PLOS ONE

RESEARCH ARTICLE

The importance of conserved amino acids in heme-based globin-coupled diguanylate cyclases

Xuehua Wan[1,2]*, Jennifer A. Saito[1,2], James S. Newhouse[2], Shaobin Hou[1,2], Maqsudul Alam[1,2†]

1 Department of Microbiology, University of Hawaii, Honolulu, Hawaii, United States of America, **2** Advanced Studies in Genomics, Proteomics and Bioinformatics, University of Hawaii, Honolulu, Hawaii, United States of America

† Deceased.
* xuehua@hawaii.edu

Citation: Wan X, Saito JA, Newhouse JS, Hou S, Alam M (2017) The importance of conserved amino acids in heme-based globin-coupled diguanylate cyclases. PLoS ONE 12(8): e0182782. https://doi.org/10.1371/journal.pone.0182782

Editor: Shihui Yang, National Renewable Energy Laboratory, UNITED STATES

Received: May 15, 2017

Accepted: July 24, 2017

Published: August 8, 2017

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This work was supported by the National Science Foundation (grant no. MCB0446431) (MA). The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

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

Abstract

Globin-coupled diguanylate cyclases contain globin, middle, and diguanylate cyclase domains that sense O_2 to synthesize c-di-GMP and regulate bacterial motility, biofilm formation, and virulence. However, relatively few studies have extensively examined the roles of individual residues and domains of globin-coupled diguanylate cyclases, which can shed light on their signaling mechanisms and provide drug targets. Here, we report the critical residues of two globin-coupled diguanylate cyclases, *Ec*GReg from *Escherichia coli* and *Bpe*GReg from *Bordetella pertussis*, and show that their diguanylate cyclase activity requires an intact globin domain. In the distal heme pocket of the globin domain, residues Phe42, Tyr43, Ala68 (*Ec*GReg)/Ser68 (*Bpe*GReg), and Met69 are required to maintain full diguanylate cyclase activity. The highly conserved amino acids His223/His225 and Lys224/Lys226 in the middle domain of *Ec*GReg/*Bpe*GReg are essential to diguanylate cyclase activity. We also identified sixteen important residues (Leu300, Arg306, Asp333, Phe337, Lys338, Asn341, Asp342, Asp350, Leu353, Asp368, Arg372, Gly374, Gly375, Asp376, Glu377, and Phe378) in the active site and inhibitory site of the diguanylate cyclase domain of *Ec*GReg. Moreover, $Bpe\text{GReg}_{266}$ (residues 1–266) and $Bpe\text{GReg}_{296}$ (residues 1–296), which only contain the globin and middle domains, can inhibit bacterial motility. Our findings suggest that the distal residues of the globin domain affect diguanylate cyclase activity and that *Bpe*GReg may interact with other c-di-GMP-metabolizing proteins to form mixed signaling teams.

Introduction

Globin-coupled diguanylate cyclases (GCDCs) form a subfamily of globin-coupled sensors (GCS) that are heme-binding sensors linked to variable signaling domains [1–5]. The N-terminal globin domains of GCDCs consist of eight alpha helices and display a myoglobin-like topology [6, 7]. The ferrous ion centered in the heme reversely binds O_2, and distal residues facilitate O_2 migration and stabilization in the hydrophobic heme pocket [6, 8–10]. In the C-

terminal diguanylate cyclase (DGC) domains of GCDCs, the highly conserved GGD/EEF (Gly-Gly-Asp/Glu-Glu-Phe) motif serves as the active site to synthesize the second messenger bis-(3'-5')-cyclic diguanosine monophosphate (c-di-GMP) [6, 8].

C-di-GMP-dependent signaling pathways regulate diverse cellular functions including motility, biofilm formation, virulence, differentiation, and the cell cycle [11]. Various c-di-GMP receptors have been identified in bacteria and the mammalian innate immune system [12–20]. C-di-GMP is synthesized by DGCs (including GCDCs) containing a GGDEF domain and degraded by phosphodiesterases (PDEs) with either an EAL or HD-GYP domain [11]. Multiple DGCs and PDEs are found in most bacteria and are often associated with sensory or regulatory domains that allow them to modulate their activities in response to internal and environmental stimuli [11]. Studies on how DGCs and PDEs sense environmental signals to regulate c-di-GMP levels will shed light on the mechanisms of bacterial behavior and provide potential drug targets to attenuate the virulence of pathogens.

Various GCDCs have been characterized from *Escherichia coli*, *Bordetella pertussis*, *Azotobacter vinelandii*, *Desulfotalea psychrophila*, *Shewanella putrefaciens*, and *Pectobacterium carotovorum* [6, 8–10, 21–23]. Two GCDCs, *Ec*GReg (also named DosC) from *E. coli* and *Bpe*GReg from the whooping cough pathogen *B. pertussis*, can sense O_2 to regulate c-di-GMP synthesis [6, 8]. They differ in respect of O_2 affinity and cooperative PDEs [6, 8, 24]. *Ec*GReg and *Ec* DosP (EAL type) couple to control c-di-GMP homeostasis, whereas *Bpe*GReg may cooperate with *Bpe* RpfG (HD-GYP type) [8, 24]. To address related questions on how globins regulate DGC activity, here we report the critical residues identified in the three domains of GCDCs and that GCDCs require an intact globin domain for their enzyme activities. We examined Phe42, Tyr43, Ala68 (*Ec*GReg)/Ser68 (*Bpe*GReg), and Met69, which are in the distal heme pocket of the globin domain. We also tested the highly conserved amino acids His223/His225 and Lys224/Lys226 in the middle domain of *Ec*GReg and *Bpe*GReg. In addition, we used *Ec*GReg as a model to examine sixteen conserved residues in the active site and inhibitory site of the DGC domains. We propose that distal globin residues facilitate O_2 binding to regulate DGC activity and *Bpe*GReg may interact with other DGCs to form mixed signaling teams.

Materials and methods

Sequence alignment and visualization of predicted critical residues

Protein sequences were aligned using EBI-MAFFT [25, 26]. The alignment was visualized by using EBI-MView [26, 27]. The homology model of *Bpe*GReg was created previously [6]. Visualization of the locations of the critical residues was carried out using UCSF Chimera [28].

Plasmid construction

The *ecGReg* (GenBank accession no. NP_416007) and *bpeGReg* (GenBank accession no. NP_882025) genes were cloned into the pTrc99A vector (primers listed in Table A in S1 File). The truncated versions of these genes were also cloned into the pTrc99A vector (primers listed in Tables B and C in S1 File).

The plasmids containing the full-length genes were used as templates for the QuickChange site-directed mutagenesis protocol (Stratagene). Briefly, *PfuTurbo* DNA polymerase (Stratagene) and the primers listed in Tables B and C in S1 File were used for polymerase chain reactions (PCR). Thermal cycling was carried out with 12 cycles as follows: 94˚C for 30 seconds, 55˚C for 30 seconds, and 68˚C for 12 minutes, with a final extension at 68˚C for 7 minutes. One microliter of *Dpn*I (10 U/μl, Promega) was then added to digest the methylated parental DNA template at 37˚C for 1 hour. The mutated plasmids were transformed into *E. coli* TOP10

cells (Invitrogen). The coding regions of the isolated plasmids were verified by Sanger sequencing on ABI 3730xl DNA analyzer.

