligands shown in full. (A) Michaelis-Menten complex with a GTP molecule bound to each active site. (B) Complex of the enzyme with the products c-di-GMP and pyrophosphate after intermolecular phosphodiester formation between 3' hydroxyl and P_α. Note the steric hindrance between the two products. Since the pyrophosphate molecules are tightly bound, it is probable that the c-di-GMP product will change its orientation in the binding site prior to dissociation. The Michaelis-Menten model is based on the GTPαS - PleD complex structure (Fig. 4), with the α-phosphate rotated manually into coordinating distance to metal A. The two DGC domains are approximately 2-fold symmetric with the symmetry axis (viewing direction) running through the center of the bound c-di-GMP in panel (B).

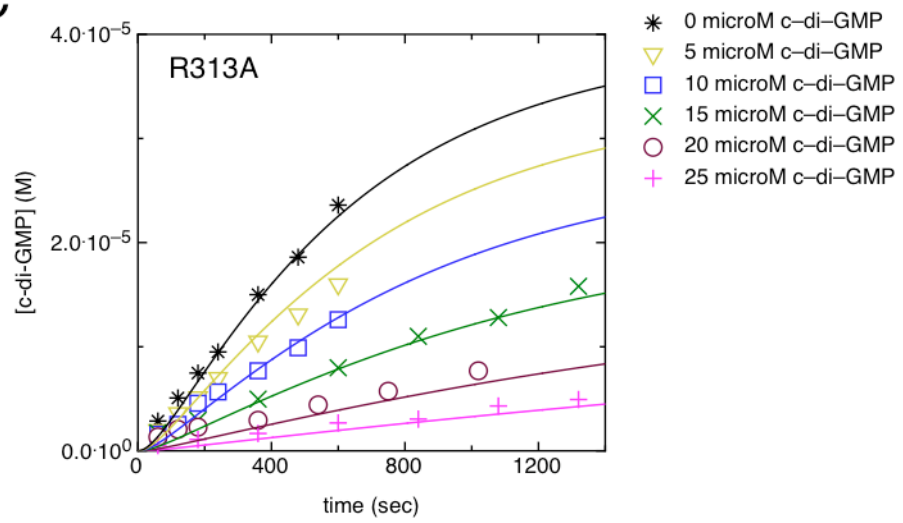
A


wt
[c-di-GMP] (M)
time(sec)
0 microM c–di–GMP
2.5 microM c–di–GM
5 microM c–di–GM
7.5 microM c–di–GM
10 microM c–di–GM
15 microM c–di–GM

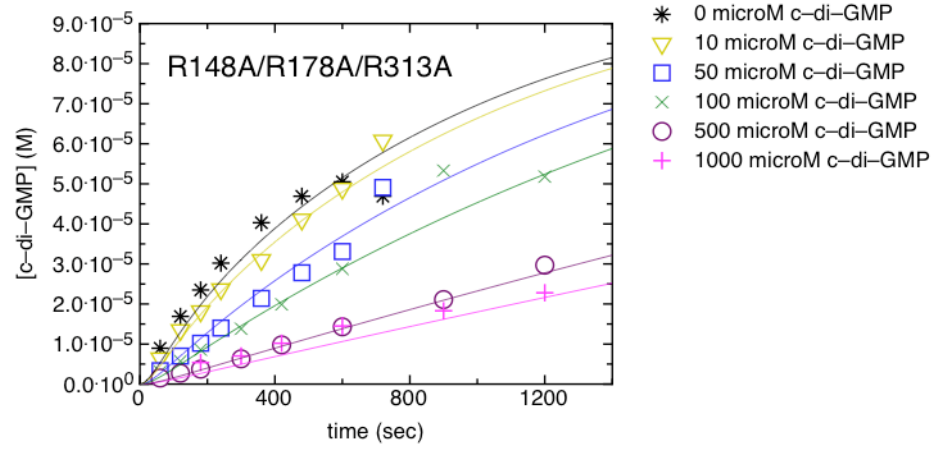
B


R148A/R178A
[c-di-GMP] (M)
time (sec)
0 microM c–di–GMP
5 microM c–di–GMP
10 microM c–di–GMP
20 microM c–di–GMP
25 microM c–di–GMP
35 microM c–di–GMP

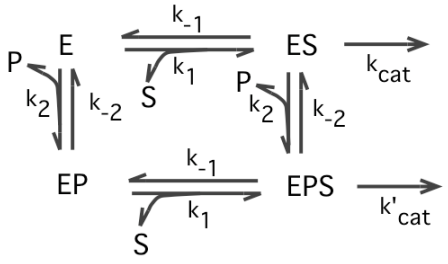
C



D



E



$K_s = k_{-1}/k_1$

$K_i = k_{-2}/k_2$

Figure S2. Time Course of c-di-GMP Production at Various c-di-GMP Concentrations

Panels A, B, and C show the data for wild-type PleD and PleD mutants R148A/R178A, R313A, R148A/R178A/R313A, respectively. The symbols indicate the experimental values obtained with a spectrophotometric assay that measures the production of pyrophosphate, the second product of the DGC reaction 2 GTP -> c-di-GMP + 2 PPi (see Material and Methods of main text). The protein concentration was 20 μM, the GTP substrate concentration 100 μM. The thin lines show simulated values based on a simple kinetic model (panel E) of non-competitive product inhibition assuming independence of substrate and product binding. The resulting parameters are given in Table 2 of the main text.

II.2 Second crystal structure of activated PleD – new insights in the mechanisms regulating dimerization and catalysis

Paul Wassmann[1], Claudia Massa[1] and Tilman Schirmer[1*]

[1] Department of Structural Biology, University of Basel, CH-4056 Basel, Switzerland

[*] Corresponding author. E-mail address: tilman.schirmer@unibas.ch

Keywords: response regulator, diguanylate cyclase, GGDEF domain, two-metal assisted catalysis.

Abbreviations: TCS, two-component system; HK, histidine kinase; HPt, histidine-containing phosphotransfer protein; RR, response regulator; Rec, respone regulator receiver domain; c-di-GMP, bis-(3'-5')-cyclic guanosine monophosphate; GTPαS, guanosine 5'-O-(1-thiotriphosphate); GMPCPP, 5'-guanylylmethylenebis-phosphonate; DGC, diguanylate cyclase; NC, nucleotidyl cyclase; AC, adenylate cyclase; POL, RNA/DNA polymerase.

Abstract

We report the crystal structure of BeF_3^--activated diguanylate cyclase PleD with the product c-di-GMP bound to one of the active sites and to the allosteric inhibition sites. PleD consists of two tandem response regulator receiver domains and a C-terminal GGDEF domain. The activated PleD dimers form a charged pocket on the molecular dyad of the $(Rec1\text{-}Rec2)_2$ stem. This binding pocket turns out to be a feature of Rec1-Rec2-GGDEF proteins.

Diguanylate cyclases, in contrast to nucleotidyl cyclases and DNA/RNA polymerases, form simultaneously two phosphodiester bonds utilizing the 'two-metal assisted' catalysis. The main question concerning this reaction is – how do the diguanylate cyclases manage to ensure the precision of the enzymatic reactions. We have analyzed the catalysis relevant features of the diguanylate cyclase PleD on its structural level and propose a catalytic mechanism for this family of proteins.

Introduction

Eubacteria, archaea and some eukaryotic classes utilize the 'two-component system' (TCS) signal transduction pathway to regulate their response to environmental traits (1,2). The archaic TCS consist of a transmembrane histidine kinase (HK) sensing the environmental stimuli, and a cytosolic response regulator (RR) responsible for the cellular response. Stimulus recognition by HK results in autophosphorylation of a conserved histidine residue in the so called H-box of the dimerization domain and in subsequent transfer of the high-energetic phosphoryl group to a conserved aspartate of the receiver domain (Rec) of RRs (3-5).

Several organisms use the more elaborated phosphorelay version of TSC, containing additionally a histidine-containing phosphotransfer (HPt) protein that shuttles phosphoryl groups between Rec domains (6,7). Evolutionary and structural analysis of HPts shows partly higher similarity to the HK dimerization domains than between proteins of the HPt class (7-9). Crystallization of HPts in complex with Rec domains of their cognate RRs has revealed the mechanisms of selectivity and phosphoryl group transfer between HKs and RRs (7,9).

Phosphorylation of Rec domains induces subtle structural changes (for details see reviews (3,5)) resulting in activation of RR's effector domains. Although several activation mechanisms are described (10-12), in many cases activation leads to dimerization of Rec domains (13,14).

High exploitation of the TCSs by bacteria can be observed in freshwater bacterium *C. crescentus*, possessing 44 different RRs (2,15). One of these RRs is PleD (16). PleD is involved in pole morphogenesis during the dimorphic cell division of *C. crescentus* producing flagellated swarmer and a holdfast bearing cells (17-19). Upon phosphorylation by its cognate HKs PleC and DivJ (20-22) PleD is sequestered to the stalked pole (20). It could be shown that dimerization and not activation is the imperative for shuttling to the pole (23). Unfortunately, neither the recruiting partner of PleD at the pole, nor the pole localization signal, which is coded by dimeric PleD, are known.

PleD consists besides the N-terminal phosphorylatable Rec1 and the non-functional Rec2 domain, which form in activated state the two-fold symmetric $(Rec1\text{-}Rec2)_2$ 'stem', of a C-terminal GGDEF effector domain (13,24). GGDEF domains have been shown to possess the diguanylate cyclase (DGC) activity and to produce the novel

second messenger c-di-GMP out of two GTP molecules (20,25). According to its cellular concentration, c-di-GMP is involved in environment-sensing and multi-cellular communication processes like sessility and motility, biofilm formation and synthesis of virulence factors in hosts (26-29).

GGDEF domains have been predicted (30) to utilize the same ferredoxin-like structural fold as monocyclic nucleotidyl cyclases (NC) and RNA/DNA polymerases (POL). This was subsequently verified by the crystal structure analysis of non-activated PleD (denoted as PleD-na throughout this paper) (24). Moreover the recent crystal structure of activated PleD in complex with substrate analoge GTPαS and two putative Mg^{2+}-ions (denoted as PleD-act1 throughout this paper) (13) demonstrate, that all three protein classes utilize the two-metal assisted catalysis mechanism (31). The nucleophilic in-line attack of (deoxy-) ribose 3'-hydroxyl group at Pα and release of Pβγ via a S_N2 (bimolecular nucleophilic substitution) mechanism is supported by two Mg^{2+}-ions. Mg^{2+}-ion (B) stabilizes the phosphates of the bound substrate as well as the intermediate pentacovalent coordinated Pα (31). Mg^{2+}-ion (A), besides stabilizing Pα, activates the attacking 3'-hydroxyl group by lowering its pK_A value either directly or through a mediating water molecule (31). Both metal ions are coordinated by the carboxyl groups of two acidic residues, which are highly conserved in all three protein clades (32). Despite performing a single phosphodiester forming reaction similarly to NCs and POLs, DGCs have the task to perform two such reactions simultaneously. Capability of a single GGDEF domain to bind only a single GTP molecule, as was shown by PleD crystal structures (13,24), results in utilization of a more sophisticated catalytic mechanism by DGCs. As to produce c-di-GMP catalytically, GGDEF domains have to form perfectly aligned dimers (23).

Here, we present the crystal structure of BeF_3^--activated PleD in complex with an active site bound c-di-GMP molecule (PleD-act2). We have compared this structure with other DGC, as well as NC and POL structures, aiming to answer the following questions. What function does the non-phosphorylatable Rec2 domain exert in the DGC PleD, being dispensable for dimerization in other Rec-GGDEF domain proteins (50)? How do the DGCs prevent monocyclization of GTP to cGMP, a reaction performed by NCs, having all necessary residues for this reaction in a single GGDEF domain? Concerning the precision of the DGCs in the process of the product biosynthesis – how do DGCs manage to carry out both cyclization reactions simultaneously, avoiding generation of product intermediates? And last but not least,

which role might the highly conserved, but functionally unassigned but indispensable glutamate 371 of the GGDEF signature motif play in c-di-GMP synthesis?

Materials and Methods

Expression and Purification

Expression and purification of C-terminally His_6-tagged wild-type and R148A R178A R313A PleD as well as wild-type DgcB (Uniprot entry: Q9A776) constructs were performed as described previously (13). Removal of bound c-di-GMP was accomplished by performing an additional size-exclusion chromatography (SEC) run using Superdex 200 HR26/60 column (GE Healthcare), 20 mM Tris-HCl (pH 8.0), 100 mM NaCl as running buffer and protein at low concentration (< 5 mg/ml). Activation procedure of PleD by $BeF_3^- \bullet Mg^{2+}$ is described elsewhere (13).

Crystallography

Prior to crystallization of activated PleD, c-di-GMP and GMPCPP (Jena Bioscience) were added with final concentrations of 0.2 mM and 2 mM, respectively. Crystallization of $BeF_3^- \bullet Mg^{2+}$ activated PleD was described previously (13). Single crystal derived diffraction data was processed with MOSFLM/SCALA(33). The crystal structure was determined by molecular replacement using the program PHASER (34) and the structure of PleD-act1 (PDB code 2v0n) as template. Model building and refinement was performed with COOT (35) and REFMAC (33), imposing two-fold NCS restraints.

Enzymology

Enzymatic activity of PleD in presence of sulfate and phosphate was measured at a protein concentration of 40 µM in a 50 mM Tris-HCl pH 8.0, 10 mM $MgCl_2$ buffer. PleD was activated adding 1 mM $BeCl_2$ and 10 mM NaF. Effects of sulfate and phosphate were measured at standard conditions adding sodium salts of both anions to a final concentration of 5 mM. Reactions were started by addition of 6.7 mM GTP and incubated for 20 minutes at room temperature. Detection of produced c-di-GMP was detected by the FPLC method (see below).

Enzymatic activity of DgcB in presence of divalent cations was performed at a protein concentration of 29.5 µM in a 50 mM Tris-HCl pH 8.0 buffer. Manganese or magnesium was added to enzymatic reactions at 0.1, 0.5, 1, 5 and 10 mM final-concentrations, respectively. Effect of calcium, that was used at same concentrations, was measured in presence of 2 mM Mg^{2+}. Reactions were started adding 0.5 mM

GTP and were incubated for 1 hour at room temperature prior to inactivation and detection by FPLC method.

Quantification of c-di-GMP levels

Quantification of c-di-GMP amount, produced by DgcB, PleD wild-type and R148A R178A R313A mutant, was performed by FPLC-anion exchange chromatography (36). In short, enzymatic reactions were stopped by denaturating the protein at 95° C for 1 minute. Heated samples were filtered (Ultrafree-MC, 0.22 µm, Millipore), diluted 1:10 in a 5 mM NH_4HCO_3 solution (= running buffer) to an end-volume of 1 ml and applied to a 1 ml Resource Q anion exchange column using an ÄKTA Purifier FPLC machine (GE Healthcare). Elution was carried out with a 5 to 1000 mM NH_4HCO_3 gradient. Nucleotides were eluted as well separated peaks applying a shallow gradient of the elution buffer, resulting in identification of GDP, GTP and c-di-GMP nucleotides at distinct positions, respectively. Identity of these nucleotides was verified by comparison to elution profiles of nucleotide standards and by ESI-TOF MS. UV-peak areas of c-di-GMP standards at different concentrations were used for standard curve creation for quantification of c-di-GMP.

HPLC-ESI TOF MS analysis

Prior to mass spectrometry the FPLC peak fractions were lyophilized in a SpeedVac (ThermoScientific) and resolubilized in 150 µl 20 mM ammoniumacetate pH 6.4, that was used as running buffer to avoid GTP degradation. Nucleotide mass detection was performed on a micrOTOF (Bruker Daltonics) ESI TOF MassSpec with preceding HPLC (Agilent) run using a SUPELCOSIL™ LC-18-T reverse phase column. Acetonitrile was used as elution buffer. MS analyses were performed in negative mode, at a detector voltage of 120 V and with a pre-pulse storage of 1 µsec.

Results

Structural overview

Crystals of BeF_3^--activated PleD were obtained under the same conditions as for PleD-act1 (13), but supplemented with GMPCPP. Data collection was performed with a single crystal diffracting initially to 2.5 Å. The crystal structure was solved by molecular replacement using the structure of PleD-act1 (PDB code: 2v0n) as search model (13) and was refined to 2.8 Å (Table 1). The observed electron density represents the whole main chain except the linker between Rec1 and Rec2 (residues 137-148) and the C-terminal 6xHis-tag of subunit A. The structure strongly resembles the search model PleD-act1 sharing the same space group (Figure 1A), although both crystals vary by about 5 Å in unit cell axes a and b. Result of the reduced unit cell of PleD-act2 is a crystal contact, which leads to unfolding of the short β3'-β3''-hairpin in subunit B. Structural comparison of the individual domains with ones from PleD-act1 (13) shows no further changes in the tertiary structure. Pair-wise superposition of the corresponding domains of PleD-act1 and PleD-act2 shows an 'all-atom'-RMSD in the range of 0.5 to 1.0 Å.

The activation of the phosphorylation sites in the Rec1 domains by the $BeF_3^- \bullet Mg^{2+}$ moieties results in identical modifications in both structures, PleD-act1 and PleD-act2. In comparison to the non-activated protein (24), the modification leads to intramolecular repositioning of the Rec domains resulting in the formation of a tight $(Rec1\text{-}Rec2)_2$ stem as was described before (13). As in PleD-act1, both GGDEF domains of PleD-act2 are cross-linked by two c-di-GMP dimers *via* the R359xxD motif of one domain and R313 of the other and vice versa.

Active site bound c-di-GMP molecule

The major difference between the two activated PleD structures is a different orientation of the c-di-GMP cross-linked GGDEF domains with respect to the $(Rec1\text{-}Rec2)_2$ stem. After superposition of the stems, the GGDEF domains have to be rotated by 8.5° with a screw component of 1.2 Å to superpose. Whereas in PleD-act1 no direct intermolecular contacts were detected between the catalytic and the Rec domains, in the present structure the GGDEF domain of protomer B forms a small contact interface (432 $Å^2$ –PISA (37)) with Rec1 and Rec2 of chain A (Figure 1b). This contact appears mainly mediated by a c-di-GMP molecule and is further

stabilized by a few H-bonds and a weak salt bridge (R148 (A) – D344 (B)). Indication that the observed intermolecular contact is not a crystallization artifact and may play some regulatory role *in vivo*, comes from previous studies on PleD (38), showing a four times higher activity of the R148A mutant in comparison to the wild type protein. The embedded c-di-GMP molecule forms rather weak contacts with the Rec domains (Figure 1B). One guanine base of the cyclic dinucleotide is stacked between R137 and P72 in van-der-Waals contact. Additionally, both bases of c-di-GMP enclose the hydrophobic G(144)AAA segment of the Rec1-Rec2 linker.

Similar interaction of the c-di-GMP molecule with the catalytic site of the GGDEF domain via the guanine moiety has been observed before (24) (Figure 2C). Besides the known H-bonding to N335 and D344, in PleD-act2 the guanine's O6-atom is in distance to form a polar contact with the highly conserved R366 (Figure 2B). Like in structure of PleD-na the 2'-hydroxyl of the ribose moiety forms a weak H-bond (3.1Å) with N335. Additional electron density was observed between one α-phosphate non-ester oxygen and E371's carboxyl, which we assigned to a putative Mg^{2+}-ion. Interaction between c-di-GMP and E371 *via* a putative Mg^{2+}-ion further increases the interface between the active site and the product (Figure 2B). Furthermore, each of the active sites bear a crystallization condition derived sulfate ion at the position of the GTPαS γ-phosphate in the structure of PleD-act1 (Figures 2A-B), interacting with the same protein residues. In contrast to PleD-act1, no electron density of the catalytic Mg^{2+}-ions A and B could be detected near the side chains of D327 and E370. Thus, the catalytic metal ions seem to bind to these acidic residues only in presence of proper substrate. To our surprise at neither of the two active sites any electron density could be detected for the substrate analog GMPCPP, although this substrate analog was present during crystallization in a two-fold excess compared to that of the successfully incorporated GTPαS of PleD-act1 (Figure 2A).

<u>Anion binding on the stem dyad</u>

As in PleD-act1, a putative sulfate ion is found bound to a deep, positively charged pocket at the "bottom" of the $(Rec1\text{-}Rec2)_2$ stem on the molecular dyad axis. The almost buried sulfate molecule interacts with the guanidinium groups of R117 and R121 and their symmetry mates (Figure 3). Figure 4A shows the positively charged binding site formed by arginines that is surrounded by a negatively charged rim (E25, E245).

All four Rec domains of a PleD dimer are involved in the formation of this deep, charged cavity. Therefore, we wondered whether such a cavity is a feature of proteins with two consecutive Rec domains that undergo dimerization. If this is the case, conservation of PleD residues involved in sulfate binding should be observed.

1070 sequences of proteins bearing two Rec domains were identified using the SMART web-server (39). To our surprise these sequences code only for proteins having either a catalytical or no output domains at all, whereas it is known that ~ 60% of all response regulators are DNA binding proteins (3,15) and only 11% are proteins with an enzymatic function (3). Of these 1070 proteins only 271 have two consecutive Rec domains. The two most abundant classes of these proteins are the DGCs with a C-terminal GGDEF domain (161 proteins) followed by 'stand-alone' Rec1-Rec2 proteins (88 proteins).

We have analyzed the conservation profiles of residues involved in phosphorylation (=activation) and formation of the charged pocket in these protein classes (Table 2). The 'stand-alone' Rec1-Rec2 proteins seem to be phosphorylated not on the Rec1 but on their Rec2 domain. Additionally, this protein class shows no significant conservation of residues, which might be involved in dimerization and charge distribution in the cavity.

In contrast to the 'stand-alone' Rec1-Rec2 proteins, Rec1-Rec2-GGDEF proteins seem to be activated mainly *via* their Rec1 domains. Although in case of PleD the Rec2 domain misses residues that would be involved in activation, high number of proteins seem to be activated at both Rec domains in the class of Rec1-Rec2-GGDEF proteins.

Y26 of PleD has been shown to be essential for dimerization and to be conserved in proteins that are homolog to PleD (23). Due to similar and partially even higher conservation of residues involved in cavity formation compared to residues involved in dimerization (Figure 4B-C, Supplemental figure 1), the observed charged pocket in PleD turns out to be a characteristic of Rec1-Rec2-GGDEF proteins.

RRs are known to interact with their cognate HKs *via* highly conserved hydrophobic residues near the phosphoacceptor site and additional polar interactions, that are HK-RR pair specific (9). To exclude the possibility that the described charged pocket might be involved in RR-HK contact formation, we have compared the Rec domain structure of PleD with structures of Rec-HPt complexes (7,9,40,41). This analysis shows that residues of PleD that are homologous to the ones involved in the Rec-HPt

interface are distinct to the ones involved in the charged cavity formation (Supplemental figure 2).

Stabilization of PleD dimers in presence of an anionic ligand

The extended interaction between α5-helices of the Rec domains is substantially contributing to the formation of PleD dimers (Figure 3). The sulfate molecule is interacting with positively charged residues on one pole of the two-fold symmetric axis of the PleD dimer. The thermodynamic stability of the $(Rec1\text{-}Rec2)_2$ dimer increases in presence of the observed sulfate molecule. As deduced from the structure with the help of the PISA server (39), the sulfate contributes -21 kcal/mol to the total ΔG of -38 kcal/mol.

This dimer stabilizing effect by the sulfate should manifest itself in an increased DGC activity, since only dimeric PleD is enzymatically active (23). Therefore enzymatic assays with non-activated and $BeF_3^- \bullet Mg^{2+}$ -activated PleD were performed in presence of sulfate and phosphate ions. To avoid product inhibition of PleD by c-di-GMP, R148A R178A R313A mutant that shows a strongly increased K_i (13) was used for this experiment. In non-activated state both, sulfate and phosphate, show a slight inhibitory effect on the catalytic activity of PleD (Figure 5), probably by competing with GTP for binding to the active site as is observed for sulfate in the present PleD structure. In contrast, the catalytic efficiency of activated PeD is increased about 1.6 fold upon addition of sulfate. Phosphate on the other hand acts inhibiting on c-di-GMP production of PleD. The observed modulation of activity is thus specific to the type of the tetrahedral oxyanion and is not a mere charge-neutralizing effect.

Divalent cation dependence of the diguanylate cyclase activity

POLs and NCs, which utilize the same structural fold as DGCs, are also known to show distinct enzymatic behavior in presence of divalent cations Mg^{2+}, Mn^{2+}, and Ca^{2+} (42-44). To observe the influence of divalent cations on the catalytic rate of DGCs, we performed enzymatic assays in presence of Mg^{2+}, Mn^{2+}, and Ca^{2+} ions. We conducted these experiments with DgcB, which is a constitutively active dimer (data not shown), instead of using PleD, which requires Mg^{2+} for dimerization in the activation process. DgcB requires Mg^{2+} or Mn^{2+} for catalytic activity and is inhibited by Ca^{2+}-ions (Figure 6). Highest activity is detected in presence of 5 mM Mg^{2+} and

0.1-0.5 mM Mn^{2+}, respectively. Interestingly, further increase in Mg^{2+}- and Mn^{2+}-ion concentrations leads to attenuation of the enzymatic activity. The same effect is observed for POLs (45-47). In case of DgcB, attenuation by manganese is considerably more potent than by magnesium.

In case of POLs usage of manganese instead of magnesium results in lower enzymatic fidelity and was discussed to be due to loose coordination requirements of manganese (43,48). Regardless the nature of used divalent cations, Mg^{2+} or Mn^{2+}, analysis of catalytic products of PleD and DgcB by ESI-TOF MS resulted in a single species with a mass of 690.1 g/mol that is in agreement with the mass of c-di-GMP. No product intermediates, such as pppGpG, have been detected, arguing for high prevalence of precise, quasi simultaneous formation of both phosphodiester bonds once the substrate-enzyme complex has formed.

Discussion

Function of the Rec2 domain in the Rec1-Rec2-GGDEF proteins

The crystal structures of the atypical RR and DGC PleD (13,24) have shown the pivotal role of the Rec1 domain in the phosphorylation dependent activation process. The Rec2 domain, on the other hand, was assigned an adaptor role, due to absence of residues that are involved in the phosphorylation process. Therefore, one of the main questions concerning structural architecture of DGCs with PleD-like Rec1-Rec2-GGDEF domain composition is, what function is fulfilled by the second Rec domain in these proteins?

