

A Multimodal Strategy Used By A Large c-di-GMP Network

Kurt M. Dahlstrom^{1,§}, Alan J. Collins^{1,§}, Georgia Doing¹, Jaclyn N Taroni², Timothy J. Gauvin¹, Casey S. Greene², Deborah A. Hogan¹ and George A. O'Toole^{1,*}

¹Department of Microbiology and Immunology, Geisel School of Medicine at Dartmouth,
Hanover, New Hampshire 03755.

²Department of Systems Pharmacology and Translational Therapeutics, Perelman School
of Medicine at the University of Pennsylvania, Philadelphia, Pennsylvania 19104.

[§]These authors contributed equally to this work.

*To whom correspondence should be addressed

Department of Microbiology and Immunology, Geisel School of Medicine at Dartmouth
Rm 202 Remsen Building, Hanover, NH 03755

E-mail: georgeo@dartmouth.edu

Tel: (603) 650-1248

Fax: (603) 650-1728

Abstract word count: 301

Manuscript word count: 8,130

1 **Abstract**

2 The *Pseudomonas fluorescens* genome encodes for 50+ proteins involved in c-di-GMP
3 signaling. Here, we demonstrate that when tested across 188 nutrients, these enzymes and
4 effectors appear capable of impacting biofilm formation. Transcriptional analysis of network
5 members across ~50 nutrient conditions indicates that altered gene expression can explain a
6 subset, but not all, of biofilm-formation responses to the nutrients. Additional organization of
7 the network is likely achieved through physical interaction, as determined via probing ~2000
8 interactions by bacterial two-hybrid assays. Our analysis revealed a multimodal regulatory
9 strategy, using combinations of ligand-mediated signals, protein-protein interaction and/or
10 transcriptional regulation to further fine-tune c-di-GMP-mediated responses. These results create
11 a profile of a large c-di-GMP network that is used to make important cellular decisions, opening
12 the door to future model building and the ability to engineer this complex circuitry in other
13 bacteria.

14

15 **Abstract Importance**

16 Cyclic diguanylate (c-di-GMP) is a key signalling molecule regulating bacterial biofilm
17 formation, and many microbes have up to dozens of proteins that make, break or bind this
18 dinucleotide. Thus, a major open question in the field is how signalling specificity is conferred
19 in this context with a soluble signalling molecule. Here, we take a systems approach, using
20 mutational analysis, transcriptional studies and bacterial two-hybrid analysis to interrogate this
21 network. We find that the network typically combines two or more modes of regulation (i.e.,
22 transcriptional control with protein-protein interaction) to generate an observed output.

23

24 **Introduction**

25 Cyclic diguanylate (c-di-GMP) is a second messenger used across a wide range of
26 bacterial species to control important life style decisions. c-di-GMP is produced by enzymes
27 called diguanylate cyclases (DGCs) with the canonical GGDEF domain, and is in turn degraded
28 by phosphodiesterases (PDEs) with the canonical EAL domain. This second messenger governs
29 cellular processes by binding to and activating effector proteins that govern such behaviors as
30 biofilm formation, motility, virulence activation, macromolecular synthesis, cell division, and
31 other processes (1-4).

32 Many bacteria have large c-di-GMP signaling networks, with dozens of DGCs, PDEs,
33 and possible effectors. Several models have been developed to help explain how large c-di-GMP
34 signaling networks may effectively communicate with specific effector proteins at the right time
35 or location in the cell. One model relies on global levels of c-di-GMP signaling to different
36 effectors based on their binding affinity. Pultz *et al.* demonstrated that two *Salmonella*
37 *Typhimurium* effector proteins with PilZ domains bound c-di-GMP with over a 40-fold
38 difference in affinity (5). This finding was generalized to *Pseudomonas aeruginosa* where over a
39 140-fold difference was found among its eight PilZ domain-containing proteins (5, 6). Consistent
40 with this “affinity model”, modulating global pools of c-di-GMP has been shown to alter
41 expression and activity of c-di-GMP-metabolizing enzymes, as well as outputs of the network
42 such as exopolysaccharide production and control of flagellar synthesis (7-11). A second model
43 put forth makes use of DGCs, PDEs, and effectors that physically interact with one another
44 allowing for local signaling. Evidence for physical interaction among c-di-GMP enzymes has
45 been found in *Yersinia pestis* (12), and examples of physical interaction being required for
46 signaling include a DGC-PDE interaction in *Escherichia coli* and a DGC-effector interaction in

