Investigating the Allosteric Regulation of YfiN from *Pseudomonas aeruginosa*: Clues from the Structure of the Catalytic Domain

Giorgio Giardina[1,2]◉, Alessandro Paiardini[1]◉, Silvia Fernicola[1,2], Stefano Franceschini[1,2], Serena Rinaldo[1,2], Valentina Stelitano[1,2], Francesca Cutruzzolà[1,2]*

1 Department of Biochemical Sciences, Sapienza University of Rome, Rome, Italy, **2** Istituto Pasteur - Fondazione Cenci Bolognetti, Department of Biochemical Sciences, Sapienza University of Rome, Rome, Italy

Abstract

Pseudomonas aeruginosa is responsible for a plethora of biofilm mediated chronic infections among which cystic fibrosis pneumonia is the most frightening. The long-term survival strategy of *P. aeruginosa* in the patients lungs is based on a fine balance of virulence *vs* dormant states and on genetic adaptation, in order to select persistent phenotypes as the small colony variants (SCVs), which strongly correlate with antibiotic resistance and poor lung function. Recent studies have coupled SCV with increased levels of the signaling molecule cyclic di-GMP, and demonstrated the central role of the diguanylate cyclase YfiN, part of the tripartite signaling module YifBNR, in c-di-GMP dependent SCV regulation. YfiN, also called TpbB, is a multi-domain membrane enzyme connecting periplasmic stimuli to cytosolic c-di-GMP production by an allosteric inside-out signaling mechanism that, due to the lack of structural data, is still largely hypothetical. We have solved the crystal structure of the catalytic domain (GGDEF), and measured the enzymatic activity of the cytosolic portion in real-time by means of a newly developed method. Based on these results we demonstrate that, unlike other diguanylate cyclase, YfiN does not undergo product feedback inhibition, and that the presence of the HAMP domain is required for dimerization and catalysis. Coupling our structural and kinetic data with an *in silico* study we are now able to propose a model for the allosteric regulation of YfiN.

Introduction

The majority of chronic infections involve a biofilm stage. In most bacteria, the synthesis of the ubiquitous second messenger cyclic di-GMP (c-di-GMP) represents a common principle in the formation of otherwise highly diverse and species-specific biofilms [1–4]. Therefore, c-di-GMP signaling pathways play a key role in chronic infections [4]. The human pathogen *Pseudomonas aeruginosa* is responsible for a plethora of biofilm-mediated chronic infections among which cystic fibrosis (CF) pneumonia is the most frightening [5]. During long-term colonization of CF lungs *P. aeruginosa* undergoes specific genotypic adaptation to the host environment and, after a yearlong persistence, it develops small-colony variants (SCVs) [6–8]. SCVs, which display high intracellular c-di-GMP levels [9–11], are characterized by enhanced biofilm formation, high fimbrial expression, repression of flagellar genes, resistance to phagocytosis, and enhanced antibiotic resistance [10–14]; their appearance correlates with a poor patient clinical outcome [6,12,15]. A direct relationship between the presence of bacterial persister cells and the recalcitrant nature of chronic infections has been proposed [16].

The c-di-GMP metabolism in *P. aeruginosa* is highly complex: 42 genes containing putative diguanylate cyclases (DGCs) and/or phosphodiesterase are present [17]. It has been shown that SCVs generated *in vitro* as well as obtained from clinical isolates contain mutations that upregulate the activity of

a specific DGC, i.e. YfiN (also called TpbB [18], encoded by the PA1120 gene), suggesting a key role of this enzyme. Since YfiN is the effector protein of a tripartite signaling module YifBNR [14,19,20], in this work we choose to use the name YfiN for coherence with the other two members of the operon *PA1119* and *PA1121*, which, in the Pseudomonas genome database (http://www.pseudomonas.com/), are called YfiB and YfiR, respectively. Formation of SCVs depends on enhanced c-di-GMP output by YfiN, which elevates transcription of the *pel* operon [11,14,21]. The YfiBNR system likely contributes to the degree of persistence of *P. aeruginosa* cells in CF lungs. Jenal and coworkers [20], have shown, by looking at mutations in the YfiBNR genes found in clinical strains of *P. aeruginosa*, that the activity of YfiN (and the occurrence of the SCV phenotype) is under continuous cycles of positive and negative selection; the same group proposed that this mechanism may contribute to the *in vivo* fitness of *P. aeruginosa* during chronic lung infections.

Figure 1 illustrates the composition of the YfiBNR system. YfiN is an inner membrane protein composed of three domains: a periplasmic PAS domain, two transmembrane helices and a cytosolic portion of the protein, which includes an HAMP domain and a cyclase domain (named GGDEF from the conserved residues in the active site). The negative regulator YfiR [14,20] is a dimeric periplasmic protein which controls the activity of YfiN by binding to the PAS domain of the DGC. YfiR has also been proposed to sense the redox state (and therefore oxygen levels) thus possibly (and intriguingly) conveying signals related to the switch to the anaerobic mode of growth (including denitrification [22]), typical of *P. aeruginosa* chronic infections. A third component of the system is the YfiB protein, spanning the outer membrane and the peptidoglycan and involved in binding YfiR, thus relieving the repression of YfiN activity (Figure 1). No structural data are available for this system and therefore several aspects of this signaling pathway are yet to be discovered in order to define its role in SCV formation during chronic infections.

We have solved the crystal structure of the cyclase domain (GGDEF) of YfiN, completed its multi-domain structure by homology modeling, and performed a biochemical characterization of two constructs of the cytoplasmic portion of YfiN. We also measured the enzymatic activity by using a new approach for c-di-GMP detection in real-time [23]. Here we show that, unlike other DGC enzymes, YfiN does not undergo product feedback inhibition, and that the *in vitro* activity depends on the presence of the HAMP domain. Moreover, we propose that the predicted PAS domain is more likely to fold as the periplasmic N-terminal domain of the receptor LapD from *P. fluorescens* [24]. Coupling structural and biochemical data, we are able to suggest a mechanistic model for the allosteric regulation of YfiN in response to YfiR binding.

Results and Discussion

Crystal structure of the GGDEF domain

Based on fold and secondary structure prediction [25,26], YfiN is organized in three domains: a N-terminal domain, spanning residues 35-161, delimited by two transmembrane helices (TM1: residues 14-34 and TM2: 162-182); a HAMP domain (residues 183-246); a C-terminal GGDEF domain (residues 249-406). In order to gain structural insights on the allosteric regulation of YfiN, we expressed the cytosolic portion of the protein, including the HAMP and the GGDEF cyclase domain (YfiN$_{HAMP-GGDEF}$; residues 183-435). This truncated construct resulted monomeric in solution. After extensive crystallization trials we could finally collect a complete data set at 2.8 Å resolution from one single hexagonal crystal. The crystal belonged to the P6$_5$22 space group. Surprisingly, analysis of the unit cell solvent content (Matthews coefficient) clearly indicated that only one of the two domains of the protein could be physically present in the crystal lattice since fitting both domains in the cell volume would result in a solvent content of 11%, which is too low for a protein crystal. The solved structure confirmed that YfiN$_{HAMP-GGDEF}$ had actually undergone proteolysis and that only the GGDEF domain had crystallized (YfiN$_{GGDEF}$). The quality of the diffraction data is good and electron density is clearly visible for all main chain atoms spanning from residue 254 to 414 of the GGDEF domain (Figure S1 and Table 1).

The crystal structure of the catalytic domain of YfiN is composed by a five-stranded β-sheet core (β2-3-1-6-7) flanked by five α-helices (αA to F) (Figure 2). YfiN$_{GGDEF}$ also displays an additional peripheral β-hairpin (β4-5), which is present in all the homologues structures (PleD from *Caulobacter crescentus* [27,28]; WspR from *P. aeruginosa* [29,30]; XCC4471 from *Xanthomonas campestris* [31] and A1U3W3 from *Marinobacter aquaeolei* [32]) with the exception of WspR that displays a long loop in a very different conformation. As expected, the overall scaffold of the structure is similar to the previously solved analogues (Figure 2). However, the cyclase domain of YfiN significantly differs from the other homologues at the level of the allosteric inhibitory site (I-site).

YfiN displays a degenerated I-site

It is a general feature of DGCs to undergo a negative feedback inhibition caused by the product binding to the so-called I-site. In particular, c-di-GMP binds as a mutually intercalated dimer with sub micro-molar affinity to the DGCs that display a conserved I-site [27,28,30] and the final effect is a cross-link between two domains that hijacks these enzymes to an inactive conformation by spatially separating the two active site. The same binding mode of dimeric c-di-GMP is also observed in receptor proteins as PelD from *P. aeruginosa*, containing a degenerated GGDEF domain [33], or PP4397 from *P. putida*, that displays a PilZ domain [34]. In all cases, enzymes or receptors, when c-di-GMP binds as an intercalated dimer an interlock between two domains is observed. These can be either identical (i.e. GGDEF/GGDEF) or different domains (i.e. GGDEF/REC, GGDEF/GAF, YcgR-N/PilZ) (Figure 3A). Among the many residues that interact with dimeric c-di-GMP in these structures, three are invariantly present: an arginine and an aspartate on one domain and a second arginine on the other domain. In particular, whilst the aspartate is probably involved in ligand recognition and binding, the two arginine residues appear to be crucial for cross-linking to take place (Figure 3A). Essentially, these