Although some of these proteins may utilize tighter control of enzymatic activity using two phosphorylatable Rec domains, a single Rec domain is sufficient for dimerization and therefore for DGC activity, as it was shown for WspR, a Rec-GGDEF domain protein (49,50). In case of PleD, residues of the Rec2 domain, which would be involved in activation by phosphorylation are missing (Table 2), rendering the domain non-functional. Ability of this Rec2 domain to interact with HKs is also lost due to the replacement of strongly conserved hydrophobic residues that are needed for interaction with HKs (9) by charged residues (R164 and R168 in α1-helix, as well as D233 in the β4-α4 loop). Therefore the non-functional Rec2 domain cannot be involved in the regulatory mechanism of signal desensitization by capturing HKs and in turn prevent the transfer of the phosphoryl group to the phophorylatable Rec1 of PleD. Additionally, involvement of Rec2 residues, R148 and R178, in the allosteric control of the DGC PleD, cross-linking Rec2 and GGDEF domains *via* a c-di-GMP dimer (24), must be an unique feature of PleD, due to a complete lack of conservation of these residues in homolog proteins (38).

We have found an evidence for the structural relevance of the Rec2 domain in the class of Rec1-Rec2-GGDEF domain proteins. Both Rec domains are involved in formation of a charged cavity in the activated PleD dimer. This cavity is spatially distinct from the HK recognition area (Supplemental figure 2). Therefore, the found charged pocket, which would be not formed in absence of Rec2, might represent a binding site.

Induction of a ligand binding site by dimerization .

Several aspects support the idea of an anionic binding site formed by the dimer

interface of the Rec1-Rec2-GGDEF domain proteins. The cavity depends structurally on protein dimerization. Therefore residues involved in dimerization and cavity formation must be similarly conserved. This is the case in PleD. The high conservation of R121 depicts the importance of this residue in the maintaining the cavity architecture as well as in the recognition of a negatively charged ligand.

Another aspect is the stabilization of the protein dimer in presence of a putative ligand, which would neutralize the positively charged area, formed by all four arginines (R117, R117', R121 and R121') (Figure 3). RRs coexist in solution in inactive and active conformations and phosphorylation shifts this equilibrium predominantly to the active conformation (51,52), which in case of PleD leads to dimer formation. In presence of the putative ligand, the equilibrium most probably shifts to the active conformation, stabilizing the dimer and thus enhancing the c-di-GMP production (Figure 5). The putative ligand could also play a role in the lifetime enhancement of the phosphorylated protein species. The intrinsic lifetime of phosphorylated RRs is in range of seconds to hours (53), but it was shown to be profoundly prolonged by interaction with other proteins (6).

Nature of the putative ligand

The presence of a specific cavity in dimeric Rec1-Rec2-GGDEF proteins raises the fundamental question about the nature of the putative ligand. The surprising result of the enzymatic assay that BeF_3^--activated PleD can be further activated by sulfate, but is inhibited by phosphate, opened the possibility of sulfate being the specific activator of PleD. But analysis of X-ray structures with bound tetrahedral oxyanions has shown crystallization condition derived sulfate ions mainly bound to protein surfaces forming one or maximally two contacts (54). The residues, which interact with these sulfate ions, are very often not conserved, due to a comparably negligible relevance of sulfate *in-vivo*. One of such sulfate ions is also found in the PleD-act1 structure (13) near the non-functional phosphorylation site of Rec2 (Figure 4A). The interface of the sulfate ion localized at the two-fold symmetry axis of the $(Rec1\text{-}Rec2)_2$ dimer in both structures of activated PleD is more specific. Such a positively charged binding pocket, which almost engulfs the anion and forms extended contacts to the ion, is typical for phosphate and phosphate containing ligands (54-56).

Some considerations deprive sulfate of the role of the seeked ligand. (I) Inappropriately for an activator, sulfate acts as an competitive inhibitor binding to the

active site (Figure 2B). (II) The sequence alignment of Rec1-Rec2-GGDEF proteins shows not only conservation of the binding site's R117 and R121, but also of residues in the rim of the cleft (E25, E245), which do not interact with the sulfate ion at all. Therefore, the putative ligand is most probably an organic molecule bearing a tetrahedral oxyanion group or a specific phospholipid or it could be a phosphorylated protein, which shows some degree of symmetry in line with the symmetric architecture of the cleft.

Possible mechanism of PleD sequestration to the stalked pole in *C. crescentus*

As stated before, PleD is involved in polar differentiation of the asymmetrically dividing bacterium *C. crescentus* (16). Sequestration of PleD to the stalked pole upon phosphorylation has been shown to be independent of DGC activity (23). Furthermore, it appears that phosphorylation induced dimerization is the prerequisite to drive PleD to the pole (23).

Therefore the interface that directs PleD to the pole has to satisfy several requirements. (I) It must be encoded by both protomers of a PleD dimer. (II) Sequestration to the pole is most probably mediated by Rec domains and not by catalytic GGDEF domains, since the later ones have to be able to form a catalytically competent and flexible dimeric subunit (24). (III) The residues that are involved in this interface have to show some degree of conservation and specificity.

Most interestingly, all these requirements are perfectly fulfilled by the charged, two-fold symmetric cleft in the dimer interface. Further *in-vivo* studies are needed to demonstrate the importance of the described cleft in the sequestration process of activated PleD as well as to identify the cognate interaction partner of PleD at the stalked pole.

Utilization of 'two-metal assisted catalysis' by DGCs

Although the crystal structures of PleD (13,24) led to remarkable insights into the enzymatic process utilized by DGCs, which leads to production of one molecule c-di-GMP and two molecules pyrophosphate by condensation of two GTP molecules, several questions stayed unanswered. Some of them concern the catalytic mechanism itself. The shared structural fold of GGDEF domains, ACs and POLs as well as the presence of two putative magnesium ions in the active site of PleD-act1 indicate utilization of the 'two-metal assisted catalysis' mechanism by DGCs.

The main feature of the 'two-metal assisted catalysis' is the presence of two appropriately coordinated Mg^{2+} -ions. In our recent PleD-act1 structure electron density for both catalytic Mg^{2+}-ions was detected (13). Unlike the accurate electron density representing the position of metal ion B, the electron density cloud that represents metal ion A, is rather diffuse (13) (Figure 2b). This is reminiscent of the situation for NCs (57,58) and POLs (44,59) that also show alternative positions for metal ion A, with an A to B ion-distance varying from 3.5 to 4.5 Å (32,60). It was proposed that the 3'OH-hydroxyl, which is coordinated by Mg^{2+}-ion A, assumes the proper attacking position at an interatomic distance of 3.5 Å between both Mg^{2+}-ions (44,60). Longer distances will keep the 3'OH-group out of range for the nucleophilic attack, rendering the enzyme in its 'resting' state.

This theory is supported by biochemical data showing inhibition by Ca^{2+} (42,61-63) and enhanced activity in presence of Mn^{2+} (42,43,64). Ca^{2+}-cations show in POLs similar coordination geometry as Mg^{2+}-ions (65), but have a larger intermetal distance of 3.9 Å in comparison to Mg^{2+}-ions (3.2 Å) (66). Mn^{2+}-ions, on the other hand, have an interatomic distance of 2.9 Å (66) and they do not obey to the strict octahedral coordination behavior of Mg^{2+}-ions.

This catalysis relevant metal ion variability in position may apply for DGCs as well. Similar to POLs, DgcB as a representative of DGCs needs magnesium or manganese for catalysis, but is inhibited by calcium (Figure 6). Both, the structure of PleD in complex with two putative magnesium ions (Figure 2A), and the catalytic behavior of DgcB in presence of divalent cations, corroborate this theory.

Additionally, the DGCs and POLs seem to share the cation induced attenuation behavior (45-47). In both cases high divalent cation concentrations lead to inhibition. Although this behavior is still not completely understood, reduced product release rates in presence of high divalent cation concentrations was proposed as a possible mechanism (44). Mutational studies in RNase H indicate involvement of a glutamate (E188) in the attenuation process (47). Substitution of this residue in *E.coli* DNA Pol I's Klenow fragment to aspartate and alanine results in 30 and 70% activity loss, respectively (45), but in unaffected substrate binding (67). This residue (E188) was structurally shown to move in and out of the active site of RNase H, interacting with metal ion A (46). It was suggested to contribute actively to the release of the product, due to its high mobility. Interestingly, E371 of PleD's GGDEF motif, which contacts the phosphate group of the c-di-GMP molecule *via* a Mg^{2+}-ion (Figure 2B), takes an

analogous position in the structure as E188 in RNase H. In PleD the conservative mutation E371Q results in complete loss of the enzymatic activity (38), showing the exclusivity of this residue for DGCs. These structural and biochemical observations suggest involvement of E371 in the process of product release.

Conformational change of the active site as prerequisite for catalysis

The recent structure of PleD-act1 in complex with two putative magnesium ions and the the substrate analog GTPαS (13) gave many insights in the mechanism of substrate binding by DGCs. A reasonable catalytic mechanism of 'two-metal assisted catalysis' was proposed that is also used by NCs and POLs.

For POLs it has been shown that for the activation of the 3'-OH group, both metal ions have to be in coordination distance with the non-bonding $R_{P\alpha}$-oxygen in a symmetrical manner (31,60,68). When this requirement is not fulfilled, the attacking hydroxide group is not in-line for the substitution reaction or it is too far away for being activated by metal ion A. Therefore the binding mode of the substrate analog GTPαS in the PleD-act1 structure (13) doesn't represent a protein-substrate complex that is ready for the catalytic reaction. The atom in the $R_{P\alpha}$ position, which is a sulfur, is not in coordination distance to metal ion A and B, possibly due to the larger atomic radius of sulfur in comparison to oxygen and the single bond character of the S-P bond (69,70). Similar orientation of R_P-αS nucleotides was observed before in the structures of RNase H (71), tmAC (58) and sAC CyaC (72). In our recent publication (13) we demonstrated that the GTPαS Pα can be brought in contact with metal ion A without disturbing any protein-ligand interactions. Nevertheless, it is impossible to bring the non-bridging $R_{P\alpha}$-oxygen in catalysis relevant coordination between both metal ions by rotation without breaking protein-nucleotide interactions, either to the Pβγ or the guanine moiety.

In the T7 RNAP structure (59) the active site bound substrate analog AMPCPP is trapped in the productive state, ready for the in-line attack. It is possible to fit the C_5'PCPP moiety of this molecule into the active site of PleD (Figure 2D), keeping in mind the following restrains. (I) The position of the Pβγ moiety has to be retained, due to the extended contacts of these phosphates with the P-loop (13). (II) The Pα R_P-oxygen has to replace the imperfect sixth coordination of E370's second carboxyl group to Mg^{2+}-ion B, taking the apical position of the octahedral Mg^{2+}-coordination

opposite to the main-chain carbonyl 328. (III) The Pα R_P-oxygen must be symmetrically coordinated by both metal ions as postulated for the productive state (31). As a result, the C5' position of the fitted C5'PCPP and the GTPαS moieties would deviate substantially (2.7 Å, Figure 2D). This implements not only a new conformation for the substrate, but also for the active site in the catalysis relevant GGDEF domain dimer.

In comparison to the extensive coordination of β- and γ-phosphate moieties and the predefined positions of both Mg^{2+}-ions by the coordinating residues D327 and E370 (Figure 2A), the interaction of the ribose and the guanine moieties with the protein is less elaborated. The main contact is established *via* N335 and D344, which are part of the helices α1 and α2.

In NCs, homolog helices show large flexibility and they are involved in dimerization (32,57,73), in the mechanism of induced fit of the active site (72,74,75), and in the activation process by interaction with the $G_S\alpha$ subunit of G-proteins (58,76). Similarly, in POLs the spatially corresponding "O"-helix is responsible for the transition between an "open" and a "closed" (=catalytically relevant) conformations of the active site (59,77). These conformational rearrangements in the active sites of NCs and POLs observed in the substrate analog bound state led to postulation of a substrate induced "active site closure"-mechanism (57-59). Structural comparison of non-ligated GGDEF domain structures, WspR (50) and NP_951600.1, with the ligated structure of PleD-act1, shows minor delocalization (1-2 Å) of helices α1 and α2 with respect to the well superposing β-sheet core as a rigid body (Supplemental figure 3). Interestingly, the alternative position of α2 doesn't induce changes in positions of the catalytically relevant residues. Side chains of N335 and D344 occupy almost identical positions in all three structures. Therefore, it seems that substrate binding doesn't induce conformational changes in the active site of the GGDEF domains.

An alternative mechanism of "active site closure", which is driven by catalysis, was proposed for NCs (74,75). It is inspired by the fact that some structures of substrate analog bound ACs show an "open" conformation while some ligand-free structures in complex with activators exhibit a "closed" one (58).

How probable is utilization of a similar mechanism by the GGDEF domain proteins, which form a single, dimeric active site to be catalytically competent? Several facts suggest dimerization induced conformational changes in the active site. (I) Although

the GGDEF domains possess all necessary residues for the intramolecular cyclization reaction (the reaction that is performed by NCs), there are no reports of a DGC producing cGMP as side product. (II) The substrate analog GTPαS bound to the active site of PleD (13) cannot adopt the productive state conformation without loosing its coordinating interactions to the protein. (III) The α1/α2 helices of the GGDEF domain, which similarly to the NCs are most probably involved in dimerization of the catalytic domains, exhibit some degree of flexibility. Moreover conformational changes in this region would result in an alternate position for the guanine moiety of the substrate.

Modeling of a processive substrate-GGDEF domain complex

Although Diadenylate Cyclases (DACs) (78) and DGCs do not share any structural similarities the products of both enzymes, c-di-AMP (79) and c-di-GMP, differ only in the kind of the nucleobase. Both products utilize the "U"-shape conformation, where the parallel nucleobases are perpendicularly arranged to the ribosyl-phosphate macrocycle. DisA was crystallized in presence of both, its product c-di-AMP and the substrate analog 3'-deoxy-ATP bound to the active site, respectively (78). Perfect superposition between the 2-fold symmetrically arranged substrate analog molecules and the product, c-di-AMP, are observed. This means that the conformation of c-di-AMP is pre-coded by the orientation of both substrate molecules with respect to each other and that c-di-AMP experiences no structural changes post-catalytically.

The catalytic mechanism of DACs is distinct from the 'two-metal ion assisted' catalysis due to the absence of the putative metal ions as well as the putative residues coordinating them. Therefore, it is clear that proper geometry of the substrates required for the nucleophilic in-line attack in DGCs will differ from the one in DACs. Nevertheless the global, pre-catalytic conformation and the intramolecular orientation of both substrate molecules in DGCs and DACs must be very similar to generate products with identical conformation.

We have fitted both 3'deoxy-ATP molecules as a single rigid body in the active site of PleD. Since the triphosphate moieties of these ligands don't show the predicted ideal conformation for the "two-metal assisted catalysis" (Figure 2D), restrains had to be imposed in symmetrical manner for both nucleotides. The phosphates-moiety of the "cis" nucleotide (nucleotide that forms direct contacts with the GGDEF domain) was aligned in coordination range to metal ion B. Additionally, the α-phosphate's

non-bridging oxygen atoms were restrained to coordinate both metal ions symmetrically. The imaginary 3'OH-group (missing in this substrate analog) of the "trans" nucleotide was aligned to be in coordination distance to metal ion A, which activates this group for the nucleophilic attack.

Our model doesn't show any prominent clashes between the nucleotides and the protein, but two prominent features (Figure 2E). The first feature is the distinct position of the nucleoside moiety in the model and the PleD-act1 structure. Atom O6 of the nucleobase comes hereby in contact distance with the highly conserved R366, as is observed for the c-di-GMP molecule in the PleD-act2 structure. Additionally, the C5' atoms of the fitted C5'PCPP and the 3'deoxy-ATP moieties take an almost identical position (Figure 2D-E). Thus, the substrate/Mg^{2+}-ions loaded half-active sites undergo quite likely a conformational change upon dimerization of the GGDEF domains to allow proper substrate alignment for the activation of the 3'OH-hydroxyls and the subsequent in-line attack. As stated before the rather flexible helices α1 and α2 are most probably involved in this process.

The second feature of this model is the position of the ribose in "trans". Its 2'OH group is localized in H-bond distance to the carboxyl group of E371 in "cis" (Figure 2E), which is simultaneously close to metal ion A (~3.3 Å) (Figure 2A). For the in-line attack to occur, the "trans" 3'OH-hydroxyl has to be properly aligned with respect to the metal ion A and Pα in "cis", and vice versa. Actually, no other residues than E371 are in proximity to fulfill this task.

Taking into account that for the production of a c-di-GMP molecule two nucleophilic attacks have to be performed simultaneously to avoid production of unwanted intermediates and site-products, E371's possible 'guiding role' appears even more attractive. Alignment of the 3'OH-hydroxyls for the in-line attack *via* the "cis"-"trans" cross-interaction of both E371 with both 2'OH- and not 3'OH-groups, would allow only simultaneous nucleophilic attacks. Misalignment of one of the attacking groups would hereby result in inappropriate distances between each of the "trans" 3'OH-"cis" Pα groups and in inappropriate angles for the in-line attack (Figure 8A).

<u>Modeling of an active site forming GGDEF domain dimer</u>

Fitting of both substrate analogs in the active site of PleD's GGDEF domain allowed us to generate the missing two-fold symmetry related GGDEF domain (Figure 7). Lack of serious clashes in the model of the GGDEF domain dimer affirms our rigid

body fitting of the substrate analogs to the substrate binding site (Figure 2E). As expected for an activated PleD dimer forming the $(Rec1\text{-}Rec2)_2$ stem, the model shows a distance of less than 6 Å between the N-termini of both GGDEF domains. Additionally, the N-terminal β0-β0'-hairpins, being conserved only in DGCs (32), are arranged as a single β-strand. As it was proposed before, helices α1 and α2 are involved in dimer contact formation and it is tempting to envisage that closure of these helices around the active site will lead to contact formations with the highly conserved residues of the helix α4.

Hypothetical reaction scheme for DGCs

Our biochemical, structural, computational and comparative investigations in the activity of PleD let us propose a putative catalytic mechanism for PleD. The catalytic reaction starts with dimerization of the GTP/Mg^{2+} bound GGDEF domains, which is induced upon phosphorylation dependent dimerization of the $(Rec1\text{-}Rec2)_2$ stem. Dimerization of the GGDEF domains induces conformational changes in helices α1 and α2 with respect to the core β-sheet, resulting in rearrangements in the substrate binding site. This and the bridging of E371 between metal A and the 2'OH-group of the "trans"-substrate lead to a proper alignment of both 3'OH-groups for the nucleophilic in-line attack on the α-phosphorus atoms (Figure 8A). The occurring of the "double"-reaction action is probably quasi simultaneous.

As suggested earlier (13,24), the subsequently formed pentacovalent phosphate intermediate (Figure 8B) could be stabilized by K332, compensating for its negative charge. Most recently it was shown that positively charged residues (mainly lysines) in the "O"-helix of POLs are actively involved in the mechanism of pentacovalent phosphate intermediate protonation as general acids (80). Hereby, the chemical reaction of protonation turns out to be the rate limiting step for nucleotide incorporation by polymerases. The same role may be attributed to K332 in PleD. In the case of protonation reaction being equally the rate-limiting step in DGCs, formation of both pentacovalent intermediates would occur faster than the protonation reaction. This would additionally enhance the enzymatic precision of DGCs.

The protonation reaction leads to a breakdown of the intermediate into c-di-GMP and two pyrophosphate molecules. In this state the joint coordination of both metal ions by the cleaved phosphate is not feasible any longer due to charge repulsion between the products (Figure 8C). In DGCs charge repulsion between the products will most

probably force the dimeric active site to open up. Similarly to NCs (32,81) and POLs (43,59,60), it is believed that c-di-GMP will leave the active site first due to a stronger interaction of the pyrophosphate molecules with the active site. In POLs release of the α-phosphate bearing product is concomitant with release of metal ion A (43,59,60). The E371-Mg^{2+}-ion-phosphate bonding in PleD-act2 may represent such a state in PleD (Figure 2C). Further movement of c-di-GMP out of the active site, alike the observed situation in the structure of PleD-na (24), would lead to the release of the metal ion (Figure 8C).

Figure legends

Figure 1. Ribbon diagram of activated PleD dimer in complex with c-di-GMP

The domains Rec1, Rec2 and GGDEF are colored red, yellow and green, respectively. The GGDEF signature motif is highlighted in blue. Ligands and interacting residues are shown in full. (A) The main difference in both structures of activated PleD - rigid body repositioning of cross-linked GGDEF domains with respect to the stem is represented by the hinge as straight black line. Bound sulfate molecules are detected at both active sites, as well as in the two-fold symmetric $(Rec1\text{-}Rec2)_2$ dimer interface. Phosphate mimicking $BeF_3^- \bullet Mg^{2+}$ ligands are localized at both phosphorylation sites. Intercalated c-di-GMP dimers are found cross-linking the GGDEF domains, binding to both allosteric sites (13). A single c-di-GMP molecule interacts with the GGDEF active site and Rec domains of second protomer (B).

Figure 2. Substrate analog and product binding to the active site of PleD

The ribbon representation shows the protein in green and the GGDEF signature motif in dark blue. Ligands, interacting residues as well as the P loop main chain (328-331) are shown in full. (A) Substrate analog GTPαS and two putative magnesium ions (labeled A and B) bound to PleD-act1 (PDB code 2v0n). (B) Product c-di-GMP and a putative Mg^{2+} ion bound to PleD-act2. (C) C-di-GMP bound to PleD-na (PDB code 1w25). (D) C5'PCPP moiety of a AMPCPP molecule from the T7 RNAP structure (PDB code 1s76) is fitted in the active site of PleD-act1 and restrained with respect to the positions of both Mg^{2+} ions (see main text). (E) Both 3'deoxy-ATP molecules from the DisA structure (PDB code 3c23) are fitted as a rigid body in the active site of PleD-act1 and restrained (see main text). For clarity GTPαS is drawn in thinner lines.

Figure 3. Sulfate binding on the molecular dyad of the $(Rec1\text{-}Rec2)_2$ stem

Secondary structure elements, which are involved in dimerization and sulfate binding in the $(Rec1\text{-}Rec2)_2$ dyad interface, are shown as ribbon. Residues involved in sulfate binding and formation of the symmetric charged cleft are shown in full. Green surface around dyad-related L124 and V125 represents the hydrophobic contact between α5-helices (Rec1).

Figure 4. Conservation and charge distribution along dimer interface of the stem

The GGDEF domains are not shown for clarity. (A-B) View onto the $(Rec1\text{-}Rec2)_2$ stem along its two-fold symmetry axis, from the bottom of the view in figure 1 and 2. (A) Surface charge distribution, generated by APBS tools (82), in the stem of PleD-act1 (PDB code: 2v0n) shows BeF_3^-•Mg^{2+} binding to the acidic pocket of the phosphorylation sites, a sulfate deeply buried at positively charged dimer interface and a sulfate binding to an exposed, basic patch in Rec2'.

Mapping of residue conservation score; using the ConSurf web-server (83), onto the dimeric stem (B) and its dimer forming face (C) of PleD-act2 according conservation of residues in Rec1-Rec2-GGDEF proteins (magenta = strong, white = medium, blue = not conserved).

Figure 5. Characterization of PleD DGC activity in presence of tetrahedral oxyanions, sulfate and phosphate

The product quantity was normalized taking the c-di-GMP amount produced by BeF_3^-•Mg^{2+} activated PleD R148A R178A R313A as reference. For details see Materials&Methods.

Figure 6. Characterization of DgcB DGC activity in presence of divalent cations

Catalytic activity varying the concentrations of magnesium (full circles) and manganese (open triangles) and calcium (+ 2 mM Mg^{2+}) (full rectangles) is present. The product quantity was normalized according the c-di-GMP amount produced by DgcB in presence of 5 mM Mg^{2+}.

Figure 7. Superposition of GGDEF domains on fitted 3'deoxy-ATP molecules

The GGDEF domains are represented as ribbon (A) and surface (B) and are colored according their residue conservation (magenta = strong, white = medium, blue = not conserved). Secondary structure elements that are involved in dimerization are labeled. The GGDEF domains were superposed on the fitted substrate analogs (see main text), which together with catalytic metal ions are shown in full.

Figure 8. Schematic representation of the catalytic cycle of DGCs

(A) Prior to catalysis GTP and magnesium ions bind to the DGC active sites. Besides minor coordination of the nucleotides by its base (see main text), the ligand is interacting mainly via the triphosphate moiety with main chain amides of the P loop and the catalytic metal ions A and B. Metal ion B is octahedrally coordinated by the α-, β- and γ-phosphates of the nucleotide, the carboyl groups of D327 and E370 as well as main chain carbonyl 328. The role of Mg^{2+} (B) is proper positioning of the triphosphate moiety for the reaction and neutralization of charge during catalysis. The second Mg^{2+} (metal ion A), which is coordinated by α-phosphate and carboyl groups of D327 and E370, initiates the reaction by lowering the pK_a values of the 3'-hydroxyl groups. The proton acceptor is not known, but it is probably a water molecule. For the enzymatic reaction to happen, both nucleotides of both active sites have to be properly aligned. First, the non-bridging oxygen of Pα has to form symmetrical contacts to both metal ions, to bring the attacking group in contact range of metal ion A. Secondly, the attacking 3'OH-group must be in-line with the scissile Pα-O bond. Upon encounter of both substrate loaded active sites the ribosyl 2'OH-group probably interacts with E371 of the other GGDEF domain, aligning the attacking group for the in-line reaction. Proper positioning of the attacking ribosyl via 2'OH would result in high product precision of DGCs. Only when both E371-2'OH contacts are formed, then the attacking groups have the proper distance and are in-line for the nucleophilic attacks, which then happen simultaneously. (B) We propose that the charge of the pentacoordinated intermediate is stabilized by K332, which subsequently protonates the leaving pyrophosphate group. (C) Product release may be initiated by opening of the GGDEF domain dimers, followed by E371-metal ion A guided dissociation of the c-di-GMP molecule.

Figure S1. Conservation score of R117 and R121

Conservation score of R117 and R121 in Rec1-Rec2-GGDEF proteins generated by the web based application WEBLOGO (84).

Figure S2. Surface mapping of residues contributing to the Rec-HK interaction

Residues of PleD, which could be involved in interaction with its cognate histidine kinases, are mapped black on the surface of the stem (same view as in figure 3 and color code as in figure 1). These residues correspond to homologous residues of RR

Spo0F, which are involved in interaction with HPt of Spo0B (40). Mapping was performed on all four Rec domains, since many Rec1-Rec2-GGDEF domain proteins utilize phosphorylation of both Rec domains (see main text). Rec domains contribute mainly via α1-helix and the phosphorylation site surrounding βα-loops to the HK-Rec interface.

Figure S3. Superposition of GGDEF domains of PleD, WspR and NP_951600.1

The GGDEF domains of PleD, WspR and NP_951600.1 (PDB codes: 2v0n, 3bre and 3ezu) were aligned superposing Cα-atoms of the central β-sheet and are shown as ribbon in green, black and magenta, respectively. N335, D344, E370 and E371 of PleD and homologous residues of WspR and NP_951600.1 are shown in full. For better understanding ligands of PleD – the GTPαS molecule and the two putative magnesium ions are included.