For protein expression and purification, full-length and truncated *Bpe*GReg proteins were engineered with an N-terminal hexahistidine tag by PCR. Primers are listed in Table D in S1 File. The PCR products were cloned into the pCR4Blunt-TOPO vector (Invitrogen) and then subcloned into the pET-3a expression vector (Novagen).

The laboratory protocol of the molecular cloning has been deposited in protocols.io (DOI: http://dx.doi.org/10.17504/protocols.io.ijxccpn).

Phenotypic assays

Plasmids were transformed into *Salmonella typhimurium* ATCC 14028 by electroporation and maintained in LB broth or plates with ampicillin (100 μg/ml) at 37˚C. For biofilm formation, cells were grown on LB without salt plates containing Congo red (40 μg/ml) for 40 hours at 37˚C. Swimming motility was assayed on 0.3% agar plates (1% tryptone, 0.5% NaCl, 1 μM thiamine) at 28˚C for 6 hours. Due to leaky expression of the pTrc99A vector, isopropyl β-D-thiogalactopyranoside (IPTG) was not used to induce protein expression during these assays.

Overexpression, purification and absorption spectra

His-tagged *Bpe*GReg, *Bpe*GReg$_{155}$, *Bpe*GReg$_{266}$, and *Bpe*GReg$_{296}$ were overexpressed in *E. coli* Rosetta2(DE3)pLysS cells for 6–8 h at room temperature. Protein expression was induced with 0.05 mM isopropyl-β-D-thiogalactopyranoside. Proteins were purified by Co^{2+}-affinity chromatography according to Piatibratov et al. [29]. The detailed protocol has been deposited in protocols.io (DOI: http://dx.doi.org/10.17504/protocols.io.ikhcct6). Spectra were measured using a Cary 1E UV-Visible spectrophotometer (Varian).

Results and discussion

Prediction of critical residues of globin-coupled diguanylate cyclases

The open reading frames of *ecGReg* and *bpeGReg* encode multidomain proteins containing 460 and 475 amino acids, respectively (Fig 1A). The region linking the N-terminal globin domain and C-terminal DGC domain could not be classified but is highly conserved in many GCDCs, including *Ec*GReg and *Bpe*GReg [3]. We refer to this region as the middle domain (Fig 1A). Fig 1B shows the protein sequence alignment of *Ec*GReg, *Bpe*GReg, and other GCDCs with the same domain architecture. We predicted a series of residues in the globin, middle, and DGC domains of *Ec*GReg and *Bpe*GReg that may affect their enzymatic activities, based on amino acid sequence alignment and the homology model of *Bpe*GReg [6]. The crystal structure of *Ec*GReg [7] was not available at the time this study was performed, so it was only used for subsequently visualizing the positions of the residues we tested rather than aiding our predictions (Fig 2).

In the *Ec*GReg globin domain, the distal residue Tyr43 stabilizes O_2 to ferrous ion [10]. Our unpublished data suggests a similar role for Tyr43 in *Bpe*GReg. We selected Tyr43 and three other residues in the distal heme pocket of both proteins for site-directed mutagenesis. In the middle domain, *Bpe*GReg His225 was examined previously [6]. Here we selected His223 in *Ec*GReg and nearby conserved residues Lys224 (*Ec*GReg)/Lys226 (*Bpe*GReg) for mutagenesis. To analyze the conserved residues in the DGC domain, we used *Ec*GReg as a model and selected 16 residues in the active and inhibitory sites.

Mutated proteins were expressed in *Salmonella typhimurium* ATCC 14028 to examine their *in vivo* DGC activities by rdar (red, dry, and rough) formation and motility assays (Figs 3–6,

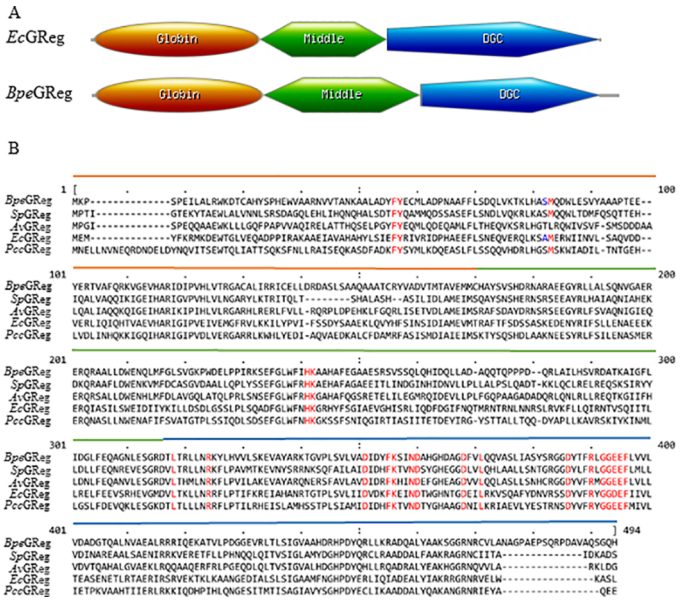

Fig 1. Two globin-coupled diguanylate cyclases consist of three domains. (A) Domain structures of *Ec*GReg and *Bpe*GReg. BLASTP searches against the non-redundant protein sequences database at the National Center for Biotechnology Information showed that *Ec*GReg residues 5–153 represent a globin sensor domain and residues 268–458 represent a DGC domain, while *Bpe*GReg residues 5–155 and residues 297–457 represent a globin sensor domain and a DGC domain, respectively. (B) Amino acid sequence alignment of five globin-coupled diguanylate cyclases: *Ec*GReg, *Bpe*GReg, *Av*GReg (*Azotobacter vinelandii*), *Sp*GReg (*Shewanella putrefaciens*), and *Pcc*GReg (*Pectobacterium carotovorum*). The sequence order is based on the alignment. The residues highlighted in red are highly conserved and discussed in this work. The residues highlighted in blue are not highly conserved and discussed in this work. The line colors indicate the globin (orange), middle (green), and DGC (blue) domains.

