Research Article



The GDP-switched GAF domain of DcpA modulates the concerted synthesis/hydrolysis of c-di-GMP in *Mycobacterium smegmatis*

Hui-Jie Chen^{1,2}, Na Li³, Ye Luo^{1,2}, Yong-Liang Jiang^{1,2}, Cong-Zhao Zhou^{1,2}, Yuxing Chen^{1,2} and Qiong Li^{1,2}

¹Hefei National Laboratory for Physical Sciences at the Microscale and School of Life Sciences, University of Science and Technology of China, Hefei, Anhui 230027, China; ²Key Laboratory of Structural Biology, Chinese Academy of Science, Hefei, Anhui 230027, China; ³National Center for Protein Science Shanghai, Shanghai 201203, China

Correspondence: Qiong Li (liqiong@ustc.edu.cn) or Na Li (nli@sibcb.ac.cn)

The second messenger c-di-GMP [bis-(3'-5')-cyclic dimeric guanosine monophosphate] plays a key role in bacterial growth, survival and pathogenesis, and thus its intracellular homeostasis should be finely maintained. Mycobacterium smegmatis encodes a GAF (mammalian cGMP-regulated phosphodiesterases, Anabaena adenylyl cyclases and Escherichia coli transcription activator FhIA) domain containing bifunctional enzyme DcpA (diguanylate cyclase and phosphodiesterase A) that catalyzes the synthesis and hydrolysis of c-di-GMP. Here, we found that *M. smegmatis* DcpA catalyzes the hydrolysis of c-di-GMP at a higher velocity, compared with synthetic activity, resulting in a sum reaction from the ultimate substrate GTP to the final product pGpG [5'-phosphoguanylyl-(3'-5')-guanosine]. Fusion with the N-terminal GAF domain enables the GGDEF (Gly-Gly-Asp-Glu-Phe) domain of DcpA to dimerize and accordingly gain synthetic activity. Screening of putative metabolites revealed that GDP is the ligand of the GAF domain. Binding of GDP to the GAF domain down-regulates synthetic activity, but up-regulates hydrolytic activity, which, in consequence, might enable a timely response to the transient accumulation of c-di-GMP at the stationary phase or under stresses. Combined with the crystal structure of the EAL (Glu-Ala-Leu) domain and the small-angle X-ray scattering data, we propose a putative regulatory model of the GAF domain finely tuned by the intracellular GTP/GDP ratio. These findings help us to better understand the concerted control of the synthesis and hydrolysis of c-di-GMP in M. smegmatis in various microenvironments.

Introduction

Bis-(3',5')-cyclic dimeric guanosine monophosphate (c-di-GMP) is an important second messenger produced by a variety of bacteria, which was first identified as an allosteric activator of cellulose synthase in *Gluconacetobacter xylinus* 30 years ago [1]. As a ubiquitous intracellular signaling molecule, c-di-GMP plays key roles in many cellular processes to help bacteria adapt to diverse environments [2]. For instance, it stimulates the synthesis of various adhesins and exopolysaccharide matrix substances in biofilms [3], or reduces cell motility by down-regulating flagellar expression or assembly [4]. Moreover, c-di-GMP regulates the expression of some bacterial virulence genes, as demonstrated in *Vibrio cholerae* [5,6]. Upon activating the stimulator of interferon genes (STING), c-di-GMP induces the secretion of type I interferon in response to microbial invasion [7]. In addition, the c-di-GMP signaling network usually cross-talks with other global regulatory pathways, such as the phosphorylation networks [8], quorum-sensing pathways [9] and other second messenger (e.g. cyclic GMP, cyclic AMP and guanosine tetraphosphate) involved pathways [10–12]. Altogether, c-di-GMP is essential for the growth, survival and pathogenesis of bacteria.

Received: 28 January 2018 Revised: 12 March 2018 Accepted: 16 March 2018

Accepted Manuscript online: 19 March 2018 Version of Record published: 9 April 2018



In fact, the intracellular homeostasis of c-di-GMP is finely tuned by a couple of enzymes with opposing activities, namely diguanylate cyclases (DGCs) and phosphodiesterases (PDEs) [13]. One molecule of c-di-GMP is synthesized by DGCs with two molecules of GTP, or degraded into one molecule of 5'-phosphoguanylyl-(3'-5')-guanosine (pGpG) or two molecules of GMP by PDEs. DGCs typically harbor a GGDEF (Gly-Gly-Asp-Glu-Phe) domain, named after the conserved GG(D/E)EF active site [14], which is critical for enzymatic activity [15]. PDEs containing an EAL (Glu-Ala-Leu) domain can hydrolyze c-di-GMP into the linear pGpG [16], which in some cases could be further cleaved into GMP by oligoribonucleases [17], whereas the HD-GYP-containing PDEs produce the ultimate products of two molecules of GMP [18]. Likewise, the EAL domain or HD-GYP domain is also named after the conserved motif at the active site.

The GGDEF domain and the EAL (or HD-GYP) domain could be encoded by two individual genes or a single fusion gene [19]. Interestingly, a fusion protein with both GGDEF and EAL domains usually harbors one or several regulatory domains at the N-terminus (Figure 1a), which are supposed to modulate the enzymatic activities in response to the internal or external stresses [2]. The regulatory domains fused with the isolated GGDEF domain or the EAL domain have been extensively investigated [20,21]; however, the regulatory mechanisms that co-ordinate the combined GGDEF and EAL domains in a single protein remain largely unknown. Among the six representatives in Figure 1a, *Legionella pneumophila* Lpl0329 possesses two regulatory domains, namely a receiver (REC) domain and a Per-Arnt-Sim (PAS) domain. Phosphorylation of the REC domain of Lpl0329 results in a decrease of the synthetic activity toward c-di-GMP, but no effect to the hydrolytic activity [22]. In addition, the PAS and PAS-associated C-terminal motif (PAC) regulatory domains of *Streptomyces coelicolor* RmdA were shown to specifically bind to hemin, the regulatory mechanism of which remains unclear [23].

The intracellular human pathogen *Mycobacterium tuberculosis* is a causative agent of contagious tuberculosis and leads to millions of deaths annually around the world. Rv1354c is composed of the GGDEF and EAL domains in tandem and an N-terminal GAF domain (Figure 1a), which was named after mammalian cGMP-regulated PDEs, *Anabaena* adenylyl cyclases and *E. coli* transcription activator FhIA [24]. Rv1354c was proposed to be responsible for the synthesis and hydrolysis of c-di-GMP in *M. tuberculosis* [25]. Owing to the ease of growth and non-pathogenicity, *Mycobacterium smegmatis* usually serves as a model organism to study *M. tuberculosis* [26]. The ortholog of Rv1354c, *M. smegmatis* DcpA (diguanylate cyclase and phosphodiesterase <u>A</u>), shares a sequence identity of 63% with the counterpart domain in *M. tuberculosis*. Deletion of *dcpA* weakens long-term survival under nutrient starvation, whereas the phenotype can be complemented by Rv1354c [27]. However, the precise regulatory role of the N-terminal GAF domain of DcpA in co-ordinating the synthesis and hydrolysis of c-di-GMP remains unknown to date.