	PleD (2WB4.PDB)
Data collection	
Space group	$P2_12_12$
Cell dimensions:	
a, b, c (Å)	123.56, 127.43, 88.13
α, β, γ (°)	90, 90, 90
Resolution (Å)	70.0 - 2.5 (2.64 – 2.5)*
R_{merge}	10.2 (28.5)
$I/\sigma I$	12.9 (5.3)
Completeness (%)	87.0 (62.4), [95.1 (81.6)]**
Redundancy	3.4 (3.3)
Refinement	
Resolution (Å)	2.8
No. reflections	31415
R_{work}/ R_{free}	24.0 / 26.8
No. Atoms	
Protein	6908
Ligand/ion	256
Water	50
B-factors (Å^2)	
Protein	17.5
Ligand/ion	12.2
Water	6.37
R.m.s. deviations	
Bond lengths (Å)	0.015
Bond angles (°)	1.6
* Highest resolution shell is shown in parentheses. ** Second pair of values represents completeness to a resolution of 2.8Å	

Table 1: Data collection and refinement statistics

Residue function	PleD residues	stand-alone Rec1-Rec2 proteins	Rec1-Rec2-GGDEF proteins
phosphoacceptor (Rec1)	D53	25% (D)	93% (D)
"Y-T"-coupling (Rec1)	T83	73% (T/S)	95% (T/S)
	F102	26% (F/Y)	95% (F/Y)
phosphoacceptor (Rec2)	N202	94% (D)	72% (D)
"Y-T"-coupling (Rec2)	V232	88% (S/T)	65% (S/T)
	I251	31% (F/Y)	89% (F/Y)
dimerization	Y26	np	34% (Y)
charged cavity formation in the $(Rec1/Rec2)_2$ stem interface	R117	27% (RK)	47% (RK)
	R121	5% (R)	63% (R)
	E25	25% (E)	32% (E)
	E245	5% (E/D); 9% (K)	31% (E/D); 39% (K)

Table 2: Conservation of Rec domain residues in PleD homolog proteins

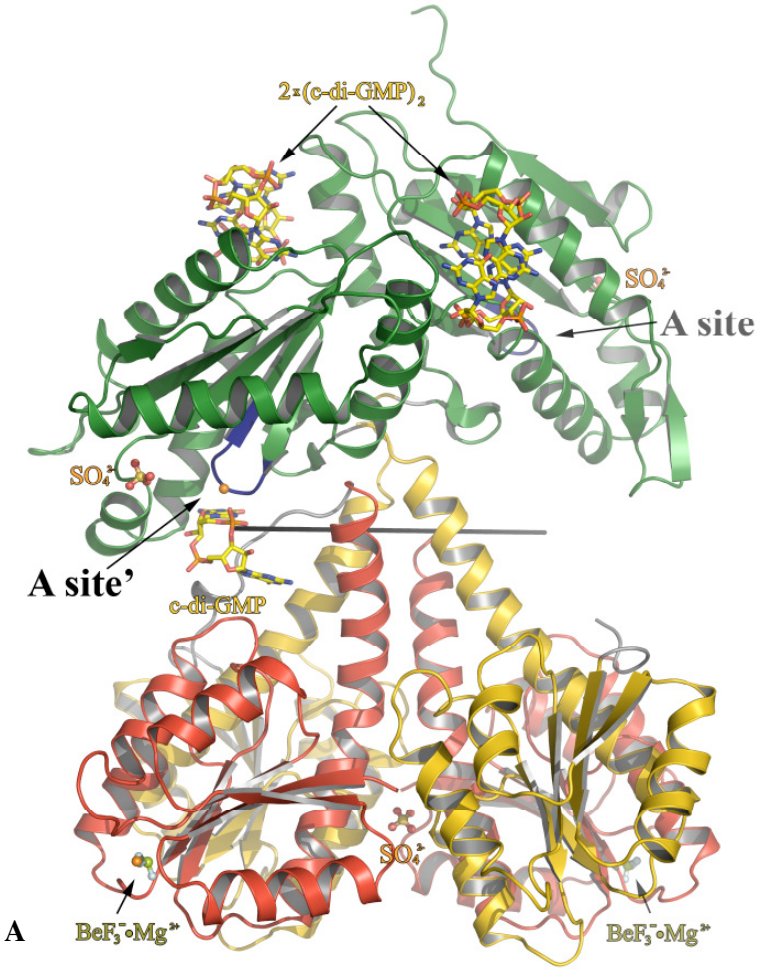

2x(c-di-GMP)2
SO4 2-
A site
SO4 2-
A site'
c-di-GMP
SO4 2-
BeF3-•Mg2+
BeF3-•Mg2+
A
E371
K442
R366
328
D344
R446
Mg2+
SO4 2-
N335
R148
A145
R137
c-di-GMP
P72
B

Figure 1

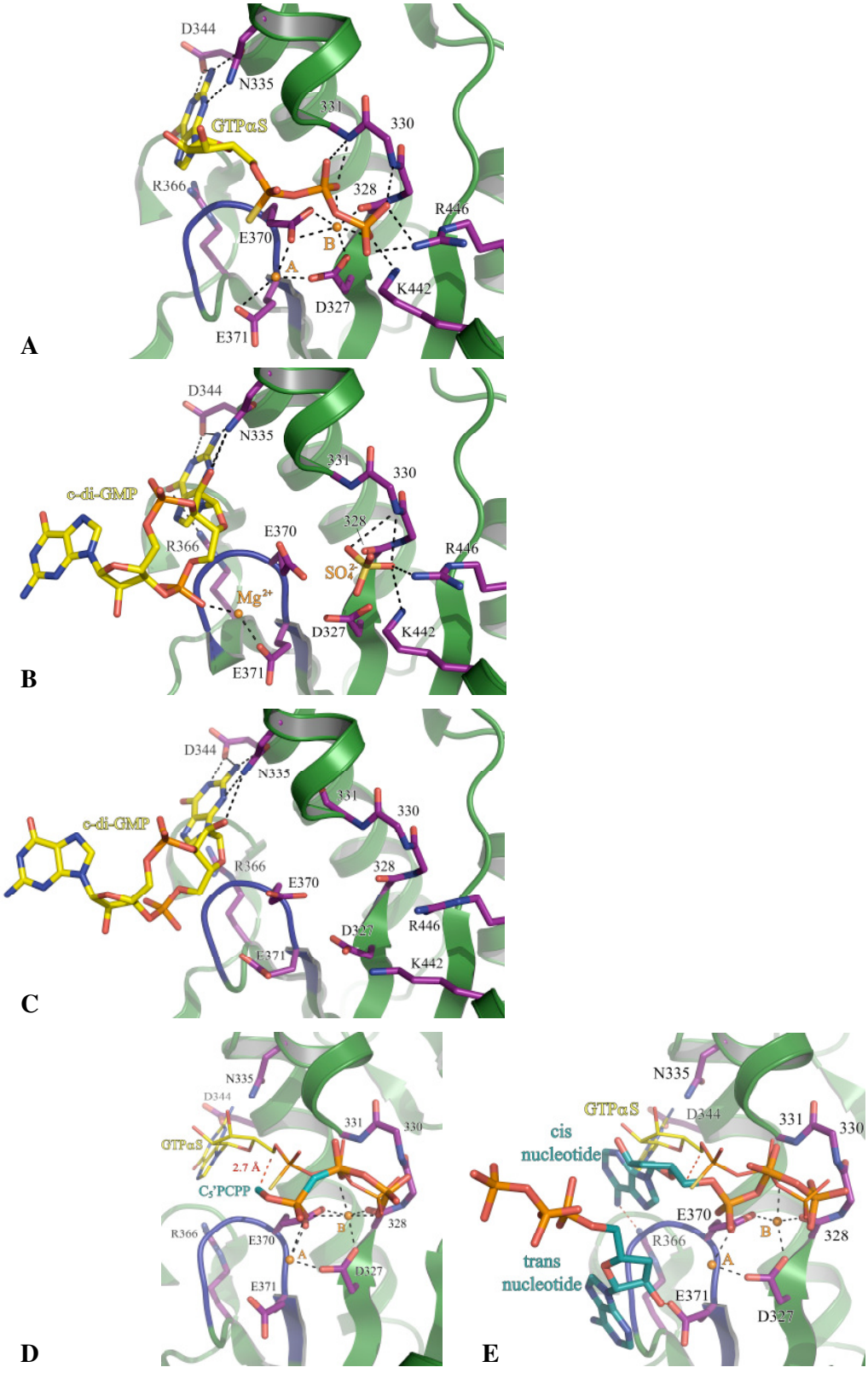

A
D344
N335
331
330
GTPαS
R366
328
R446
E370
B
A
K442
D327
E371
B
D344
N335
331
330
c-di-GMP
E370
328
R366
R446
SO4 2-
Mg2+
D327
K442
E371
C
D344
N335
331
330
c-di-GMP
R366
E370
328
D327
R446
E371
K442
D
N335
D344
331
330
GTPαS
2.7 Å
C3'PCPP
R366
E370
B
328
A
D327
E371
E
N335
GTPαS
D344
331
330
cis
nucleotide
E370
B
328
R366
A
trans
nucleotide
E371
D327

Figure 2

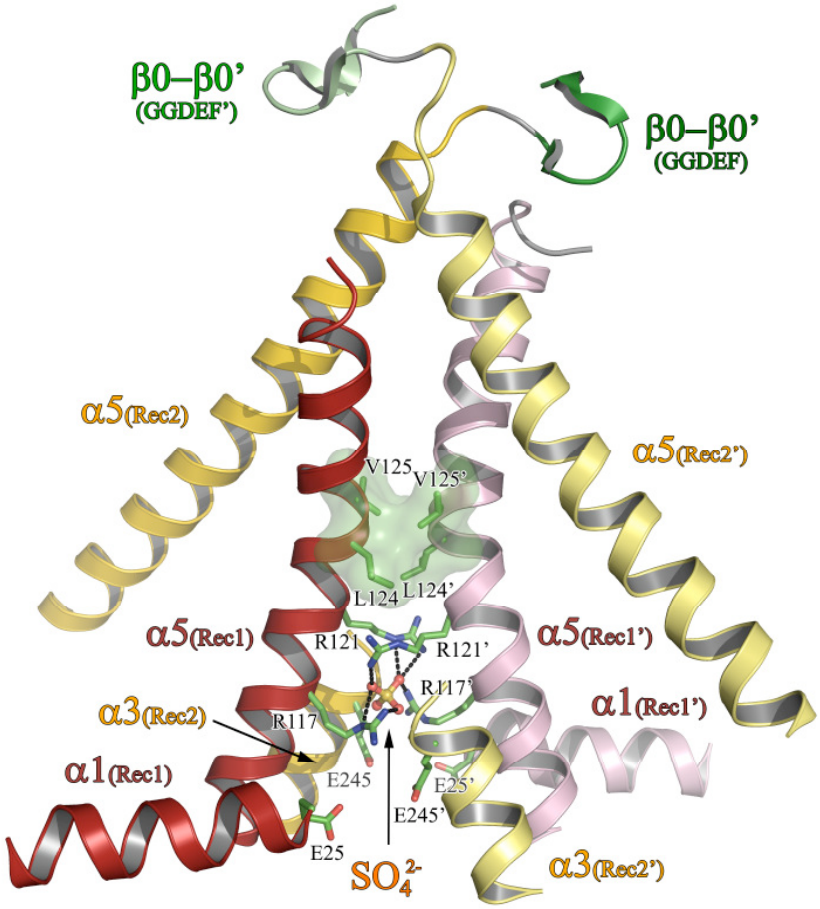

β0–β0'
(GGDEF')
β0–β0'
(GGDEF)
α5(Rec2)
α5(Rec2')
V125
V125'
L124
L124'
α5(Rec1)
α5(Rec1')
R121
R121'
R117'
α3(Rec2)
α1(Rec1')
R117
α1(Rec1)
E245
E25'
E245'
E25
SO4 2-
α3(Rec2')

Figure 3

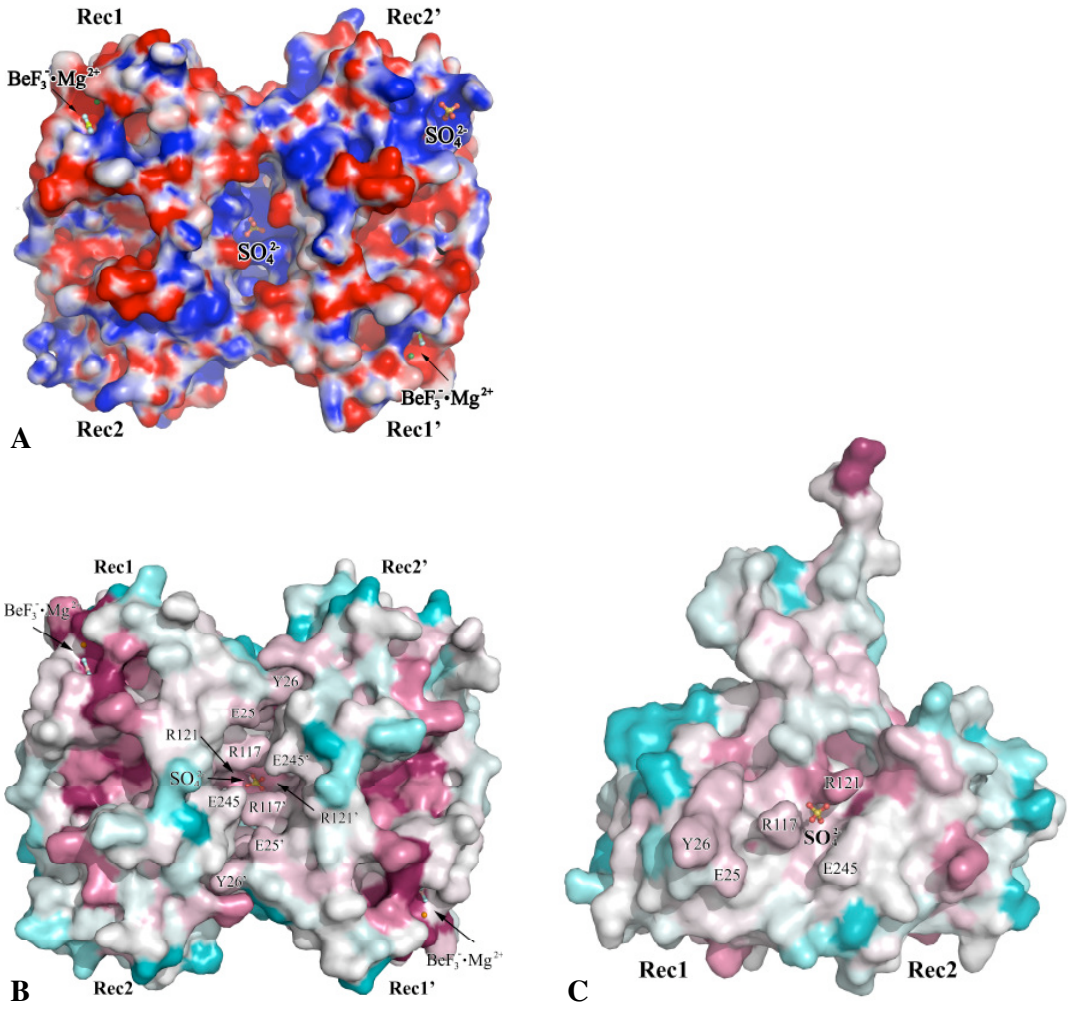

Rec1
Rec2'
BeF_3^-·Mg^{2+}
SO_4^{2-}
SO_4^{2-}
BeF_3^-·Mg^{2+}
Rec2
Rec1'
A
Rec1
Rec2'
BeF_3^-·Mg^{2+}
Y26
E25
R121
R117
E245'
SO_4^{2-}
E245
R117'
R121'
E25'
Y26'
BeF_3^-·Mg^{2+}
Rec2
Rec1'
B
R121
R117
SO_4^{2-}
Y26
E25
E245
Rec1
Rec2
C

Figure 4

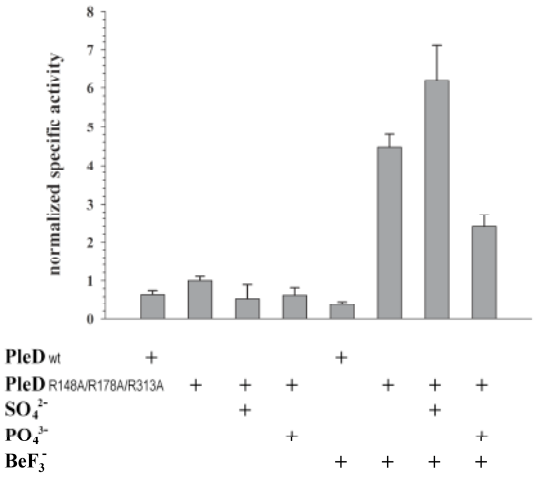


Figure 5

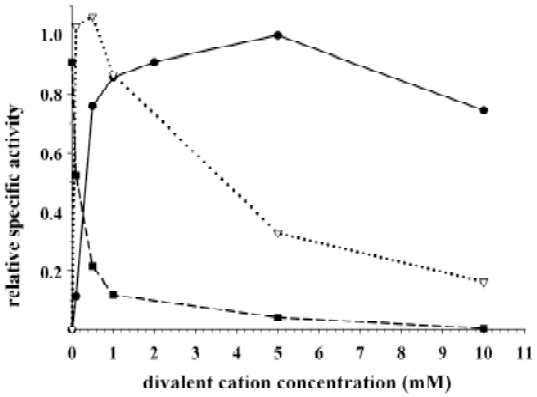


Figure 6

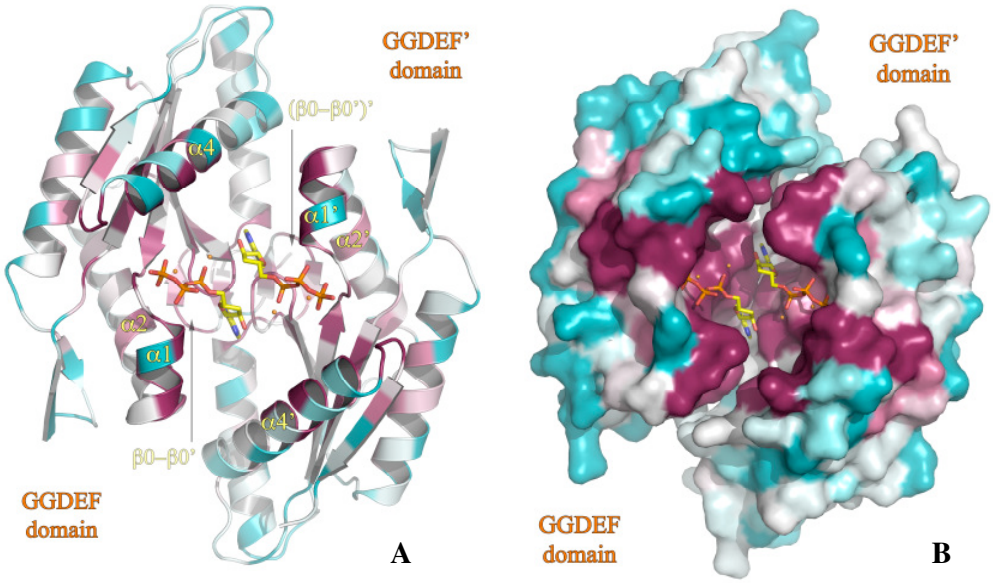


Figure 7

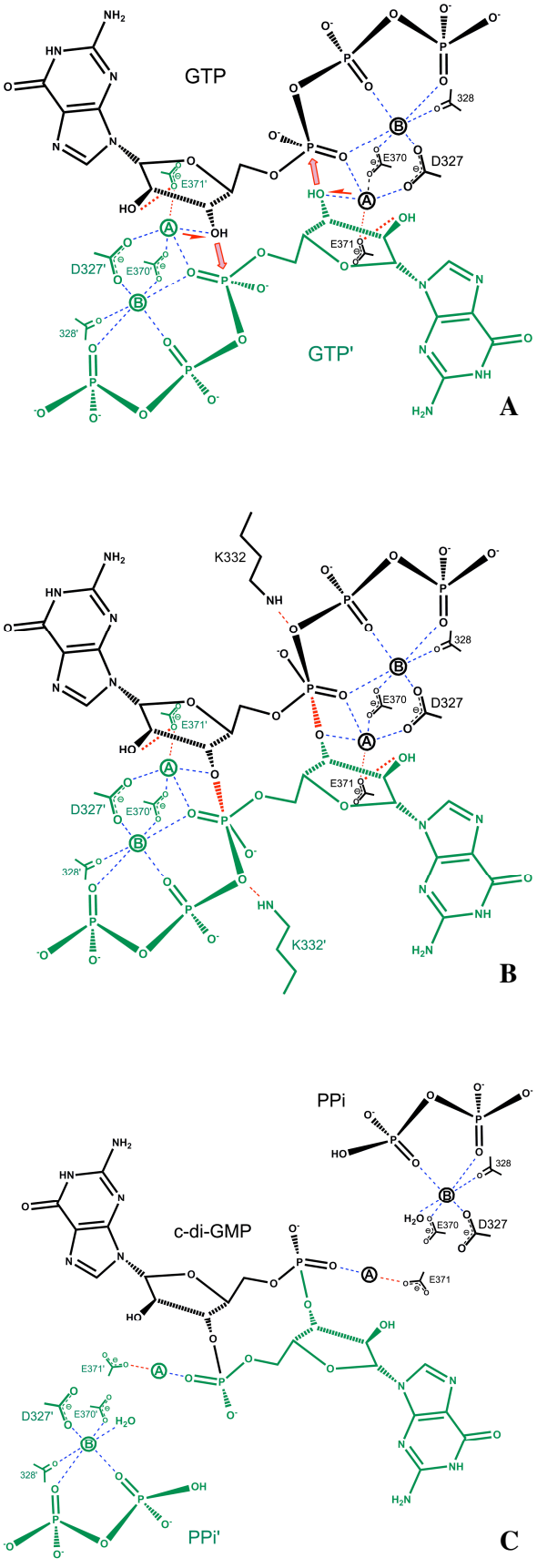
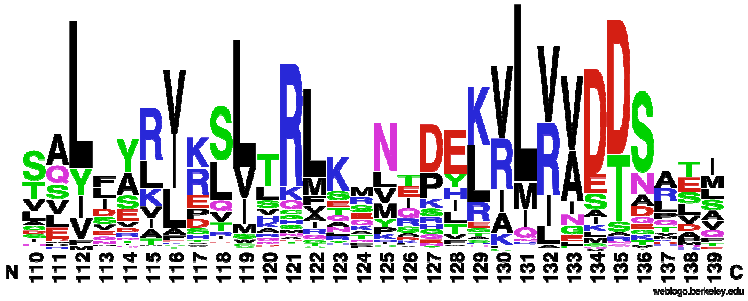
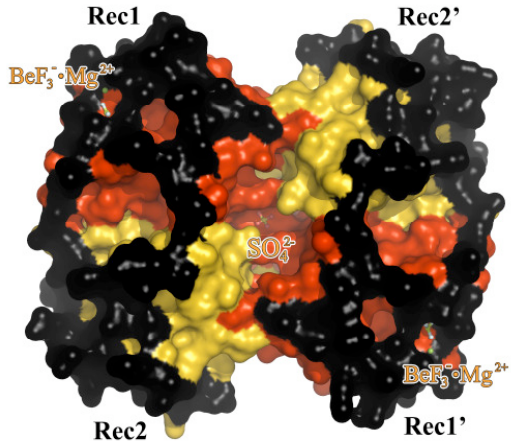

GTP
GTP'
D327
D327'
E370
E370'
E371
E371'
328
328'
A
K332
K332'
B
PPi
PPi'
c-di-GMP
C

Figure 8



Supplemental figure 1



Supplemental figure 2

Supplemental figure 3

Bibliography

1. Beier, D., and Gross, R. (2006) *Curr Opin Microbiol* **9**, 143-152
2. Skerker, J. M., Prasol, M. S., Perchuk, B. S., Biondi, E. G., and Laub, M. T. (2005) *PLoS Biol* **3**, e334
3. Gao, R., Mack, T. R., and Stock, A. M. (2007) *Trends Biochem Sci* **32**, 225-234
4. Wolanin, P. M., Thomason, P. A., and Stock, J. B. (2002) *Genome Biol* **3**, REVIEWS3013
5. Stock, A. M., Robinson, V. L., and Goudreau, P. N. (2000) *Annu Rev Biochem* **69**, 183-215
6. Janiak-Spens, F., Sparling, D. P., and West, A. H. (2000) *J Bacteriol* **182**, 6673-6678
7. Zhao, X., Copeland, D. M., Soares, A. S., and West, A. H. (2008) *J Mol Biol* **375**, 1141-1151
8. Marina, A., Waldburger, C. D., and Hendrickson, W. A. (2005) *Embo J* **24**, 4247-4259
9. Varughese, K. I., Tsigelny, I., and Zhao, H. (2006) *J Bacteriol* **188**, 4970-4977
10. Djordjevic, S., Goudreau, P. N., Xu, Q., Stock, A. M., and West, A. H. (1998) *Proc Natl Acad Sci U S A* **95**, 1381-1386
11. Maris, A. E., Sawaya, M. R., Kaczor-Grzeskowiak, M., Jarvis, M. R., Bearson, S. M., Kopka, M. L., Schroder, I., Gunsalus, R. P., and Dickerson, R. E. (2002) *Nat Struct Biol* **9**, 771-778
12. De Carlo, S., Chen, B., Hoover, T. R., Kondrashkina, E., Nogales, E., and Nixon, B. T. (2006) *Genes Dev* **20**, 1485-1495
13. Wassmann, P., Chan, C., Paul, R., Beck, A., Heerklotz, H., Jenal, U., and Schirmer, T. (2007) *Structure* **15**, 915-927
14. Toro-Roman, A., Wu, T., and Stock, A. M. (2005) *Protein Sci* **14**, 3077-3088
15. Galperin, M. Y. (2006) *J Bacteriol* **188**, 4169-4182
16. Hecht, G. B., and Newton, A. (1995) *J Bacteriol* **177**, 6223-6229
17. Aldridge, P., and Jenal, U. (1999) *Mol Microbiol* **32**, 379-391
18. Aldridge, P., Paul, R., Goymer, P., Rainey, P., and Jenal, U. (2003) *Mol Microbiol* **47**, 1695-1708
19. Levi, A., and Jenal, U. (2006) *J Bacteriol* **188**, 5315-5318
20. Paul, R., Weiser, S., Amiot, N. C., Chan, C., Schirmer, T., Giese, B., and Jenal, U. (2004) *Genes Dev* **18**, 715-727
21. Paul, R., Jaeger, T., Abel, S., Wiederkehr, I., Folcher, M., Biondi, E. G., Laub, M. T., and Jenal, U. (2008) *Cell* **133**, 452-461
22. Matroule, J. Y., Lam, H., Burnette, D. T., and Jacobs-Wagner, C. (2004) *Cell* **118**, 579-590
23. Paul, R., Abel, S., Wassmann, P., Beck, A., Heerklotz, H., and Jenal, U. (2007) *J Biol Chem* **282**, 29170-29177
24. Chan, C., Paul, R., Samoray, D., Amiot, N. C., Giese, B., Jenal, U., and Schirmer, T. (2004) *Proc Natl Acad Sci U S A* **101**, 17084-17089
25. Ryjenkov, D. A., Tarutina, M., Moskvin, O. V., and Gomelsky, M. (2005) *J Bacteriol* **187**, 1792-1798
26. Jenal, U., and Malone, J. (2006) *Annu Rev Genet* **40**, 385-407
27. Romling, U., and Amikam, D. (2006) *Curr Opin Microbiol* **9**, 218-228
28. Ryan, R. P., Fouhy, Y., Lucey, J. F., and Dow, J. M. (2006) *J Bacteriol* **188**, 8327-8334