https://doi.org/10.1371/journal.pone.0182782.g001

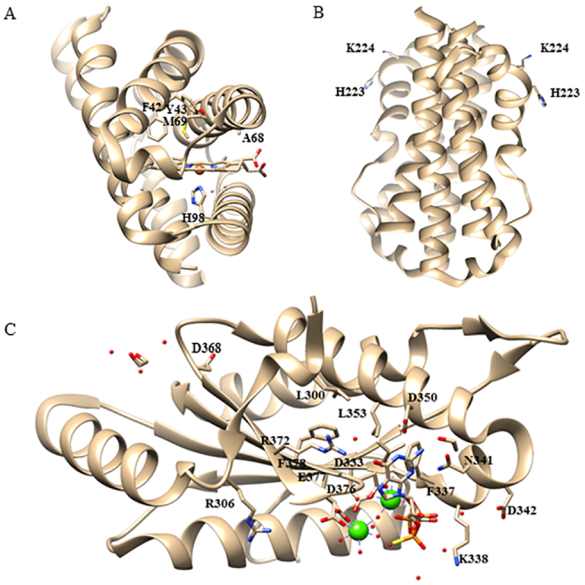

Fig 2. Positions of predicted critical amino acids in the crystal structure of *Ec*GReg/DosC. (A) Crystal structure of the globin domain of *Ec*GReg/DosC (pdb: 4zvb, ferrous form) shows the locations of the critical residues in the heme pocket. H98 is the proximal histidine which binds the heme. The side chains of F42, Y43, and M69 surround the heme center. The side chain of A68 points away from the heme center. (B) Crystal structure of the middle domain of *Ec*GReg/DosC (pdb: 4zvc, form I) shows the locations of critical residues H223 and K224, (C) Crystal structure of the DGC domain of *Ec*GReg/DosC (pdb: 4zvf, GTPαS-bound) shows the locations of mutated residues in this work.

https://doi.org/10.1371/journal.pone.0182782.g002

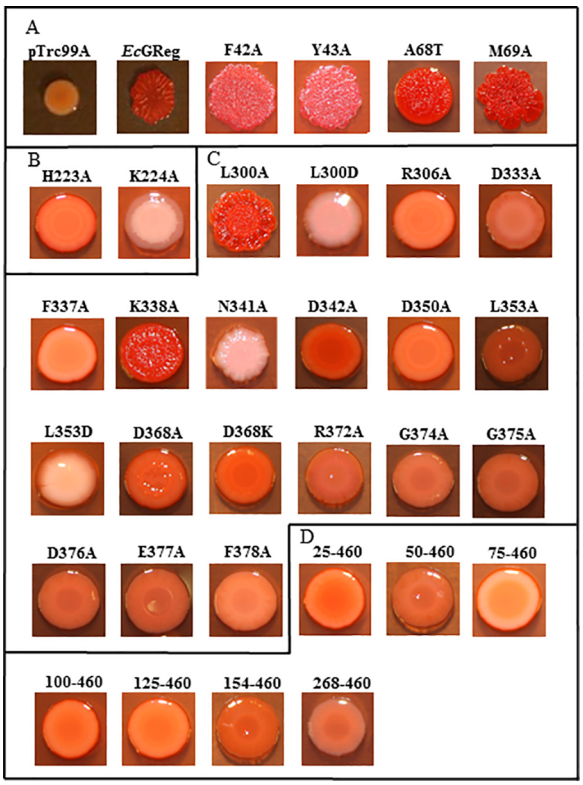

Fig 3. Rdar morphotype development of *Ec*GReg mutants on Congo red agar plates. (A) Rdar morphotypes of *Ec*GReg globin domain mutants. (B) Rdar morphotypes of *Ec*GReg middle domain mutants. (C) Rdar morphotypes of *Ec*GReg DGC domain mutants. (D) Rdar morphotypes of truncated *Ec*GReg proteins.

https://doi.org/10.1371/journal.pone.0182782.g003

Tables 1 and 2). High levels of intracellular c-di-GMP can inhibit motility, increase the production of exopolysaccharides (EPS) (*e.g.*, cellulose) and adhesion factors (*e.g.*, curli fimbriae), and enhance biofilm formation [30]. *S. typhimurium* develops an rdar morphotype on Congo red agar plates at 28˚C but not at 37˚C, indicating the expression of cellulose and curli fimbriae [31, 32]. At 37˚C, the temperature regulation of rdar morphotype development can be overcome by overexpression of the DGC AdrA as well as heterologous expression of other DGCs such as *Ec*GReg and *Bpe*GReg [6, 30, 33].

Phe42, Tyr43, Ala68/Ser68 and Met69 in the globin distal heme pockets of *Ec*GReg and *Bpe*GReg are required for DGC activity

In the N-terminal globin domains, we selected residues because they were located near the distal heme pocket, affecting the orientation and migration of ligands. Expression of the *Ec*GReg distal residue mutants, F42A, Y43A, A68T, and M69A, produced the rdar morphotype but failed to inhibit *S. typhimurium* motility (Figs 3A and 4A). Similar results were also observed for the *Bpe*GReg mutants F42A and M69A (Figs 5A and 6A). On the other hand, the *Bpe*GReg Y43A mutation severely impaired rdar development but did not affect the ability to inhibit motility (Figs 5A and 6A). Only the *Bpe*GReg S68A mutant could neither confer the rdar morphotype nor inhibit motility (Figs 5A and 6A). These results indicate that residues in the heme pocket of the globin domain are required for *Ec*GReg and *Bpe*GReg to be fully active. Particularly, in *Bpe*GReg, the hydroxyl group of Ser68 is required for full protein function.

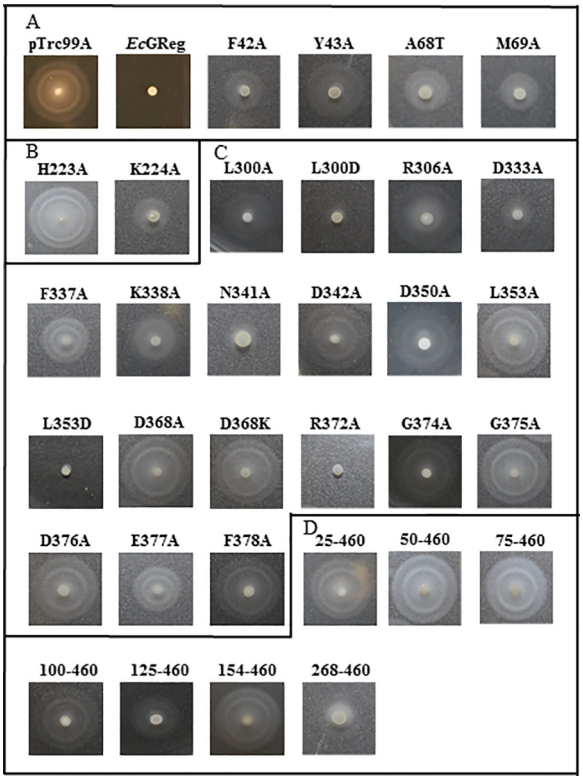

Fig 4. Swimming motility of *Ec*GReg mutants on 0.3% tryptone agar plates. (A) Swimming motility of *Ec*GReg globin domain mutants. (B) Swimming motility of *Ec*GReg middle domain mutants. (C) Swimming motility of *Ec*GReg DGC domain mutants. (D) Swimming motility of truncated *Ec*GReg proteins.

https://doi.org/10.1371/journal.pone.0182782.g004

His223/His225 and Lys224/Lys226 in the middle domains of *Ec*GReg and *Bpe*GReg are required for DGC activity

We previously demonstrated that *Bpe*GReg could be inactivated by an H225A mutation in the middle domain, as no detectable c-di-GMP was produced *in vitro* and no c-di-GMP-dependent