YfiBNR tripartite system
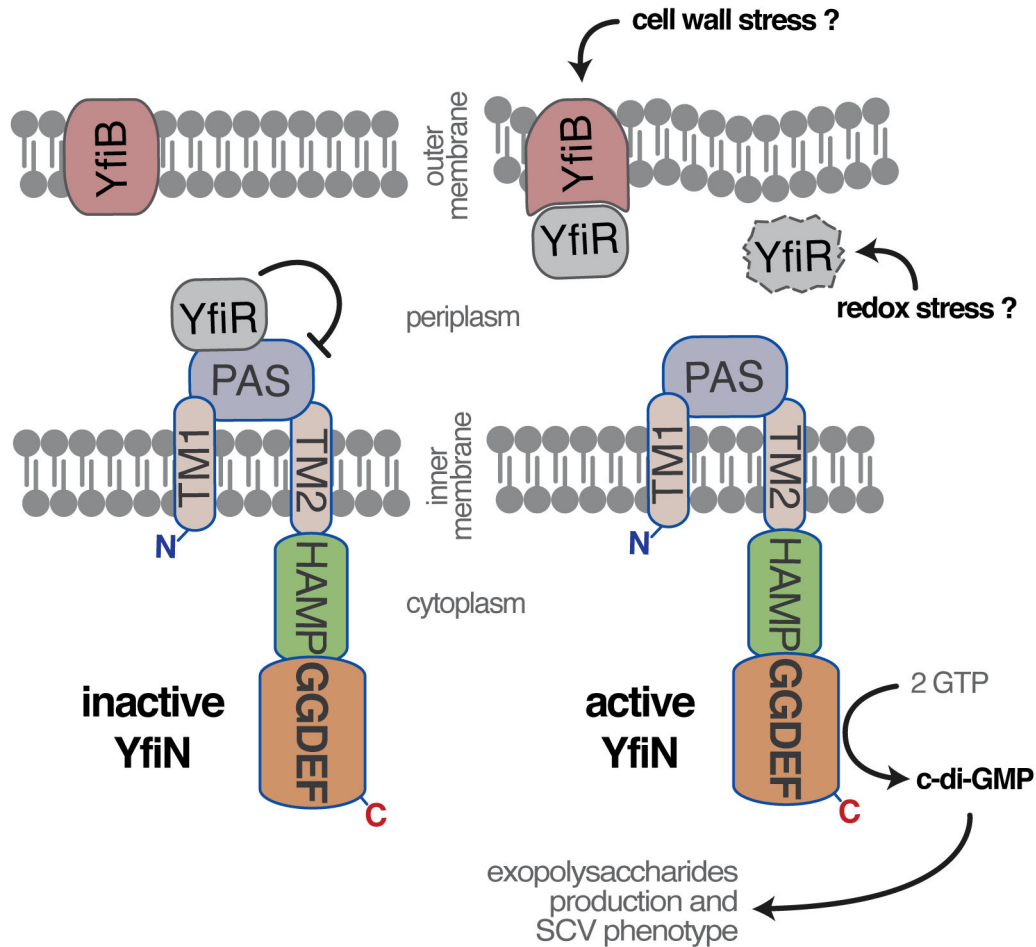
Figure 1. YfiBNR tripartite system organization. Schematic representation of the localization the YfiBNR system. YfiN is repressed by the specific interaction of YfiR with its periplasmic domain, while dissociation of the complex, and the consequent activation of YfiN, may be induced by a YfiB-mediated cell wall stress sensing mechanism and/or by redox driven misfolding of YfiR [20].

arginine residues bind c-di-GMP making a π-cation interaction with one guanine while H-bonding a second one. This peculiar binding mode is called *stair-motif* interaction and is recurrent in protein/DNA complexes [35,36]. Each arginine residues interacts with both c-di-GMP molecules. Therefore, since each domain provides one of the two key arginines, dimeric c-di-GMP is able to *glue* two domains through a double stair-motif interaction. In the case of the I-site of DGCs the first arginine is provided by the primary I-site (I_p) of the GGDEF domain (the conserved RxxD motif), while the second may be recruited from the secondary I-site (I_s) of another GGDEF domain [28,30,32] or from a different one (i.e. the REC domain of PleD [27] or the receptors PelD [33] and PP4397 [34]). Consequently, it must be clarified that the presence of the RxxD motif in the primary I-site of a DGC domain is a necessary but not sufficient condition for feedback inhibition, since a second arginine, provided by the I_s or another domain, is also needed.

The GGDEF domain of YfiN displays a conserved RxxD motif in the I_p, while the I_s appears degenerated. In particular, the second arginine necessary to form an inactive GGDEF/ GGDEF dimer, is substituted with Asp-273 (Figure 3B and 2B). Moreover, another important arginine is missing in YfiN I_s. This residue, which in PleD is Arg-390 and buttress (c-di-GMP)$_2$ by an additional stair-motif interaction [28], in YfiN is substituted with Asn-351. Finally, the α-helix harboring the I_s (α-A) is shifted with respect to the corresponding helix of PleD, WspR and A1U3W3, which all display product feedback inhibition. The shift is due to the hindrance of Tyr-379 side chain (Figure 3B). A similar shift, which hampers potential binding of (c-di-

Table 1. Data collection and refinement statistics for YfiN_GGDEF.

Coordinates	4IOB
Data collection	
Beamline	ESRF (ID14-1)
Space group	P 6_5 2 2
Cell dimensions	
$a = b, c$ (Å)	70.35, 106.87
Resolution (Å)	40.0-2.78 (2.94-2.78)
R_{factor}	8.3 (68.2)
I / sigma I	31.1 (3.3)
Completeness (%)	99.6 (98.2)
Reflections	
Observed	59914 (6510)
Unique	4343 (666)
B Wilson	57.9
Refinement	
Resolution (Å)	40.17-2.78
No. unique reflections	4095
R_{work} / R_{free}	27.8 / 28.0
Mean B-factor (Å^2) (atoms)	
Protein	48.1 (1235)
tert-butanol	33.3 (5)
Glycerol	54.9 (6)
R.m.s. deviations	
Bond lengths (Å)	0.0014
Bond angles (°)	0.460
Ramachandran: (%)	
Favored	93.1
Allowed	6.9

Values in parentheses refer to highest-resolution shell.

doi: 10.1371/journal.pone.0081324.t001

GMP)$_2$ to the I-site for sterical reasons, is observed only in the structure of XCC4471 that also displays a degenerated I-site [31].

These evidences suggest that YfiN is not able to undergo canonical product inhibition of DGCs, implying homodimer formation between the two catalytic domains. However, since the RxxD motif is conserved, the enzyme could still bind dimeric c-di-GMP and display product inhibition through an eventual cross-link of the GGDEF and HAMP domain, with the second arginine provided by the latter. To verify this possibility we measured the binding affinity of YfiN_HAMP-GGDEF for c-di-GMP.

YfiN_HAMP-GGDEF does not bind c-di-GMP

Binding of c-di-GMP to YfiN_HAMP-GGDEF was directly measured using isothermal titration calorimetry (ITC) and no binding was observed (Figure 4A). Of course an eventual misfolding of the soluble truncated construct could bias this result. To exclude this possibility we also measured the binding affinity of YfiN_HAMP-GGDEF for the substrate. Binding of GTP was carried out in the presence of CaCl$_2$, which does not allow hydrolysis after substrate binding. YfiN_HAMP-GGDEF binds GTP with submicromolar affinity and a stoichiometry close to one (Figure 4B). As

expected for specific binding, integration of the titration peaks produced a sigmoidal enthalpy curve for the interaction (the corresponding results are summarized in Table 2). It is worth mentioning that the K_d measured in this experiment could not correspond to the K_M value, since no catalysis has followed the binding event; moreover, it is not excluded that the affinity of GTP for the active site may be slightly altered by the calcium ion, with respect to the physiological metal (i.e. magnesium or manganese). To verify whether c-di-GMP could in any way hamper or negatively affect substrate binding to YfiN_HAMP-GGDEF, the GTP binding experiment was also repeated in the presence of an excess of product: no influence of c-di-GMP on the binding affinity of the substrate was observed (Figure S2 and Table 2). Taking these data together we can also exclude an eventual feedback inhibition mechanism involving hetero-domain cross-linking. To further verify whether these results could be affected by the truncation of the N-terminal portion of the enzyme, we measured the enzymatic activity of purified YfiN_HAMP-GGDEF.

YfiN_HAMP-GGDEF is active in vitro

The enzymatic activity of YfiN_HAMP-GGDEF could be measured using a new method for *in vitro* real-time quantification of c-di-GMP recently developed in our group [23]. We observed complete conversion of GTP to c-di-GMP (Figure 4C). It may be assumed that in order to condensate two GTP molecules, the GGDEF domains must come together at a certain time during catalysis. In this sense, it is important to notice that, although monomeric in solution, the purified YfiN_HAMP-GGDEF is still able to catalyze the condensation reaction of two molecules of GTP to c-di-GMP *in vitro*. Therefore, since neither the presence of the substrate nor that of the product changes the oligomeric state of the enzyme (data not shown), the formation of a transient catalytic dimer must be assumed. Indeed, the real time kinetics, as monitored with this new method, displays an interesting sigmoidal behavior (currently under investigation), which may well be related with such a mechanism. To verify the role of the HAMP domain in transient dimer formation, we produced a shorter construct containing only the GGDEF domain (YfiN_GGDEF; residues 232-435). This construct, which as expected is monomeric (Figure S5), although still able to bind GTP with micro-molar affinity, is completely inactive (Figure 4C and 4D), indicating that the HAMP domain is crucial for transient dimerization and catalysis to occur. On the other hand, the activity of YfiN_HAMP-GGDEF confirms that YfiN does not undergo product feedback inhibition, at least *in vitro* and in the micromolar range that we explored (up to 50 μM c-di-GMP). Likewise, Wood and co-workers have shown that *in vitro* feedback inhibition for full-length YfiN is observed only at c-di-GMP concentration higher than 200 μM [18].

Thus, the YfiBNR signaling system appears to be an ON/OFF switch, with the output of the module (i.e. c-di-GMP production) responding only to external stress signals and not to endogenous c-di-GMP levels. It as been shown that the domain architecture of YfiN represents a widespread module to connect periplasmic stimuli to a cytosolic response or *vice*
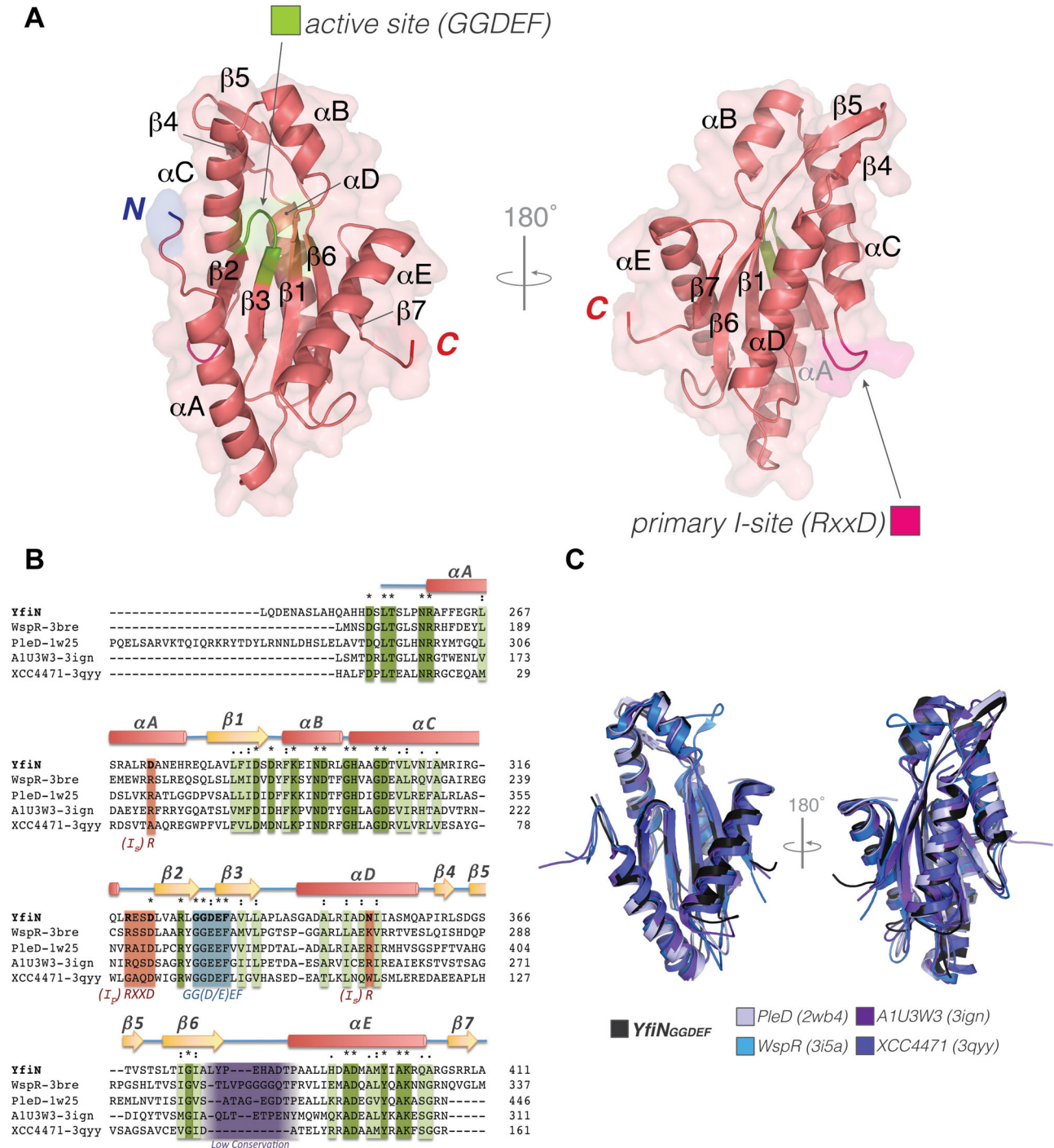
Figure 2. Cristal structure of YfiN$_{GGDEF}$. **A)** Cartoon representation of the YfiN$_{GGDEF}$ structure. The active site and primary inhibitory site (I$_p$) signature residues (GGDEF and RxxD) are shown in green and magenta respectively. **B)** Sequence alignment of the GGDEF domain of YfiN with the other DGCs of known structure; PleD from *C. crescentus* [27,28]; WspR from *P. aeruginosa* [29]; A1U3W3 from *M. aquaeolei* [32] and XCC4471 from *X. campestris* [31]. **C)** Structure superposition of YfiN$_{GGDEF}$ with the other DGC. YfiN$_{GGDEF}$ (black); PleD from *C. crescentus* [27,28] (grey - PDB: 2wb4 – rmsd: 1.23 Å); WspR from *P. aeruginosa* [29] (cyan - PDB: 3i5a - rmsd: 1.31 Å); XCC4471 from *X. campestris* [31] (light purple - PDB: 3qyy - rmsd: 1.64 Å) and A1U3W3 from *M. aquaeolei* [32] (dark purple - PDB: 3ign - rmsd: 1.34 Å).