29. Cotter, P. A., and Stibitz, S. (2007) *Curr Opin Microbiol* **10**, 17-23
30. Pei, J., and Grishin, N. V. (2001) *Proteins* **42**, 210-216
31. Steitz, T. A., and Steitz, J. A. (1993) *Proc Natl Acad Sci U S A* **90**, 6498-6502
32. Sinha, S. C., and Sprang, S. R. (2006) *Rev Physiol Biochem Pharmacol* **157**, 105-140
33. (1994) *Acta Crystallogr D Biol Crystallogr* **50**, 760-763
34. McCoy, A. J. (2007) *Acta Crystallogr D Biol Crystallogr* **63**, 32-41
35. Emsley, P., and Cowtan, K. (2004) *Acta Crystallogr D Biol Crystallogr* **60**, 2126-2132
36. Wright, M., Tanner, J. A., and Miller, A. D. (2003) *Anal Biochem* **316**, 135-138
37. Krissinel, E., and Henrick, K. (2007) *J Mol Biol* **372**, 774-797
38. Christen, B., Christen, M., Paul, R., Schmid, F., Folcher, M., Jenoe, P., Meuwly, M., and Jenal, U. (2006) *J Biol Chem* **281**, 32015-32024
39. Letunic, I., Copley, R. R., Schmidt, S., Ciccarelli, F. D., Doerks, T., Schultz, J., Ponting, C. P., and Bork, P. (2004) *Nucleic Acids Res* **32**, D142-144
40. Zapf, J., Sen, U., Madhusudan, Hoch, J. A., and Varughese, K. I. (2000) *Structure* **8**, 851-862
41. Xu, Q., Porter, S. W., and West, A. H. (2003) *Structure* **11**, 1569-1581
42. Rauch, A., Leipelt, M., Russwurm, M., and Steegborn, C. (2008) *Proc Natl Acad Sci U S A* **105**, 15720-15725
43. Vaisman, A., Ling, H., Woodgate, R., and Yang, W. (2005) *Embo J* **24**, 2957-2967
44. Nowotny, M., and Yang, W. (2006) *Embo J* **25**, 1924-1933
45. Gangurde, R., Kaushik, N., Singh, K., and Modak, M. J. (2000) *J Biol Chem* **275**, 19685-19692
46. Nowotny, M., Gaidamakov, S. A., Crouch, R. J., and Yang, W. (2005) *Cell* **121**, 1005-1016
47. Keck, J. L., Goedken, E. R., and Marqusee, S. (1998) *J Biol Chem* **273**, 34128-34133
48. Allingham, J. S., and Haniford, D. B. (2002) *J Mol Biol* **319**, 53-65
49. Malone, J. G., Williams, R., Christen, M., Jenal, U., Spiers, A. J., and Rainey, P. B. (2007) *Microbiology* **153**, 980-994
50. De, N., Pirruccello, M., Krasteva, P. V., Bae, N., Raghavan, R. V., and Sondermann, H. (2008) *PLoS Biol* **6**, e67
51. Volkman, B. F., Lipson, D., Wemmer, D. E., and Kern, D. (2001) *Science* **291**, 2429-2433
52. Nohaile, M., Kern, D., Wemmer, D., Stedman, K., and Kustu, S. (1997) *J Mol Biol* **273**, 299-316
53. Zapf, J., Madhusudan, M., Grimshaw, C. E., Hoch, J. A., Varughese, K. I., and Whiteley, J. M. (1998) *Biochemistry* **37**, 7725-7732
54. Copley, R. R., and Barton, G. J. (1994) *J Mol Biol* **242**, 321-329
55. Quiocho, F. A. (1996) *Kidney Int* **49**, 943-946
56. Hirsch, A. K., Fischer, F. R., and Diederich, F. (2007) *Angew Chem Int Ed Engl* **46**, 338-352
57. Sinha, S. C., Wetterer, M., Sprang, S. R., Schultz, J. E., and Linder, J. U. (2005) *Embo J* **24**, 663-673
58. Tesmer, J. J., Sunahara, R. K., Johnson, R. A., Gosselin, G., Gilman, A. G., and Sprang, S. R. (1999) *Science* **285**, 756-760
59. Yin, Y. W., and Steitz, T. A. (2004) *Cell* **116**, 393-404
60. Yang, W., Lee, J. Y., and Nowotny, M. (2006) *Mol Cell* **22**, 5-13

61. Cooper, D. M. (2003) *Biochem J* **375**, 517-529
62. Savilahti, H., Rice, P. A., and Mizuuchi, K. (1995) *Embo J* **14**, 4893-4903
63. Vipond, I. B., Baldwin, G. S., and Halford, S. E. (1995) *Biochemistry* **34**, 697-704
64. Litvin, T. N., Kamenetsky, M., Zarifyan, A., Buck, J., and Levin, L. R. (2003) *J Biol Chem* **278**, 15922-15926
65. Ling, H., Boudsocq, F., Woodgate, R., and Yang, W. (2004) *Mol Cell* **13**, 751-762
66. Pauling, L. (1947) *J. Am. Chem. Soc.* **69**, 542-553
67. Gangurde, R., and Modak, M. J. (2002) *Biochemistry* **41**, 14552-14559
68. Steitz, T. A. (1998) *Nature* **391**, 231-232
69. Frey, P. A., and Sammons, R. D. (1985) *Science* **228**, 541-545
70. Eckstein, F. (1985) *Annu Rev Biochem* **54**, 367-402
71. Haruki, M., Tsunaka, Y., Morikawa, M., Iwai, S., and Kanaya, S. (2000) *Biochemistry* **39**, 13939-13944
72. Steegborn, C., Litvin, T. N., Levin, L. R., Buck, J., and Wu, H. (2005) *Nat Struct Mol Biol* **12**, 32-37
73. Tews, I., Findeisen, F., Sinning, I., Schultz, A., Schultz, J. E., and Linder, J. U. (2005) *Science* **308**, 1020-1023
74. Kamenetsky, M., Middelhaufe, S., Bank, E. M., Levin, L. R., Buck, J., and Steegborn, C. (2006) *J Mol Biol* **362**, 623-639
75. Steegborn, C., Litvin, T. N., Hess, K. C., Capper, A. B., Taussig, R., Buck, J., Levin, L. R., and Wu, H. (2005) *J Biol Chem* **280**, 31754-31759
76. Tesmer, J. J., Sunahara, R. K., Gilman, A. G., and Sprang, S. R. (1997) *Science* **278**, 1907-1916
77. Doublie, S., and Ellenberger, T. (1998) *Curr Opin Struct Biol* **8**, 704-712
78. Witte, G., Hartung, S., Buttner, K., and Hopfner, K. P. (2008) *Mol Cell* **30**, 167-178
79. Romling, U. (2008) *Sci Signal* **1**, pe39
80. Castro, C., Smidansky, E. D., Arnold, J. J., Maksimchuk, K. R., Moustafa, I., Uchida, A., Gotte, M., Konigsberg, W., and Cameron, C. E. (2009) *Nat Struct Mol Biol* **16**, 212-218
81. Linder, J. U. (2006) *Cell Mol Life Sci* **63**, 1736-1751
82. Baker, N. A., Sept, D., Joseph, S., Holst, M. J., and McCammon, J. A. (2001) *Proc Natl Acad Sci U S A* **98**, 10037-10041
83. Landau, M., Mayrose, I., Rosenberg, Y., Glaser, F., Martz, E., Pupko, T., and Ben-Tal, N. (2005) *Nucleic Acids Res* **33**, W299-302
84. Crooks, G. E., Hon, G., Chandonia, J. M., and Brenner, S. E. (2004) *Genome Res* **14**, 1188-1190

II.3 Biochemical and biophysical analysis of c-di-GMP dependent inhibition of the diguanylate cyclase PleD

Paul Wassmann[1], Eric Kusznir[2], Francis Mueller[2] and Tilman Schirmer[1*]

[1] Core program of Structural Biology & Biophysics, University of Basel, CH-4056 Basel, Switzerland

[2] Department of Molecular Structure Research, F. Hoffmann-La Roche AG, CH-4070 Basel, Switzerland

[*] Corresponding author. E-mail address: tilman.schirmer@unibas.ch

Keywords: diguanylate cyclase, GGDEF domain, c-di-GMP, regulatory mechanism.

Abbreviations: c-di-GMP, bis-(3'-5')-cyclic guanosine monophosphate; DGC, diguanylate cyclase; PDE, phosphodiesterase; RR, response regulator; Rec, respone regulator receiver domain; CD, circular dichroism; SE AUC, sedimentation equilibrium analytical ultracentrifugation; SPR, surface plasmon resonance.

Abstract

The diguanylate cyclase PleD, which produces the bacterial second messenger c-di-GMP by a productive GGDEF domain dimer, is known to be tightly regulated. One of the regulatory mechanisms is the phosphorylation dependent dimerization of PleD *via* its N-terminal response regulator domains. Furthermore, PleD's activity is controlled by inhibition, binding allosterically its product c-di-GMP.

Enzymatic and biophysical techniques (SPR, AUC and CD) were used to shed light on details of the allosteric product inhibition on the molecular level. Our SPR experiments confirm binding of c-di-GMP to the allosteric site in the dimeric form. The binding itself depends strongly on the intact primary inhibition site (I_P-site: RxxD motif and R390) and to a lesser amount on the secondary I-sites ($I_{S,DGC}$: R313; $I_{S,D2}$: R148, R178). Importance of the I_S-sites in the c-di-GMP dependent domain-cross-linking was shown by thermal unfolding of PleD. Additionaly, in combination with previously published results our new I-site mutants show unambiguously that binding of c-di-GMP to the I_P alone is insufficient for the inhibition. Therefore, all these

experiments clearly prove the previously proposed “inhibition by domain immobilization” mechanism.

Introduction

The dawning of the bis-(3'-5')-cyclic dimeric guanosine monophosphate (c-di-GMP) as a global second messenger in bacteria was postponed for almost twenty years since the day of its discovery (1). The ubiquitous bacterial second messenger is nowadays known to be involved in processes like cell cycle progression (2), cell differentiation (3), multicellular behavior (4) as well as interaction between bacteria and their eukaryotic hosts (5,6). The main task of c-di-GMP, hereby, seems to be switching the bacterial faith between two lifestyles: the sessile, environment-persistent, surface-attached one on the one hand and the motile, often virulent one on the other. The inverse regulation of biofilm formation and virulence was shown by knock-outs of the c-di-GMP degrading phophodiesterase VieA in *Vibrio cholerae* (7,8).

Although c-di-GMP takes part in such a high variety of cellular processes, the downstream targets of the second messenger are only starting to emerge, e.g. the PiLZ domain proteins (9) or riboswitches (10). The identification of the proteins involved in biosynthesis and breakdown of c-di-GMP turned out to be less problematic. This task is performed by diguanylate cyclases (DGCs) and phosphodiesterases (PDEs) containing enzymatically active GGDEF and EAL/HD-GYP domains, respectively. The domains are named after the conserved consensus sequences in their active sites. The interesting aspect, hereby, is the high number of these proteins in the genomes of single bacterial species, peaking in the genome of *Vibrio vulnificus* coding for nearly 100 GGDEF and EAL domain proteins.

In order to produce a c-di-GMP molecule the DGCs consume two molecules of GTP. The PleD from *C. crescentus* (11) represents one of the best studied DGCs. The PleD's DGC activity is strongly regulated by two distinct mechanisms – the N-terminal, regulatory domain mediated dimerization and the allosteric product inhibition by c-di-GMP (12-14).

Crystal structures of PleD (12,14) have shown binding-capacity for only one GTP molecule per a GGDEF domain. Subsequent experiments have shown the necessity of the dimeric form for DGCs to be able to produce c-di-GMP (15). Since the GGDEF domains are unable to dimerize on their own (P. Wassmann unpublished data), the DGCs use additional N-terminal domains for this task. Often, these dimerization domains exhibit regulatory functions. The GGDEF domain proteins are frequently found in combination with GAF, PAS, Rec and other regulatory domains. PleD

consists besides the C-terminal GGDEF domain of one phosphorylatable and one degenerate response regulator domain (Rec1 and Rec2). Upon phosphorylation by its cognate His-kinases PleC and DivJ (16) PleD dimerizes *via* its Rec domains forming a $(Rec1\text{-}Rec2)_2$ stem (14). In the activated state the flexible GGDEF domains form a joint active site that is able to produce c-di-GMP.

Inhibition of PleD by allosterically binding c-di-GMP represents another regulatory mechanism. The crystal structures of PleD (12,14) have shown a c-di-GMP dimer bound to an allosteric site of the GGDEF domain (primary inhibition site = I_P). In case of PleD the I_P consists of the RxxD motif and arginine 390. Moreover, the c-di-GMP dimer cross-links the GGDEF domain either with the Rec2 domain of the same protomer or with the GGDEF domain of the second protomer in a PleD dimer ((14)–Figure 7). Both arginines, R148 and R178, of the Rec2 domain interacting with c-di-GMP are called the secondary inhibition site $I_{S,D2}$. And arginine, R313, in the GGDEF domain, which is involved in the cross-link mechanism, is analogously called $I_{S,DGC}$. Due to a high conservation score of the RxxD motif, its presence in the sequences of DGCs was proposed to be an indicator for c-di-GMP dependent regulation in these DGCs (13).

Combining biochemical and biophysical approaches studying PleD, we sought a closer view on these regulatory mechanisms. We were particularly interested in the following questions. Is the sole presence of the RxxD motif sufficient to bind and to inhibit DGCs, e.g. PleD? What are the details of the allosteric binding process of c-di-GMP to PleD concerning the binding stoichiometry as well as the order of the binding events? And which functions do the secondary inhibition sites (I_S-sites) as well as R390 have in the c-di-GMP dependent inhibition process, exactly?

Materials and Methods

PleD cloning

A QuickChangeII Site-Directed Mutagenesis Kit (Stratagene) was used to create PleD mutations, R390S and R148A R178A R313A R390S (= $\Delta I_{S,all/R390}$ mutant), taking as templates pRUN plasmids (derived from the pBR322 vector) containing the C-terminally His_6-tagged wild-type gene (12) or the R148A R178A R313A triple mutant(= $\Delta I_{S,all}$ mutant) (14). The wild-type PleD and mutants were expressed and purified as described elsewhere (14).

Enzymatic assays

The DGC activity of wild-type and mutant PleD at a nominal concentration of 20 µM in presence of c-di-GMP was analyzed by detection of the site-product pyrophosphate using a pyrophosphatase-coupled spectrophotometric assay. The details of the assay were reported before (12). The PleD samples were incubated with c-di-GMP for at least thirty minutes prior to the start of the enzymatic reaction adding 100 µM GTP. To obtain the IC50, the specific activity of PleD mutant R390S as a function of [c-di-GMP] were fitted by a single exponential with the program SigmaPlot (SPSS Science).

Thermal unfolding monitored by circular dichroism

The far-UV CD spectra were measured using an Aviv62A DS (Lakewood, N.J.) circular dichroism spectrophotometer with a computer controlled Kryo-Thermat 650 (Haake) water bath. The samples of wild-type PleD and the $\Delta I_{S,all}$ mutant at 10 µM were measured in a 50 mM Tris pH 8.0 buffer ±120 µM c-di-GMP. For PleD activation 1 mM $BeCl_2$, 10 mM NaF and 10 mM $MgCl_2$ were added to the buffer. The thermal stability of PleD was monitored by heating samples over a range of 10-75 °C at a rate of 30 °C/h in a water-jacketed cell with a 0.1 cm light path. The ellipticity at 222 nm was recorded at 1 °C intervals, collecting data at each point for 3-6 s. The observed ellipticity θ_{obs} was normalized according to the equation:

$\theta_{norm} = (\theta_{obs} - \theta_{min}) \times 100 / (\theta_{max} - \theta_{min})$,

whereby θ_{min} and θ_{max} are the ellipticities of unfolded and folded PleD, respectively. To determine the apparent T_M of the unfolding, data were fit using the commercial

program SigmaPlot (SPSS Science) assuming a two-state transition between the folded and the unfolded protein.

Analytical ultracentrifugation

A Beckman-Coulter XL-A analytical ultracentrifuge was used for the experiments, which were conducted at 10 °C. Where appropriate protein samples were buffered either in the non-activating (10 mM HEPES pH 7.5, 100 mM NaCl) or the activating buffer (10 mM HEPES pH 7.5, 100 mM NaCl, 10% N-Methylformamide, 10 mM $MgCl_2$, 1 mM $BeCl_2$, 10 mM NaF). The measurements of the density of the non-activating (1.005 g/cm³) and the activating buffer (1.012 g/cm³) were performed at 10 °C with an Anton Paar DMA 4500 density meter. The buffer viscosity of 1.31 mPa × s was measured with a Brookfield DV-II cone-plate viscometer. The partial specific volume of the PleD samples at 10° C (0.734 cm³/g) was calculated using the amino acid sequence. All measurements were analyzed using the absorption coefficients of 10190 $M^{-1}cm^{-1}$ at 275 nm and 3957 $M^{-1}cm^{-1}$ at 253 nm for wild-type PleD (M_W = 50447 Da), of 10189 $M^{-1}cm^{-1}$ at 275 nm and 3957 $M^{-1}cm^{-1}$ at 253 nm for the $\Delta I_{S,all}$ mutant (M_W = 50192 Da) as well as of 14945 $M^{-1}cm^{-1}$ at 275 nm and 23700 $M^{-1}cm^{-1}$ at 253 nm for c-di-GMP (M_W = 690 Da). The sedimentation equilibrium experiments were performed at 12000 rpm with 20 µM wild-type or mutant PleD in the non-activating and the activating buffers as well as with 20 µM wild-type or mutant PleD ± 40 µM c-di-GMP in the activating buffer. Data collection was performed at 253 and 275 nm after reaching the equilibrium (20 hours). Using two different wavelengths allowed the determination of the PleD concentration in presence of the light absorbing ligand c-di-GMP. The data analysis was performed with the DISCREEQ software (17,18). In short, fits were performed according to the review (19) equation [9], using the calculated molecular weights of six theoretical oligomeric species of PleD as constants. The monomer-dimer dissociation constants were calculated applying the law of mass action.

Surface plasmon resonance

The NTA Sensor Chips, the HBS-P and the HBS-EP buffers were purchased from GE Healthcare, $NiCl_2$ and urea from Sigma-Aldrich. All experiments were conducted on a BIACORE 3000 (GE Healthcare) at 10 °C and a constant flow of 30 µl/min. The protein and the c-di-GMP samples were diluted in the running buffer (10 mM HEPES

pH 7.4, 0.15 M NaCl, 50 µM EDTA and 0.005% Surfactant P20). The activation of the sensor chips was performed reproducibly applying a 500 µM $NiCl_2$ solution for one minute. The wild-type and mutant PleD samples (30 µg/ml) were immobilized in three of the four channels applying the protein solutions for one minute, resulting in a signal of approximately 1000 response units (RU). The protein-free channel was used as a reference channel.

For the analysis of the analyte binding, c-di-GMP was applied at concentrations of 50, 25, 12.5, 6.25, 3.125, 1.5625 and 0 (=buffer) µM, respectively. After each c-di-GMP injection and simultaneous record of the association and dissociation phases the protein was striped from the chip applying an 8 M urea solution. The chip surface was activated by $NiCl_2$ and the proteins were immobilized on the channels again before applying a new c-di-GMP sample.

The analysis of c-di-GMP-protein binding kinetics was performed using the BIAevaluation software Ver. 4.1. The sensorgrams, subtracted by the response on the reference channel, were normalized with respect to the immobilized protein levels. The responses of the 0 µM c-di-GMP (= buffer) injections were subtracted from other sensorgrams prior to the performance of the fitting procedure. The processed sensorgrams were fitted using a simple 1:1 binding model with a global parameter for the drift.

Calculation of dimer concentration

The concentration of dimers [D] can be calculated from the absolute concentration $[M_0]$ knowing the dimerization dissociation constant K_D using the following equation:

$$[D] = 0.5 \times [M_0] + 0.125 \times (K_D - \sqrt{(8 \times [M_0] \times K_D + K_D^2)}),$$

whereby the concentration of monomers [M] can be derived from:

$$[M_0] = 2 \times [D] + [M].$$

Results

Thermal unfolding of PleD monitored by CD

Our recent publication has verified the crucial role of the secondary inhibition sites, $I_{S,D2}$ and $I_{S,DGC}$, in the inhibition process by site directed mutagenesis and enzymatic characterization (14). C-di-GMP mediated domain cross-linking should manifest itself in a higher rise of the protein stability compared to PleD that binds c-di-GMP solely to the primary I-site.

Thermal unfolding of wild-type PleD monitored by CD at 222 nm in absence or presence of c-di-GMP (I_P-$I_{S,D2}$ cross-linking) reveals a T_M increase of 2.8 °C in the ligand bound state (Figure 1A, Table 1). The bad fitting to the two-state model in the initial phase of the heat-denaturation, might represent unfolding of the single domains. Addition of c-di-GMP seems to decrease this atypical behavior. To investigate the effect of protein stabilization by c-di-GMP cross-linking I_P and $I_{S,DGC}$, which is observed in PleD dimers, the protein was activated by $BeF_3^- \bullet Mg^{2+}$. The activation process leads to destabilization of PleD that can be attributed to the Mg^{2+}-ions (Figure 1A). The subsequent addition of c-di-GMP to the activated PleD dramatically increases the T_M value by 11.9 °C (Table 1).

The mutation of the three arginine residues to alanine in the secondary I-sites deficient PleD mutant (R148A R178A R313A = PleD-$\Delta I_{S,all}$) leads to a destabilization of the protein by 3.3 °C (Figure 1B). In contrast to the wild-type protein, addition of c-di-GMP increases the stability of the mutant to a much lesser amount in both states, the non-activated and the activated one (Table 1). Due to the absence of the secondary I-sites - disabling the domain cross-linking, this decreased stabilization effect must be attributed to the binding of c-di-GMP to the primary I-site.

Therefore, these results show stabilization of PleD by c-di-GMP binding to the primary I-site (a) and especially to the secondary I-sites (b). The later can be attributed to long range domain interactions.

Investigation of the PleD dimerization process by AUC

Dimerization of PleD is used as one of the regulatory mechanisms for this protein (14,15). Additionally, it is known that mutations in the $I_{S,D2}$-site result in a 4-10 fold higher enzymatic activity compared to the wild-type protein (13,14). To rule out the

involvement of the secondary I-sites in dimerization of the protein as well as to elucidate the effect of c-di-GMP in this process we have performed sedimentation equilibrium AUC experiments with PleD wild-type and the $\Delta I_{S,all}$ mutant in non-activating and activating ($+BeF_3^- \bullet Mg^{2+}$) conditions, respectively (Figure 2). Due to instability of PleD protein losses were detected during these experiments (Table 2).

Under non-activating conditions, both, wild-type and mutant PleD are found exclusively in the monomeric state at a concentration of 20 µM (Table 2). Activation by $BeF_3^- \bullet Mg^{2+}$ leads to a monomer/dimer equilibrium that translates into a K_D for dimerization of 10-15 µM for both wild-type and mutant PleD (Table2). Therefore, mutation of the $I_{S,D2}$-site does not alter the dimerization process.

Surprisingly, addition of c-di-GMP to activated wild-type or mutant reduced the dimerization K_D by an order of magnitude in both cases (Table2). These results suggest stabilization of the dimerization in presence of c-di-GMP independent of the secondary I-sites. Unfortunately, we could not perform similar experiments with the non-activated protein samples, for which much higher protein and c-di-GMP concentrations are needed. Due to the high absorbance of c-di-GMP, utilization of higher concentrations is not feasible.

In these experiments c-di-GMP was added in non-saturating concentrations to the PleD samples, due to its very high extinction coefficient. Nevertheless, the calculated protein-ligand stoichiometry (Table 2) indicates increased binding of c-di-GMP to PleD wild-type than to the $\Delta I_{S,all}$ mutant.

Determination of thermodynamic and kinetic rate constants of PleD by surface plasmon resonance (SPR)

ITC experiments have shown binding of two molecules c-di-GMP per PleD protomer (15). Additionally, the crystal structures of PleD show a c-di-GMP dimer cross-linking the primary with one of the secondary I-sites (12,14).

Two ways of binding of these two c-di-GMP molecules are conceivable. They could bind sequentially as monomers or as a dimer.

We have used the SPR technique to get further insight into the binding process. Association phases of 90 s and dissociation phases of 300 s of c-di-GMP binding to the PleD wild-type as well as to the mutants R313A ($\Delta I_{S,DGC}$), R148A R178A ($\Delta I_{S,D2}$) and $\Delta I_{S,all}$ were recorded (Figure 3).

The fittings of the obtained curves using the nominal concentrations of applied c-di-GMP and the fitting algorithms of the BIAevaluation software for a 1:1 binding or sequential binding of two c-di-GMP molecules resulted in very bad fits ($\chi^2 > 1$).

The rich oligomeric behavior of c-di-GMP is well known, in particular to form dimers with intercalated bases (20,21). Due to the utilization of extremely high concentrations (millimolar range) in these studies the dimerization K_D was not determined (21). Recent NMR studies performed at lower c-di-GMP concentrations suggest a value of about 50 µM for the K_D (M.Allan, unpublished data).

This dimerization K_D can be used for the concentration calculations of the monomeric and the dimeric species out of the nominal concentrations (see Materials&Methods). If in such a way calculated concentrations of monomeric c-di-GMP are used in the previously described fittings, the fits do not improve (data not shown).

On the other hand, performance of the fitting procedures with the calculated dimer concentrations of c-di-GMP (dimerization K_D = 50 µM) using a simple 1:1 binding algorithm leads to trustable fits (χ^2 = 0.1 for PleD wild-type) (Figure 3). The quality of the fits deteriorated, if for the assumed dimerization K_D of c-di-GMP, values below 10 µM or above 100 µM were used.