Here, we found that *M. smegmatis* DcpA indeed possesses dual activities to catalyze both the synthesis and hydrolysis of c-di-GMP using GTP as the ultimate substrate; however, the hydrolytic activity is higher than the synthetic activity, resulting in a sum reaction of producing only the final product pGpG. Notably, the dimerization mediated by the N-terminal GAF domain is indispensable for the synthetic activity of the GGDEF domain, whereas the full-length DcpA possesses a much higher hydrolytic activity compared with the isolated EAL domain. After screening the putative metabolites, we identified that GDP is the ligand of the GAF domain. Binding of GDP to the GAF domain decreases synthetic activity, but increases hydrolytic activity, which in consequence leads to the accelerated consumption and limited production of c-di-GMP, reminiscent of the decrease of the GTP/GDP ratio that enables the timely elimination of the accumulated c-di-GMP at the stationary phase or under stresses. In addition, we solved the crystal structure of the EAL domain, each subunit of which folds into a classic triosephosphate isomerase (TIM) barrel structure. Combining with the small-angle X-ray scattering (SAXS) data, we propose a putative model in which GDP binding to the GAF domain triggers conformational changes in the overall dimeric structure, which in consequence co-ordinates the dual activities of DcpA. These findings help us to better understand the fine control of the synthesis and hydrolysis of c-di-GMP in M. smegmatis and might provide hints for studying its homolog in M. tuberculosis.

Experimental procedures Cloning, expression and purification of DcpA and mutants

The coding region of DcpA ($MSMEG_{2196}$) was amplified from the genomic DNA of M. smegmatis str. mc² 155 using the primers in Supplementary Table S1, and was cloned into a pET28a-derived vector with an





Figure 1. DcpA possesses dual activities.

(a) Domain organization of several proteins from Gram-positive and Gram-negative bacteria that contain tandem GGDEF and EAL domains. The boundaries of each domain of DcpA were defined according to the multi-sequence alignment. The NCBI accession numbers of all the proteins are as follows: *Mycobacterium smegmatis* DcpA, WP_003893571.1; *Mycobacterium tuberculosis* Rv1354c, NP_215870.1; *Streptomyces coelicolor* RmdA, NP_625226.1; *Shewanella woodyi* Swoo_2750, ACA87026.1; *Rhodobacter sphaeroides* BphG1, YP_345384.1; *Legionella pneumophila* Lpl0329, CAH14560.1. REC, receiver; PAS, Per-Arnt-Sim; PAC, PAS-associated C-terminal motif; GAF, cGMP-regulated PDEs, adenylyl cyclases and FhIA.
(b) Multiple-sequence alignment of the catalytic residues of DcpA and homologs. The alignment was performed with the programs Multalin and Espript. The corresponding secondary structural elements of DcpA are displayed above the sequences and the catalytic residues are marked with red triangles. All sequences were downloaded from the NCBI database (www.ncbi. nlm.nih.gov) with the following accession numbers: *Mycobacterium smegmatis*, WP_003893571.1; *Mycobacterium tuberculosis*, NP_215870.1; *Rhodococcus fascians*, WP_052067107.1; *Thalassobacillus devorans*, WP_084315756.1; *Pseudorhodoferax* sp. Leaf274, WP_056901087.1; *Bacillus cereus*, AlE81798.1; *beta proteobacterium*, KPF59561.1; *Oxalobacteraceae bacterium*, WP_020703251.1; *Desulfovibrio magneticus*, WP_024825048.1. (c) HPLC spectra to show the dual activities of DcpA at different time courses. (d) HPLC spectra of mutants DcpA^{E259Q} and DcpA^{D506N&D507N} toward GTP and c-di-GMP, respectively.

N-terminal 6×His-tag. The construct was overexpressed in *E. coli* BL21 (DE3) strain (Novagen) at 16°C for 20 h after induction by 0.2 mM isopropyl β -D-1-thiogalactopyranoside at an $A_{600 \text{ nm}}$ of 0.8. Cells were harvested and resuspended in the lysis buffer (20 mM HEPES (pH 8.0), 100 mM NaCl and 5% glycerol), and disrupted by sonication. After centrifugation, the supernatant was loaded onto a Ni-NTA column (Qiagen) equilibrated with the binding buffer (20 mM HEPES (pH 8.0), 100 mM NaCl and 5% glycerol). The target protein was eluted with 400 mM imidazole and further purified by gel filtration (Superdex 200, GE Healthcare) in the binding buffer plus 2 mM MgCl₂. The fractions containing the target protein were collected and concentrated to 10 mg/ml for crystallization. For enzymatic activity assays, the protein samples were pooled at the highest peak fractions without concentration and stored at -80° C with 30% glycerol. The purity of protein was assessed by 12% (w/v) gel electrophoresis.

The selenomethionine (SeMet)-substituted DcpA was overexpressed in *E. coli* B834 (DE3) (Novagen). The transformed cells were grown at 37°C in SeMet medium (M9 medium supplemented with 50 mg/l SeMet and other essential amino acids at 50 mg/l) to an $A_{600 \text{ nm}}$ of 0.8. The following steps in protein expression and purification were the same as those for the native protein.

For the enzymatic activities, we constructed several truncated versions of DcpA, namely DcpA_{GAF} (Met1-Asp180), DcpA_{GGDEF} (Asp180-Ser339), DcpA_{EAL} (Asn350-Met615) and DcpA_{GAF-GGDEF} (Met1-Asp339). They were overexpressed and purified in the same manner as the full-length DcpA, except that DcpA_{EAL} was overexpressed in BL21 pKY206 (DE3) (Novagen). Mutants DcpA^{E259Q} and DcpA^{D506N&D507N} were obtained using the MutExpressTM Fast Mutagenesis Kit with the plasmid encoding the wild-type DcpA as the template. They were expressed, purified and stored as described above. The sequences of the primers for cloning the mutants of DcpA are listed in Supplementary Table S1.

Analytical gel filtration chromatography was performed to estimate the molecular mass of DcpA and mutants in the absence or presence of GDP in solution by a Superdex 200 Increase 10/300 column (GE Healthcare). Standard molecular markers used for calibration are as follows: ribonuclease A (13.7 kDa), ovalbumin (43.0 kDa), conalbumin (75.0 kDa), aldolase (158.0 kDa), ferritin (440.0 kDa) and thyroglobulin (669.0 kDa).

Crystallization, data collection and processing

Crystals of the SeMet-substituted proteins were grown at 16°C using the hanging-drop vapor-diffusion method, with a drop of 1 μ l protein solution mixed with an equal volume of the reservoir solution (8% polyethylene glycol 6000, 0.1 M MgCl₂ and 0.1 M MES (pH 6.0)). After optimization, crystals were transferred to cryoprotectant (reservoir solution supplemented with 30% glycerol) and flash-cooled with liquid nitrogen. The X-ray diffraction data were collected at 100 K in a liquid nitrogen stream using beamline BL19U1 with a PILATUS 6M detector at the Shanghai Synchrotron Radiation Facility (SSRF). All diffraction data were integrated and scaled with the program HKL2000 [28].

Structure determination and refinement

The structure of MorA (PDB ID: 4RNH) was used as the search model to determine the structure of the EAL domain by molecular replacement using the Molrep program [29] in the CCP4i program suite [30]. The model was further refined by the maximum likelihood method implemented in REFMAC5 [31] as part of the CCP4i program suite and rebuilt interactively with the program COOT [32]. Subsequent refinements were performed by REFMAC5 and COOT. The final model was evaluated with the web service MolProbity [33]. The crystallographic parameters are listed in Table 1. The interface area was calculated by PDBePISA [34]. All structure figures were prepared with PyMOL (https://pymol.org).

Surface plasmon resonance assays

All surface plasmon resonance (SPR) experiments were performed at 25°C using a Biacore 3000 instrument (GE Healthcare) in phosphate-buffered saline (PBS), pH 7.4. The GAF domain was covalently immobilized on the carboxymethyl-dextran surface of the CM5 chip. Putative metabolites or different concentrations of GDP in the running buffer were serially injected over the sensor chips with the immobilized protein and the blank for 2 min at a flow rate of 30 μ l/min. Afterwards, the sensor chip surface was regenerated after washing with PBS. The equilibrium responses were fitted with the linear fit and steady-state affinity mode using the BIA evaluation software. The graphs were plotted using the Origin 8.1 software.