doi: 10.1371/journal.pone.0081324.g002
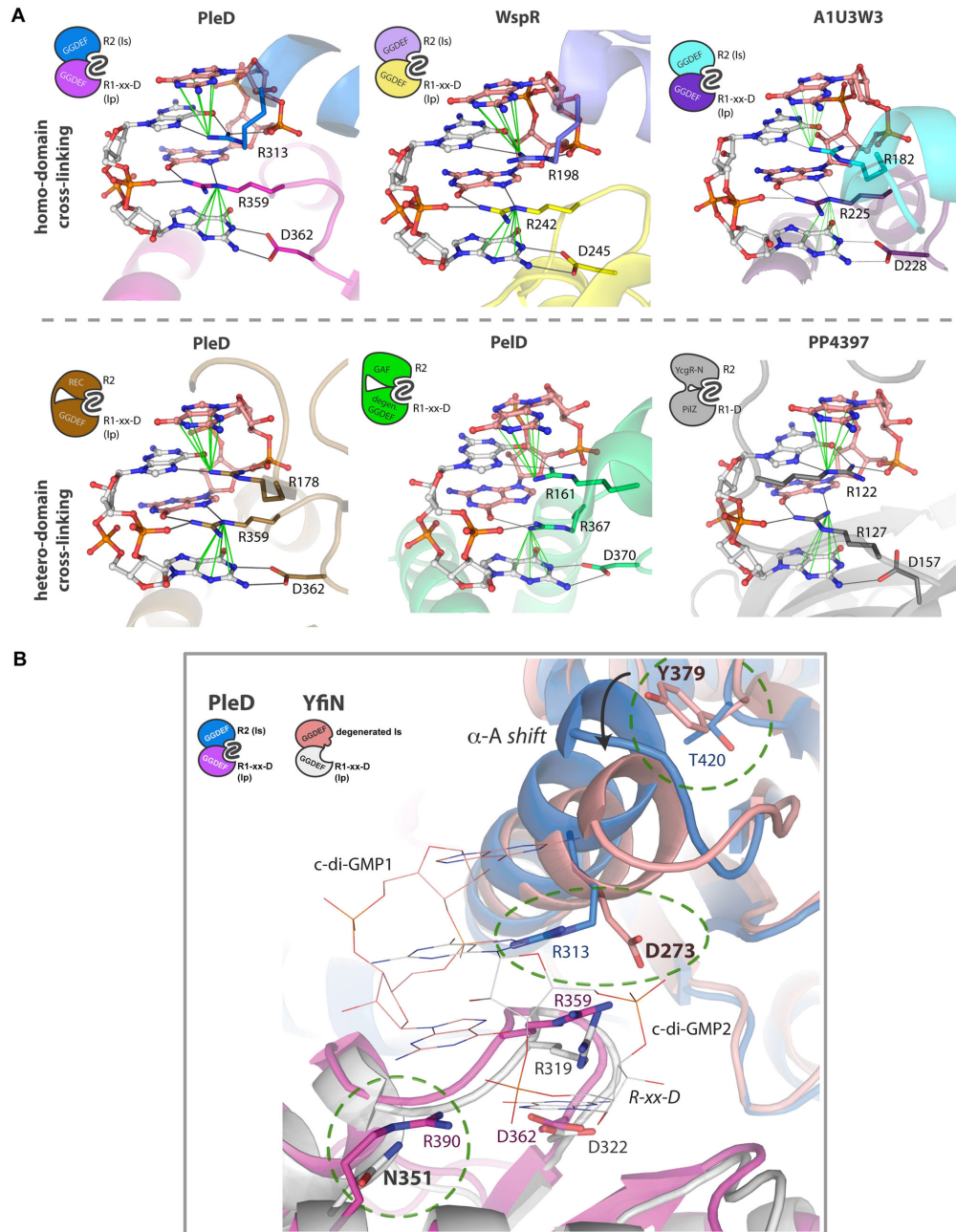
Figure 3. YfiN displays a degenerated I_s-Site. A) Binding mode of dimeric c-di-GMP to the I-site of DGCs or to receptor proteins. The first row shows the homo-domain cross-linking (GGDEF/GGDEF), while the second shows the hetero-domain cross-linking (within the same chain) of inhibited PleD and two c-di-GMP receptors. For all structures different colors are used to illustrate domains belonging to different subunits, the side chains of the two arginines and the aspartic acid (R1; R2 and D) are shown as sticks, while the two bound c-di-GMP molecules as balls and sticks. Grey continuous lines indicate H-bonds, while green continuous lines highlight the π-cation interaction among a charged nitrogen atom of the arginine residues and the guanine delocalised π-system. I_p and I_s indicate primary and secondary inhibitory sites respectively. Starting from top left, the reported structure are: PleD (PDB: 2v0n [28]); WspR (PDB: 3bre [30]); A1U3W3 (PDB: 3ign [32]); PleD (PDB: 1w25 [27]); PelD (PDB: 4dn0) [33]. and PP4397 (PDB: 3kyf [34]). **B**) Comparison of the I-site of YfiN and (PDB: 2v0n [28]). The two subunits of a hypothetical inhibited dimer of YfiN (superposed on the structure of PleD) are shown in white and pink, while the same color code of panel A is used for PleD. C-di-GMP molecules (bound to PleD) are shown as lines. YfiN lacks two of the three arginine residues binding to c-di-GMP through the stair motif interaction (D273 and N351 - bold labels). Moreover, the presence of a bulky side chain (Y379) yields a shift of helix-A, implying a reduced, sub optimal, volume of the I-site.

doi: 10.1371/journal.pone.0081324.g003
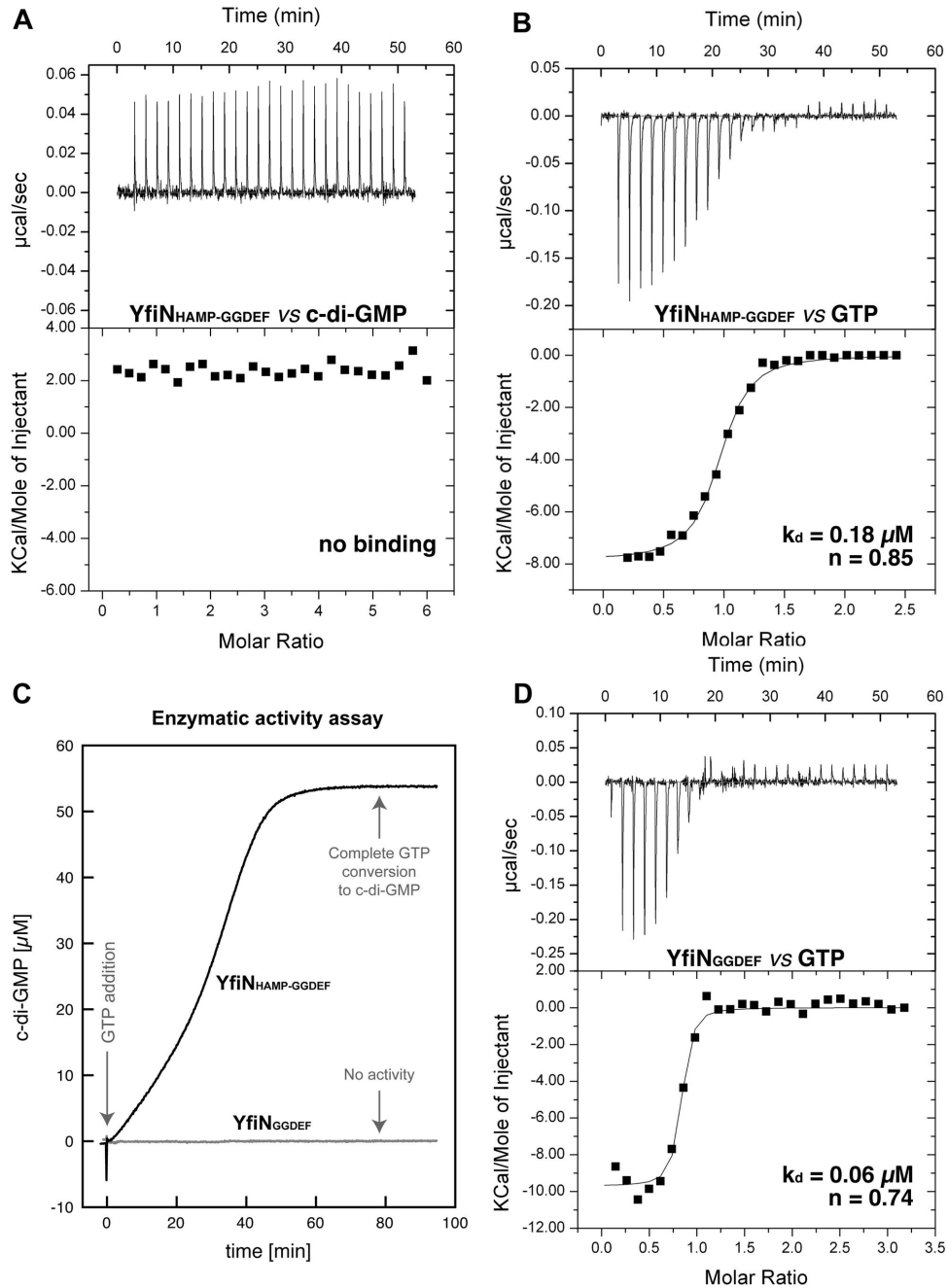
Figure 4. Binding affinity for nucleotides and enzymatic activity of YfiN_{HAMP-GGDEF} and YfiN_{GGDEF}. For all ITC experiments upper panels show the Raw ITC data, while lower panels show the integrated peak areas (black square) fitted with the one-binding-site model of ORIGIN provided by MicroCal (continuous lines). Derived thermodynamic parameters are listed in Table 2 A) Microcalorimetric titration of 3 µM YfiN_{HAMP-GGDEF} with c-di-GMP (90 µM in the syringe). No binding was observed either in the presence of $CaCl_2$ or in the presence of $MgCl_2/MnCl_2$ (data not shown). No thermodynamic parameters were derived. **B)** Microcalorimetric titrations of 14 µM enzyme solution with GTP (170 µM in the syringe). The thermodynamic profile indicates that the interaction of YfiN_{HAMP-GGDEF} with GTP presents favorable binding enthalpy and entropy, which suggests that hydrogen bonding and hydrophobic interactions are mainly involved in the binding event, rather than conformational changes. **C)** Cyclase activity of 10 µM YfiN_{HAMP-GGDEF} or YfiN_{GGDEF} assayed in real time by circular dichroism spectroscopy after addition of 100 µM GTP. For YfiN_{HAMP-GGDEF} (Black) The final c-di-GMP concentration corresponds to complete conversion of 100 µM GTP, whilst for YfiN_{GGDEF} (grey) no product is detected even if the sample is allowed to react for 24 h (not shown). **D)** Microcalorimetric titrations of 11 µM YfiN_{GGDEF} with GTP (170 µM in the syringe).
doi: 10.1371/journal.pone.0081324.g004