Although binding of c-di-GMP monomers and dimers could apply in parallel we do not approach this option. The very low propensity of such a scenario, the inability to prove the correctness of a corresponding fit, taking in account too many variables, and the very good fits with c-di-GMP dimers as the binding species let us discard this possibility. Thus, we propose that c-di-GMP binds as a dimer to PleD with a dimerization K_D of 50 µM for c-di-GMP. The binding is hereby very strong, with an apparent K_D of 25 nM (Table 3). A c-di-GMP solution with the nominal concentration of 1.2 µM contains this dimer concentration.

To elucidate the interaction properties between PleD and $[c\text{-di-GMP}]_2$ regarding the secondary I-sites, the SPR experiments were conducted with the I_S mutants (Table 3). The absence of the $I_{S,DGC}$ (=R313A) has almost no effect on the apparent binding properties. Whereas, the absence of the $I_{S,D2}$ (=R148A R178A) or both secondary I-sites results in decreased binding with a four-fold higher dissociation constant.

We have tried to estimate the effect of the subsequent cross-linking process fitting the curves to a “two-state reaction” algorithm of the BiaEval software assuming the following reactions:

(1) $(c\text{-di-GMP})_2$ + PleD ⇆ PleD-$(c\text{-di-GMP})_2$

(2) PleD-(c-di-GMP)$_2$ ⇆ PleD*-(c-di-GMP)$_2$,

whereby the first reaction represents the binding of c-di-GMP dimers to the primary I-site, and the second reaction represents the cross-linking mechanism. The quality of the fits stayed hereby unchanged. Additionally, the obtained rate constants for the second reaction were negligible in comparison to the ones of the first reaction and are therefore not trustable (data not shown). The kinetic parameters for a potential second step can therefore not be derived by SPR.

Thus, binding of c-di-GMP to the immobilized PleD is not strongly affected by secondary I-sites in the non-activated state i.e. monomeric, where c-di-GMP dependent cross-linking can take place solely between $I_{S,D2}$ and I_P.

Enzymatic activity of the PleD R390S mutants

The fact that c-di-GMP is to our knowledge the first inhibitor, working as domain/protein cross-linker, and that lack of the secondary I-sites doesn't result in the complete abolishment of the inhibition (14) led us investigate this regulatory mechanism in closer detail by the means of site directed mutagenesis.

In case of PleD the primary I-site is composed of the strongly conserved RxxD motif and the arginine 390. Unfortunately, mutations in the RxxD motif result in proteins lacking any enzymatic activity (13) and are therefore inconclusive.

We have mutated arginine 390 of the primary I-site to a serine in the wild-type and in the secondary I-sites deficient (R148A R178A R313A) background to investigate the role of this residue in the process of inhibition.

The strong involvement of the primary I-site mutant R390S in the c-di-GMP dependent inhibition is evident in the strongly increased K_i (140 µM) (Figure 4). The combination of this mutation with the $\Delta I_{S,all}$ mutant (=> $\Delta I_{S,all/R390}$ mutant) leads to a complete independence from the regulatory effect of c-di-GMP. Additionally, we couldn't detect any binding of c-di-GMP (c = 100 µM) by this mutant using the SPR technique (data not shown). To our surprise this mutant doesn't exhibit anymore the higher enzymatic activity that was observed with the $\Delta I_{S,all}$ mutant (Figure 4).

Thus, R390 of the primary I-site turns out to be strongly involved in c-di-GMP binding and in the inhibitory process.

Discussion

C-di-GMP binds as a dimer to the DGC PleD

One of the interesting aspects concerning the second messenger c-di-GMP is its ability to form multimers (20,21). Due to intracellular concentrations in the sub-micromolar range (7,22,23) only monomeric and dimeric forms of c-di-GMP seem to predominate *in-vivo*. Moreover, both species are relevant for the signal transduction process.

Whereas PDEs were shown to bind c-di-GMP in the monomeric form (24), the crystal structures of PleD and WspR have shown c-di-GMP dimers bound to the allosteric sites of the DGCs (12,14,25). In contrast to the general opinion of sequential binding of c-di-GMP monomers to the allosteric sites of DGCs, our experiments show c-di-GMP dimers as being the relevant species in this process. The very low dissociation constant between $(c\text{-di-GMP})_2$ and PleD (25 nM) is actually in near agreement with the previously described K_i and K_D of 0.6 µM c-di-GMP (12,15). Assuming a dimerization K_D of 50 µM (M. Allan, unpublished data) a 0.6 µM c-di-GMP solution consists of 7 nM dimers.

Allosteric regulation of DGCs by c-di-GMP dimers might additionally lead to a higher stringency of the inhibitory process. At nanomolar levels of c-di-GMP, which are normally found in bacterial cells (26,27), the DGCs are almost unaffected by c-di-GMP due to very low concentrations of $(c\text{-di-GMP})_2$. The rise in the overall c-di-GMP concentration from ~50 nM to ~5 µM (the physiological levels) leads to an quasi exponential increase in the $(c\text{-di-GMP})_2$ levels (Figure 5). Due to a very low binding K_D to PleD, inhibition by $(c\text{-di-GMP})_2$ must have an exponential character, accordingly. This is in agreement with the regulatory role of this process, which should shut down DGCs at higher cellular levels of the second messenger.

The future has to show which of the described c-di-GMP species play a role in the downstream signaling processes. Both species, monomers and dimers, are imaginable to be involved in these, according to the expected downstream effects. If the c-di-GMP sensor is a part of a signaling pathway responding to low concentrations of the second messenger, the sensor may recognize monomers. The signal circuitries responding to higher c-di-GMP concentrations might rely on c-di-GMP sensors, which recognize only dimers. Such situation could actually explain the structural

observation of PilZ domain proteins in complex with monomeric (28) and dimeric (29) c-di-GMP.

Regulation of PleD by the "domain-immobilization" mechanism

The discovery of two separate cross-linking modes of PleD by c-di-GMP dimers binding to one primary and one of the two secondary I-sites, respectively, as well as the resulting regulation by the "domain-immobilization" mechanism has often been met with skepticism. Although more and more structural evidences appear (12,14,25,28), which affirm the relevance of the secondary I-sites for the binding of c-di-GMP, the main scientific opinion states the RxxD motif of the I_P as the sole c-di-GMP binder, resulting in protein inhibition by unknown conformational changes (13,26).

According to our results the $\Delta I_{S,all/R390}$ PleD mutant having only the RxxD motif left for interaction with c-di-GMP doesn't bind the second messenger anymore. In line with this, the $\Delta I_{S,all/R390}$ mutant shows no product inhibition. In addition to the PleD-c-di-GMP complex structures we have now the biophysical evidence that R390 is part of the I_P-site interacting directly with c-di-GMP. Other DGCs having no arginines at the corresponding position probably utilize other near-by residues in combination with the well conserved RxxD motif to bind c-di-GMP. Thus, although the RxxD motif stays the unchallenged indicator of allosterically regulated DGCs (13), its sole presence is not sufficient for the binding process.

The surprising AUC result that binding of c-di-GMP to the I_P-site enhances the dimerization of PleD independently of the I_S-sites must be verified and the putative mechanism has to be characterized in the future. It is imaginable that such protein-ligand interaction leads to a reduced flexibility of the linker-attached GGDEF domain or keep it out of the Rec1-Rec2 dimer interface, allowing a higher stochastic propensity for the Rec domains to form the $(Rec1\text{-}Rec2)_2$ stem dimer.

Although our SPR results indicate a strong preference of the I_P-site for binding c-di-GMP, the thermal unfolding experiments verify a crucial role for the secondary I-sites in the c-di-GMP dependent protein stabilization, especially in the activated state.

Thus, the allosteric product regulation mechanism is initiated by a highly affine binding of c-di-GMP dimers to the I_P followed by an interaction with the secondary I-sites cross-linking the involved domains.

The subsequent interaction of 'captivated' c-di-GMP with the secondary I-sites, at least with $I_{S,D2}$, must be (much) lower in affinity than binding to the I_P, since the dissociation constants of the wild-type PleD and the $\Delta I_{S,all}$ mutant differ only by a factor of ~4. Missing improvements in the quality of the fits, which additionally take the cross-linking mechanism in account, also speaks for this conclusion. No conclusions can be drawn concerning the I_P-$I_{S,Dgc}$ cross-linking mode from these SPR experiments, due to the incompatibility of the activation procedure with this technique. On the other hand the R390S PleD mutant, which binds c-di-GMP between the RxxD motif and the secondary I-sites, exhibits an extreme rise in the K_i. This confirms the lower binding affinity of all secondary I-sites in comparison to the very strong binding affinity of the I_P.

According to these findings we were mistaken in our recent publication describing the secondary I-sites mutant $\Delta I_{S,all}$ as partially inhibitable (14). The ~ 4-10 fold higher activity of this mutant compared to the wild-type protein is due to the mutations of arginines 148 and 178 to alanines as was described before (13). Our AUC data clearly demonstrate the independence of this 'activation' mechanism from the dimerization process. Addition of c-di-GMP to the $\Delta I_{S,all}$ mutant results in an asymptotic decrease of the enzymatic activity that converges to a rate that is equal to rate of the not-inhibited wild-type PleD. Surprisingly, elimination of arginine 390 in this background (= $\Delta I_{S,all/R390}$) abolishes completely the described activation, but also the observed 'inhibition' by c-di-GMP. The enzymatic activity of this quadruple mutant is equal to the activity of the not-inhibited wild-type PleD and of the inhibited $\Delta I_{S,all}$. Therefore, addition of c-di-GMP to the $\Delta I_{S,all}$ mutant seems to counteract the observed increased activity. We propose that R390 is involved in this activation process, but can be withdrawn from it by binding c-di-GMP to the I_P-site.

Summing up, we have shown that: (I) c-di-GMP binds in its dimeric form to PleD; (II) the recruiting of $(c\text{-di-GMP})_2$ to PleD is performed mainly by the I_P-site; (III) for a I_P-c-di-GMP interaction a complete I_P-site is needed, namely R390 and the RxxD motif; (IV) the I_P-c-di-GMP interaction doesn't result in inhibition; (V) cross-link of c-di-GMP loaded I_P-site with one of the secondary I-sites causes the inhibition of PleD. Further investigations must have the aim to quantify the thermodynamic and kinetic binding parameters of the interaction between the c-di-GMP bound to the I_P-site and the secondary I-sites to gain the complete understanding of the allosteric product inhibition process.

Acknowledgements

We thank PD Walter Huber and Josiane Kohler from F. Hoffmann-La Roche AG for giving us the opportunity to perform the SPR experiments using their instruments and consumables as well as for introducing in and teaching of the SPR technology.

Figure legends

Figure 1. Thermal unfolding profiles of PleD wild type (A) and the PleD-$\Delta I_{S,all}$ mutant (B)

The thermal stability was assayed of non-activated (- c-di-GMP [black], + c-di-GMP [red]) and $BeF_3^- \bullet Mg^{2+}$-activated protein (- c-di-GMP [green], + c-di-GMP [yellow]). The unfolding of wild type protein was additionally measured in presence of Mg^{2+}-ions [blue]. The elipticity at 222 nm was monitored as a function of temperature. The data have been fitted with a two-state model (continuous lines).

Figure 2. Analysis of the quaternary state of PleD by sedimentation equilibrium AUC

Representative plots of the SE runs of PleD wild-type (circles) and PleD-$\Delta I_{S,all}$ mutant (triangles). (A) Absorbance of $BeF_3^- \bullet Mg^{2+}$-activated protein samples in presence of c-di-GMP [red] measured at 253 nm. (B) Absorbance of non-activated [blue] and activated protein samples (- c-di-GMP [green], + c-di-GMP [red]) measured at 275 nm. The corresponding plots of the residuals are shown at the upper panels (wild-type as lines, mutant as dots). The black lines represent the global fits to an equilibrium model.

Figure 3. Analysis of c-di-GMP binding to immobilized PleD samples by SPR

Representation of association and dissociation phases of c-di-GMP at nominal concentrations of 50, 25, 12.5, 6.25, 3.125 and 1.5625 µM, respectively, binding to immobilized PleD wild-type (A), the $I_{S,Dgc}$ mutant (C), the $I_{S,D2}$ mutant (E) and the $I_{S,all}$ mutant (G). The fits of the data are shown as stripped red lines. The corresponding residuals are shown in the lower panels (B, D, F and H).

Figure 4. Analysis of c-di-GMP dependent inhibition of PleD

Plots of the specific activity of PleD wild-type (blue), the $I_{S,DGC}$ mutant (black), the $I_{S,D2}$ mutant (red), the $I_{S,all}$ mutant (green), the R390S mutant (magenta) and the combined $I_{S,all/R390}$ mutant (yellow) as a function of c-di-GMP concentrations. Note, the ordinate is represented logarithmically!

Figure 5. Simulation of dimeric (and monomeric) c-di-GMP concentrations at physiological concentrations of c-di-GMP (overall)

The concentrations of the monomeric (blue) and the dimeric (red) species of c-di-GMP were calculated according to the equation given in the Materials&Methods section assuming a dimerization K_D of 50 µM. Note the different scaling of the concentrations of both species!

sample	apparent T_M (°C)
PleD wt	48.2 ± 0.6
PleD wt + c-di-GMP	51.0 ± 0.3
PleD wt + Mg^{2+}	36.2 ± 0.5
PleD wt + BeF_3^- • Mg^{2+}	33.6 ± 0.7
PleD wt + BeF_3^-•Mg^{2+} + c-di-GMP	48.1 ± 1.1
PleD $\Delta I_{S,all}$	44.9 ± 0.6
PleD $\Delta I_{S,all}$ + c-di-GMP	44.4 ± 0.5
PleD $\Delta I_{S,all}$ + BeF_3^- • Mg^{2+}	35.8 ± 0.5
PleD $\Delta I_{S,all}$ + BeF_3^- • Mg^{2+} + c-di-GMP	41.3 ± 0.6

Table 1: Thermal unfolding of PleD samples

sample	protein concentration in solution (µM)	abundance (%)			dimer. K_D* (µM)	protein: ligand stoichio-metry
		monomer	dimer	multimers		
PleD wt	10.81	100	0	0	-	-
activated PleD wt	8.44	66.9	32.5	0.6	11.6	-
activated PleD wt + c-di-GMP	10.86	11.6	85.1	3.3	0.16	1:1.17
PleD $\Delta I_{S,all}$	9.91	97.6		2.4	-	-
activated PleD $\Delta I_{S,all}$	8.44	68.1	25.6	6.3	15.3	-
activated PleD $\Delta I_{S,all}$ + c-di-GMP	2.59	21.7	58.6	19.7	0.15	1:0.46

Table 2: Results of SE AUC experiments with PleD samples

* The K_D was calculated neglecting the higher oligomer species

sample	k_{on} (s^{-1} M^{-1})	k_{off} (s^{-1})	K_D (nM)
PleD wt	$7.85 \pm 0.04 \times 10^4$	$1.96 \pm 0.02 \times 10^{-3}$	25 ± 1
PleD $\Delta I_{S,DGC}$	$8.92 \pm 0.06 \times 10^4$	$5.49 \pm 0.02 \times 10^{-3}$	61 ± 2
PleD $\Delta I_{S,D2}$	$3.28 \pm 0.04 \times 10^4$	$3.71 \pm 0.06 \times 10^{-3}$	110 ± 6
PleD $\Delta I_{S,all}$	$3.72 \pm 0.04 \times 10^4$	$4.03 \pm 0.05 \times 10^{-3}$	108 ± 5

Table 3: SPR derived kinetic and thermodynamic rate constants

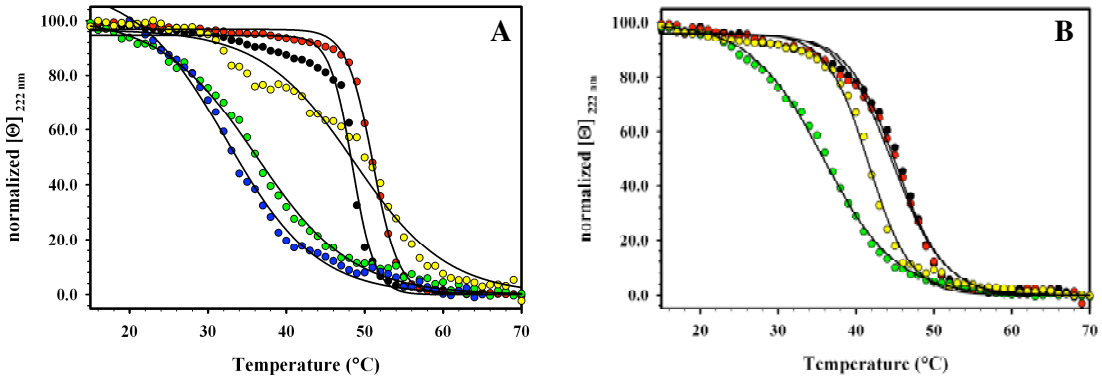


Figure 1

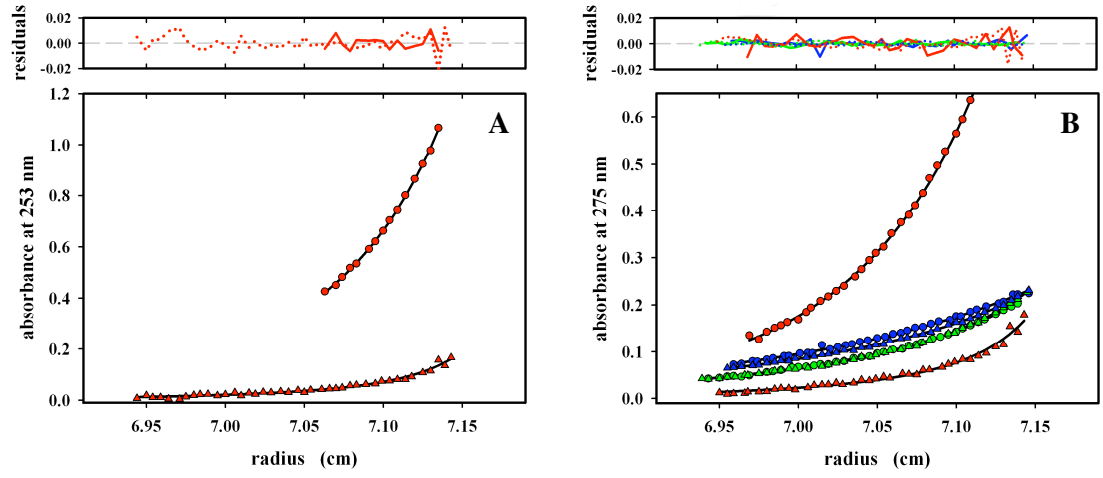


Figure 2

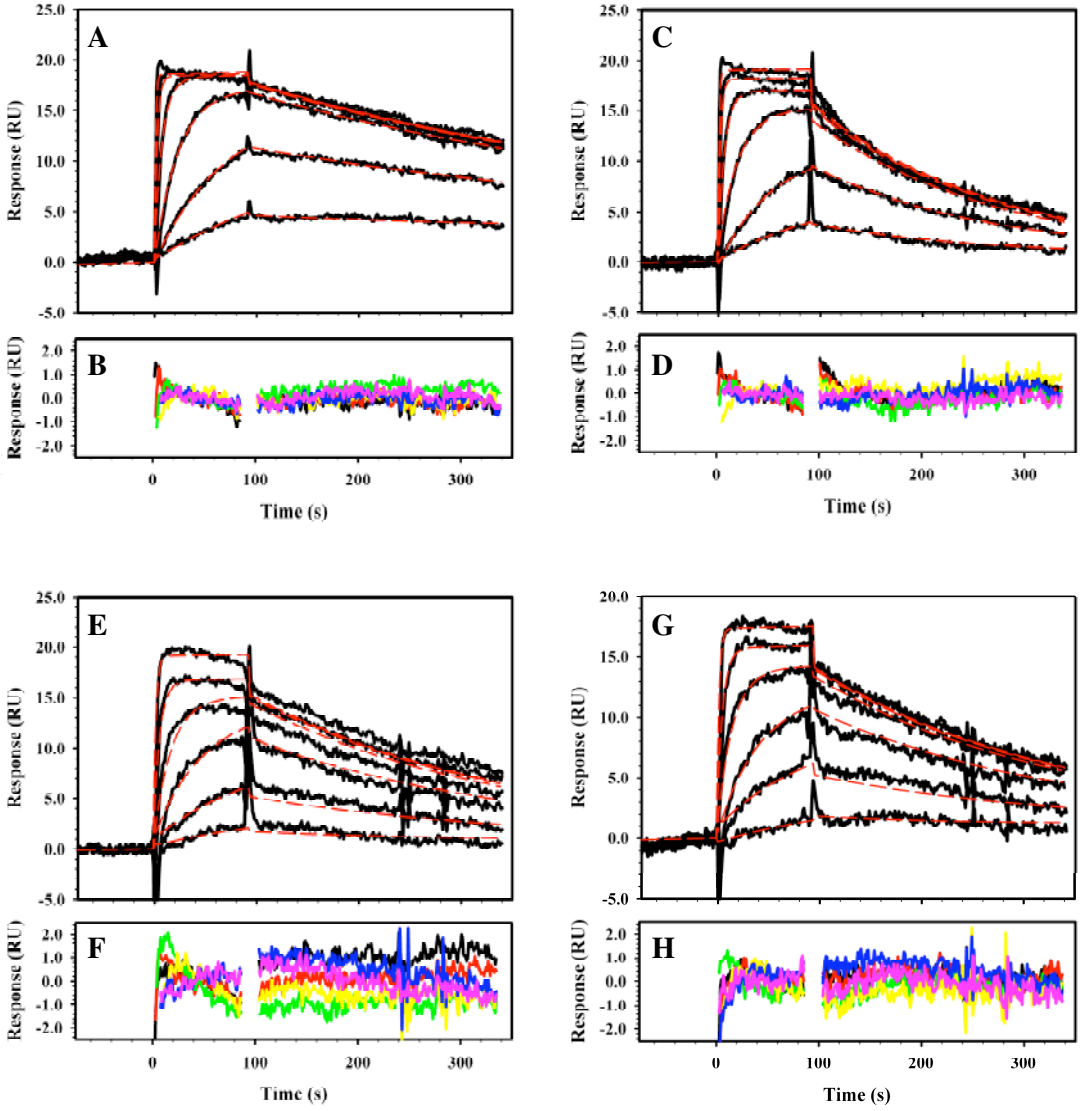


Figure 3

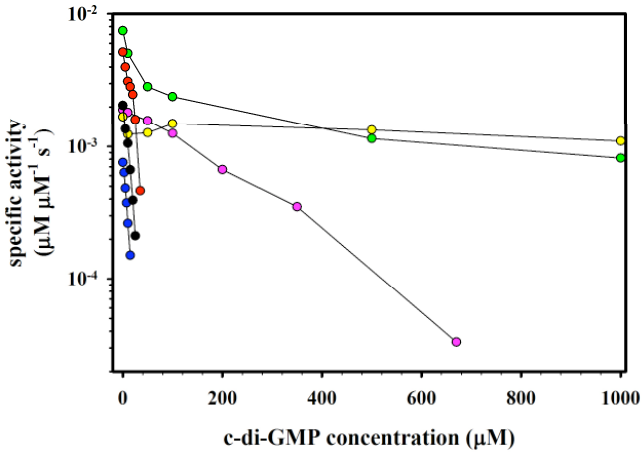


Figure 4

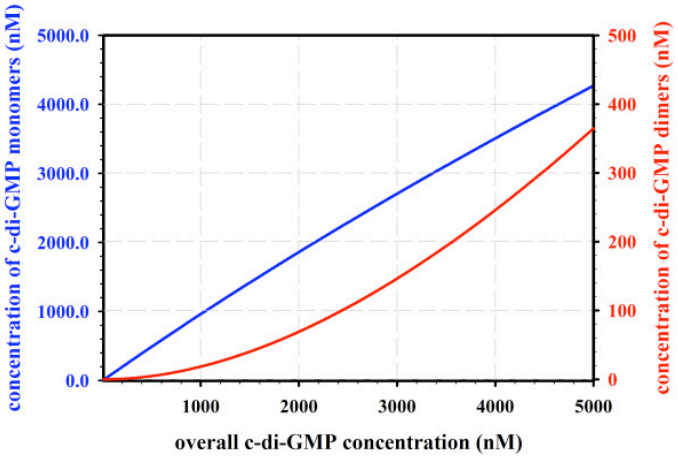


Figure 5

Bibliography

1. Ross, P., Weinhouse, H., Aloni, Y., Michaeli, D., Weinberger-Ohana, P., Mayer, R., Braun, S., de Vroom, E., van der Marel, G. A., van Boom, J. H., and Benziman, M. (1987) *Nature* **325**, 279-281
2. Duerig, A., Abel, S., Folcher, M., Nicollier, M., Schwede, T., Amiot, N., Giese, B., and Jenal, U. (2009) *Genes Dev* **23**, 93-104
3. Paul, R., Weiser, S., Amiot, N. C., Chan, C., Schirmer, T., Giese, B., and Jenal, U. (2004) *Genes Dev* **18**, 715-727
4. Aldridge, P., Paul, R., Goymer, P., Rainey, P., and Jenal, U. (2003) *Mol Microbiol* **47**, 1695-1708
5. Kim, Y. R., Lee, S. E., Kim, C. M., Kim, S. Y., Shin, E. K., Shin, D. H., Chung, S. S., Choy, H. E., Progulske-Fox, A., Hillman, J. D., Handfield, M., and Rhee, J. H. (2003) *Infect Immun* **71**, 5461-5471
6. Ausmees, N., Jonsson, H., Hoglund, S., Ljunggren, H., and Lindberg, M. (1999) *Microbiology* **145 (Pt 5)**, 1253-1262
7. Tischler, A. D., and Camilli, A. (2004) *Mol Microbiol* **53**, 857-869
8. Tischler, A. D., and Camilli, A. (2005) *Infect Immun* **73**, 5873-5882
9. Amikam, D., and Galperin, M. Y. (2006) *Bioinformatics* **22**, 3-6
10. Sudarsan, N., Lee, E. R., Weinberg, Z., Moy, R. H., Kim, J. N., Link, K. H., and Breaker, R. R. (2008) *Science* **321**, 411-413
11. Aldridge, P., and Jenal, U. (1999) *Mol Microbiol* **32**, 379-391
12. Chan, C., Paul, R., Samoray, D., Amiot, N. C., Giese, B., Jenal, U., and Schirmer, T. (2004) *Proc Natl Acad Sci U S A* **101**, 17084-17089
13. Christen, B., Christen, M., Paul, R., Schmid, F., Folcher, M., Jenoe, P., Meuwly, M., and Jenal, U. (2006) *J Biol Chem* **281**, 32015-32024
14. Wassmann, P., Chan, C., Paul, R., Beck, A., Heerklotz, H., Jenal, U., and Schirmer, T. (2007) *Structure* **15**, 915-927
15. Paul, R., Abel, S., Wassmann, P., Beck, A., Heerklotz, H., and Jenal, U. (2007) *J Biol Chem* **282**, 29170-29177
16. Paul, R., Jaeger, T., Abel, S., Wiederkehr, I., Folcher, M., Biondi, E. G., Laub, M. T., and Jenal, U. (2008) *Cell* **133**, 452-461
17. Schuck, P. (1994) in *Progress in colloid and polymer science* Vol. 94, pp. 1-13, Kremer, F., Richtering, W.
18. Schuck, P., Legrum, B., Passow, H., and Schubert, D. (1995) *Eur J Biochem* **230**, 806-812
19. Lebowitz, J., Lewis, M. S., and Schuck, P. (2002) *Protein Sci* **11**, 2067-2079
20. Egli, M., Gessner, R. V., Williams, L. D., Quigley, G. J., van der Marel, G. A., van Boom, J. H., Rich, A., and Frederick, C. A. (1990) *Proc Natl Acad Sci U S A* **87**, 3235-3239
21. Zhang, Z., Kim, S., Gaffney, B. L., and Jones, R. A. (2006) *J Am Chem Soc* **128**, 7015-7024
22. Weinhouse, H., Sapir, S., Amikam, D., Shilo, Y., Volman, G., Ohana, P., and Benziman, M. (1997) *FEBS Lett* **416**, 207-211
23. Simm, R., Morr, M., Kader, A., Nimtz, M., and Romling, U. (2004) *Mol Microbiol* **53**, 1123-1134
24. Minasov, G., Padavattan, S., Shuvalova, L., Brunzelle, J. S., Miller, D. J., Basle, A., Massa, C., Collart, F. R., Schirmer, T., and Anderson, W. F. (2009) *J Biol Chem* **284**, 13174-13184
25. De, N., Pirruccello, M., Krasteva, P. V., Bae, N., Raghavan, R. V., and Sondermann, H. (2008) *PLoS Biol* **6**, e67