Enzymatic activity assays

The standard samples were bought from Sigma–Aldrich and Biolog. The activities of DcpA and its mutants were assayed by high-performance liquid chromatography (HPLC) and performed in the buffer containing 100 mM HEPES (pH 7.5), 100 mM NaCl and 10 mM MgCl₂. A 20 μ l reaction mixture containing 0.5 mM GTP and 5 μ M full-length DcpA was incubated at 37°C for different time courses. Then, the reactions were terminated by heating at 95°C for 10 min. After centrifugation, 10 μ l of the supernatant was applied to the HPLC system (Agilent 1200 Series, U.S.A.). The buffer containing 20 mM KH₂PO₄/K₂HPO₄ (pH 7.0) and 5% methanol was used for pre-equilibration of the column (Zorbax 300 SB-C18 column, 4.6 × 250 mm, Agilent, U.S.A.) and separation of the components at a flow rate of 1 ml/min. The column temperature was kept at 25°C. The eluted fractions were detected and quantitated by the absorbance at 254 nm. The details of the enzymatic



	Se-DcpA _{EAL}
Data collection	
Space group	1422
Unit cell parameters	
a, b, c (Å)	129.1, 129.1, 165.5
α, β, γ (°)	90, 90, 90
Resolution range (Å)	101.8–3.0 (3.2–3.0) ¹
Unique reflections	14 428 (2235)
Completeness (%)	99.4 (97.4)
$\langle l/\sigma(l)\rangle$	11.1 (2.1)
R ² merge (%)	18.7 (89)
Average redundancy	8.3 (8.7)
Structure refinement	
Resolution range (Å)	101.8–3.0
No. of reflections	13688
R-factor ³ /R-free ⁴ (%)	25.6/31.0
No. of protein atoms	3370
No. of water atoms	29
Mean <i>B</i> factors (Ų)	69.6
RMSD ⁵	
Bond lengths (Å)	0.0047
Bond angles (°)	0.9585
Ramachandran plot ⁶ (% of residues)	
Most favored	96.8
Additional allowed	3.2
Outliers	0
PDB entry	5YRP

Table 1 Crystal parameters, data collection and structure refinement

¹The values in parentheses refer to statistics in the highest bin. ${}^{2}R_{merge} = \sum_{hkl} \sum_{i} |l_i(hkl) - \langle l(hkl) \rangle| / \sum_{hkl} \sum_{i} l_i(hkl)$, where $l_i(hkl)$ is the intensity of an observation and $\langle l(hkl) \rangle$ is the mean value for its unique reflection. Summations are over all reflections.

 ${}^{3}R$ -factor = $\sum_{h} |F_{o}(h) - F_{c}(h)| / \sum_{h} F_{o}(h)$, where F_{o} and F_{c} are the observed and

calculated structure factor amplitudes, respectively.

 ${}^{4}R$ -free was calculated with 5% of the data excluded from the refinement.

⁵Root-mean-square deviation from ideal values.

⁶Categories were defined by Molprobity.

reactions concerning $DcpA_{GGDEF}$, $DcpA_{GAF-GGDEF}$, $DcpA_{EAL}$, $DcpA^{E259Q}$ and $DcpA^{D506N\&D507N}$ are shown in corresponding figure legends. The detection and quantitation of each reaction were the same as for DcpA. All assays were performed in three independent experiments to calculate the means and standard deviations.

Circular dichroism spectrometry

DcpA was purified in a PBS (pH 7.4), using the protocol used to purify proteins for crystallization. DcpA in the absence of GDP was diluted in the same buffer to 0.6 mg/ml, and its secondary-structure compositions were analyzed using a Jasco J-810 circular dichroism (CD) spectropolarimeter. Spectra were recorded from 190 to 280 nm with an interval of 1 nm at room temperature. The final spectrum represents the average data of



three consecutive scans. The secondary structures of DcpA in the presence of GDP were analyzed in the same manner.

SAXS experiments

SAXS was used to investigate the overall conformation of DcpA in the absence or presence of GDP. SAXS experiments were performed at beamline BL19U2 of the National Center for Protein Science Shanghai (NCPSS) at SSRF [35]. Scattered X-ray intensities were collected using a Pilatus 1 M detector (DECTRIS Ltd). The measurements were carried out with 20×1 s exposures and scattering profiles for 20 frames were compared to evade the possible sample radiation damage. To reduce the radiation damage, a flow cell made of a cylindrical quartz capillary with a diameter of 1.5 mm and a wall of 10 µm was used during the measurements. The scattering was recorded in the range of the momentum transfer $0.01 < q < 0.45 \text{ Å}^{-1}$, where $q = (4\pi \sin\theta)/\lambda$, 2θ is the scattering angle and $\lambda = 1.033$ Å is the X-ray wavelength. To exclude concentration dependence and a possible influence of the interaction of the protein macromolecules in solution, 1.0, 2.5 and 5.0 mg/ml protein concentrations were prepared and measured. The 2D scattering images were converted into 1D SAXS curves through azimuthal averaging after solid-angle correction and then normalized with the intensity of the transmitted X-ray beam, using the software package BioXTAS RAW [36]. The scattering curves were primarily processed using the program PRIMUS [37] in the ATSAS software package [38]. After subtraction from the buffer (20 mM HEPES (pH 7.5), 100 mM NaCl, 5% glycerol and 2 mM MgCl₂, with or without GDP) scattering, the data were averaged and extrapolated to zero concentration. A linear Guinier plot was used to evaluate the radius of gyration (R_g) in the low q range ($q \times R_g < 1.3$). The maximal diameter of the scattering objects D_{max} and the pair distance distribution function P(r) were calculated by the program GNOM [39]. The *ab initio* reconstruction of each sample was performed using the program DAMMIN [40] with 20 runs for each experimental group, and DAMAVER [41] was used to analyze the normalized spatial discrepancy between 20 models. The filtered SAXS model by DAMFILT [41] was shown in PyMOL. Rigid-body modeling structures of the multi-domain protein DcpA and DcpA-GDP complex against the scattering curve were calculated with the program CORAL [38]. The theoretical scattering curves from the calculated rigid-body modeling structure of apo DcpA and liganded DcpA were calculated with the program CRYSOL [42] and compared with relative scattering profile with a χ^2 value to indicate the discrepancy. The crystal structures and the SAXS models were matched using SUPCOMB [43]. The SAXS parameters were listed in Supplementary Table S2.

Results

DcpA catalyzes the synthesis and hydrolysis of c-di-GMP in the presence of GTP

The 615-residue DcpA from *M. smegmatis* contains three distinct domains: an N-terminal GAF domain, followed by a GGDEF domain and a C-terminal EAL domain (Figure 1a). The boundaries of each domain were defined according to the secondary-structure-based multiple-sequence alignment, and the catalytic residues of the GGDEF/EAL domain were deduced from the homologs (Figure 1b). Notably, the pattern of domain organization and putative catalytic residues are highly conserved in most mycobacterial species [27], suggesting that these fusion proteins should all possess both synthetic and hydrolytic activities. Nevertheless, most previously identified fusion proteins with both GGDEF and EAL domains only exhibit either synthetic or hydrolytic activity due to the mutation at one active site [19]. To verify if DcpA has dual activities, we applied the purified DcpA to the activity assays. As shown in Figure 1c, both c-di-GMP and pGpG could be detected at a proper time point, indicating that DcpA can use GTP to synthesize c-di-GMP, which is, in turn, hydrolyzed into pGpG, in agreement with the previous report [27]. However, only pGpG is accumulated in the reaction mixture with time, and c-di-GMP is maintained at a residual level, indicating that the reaction velocity of hydrolysis is much higher than that of synthesis (Figure 1c).