Table 2. Thermodynamic parameters derived from Microcalorimetric titrations of YfiN$_{HAMP-GGDEF}$ and YfiN$_{GGDEF}$ with nucleotides.

Protein	Ligand	n	K$_a$ x 10^6 M^{-1}	K$_d$ μM	ΔH kcal/mol	−TΔS kcal/mol	ΔG kcal/mol
YfiN$_{HAMP-GGDEF}$	GTP	0.85 ± 0.1	5.62 ± 1.9	0.18	-8.1 ± 0.3	-1.29	-9.36
YfiN$_{HAMP-GGDEF}$ [a]	GTP	0.73 ± 0.03	6.46 ± 2.7	0.15	-7.1 ± 0.3	-2.24	-9.30
YfiN$_{HAMP-GGDEF}$	c-di-GMP	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
YfiN$_{GGDEF}$	GTP	0.74 ± 0.04	18.1 ± 7.5	0.067	-9.9 ± 0.9	-5.31	-10.4

Values are the means of three independent experiments.

a. This experiment was done after incubation of both GTP and protein samples with 40 μM c-di-GMP.

doi: 10.1371/journal.pone.0081324.t002

versa [14,37–39]. It is, therefore, compelling to clarify the molecular detail of this allosteric inside-out signaling system.

Homology modeling of full-length YfiN

To gain insights into the mechanism of allosteric regulation of YfiN and how modifications affecting the periplasmic domain are transmitted into the cytoplasm, homology modeling of the full-length dimeric protein was attempted. Figure 5 shows the predicted domain organization of YfiN along with the most significant structural templates found, according to two different fold prediction servers (i.e., Phyre2 [25] and HHPRED [26]), and the dimeric model of YfiN. The N-terminal region of YfiN has been previously predicted to fold as a PAS domain, and consequently modeled [20] using as structural template the Sensor Kinase CitA binding domain (PDB Code: 1p0z [40]). However, the recent finding that the N-terminal domain of the HAMP-GGDEF-EAL protein LapD from *P. fluorescens* adopts a novel fold, consisting of a V-shaped, domain-swapped dimer (PDB Code: 3pjv [24]) that shows only weak structural similarity to the PAS fold (RMSD ~2.5 Å), prompted us to investigate further this issue by resubmitting the N-terminal region of YfiN to HHPRED and another fold prediction method, Phyre2 [25]. Consistent with our premise, residues 35-161 of YfiN are predicted to fold as a swapped LapD-like domain with a score and significance (HHPRED: E-value = 5.1 e-04, score = 53.05, confidence = 98.2%; Phyre2: confidence = 97.2%) higher compared to the Sensor kinase CitA (HHPRED: E-value = 1.3, score = 33.59, confidence = 91.2%). Each arm of this fold consists of two α-helices and two β-strands contributed by one of the two protomers, complemented by two β-strands flanked by helical segments from the other [24]. As in LapD, the N- and C-terminal helices of the LapD-like domains presumably connect directly to the transmembrane helices (TM2) and the HAMP domains. To model the later domain (residues 182-246) we used as structural template the HAMP domain of the aerotaxis transducer AER2 (PDB Code: 4l3M [39]), while transmembrane helices and neighboring positively charged loop regions (residues 11-34; 162-184) were modeled based on Sensor protein QSEC (PDB Code: 2KSE [41]), for all alignments see Figure S3. Finally, the model was connected to the crystal structure of the C-terminal GGDEF domain by modeling the linker region (residues 247-253) on the basis of the template diguanylate cyclase response regulator WspR (PDB Code: 3I5C [29]).
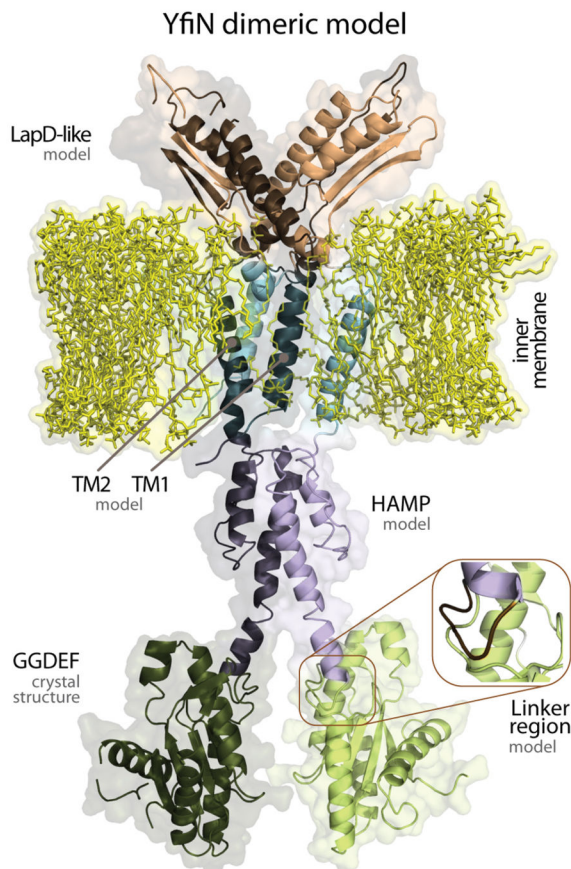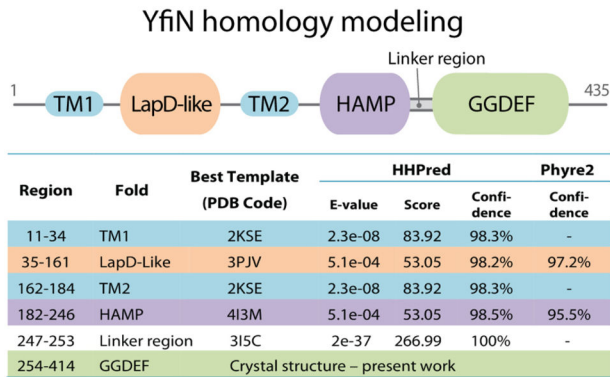
Following the results of the homology modeling it is likely that the allosteric switch of YfiN resembles that suggested for the LapD receptor [24]. In particular, as illustrated in Figure 6, YfiR would bind in the central gorge of the V-shaped PAS domain of YfiN's dimer. The release of the complex should produce a conformational change of the two arms of the PAS domains resulting in a shift of the TM2 helices, which are pushed towards the cytosolic side of the inner membrane. This movement of the TM2 should then be transmitted through a torsion of the HAMP domains helices to the terminal of this allosteric chain that is the conserved linker region connecting the last α-helix of the HAMP (*stalk helix*) to the GGDEF domain. The final effect is the unlocking of the C-terminal domains, which are now able to adopt a catalytically competent dimeric conformation (Figure 6).

Normal modes and sequence conservation analyses are in agreement with the allosteric regulation model of YfiN

To support this hypothetical mechanism, we analyzed the conformational changes and hinge regions of YfiN, underpinning its allosteric regulation. To this end, we applied coarse-grained, residue-level elastic network models (namely, the Gaussian Network Model [GNM] and its extension Anisotropic Network Model [ANM] [42,43]) to the full dimeric model of YfiN. Movie S1 provides a convenient visualization of the obtained results. The predicted LapD-like domain of YfiN undergoes a very large conformational bending, varying the angle between the arms of the V-shaped fold, most likely as a consequence of YfiR binding. Such a bending triggers, through the movement of the TM2 helices and the first predicted hinge region (residues 153-154), a torsional rotation of the downstream HAMP domain, which could form therefore the structural basis for modulating the interaction between the C-terminal GGDEF domains, possibly through an unlocking of the second predicted hinge, the linker region (residues 247-253).

As an additional indirect support to this hypothetical mechanism, we mapped the sequence conservation of YfiN and the position of known activating/inactivating mutations [20] on the full length model of YfiN, to confirm the potentially important regions for activity and/or allosteric regulation (Figure 7). Therefore, a multiple sequence alignment of 53 non-redundant orthologous of YfiN sequences was constructed

YfiN homology modeling

Region	Fold	Best Template (PDB Code)	HHPred			Phyre2
			E-value	Score	Confidence	Confidence
11-34	TM1	2KSE	2.3e-08	83.92	98.3%	-
35-161	LapD-Like	3PJV	5.1e-04	53.05	98.2%	97.2%
162-184	TM2	2KSE	2.3e-08	83.92	98.3%	-
182-246	HAMP	4I3M	5.1e-04	53.05	98.5%	95.5%
247-253	Linker region	3I5C	2e-37	266.99	100%	-
254-414	GGDEF	Crystal structure – present work				

YfiN dimeric model

Figure 5. Dimeric model of YfiN. Predicted domain organization of YfiN along with the most significant structural templates found, according to two different fold prediction servers (i.e., Phyre2 [25] and HHPRED [26]) used for homology modeling. The final model including the crystal structure of the catalytic domain is also shown.

doi: 10.1371/journal.pone.0081324.g005

from other Pseudomonas strains and from more distantly related sequences from other bacteria (Figure S4).