26. Hengge, R. (2009) *Nat Rev Microbiol* **7**, 263-273
27. Simm, R., Morr, M., Remminghorst, U., Andersson, M., and Romling, U. (2009) *Anal Biochem* **386**, 53-58
28. Benach, J., Swaminathan, S. S., Tamayo, R., Handelman, S. K., Folta-Stogniew, E., Ramos, J. E., Forouhar, F., Neely, H., Seetharaman, J., Camilli, A., and Hunt, J. F. (2007) *Embo J* **26**, 5153-5166
29. Christen, M., Christen, B., Allan, M. G., Folcher, M., Jeno, P., Grzesiek, S., and Jenal, U. (2007) *Proc Natl Acad Sci U S A* **104**, 4112-4117

II.4 C-di-GMP dependent regulation of the 'stand-alone' diguanylate cyclases DgcA and DgcB from *C. crescentus*

Paul Wassmann[1], Andreas Beck[1], Eric Kusznir[2], Francis Mueller[2] and Tilman Schirmer[1*]

[1] Core program of Structural Biology & Biophysics, University of Basel, CH-4056 Basel, Switzerland

[2] Department of Molecular Structure Research, F. Hoffmann-La Roche AG, CH-4070 Basel, Switzerland

[*] Corresponding author. E-mail address: tilman.schirmer@unibas.ch

Keywords: c-di-GMP, diguanylate cyclase, regulatory mechanism, stand-alone DGCs

Abbreviations: Rec, respone regulator receiver domain; DGC, diguanylate cyclase; PDE, phosphodiesterase; c-di-GMP, bis-(3'-5')-cyclic dimeric guanosine monophosphate; GTPαS, guanosine 5'-O-(1-thiotriphosphate); ITC, isothermal titration calorimetry; SEC-MALS, size exclusion chromatography coupled multi angle light scattering; SLS, static light scattering; SPR, surface plasmon resonance

Abstract

Most diguanylate cyclases are composed of a C-terminal catalytic GGDEF domain and N-terminal dimerization/regulatory domains. A minor class of diguanylate cyclases shows non-annotated sequences of various lengths at its N-termini. These proteins are therefore called 'stand-alone' DGCs. We have analyzed two of these proteins, DgcA and DgcB from *Caulobacter cescentus*. Both are forming permanent N-terminal 'domain'-mediated dimers. In contrast to DgcA, the N-terminal 'domain' of DgcB seems additionally to have a regulatory function.

Another regulatory mechanism utilized by diguanylate cyclases is the allosteric product inhibition. Hereby, the main theory postulates binding of c-di-GMP to the RxxD sequence motif in the GGDEF domain. DgcA, which possesses these residues,

was previously shown to be inhibitable. DgcB, on the other hand, lacks the RxxD motif. Nevertheless, we could show inhibition of DgcB by c-di-GMP. SPR and ITC experiments show binding of c-di-GMP dimers to a non-defined allosteric site. Additionally, SEC-MALS and AUC experiments demonstrate formation of DgcB tetramers in presence of c-di-GMP.

Introduction

The role of the eubacterial second messenger bis-(3'-5')-cyclic dimeric guanosine monophosphate (c-di-GMP) stayed for a long time unrecognized (1). The interest in this signaling pathway awoke with the realization that the c-di-GMP producing diguanylate cyclases (DGCs) and degrading phosphodiesterases (PDEs) are ubiquitously coded in the genomes of bacteria (2). The second messenger gained further attention by emerging evidences showing involvement of c-di-GMP in several crucial cellular functions. It regulates the transition between the sessile and the motile states in bacteria, controlling biofilm formation (3-5) on the one hand and virulence of pathogens on the other (6-8). Additionally, it was shown to be involved in control of the cell cycle (9) and antibiotic biosynthesis (10).

Despite the high versatility of cellular responses that are regulated by c-di-GMP, the number of identified direct downstream targets stays relatively low. The second messenger is known to interact with proteins containing the PilZ domain (11,12), but the downstream effects of this binding is not resolved, yet. Most recently c-di-GMP was shown to play a role on the DNA/RNA level interacting with riboswitches (13).

The few identified effectors of c-di-GMP are facing an abundance of DGCs and PDEs, which are often present in high numbers in a single bacterial species. The enzymatic activity of DGCs, the biosynthesis of c-di-GMP consuming two molecules GTP, is performed by the GGDEF domains, named according to the conserved consensus motif of the active site. The DGCs were shown to act as dimers (14) forming a single active site at the interface of the GGDEF domains. These results are in agreement with the fact that the GGDEF domains bind only one GTP molecule (15,16).

Enzymatic activity of DGCs is highly regulated. One of the regulatory mechanisms is the allosteric product inhibition of DGCs. Crystal structures of PleD (15,16) and WspR (17) have shown binding of c-di-GMP dimers to an RxxD motif in the GGDEF domain (referred to as the primary I-site) and to additional arginines (secondary I-sites) from another domain. It was proposed that such relative domain immobilization prevents the GGDEF domains to form a functional active site. The RxxD motif is conserved in the DGCs and was, therefore, postulated to be an indicator of DGCs, which are regulated by product inhibition (18).

Dimerization of DGCs is used as another regulatory mechanism. Many DGCs use N-terminal domains for dimerization, e.g. response regulator receiver domains (Rec). In case of PleD dimerization occurs upon phosphorylation of the N-terminal Rec domain (14,16).

In addition to the main class of DGCs consisting of C-terminal GGDEF domains and N-terminal regulatory dimerization domains, a high number of GGDEF domain proteins contains N-terminal 'domains' of ~ 25 to 500 amino acids length. These 'domains' are not annotated. Therefore, such proteins were misleadingly named 'stand-alone' DGCs. Two contradicting theories are attributed to the function of these N-terminal domains. One assigns no specific function to them. This leads to a theoretical ability of the GGDEF domains to dimerize. The other theory sees in them the necessary dimerization domain resulting in formation of permanent DGC dimers. The last theory implies the inability of the GGDEF domains to dimerize on their own.

We have analyzed, biochemically and biophysically, two of such 'stand-alone' DGCs from *C. crescentus*, DgcA and DgcB. According to the present RxxD motif, DgcA was shown to be regulated by product inhibition (18). DgcB misses the RxxD motif and is therefore believed to be unaffected by product inhibition.

In this study we have tried to answer the following questions. (I) What function could the N-terminal 'domains' have in the 'stand-alone' DGCs? (II) Are the GGDEF domains able to dimerize by their own or do they require N-terminal dimarization domains? (III) Several 'stand-alone' DGCs miss the RxxD motif like DgcB. Are these proteins permanently active, or do they utilize a distinct regulatory mechanism?

Materials and Methods

Expression and Purification

Instructions of DNA construct generation can be provided upon request. C-terminally His_6-tagged versions of full-length DgcA, ΔDgcA (aa 83 – 237) and full length DgcB were overexpressed in *E.coli* BL21(DE3)pLysS (Novagen) strain. Cells carrying the appropriate plasmid were grown in Luria Broth (LB) media supplemented with ampicillin (100 µg/ml) and chloramphenicol (20 µg/ml). Induction of protein expression was performed adding IPTG to a final concentration of 0.5 mM. Clarified crude extracts were loaded on a HP HisTrap column (GE Healthcare) using 20 mM HEPES pH 7.5, 100 mM KCl, 20 mM imidazole as loading buffer. Elution of His_6-tagged proteins was performed raising the imidazole concentration. Concentrated protein samples were applied to the Superdex200 HR26/60 size-exclusion-chromatography (SEC) column using 20 mM HEPES pH 7.5, 100 mM KCl as running buffer.

Separation of c-di-GMP from DgcB was accomplished by a subsequent SEC run keeping the protein concentration below 5 mg/ml.

The phosphodiesterase YahA was purified according a published protocol (19) and was a generous gift from A. Basle (University of Basel, Switzerland).

Diguanylate cyclase activity assay

Enzymatic experiments with DgcA, ΔDgcA and DgcB were performed in a 20 mM Tris/HCl pH 8.0, 100 mM NaCl buffer at 1.6, 2.3 and 26 µM protein concentration, respectively, measuring the concentration of the second DGC product pyrophosphate. Degradation of pyrophosphate to phosphate was performed by the pyrophophatase. Detection of the phosphate using a coupled spectrophotometric assay is described elsewhere (15). Where necessary, PDE YahA (2 mg/ml) was added to prevent product inhibition of the DGCs. Measurement of DgcB inhibition was performed using apo-protein that was incubated with c-di-GMP (10, 25, 50, 100 µM) for at least thirty minutes prior to adding the substrate, GTP (c = 500 µM).

Isothermal titration calorimetry

Buffer (20 mM Tris-HCl pH 8.0, 100 mM NaCl, 5 mM $MgCl_2$, 2 mM $CaCl_2$) equilibrated protein and ligand solutions were thoroughly degassed at 25° C prior to

the performance of the experiments. Determination of dissociation constants of c-di-GMP and GTPαS binding to DgcB was conducted using a VP-ITC isothermal titration calorimeter (Microcal, Northhampton MA). Experiments were performed at 25° C with c-di-GMP (120 µM) or GTPαS (200 µM) in the syringe and DgcB (5 µM) in the cell. For the competition studies DgcB was saturated with the appropriate ligand at a concentration of 200 µM prior to the injection into the cell. To ensure complete re-equilibration of the solutions, delay between sequential injections was kept at 5-10 min. The integrated titration peaks were fitted as described in (20).

SEC-coupled multiangle light scattering (SEC-MALS)

SEC-MALS experiments were carried out using the ÄKTA*purifier* (GE Healthcare) connected Superdex200 HR10/300 column (GE Healthcare), the three-angle static light scattering detector miniDAWN (Wyatt Technology Corporation) and the differential refractive index detector Optilab rEX (Wyatt Technology Corporation). Bovine serum albumin (Sigma) was used for normalization of the SEC-MALS hardware. Where appropriate, protein samples (100 µl) were incubated with c-di-GMP for thirty minutes prior to the injection to the column. Measurements were performed at a flow rate of 0.7 ml/min. Concentrations of the eluted protein was monitored by the differential refractive index, which in combination with the light scattering signal was used for the molecular weight calculation using the program ASTRA5.3.

Analytical ultracentrifuge (AUC)

All experiments were performed at 10 °C on a Beckman-Coulter XL-A analytical ultracentrifuge with absorbance optics. The buffer was 10 mM HEPES, pH 7.5, 100 mM NaCl for all protein samples. At 10 °C the density was 1.005 g/cm3, measured with an Anton Paar DMA 4500 density meter, and the viscosity was 1.31 mPas measured with a Brookfield DV-II cone-plate viscometer. The partial specific volume of DgcB at 10° C was calculated using the amino acid sequence. All calculations were performed using the absorption coefficients of 23968 $M^{-1}cm^{-1}$ at 280 nm and 10592 $M^{-1}cm^{-1}$ at 253 nm for DgcB as well as of 12852 $M^{-1}cm^{-1}$ at 280 nm and 23700 $M^{-1}cm^{-1}$ at 253 nm for c-di-GMP. Sedimentation velocity runs with DgcB (26 µM) ± c-di-GMP (60µM) were conducted at 40000 rpm and recorded at 280 nm. The data were analyzed with Sedfit (21). Sedimentation equilibrium experiments were

performed at 12000 rpm with 13 µM DgcB ± 40 µM c-di-GMP. Equilibrium was reached after 20 hours. Using two different wavelengths, 253 nm and at 280 nm, allowed the determination of the DgcB concentration in presence of the likewise light absorbing ligand c-di-GMP. The data were analyzed with DISCREEQ (22-24). In short, fits were performed according to the review (23) equation [9], using the calculated molecular weights of six theoretical oligomeric species of DgcB as constants. The dimer-tetramer dissociation constants were calculated applying the law of mass action.

Surface plasmon resonance (SPR)

All experiments were performed on a BIACORE 3000 (GE Healthcare). The HBS-EP and HBS-P buffers and the NTA Sensor Chips were purchased from GE Healthcare. $NiCl_2$ and urea were obtained from Sigma-Aldrich.

10 mM HEPES pH 7.4, 0.15 M NaCl, 50 µM EDTA and 0.005% Surfactant P20 was used as running and dilution buffer. The experiments were performed at a constant flow of 30 µl/min and at 10° C. The Sensor Chip NTA was reproducibly activated by addition of a 500 µM $NiCl_2$ solution for one minute. DgcB was used at a concentration of 30 µg/ml and was immobilized on the chip applying the protein solution for one minute, resulting in a signal of approximately 1000 response units (RU). The interaction of c-di-GMP with DgcB in the association and the dissociation phases was performed applying the analyte to a reference channel and to the channel with the immobilized protein. The protein was striped with a 8 M Urea solution after each measurement cycle and was freshly immobilized on the chip prior to the application of the next c-di-GMP sample. Such procedure was performed due to a drifting baseline and a regeneration-prone binding of c-diGMP.

Analysis of c-di-GMP binding was performed using the BIAevaluation software Ver. 4.1. The reference channel subtracted sensorgrams were normalized according to the immobilized protein level. The sensorgram of injected 0 µM c-di-GMP (= buffer) was subtracted from all other sensorgrams prior to the performance of the fitting procedure. The curated sensorgrams were fitted using a simple 1:1 binding model with a global parameter for the drift. The maximum response upon binding of one molecule c-di-GMP per molecule DgcB, Rmax, was calculated according to the following equation:

$$R_{max,\ c\text{-}di\text{-}GMP} = R_{DgcB} \times (M_{W,\ c\text{-}di\text{-}GMP} / M_{W,\ DgcB}),$$

Where R_{DgcB} is the amount of immobilized protein, $M_{W,\ c\text{-}di\text{-}GMP}$ and $M_{W,\ DgcB}$ are the molecular weights of c-di-GMP and DgcB, respectively.

Estimation of the dimeric species concentration

Concentration of the dimeric species [D] can be calculated from the total molar concentration $[M_0]$ when the dissociation constant K_D is known using the following equation:

$$[D] = 0.5 \times [M_0] + 0.125 \times (K_D - \sqrt{(8 \times [M_0] \times K_D + K_D^2)}),$$

whereby

$$[M_0] = 2 \times [D] + [M].$$

Results

Sequence analysis of DgcA and DgcB

Both proteins are composed of a C-terminal GGDEF domain and an N-terminal domain of unknown structure and functional role (Figure 1A). The PCOILS web server (25) predicts coiled-coil regions in the N-terminal domains of both proteins with a > 50% probability (DgcA – aa 49-80, DgcB – aa 80-123, aa 146-168). To predict the DGC activity as well as the product inhibition behavior GGDEF domain sequences of DgcA and DgcB were compared with sequences of structurally known DGCs WspR and PleD (Figure 1B). As PleD and WspR, DgcA and DgcB bear the DGC consensus motif GGDEF/GGEEF. Protein sequence of DgcA contains the RxxD motif that has been shown to be indispensable for the c-di-GMP dependent inhibition of the diguanylate cyclase activity (18). Consistent with the domain-cross-linking mechanism (16) DgcA contains besides the primary I-site (RxxD motif) also an arginine (R110) of helix α0 representing a secondary I site candidate. In WspR, an arginine at an equivalent position is involved in binding of c-di-GMP (17). Additionally, the DgcA sequence shows an arginine (R103) two helical turns apart from R110 that could synergistically interact with c-di-GMP.

The RxxD motif of DgcB is degenerated to APPR, but DgcB bears an arginine at an equivalent position to R103 of DgcA. According to the main role of the RxxD motif in binding c-di-GMP, the protein sequence analysis leads to the assumption that DgcB is unaffected by the allosteric inhibition process. Thus, the presented analysis suggests that both proteins are functional diguanylate cyclases and that DgcA might be allosterically regulated by c-di-GMP.

The N-terminal domain mediates dimerization in DgcA

Both, DgcA and DgcB, elute from size-exclusion chromatography (SEC) columns as dimers (results not shown for DgcA). To see whether the N-terminal domains are involved in dimerization, a *dgcA* construct coding for a mutant (ΔDgcA = aa 83-237) that lacks the N-terminal domain was created. The correct fold of the resulting protein was assured by SPR experiments, showing similar c-di-GMP binding characteristics for the wild-type and the truncated versions of DgcA (Supplementary figure 1). As judged by SEC ΔDgcA is monomeric (results not shown) suggesting that the N-terminal domain is responsible for dimerization of DgcA. Dimerization is a

prerequisite for DGC activity (14), therefore enzymatic activity of both proteins was measured. The analysis showed no enzymatic activity for ΔDgcA, whereas full-length DgcA shows a high specific activity for DGCs of 0.61 μM μM^{-1} s^{-1} (see next paragraph) (Figure 2A).

Inhibition of DgcA and DgcB by c-di-GMP

DGCs with an allosteric product binding site are known to co-purify in presence of c-di-GMP (16). This is the case for DgcA and surprisingly also for DgcB. To be able to measure the enzymatic activity of ΔDgcA, DgcA and DgcB in absence of the inhibitory effect of c-di-GMP, measurements were performed in presence of c-di-GMP degrading phosphodiesterase YahA. DgcA shows a much higher specific activity (0.61 μM μM^{-1} s^{-1}) in comparison to DgcB (8.8×10^{-4} μM μM^{-1} s^{-1}) (Figure 2). Inhibition of DgcA and DgcB by co-purified and enzymatically produced c-di-GMP can be observed in enzymatic assays lacking YahA (Figure 2). Separation of c-di-GMP from DgcB can be achieved by SEC (HR16/60 Superdex 75) using protein samples of concentrations lower than 75 μM. The RxxD motif deficient DgcB (Figure 1) is fully inhibitable (Figure 3), with a rather high IC_{50} value of 20 μM (at a protein concentration of 26 μM).

Non-competitive inhibition of DgcB by c-di-GMP

DgcB lacks the predicted c-di-GMP binding residues of the primary I-site (16,18). Nevertheless it is affected by product inhibition. To understand this molecular process, competitive binding studies were performed using substrate analogue and product. For the determination of the stoichiometry as well as of the K_D-values (=K_i), isothermal titration calorimetry (ITC) technique was used. The integrated binding enthalpies of c-di-GMP added to DgcB resulted in readily fitable sigmoidal curves (Figure 4). Deviation from ideality was observed for the first 3-5 injections of c-di-GMP, as reported before for PleD (14). To avoid complication of our fits by introduction of terms taking in account e.g. the multimerization state of c-di-GMP (26), the value of the first injection was excluded from the fit. To study, whether c-di-GMP competes with the substrate for binding to the active site the non-hydrolysable substrate analogue GTPαS was used as competitor (Figure 4B). Both, in presence and absence of GTPαS, the dissociation constants of c-di-GMP were found to be virtually identical (0.66 ± 0.1 μM vs. 0.69 ± 0.1 μM). Both experiments show a stoichiometry

of 2 ± 0.4, which is in agreement with the presence of c-di-GMP dimers in structures of PleD and WspR (15-17,27).

Similarly, binding of the substrate analogue was also not influenced by the presence of c-di-GMP. Injections of GTPαS to apo and c-di-GMP incubated DgcB showed in both cases dissociation constants of 5 ± 3 μM (Figure 4C).

Thus, two molecules of c-di-GMP bind to an allosteric site of DgcB that is distinct from the active site. Binding of the inhibitor doesn't influence substrate binding, and *vice versa.*

Kinetics of c-di-GMP binding monitored by SPR

To get further insights in the binding mechanism between c-di-GMP and DgcB surface plasmon resonance (SPR) experiments were performed instrument. Three main questions were pursued by this approach. Can the SPR technique verify a c-di-GMP to DgcB binding ratio of 2:1, which was measured in solution by ITC? If so, does c-di-GMP bind to DgcB sequentially or as a dimer? And what are the kinetic parameters of this interaction?

To avoid obstruction of the binding site and hampered flexibility immobilizing the protein *via* one of the randomly working covalent immobilization techniques, NTA sensor chips were used to immobilize the protein *via* its C-terminal His_6-tag. Due to the dimeric nature of DgcB the immobilization occurs *via* both His_6-tags, which was shown in other cases to increase the stability of the sensor chip trapped proteins (28).

To obtain the binding stoichiometry steady state equilibrium experiments were conducted. For the purpose to reach the equilibrium an association phase of six minutes was chosen followed by a seven minutes long dissociation phase (Figure 5). Steady state responses were achieved after 250 seconds and were used to estimate the maximum response of the c-di-GMP binding to the trapped DgcB in a Scatchard plot analysis (Figure 6A). Analysis of the data resulted in an R_{max} of 40.4 RU (slope = $-K_A$, intercept = $K_A \times R_{max}$. Comparison of this value with the expected R_{max} of 20.5 RU taking in account the response of the immobilized protein (1170 RU) and a theoretical protein-ligand stoichiometry of 1:1 suggests binding of two molecules c-di-GMP per molecule DgcB. The non-linearity of the Scatchard plot (Figure 6A) clearly indicates that the 1:1 model is not adequate.

This result enables c-di-GMP to be involved in several possible binding scenarios. Monomeric c-di-GMP molecules could bind independently to two distinct binding

sites on DgcB or they could bind sequentially to a single site. Another possibility is binding of a c-di-GMP dimer to a single site on DgcB. C-di-GMP is known to form multimers (26). Additionally, crystal structures of DGCs show I-site bound c-di-GMP dimers (15-17). Recent NMR experiments indicate a dimerization dissociation constant for c-di-GMP in the range of 10 to 50 µM (M. Allan, unpublished data). Although c-di-GMP might bind to DgcB as monomer and dimer at the same time, it is highly unlikely in our eyes. Additionally we tried to prevent any overfitting introducing too many variables. Therefore, the last binding scenario is not followed any further.

To be able to identify the appropriate binding mode of c-di-GMP, kinetic measurements were performed. Analysis of the binding sensorgrams with eight distinct c-di-GMP concentrations was performed using the fitting algorithms of the BIAevaluation software estimating the k_{on} and the k_{off} values simultaneously (Figure 7). A negative slope can be observed in the association phase using higher analyte concentrations, which lead to deviations between the measured data and the fits. Although a physiological relevance of the observed behavior cannot be excluded, e.g. conformational change in the analyte-bound state, a methodological or a technical malfunction seems to be more probable (29,30).

Initially we have tried to fit the data not taking in account the dimerization of c-di-GMP. Fittings of the monomeric c-di-GMP binding sequentially to a single site or independently to two distinct sites, which are represented by the 'Bivalent analyte' and the 'Heterogeneous ligand – parallel reactions' fitting algorithms in the BIAevaluation software, result in very bad predictions with too high χ^2 values ($\chi^2 >$ 1.5) and in residuals, which are not tendency-free (Figure 7A, 7C, 7D). No drastic improvement could be observed performing the fits using the calculated concentrations of monomeric c-di-GMP at the assumed dimerization K_D's of 10 and 50 µM (see Materials&Methods). The fits improved drastically using calculated concentrations of dimeric c-di-GMP with assumed dimerization K_D's of 10 and 50 µM, respectively, and a 1:1 binding stoichiometry (Figures 7B, 7E, 7F). (The fits worsened using dimerization K_D's of 100 µM and 1 µM).

These results are supported by the Scatchard plot analysis of the equilibrium experiments. Usage of the calculated c-di-GMP dimer concentrations instead of the nominal concentrations results in data points lying on a straight (Figure 6B). Additionally, the independently from each other derived K_D values of c-di-GMP

dimers binding to DgcB from the equilibrium experiments and from the kinetics experiments (calculated from the rate constants) are almost identical (Table 1).

Thus, c-di-GMP binds as a dimer to a DgcB protomer with a K_D of 21 ± 14 nM assuming the c-di-GMP dimerization K_D being in range of 10 to 50 µM. Such c-di-GMP dimer concentration is present in a c-di-GMP solution with the nominal concentration of 0.51 ± 0.19 µM and 1.09 ± 0.38 µM, respectively (c-di-GMP dimerization K_D = 10 and 50 µM, respectively).

Binding of c-di-GMP induces tetramerization of DgcB

SEC elution profiles of DgcB showing two distinct peaks indicate propensity of DgcB to form higher oligomers. To analyze this behavior SEC-MALS experiments were performed. To take the dilution of proteins during the SEC in account the concentration of protein was measured at the SLS detector by a refractive index diffractometer (Table 2). The obtained molecular masses indicate presence of dimeric and tetrameric DgcB species (Table 2). Measurements of apo DgcB at three different concentrations of 26, 390 and 676 µM (initial!), respectively, show appearance of tetramers at rather high, non-physiological concentrations (Figures 8A, 8C and 8E, Table 2).