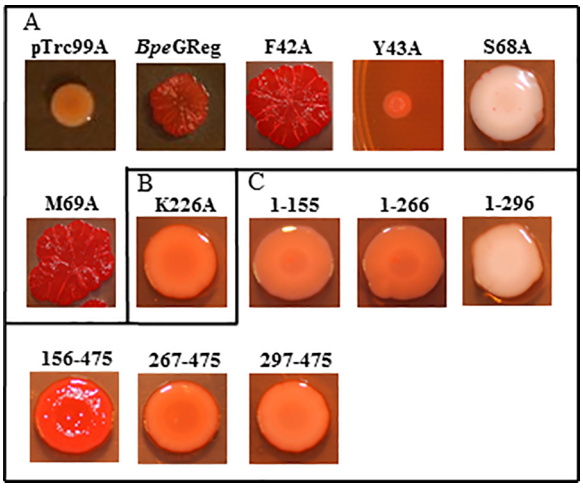

Fig 5. Rdar morphotype development of *Bpe*GReg mutants on Congo red agar plates. (A) Rdar morphotypes of pTrc99A vector control, *Bpe*GReg, and *Bpe*GReg globin domain mutants. (B) Rdar morphotypes of *Bpe*GReg middle domain mutant. (C) Rdar morphotypes of truncated *Bpe*GReg proteins.

https://doi.org/10.1371/journal.pone.0182782.g005

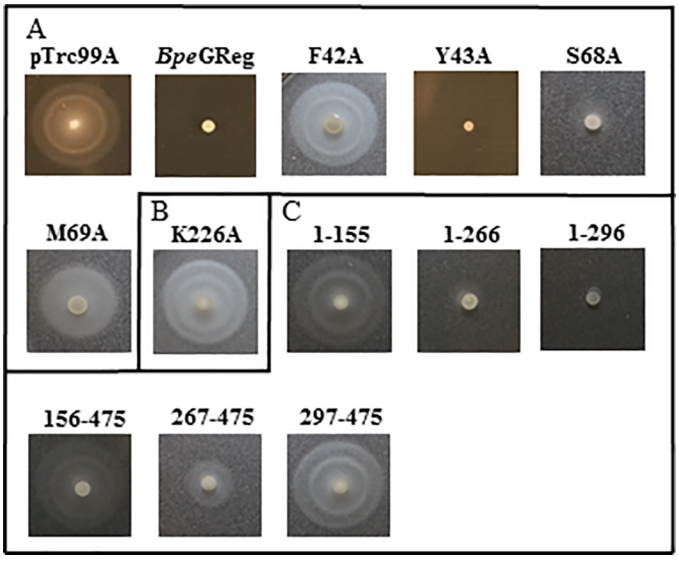

Fig 6. Swimming motility of *Bpe*GReg mutants on 0.3% tryptone agar plates. (A) Swimming motility of pTrc99A vector control, *Bpe*GReg, and *Bpe*GReg globin domain mutants. (B) Swimming motility of *Bpe*GReg middle domain mutant. (C) Swimming motility of truncated *Bpe*GReg proteins.

https://doi.org/10.1371/journal.pone.0182782.g006

physiological effects were observed when expressed in *S. typhimurium* [6]. When the corresponding residue in *Ec*GReg was mutated (H223A), this protein was also unable to induce the rdar morphotype or inhibit motility in *S. typhimurium* (Figs 3B and 4B). Mutation of another highly conserved residue, K224A (*Ec*GReg)/K226A (*Bpe*GReg), led to inactive phenotypes as well (Figs 3B–6B). Taken together, these results suggest an essential role of the middle domain for DGC activation. *Ec*GReg His223 and Lys224 locate in a π-helix (residues 221–225) that constitutes the middle region of the α-helix B (residues 216–228) of the middle domain [7]. Although the α-helices-associated π-helices are overlooked and rarely annotated, they are present in ~15% of all proteins and tend to be associated with protein function, *e.g.* peristaltic-like shifts to extend the binding cavity of a substrate [34]. In the homodimer structure of the *Ec*GReg middle domain (Fig 2B) [7], the side chains of His223 and Lys224 point towards the surrounding environment, contributing to the positive charges on the protein surface. Charged side chains on protein surfaces may play roles in allosteric regulation, protein-protein interaction, or folding/stability of protein. Further investigation will help us to understand the specific roles of the middle domain and the conserved π-helix of GCDCs.

Critical residues in the A-site and I-site of the DGC domain are required for enzyme activity

DGCs function as homodimers, with the two monomers forming the active site (A-site) at the dimer interface [35, 36]. The signature GG(D/E)EF motif constitutes part of the A-site, and several studies have suggested that absolute conservation of all five residues in this motif is required for catalysis [37, 38]. However, the first residue of the GG(D/E)EF motif appears to be flexible to substitutions in some DGCs. A DGC from *Pectobacterium atrosepticum* was observed to be active with Gly, Ser, or Ala in the first position, while a DGC from *Vibrio cholerae* was active with Gly, Ala, Met, or His [39, 40]. Using *Ec*GReg as a model, we wanted to determine which residues of the GGDEF motif were essential for its catalytic function. We therefore mutated each residue of the GGDEF motif to Ala (G374A, G375A, D376A, E377A,

and F378A). All of these mutants failed in conferring the rdar morphotype or inhibiting motility in *S. typhimurium* (Figs 3C and 4C), indicating that the canonical GGDEF motif is required for the DGC activity of *Ec*GReg. We further examined the roles of other residues within the A-site. Six mutations (R306A, D333A, F337A, D342A, D350A, and L353A) resulted in a lack of *in vivo* function (Figs 3C and 4C). Six other mutants retained partial function: cells harboring the L300A and K338A mutants formed a partial rdar morphotype (Fig 3C), while the L300D, N341A, L353D, and R372A mutants were able to inhibit or partially inhibit motility (Fig 4C). These results highlight the requirement of additional A-site residues for DGC activity, besides those of the GGDEF motif. Many of the residues (Asp333, Phe337, Lys338, Asn341, and Asp350) directly interact with either the substrate or metal ions in the A-site [7]. The other residues may help maintain the functional structure of the DGC domain.

An inhibitory site (I-site), consisting of an R*XX*D motif, is found in DGC domains, including those of *Ec*GReg and *Bpe*GReg [41]. The I-site is located five residues upstream of the GG(D/E)EF motif (Fig 1B). Feedback inhibition occurs when a c-di-GMP dimer binds to this allosteric site, thus decreasing DGC activity [35, 41]. We tested the effect of mutating the conserved Asp in the $R_{365}XXD_{368}$ motif of *Ec*GReg. The D368A and D368K variants both lacked the ability to produce the rdar morphotype or inhibit motility in *S. typhimurium* (Figs 3C and 4C). This is consistent with the findings of Kitanishi *et al.* [10], who observed that *Ec*GReg D368A did not have biofilm formation activity when expressed in *E. coli*. Furthermore, Burns *et al.* [42] found that mutation of the first Arg in the *Bpe*GReg I-site (R364A) affected oligomerization and decreased catalytic activity, leading to the conclusion that the I-site plays a role in controlling the conformation/dynamics of DGC domains, in addition to product inhibition.