Strikingly, the accessible central gorge of the LapD-like periplasmic domain, presumably involved into the interaction of the periplasmic domain with YfiR, is characterized by a well-

conserved helix spanning residues 44-72 (aLrxYaxxNlxLiaRsxxYTxEaavvFxD; Figure 7A). This region not only is highly exposed but also includes 90% of the identified mutations in the periplasmic domain of YfiN that produce YfiR-independent alleles (residues 51, 58-59, 62, 66-68, 70) [20].

The folding of the dimeric HAMP domains as a four-helices bundle is also supported by the strict conservation of the core of the helix-loop-helix motif putatively involved in dimerization with the other monomer (residues 216-235: ELxxlxxDFNxLxdElexWq; (Figure 7B). Interestingly, since both YfiN$_{HAMP-GGDEF}$ and YfiN$_{GGDEF}$ constructs are monomeric in *in vitro* and bind GTP with similar affinity, but only the first is able to further condensate it to c-di-GMP, we must assume that, for YfiN$_{HAMP-GGDEF}$, catalysis proceeds through a HAMP-mediated transient dimerization. Therefore, we can speculate that the periplasmic domain of YfiN may not only play a regulatory role, but would also be essential to maintain the enzyme in a dimeric state, allowing the HAMP domains to form a stable four-helices bundle, thus keeping the two GGDEF domains in close proximity.

The linker region between the C-terminal GGDEF domain and the stalk helix of the HAMP domain, that we suggest to be crucial in the allosteric regulation, is also highly conserved (residues 249-260: AxHDxLTgLxNR) (Figure 7C). The importance of this region is confirmed by the deletion mutant Δ255-257, which is inactive and is dominant over the activating substitution G173D [20]. We have modeled this loop on the basis of the inhibited structure of WspR (PDB Code: 3I5C [29]) but, based on the location of the GTP binding site, this conformation would be incompatible with a catalytic encountering of the two GGDEF domains. Therefore, a severe rearrangement of this region, as a consequence of the HAMP domains torsion, must be assumed for catalysis to take place. Thereby, the role of the linker region would be to allosterically allow or deny the encountering of the two GGDEF domains depending on the HAMP conformation. Moreover, since this linker loop is located near the substrate binding site, it is not excluded that GTP binding may also play a role in the conformational change of this region of the enzyme.

Finally, the C-terminal GGDEF domain is also characterized by a large evolutionarily conserved surface region, which comprise the active site GGDEF motif (residues 319-338: RexDxVaRlGGDEFavllxp), and the adjacent helix-turn-helix region (residues 290-310: DxDxFKxxNDxxGHaxGDxVL;) (Figure 7C). These are presumably involved in GTP binding and monomer-monomer contacts upon formation of the catalytically competent GGDEF dimer.

Conclusions

We have shown that YfiN displays a degenerated secondary I-site and that the conserved primary I-site (RxxD) has no counterpart supplied by the HAMP domain, since YfiN$_{HAMP-GGDEF}$ is not able to bind c-di-GMP. On the other hand, YfiN$_{HAMP-GGDEF}$ binds GTP with sub-micromolar affinity, and is able to condensate it into c-di-GMP. These data point to the conclusion that YfiN does not undergo product feedback
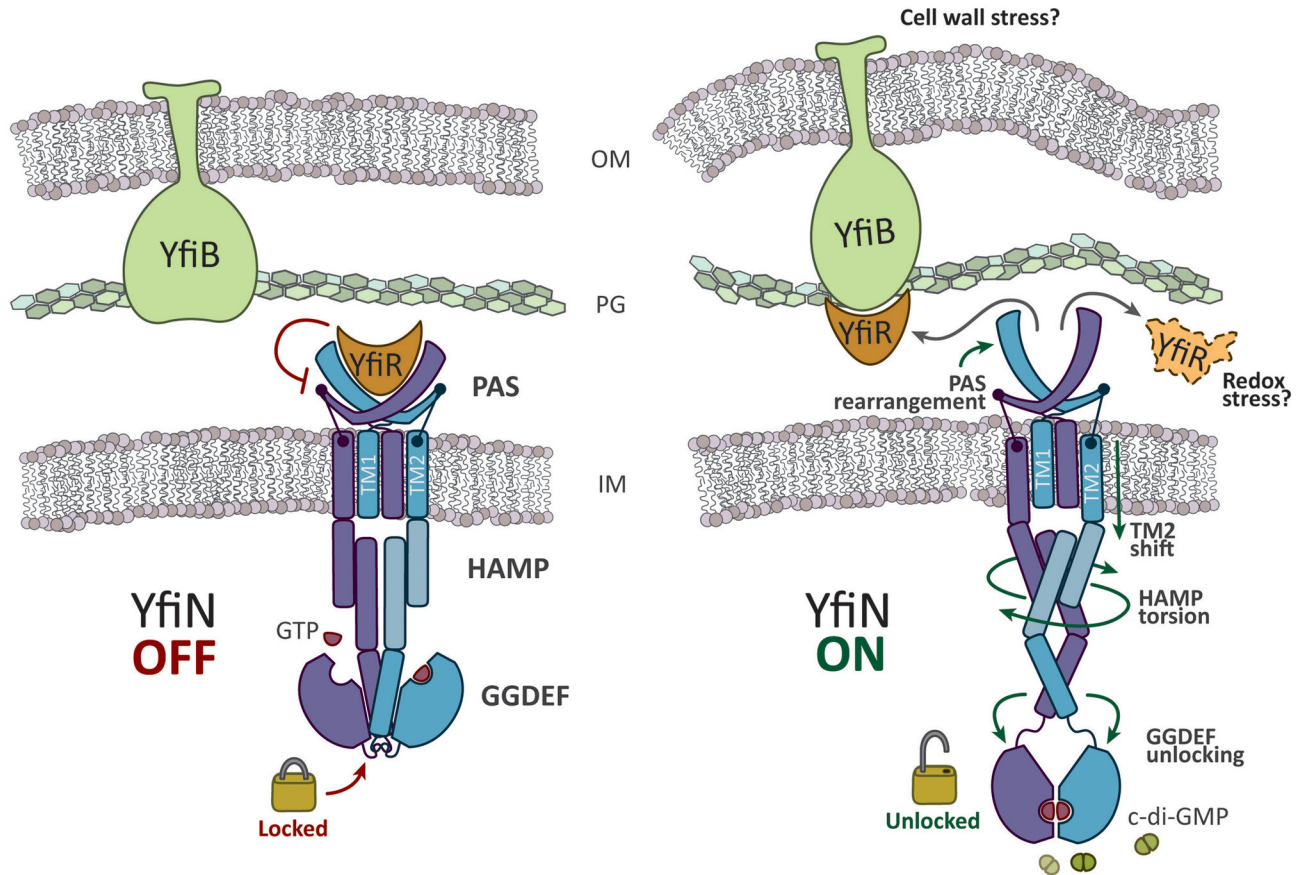
Figure 6. Scheme of allosteric regulation of YfiN. Schematic representation of the putative allosteric regulation of YfiN based on homology modeling pointing to a LapD-like allosteric communication between the periplasmic and the cytosolic portions of the enzyme that is mediated by a conformational change of the HAMP domain.
doi: 10.1371/journal.pone.0081324.g006

inhibition as other DGCs and, therefore, functions as ON/OFF cyclase responding solely to periplasmic signals.

It is becoming clear that the regulation of different DGCs depends firmly on the architecture of the accessory domains of each enzyme. Therefore, targeting the allosteric modules (e.g. the regulatory domains) together with of the catalytic domain could become a winning strategy to fight biofilm-mediated infections. This is especially true in the case of the YfiBNR system, which functions as an entry point for different environmental signals during *Pseudomonas* adaptation. Of course, availability of structural data represents the bottleneck for an efficient drug design approach: understanding the structural details of the allosteric control of DGC activity is highly desirable yet challenging. By assuming a LapD-like fold for YfiN periplasmic portion, we could speculate that its allosteric regulation is similar to the P. *fluorescence* receptor [24]. Normal modes and sequence conservation analyses, as well as mapping of the activating/inactivating mutations on the homology model are in agreement with a LapD-like activating mechanism, solely depending on the interaction between YfiR and YfiN in the periplasmic space. Based on our biochemical