47 *Pseudomonas fluorescens* (13, 14). Transcriptional control of the genes coding for c-di-GMP
48 related enzymes, including *rapA* in *P. fluorescens* and several DGCs and PDEs in *Vibrio*
49 *cholerae*, provides a third mechanism of signaling specificity (15-18). Finally, ligand binding
50 may mediate the activity state of various DGCs and PDEs as is the case in *E. coli* for the oxygen-
51 sensing DosC-DosP DGC-PDE pair, or the inactivating effect of zinc on DgcZ (19, 20). Both
52 transcriptional regulation and ligand-mediate activation could conceivable fit in either a global or
53 local signaling model. Thus, there does not appear to be any single mechanism used by bacteria
54 to assure a specific output from this complex network.

55 In order to understand how large c-di-GMP networks are organized and utilized, we
56 focused on *P. fluorescens* Pf01. The mechanism of biofilm formation in *P. fluorescens* is one of
57 the best understood of such signaling systems, from input to output (21, 22). The effector protein
58 LapD binds cytoplasmic c-di-GMP that in turns leads to an accumulation of the large adhesion
59 LapA on the cell surface via an inside-out signaling mechanism (23, 24). *P. fluorescens* contains
60 a large c-di-GMP network, with 21 GGDEF domain-containing proteins, 5 EAL-domain
61 containing proteins, and 17 dual-domain proteins containing both a GGDEF and EAL domain
62 (**Figure 1**). These dual-domain proteins may function as either DGCs, PDEs, both, or as
63 effectors that can bind c-di-GMP (23, 25-27). *P. fluorescens* also contains 6 proteins with PilZ
64 domains, a domain previously characterized as a c-di-GMP binding effector (28, 29). In a
65 previous study, a mutant library of *P. fluorescens* was assayed for a majority of these genes'
66 impact on biofilm formation on a single medium containing glycerol and tryptone. Four DGCs
67 and five putative PDEs appeared to impact biofilm formation under this minimal medium
68 condition, leaving most enzymes without an apparent function (30). This conundrum of "extra
69 enzymes" is not unique to *P. fluorescens*. Similar results have been obtained in *V. cholerae*, *P.*

70 *aeruginosa* and *E. coli* (31-33). This observation leads to two important questions. First, can
71 these other enzymes provide fine control over biofilm formation under non-laboratory
72 conditions? And second, in what ways can such a large network based around a diffusible
73 molecule be organized to provide order to the signaling process?

74 Here, we attempt resolve these questions using *P. fluorescens* Pf01 as a model system for
75 a large c-di-GMP signaling network. We find that when tested under a broad spectrum of
76 conditions, a large majority of the c-di-GMP-related enzymes and effectors of *P. fluorescens* can
77 impact biofilm formation. Further, we find that transcriptional regulation, non-transcriptional
78 responses to putative ligands and protein-protein interaction are common features among
79 network members. Furthermore, these modes of regulation can be combined, for example, with
80 some pairs of proteins interacting while also responding to common extracellular cues. These
81 findings provide a roadmap for approaching c-di-GMP signaling in bacteria of clinical and
82 environmental relevance, with particular emphasis on understanding how multimodal regulatory
83 strategies modulate biofilm formation in the context of a complex network.