The mutations characterized in this study can be classified into three groups: 1) mutants that could no longer confer the rdar morphotype or inhibit motility; 2) mutants that could confer the rdar morphotype but not inhibit motility; and 3) mutants that could inhibit motility but not confer the rdar morphotype (Tables 1 and 2). We conclude that the group 1 mutants, which are mainly located within the middle and DGC domains, include the most crucial residues required for DGC activity. The group 2 and 3 mutants are found in the globin and DGC domains. Particularly, the globin distal residue mutants Y43A are classified into different groups (*Ec*GReg in group 2 versus *Bpe*GReg in group 3). The differential effects of the group 2 and 3 mutants on rdar formation and motility could be due to signaling specificity. Rdar formation and motility inhibition involve a number of processes that are regulated at multiple levels, and particular DGCs and/or PDEs are often attributed to regulating specific processes [11]. The molecular mechanisms by which *Ec*GReg and *Bpe*GReg exert their specificities are not yet clearly understood. We presume that the group 2 and 3 mutations led to decreased c-di-GMP production, resulting in threshold levels sufficient for activating one phenotype over the other. This could involve c-di-GMP receptors with different binding affinities [43]. However, we cannot exclude the possibility that proximity to the appropriate targets was disrupted, as some of the mutations may have affected localization of the GCDC proteins in the cell or their association into protein complexes. It was shown that *Ec*GReg, along with *Ec* DosP, associates with PNPase in a ribonucleoprotein complex to regulate RNA turnover [15], but it is unknown how this may be connected to biofilm formation or motility.

The biofilm formation (rdar) activities of a few *Ec*GReg mutant variants have been previously evaluated in *E. coli* [10, 44], and the results were, for the most part, in agreement with what we observed for similar mutant proteins expressed in *S. typhimurium*. However, there was one discrepancy. While the H223A mutant displayed a wild-type phenotype in *E. coli* [10], our result indicated that the activity of the mutant was impaired. Differences in the c-di-GMP signaling networks of *E. coli* and *S. typhimurium* may account for the different phenotypic outcomes [45]. Contributions of species-specific DGCs/PDEs, interactions with distinct sets of

Table 1. Summary of phenotypes of *S. typhimurium* expressing *Ec*GReg mutants.

Residue	Mutant	RDAR	Motility inhibition	Group
Globin domain				
F42	A	+	-	2
Y43	A	+	-	2
A68	T	partial	-	2
M69	A	+	-	2
Middle domain				
H223	A	-	-	1
K224	A	-	-	1
DGC domain				
L300	A	partial	-	2
	D	-	+	3
R306	A	-	-	1
D333	A	-	-	1
F337	A	-	-	1
K338	A	partial	-	2
N341	A	-	partial	3
D342	A	-	-	1
D350	A	-	-	1
L353	A	-	-	1
	D	-	+	3
D368	A	-	-	1
	K	-	-	1
R372	A	-	+	3
G374	A	-	-	1
G375	A	-	-	1
D376	A	-	-	1
E377	A	-	-	1
F378	A	-	-	1

https://doi.org/10.1371/journal.pone.0182782.t001

proteins, or differences in the c-di-GMP thresholds needed to elicit a phenotypic response are among the possible factors that could lead to the variable effects seen in these organisms.

S. typhimurium expressing individual domains of *Bpe*GReg and *Ec*GReg cannot form the full rdar morphotype, while *Bpe*GReg$_{266}$ and *Bpe*GReg$_{296}$ can inhibit motility

Individual DGC domains have been shown to possess low-level enzymatic activity *in vitro*, which is several orders of magnitude lower than that of the full-length proteins [46]. A

Table 2. Summary of phenotypes of *S. typhimurium* expressing *Bpe*GReg mutants.

Residue	Mutant	RDAR	Motility inhibition	Group
Globin domain				
F42	A	+	-	2
Y43	A	partial	+	3
S68	A	-	-	1
M69	A	+	-	2
Middle domain				
K226	A	-	-	1

https://doi.org/10.1371/journal.pone.0182782.t002

construct containing the middle and DGC domains of *Ec*GReg (residues 173–460) exhibited four-fold less activity than the full-length protein, whereas the DGC domain alone (residues 297–460) did not have any detectable activity due to it being monomeric in solution [7]. In order to assess whether truncated forms of *Ec*GReg and *Bpe*GReg have DGC activity *in vivo*, we tested the functions of various constructs in *S. typhimurium* (Fig 7). We used *Ec*GReg as a model to examine whether it could maintain its function when the globin domain was gradually truncated. All of the *Ec*GReg truncated proteins could not confer a full rdar morphotype (Fig 3D), and were unable to inhibit motility as well (Fig 4D). We also tested the individual or combined domains of *Bpe*GReg. Intriguingly, two of the *Bpe*GReg constructs, *Bpe*GReg$_{266}$ and *Bpe*GReg$_{296}$ (residues 1–266 and 1–296), which only contain the globin and middle domains, were able to inhibit motility (Fig 6C). Compared to wild-type *Bpe*GReg, the absorption spectra of *Bpe*GReg$_{155}$, *Bpe*GReg$_{266}$ and *Bpe*GReg$_{296}$ showed that they were able to bind heme (Fig A in S1 File), indicating that the heme-based globin was present. The presence of the globin