The fourth residue (Glu) at the conserved GG(D/E)EF active site was proposed to co-ordinate the metal ion that drives the synthetic activity [20], whereas the conserved DDFGTGYSS motif of the EAL domain is essential for the hydrolytic activity beyond the EAL motif [44]. Indeed, a single mutation E259Q (termed DcpA^{E259Q}) at the GG(D/E)EF active site or double mutation D506N&D507N (termed DcpA^{D506N&D507N}) at the DDFGTGYSS motif of DcpA completely abolished the corresponding activity, further proving that these residues are catalytic residues (Figure 1d).



The GAF domain is essential for the activities of DcpA

In many cases, the fused regulatory domain could induce the dimerization of DGC, which is indispensable for the synthetic activity [45]. Analytical gel filtration indicated that the isolated GAF domain (termed DcpA_{GAF}) or GGDEF domain (termed DcpA_{GGDEF}) is monomeric in solution, whereas the full-length DcpA and the EAL domain-deleted version (termed DcpA_{GAF-GGDEF}) are dominantly dimeric (Figure 2a). In addition, the results of multiple-angle light-scattering (MALS) assays further supported that DcpA_{GAF} and DcpA_{GGDEF} are monomeric in solution, whereas DcpA_{GAF-GGDEF} is most probably a dimer (Supplementary Figure S1). The results suggested that the GAF domain is required for the dimerization of the GGDEF domain. Moreover, activity assays indicated that DcpA_{GAF-GGDEF}, but not DcpA_{GGDEF} possesses synthetic activity (Figure 2b). Therefore, it suggested that the GAF-mediated dimerization is indispensable for the synthetic activity of DcpA.

Furthermore, it showed that the hydrolytic activity of the isolated EAL domain (termed $DcpA_{EAL}$) is much lower than that of the full-length DcpA (Figure 2c), suggesting that fusion domains contribute to the higher hydrolytic activity of DcpA. A previous study reported that the tandem GGDEF and EAL domains of DcpA, after deleting the GAF domain, showed no hydrolytic activity *in vitro* [27]. This indicated that the GAF domain could augment the hydrolytic activity of the EAL domain. Notably, the role of the GAF domain in the dual activities of DcpA is similar to that of the previously reported regulatory domain of *Klebsiella pneumoniae* BlrP1 [2].

GDP is the ligand of the GAF domain

The GAF domains, which are distributed widely from Archaea to mammals, represent one of the largest families of small-molecule-binding regulatory domains [46]. Though the GAF domains from different species display a similar three-dimensional structure, the binding ligands differ a lot, making it difficult to identify the ligand of the GAF domain only based on sequence alignment and/or structural comparison. It is known that most of the reported ligands are closely related to the corresponding signaling pathways regulated by the GAF domains. The up-regulation of the *dcpA* gene upon starvation or depletion of carbon source [47] strongly indicated that the ligand of the GAF domain is most probably involved in energy metabolism. Following these hints and taking the previously identified ligands into account, 26 putative metabolites were applied to the binding affinity screening using SPR. As displayed in Figure 3a, $DcpA_{GAF}$ only binds to GDP, but neither the previously reported ligands valine, isoleucine and pyruvate, nor the substrate GTP and analog GMP or cAMP. Furthermore, we determined the binding affinity of $DcpA_{GAF}$ toward GDP and revealed an equilibrium dissociation constant (K_d) of 2.25 mM (Figure 3b), which is quite similar to that with the full-length DcpA (K_d of 2.21 mM) (Figure 3c). However, the individual GGDEF or EAL domain does not interact with GDP.

GDP regulates the activities of DcpA

To investigate the putative effect of GDP on the synthetic and hydrolytic activities, we compared the corresponding activities of DcpA in the presence and absence of GDP. The double-mutant $DcpA^{D506N\&D507N}$ that





(a) Analytical gel filtration profiles of the full-length and truncated DcpA. The molecular mass of each sample was calculated according to the standard curve. (b) HPLC spectra of the synthetic activity of DcpA_{GGDEF} and DcpA_{GAF-GGDEF}. DcpA_{GGDEF} (20 μ M) was incubated with 0.2 mM GTP at 37°C for 12 h, whereas the reaction time for DcpA_{GAF-GGDEF} was 1 h. (c) HPLC spectra of the hydrolytic activity of DcpA_{EAL} and DcpA. The reaction lasted for 2 h at 37°C in the presence of 10 μ M enzyme and 0.2 mM c-di-GMP. The concentrations of substrates were chosen to ensure the reaction is in the linear interval.





Figure 3. GDP is the ligand of the GAF domain and could regulate the activities of DcpA.

(a) SPR sensor diagram for the putative ligands against the immobilized DcpA_{GAF}. The concentration of each small molecule is 1 mM. The steady-state equilibrium curves of (b) DcpA_{GAF} and (c) the full-length DcpA against GDP. The binding affinity of DcpA_{GAF} or full-length DcpA toward GDP was calculated from the data at the concentration varying from 0.2 to 3.2 mM using the linear fit and steady-state affinity mode. The relative (d) synthetic and (e) hydrolytic activity of DcpA in the presence of various concentrations of GDP. For the synthetic activity, 5 μ M DcpA^{D506N&D507N} was incubated with 0.2 mM GTP for 45 min at 37°C. For the hydrolytic activity, 5 μ M DcpA^{E259Q} was incubated with 0.5 mM c-di-GMP for 30 min at 37°C. (f) The relative synthetic activity of DcpA at different GTP/GDP ratios. Various concentrations of GDP were applied to the activity assays, with the concentration of GTP fixed at 0.5 mM. The reactions were started by adding 5 μ M enzyme and terminated after 45 min. The experimental conditions for two mutants were chosen to make the reaction locate in the linear interval. The assays were performed by HPLC, and the relative activity was defined based on the peak areas of product. The error bars represent SD from three independent assays with three different protein preparations. Two-tailed Student's *t*-test is used for the comparison of statistical significance. The *P*-value of <0.001 is indicated with ***.

only possesses the synthetic activity and the single-mutant $DcpA^{E259Q}$ that just retains the hydrolytic activity (Figure 1d) were used as the positive controls for the comparisons. Upon the addition of GDP to 0.05 and 0.2 mM, the synthetic activity of $DcpA^{D506N\&D507N}$ decreased to two-third and one-quarter, respectively (Figure 3d). Once the GDP concentration reached 1 mM, no synthetic activity could be detected (Figure 3d). In contrast, the addition of GDP sharply increased the hydrolytic activity of $DcpA^{E259Q}$ to 4–30-fold, in a GDP concentration-dependent manner (Figure 3e). These results revealed that binding of GDP to the GAF domain could down-regulate the synthetic activity, but up-regulate the hydrolytic activity of DcpA, which in consequence might lead to the consumption of intracellular c-di-GMP in *M. smegmatis* at a much faster rate in response to some given environmental stimuli.