84

85

86 **Results**

87 **A majority of proteins with predicted roles as c-di-GMP metabolizing enzymes or effectors**

88 **participate in biofilm formation.** A previous study from our group constructed a mutant library

89 of a majority of DGCs and PDEs in *P. fluorescens* and tested these mutants for biofilm formation

90 on our standard laboratory minimal medium with glycerol and tryptone (30). Four DGCs –

91 GcbA, GcbB, GcbC, and WspR – were found to positively contribute to biofilm formation under

92 these conditions. Mutations in the genes coding for five dual domain proteins – Pfl01_0192,

93 Pfl01_1887, Pfl01_2709, Pfl01_4086, and Pfl01_4876 – were found to increase biofilm

94 formation under these conditions, indicating that they likely act as PDEs. Given the large array

95 of remaining GGDEF and EAL domain-containing proteins that demonstrated no phenotype

96 under this single growth condition, we hypothesized that some number of these c-di-GMP

97 metabolizing enzymes and/or effectors may also contribute to biofilm formation under

98 conditions not previously tested. Further, in *P. fluorescens* most c-di-GMP related enzymes are

99 predicted to be fused to a variety of ligand-binding domains, including CACHE, CHASE, and

100 GAF domains (**Figure 1**). It therefore seemed possible that the c-di-GMP metabolizing enzymes

101 could be activated by growing *P. fluorescens* under different conditions. To this end, we tested

102 the WT and the 50 strains each carrying a mutation in an individual c-di-GMP-related gene

103 (**Table S1**) using Biolog plates (Biolog, Inc) to test 188 different nutrients (**Table S2**) for their

104 impact on biofilm formation. While we refer to the compounds in each of the wells of the Biolog

105 plates as “nutrients” because they are all organic compounds, these compounds could

106 conceivably serve as carbon sources, input signals, both or neither. The term “nutrients” is used

107 here as a convenient shorthand to generally describe the chemical compounds in the Biolog

108 plates.

109 The minimal medium typically used to grow *P. fluorescens* biofilms for our studies,
110 referred to as K10-T1 (K10), is a high phosphate medium containing tryptone and glycerol as
111 carbon and energy sources (30). For the purposes of testing biofilm formation in the Biolog
112 plates, we used a similar medium lacking the tryptone with its diverse array of carbon sources.
113 Because the nutrient sources tested here may not be metabolized by *P. fluorescens*, yet may still
114 impact the activity of the c-di-GMP network, glycerol was retained in this medium to permit
115 growth. We refer to this glycerol-containing medium as “base minimal medium” (BMM, see
116 Materials and Methods for details).

117 In-frame deletions of genes encoding the predicted DGCs, PDEs, and effectors as well as
118 the six genes predicted to encode PilZ domain-containing proteins and a FleQ homologue were
119 constructed. The remainder of tested strains originate from a previously reported (30) single-
120 crossover mutant library (**Table S1**). Each strain was grown overnight in LB with aeration,
121 normalized to OD₆₀₀ of ~0.716, and added to the BMM with each carbon source found in the
122 Biolog plates before allowing biofilm formation to commence. The nutrients that were toxic to *P.*
123 *fluorescens*, as indicated by lack of growth of the wild-type strain (not shown), were eliminated
124 from the analysis. Of the remaining nutrients, it was clear that different nutrients have
125 differential effects on biofilm formation (**Figure S1**).

126 Of the 50 mutants tested, 44 mutants demonstrated a significant difference ($p > 0.001$)
127 from wild type when comparing their median ability to promote biofilm formation across all
128 nutrients tested (**Figure 2A**). Among strains lacking the GGDEF-containing proteins, 7 showed a
129 significant reduction in biofilm formation as might canonically be expected for this class of
130 mutants. Surprisingly, 11 of these mutants with mutations in DGC-encoding genes showed a
131 significant increase in biofilm formation, suggesting that some putative DGCs may be able to

132 impact biofilm formation contrary to the manner canonically associated with their enzymatic
133 activity. Another 3 mutants in DGC-encoding genes showed no significant impact. Results from
134 strains lacking EAL-containing proteins were more straightforward, with 4 such mutants
135 showing elevated levels of biofilm formation, and a single strain that did not significantly differ
136 from wild type.