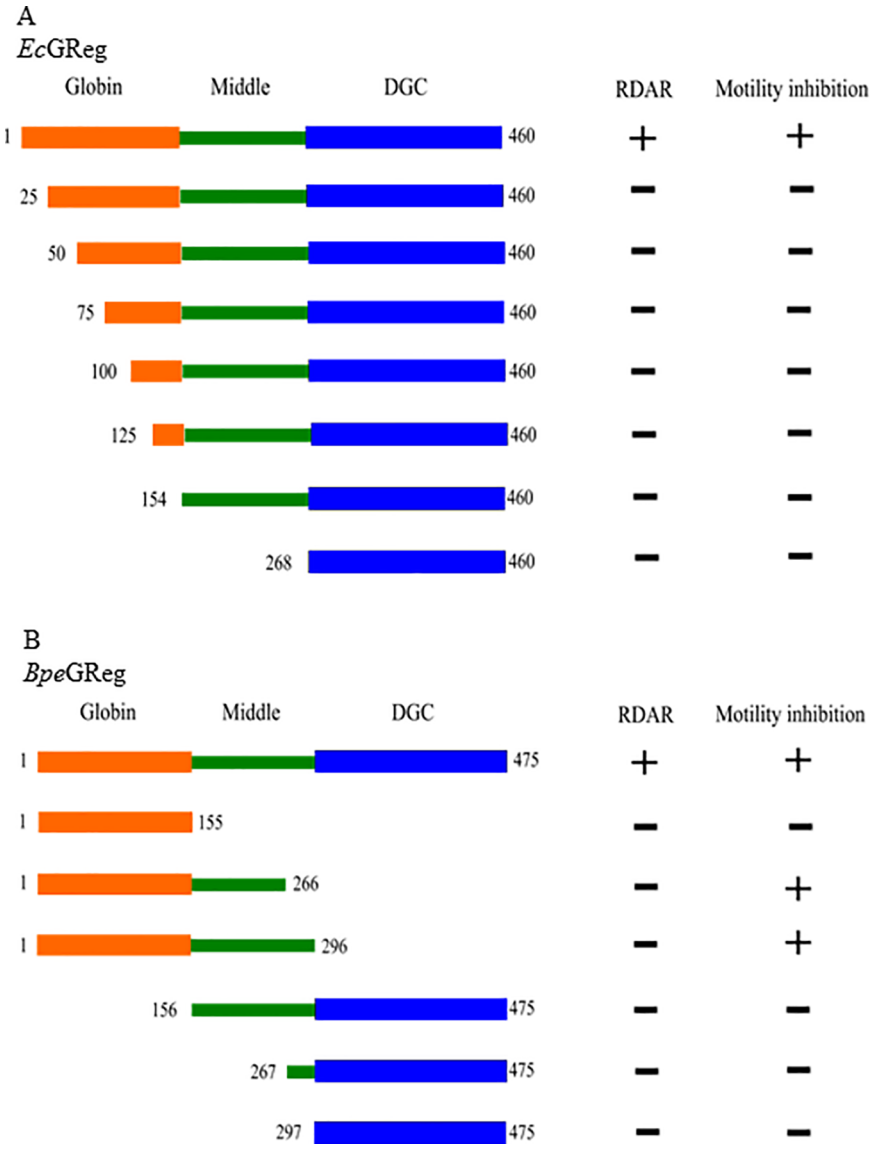

Fig 7. Schematic diagram of truncated *Ec*GReg (A) and *Bpe*GReg (B).

https://doi.org/10.1371/journal.pone.0182782.g007

domain alone (*Bpe*GReg$_{155}$), however, was not sufficient for eliciting a physiological response. Our data suggest that both the heme-based globin and middle domain are required for signaling and they may interact with other DGCs to form signaling clusters.

Our analysis of the truncated proteins suggests that the complete structures of *Ec*GReg and *Bpe*GReg are required for full DGC activity. Previous studies have demonstrated that the cyclase activity of GCDCs is regulated by O_2 binding to the globin sensor domain [6, 8, 21–23], so it is not surprising that the presence of this domain would be necessary for optimal function. However, it is unexpected to find that expression of two *Bpe*GReg constructs lacking the DGC domain resulted in motility inhibition. As these proteins did not have the ability to synthesize c-di-GMP, we postulate that the globin and middle domains of *Bpe*GReg may be able to interact with other DGCs, PDEs, or effector proteins to coordinate c-di-GMP signaling. Thus, the function of the truncated *Bpe*GReg proteins may be partially rescued by its partners in mixed signaling teams. There is growing evidence for the involvement of protein-protein interactions in regulating c-di-GMP networks. For example, *Ec*GReg forms a complex with the PDE DosP [8]; the stand-alone NO sensor H-NOX binds to HaCE, a dual-functioning DGC and PDE in *Shewanella woodyi* [47, 48]; and the DGC GcbC associates with its target receptor LapD in *Pseudomonas fluorescens* [49, 50].

Conclusion

Characterization of *Ec*GReg and *Bpe*GReg will provide general insights into the structure and function of GCSs. Here we provided evidence that residues Phe42, Tyr43, Ala68/Ser68 and Met69 in the distal heme pockets of *Ec*GReg and *Bpe*GReg could affect C-terminal DGC activity. His223/His225 and Lys224/Lys226 in the middle domains of *Ec*GReg and *Bpe*GReg were also required, and a number of critical residues in the A-site and I-site of the *Ec*GReg DGC domain were identified. In addition, the full globin fold is required for GCDC activity. We further hypothesize that *Bpe*GReg, via its globin and middle domains, may be able to form clusters with other c-di-GMP-metabolizing proteins. This may shed light on the functions of other globins, especially those single-domain proteins with unknown functions.

Supporting information

S1 File. Table A. Primers used for cloning *Ec*GReg and *Bpe*GReg in pTrc99A. Table B. Primers used for the construction of *Ec*GReg mutants and truncated *Ec*GReg. Table C. Primers used for the construction of *Bpe*GReg mutants and truncated *Bpe*GReg. Table D. Primers used for expression of truncated *Bpe*GReg in pET-3a. Figure A. Absorption spectra of truncated *Bpe*GReg proteins. a) Wild-type *Bpe*GReg (solid red line) showed heme-bound absorption spectra. b) *Bpe*GReg$_{155}$, *Bpe*GReg$_{266}$, and *Bpe*GReg$_{296}$ showed heme-bound absorption spectra, similar to that of wild-type *Bpe*GReg.
(PDF)

Acknowledgments

Dr. Maqsudul Alam passed away before the submission of the final version of this manuscript. Dr. Xuehua Wan accepts responsibility for the integrity and validity of the data collected and analyzed.

Author Contributions

Conceptualization: Xuehua Wan, Jennifer A. Saito, James S. Newhouse, Shaobin Hou, Maqsudul Alam.

Data curation: Xuehua Wan, Jennifer A. Saito, James S. Newhouse, Shaobin Hou.

Formal analysis: Xuehua Wan, Jennifer A. Saito, James S. Newhouse.

Funding acquisition: Maqsudul Alam.

Investigation: Xuehua Wan, Jennifer A. Saito, James S. Newhouse, Shaobin Hou, Maqsudul Alam.

Validation: Xuehua Wan, Jennifer A. Saito.

Writing – original draft: Xuehua Wan, Jennifer A. Saito.

Writing – review & editing: Xuehua Wan, Jennifer A. Saito.