It is known that the intracellular GTP/GDP ratio is strongly correlated to the growth phase and supply of environmental nutrients, and thus involved in bacterial growth and survival at starvation [48]. Previous reports showed that during the exponential growth phase, the intracellular concentration of c-di-GMP is maintained at a basal level [49], which is correlated to the high GTP/GDP ratio at this phase. In contrast, at the stationary phase or upon starvation, the intracellular c-di-GMP is accumulated and reaches a relatively high level [27,49], which in turn slows down the growth rate and initiates stress responses, such as biofilm formation, to help bacterial survival and adaption [27,50]. However, the accumulated c-di-GMP should be timely degraded to keep



its intracellular homeostasis. On one hand, although the bifunctional enzyme DcpA consumes c-di-GMP at a higher velocity compared with its production in the normal condition, a decreasing intracellular GTP/GDP ratio will trigger an accelerated degradation of c-di-GMP (Figures 1c and 3e). On the other hand, the synthetic activity of DcpA is down-regulated along with the decrease in the intracellular GTP/GDP ratio (Figure 3f), leading to the very limited production of c-di-GMP. These dual mechanisms will guarantee the timely down-regulation of intracellular c-di-GMP to the basal level. It is also consistent with the previous finding of increased promoter activity of *dcpA* upon starvation [27].

Overall structure of the EAL domain

To obtain structural insights, we applied the full-length DcpA to crystallization. However, the refined structure from the optimized crystal only contains the EAL domain (DcpA_{EAL}, Ala364-Pro606), which is probably due to the degradation of the other two domains during crystallization. Each asymmetric unit contains two molecules of the EAL domain, which form a homodimer with a total buried interface area of 895 Å². The interface is stabilized by helices $\alpha 4$ and $\alpha 6$ in addition to the loop $\beta 5-\alpha 4$ of each subunit (Figure 4a), sharing a dimerization pattern similar to other EAL-containing proteins, such as *Thiobacillus denitrificans* TBD1265, *K. pneumoniae* BlrP1 and *Pseudomonas aeruginosa* MorA [51–53].



Figure 4. Structural analyses of DcpA_{EAL}.

(a) The overall structure of DcpA_{EAL}. The structure is shown in cartoon representation. The two subunits are colored in cyan and red, respectively. The secondary structural elements (helices $\alpha 4$, $\alpha 6$ and strand $\beta 5$) at the dimeric interface are labeled. The dash lines indicate the residues missing in the electron density map. The Mg²⁺ ions are shown as spheres in the same color with the corresponding subunit of DcpA_{EAL}, whereas the water molecule (Wat) is displayed as gray spheres. (b) Structural superposition of DcpA_{EAL} (cyan) on that of *T. denitrificans* TBD1265 (wheat; PDB ID: 2R6O). The water molecule of TBD1265 is shown as light blue spheres. The inset shows a close-up view of the active site. The metal-coordinated and catalytic residues of two structures are shown as sticks; however, only those of DcpA_{EAL} are labeled. (c) Structural superposition of DcpA_{EAL} (cyan) on the c-di-GMP-complexed EAL domain of TBD1265 (pink; PDB ID: 3N3T). c-di-GMP is shown as green sticks. (d) A schematic comparison of the DDFGTGYSS motifs. The helix $\alpha 5$ and loop $\beta 5-\alpha 5$ of monomeric MorA (PDB ID: 4RNJ)/ dimeric MorA (PDB ID: 4RNI), and helix $\alpha 4$ and loop $\beta 5-\alpha 4$ of dimeric DcpA_{EAL} in different lengths are shown.



Each subunit displays a classic TIM barrel structure composed of eight β -strands surrounded by eight α -helices, leaving 19 residues (Glu390-Ser408) of one subunit and 24 residues (Arg388-Leu411) of the other subunit missing in the electron density map (Figure 4a). Compared with the known structures of EAL domains from other species, they appear to have a very similar tertiary structure (Supplementary Figure S2). The DALI search [54] revealed that DcpA_{EAL} shares a highest structural homology with that of *T. denitrificans* TBD1265 (PDB ID: 2R6O, *Z*-score 30.5, sequence identity 38%), with a root-mean-square deviation of 2.86 Å over 395 C α atoms. Superposition of the two structures revealed that DcpA_{EAL} should also possess a two metal-coordinated active site, though only one Mg²⁺ ion could be defined in our structure (Figure 4b). This Mg²⁺ ion is positioned in a place similar to the primary one of TBD1265 and co-ordinated by the side chains of Glu384, Asn444, Glu476 and Asp506 (Figure 4b). Despite the co-ordinated residues Asp507 and Asp529 at the secondary metal-binding site, or the so-called low-affinity site, being conserved, the putative Mg²⁺ ion is missing in our structure (Figure 4b). Notably, Asp506 and Asp507 are part of the conserved DDFGTGYSS motif, which constitutes the loop β 5- α 4 at the dimeric interface. Besides, the conserved Glu563 is thought to be a general base catalyst that facilitates the water-mediated nucleophilic attack toward c-di-GMP (Figure 4b).

To date, 22 c-di-GMP-complexed structures of the EAL domain from nine species have been reported. Superimposing our structure with the c-di-GMP-complexed EAL domain of *T. denitrificans* TBD1265 (PDB ID: 3N3T) showed that c-di-GMP is supposed to bind to the C-terminal segment of the TIM barrel in DcpA_{EAL} (Figure 4c), a similar site of known TIM barrel enzymes [55]. A previous report showed that site-directed mutagenesis preventing dimerization did not alter the substrate-binding ability of the EAL domain, but resulted in drastically reduced hydrolytic activity, indicating that dimerization leads to a shorter helix α 5 of only six residues (corresponding to helix α 4 in our structure) accompanying with an elongated loop β 5- α 5 of nine residues (corresponding to the loop β 5- α 4 in our structure) at the dimeric interface, eventually allowing the two aspartate residues to enter the active site for co-ordinating the catalytic metal ions [52]. A structure-based comparison showed that DcpA_{EAL} is inclined to be an activated dimer (Figure 4d).

Overall conformational change of DcpA induced by GDP using SAXS

To investigate how the GAF domain regulates the dual activities upon GDP binding, we applied SAXS to compare the overall three-dimensional structures of DcpA in the presence and absence of GDP. The Guinier regions for both apo and liganded DcpA experimental groups are linear (Figure 5a), indicating that the measured system is monodispersed and homogeneous in solution. In addition, the experimental molecular mass obtained in SAXS indicated that DcpA adopts a dimeric structure either in the presence or absence of GDP, which is consistent with the result of analytical gel filtration (Figure 5b), indicating that there is no contribution of aggregates to the scatter. Based on the SAXS data, the calculated maximum particle dimension (D_{max}) for apo DcpA is 169.5 Å, whereas that of the liganded is 200.5 Å (Figure 5c), revealing that DcpA becomes somewhat extended in the presence of GDP. Notably, the CD spectra showed that apo and liganded DcpA exhibit a similar secondary-structure composition (Figure 5d). These results suggested that GDP binding to the GAF domain might induce the rearrangements of the overall conformation of dimeric DcpA.

Discussion

The GAF domains are widespread in all kingdoms of life through evolution for over 2 billion years [57]. They usually coexist with other catalytic domains, which are involved in a variety of cellular signaling pathways. It is well known that the GAF domains function through binding to small ligands, even though only a few of them have been identified. The hydrolytic activity of mammalian PDE5 is significantly increased upon binding of cGMP to the GAF domain, which might trigger the rearrangements of the overall conformation of the full-length protein [58]. In addition, the winged helix-turn-helix domain of *Bacillus subtilis* CodY splays when the N-terminal GAF domain binds to isoleucine, which in turn increases the DNA-binding affinity of CodY [59].