data on the truncated constructs, indicating that the presence of the HAMP domain is essential to induce the transient dimerization of the monomeric YfiN$_{HAMP-GGDEF}$, we suggest that the periplasmic domain of the full-length protein, by assuming a LapD-like fold that is based on domain-swapping, could function as the driving force for dimerization. A key role in the conformational transition appears to be played by the region connecting the HAMP to the GGDEF domain. We propose that this linker loop may act as a hinge whose locking/unlocking equilibrium, driven by the conformation of the HAMP domain helices, controls the catalysis by keeping the two GGDEF domains separated or allowing their facing (Figure 6). Catalysis through transient encountering of the GGDEF domains could be a general feature of DGCs, which have evolved different regulatory modules that inhibit catalysis always by spatially separating the two GGDEF domains [27,29]. On the other hand, the GGDEF domains are dynamically exploring their allowed conformational space looking for each other like lovers do, waiting for activation and substrate to come and let them finally meet.
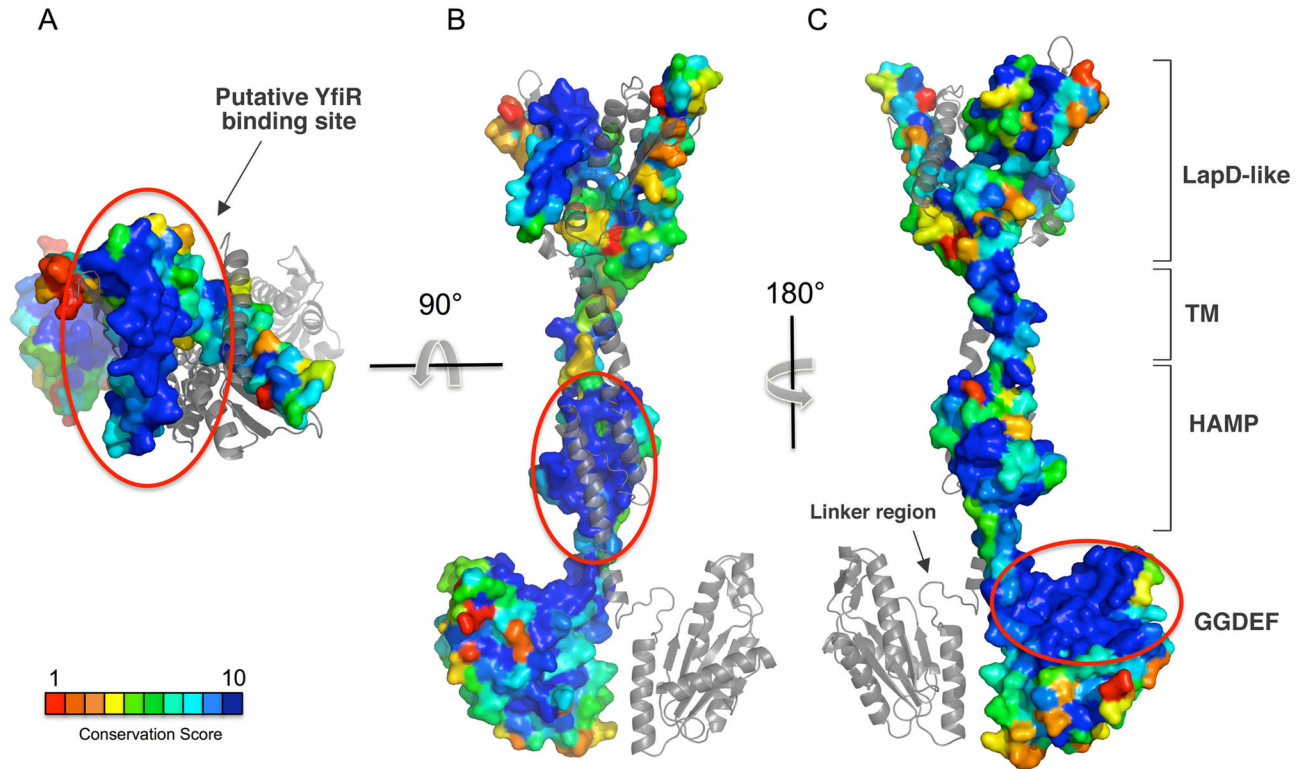
Figure 7. Mapping sequence conservation on YfiN model. Location of strictly conserved regions (grading from cyan to blue) mapped on the model of YfiN. **A**) The central V-shaped gorge of the periplasmic domain is fully conserved. Since this region is solvent exposed a similar conservation degree suggests that this is the putative binding site of YfiR. **B**) The core of the four-helices bundle of the HAMP domain is conserved, as expected. **C**) The most conserved region of the GGDEF domain comprises the region of the active site (highlighted in red) and the linker region, the small loop connecting the catalytic and the HAMP domains. The conformation of the linker region, as modeled on the structure of WspR [29]), would not allow the two GGDEF domain to assume catalytically competent conformation (i.e. with the two active sites facing each other). Therefore a severe rearrangement of the linker region (unlocking) must be assumed in order for catalysis to occur.

doi: 10.1371/journal.pone.0081324.g007

Materials and Methods

Protein cloning, expression and purification

Both the $YfiN_{HAMP\text{-}GGDEF}$ and $YfiN_{GGDEF}$ fragments were amplified from a pET24b plasmid harboring a synthetic $YfiN_{fl}$ gene (Geneart). The purified PCR products, verified by sequencing, were ligated (NdeI, XhoI) in frame with a C-terminal His-tag into a pET24b vector (Novagen) and transformed into BL21-(DE3) *E. coli* strain for expression.

Both construct were expressed as described in [14]. Briefly: cells from a single colony were used to inoculate 5 mL of Luria-Bertani (LB) medium containing 30 μg/mL of kanamycin and grown at 37° C. After 10 h cells were diluted into 300 mL of LB and grown at 37° C over night before final dilution in 3x1 L of LB. Cells were grown for 2.5 h at 37° C before induction with 100 μM isopropyl β-D-1-thiogalactopyranoside (IPTG). After 2.5 h at 30° C cells were harvested by centrifugation and stored at -20 °C.

Cells were lysed by sonication and proteins were purified using an Ni-HiTrapTM Chelating HP column (GE Healthcare)

equilibrated with 10 mM Tris–HCl, pH 8.0, 250 mM NaCl, 10% glycerol; the proteins were eluted with 100 mM imidazole, in the same buffer. Finally, the purified proteins were loaded on an FPLC column (Superdex 75 10/300, GE Healthcare), and eluted with 10 mM Tris–HCl pH 8.0, 100 mM NaCl, 2% glycerol. Size exclusion chromatography (SEC) analysis for the shorter construct ($YfiN_{GGDEF}$; Mw = 23.5 kDa) indicated an apparent molecular mass of 28 kDa consistent with a monomeric state, while for the $YfiN_{HAMP\text{-}GGDEF}$ resulted in an ambiguous apparent molecular mass of 41 kDa, in between a monomeric (28 kDa) and a dimeric (56 kDa) form in solution. Therefore, further investigation of the aggregation state of was conducted on $YfiN_{HAMP\text{-}GGDEF}$ by analytical ultracentrifugation (AUC) (Figure S5).

Analytical Ultracentrifugation

Size distribution of $YfiN_{HAMP\text{-}GGDEF}$ in solution was assessed in sedimentation velocity experiments carried out on a Beckman XLI analytical ultracentrifuge using absorbance optics. The experiments were conducted at 35,000 rpm and 20 °C at a

protein concentration of 2 mg/mL in 250 mM NaCl, 10 mM Tris-HCl pH 8.0, 10% glycerol. Radial absorbance scans were obtained at 280 nm at a spacing of 0.003 cm with three averages in a continuous scan mode. Sedimentation coefficients were calculated using the software Sedfit [44] and were reduced to water and 20 °C ($s_{20,w}$) using standard procedures. Sednterp software (http://sednterp.unh.edu/) was used to calculate the buffer density and viscosity. The sedimentation coefficient (S) of YfiN$_{HAMP-GGDEF}$ was 2.3 for 98% of the protein, consistent with a molecular mass of 21 kDa, pointing to a monomeric state of YfiN$_{HAMP-GGDEF}$ in solution.

Crystallization - data collection and refinement

Crystallization condition for YfiN$_{HAMP-GGDEF}$ were screened using a crystallization robot (Phoenix, Art Robbins), by mixing 300 nL of 3.7 mg/mL protein solution in 0.1 M NaCl, 10 mM Tris pH 8 and 2 % glycerol with equal volumes of screen solution. No positive hit was observed during the first three month. After seven month one single hexagonal crystal was observed in the droplet corresponding to solution n.17 of Crystal-Screen2 (Hampton) containing 0.1 M Sodium Citrate dehydrate pH 5.6 and 35% v/v tert-butanol. The crystal was flash frozen in liquid nitrogen, without any cryoprotectant, and diffracted to 2.77 Å resolution (ESRF, ID 14.1). Data were processed with XDS [45]. The crystal belonged to the P6$_5$22 space group with the following unit cell constants: a=b=70.87 Å; c=107.62 Å.

The Matthews coefficient for YfiN$_{HAMP-GGDEF}$ was 1.38 Å^{3}Da^{-1} with a solvent fraction of 0.11, pointing to the assumption that only the GGDEF domain (YfiN$_{GGDEF}$) was present in the crystal lattice (Matthews coefficient for YfiN$_{GGDEF}$ was 1.93 Å^{3}Da^{-1} with a solvent fraction of 0.36). Phases were obtained by molecular replacement using the GGDEF domain of PleD (PDB ID: 2wb4) as template with Molrep [46]. Cycles of model building and refinement were routinely carried out with Coot [47] and Refmac5.6 [48], model geometry was assessed by ProCheck [49] and MolProbity [50]. Final statistics for data collection and model building are reported in Table 1. Coordinates have been deposited in the Protein Data Bank (PDB: 4iob).

ITC analysis

ITC experiments were carried out using an iTC200 microcalorimeter (MicroCal), by titrating YfiN$_{HAMP-GGDEF}$ protein sample with either GTP or c-di-GMP, and YfiN$_{GGDEF}$ with GTP. Nucleotide stock solutions were prepared in water and diluted into ITC buffer (final concentrations: 10 mM Tris pH 8, 250 mM NaCl, 1,7 % glycerol, 5 mM CaCl$_2$). Protein solution was diluted into the same buffer lacking glycerol. Titration with c-di-GMP were carried out by injecting 1.5 µL aliquots of 90 µM c-di-GMP to a 3 µM protein solution at 25° C; titration with GTP was carried out by injecting 1.5 µL aliquots of 170 µM GTP to 14 µM protein solution at 25° C. The same experiment has been repeated by incubating both GTP and protein samples with 40 µM c-di-GMP. Injection of nucleotides into buffer was also performed as control, under the same experimental conditions. If indicated, data were fitted as described in [51]. All measurements were done in duplicate and the derived thermodynamic parameters are reported in Table 2.

Real-time enzymatic essay

YfiN activity was measured by circular dichroism (CD) spectroscopy as described in [23]. In brief: c-di-GMP concentration in solution can be deduced by the specific CD signal of the intercalated c-di-GMP dimer at 282 nm. This signal is enhanced in the presence of manganese, which forms a stable complex with c-di-GMP cis-dimer that is linearly dependent on c-di-GMP concentration. The condensation reaction was started by adding 100 µM GTP (Sigma) to a 10 µM solution of YfiN$_{HAMP-GGDEF}$ or YfiN$_{GGDEF}$ in 150 mM NaCl, 20 mM Tris/HCl pH 7.5, 10 mM MgCl$_2$, 2.5 mM MnCl$_2$ and 1% glycerol. C-di-GMP formation was monitored following the CD signal at 282 nm, using a 1 cm quartz cuvette (Hellma) on a JASCO J-710 spectropolarimeter at 20° C.

Homology modeling and in silico analysis

The YfiN protein sequence from *Pseudomonas aeruginosa* was retrieved from the Uniprot database (http://www.uniprot.org; accession number: Q9I4L5). UniRef50 was used to find sequences closely related to YfiN from the Uniprot database. 123 orthologous sequences displaying a minimum percentage of sequence identity of 50% were obtained. Each sequence was then submitted to PSI-Blast (www.ncbi.nlm.nih.gov/blast; number of iterations, 3; E-Value cutoff, 0.0001 [52]), to retrieve orthologous sequences from the NR_PROT_DB database. Sequence fragments, redundancy (>95%) and too distant sequences (<35%) were then removed from the dataset. At the end of this procedure, 53 sequences were retrieved (Figure S4). The conservation of residues and motifs within the YfiN sequences was assessed through a multiple sequence alignment, using the ClustalW tool [53] at EBI (http://www.ebi.ac.uk/clustalw).

Secondary structure predictions were performed using several tools available, including DSC [54] and PHD [55], accessed through NPSA at PBIL (http://npsa-pbil.ibcp.fr/), and Psi-Pred (http://bioinf.cs.ucl.ac.uk/psipred [56]). A consensus of the predicted secondary structures was then derived for further analysis.