137 Analysis of strains lacking the dual-domain proteins revealed a range of results. 10 of the
138 mutant strains demonstrated elevated levels of biofilm formation, three showed a decrease, and
139 four had no significant impact (**Figure 2A**). While prediction of catalytic activity from sequence
140 alone can be unreliable, it is noteworthy that the bulk of mutants lacking dual-domain proteins
141 show elevated biofilm formation (**Figure 2A**). Given that *P. fluorescens* harbors 21 GGDEF
142 containing-proteins while having only five EAL domain-containing proteins, it is perhaps not
143 surprising that most of the dual domain proteins participate in c-di-GMP turnover.

144 Of the putative effectors, the PilZ domain-containing mutants of Pfl01_0958,
145 Pfl01_4008, and Pfl01_4884 showed significant increases across all carbon sources, mutants of
146 Pfl01_3860 and Pfl01_4257 showed no significant variation, and Pfl01_4150 mutant showed a
147 modest but significant decrease across all carbon sources compared to the wild type. When their
148 encoding genes were mutated, two of the effectors, the FleQ homolog Pfl01_1532 and LapD,
149 both showed significant loss of biofilm formation across a majority of the conditions tested,
150 demonstrating their general relevance to promoting biofilm formation.

151 Interestingly, 11 mutants that showed no aggregate change in biofilm formation across all
152 conditions. These include the GGDEF domain-containing Pfl01_0692, Pfl01_2176, Pfl01_2295,
153 Pfl01_4451 and Pfl01_5168; the EAL domain-containing Pfl01_2920; the dual domain-
154 containing Pfl01_1887, Pfl01_2709, Pfl01_5518, and Pfl01_5643; and the PilZ domain-

155 containing Pfl01_4884. However, while there may be no global effect detected, there are
156 individual nutrients that appear to impact biofilm formation for these mutants. For example, the
157 Pfl01_1887 mutant showed no significant difference from wild-type in this overall comparison,
158 despite being previously identified as a critical contributor to biofilm formation in this organism
159 in the minimal K10 medium where its mutation results in elevated biofilm formation (30).
160 Taken together, our data suggest that while most genes encoding components of the c-di-GMP
161 network when mutated show a biofilm phenotype under at least some small set of nutrients, other
162 genes appear to play a much broader role in biofilm formation across many environments.

163

164 **Biolog data reveals heterogeneity in mutant-by-mutant response to nutrients.** To assess
165 trends across nutrient conditions and potential patterns across genes, all values for each genotype
166 were expressed as a ratio relative to that genotype's median biofilm formation across all
167 conditions. This treatment of the data converts each reading to a measure of how the biofilm in
168 that condition differed from a baseline biofilm formation for that strain. Each of these ratios was
169 then compared to that of WT, and any mutants that show a response to a particular condition as
170 evidenced by a difference in their ratio to baseline by a statistically significant amount ($p < 0.05$,
171 See Materials and Methods for details) are highlighted in **Figure S2** (red: higher ratio than WT;
172 blue: lower ratio than WT; black: not significantly different from WT). We therefore are able to
173 view data points from mutants which differ from how the wild-type behaved in the same carbon
174 source, while downplaying the global effect a mutation might have across all carbon sources
175 from driving the outcome of the analysis (see **Figure 2A**). Squares that are black therefore do not
176 necessarily represent nutrients conditions with no phenotype for that mutant, but represent
177 instances where the mutant did not respond significantly differently from WT to a given

178 condition. This analysis thus provided us an opportunity to examine how mutants behaved in
179 specific nutrients irrespective of their general phenotype.

180 An immediately apparent finding is the heterogeneity of phenotype that a given mutant
181 can have in particular carbon sources. No class of mutations in GGDEF-, EAL-, Dual- or PilZ-
182 encoding genes was exempt from producing higher and lower biofilm than expected compared to
183 wild type under particular nutrient conditions (**Figure S2**). Further, no class of nutrients tended
184 to generally promote significantly different biofilm levels. The lack of association between
185 particular kinds of nutrients and the phenotypes of particular classes of mutants may indicate that
186 different enzymes in the network respond to different input conditions.