Bioinformatic analyses revealed that *M. smegmatis* encoded two enzymes MSMEG_2196 (DcpA) and MSMEG_2774 (MSDGC-2), which probably play a role in c-di-GMP synthesis/hydrolysis. We showed here DcpA that contains the tandem GGDEF and EAL domains indeed possesses both synthetic and hydrolytic activities (Figure 1c). However, it was reported that MSDGC-2 that contains only a putative GGDEF domain loses the synthetic activity due to the mutations of the catalytic residues [27]. Thus, DcpA becomes the only enzyme responsible for the synthesis and hydrolysis of c-di-GMP in *M. smegmatis*. Our results showed that the GAF domain is indispensable for the synthetic activity and could increase the hydrolytic activity of DcpA





Figure 5. Conformational changes between apo and liganded DcpA.

(a) SAXS intensity profiles of apo (red line) and liganded DcpA (gray line). The inset shows the Guinier plots of apo and liganded DcpA. (b) Analytical gel filtration profiles of DcpA in the absence and presence of GDP. The molecular mass of each sample was calculated according to the standard curve. (c) Electron pair distribution function P(r) profiles of the apo (red line) and liganded DcpA (gray line). The D_{max} of apo DcpA is 169.5 Å, and significantly increased to 200.5 Å upon the addition of GDP, showing a structural extension. (d) CD spectra of DcpA in the absence and presence of GDP.

(Figure 2b,c). Moreover, screening of putative metabolites revealed that GDP is the ligand of the GAF domain (Figure 3a). The binding of GDP to the GAF domain down-regulates synthetic activity, but up-regulates hydro-lytic activity (Figure 3d,e), thus in sum contributing to the accelerated assimilation of c-di-GMP upon



Figure 6. Superposition of the crystal structures against the SAXS models.

The averaged and filtered *ab initio* SAXS envelopes of DcpA in the (**a**) absence and (**b**) presence of GDP were superimposed with the rigid-body modeling structures composed of GAF and GGDEF–EAL domains. The dimeric profile of DcpA is shown in green and yellow for two subunits, respectively.



environmental stresses. Then, using SAXS, we found that GDP binding to the GAF domain might induce the overall conformational change of dimeric DcpA (Figure 5).

Furthermore, we fitted the structures of the GGDEF-EAL and GAF domains into the *ab initio* molecular envelopes of apo and liganded DcpA, respectively. The envelopes were reconstructed from SAXS data by dummy bead modeling, and the one having the lowest normalized spatial discrepancy was picked as the most representative model. The structure of the fused GGDEF and EAL domains was gained by homology modeling based on the EAL domain of DcpA itself, combined with *P. aeruginosa* MorA (PDB ID: 4RNH), whereas the structure of the GAF domain was deduced from a cyanobacterial protein (PDB ID: 3W2Z). The positions of the input domains of apo and liganded DcpA were searched and optimized for 10 rounds based on the SAXS envelopes, respectively. In the model of apo DcpA (χ^2 value 3.037), the GAF domain of one subunit interacts with the GGDEF domain of the other subunit, making DcpA a compact dimer (Figure 6a), in accordance with the results of analytical gel filtration (Figure 2a). However, in the presence of GDP (χ^2 value 2.536), DcpA forms an extended dimer with an interface between the two EAL domains, leaving the two GAF domains kicked outwards (Figure 6b).

All together, these findings not only help us better understand the fine control of synthesis and hydrolysis of c-di-GMP in *M. smegmatis*, but also provide hints for exploration of the homolog Rv1354c in *M. tuberculosis*. However, more structural information and biochemical investigations are needed to clearly elucidate the fine catalytic mechanism of these dual-functional enzymes.

Abbreviations

c-di-GMP, bis-(3'-5')-cyclic dimeric guanosine monophosphate; DcpA, diguanylate cyclase and phosphodiesterase <u>A</u>; DGC, diguanylate cyclase; EAL, Glu-Ala-Leu; GAF, mammalian cGMP-regulated PDEs, *Anabaena* adenylyl cyclases and *Escherichia coli* transcription activator <u>FhIA</u>; GGDEF, Gly-Gly-Asp-Glu-Phe; HPLC, high-performance liquid chromatography; *K*_d, dissociation constant; PAS, Per-Arnt-Sim; PBS, phosphate-buffered saline; PDE, phosphodiesterase; pGpG, 5'-phosphoguanylyl-(3'-5')-guanosine; REC, receiver; SAXS, small-angle X-ray scattering; SeMet, selenomethionine; SPR, surface plasmon resonance; SSRF, Shanghai Synchrotron Radiation Facility; TIM, triosephosphate isomerase.

Author Contribution

H.-J.C., Q.L., N.L., Y.C. and C.-Z.Z. designed the study, analyzed the data and wrote the paper. H.-J.C. performed the enzymatic experiments. H.-J.C., Y.L. and Y.-L.J. performed crystal screening and optimization, data collection and structural determination. H.-J.C. and Q.L. did the SAXS experiments and N.L. processed the SAXS data.

Funding

This work is supported by the National Natural Science Foundation of China (grant no. U1332114), the Ministry of Science and Technology of China (grant no. 2016YFA0400900) and the Youth Innovation Promotion Association of Chinese Academy of Science (grant no. 2017319).

Acknowledgements

We appreciate the assistance of the staff at the Shanghai Synchrotron Radiation Facility (SSRF) and the Core Facility Center for Life Sciences at the University of Science and Technology of China. We are grateful to all of the developers of the CCP4 suite, PyMOL and ESPript.

Competing Interests

The Authors declare that there are no competing interests associated with the manuscript.

References

- 1 Ross, P., Weinhouse, H., Aloni, Y., Michaeli, D., Weinberger-Ohana, P., Mayer, R. et al. (1987) Regulation of cellulose synthesis in *Acetobacter xylinum* by cyclic diguanylic acid. *Nature* **325**, 279–281 https://doi.org/10.1038/325279a0
- 2 Schirmer, T. and Jenal, U. (2009) Structural and mechanistic determinants of c-di-GMP signalling. *Nat. Rev. Microbiol.* **7**, 724–735 https://doi.org/10. 1038/nrmicro2203
- 3 Römling, U. and Amikam, D. (2006) Cyclic di-GMP as a second messenger. Curr. Opin. Microbiol. 9, 218–228 https://doi.org/10.1016/j.mib.2006.02. 010