A fold prediction-based approach was utilized to gain some structural insights into the domain organization of YfiN and related proteins. Although three-dimensional modeling performed using such techniques is seldom accurate at the atomic level, the recognition of a correct fold, which takes advantage of the knowledge available in structural databases, is often successful. The programs Phyre2 [25] and HHPRED [26] were used to detect domain organization and to find a suitable template fold for YfiN. All the programs options were kept at default.

A three-dimensional model of YfiN (residues 11-253) was constructed using the MODELLER-8 package [57], using as structural templates the following crystal structures: the N-terminal domain of the HAMP/GGDEF/EAL protein LapD from *P. fluorescens* (residues 35-161; PDB Code: 3pjv [24]); the HAMP domain of Aerotaxis transducer AER2 (residues 182-246; PDB Code: 4i3m [39]); Sensor protein QSEC (residues 11-34; 162-184; PDB Code: 2kse [41]); diguanylate cyclase response regulator WspR (residues 247-253; PDB Code: 3i5c [29]).

A full YfiN dimeric model was built starting from the crystal structure of the cyclase domain (GGDEF – present work) and performing a backward multi-step homology modeling approach, in which each new predicted domain has been linked to the previously obtained model by following the orientation of its structural template. The structural templates were oriented as follows: 1) GGDEF domain of YfiN (residues 254-414) was initially superposed to the GGDEF domain of WspR from *Pseudomonas aeruginosa* (PDB Code: 3i5c) to predict the structure and orientation of the linker region (residues 247-253 of YfiN, corresponding to residues 170-176 of 3i5c); 2) the helical stalk motif of 3i5c (residues 157-170) was then superposed to the C-terminal helix of the HAMP domain of the aerotaxis transducer Aer2 (residues 138-156), to predict the structure and orientation of the HAMP domain of Yfin (residues 182-146); 3) the orientation of the TM helices of Sensor protein qseC (PDB Code: 2KSE) with respect to the hydrocarbon core of the lipid bilayer was derived from the OPM server [58]; the N-terminal domain of LapD (PDB Code: 3pjv) was roughly oriented perpendicular to the lipid bilayer, following the relative position of the inner cell membrane and connection to the flanking TM helices as indicated by [24]. Ten different models were built and evaluated using Prosa2003 [59]: the model displaying the lowest energy profile (Z-Score= -4.86) was taken as the representative one. The initial alignment, obtained from threading methods, was then subjected to minor changes in the attempt to increase low score-regions.

Normal mode analysis and hinge regions predictions were carried out by using the "HingeProt" server, using as cutoff distances for GNM and ANM the default values 10 Å and 18 Å, respectively [60].

Evolutionary sequence conservation was mapped onto the accessible surface of the best model by means of CAMPO [61], using the previously obtained alignment.

Supporting Information

Figure S1. Residues visible in the crystal structure of YfiN$_{GGDEF}$. The predicted HAMP and GGDEF domains are underlined in purple and orange respectively. The residues that are visible in the electron density are highlighted in green (254-414). The linker region between the HAMP and the GGDEF domains, where proteolysis conceivably occurred, is coloured in blue.
(TIFF)

Figure S2. Binding of GTP to YfiN$_{HAMP-GGDEF}$ in the presence of c-di-GMP. Representative microcalorimetric titration of 14 μM enzyme with GTP (170 μM in the syringe) in the presence of 40 μM c-di-GMP in both solutions. *Upper panel*: Raw ITC data. *Lower panel*: Integrated peak areas (black square). Fit with the one-binding-site model of ORIGIN provided by MicroCal (continuous lines) is also depicted.
(TIFF)

Figure S3. Best templates for homology modelling of full length YfiN. Sequence alignments based on secondary

structure prediction of the different domains of YfiN with the most significant structural templates according to two different fold prediction servers (Phyre2 and HHPRED).
(TIF)

Figure S4. Sequence conservation. Multiple sequence alignment of 53 non-redundant orthologous of YfiN sequences, from other *Pseudomonas* strains and from more distantly related sequences from other bacteria.
(PDF)

Figure S5. Determination of the aggregation state of YfiN$_{HAMP-GGDEF}$ and YfiN$_{HAMP-GGDEF}$ in solution. A) Size exclusion chromatography (SEC) of YfiN$_{HAMP-GGDEF}$ (*green*) and YfiN$_{GGDEF}$ (*blue*) after the affinity chromatography purification step. The proteins elutes with an apparent molecular mass of 41 kDa and 28 kDa respectively. **B**) Calibration curve obtained using the following standards: BSA 66 kDa; Carbonic Anhydrase 29 kDa; Myoglobin 18 kDa; Ribonuclease A 13.7 kDa and Aprotinin 6.5 kDa. **C**) Sedimentation velocity experiment to determine the size distribution of YfiN$_{HAMP-GGDEF}$ in solution. The sedimentation coefficient (S) was 2.3 for 98% of the protein, consistent with a molecular mass of 21 kDa, and indicating a monomeric state of YfiN$_{HAMP-GGDEF}$ in solution. **D**) The YfiN$_{HAMP-GGDEF}$, the results of the SEC analysis indicates that the two domains of the protein are mobile, thus displaying a large hydrodynamic volume. On the contrary, YfiN$_{GGDEF}$ displays an apparent molecular mass consistent with a monomer, as illustrated in the scheme.
(TIF)

Movie S1. Normal Modes Analysis on YfiN model. The animation illustrates the rigid parts of YfiN the hinges connecting them, together with the direction of the fluctuation of each residue in the slowest two modes as predicted by the server *HingeProt* [60]. Two orthogonal points of view of the predicted protein motion are shown on the left and on the right respectively.
(MOV)

Acknowledgements

Author Contributions

Conceived and designed the experiments: GG AP SR S. Franceschini FC. Performed the experiments: GG AP S. Franceschini S. Fernicola VS SR. Analyzed the data: GG AP S. Fernicola SR FC. Contributed reagents/materials/analysis tools: AP SR FC. Wrote the manuscript: GG AP SR FC.

References

1. Schirmer T, Jenal U (2009) Structural and mechanistic determinants of c-di-GMP signalling. Nat Rev Microbiol 7: 724–735. doi:10.1038/nrmicro2203. PubMed: 19756011.

2. Sondermann H, Shikuma NJ, Yildiz FH (2012) You've come a long way: c-di-GMP signaling. Curr Opin Microbiol 15: 140–146. doi:10.1016/j.mib.2011.12.008. PubMed: 22226607.

3. Chua SL, Tan SY-Y, Rybtke MT, Chen Y, Rice SA et al. (2013) Bis-(3'-5')-Cyclic Dimeric GMP Regulates Antimicrobial Peptide Resistance in Pseudomonas aeruginosa. Antimicrob Agents Chemother 57: 2066–2075. doi:10.1128/AAC.02499-12. PubMed: 23403434.

4. Römling U, Galperin MY, Gomelsky M (2013) Cyclic di-GMP: the first 25 years of a universal bacterial second messenger. Microbiol Mol Biol Rev 77: 1–52. doi:10.1128/MMBR.00043-12. PubMed: 23471616.

5. Govan JR, Deretic V (1996) Microbial pathogenesis in cystic fibrosis: mucoid Pseudomonas aeruginosa and Burkholderia cepacia. Microbiol Rev 60: 539–574. PubMed: 8840786.

6. Häussler S, Tümmler B, Weissbrodt H, Rohde M, Steinmetz I (1999) Small-colony variants of Pseudomonas aeruginosa in cystic fibrosis. Clin Infect Dis 29: 621–625. doi:10.1086/598644. PubMed: 10530458.

7. Häussler S (2003) Highly adherent small-colony variants of Pseudomonas aeruginosa in cystic fibrosis lung infection. J Med Microbiol 52: 295–301. doi:10.1099/jmm.0.05069-0. PubMed: 12676867.

8. Häussler S (2004) Biofilm formation by the small colony variant phenotype of Pseudomonas aeruginosa. Environ Microbiol 6: 546–551. doi:10.1111/j.1462-2920.2004.00618.x. PubMed: 15142242.

9. Smith EE, Buckley DG, Wu Z, Saenphimmachak C, Hoffman LR et al. (2006) Genetic adaptation by Pseudomonas aeruginosa to the airways of cystic fibrosis patients. Proc Natl Acad Sci U S A 103: 8487–8492. doi:10.1073/pnas.0602138103. PubMed: 16687478.

10. Starkey M, Hickman JH, Ma L, Zhang N, De Long S et al. (2009) Pseudomonas aeruginosa rugose small-colony variants have adaptations that likely promote persistence in the cystic fibrosis lung. J Bacteriol 191: 3492–3503. doi:10.1128/JB.00119-09. PubMed: 19329647.

11. Meissner A, Wild V, Simm R, Rohde M, Erck C et al. (2007) Pseudomonas aeruginosa cupA-encoded fimbriae expression is regulated by a GGDEF and EAL domain-dependent modulation of the intracellular level of cyclic diguanylate. Environ Microbiol 9: 2475–2485. doi:10.1111/j.1462-2920.2007.01366.x. PubMed: 17803773.

12. Rohde M (1999) Small-Colony Variants of Pseudomonas aeruginosa in Cystic Fibrosis. Clin Infect Dis 29: 621–625. doi:10.1086/598644. PubMed: 10530458.

13. Drenkard E, Ausubel FM (2002) Pseudomonas biofilm formation and antibiotic resistance are linked to phenotypic variation. Nature 416: 740–743. doi:10.1038/416740a. PubMed: 11961556.

14. Malone JG, Jaeger T, Spangler C, Ritz D, Spang A et al. (2010) YfiBNR mediates cyclic di-GMP dependent small colony variant formation and persistence in Pseudomonas aeruginosa. PLoS Pathog 6: e1000804. doi:10.1371/journal.ppat.1000804. PubMed: 20300602.

15. Reinhardt A, Köhler T, Wood P, Rohner P, Dumas J-L et al. (2007) Development and persistence of antimicrobial resistance in Pseudomonas aeruginosa: a longitudinal observation in mechanically ventilated patients. Antimicrob Agents Chemother 51: 1341–1350. doi:10.1128/AAC.01278-06. PubMed: 17261619.

16. Mulcahy LR, Burns JL, Lory S, Lewis K (2010) Emergence of Pseudomonas aeruginosa strains producing high levels of persister cells in patients with cystic fibrosis. J Bacteriol 192: 6191–6199. doi:10.1128/JB.01651-09. PubMed: 20935098.

17. Kulasakara H, Lee V, Brencic A, Liberati N, Urbach J et al. (2006) Analysis of Pseudomonas aeruginosa diguanylate cyclases and phosphodiesterases reveals a role for bis-(3'-5')-cyclic-GMP in virulence. Proc Natl Acad Sci U S A 103: 2839–2844. doi:10.1073/pnas.0511090103. PubMed: 16477007.

18. Ueda A, Wood TK (2010) Tyrosine Phosphatase TpbA of Pseudomonas aeruginosa Controls Extracellular DNA via Cyclic Diguanylic Acid Concentrations. Environ Microbiol 2: 449–455. doi:10.1111/j.1758-2229.2010.00171.x. PubMed: 21552365.