187

188 **Differential impact of nutrients as a function of the genes in the c-di-GMP network.** To
189 display the relative impact of the different nutrients tested here on biofilm formation by strains
190 carrying mutations in the network, we plotted the percentage of conditions with a biofilm
191 phenotype (decreased or increased) as a function of each mutant (**Figure 2B**). These data show
192 that some mutants display a biofilm phenotype in a large number of the conditions tested. For
193 example, the *lapD* mutant shows significantly reduced biofilm formation in >20% of the
194 conditions tested, confirming its broad importance in biofilm formation. In contrast, other
195 mutants showed changes in biofilm formation in <5% of the conditions tested, with the strain
196 carrying a mutation in Pfl01_5643 showing a biofilm phenotype in only 1 of the 188 nutrients
197 tested. Interestingly, the mutants with increased biofilm formation are not uniformly disrupting
198 genes predicted to code for PDEs, and thus predicted to result in increased c-di-GMP and
199 enhanced biofilm formation. Similarly, the classes of proteins represented on the right side of
200 the figure also include DGCs (**Figure 2B**). These data speak to the differential impact of

201 mutating different components of the network across the possible environments this microbe
202 may encounter.

203
204 **c-di-GMP-related genes are transcriptionally regulated in response to nutrients.** Given that
205 different DGCs, PDEs and effectors appeared to play a role in biofilm formation depending on
206 the conditions tested, we hypothesized that transcriptional regulation may be partly responsible
207 for determining which enzymes are present to contribute to c-di-GMP signaling. While very few
208 transcriptional studies of DGCs/PDEs/effectors have previously been conducted, there is some
209 precedence for transcriptional regulation. For example, transcription of the *rapA* gene, which
210 codes for a dual domain protein of *P. fluorescens* shown to have PDE activity, has been
211 previously demonstrated to be up-regulated when cells are grown in a low phosphate medium
212 (15).

213 Our goal was to capture possible roles for transcriptional regulation that may be
214 responsible for phenotypes observed in the Biolog plates. To this end, we ranked carbon sources
215 based on their ability to promote biofilm across all tested strains, and selected 14 “low” and
216 “medium” biofilm-promoting nutrients, 13 “high” biofilm-promoting carbon sources, and 18
217 “very high” biofilm-promoting carbon sources (**Table S3, Figure S1**). To assess transcription of
218 all the genes encoding DGCs, PDE, dual domain and PilZ proteins, cells were grown on the
219 BMM with the selected nutrients at a concentration of 2 mM unless, otherwise indicated (see
220 **Table S3**), for six hours before RNA was extracted. Expression analysis was conducted using the
221 Nanostring nCounter system, which directly measures transcript abundance without an
222 amplification step, and raw data were expressed as reads per 1000 counts (see Materials and
223 Methods, **Table S4 & S5**). This normalization strategy has been successfully employed

224 previously in *P. aeruginosa* (34), and we have adopted it here for the purposes of comparing
225 transcripts across many conditions.

226 To assess the impact of nutrients on expression of the genes in the c-di-GMP network, we
227 plotted the ratio of the transcript reads produced by each genes from each of the 46 nutrients
228 tested over the reads obtained in the BMM (e.g., ratio = reads on nutrient X/reads on base
229 medium). **Figure 3A** shows the individual data points with significant changes in expression as
230 determined indicated by the blue dots (see Materials and Method as for statistical analysis,
231 **Figure S3** for the data summarized as a heatmap, and **Figure S4A** for the figure with a legend
232 showing the nutrients). **Figure 3A** and **Figure S4A** shows that for most genes and for most
233 nutrients, the change in gene expression is <2-fold. However, there are several instances where
234 specific nutrients increase or decrease gene expression from 2-8-fold. We found that the gene
235 showing the highest fold-change encoded RapA, a verified PDE. The expression of *rapA* was
236 previously shown to be strongly up-regulated under low phosphate conditions (15), thus serving
237 to validate our transcriptional analysis. These data indicate that the degree of transcriptional
238 response to any particular nutrient, when such a response is detected, is quite modest (~2-fold).
239