- 4 Wolfe, A.J. and Visick, K.L. (2008) Get the message out: cyclic-di-GMP regulates multiple levels of flagellum-based motility. J. Bacteriol. **190**, 463–475 https://doi.org/10.1128/JB.01418-07
- 5 Tamayo, R., Pratt, J.T. and Camilli, A. (2007) Roles of cyclic diguanylate in the regulation of bacterial pathogenesis. *Annu. Rev. Microbiol.* **61**, 131–148 https://doi.org/10.1146/annurev.micro.61.080706.093426
- 6 Tischler, A.D. and Camilli, A. (2005) Cyclic diguanylate regulates Vibrio cholerae virulence gene expression. Infect. Immun. 73, 5873–5882 https://doi. org/10.1128/IAI.73.9.5873-5882.2005
- 7 McWhirter, S.M., Barbalat, R., Monroe, K.M., Fontana, M.F., Hyodo, M., Joncker, N.T. et al. (2009) A host type I interferon response is induced by cytosolic sensing of the bacterial second messenger cyclic-di-GMP. J. Exp. Med. 206, 1899–1911 https://doi.org/10.1084/jem.20082874
- 8 Lori, C., Ozaki, S., Steiner, S., Böhm, R., Abel, S., Dubey, B.N. et al. (2015) Cyclic di-GMP acts as a cell cycle oscillator to drive chromosome replication. *Nature* **523**, 236–239 https://doi.org/10.1038/nature14473
- 9 Srivastava, D. and Waters, C.M. (2012) A tangled web: regulatory connections between quorum sensing and cyclic di-GMP. J. Bacteriol. 194, 4485–4493 https://doi.org/10.1128/JB.00379-12
- 10 An, S.Q., Chin, K.H., Febrer, M., McCarthy, Y., Yang, J.G., Liu, C.L. et al. (2013) A cyclic GMP-dependent signalling pathway regulates bacterial phytopathogenesis. *EMBO J.* 32, 2430–2438 https://doi.org/10.1038/emboj.2013.165
- 11 Almblad, H., Harrison, J.J., Rybtke, M., Groizeleau, J., Givskov, M., Parsek, M.R. et al. (2015) The cyclic AMP-Vfr signaling pathway in *Pseudomonas* aeruginosa is inhibited by cyclic di-GMP. J. Bacteriol. **197**, 2190–2200 https://doi.org/10.1128/JB.00193-15
- 12 Gupta, K.R., Kasetty, S. and Chatterji, D. (2015) Novel functions of (p)ppGpp and cyclic di-GMP in mycobacterial physiology revealed by phenotype microarray analysis of wild-type and isogenic strains of *Mycobacterium smegmatis. Appl. Environ. Microbiol.* **81**, 2571–2578 https://doi.org/10.1128/ AEM.03999-14
- 13 Tal, R., Wong, H.C., Calhoon, R., Gelfand, D., Fear, A.L., Volman, G. et al. (1998) Three *cdg* operons control cellular turnover of cyclic di-GMP in *Acetobacter xylinum*: genetic organization and occurrence of conserved domains in isoenzymes. *J. Bacteriol.* **180**, 4416–4425 PMID:9721278
- 14 Ausmees, N., Mayer, R., Weinhouse, H., Volman, G., Amikam, D., Benziman, M. et al. (2001) Genetic data indicate that proteins containing the GGDEF domain possess diguarylate cyclase activity. *FEMS Microbiol. Lett.* 204, 163–167 https://doi.org/10.1111/j.1574-6968.2001.tb10880.x
- 15 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**(Pt 4), 980–994 https://doi.org/10.1099/mic.0.2006/002824-0
- 16 Schmidt, A.J., Ryjenkov, D.A. and Gomelsky, M. (2005) The ubiquitous protein domain EAL is a cyclic diguanylate-specific phosphodiesterase: enzymatically active and inactive EAL domains. J. Bacteriol. **187**, 4774–4781 https://doi.org/10.1128/JB.187.14.4774-4781.2005
- 17 Cohen, D., Mechold, U., Nevenzal, H., Yarmiyhu, Y., Randall, T.E., Bay, D.C. et al. (2015) Oligoribonuclease is a central feature of cyclic diguanylate signaling in *Pseudomonas aeruginosa. Proc. Natl Acad. Sci. U.S.A.* **112**, 11359–11364 https://doi.org/10.1073/pnas.1421450112
- 18 Ryan, R.P., Fouhy, Y., Lucey, J.F., Crossman, L.C., Spiro, S., He, Y.W. et al. (2006) Cell-cell signaling in *Xanthomonas campestris* involves an HD-GYP domain protein that functions in cyclic di-GMP turnover. *Proc. Natl Acad. Sci. U.S.A.* **103**, 6712–6717 https://doi.org/10.1073/pnas.0600345103
- Hengge, R. (2009) Principles of c-di-GMP signalling in bacteria. *Nat. Rev. Microbiol.* 7, 263–273 https://doi.org/10.1038/nrmicro2109
 Chan, C., Paul, R., Samoray, D., Amiot, N.C., Giese, B., Jenal, U. et al. (2004) Structural basis of activity and allosteric control of diguanylate cyclase.
- 20 Gran, C., Pati, K., Samoray, D., Anno, N.C., Giese, D., Jena, O. et al. (2004) Structural basis of activity and anosteric control of diguariyate cyclase. Proc. Natl Acad. Sci. U.S.A. 101, 17084–17089 https://doi.org/10.1073/pnas.0406134101
- 21 Hasegawa, K., Masuda, S. and Ono, T.A. (2006) Light induced structural changes of a full-length protein and its BLUF domain in YcgF (BIrp), a blue-light sensing protein that uses FAD (BLUF). *Biochemistry* **45**, 3785–3793 https://doi.org/10.1021/bi051820x
- 22 Levet-Paulo, M., Lazzaroni, J.C., Gilbert, C., Atlan, D., Doublet, P. and Vianney, A. (2011) The atypical two-component sensor kinase LpI0330 from Legionella pneumophila controls the bifunctional diguanylate cyclase-phosphodiesterase LpI0329 to modulate bis-(3'-5')-cyclic dimeric GMP synthesis. J. Biol. Chem. 286, 31136–31144 https://doi.org/10.1074/jbc.M111.231340
- 23 Hull, T.D., Ryu, M.H., Sullivan, M.J., Johnson, R.C., Klena, N.T., Geiger, R.M. et al. (2012) Cyclic di-GMP phosphodiesterases RmdA and RmdB are involved in regulating colony morphology and development in *Streptomyces coelicolor. J. Bacteriol.* **194**, 4642–4651 https://doi.org/10.1128/JB.00157-12
- 24 Aravind, L. and Ponting, C.P. (1997) The GAF domain: an evolutionary link between diverse phototransducing proteins. *Trends Biochem. Sci.* 22, 458–459 https://doi.org/10.1016/S0968-0004(97)01148-1
- 25 Cui, T., Zhang, L., Wang, X. and He, Z.G. (2009) Uncovering new signaling proteins and potential drug targets through the interactome analysis of *Mycobacterium tuberculosis. BMC Genomics* **10**, 118 https://doi.org/10.1186/1471-2164-10-118
- 26 Snapper, S.B., Melton, R.E., Mustafa, S., Kieser, T. and Jacobs, Jr, W.R. (1990) Isolation and characterization of efficient plasmid transformation mutants of *Mycobacterium smegmatis*. *Mol. Microbiol.* **4**, 1911–1919 https://doi.org/10.1111/j.1365-2958.1990.tb02040.x
- 27 Bharati, B.K., Sharma, I.M., Kasetty, S., Kumar, M., Mukherjee, R. and Chatterji, D. (2012) A full-length bifunctional protein involved in c-di-GMP turnover is required for long-term survival under nutrient starvation in *Mycobacterium smegmatis*. *Microbiology* **158**(Pt 6), 1415–1427 https://doi.org/10.1099/mic.0.053892-0
- 28 Otwinowski, Z. and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* 276, 307–326 https://doi.org/ 10.1016/S0076-6879(97)76066-X
- 29 Vagin, A. and Teplyakov, A. (1997) MOLREP: an automated program for molecular replacement. J. Appl. Crystallogr. **30**, 1022–1025 https://doi.org/10. 1107/S0021889897006766
- 30 Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R. et al. (2011) Overview of the CCP4 suite and current developments. Acta Crystallogr. D Biol. Crystallogr. 67(Pt 4), 235–242 https://doi.org/10.1107/S0907444910045749
- 31 Murshudov, G.N., Skubák, P., Lebedev, A.A., Pannu, N.S., Steiner, R.A., Nicholls, R.A. et al. (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 355–367 https://doi.org/10.1107/S0907444911001314
- 32 Emsley, P. and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. 60, 2126–2132 https://doi.org/10.1107/S0907444904019158
- 33 Chen, V.B., Arendall, W.B., Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J. et al. (2010) Molprobity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr. D Biol. Crystallogr. 66, 12–21 https://doi.org/10.1107/S0907444909042073
- 34 Krissinel, E. and Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state. J. Mol. Biol. 372, 774–797 https://doi.org/10.1016/ j.jmb.2007.05.022