To see whether c-di-GMP has any influence on the quaternary state of DgcB, SEC-MALS experiments were carried out with DgcB samples, which were preincubated with three fold higher concentration of c-di-GMP. The buffer was not supplemented with c-di-GMP to save costs. Thus, complex dissociation during the chromatography run had to be expected. C-di-GMP complexed DgcB was assayed at two different initial concentrations (26 and 390 µM, respectively) (Figures 8B and 8D, Table 2). Indeed, both runs show dissociation of c-di-GMP from the protein, which can be identified by UV signal following peak 1. At lower protein concentration c-di-GMP completely separates from protein and elutes as a separate peak at the total column volume. Therefore, this protein sample shows the elution profile of apo-protein due to complete separation of c-di-GMP during the SEC run. C-di-GMP is also ablating from DgcB used at the higher concentration (390 µM, initially), which is indicated by the tailing off of the UV-signal, but the nucleotide doesn't completely dissociate from the protein. In contrast to the measurements of the apo-protein, which show co-presence of dimeric and tetrameric DgcB species (samples of 390 and 676 µM), the c-di-GMP preincubated protein (sample of 390 µM) is entirely tetrameric. Thus, apo-

DgcB shows a propensity to form tetramers at non-physiological concentrations (31,32), whereas addition of c-di-GMP strongly reduces the K_D of tetramer-dimer equilibrium.

Quantification of c-di-GMP induced tetramerization of DgcB monitored by analytical ultracentrifugation (AUC)

The effect of c-di-GMP inducing a shift in the DgcB dimer-tetramer equilibrium to lower concentrations was investigated using apo- and c-di-GMP saturated DgcB samples in the sedimentation equilibrium AUC experiments (Figure 9). Although DgcB is stable at concentrations used for this experiment, all AUC runs show protein losses due to aggregation, especially in presence of c-di-GMP. Neglecting the higher aggregates apo-DgcB is quasi dimeric (Table 3). In contrast, the c-di-GMP saturated DgcB sample reveals a dimer-tetramer equilibrium with a dissociation constant of 0.02 µM. Consistent with the ITC and SPR results, the AUC experiment yields a 1:1.8 stoichiometry between DgcB and c-di-GMP.

To monitor the influence of c-di-GMP on the hydrodynamic parameters of DgcB, sedimentation velocity AUC experiments were conducted (Figure 10). Addition of c-di-GMP turns out to influence the shape of both, dimeric and tetrameric, DgcB species (Table 4). Binding of c-di-GMP to DgcB results in modification of the axial ratio that is equivalent to elongation of the protein dimers. Increased spheroid a/b ratio of DgcB tetramers in comparison to the dimeric species indicates further elongation of the protein molecules.

Activation of DgcB by an unknown effector (small ligand)

The specific activity of DgcB decreases with a simultaneous release of phosphate as measured by a photometric assay (33). A four times lower enzymatic activity is observed two days after the purification compared to measurements with freshly purified DgcB. The protein is stable at 4 °C for weeks retaining its reduced enzymatic rate. Addition of phosphate doesn't restore the specific activity and also no specific binding can be detected by ITC (data not shown). Mass analysis of DgcB samples either treated or not with alkaline phosphatase yields identical masses. Therefore, activation by protein phosphorylation can be ruled out. The released phosphate must be part of a bigger and rather instable ligand, which must have a very low binding K_D. Since the well studied GGDEF domain is not known to interact with activators, it

must be a feature of the N-terminal dimerization domain. Further experiments are needed to elucidate the identity of this activator.

Discussion

Inability of GGDEF domains to dimerize autonomously

Cyclization of two GTP to a single c-di-GMP molecule by DGCs poses several requirements on this class of enzymes. Several studies have shown the necessity of a dimeric state for DGCs to be enzymatically active (14,34). This is in agreement with X-ray structures of DGCs showing GGDEF domains being able to bind only one substrate molecule (16). One of the main questions has always been – are GGDEF domains able to form dimers by themselves and so to produce c-di-GMP independently of associated regulatory domains?

Most proteins bearing the GGDEF domain possess regulatory dimerization domains (e.g. REC, PAS, PAC, GAF), transmembrane domains or both. On the other hand there are several GGDEF domain proteins with non-annotated N-terminal domains of ~ 25 to 500 aa length like DgcA and DgcB from *C. crescentus*. Although it is tempting to consider these as 'GGDEF domain stand-alone proteins', in case of DgcA and DgcB these N-terminal segments turn out to be involved in the dimerization process. Deletion of the N-terminal segment in DgcA resulting in a real 'GGDEF domain stand-alone protein' disables the dimer formation and therefore also the enzymatic activity of this protein.

The example of the two so-called 'stand-alone DGCs', DgcA and DgcB, reveals at least two different functions for the N-terminal parts of these proteins. In DgcA the role of the N-terminal segment is dedicated to formation of permanent protein dimers. Proteins like DgcA are regulated solely by its product c-di-GMP. In case of DgcB preliminary experiments indicate a regulatory role for the N-terminal 'domain' besides its function in dimerization.

It stays questionable whether there are true stand-alone DGCs at all and what function they might have. Interestigly there are no proteins in the SMART database (35) beginning N-terminally with a GGDEF domain. All of the so-called 'stand-alone' DGCs exhibit at least ~ twenty not-annotated amino acids in front of their GGDEF domains.

Regulation of DgcA and DgcB by c-di-GMP

Besides dimerization dependent regulation many DGCs are regulated through allosteric inhibition by their product c-di-GMP. The same has been observed for

DgcA, which is in agreement with the presence of primary and secondary I-sites in its amino acid sequence. Mutation studies have shown the necessity of the primary I-site RxxD motif for product inhibition in this protein (18). Our comparative sequence analysis shows presence of two arginins in the putatative secondary I-site. One of these residues at a topologically identical position in WspR was shown to interact with c-di-GMP (17). Therefore, inhibition of DgcA *via* cross-linking of the GGDEF domains is likely.

Non-competitive inhibition of DgcB that possesses a degenerated RxxD motif on the other hand contradicts the general opinion of the primary I-site's RxxD motif to be indispensable for c-di-GMP binding. This result is even more puzzling regarding the binding properties of c-di-GMP. The similar K_i/K_D values of c-di-GMP binding to DgcB, PleD (14) and DgcA (18) as well as binding of c-di-GMP in dimeric form to DgcB and PleD (14,16, P.Wassmann unpublished data) suggest a similar binding site in all three proteins. This would implement binding of c-di-GMP to the modified RxxD motif (**A**PP**R**). The consequence hereby would be binding of c-di-GMP to primary I-sites, which are lacking the exclusive RxxD motif. This would result in an extended number of c-di-GMP regulated DGCs. Mutational studies might prove this idea.

Additionally, DgcB exhibits c-di-GMP induced tetramerization. Our results show for the first time proofs for c-di-GMP induced protein inhibition by an intermolecular cross-linking mechanism. We propose in analogy to the intramolecular cross-linking mechanism that was observed in PleD (16) inhibition by GGDEF domain immobilization. The question arises how does c-di-GMP manage to stabilize the DgcB tetramers? Binding of c-di-GMP dimers to DgcB in a 1:1 stoichiometry implies binding of four c-di-GMP dimers to a DgcB tetramer. The same stoichiometry was observed in solution (AUC and ITC) assaying DgcB tetramers as well as in SPR experiments with immobilized protein that cannot tetramerize. Binding of c-di-GMP between identical sites in a DgcB tetramer should result in a decreasing stoichiometry – something that we do not observe. Therefore, such a behavior is only feasible in presence of two binding sites for c-di-GMP per a DgcB monomer, an independent one and one that binds only in combination with the other. As mentioned before, besides using the modified primary I-site for this task, DgcB could utilize R200 as the secondary I-site. It is localized in the α0-helix, one turn in front of the topologically analogous R313 of PleD's secondary I-site ($I_{S,DGC}$) (16). This arginine could cross-

link with the c-di-GMP loaded primary I-site leading to formation of tetramers. To prove this possibility mutagenesis might be the appropriate technique.

Elongation of the spheroid dimension ratio of DgcB tetramers in comparison to the dimeric species (Table 3) indicates an antiparallel arrangement of the dimers in a two-fold symmetric tetramer. Two distinct extremes are hereby imaginable. The DgcB dimers could form tetramers interacting solely *via* the cross-linking c-di-GMP molecules (kissing complex) or *via* extended protein-protein contacts in addition to the c-di-GMP cross-linking (hugging complex). The later extreme is in agreement with the very low tetramerization K_D of c-di-GMP loaded DgcB and is therefore a more probable one for DgcB. A DGC with such an arrangement has been observed before. The dimeric WspR forms in its crystal structure intercalated tetramers (17), but is stated by the authors as a crystal artifact. Similar arrangement might be utilized by DgcB in the c-di-GMP loaded state.

Intermolecular cross-linking of proteins via c-di-GMP opens new possibilities in the field of 'c-di-GMP-signaling'. Although c-di-GMP is known to be involved in regulation of several critical processes as virulence, biofilm formation and cell cycle progression (5,32,36) the amount of the identified c-di-GMP targets is rare, including PilZ domain proteins and riboswitches (11,13). One of the downstream effects of the second messenger c-di-GMP could be homo- and heteromolecular cross-linking of proteins modulating so their activity. Finding of such interactions with common techniques is a rather difficult task and might represent the reason for the little number of known downstream targets of c-di-GMP.

Acknowledgements

We thank PD Walter Huber and Josiane Kohler from F. Hoffmann-La Roche AG for giving us the opportunity to perform the SPR experiments using their instruments and consumables as well as for introducing in and teaching of the SPR technology.

Figure legends

Figure 1. Partial sequence alignment of DgcA and DgcB with diguanylate cyclases PleD and WspR,

(A) Schematic representation of the domain architecture of DgcA (Uniprot: Q9A3B9) and DgcB (Uniprot: Q9A776). The GGDEF domains are shown in green and the N-terminal in yellow.

(B) Truncated sequence alignments of GGDEF domains of DgcA DgcB and PleD from *C. crescentus* and WspR from *P. aeruginosa* [by CLUSTALX (37)]. The consensus sequence motif of the active site, GGDEF, is shaded in blue. The primary I-site, composed of the RxxD motif and in case of PleD of R390, are shaded in red and brown, respectively. Arginines in the α0-helix, which are or might be part of the secondary I-site, are shaded in green. Secondary structure elements of PleD derived from the crystal structure are shown above the sequence alignment. Arrows and bars represent β-strands and α-helices, respectively.

Figure 2. Time course of c-di-GMP production

C-di-GMP production of DgcA (A) and DgcB (B) in presence (circles) and absence (triangles) of the PDE YahA and of ΔDgcA (A) in presence of YahA (diamonds).

Figure 3. Inhibition of DgcB by c-di-GMP

Plot of specific activity of DgcB at various c-di-GMP concentrations. The experimental data (diamonds) were fitted with an exponential decay function (continuous line) yielding an IC_{50} of 20 µM at a protein concentration of 26 µM.

Figure 4. Binding of c-di-GMP to an allosteric site of DgcB

(A) Raw ITC data with 120 µM c-di-GMP titrated into 5 µM DgcB at 25 °C. (B) Integrated peaks of c-di-GMP titration to DgcB (circles) and to DgcB pre-incubated with 200 µM GTPαS (rectangles). (C) Integrated peaks of GTPαS titration to DgcB (rectangles) and to DgcB pre-incubated with 200 µM c-di-GMP (circles). Fits representing the binding of the ligands are shown as solid lines.

Figure 5. SPR experiment representing c-di-GMP steady-state binding to immobilized DgcB

Representation of a sensorgram plot showing steady-state binding of different c-di-GMP concentrations (1.30, 1.00, 0.78 and 0.65 µM shown in black, red, green and blue, respectively) to DgcB, which is immobilized on a NTA chip. The analyte samples were injected for six minutes and the dissociation was recorded for seven minutes. The steady-state was reached approximately 200 s after c-di-GMP injection initiation.

Figure 6. Scatchard plot analysis of the SPR steady-state experiment

Scatchard plot of c-di-GMP binding to immobilized DgcB. R_{eq} are measured RU values 250 seconds after the injection initiation. For the calculation of the $R_{eq}/c_{analyte}$ values nominal (A) and dimer (B) concentrations of c-di-GMP were used assuming a c-di-GMP dimerization K_D of 50 µM. The slope of the resulting straight is equal to the negative value of K_A and the intercept to $K_A \times R_{max}$.

Figure 7. Kinetic analysis of c-di-GMP binding to immobilized DgcB

Binding sensorgrams of c-di-GMP to DgcB. C-di-GMP was used at nominal concentrations of 50, 25, 12.5, 6.25, 3.13, 1.56, 0.78 and 0.39 µM. (A) Fits of the sensorgrams using the nominal c-di-GMP concentrations and the 'Heterogeneous ligand – parallel reactions' or the 'Bivalent analyte' fitting algorithm of the BIAevaluation software (green and red stripped lines, respectively). (B) Fits of the same sensorgrams using a simple '1:1 binding with drifting baseline' algorithm of the BIAevaluation software and the calculated concentrations of c-di-GMP dimers (see Materials&Methods) assuming a dimerization K_D of 50 or 10 µM (green and red stripped lines, respectively).

Residuals of the fits are shown in the lower panels (C, D, E, F).

Figure 8. Analysis of DgcB oligomerization state by SEC MALS

SEC MALS experiments were performed with DgcB at initial concentrations of 26 (A, B), 390 (C, D) and 676 µM (E) with apo-DgcB (A, C, E) and c-di-GMP pre-incubated DgcB (B, D). The analysis includes monitoring of the UV signal (red stripped line), differential refractive index signal (black dotted line) and the light scattering signal (black line) in the range between the column void volume (V_0) and

the total column volume (V_t). Blue lines indicate the calculated molecular weights of both peaks.

Figure 9. Sedimentation equilibrium runs of apo and c-di-GMP loaded DgcB

Representative plots of the SE runs of apo (red symbols) and c-di-GMP loaded (blue symbols) DgcB detecting the absorbance at 253 (A) and 280 nm (B), respectively. The black lines represent the global fits. Corresponding plots of the residuals are shown in the upper part of the figure.

Figure 10. Sedimentation coefficient distribution plot for DgcB

Sedimentation velocity analytical ultracentrifugation experiments were performed with DgcB (26 µM) in absence (stripped line) and presence of c-di-GMP (60µM) (full line).

Supplemental figure 1. BIACORE analysis of c-di-GMP binding to DgcA and ΔDgcA

Both His_6-tagged proteins were immobilized to two different channels of an NTA-Chip (GE Healthcare). Injections of c-di-GMP (31.25, 62.5, 125, 250, 500 µM, respectively) were applied to the channels with immobilized protein as well as to a reference channel. Due to a constant drift in the base line and lack of an appropriate regeneration condition for the immobilized protein c-di-GMP was injected in kinetic titration series (38). Sensorgrams, which are subtracted by the reference channel signal, show reproducible binding of c-di-GMP to DgcA (red line) and ΔDgcA (black and green lines). Neither the binding constants nor the thermodynamic parameters could be deduced from these experiments. The analysis failed due to a combination of erroneous factors: a complex drifting base line and a c-di-GMP solution containing additional, unknown ingredients, resulting in a solution effect that couldn't be compensated for.

Fit relevant c-di-GMP species	**SPR equilibrium experiment** (Scatchard)	**SPR kinetics experiments** (1:1 fitting algorithm)			
	binding K_D of c-di-GMP to DgcB (M)		**k_{on} (M^{-1} s^{-1})**	**k_{off} (s^{-1})**	**χ^2**
monomeric (nominal [c-di-GMP])	1.7×10^{-6}	2.4×10^{-7}	8.5×10^{3}	2.1×10^{-3}	1.3
dimeric (c-di-GMP dimerization K_D = 10 µM)	3.6×10^{-8}	2.8×10^{-8}	1.2×10^{5}	3.3×10^{-3}	0.5
dimeric (c-di-GMP dimerization K_D = 50 µM)	7.8×10^{-9}	8.1×10^{-9}	4.4×10^{5}	3.6×10^{-3}	0.5

Table 1: SPR derived kinetic and thermodynamic rate constants

Initial DgcB concentration (µM)	Incubation with c-di-GMP (µM)		Peak 1	Peak 2
26	**-**			**n.d.**
		peak limits (ml)	12.81 – 14.02	
		M_W (kDa)	71.7 ± 0.6	
		polydispersity (M_W / M_N)	1.000 ± 0.008	
		max. peak concentration (µM)	3.2	
26	**+ 80**			
		peak limits (ml)	12.92 – 13.76	10.12 – 12.17
		M_W (kDa)	72.5 ± 0.7	124.6 ± 12.5
		polydispersity (M_W / M_N)	1.000 ± 0.010	1.052 ± 0.142
		max. peak concentration (µM)	2.4	0.2
390	**-**			
		peak limits (ml)	13.21 – 13.58	12.15 – 12.63
		M_W (kDa)	78.2 ± 0.3	104.2 ± 0.5
		polydispersity (M_W / M_N)	1.002 ± 0.010	1.001 ± 0.009
		max. peak concentration (µM)	21.8	27.8
390	**+ 1000**		**n.d.**	
		peak limits (ml)		11.66 – 12.33
		M_W (kDa)		139.0 ± 0.1
		polydispersity (M_W / M_N)		1.002 ± 0.002
		max. peak concentration (µM)		29.4
676	**-**			
		peak limits (ml)	13.28 – 13.79	11.95 – 12.46
		M_W (kDa)	73.8 ± 0.2	121.7 ± 0.4
		polydispersity (M_W / M_N)	1.000 ± 0.004	1.003 ± 0.007
		max. peak concentration (µM)	50.5	32.5

Table 2: SEC-MALS results DgcB in absence and presence of c-di-GMP (n.d. not detected)

sample	residual protein concentration [µM]	abundance [%]		K_D of tetramer. [µM]
		dimer	tetramer	
DgcB	7.07	98.5	1.2	571
DgcB + c-di-GMP	2.56	9.1	83.8	0.02

Table 3: SE-AUC results of the DgcB ± c-di-GMP runs

sample	quaternary state	abundance [%]	S [10^{-13} s]	R_H [nm]	D [10^{-7} cm^2/s]	f/f_0	Ratio of spherical axes (a/b)	
							oblate	prolate
DgcB	dimer	98.4	3.81	3.8	4.167	1.35	4.24	4.52
DgcB + c-di-GMP	dimer	8.4	3.55	4.1	3.882	1.45	5.57	6.05
	tetramer	73.3	5.21	5.6	2.849	1.56	7.27	8.04
	higher aggregate	13.0	shoulder					