240 **A few nutrients drive the largest changes in expression of genes in the network.** In a large
241 network of many related c-di-GMP proteins, one may consider two general methods of gene
242 regulation. The first is to have one or a small number of c-di-GMP-related genes dramatically
243 up- or down-regulated to a given condition, such as the response of *rapA* in low phosphate. An
244 alternative strategy would be to finely regulate a larger number of genes in the network at the
245 same time to produce a desired c-di-GMP output. To test this idea, we examined which nutrients
246 cause the greatest fold-increase and decrease in each gene compared to the BMM. If groups of

247 genes are not regulated together under the same conditions, we would expect that the nutrients
248 causing the largest changes compared to the base medium to be different for each gene.
249 Conversely, if genes are regulated together, we would expect the same carbon sources to be
250 responsible for the largest fold changes in several genes across the network.

251 We found that six nutrients – m-tartaric acid, glycogen, D-ribose, D-galactose, L-proline,
252 and the low phosphate condition – are responsible for the highest recorded expression of 26
253 genes in the network (**Table S6** – bolded nutrients). Intriguingly, this same group of nutrients is
254 responsible for the lowest relative expression of an overlapping group of 29 genes. Fold-change
255 compared to the BMM among these genes varied greatly, from near baseline to up to five-fold.
256 Furthermore, m-tartaric acid and glycogen were found to be among the highest and lowest
257 biofilm-promoting nutrients we tested (**Table S3**). Notably, these two nutrients were responsible
258 for a majority of the peak high and low transcriptional values, impacting 18 and 10 genes,
259 respectively (**Table S6**). Overall, 12% of the tested nutrients appear to create the largest
260 transcriptional changes in half of the genes in the network, suggesting that at least for some
261 nutrients that do impact transcription, they may do so across many genes simultaneously.

262

263 **Growth on a surface minimally impacts expression of genes in the c-di-GMP network.**

264 Previous studies in pseudomonads showed that growth on a surface increases c-di-GMP levels
265 compared to growth as planktonic cells (35-37). Thus, we compared the expression of all genes
266 in the c-di-GMP network for a select set of nutrients in liquid medium versus the same medium
267 solidified with 1.5% agar. Specifically, we investigated gene expression on K10T-1, K10T-1
268 with low phosphate, BMM, as well as the BMM supplemented with citrate, pyruvate, and L-
269 methionine (**Table S7**). The data were plotted as a ratio of the high c-di-GMP to the low c-di-

270 GMP condition on a per-nutrient basis and analyzed by a Wilcoxon signed rank test with a
271 multiple comparisons correction (**Figure S4B**). Only three genes were found to have a
272 significantly different expression on a surface compared to a liquid (**Figure S4B**). The *gcbA*
273 gene, which has previously been described as being important in early surface attachment (38)
274 was significantly more highly expressed when cells were grown in liquid, while the genes
275 Pfl01_0050 and Pfl01_1336 were significantly more highly expressed when cells were grown on
276 a solid. The *gcbA* and Pfl01_1336 genes were both differentially regulated by ~2-4-fold, while
277 Pfl01_0050 exhibited less than a 2-fold change. Based on the modest dynamic range we
278 observed for most genes in the c-di-GMP network when exposed to different chemical
279 compounds (**Figure 3A**), it may possible that these differences in expression could be sufficient
280 to impact the role of these genes in the response of cells to a surface.