- 35 Li, N., Li, X., Wang, Y., Liu, G., Zhou, P., Wu, H. et al. (2016) The new NCPSS BL19U2 beamline at the SSRF for small-angle X-ray scattering from biological macromolecules in solution. *J. Appl. Crystallogr.* **49**, 1428–1432 https://doi.org/10.1107/S160057671601195X
- 36 Nielsen, S., Toft, K.N., Snakenborg, D., Jeppesen, M.G., Jacobsen, J., Vestergaard, B. et al. (2009) BioXTAS RAW, a software program for high-throughput automated small-angle X-ray scattering data reduction and preliminary analysis. J. Appl. Crystallogr. 42, 959–964 https://doi.org/10. 1107/S0021889809023863
- 37 Konarev, P.V., Volkov, V.V., Sokolova, A.V., Koch, M.H. and Svergun, D.I. (2003) PRIMUS: a windows PC-based system for small-angle scattering data analysis. J. Appl. Crystallogr. 36, 1277–1282 https://doi.org/10.1107/S0021889803012779
- 38 Petoukhov, M.V., Franke, D., Shkumatov, A.V., Tria, G., Kikhney, A.G., Gajda, M. et al. (2012) New developments in the ATSAS program package for small-angle scattering data analysis. J. Appl. Crystallogr. 45, 342–350 https://doi.org/10.1107/S0021889812007662
- 39 Svergun, D. (1992) Determination of the regularization parameter in indirect-transform methods using perceptual criteria. J. Appl. Crystallogr. 25, 495–503 https://doi.org/10.1107/S0021889892001663
- 40 Fischer, H., Oliveira Neto, M., Napolitano, H., Polikarpov, I. and Craievich, A. (2010) Determination of the molecular weight of proteins in solution from a single small-angle X-ray scattering measurement on a relative scale. J. Appl. Crystallogr. 43, 101–109 https://doi.org/10.1107/S0021889809043076
- 41 Volkov, V.V. and Svergun, D.I. (2003) Uniqueness of *ab initio* shape determination in small-angle scattering. J. Appl. Crystallogr. 36, 860–864 https://doi.org/10.1107/S0021889803000268
- 42 Svergun, D., Barberato, C. and Koch, M.H.J. (1995) CRYSOL—A program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. J. Appl. Crystallogr. 28, 768–773 https://doi.org/10.1107/S0021889895007047
- 43 Kozin, M.B. and Svergun, D.I. (2001) Automated matching of high- and low-resolution structural models. J. Appl. Crystallogr. 34, 33–41 https://doi.org/ 10.1107/S0021889800014126
- 44 Rao, F., Yang, Y., Qi, Y. and Liang, Z.X. (2008) Catalytic mechanism of cyclic di-GMP-specific phosphodiesterase: a study of the EAL domain-containing RocR from *Pseudomonas aeruginosa. J. Bacteriol.* **190**, 3622–3631 https://doi.org/10.1128/JB.00165-08
- 45 Schirmer, T. (2016) C-di-GMP synthesis: structural aspects of evolution, catalysis and regulation. J. Mol. Biol. 428, 3683–3701 https://doi.org/10. 1016/j.jmb.2016.07.023
- 46 Zoraghi, R., Corbin, J.D. and Francis, S.H. (2004) Properties and functions of GAF domains in cyclic nucleotide phosphodiesterases and other proteins. *Mol. Pharmacol.* **65**, 267–278 https://doi.org/10.1124/mol.65.2.267
- 47 Bharati, B.K., Swetha, R. and Chatterji, D. (2013) Identification and characterization of starvation induced *msdgc-1* promoter involved in the c-di-GMP turnover. *Gene* **528**, 99–108 https://doi.org/10.1016/j.gene.2013.07.043
- 48 Bittner, A.N., Kriel, A. and Wang, J.D. (2014) Lowering GTP level increases survival of amino acid starvation but slows growth rate for *Bacillus subtilis* cells lacking (p)ppGpp. *J. Bacteriol.* **196**, 2067–2076 https://doi.org/10.1128/JB.01471-14
- 49 Spangler, C., Böhm, A., Jenal, U., Seifert, R. and Kaever, V. (2010) A liquid chromatography-coupled tandem mass spectrometry method for quantitation of cyclic di-guanosine monophosphate. *J. Microbiol. Methods* **81**, 226–231 https://doi.org/10.1016/j.mimet.2010.03.020
- 50 Armitage, J.P. and Berry, R.M. (2010) Time for bacteria to slow down. *Cell* **141**, 24–26 https://doi.org/10.1016/j.cell.2010.03.023
- 51 Tchigvintsev, A., Xu, X., Singer, A., Chang, C., Brown, G., Proudfoot, M. et al. (2010) Structural insight into the mechanism of c-di-GMP hydrolysis by EAL domain phosphodiesterases. *J. Mol. Biol.* **402**, 524–538 https://doi.org/10.1016/j.jmb.2010.07.050
- 52 Phippen, C.W., Mikolajek, H., Schlaefli, H.G., Keevil, C.W., Webb, J.S. and Tews, I. (2014) Formation and dimerization of the phosphodiesterase active site of the *Pseudomonas aeruginosa* MorA, a bi-functional c-di-GMP regulator. *FEBS Lett.* **588**, 4631–4636 https://doi.org/10.1016/j.febslet.2014.11. 002
- 53 Barends, T.R., Hartmann, E., Griese, J.J., Beitlich, T., Kirienko, N.V., Ryjenkov, D.A. et al. (2009) Structure and mechanism of a bacterial light-regulated cyclic nucleotide phosphodiesterase. *Nature* **459**, 1015–1018 https://doi.org/10.1038/nature07966
- 54 Holm, L. and Rosenström, P. (2010) Dali server: conservation mapping in 3D. Nucleic Acids Res. 38(Suppl 2), W545–W549 https://doi.org/10.1093/ nar/gkq366
- 55 Nagano, N., Orengo, C.A. and Thornton, J.M. (2002) One fold with many functions: the evolutionary relationships between TIM barrel families based on their sequences, structures and functions. *J. Mol. Biol.* **321**, 741–765 https://doi.org/10.1016/S0022-2836(02)00649-6
- 56 Sundriyal, A., Massa, C., Samoray, D., Zehender, F., Sharpe, T., Jenal, U. et al. (2014) Inherent regulation of EAL domain-catalyzed hydrolysis of second messenger cyclic di-GMP. J. Biol. Chem. 289, 6978–6990 https://doi.org/10.1074/jbc.M113.516195
- 57 Martinez, S.E., Beavo, J.A. and Hol, W.G. (2002) GAF domains: two-billion-year-old molecular switches that bind cyclic nucleotides. *Mol. Interv.* 2, 317 https://doi.org/10.1124/mi.2.5.317
- 58 Heikaus, C.C., Stout, J.R., Sekharan, M.R., Eakin, C.M., Rajagopal, P., Brzovic, P.S. et al. (2008) Solution structure of the cGMP binding GAF domain from phosphodiesterase 5 insight into nucleotide specificity, dimerization, and cGMP-dependent conformational change. J. Biol. Chem. 283, 22749–22759 https://doi.org/10.1074/jbc.M801577200
- 59 Levdikov, V.M., Blagova, E., Young, V.L., Belitsky, B.R., Lebedev, A., Sonenshein, A.L. et al. (2017) Structure of the branched-chain amino acid and GTP-sensing global regulator, CodY, from *Bacillus subtilis. J. Biol. Chem.* **292**, 2714–2728 https://doi.org/10.1074/jbc.M116.754309