Table 4: SV-AUC results of the DgcB ± c-di-GMP runs

Figure 1A

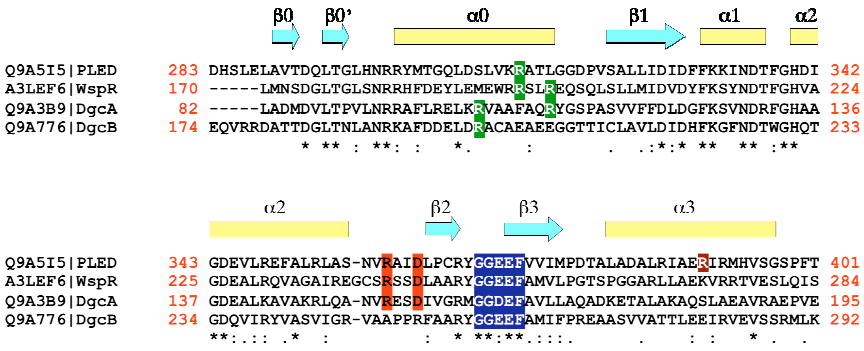


Figure 1B

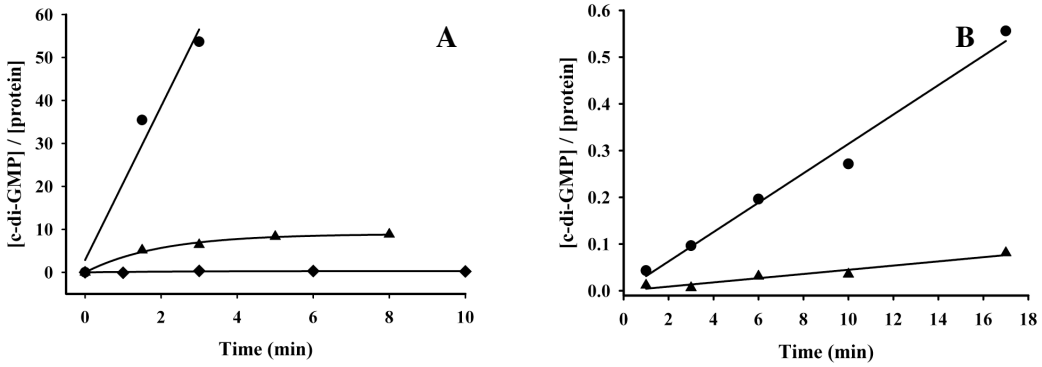


Figure 2

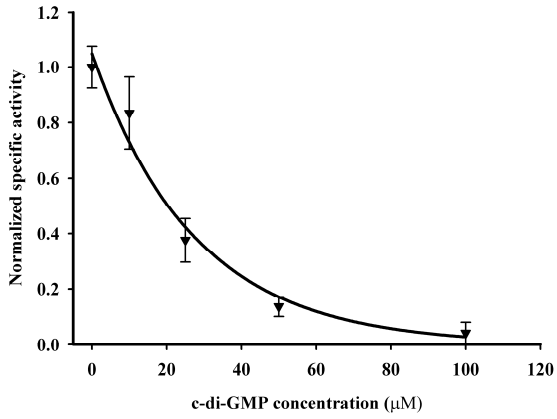


Figure 3

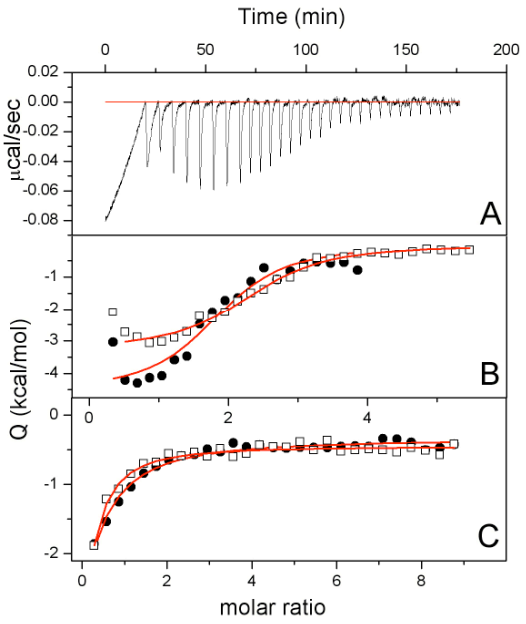


Figure 4

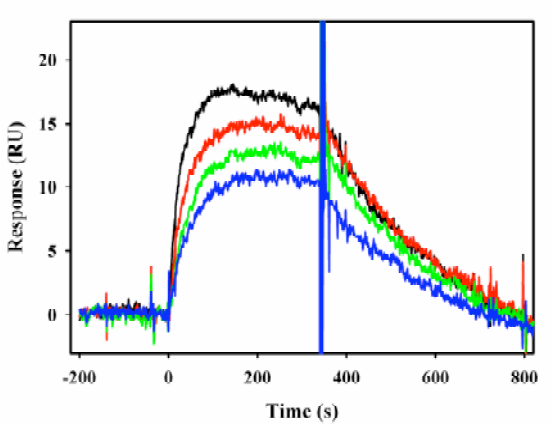


Figure 5

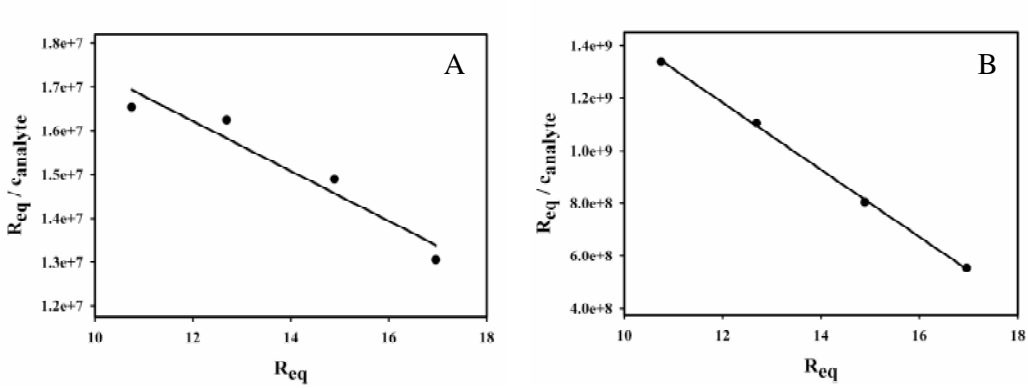


Figure 6

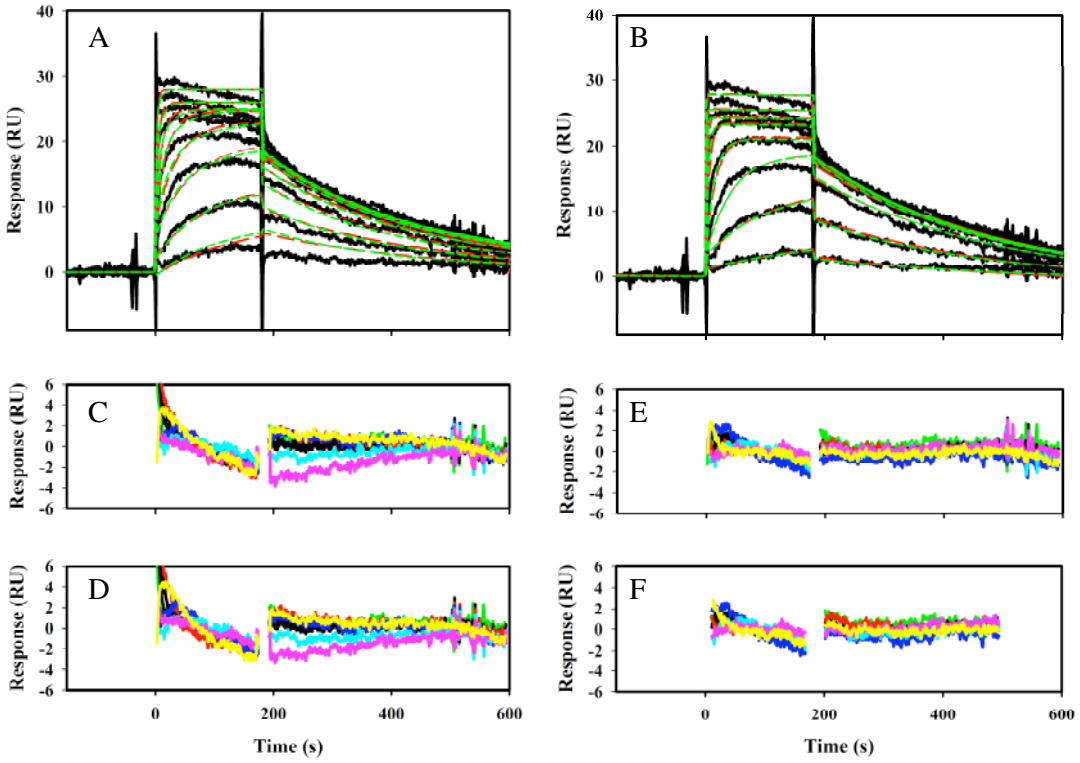


Figure 7

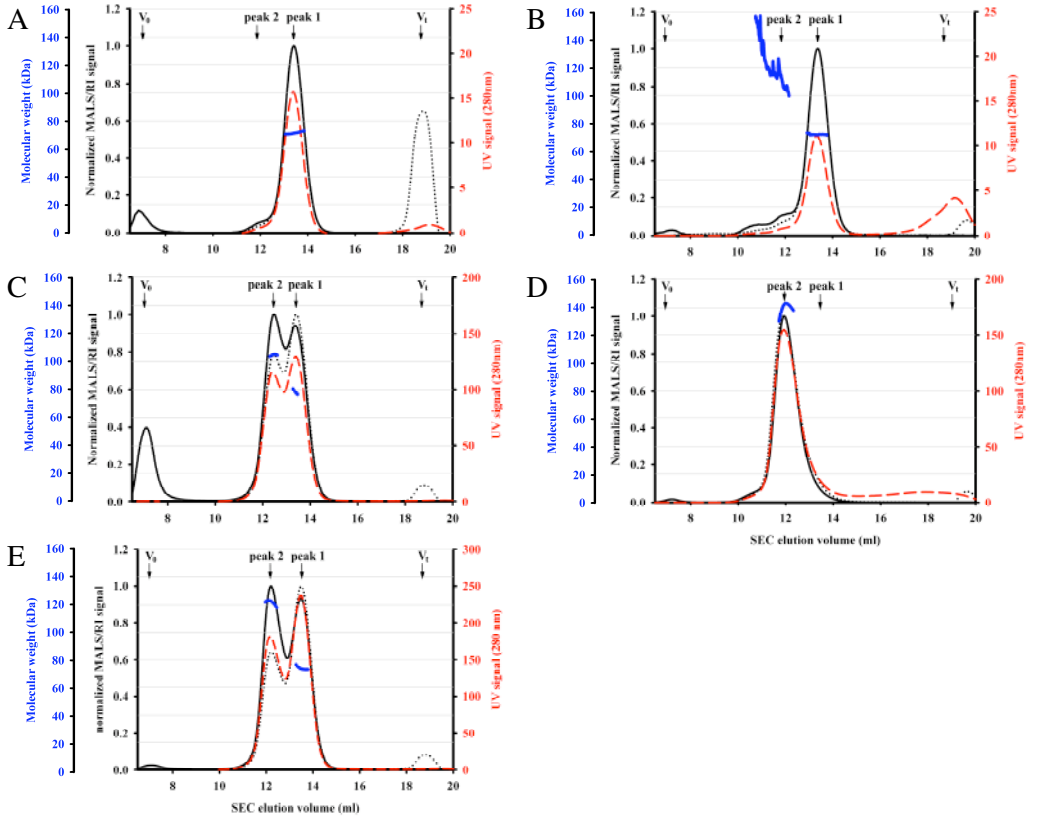


Figure 8

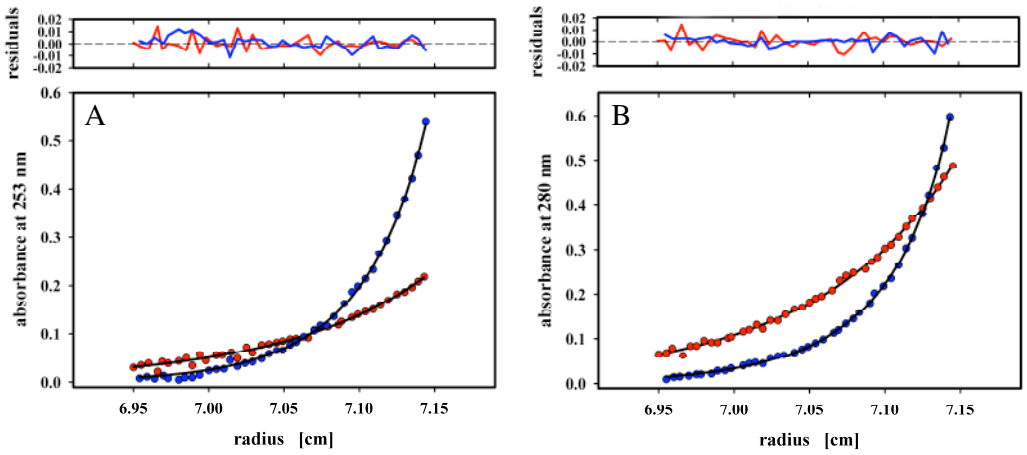


Figure 9

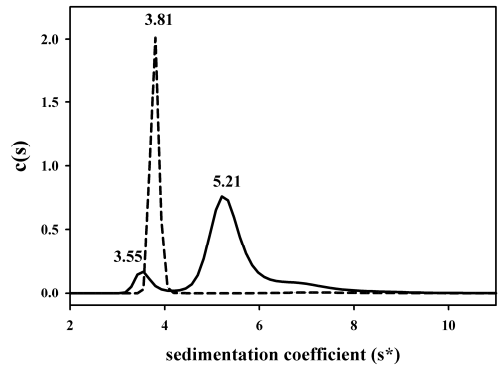
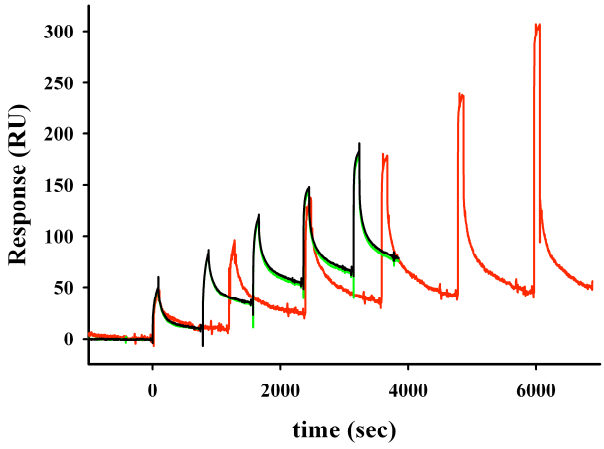


Figure 10



Supplemental figure 1

Bibliography

1. Ross, P., Weinhouse, H., Aloni, Y., Michaeli, D., Weinberger-Ohana, P., Mayer, R., Braun, S., de Vroom, E., van der Marel, G. A., van Boom, J. H., and Benziman, M. (1987) *Nature* **325**, 279-281
2. Galperin, M. Y. (2004) *Environ Microbiol* **6**, 552-567
3. Aldridge, P., Paul, R., Goymer, P., Rainey, P., and Jenal, U. (2003) *Mol Microbiol* **47**, 1695-1708
4. Romling, U., Gomelsky, M., and Galperin, M. Y. (2005) *Mol Microbiol* **57**, 629-639
5. Jenal, U., and Malone, J. (2006) *Annu Rev Genet* **40**, 385-407
6. Ryan, R. P., Fouhy, Y., Lucey, J. F., Jiang, B. L., He, Y. Q., Feng, J. X., Tang, J. L., and Dow, J. M. (2007) *Mol Microbiol* **63**, 429-442
7. Cotter, P. A., and Stibitz, S. (2007) *Curr Opin Microbiol* **10**, 17-23
8. Ryan, R. P., Fouhy, Y., Lucey, J. F., Crossman, L. C., Spiro, S., He, Y. W., Zhang, L. H., Heeb, S., Camara, M., Williams, P., and Dow, J. M. (2006) *Proc Natl Acad Sci U S A* **103**, 6712-6717
9. Duerig, A., Abel, S., Folcher, M., Nicollier, M., Schwede, T., Amiot, N., Giese, B., and Jenal, U. (2009) *Genes Dev* **23**, 93-104
10. Fineran, P. C., Williamson, N. R., Lilley, K. S., and Salmond, G. P. (2007) *J Bacteriol* **189**, 7653-7662
11. Amikam, D., and Galperin, M. Y. (2006) *Bioinformatics* **22**, 3-6
12. Benach, J., Swaminathan, S. S., Tamayo, R., Handelman, S. K., Folta-Stogniew, E., Ramos, J. E., Forouhar, F., Neely, H., Seetharaman, J., Camilli, A., and Hunt, J. F. (2007) *Embo J* **26**, 5153-5166
13. Sudarsan, N., Lee, E. R., Weinberg, Z., Moy, R. H., Kim, J. N., Link, K. H., and Breaker, R. R. (2008) *Science* **321**, 411-413
14. Paul, R., Abel, S., Wassmann, P., Beck, A., Heerklotz, H., and Jenal, U. (2007) *J Biol Chem* **282**, 29170-29177
15. Chan, C., Paul, R., Samoray, D., Amiot, N. C., Giese, B., Jenal, U., and Schirmer, T. (2004) *Proc Natl Acad Sci U S A* **101**, 17084-17089
16. Wassmann, P., Chan, C., Paul, R., Beck, A., Heerklotz, H., Jenal, U., and Schirmer, T. (2007) *Structure* **15**, 915-927
17. De, N., Pirruccello, M., Krasteva, P. V., Bae, N., Raghavan, R. V., and Sondermann, H. (2008) *PLoS Biol* **6**, e67
18. Christen, B., Christen, M., Paul, R., Schmid, F., Folcher, M., Jenoe, P., Meuwly, M., and Jenal, U. (2006) *J Biol Chem* **281**, 32015-32024
19. Schmidt, A. J., Ryjenkov, D. A., and Gomelsky, M. (2005) *J Bacteriol* **187**, 4774-4781
20. Pierce, M. M., Raman, C. S., and Nall, B. T. (1999) *Methods* **19**, 213-221
21. Schuck, P. (2000) *Biophys J* **78**, 1606-1619
22. Schuck, P. (1994) in *Progress in colloid and polymer science* Vol. 94, pp. 1-13, Kremer, F., Richtering, W.
23. Lebowitz, J., Lewis, M. S., and Schuck, P. (2002) *Protein Sci* **11**, 2067-2079
24. Schuck, P., Perugini, M. A., Gonzales, N. R., Howlett, G. J., and Schubert, D. (2002) *Biophys J* **82**, 1096-1111
25. Lupas, A. (1996) *Methods Enzymol* **266**, 513-525
26. Zhang, Z., Kim, S., Gaffney, B. L., and Jones, R. A. (2006) *J Am Chem Soc* **128**, 7015-7024

27. Egli, M., Gessner, R. V., Williams, L. D., Quigley, G. J., van der Marel, G. A., van Boom, J. H., Rich, A., and Frederick, C. A. (1990) *Proc Natl Acad Sci U S A* **87**, 3235-3239
28. Nieba, L., Nieba-Axmann, S. E., Persson, A., Hamalainen, M., Edebratt, F., Hansson, A., Lidholm, J., Magnusson, K., Karlsson, A. F., and Pluckthun, A. (1997) *Anal Biochem* **252**, 217-228
29. Cannon, M. J., Papalia, G. A., Navratilova, I., Fisher, R. J., Roberts, L. R., Worthy, K. M., Stephen, A. G., Marchesini, G. R., Collins, E. J., Casper, D., Qiu, H., Satpaev, D., Liparoto, S. F., Rice, D. A., Gorshkova, II, Darling, R. J., Bennett, D. B., Sekar, M., Hommema, E., Liang, A. M., Day, E. S., Inman, J., Karlicek, S. M., Ullrich, S. J., Hodges, D., Chu, T., Sullivan, E., Simpson, J., Rafique, A., Luginbuhl, B., Westin, S. N., Bynum, M., Cachia, P., Li, Y. J., Kao, D., Neurauter, A., Wong, M., Swanson, M., and Myszka, D. G. (2004) *Anal Biochem* **330**, 98-113
30. Nordin, H., Jungnelius, M., Karlsson, R., and Karlsson, O. P. (2005) *Anal Biochem* **340**, 359-368
31. Simm, R., Morr, M., Remminghorst, U., Andersson, M., and Romling, U. (2009) *Anal Biochem* **386**, 53-58
32. Hengge, R. (2009) *Nat Rev Microbiol* **7**, 263-273
33. Baykov, A. A., Evtushenko, O. A., and Avaeva, S. M. (1988) *Anal Biochem* **171**, 266-270
34. Paul, R., Weiser, S., Amiot, N. C., Chan, C., Schirmer, T., Giese, B., and Jenal, U. (2004) *Genes Dev* **18**, 715-727
35. Letunic, I., Copley, R. R., Schmidt, S., Ciccarelli, F. D., Doerks, T., Schultz, J., Ponting, C. P., and Bork, P. (2004) *Nucleic Acids Res* **32**, D142-144
36. Romling, U., and Amikam, D. (2006) *Curr Opin Microbiol* **9**, 218-228
37. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) *Nucleic Acids Res* **22**, 4673-4680
38. Karlsson, R., Katsamba, P. S., Nordin, H., Pol, E., and Myszka, D. G. (2006) *Anal Biochem* **349**, 136-147

III CONCLUSIONS AND PERSPECTIVES

The DGCs are known to be enzymatically active as dimers (43). This work has shown that isolated GGDEF domains are not able to form dimers and that therefore the so-called 'stand-alone' DGCs use their N-terminal segments for dimerization.

Hereby, two different kinds of 'stand-alone' DGCs were identified. The permanently dimeric and active 'steady-state keepers' like DgcA produce c-di-GMP independently of any environmental cues sustaining a constant c-di-GMP level. The enzymatic activity of these proteins is limited solely by the allosteric product inhibition. The dimeric 'signal amplifiers' like DgcB, on the other hand, use their N-terminal dimerization segments/domains additionally for modulation of their enzymatic activity in response to environmental cues. Future investigations must aim at the identification and the characterization of these regulatory dimerization domains in order to get insights into the upstream signaling cascade components.

To use the dimerization process itself as a regulatory mechanism is the concern of several DGCs, e.g. the ones bearing N-terminal Rec-domains. The crystal structure of $BeF_3^- \bullet Mg^{2+}$ activated PleD is one of the first structures showing a full-length response regulator in its activated state. Comparison with the structure of non-activated PleD (22) resulted in the elucidation of the molecular mechanisms of the dimerization process. Additionally, the formation of a two-fold symmetric, charged pocket at the $(Rec1\text{-}Rec2)_2$ stem interface probably explains the role of the non-phosphorylatable Rec2 domain in PleD and might represent the long-sought 'pole-localization'-signal for PleD and other Rec1-Rec2-GGDEF domain proteins.

Besides giving insights in the substrate binding mode of the DGCs, the obtained crystal structures of activated PleD shed light on the catalytic mechanism of DGCs. In combination with biochemical data the structures verified the 'two-metal assisted' catalysis mechanism for the DGCs. In future, combination of mutational and structural examinations should prove the relevance of the proposed mechanistic aspects for the catalysis: (I) the occurrence of conformational changes during the formation of the dimeric active site in the GGDEF domains; (II) the crucial role of the GGDEF sequence motif's glutamate in the process of quasi simultaneous phosphodiester-bond formation as well as in the process of the product-release.

The most-inspiring scientific results derive from the investigations concerning the allosteric product inhibition mechanism. The crystal structures of activated PleD

revealed a new, c-di-GMP dimer dependent domain-cross-linking mode that is generally applicable to DGCs, involving in the process merely the GGDEF domains. It turned out that the successful inhibition of the DGC PleD relies on the presence of primary- and secondary I-sites, whereas the initial binding of c-di-GMP depends mainly on the primary I-site.

The minimal composition of the primary I-site stays yet unclear. Whereas PleD was shown to need both the RxxD motif and the R390 in order to bind c-di-GMP, other DGCs don't seem to utilize the arginine at an analogue position to PleD's R390 for the binding process. Further complication of the c-di-GMP binding process poses the allosteric inhibition of DgcB that has a degenerated RxxD motif. Combination of mutagenesis and structural elucidation of the c-di-GMP-DgcB complex will be an imperative step in the global understanding of the allosteric product inhibition.

Finally, the shown ability of c-di-GMP to cross-link proteins, intramolecularly in PleD and intermolecularly in DgcB, opens new possibilities for the second messenger in response-modulation of the signal transduction components. If c-di-GMP turns out to exert its role in cross-linking downstream effectors leading subsequently to a specific cellular response, it will be inevitable to combine microbiological, biochemical, biophysical (SPR) and structural (x-ray crystallography and electron microscopy) techniques to elucidate the underlying mechanisms.

IV BIBLIOGRAPHY

1. Camilli, A., and Bassler, B. L. (2006) *Science* **311**, 1113-1116
2. Linder, J. U., and Schultz, J. E. (2008) *Curr Opin Struct Biol* **18**, 667-672
3. Shenoy, A. R., and Visweswariah, S. S. (2006) *Trends Microbiol* **14**, 543-550
4. Shenoy, A. R., and Visweswariah, S. S. (2006) *FEBS Lett* **580**, 3344-3352
5. Potrykus, K., and Cashel, M. (2008) *Annu Rev Microbiol* **62**, 35-51
6. Srivatsan, A., and Wang, J. D. (2008) *Curr Opin Microbiol* **11**, 100-105
7. Witte, G., Hartung, S., Buttner, K., and Hopfner, K. P. (2008) *Mol Cell* **30**, 167-178
8. Romling, U. (2008) *Sci Signal* **1**, pe39
9. Hengge, R. (2009) *Nat Rev Microbiol* **7**, 263-273
10. Tamayo, R., Pratt, J. T., and Camilli, A. (2007) *Annu Rev Microbiol* **61**, 131-148
11. Pesavento, C., and Hengge, R. (2009) *Curr Opin Microbiol* **12**, 170-176
12. Cotter, P. A., and Stibitz, S. (2007) *Curr Opin Microbiol* **10**, 17-23
13. Ryan, R. P., Fouhy, Y., Lucey, J. F., and Dow, J. M. (2006) *J Bacteriol* **188**, 8327-8334
14. Jenal, U., and Malone, J. (2006) *Annu Rev Genet* **40**, 385-407
15. Romling, U., and Amikam, D. (2006) *Curr Opin Microbiol* **9**, 218-228
16. Romling, U., Gomelsky, M., and Galperin, M. Y. (2005) *Mol Microbiol* **57**, 629-639
17. Ross, P., Weinhouse, H., Aloni, Y., Michaeli, D., Weinberger-Ohana, P., Mayer, R., Braun, S., de Vroom, E., van der Marel, G. A., van Boom, J. H., and Benziman, M. (1987) *Nature* **325**, 279-281
18. Egli, M., Gessner, R. V., Williams, L. D., Quigley, G. J., van der Marel, G. A., van Boom, J. H., Rich, A., and Frederick, C. A. (1990) *Proc Natl Acad Sci U S A* **87**, 3235-3239
19. Zhang, Z., Kim, S., Gaffney, B. L., and Jones, R. A. (2006) *J Am Chem Soc* **128**, 7015-7024
20. Beyhan, S., and Yildiz, F. H. (2007) *Mol Microbiol* **63**, 995-1007
21. Kader, A., Simm, R., Gerstel, U., Morr, M., and Romling, U. (2006) *Mol Microbiol* **60**, 602-616
22. Chan, C., Paul, R., Samoray, D., Amiot, N. C., Giese, B., Jenal, U., and Schirmer, T. (2004) *Proc Natl Acad Sci U S A* **101**, 17084-17089
23. De, N., Pirruccello, M., Krasteva, P. V., Bae, N., Raghavan, R. V., and Sondermann, H. (2008) *PLoS Biol* **6**, e67
24. Minasov, G., Padavattan, S., Shuvalova, L., Brunzelle, J. S., Miller, D. J., Basle, A., Massa, C., Collart, F. R., Schirmer, T., and Anderson, W. F. (2009) *J Biol Chem* **284**, 13174-13184
25. Benach, J., Swaminathan, S. S., Tamayo, R., Handelman, S. K., Folta-Stogniew, E., Ramos, J. E., Forouhar, F., Neely, H., Seetharaman, J., Camilli, A., and Hunt, J. F. (2007) *Embo J* **26**, 5153-5166
26. Christen, M., Christen, B., Allan, M. G., Folcher, M., Jeno, P., Grzesiek, S., and Jenal, U. (2007) *Proc Natl Acad Sci U S A* **104**, 4112-4117
27. Tal, R., Wong, H. C., Calhoon, R., Gelfand, D., Fear, A. L., Volman, G., Mayer, R., Ross, P., Amikam, D., Weinhouse, H., Cohen, A., Sapir, S., Ohana, P., and Benziman, M. (1998) *J Bacteriol* **180**, 4416-4425
28. Galperin, M. Y., Nikolskaya, A. N., and Koonin, E. V. (2001) *FEMS Microbiol Lett* **203**, 11-21

29. Duerig, A., Abel, S., Folcher, M., Nicollier, M., Schwede, T., Amiot, N., Giese, B., and Jenal, U. (2009) *Genes Dev* **23**, 93-104
30. Fineran, P. C., Williamson, N. R., Lilley, K. S., and Salmond, G. P. (2007) *J Bacteriol* **189**, 7653-7662
31. Whitchurch, C. B., and Mattick, J. S. (1994) *Mol Microbiol* **13**, 1079-1091
32. Comolli, J. C., Hauser, A. R., Waite, L., Whitchurch, C. B., Mattick, J. S., and Engel, J. N. (1999) *Infect Immun* **67**, 3625-3630
33. Lynch, A. S., and Robertson, G. T. (2008) *Annu Rev Med* **59**, 415-428
34. Huang, D. B., Mohanty, A., DuPont, H. L., Okhuysen, P. C., and Chiang, T. (2006) *J Med Microbiol* **55**, 1303-1311
35. Huang, B., Whitchurch, C. B., and Mattick, J. S. (2003) *J Bacteriol* **185**, 7068-7076
36. Kazmierczak, B. I., Lebron, M. B., and Murray, T. S. (2006) *Mol Microbiol* **60**, 1026-1043
37. Simm, R., Morr, M., Kader, A., Nimtz, M., and Romling, U. (2004) *Mol Microbiol* **53**, 1123-1134
38. Garcia, B., Latasa, C., Solano, C., Garcia-del Portillo, F., Gamazo, C., and Lasa, I. (2004) *Mol Microbiol* **54**, 264-277
39. Hickman, J. W., Tifrea, D. F., and Harwood, C. S. (2005) *Proc Natl Acad Sci U S A* **102**, 14422-14427
40. Kirillina, O., Fetherston, J. D., Bobrov, A. G., Abney, J., and Perry, R. D. (2004) *Mol Microbiol* **54**, 75-88
41. Lee, S. H., Butler, S. M., and Camilli, A. (2001) *Proc Natl Acad Sci U S A* **98**, 6889-6894
42. Tischler, A. D., and Camilli, A. (2005) *Infect Immun* **73**, 5873-5882
43. Paul, R., Weiser, S., Amiot, N. C., Chan, C., Schirmer, T., Giese, B., and Jenal, U. (2004) *Genes Dev* **18**, 715-727
44. Ryjenkov, D. A., Tarutina, M., Moskvin, O. V., and Gomelsky, M. (2005) *J Bacteriol* **187**, 1792-1798
45. Pei, J., and Grishin, N. V. (2001) *Proteins* **42**, 210-216
46. Christen, B., Christen, M., Paul, R., Schmid, F., Folcher, M., Jenoe, P., Meuwly, M., and Jenal, U. (2006) *J Biol Chem* **281**, 32015-32024
47. Christen, M., Christen, B., Folcher, M., Schauerte, A., and Jenal, U. (2005) *J Biol Chem* **280**, 30829-30837
48. Schmidt, A. J., Ryjenkov, D. A., and Gomelsky, M. (2005) *J Bacteriol* **187**, 4774-4781
49. Tamayo, R., Tischler, A. D., and Camilli, A. (2005) *J Biol Chem* **280**, 33324-33330
50. Galperin, M. Y. (2004) *Environ Microbiol* **6**, 552-567
51. Delgado-Nixon, V. M., Gonzalez, G., and Gilles-Gonzalez, M. A. (2000) *Biochemistry* **39**, 2685-2691
52. Hasegawa, K., Masuda, S., and Ono, T. A. (2006) *Biochemistry* **45**, 3785-3793
53. Nakasone, Y., Ono, T. A., Ishii, A., Masuda, S., and Terazima, M. (2007) *J Am Chem Soc* **129**, 7028-7035
54. Hurley, J. H. (2003) *Sci STKE* **2003**, PE1
55. Amikam, D., and Galperin, M. Y. (2006) *Bioinformatics* **22**, 3-6
56. Ryjenkov, D. A., Simm, R., Romling, U., and Gomelsky, M. (2006) *J Biol Chem* **281**, 30310-30314
57. Lee, V. T., Matewish, J. M., Kessler, J. L., Hyodo, M., Hayakawa, Y., and Lory, S. (2007) *Mol Microbiol* **65**, 1474-1484

58. Sudarsan, N., Lee, E. R., Weinberg, Z., Moy, R. H., Kim, J. N., Link, K. H., and Breaker, R. R. (2008) *Science* **321**, 411-413
59. Paul, R., Abel, S., Wassmann, P., Beck, A., Heerklotz, H., and Jenal, U. (2007) *J Biol Chem* **282**, 29170-29177
60. Paul, R., Jaeger, T., Abel, S., Wiederkehr, I., Folcher, M., Biondi, E. G., Laub, M. T., and Jenal, U. (2008) *Cell* **133**, 452-461

V APPENDIX

Acknowledgements

This dissertation was conducted at the Structural Biology unit of the Biozentrum Basel under the supervision of Prof. Tilman Schirmer and in a close collaboration with Prof. Urs Jenal.

Formost, I want to express my special gratitude to my supervisor, Prof. Tilman Schirmer, giving me the opportunity and the freedom to work on such a demanding and at the same time exciting project. He guided me through my PhD as a constructive critic, being highly supportive during the process of structure elucidation and in cases concerning theoretical issues.

The success of this dissertation is affiliated to its interdisciplinary character. I am grateful to all of my collaborators. Many thanks go to Andreas Beck for performance of the ITC experiments as well as to Francis Mueller and Eric Kusznir for carrying out the AUC experiments. Additionally, I want to sincerely thank people for their support teaching me several techniques: Walter Huber and Josianne Kohler (SPR); Prof. Dagmar Klostermeier and Martin Linden (FRET); Therese Schulthess (CD); and Marco Rogowski (ESI-TOF MS). Besides, I, "the bad boy", am strongly indebted to Arnaud Basle, who familiarized me with the lab specific pecularities in the initial phase of my work, and who had taught me molecular biology.

I am grateful to all present and former members of the lab for their continuous hearty and often psychological support: Arnaud Basle, Christophe Wirth, Marcel Meury, Claudia Massa, Ludwig Zumthor, Franziska Zaehringer, Arnaud Goepfert, Caroline Peneff and Zora Marcovic-Housley. For the breaks in form of exhausting disputations concerning "all the world and his wife" I want to thank Dietrich Samoray.

My special thanks go to my wife Elena, who has supported me with her love and faith sharing all the ups and downs, indulgent to my every temper, and who has always encouraged me to go on.

Last but not least I want to express my gratitude to my parents and my sister for their unconditional assistance.