19. McDonald MJ, Gehrig SM, Meintjes PL, Zhang X-X, Rainey PB (2009) Adaptive divergence in experimental populations of Pseudomonas fluorescens. IV. Genetic constraints guide evolutionary trajectories in a parallel adaptive radiation. Genetics 183: 1041–1053. doi:10.1534/genetics.109.107110. PubMed: 19704015.

20. Malone JG, Jaeger T, Manfredi P, Dötsch A, Blanka A et al. (2012) The YfiBNR Signal Transduction Mechanism Reveals Novel Targets for the Evolution of Persistent Pseudomonas aeruginosa in Cystic Fibrosis Airways. PLoS Pathog 8: e1002760. doi:10.1371/journal.ppat.1002760. PubMed: 22719254.

21. Irie Y, Borlee BR, O'Connor JR, Hill PJ, Harwood CS et al. (2012) Self-produced exopolysaccharide is a signal that stimulates biofilm formation in Pseudomonas aeruginosa. Proc Natl Acad Sci U S A 109: 20632–20636. doi:10.1073/pnas.1217993109. PubMed: 23175784.

22. Giardina G, Castiglione N, Caruso M, Cutruzzolà F, Rinaldo S (2011) The Pseudomonas aeruginosa DNR transcription factor: light and shade of nitric oxide-sensing mechanisms. Biochem Soc Trans 39: 294–298. doi:10.1042/BST0390294. PubMed: 21265791.

23. Stelitano V, Brandt A, Fernicola S, Franceschini S, Giardina G et al. (2013) Probing the activity of diguanylate cyclases and c-di-GMP phosphodiesterases in real-time by CD spectroscopy. Nucleic Acids Res 41: e79. doi:10.1093/nar/gkt028. PubMed: 23358823.

24. Navarro MV, Newell PD, Krasteva PV, Chatterjee D, Madden DR, et al. (2011) Structural basis for c-di-GMP-mediated inside-out signaling controlling periplasmic proteolysis. PLoS Biol 9: e1000588. doi:10.1371/journal.pbio.1000588.

25. Kelley La, Sternberg MJE (2009) Protein structure prediction on the Web: a case study using the Phyre server. Nat Protoc 4: 363–371. doi:10.1038/nprot.2009.2. PubMed: 19247286.

26. Hildebrand A, Remmert M, Biegert A, Söding J (2009) Fast and accurate automatic structure prediction with HHpred. Proteins 77 Suppl 9: 128–132. doi:10.1002/prot.22499. PubMed: 19626712.

27. Chan C, Paul R, Samoray D, Amiot NC, Giese B et al. (2004) Structural basis of activity and allosteric control of diguanylate cyclase. Proc Natl Acad Sci U S A 101: 17084–17089. doi:10.1073/pnas.0406134101. PubMed: 15569936.

28. Wassmann P, Chan C, Paul R, Beck A, Heerklotz H et al. (2007) Structure of BeF3- -modified response regulator PleD: implications for diguanylate cyclase activation, catalysis, and feedback inhibition. Structure 15: 915–927. doi:10.1016/j.str.2007.06.016. PubMed: 17697997.

29. De N, Navarro MVAS, Raghavan RV, Sondermann H (2009) Determinants for the activation and autoinhibition of the diguanylate cyclase response regulator WspR. J Mol Biol 393: 619–633. doi:10.1016/j.jmb.2009.08.030. PubMed: 19695263.

30. De N, Pirruccello M, Krasteva PV, Bae N, Raghavan RV (2008) Phosphorylation-Independent Regulation of the Diguanylate Cyclase WspR. PLoS Biol 6: 601–617. doi:10.1371/journal.pbio.0060067. PubMed: 18366254.

31. Yang C-Y, Chin K-H, Chuah ML-C, Liang Z-X, Wang AH-J et al. (2011) The structure and inhibition of a GGDEF diguanylate cyclase complexed with (c-di-GMP) 2 at the active site. Acta Crystallogr D Biol Crystallogr 67: 997–1008. doi:10.1107/S090744491104039X.

32. Vorobiev SM, Neely H, Yu B, Seetharaman J, Xiao R et al. (2012) Crystal structure of a catalytically active GG(D/E)EF diguanylate cyclase domain from Marinobacter aquaeolei with bound c-di-GMP product. J Struct Funct Genomics 13: 177–183. doi:10.1007/s10969-012-9136-4. PubMed: 22843345.

33. Whitney JC, Colvin KM, Marmont LS, Robinson H, Parsek MR et al. (2012) Structure of the cytoplasmic region of PelD, a degenerate diguanylate cyclase receptor that regulates exopolysaccharide production in Pseudomonas aeruginosa. J Biol Chem, 287: 23582–93. doi:10.1074/jbc.M112.375378. PubMed: 22605337.

34. Ko J, Ryu K-S, Kim H, Shin J-S, Lee J-O et al. (2010) Structure of PP4397 reveals the molecular basis for different c-di-GMP binding modes by Pilz domain proteins. J Mol Biol 398: 97–110. doi:10.1016/j.jmb.2010.03.007. PubMed: 20226196.

35. Rooman M, Liévin J, Buisine E, Wintjens R (2002) Cation–π/H-bond Stair Motifs at Protein–DNA. Interfaces - J Mol Biol 319: 67–76. doi:10.1016/S0022-2836(02)00263-2.

36. Sathyapriya R, Vishveshwara S (2004) Interaction of DNA with clusters of amino acids in proteins. Nucleic Acids Res 32: 4109–4118. doi:10.1093/nar/gkh733. PubMed: 15302912.

37. Hulko M, Berndt F, Gruber M, Linder JU, Truffault V et al. (2006) The HAMP domain structure implies helix rotation in transmembrane signaling. Cell 126: 929–940. doi:10.1016/j.cell.2006.06.058. PubMed: 16959572.

38. Newell PD, Monds RD, Toole GAO (2009) LapD is a bis-(3',5')-cyclic dimeric GMP-binding protein that regulates surface attachment by Pseudomonas fluorescens Pf0 – 1. Proc Natl Acad Sci U S A 106: 3461–3466. doi:10.1073/pnas.0808933106. PubMed: 19218451.

39. Airola MV, Sukomon N, Samanta D, Borbat PP, Freed JH et al. (2013) HAMP Domain Conformers That Propagate Opposite Signals in Bacterial Chemoreceptors. PLoS Biol 11: e1001479. doi:10.1371/journal.pbio.1001479. PubMed: 23424282.

40. Reinelt S, Hofmann E, Gerharz T, Bott M, Madden DR (2003) The structure of the periplasmic ligand-binding domain of the sensor kinase CitA reveals the first extracellular PAS domain. J Biol Chem 278: 39189–39196. doi:10.1074/jbc.M305864200. PubMed: 12867417.

41. Maslennikov I, Klammt C, Hwang E, Kefala G, Okamura M et al. (2010) Membrane domain structures of three classes of histidine kinase receptors by cell-free expression and rapid NMR analysis. Proc Natl Acad Sci U S A 107: 10902–10907. doi:10.1073/pnas.1001656107. PubMed: 20498088.

42. Haliloglu T, Bahar I, Erman B (1997) Gaussian Dynamics of Folded. Proteins - Phys Rev Lett 79: 3090–3093. doi:10.1103/PhysRevLett. 79.3090.

43. Atilgan AR, Durell SR, Jernigan RL, Demirel MC, Keskin O et al. (2001) Anisotropy of fluctuation dynamics of proteins with an elastic network model. Biophys J 80: 505–515. doi:10.1016/S0006-3495(01)76033-X. PubMed: 11159421.

44. Schuck P (2000) Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. Biophys J 78: 1606–1619. doi:10.1016/S0006-3495(00)76713-0. PubMed: 10692345.

45. Kabsch W (1993) Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. J Appl Crystallogr 26: 795–800. doi:10.1107/S0021889893005588.

46. Vagin A, Teplyakov A (2010) Molecular replacement with MOLREP. Acta Crystallogr D Biol Crystallogr 66: 22–25. doi:10.1107/ S0907444909042589. PubMed: 20057045.

47. Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr 60: 2126–2132. doi: 10.1107/S0907444904019158. PubMed: 15572765.

48. Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr D Biol Crystallogr 53: 240–255. doi:10.1107/ S0907444996012255. PubMed: 15299926.

49. Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Crystallography 26: 283–291. doi:10.1107/ S0021889892009944.

50. Davis IW, Leaver-Fay A, Chen VB, Block JN, Kapral GJ et al. (2007) MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. Nucleic Acids Res 35: W375–W383. doi:10.1093/nar/ gkm216. PubMed: 17452350.

51. Chi CN, Haq SR, Rinaldo S, Dogan J, Cutruzzolà F et al. (2012) Interactions outside the boundaries of the canonical binding groove of a PDZ domain influence ligand binding. Biochemistry 51: 8971–8979. doi: 10.1021/bi300792h. PubMed: 23046383.

52. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z et al. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25: 3389–3402. doi:10.1093/nar/ 25.17.3389. PubMed: 9254694.

53. Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22: 4673–4680. doi:10.1093/nar/22.22.4673. PubMed: 7984417.

54. King RD, Saqi M, Sayle R, Sternberg MJ (1997) DSC: public domain protein secondary structure predication. Comput Appl Biosci 13: 473– 474. PubMed: 9283763.

55. Rost B, Sander C, Schneider R (1994) PHD--an automatic mail server for protein secondary structure prediction. Comput Appl Biosci 10: 53– 60. PubMed: 8193956.

56. Jones DT (1999) Protein secondary structure prediction based on position-specific scoring matrices. J Mol Biol 292: 195–202. doi: 10.1006/jmbi.1999.3091. PubMed: 10493868.

57. Sali A, Potterton L, Yuan F, van Vlijmen H, Karplus M (1995) Evaluation of comparative protein modeling by MODELLER. Proteins 23: 318–326. doi:10.1002/prot.340230306. PubMed: 8710825.

58. Lomize MA, Lomize AL, Pogozheva ID, Mosberg HI (2006) OPM: orientations of proteins in membranes database. Bioinformatics 22: 623–625. doi:10.1093/bioinformatics/btk023. PubMed: 16397007.

59. Sippl MJ (1993) Recognition of errors in three-dimensional structures of proteins. Proteins 17: 355–362. doi:10.1002/prot.340170404. PubMed: 8108378.

60. Emekli U, Schneidman-Duhovny D, Wolfson HJ, Nussinov R, Haliloglu T (2008) HingeProt: automated prediction of hinges in protein structures. Proteins 70: 1219–1227. doi:10.1002/prot.21613. PubMed: 17847101.

61. Paiardini A, Bossa F, Pascarella S (2005) CAMPO, SCR_FIND and CHC_FIND: a suite of web tools for computational structural biology. Nucleic Acids Res 33: W50–W55. doi:10.1093/nar/gki416. PubMed: 15980521.