References

1. Hou S, Larsen RW, Boudko D, Riley CW, Karatan E, Zimmer M, et al. 2000. Myoglobin-like aerotaxis transducers in *Archaea* and *Bacteria*. Nature 403:540–544 https://doi.org/10.1038/35000570 PMID: 10676961
2. Hou S, Freitas T, Larsen RW, Piatibratov M, Sivozhelezov V, Yamamoto A, et al. 2001. Globin-coupled sensors: a class of heme-containing sensors in *Archaea* and *Bacteria*. Proceedings of the National Academy of Sciences of the United States of America 98:9353–9358 https://doi.org/10.1073/pnas.161185598 PMID: 11481493
3. Freitas TAK, Hou S, Alam M. 2003. The diversity of globin-coupled sensors. FEBS Letters 552: 99–104 PMID: 14527668
4. Freitas TAK, Saito JA, Wan X, Hou S, Alam M. 2008. Protoglobin and globin-coupled sensors. In: Ghosh A, ed. The smallest biomolecules: diatomics and their interactions with heme proteins. Amsterdam London: Elsevier, 175–202
5. Kitanishi K, Kobayashi K, Uchida T, Ishimori K, Igarashi J, Shimizu T. 2011. Identification and functional and spectral characterization of a globin-coupled histidine kinase from *Anaeromyxobacter* sp. Fw109-5. Journal of Biological Chemistry 286:35522–35534 https://doi.org/10.1074/jbc.M111.274811 PMID: 21852234
6. Wan X, Tuckerman JR, Saito JA, Freitas TAK, Newhouse JS, Denery JR, et al. 2009. Globins synthesize the second messenger c-di-GMP in bacteria. Journal of Molecular Biology 388:262–270 https://doi.org/10.1016/j.jmb.2009.03.015 PMID: 19285985
7. Tarnawski M, Barends TR, Schlichting I. 2015. Structural analysis of an oxygen-regulated diguanylate cyclase. Acta Crystallographica D Biological Crystallography 71:2158–2177 https://doi.org/10.1107/S139900471501545X PMID: 26527135
8. Tuckerman JR, Gonzalez G, Sousa EH, Wan X, Saito JA, Alam M, et al. 2009. An oxygen-sensing diguanylate cyclase and phosphodiesterase couple for c-di-GMP control. Biochemistry 48:9764–9774 https://doi.org/10.1021/bi901409g PMID: 19764732
9. Thijs L, Vinck E, Bolli A, Trandafir F, Wan X, Hoogewijs D, et al. 2007. Characterization of a globin-coupled oxygen sensor with a gene-regulating function. Journal of Biological Chemistry 282:37325–37340 https://doi.org/10.1074/jbc.M705541200 PMID: 17925395
10. Kitanishi K, Kobayashi K, Kawamura Y, Ishigami I, Ogura T, Nakajima K, et al. 2010. Important roles of Tyr43 at the putative heme distal side in the oxygen recognition and stability of the Fe(II)-O2 complex of YddV, a globin-coupled heme-based oxygen sensor diguanylate cyclase. Biochemistry 49:10381–10393 https://doi.org/10.1021/bi100733q PMID: 21067162
11. Römling U, Galperin MY, Gomelsky M. 2013. Cyclic di-GMP: the first 25 years of a universal bacterial second messenger. Microbiology and Molecular Biology Reviews 77:1–52 https://doi.org/10.1128/MMBR.00043-12 PMID: 23471616
12. Ryan RP, Tolker-Nielsen T, Dow JM. 2012. When the PilZ don't work: effectors for cyclic di-GMP action in bacteria. Trends in Microbiology 20:235–242 https://doi.org/10.1016/j.tim.2012.02.008 PMID: 22444828
13. Amikam D, Galperin MY. 2006. PilZ domain is part of the bacterial c-di-GMP binding protein. Bioinformatics 22:3–6 https://doi.org/10.1093/bioinformatics/bti739 PMID: 16249258
14. Fang X, Ahmad I, Blanka A, Schottkowski M, Cimdins A, Galperin MY, et al. 2014. GIL, a new c-di-GMP-binding protein domain involved in regulation of cellulose synthesis in enterobacteria. Molecular Microbiology 93:439–452 https://doi.org/10.1111/mmi.12672 PMID: 24942809

15. Tuckerman JR, Gonzalez G, Gilles-Gonzalez MA. 2011. Cyclic di-GMP activation of polynucleotide phosphorylase signal-dependent RNA processing. Journal of Molecular Biology 407:633–639 https://doi.org/10.1016/j.jmb.2011.02.019 PMID: 21320509

16. Lori C, Ozaki S, Steiner S, Böhm R, Abel S, Dubey BN, 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 PMID: 25945741

17. Sudarsan N, Lee ER, Weinberg Z, Moy R H, Kim JN, Link KH, et al. 2008. Riboswitches in eubacteria sense the second messenger cyclic di-GMP. Science 321:411–413 https://doi.org/10.1126/science.1159519 PMID: 18635805

18. Burdette DL, Monroe KM, Sotelo-Troha K, Iwig JS, Eckert B, Hyodo M, et al. 2011. STING is a direct innate immune sensor of cyclic di-GMP. Nature 478:515–518 https://doi.org/10.1038/nature10429 PMID: 21947006

19. Parvatiyar K, Zhang Z, Teles RM, Ouyang S, Jiang Y, Iyer SS, et al. 2012. The helicase DDX41 recognizes the bacterial secondary messengers cyclic di-GMP and cyclic di-AMP to activate a type I interferon immune response. Nature Immunology 13:1155–1161 https://doi.org/10.1038/ni.2460 PMID: 23142775

20. Li W, Cui T, Hu L, Wang Z, Li Z, He Z-G. 2015b. Cyclic diguanylate monophosphate directly binds to human siderocalin and inhibits its antibacterial activity. Nature Communications 6:8330

21. Sawai H, Yoshioka S, Uchida T, Hyodo M, Hayakawa Y, Ishimori K, et al. 2010. Molecular oxygen regulates the enzymatic activity of a heme-containing diguanylate cyclase (HemDGC) for the synthesis of cyclic di-GMP. Biochimica et Biophysica Acta 1804:166–172 https://doi.org/10.1016/j.bbapap.2009.09.028 PMID: 19818878

22. Wu C, Cheng YY, Yin H, Song XN, Li WW, Zhou XX, et al. 2013. Oxygen promotes biofilm formation of *Shewanella putrefaciens* CN32 through a diguanylate cyclase and an adhesin. Scientific Reports 3:1945 https://doi.org/10.1038/srep01945 PMID: 23736081

23. Burns JL, Deer DD, Weinert EE. 2014. Oligomeric state affects oxygen dissociation and diguanylate cyclase activity of globin coupled sensors. Molecular BioSystems 10:2823–2826 https://doi.org/10.1039/c4mb00366g PMID: 25174604

24. Méndez-Ortiz MM, Hyodo M, Hayakawa Y, Membrillo-Hernández J. 2006. Genome-wide transcriptional profile of *Escherichia coli* in response to high levels of the second messenger 3',5'-cyclic diguanylic acid. Journal of Biological Chemistry 281:8090–8099 https://doi.org/10.1074/jbc.M510701200 PMID: 16418169

25. Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Molecular Biology and Evolution 30:772–780 https://doi.org/10.1093/molbev/mst010 PMID: 23329690

26. Li W, Cowley A, Uludag M, Gur T, McWilliam H, Squizzato S, et al. 2015a. The EMBL-EBI bioinformatics web and programmatic tools framework. Nucleic Acids Research 43:W580–W584

27. Brown NP, Leroy C, Sander C. 1998. MView: a web-compatible database search or multiple alignment viewer. Bioinformatics 14:380–381 PMID: 9632837

28. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, et al. 2004. UCSF Chimera—a visualization system for exploratory research and analysis. Journal of Computational Chemistry 25:1605–1612 https://doi.org/10.1002/jcc.20084 PMID: 15264254

29. Piatibratov M, Hou S, Brooun A, Yang J, Chen H, Alam M. (2000). Expression and fast-flow purification of a polyhistidine-tagged myoglobin-like aero- taxis transducer. Biochim. Biophys. Acta 1524: 149–154. PMID: 11113561

30. Simm R, Morr M, Kader A, Nimtz M, Römling U. 2004. GGDEF and EAL domains inversely regulate cyclic di-GMP levels and transition from sessility to motility. Molecular Microbiology 53:1123–1134 https://doi.org/10.1111/j.1365-2958.2004.04206.x PMID: 15306016

31. Römling U, Bian Z, Hammar M, Sierralta WD, Normark S. 1998. Curli fibers are highly conserved between *Salmonella typhimurium* and *Escherichia coli* with respect to operon structure and regulation. Journal of Bacteriology 180:722–731 PMID: 9457880

32. Zogaj X, Nimtz M, Rohde M, Bokranz W, Romling U. 2001. The multicellular morphotypes of *Salmonella typhimurium* and *Escherichia coli* produce cellulose as the second component of the extracellular matrix. Molecular Microbiology 39:1452–1463 PMID: 11260463

33. Kader A, Simm R, Gerstel U, Morr M, Römling U. 2006. Hierarchical involvement of various GGDEF domain proteins in rdar morphotype development of *Salmonella enterica* serovar Typhimurium. Molecular Microbiology 60:602–616 https://doi.org/10.1111/j.1365-2958.2006.05123.x PMID: 16629664

34. Cooley RB, Arp DJ, Karplus PA. 2010. Evolutionary origin of a secondary structure: π-helices as cryptic but widespread insertional variations of α-helices enhancing protein functionality. Journal of Molecular Biology 404:232–246 https://doi.org/10.1016/j.jmb.2010.09.034 PMID: 20888342

35. Chan C, Paul R, Samoray D, Amiot NC, Giese B, Jenal U, et al. 2004. Structural basis of activity and allosteric control of diguanylate cyclase. Proceedings of the National Academy of Sciences of the United States of America 101:17084–17089 https://doi.org/10.1073/pnas.0406134101 PMID: 15569936

36. Paul R, Abel S, Wassmann P, Beck A, Heerklotz H, Jenal U. 2007. Activation of the diguanylate cyclase PleD by phosphorylation-mediated dimerization. Journal of Biological Chemistry 282:29170–29177 https://doi.org/10.1074/jbc.M704702200 PMID: 17640875

37. Kirillina O, Fetherston JD, Bobrov AG, Abney J, Perry RD. 2004. HmsP, a putative phosphodiesterase, and HmsT, a putative diguanylate cyclase, control Hms-dependent biofilm formation in *Yersinia pestis*. Molecular Microbiology 54:75–88 https://doi.org/10.1111/j.1365-2958.2004.04253.x PMID: 15458406

38. Malone JG, Williams R, Christen M, Jenal U, Spiers AJ, Rainey PB. 2007. The structure-function relationship of WspR, a *Pseudomonas fluorescens* response regulator with a GGDEF output domain. Microbiology 153:980–994 https://doi.org/10.1099/mic.0.2006/002824-0 PMID: 17379708

39. Perez-Mendoza D, Coulthurst SJ, Humphris S, Campbell E, Welch M, Toth IK, et al. 2011. A multi-repeat adhesin of the phytopathogen, *Pectobacterium atrosepticum*, is secreted by a type I pathway and is subject to complex regulation involving a non-canonical diguanylate cyclase. Molecular Microbiology 82:719–733 https://doi.org/10.1111/j.1365-2958.2011.07849.x PMID: 21992096

40. Hunter JL, Severin GB, Koestler BJ, Waters CM. 2014. The *Vibrio cholerae* diguanylate cyclase VCA0965 has an AGDEF active site and synthesizes cyclic di-GMP. BMC Microbiology 14:22 https://doi.org/10.1186/1471-2180-14-22 PMID: 24490592

41. Christen B, Christen M, Paul R, Schmid F, Folcher M, Jenoe P, et al. 2006. Allosteric control of cyclic di-GMP signaling. Journal of Biological Chemistry 281:32015–32024 https://doi.org/10.1074/jbc.M603589200 PMID: 16923812

42. Burns JL, Rivera S, Deer DD, Joynt SC, Dvorak D, Weinert EE. 2016. Oxygen and bis(3',5')-cyclic dimeric guanosine monophosphate binding control oligomerization state equilibria of diguanylate cyclase-containing globin coupled sensors. Biochemistry 55:6642–6651 https://doi.org/10.1021/acs.biochem.6b00526 PMID: 27933792

43. Pultz IS, Christen M, Kulasekara HD, Kennard A, Kulasekara B, Miller SI. 2012. The response threshold of *Salmonella* PilZ domain proteins is determined by their binding affinities for c-di-GMP. Molecular Microbiology 86:1424–1440 https://doi.org/10.1111/mmi.12066 PMID: 23163901

44. Nakajima K, Kitanishi K, Kobayashi K, Kobayashi N, Igarashi J, Shimizu T. 2012. Leu65 in the heme distal side is critical for the stability of the Fe(II)-O_2 complex of YddV, a globin-coupled oxygen sensor diguanylate cyclase. Journal of Inorganic Biochemistry 108:163–170 https://doi.org/10.1016/j.jinorgbio.2011.09.019 PMID: 22005448

45. Sommerfeldt N, Possling A, Becker G, Pesavento C, Tschowri N, Hengge R. 2009. Gene expression patterns and differential input into curli fimbriae regulation of all GGDEF/EAL domain proteins in *Escherichia coli*. Microbiology 155:1318–1331 https://doi.org/10.1099/mic.0.024257-0 PMID: 19332833

46. Ryjenkov DA, Tarutina M, Moskvin OV, Gomelsky M. 2005. Cyclic diguanylate is a ubiquitous signaling molecule in bacteria: insights into biochemistry of the GGDEF protein domain. Journal of Bacteriology 187:1792–1798 https://doi.org/10.1128/JB.187.5.1792-1798.2005 PMID: 15716451

47. Liu N, Xu Y, Hossain S, Huang N, Coursolle D, Gralnick JA, et al. 2012. Nitric oxide regulation of cyclic di-GMP synthesis and hydrolysis in *Shewanella woodyi*. Biochemistry 51:2087–2099 https://doi.org/10.1021/bi201753f PMID: 22360279

48. Lahiri T, Luan B, Raleigh DP, Boon EM. 2014. A structural basis for the regulation of an H-NOX-associated cyclic-di-GMP synthase/phosphodiesterase enzyme by nitric oxide-bound H-NOX. Biochemistry 53:2126–2135 https://doi.org/10.1021/bi401597m PMID: 24628400

49. Dahlstrom KM, Giglio KM, Collins AJ, Sondermann H, O'Toole GA. 2015. Contribution of physical interactions to signaling specificity between a diguanylate cyclase and its effector. mBio 6:e01978–15 https://doi.org/10.1128/mBio.01978-15 PMID: 26670387

50. Dahlstrom KM, Giglio KM, Sondermann H, O'Toole GA. 2016. The inhibitory site of a diguanylate cyclase is a necessary element for interaction and signaling with an effector protein. Journal of Bacteriology 198:1595–1603 https://doi.org/10.1128/JB.00090-16 PMID: 27002135