281
282 **Modulation of gene expression in the network by c-di-GMP.** In *P. aeruginosa*, increased
283 levels of c-di-GMP positively regulate the expression of exopolysaccharide production and
284 down-regulate expression of flagellar genes (8, 10, 39). Furthermore, previous reports have
285 shown c-di-GMP can regulate transcriptional levels of genes through riboswitches, although
286 there have been no reports of such structures in *P. fluorescens* (7, 40). Thus, we assessed the
287 impact of modulating c-di-GMP levels on expression of genes in the network by selecting two
288 strains of *P. fluorescens* that produce low and high levels of c-di-GMP, respectively. The
289 previously reported $\Delta 4$ DGC mutant lacks the GcbA, GcbB, GcbC, and WspR DGCs, fails to
290 make a biofilm in K10 medium, and produces 2-fold lower c-di-GMP than the wild-type strain
291 (41). Conversely, the GcbC R366E mutant contains a point mutation in the auto-inhibitory site of

292 this DGC, producing over 10-fold more c-di-GMP than the Δ 4DGC strain (41). Three biological
293 replicates of each strain were grown on BMM and analyzed via Nanostring (**Table S8**).

294 Most genes showed changes of <2-fold (**Table S8**), indicating that this large c-di-GMP
295 network does not broadly rely on a core global level of intracellular c-di-GMP to regulate the
296 transcription of other members of the network, although there is a small set of genes whose
297 expression is significant impacted by c-di-GMP levels (**Table S8**, in bold). The gene Pfl01_4451
298 (the homolog of the *P. aeruginosa sadC* gene) was up-regulated 2.4-fold in the high c-di-GMP
299 mutant compared to wild-type. This result was significant ($p < 0.01$, t-test with Benjamini-
300 Hochberg correction comparing the low and high c-di-GMP producing strains). The *sadC* gene is
301 known to be important in early surface responses. Additionally, Pfl01_0190, Pfl01_5255, and
302 Pfl01_5518 also show significant and greater than 2-fold differences in expression, between the
303 low and high c-di-GMP strains tested here, indicating that transcription of only a small portion of
304 the network is sensitive to large swings in intracellular c-di-GMP concentrations.

305
306 **Broad trends observed in the transcriptional network of *P. fluorescens* are also seen in *P.***
307 ***aeruginosa*.** Our analysis of the expression of c-di-GMP-related genes (CDG genes) in *P.*
308 *fluorescens* showed that, with the exception of *rapA* in low phosphate, these genes do not show
309 large variations across the conditions tested. To address whether a similar pattern of expression
310 of CDG genes is found in another related organism, we analyzed the large number of microarray
311 experiments that have been conducted in *P. aeruginosa*. Using the tools developed by Tan *et al.*
312 (42) we assembled a compendium of all published microarrays in *P. aeruginosa*. This
313 compendium includes data for 5549 genes in 1185 experiments including 78 different growth
314 media.

315 First, to assess whether the broad characteristics of expression of CDG genes in *P.*
316 *fluorescens* Pf01 are also seen in *P. aeruginosa* we compared the overall expression of CDG
317 genes in both organisms. To do this analysis, we normalised the expression of genes between
318 experiments in the compendium by converting all expression data to values between 0 and 1
319 according to how they compare to the highest and lowest reported value for each experiment,
320 with the highest value becoming 1 and the lowest value becoming 0. We compared summary
321 statistics of these normalized data to our counts per 1000 normalized Nanostring data (see **Table**
322 **S5**). We calculated the median expression for CDG genes in both datasets, as well as the median
323 of non-CDG genes, which we defined as genes coding for proteins without a GGDEF, EAL, HD-
324 GYP, or PilZ domain. We also included the Pho-regulon in the *P. aeruginosa* PAO1
325 compendium for a comparison with a group of genes that are under robust transcriptional
326 regulation. The pattern of expression observed for *P. fluorescens* Pf01, wherein most genes
327 encoding cdG-metabolizing enzymes show relatively low levels of expression while PilZ-
328 domain-encoding genes are more highly expressed (**Figure S5A**), is also observed for *P.*
329 *aeruginosa* PAO1 (**Figure S5B**), although the difference in magnitude of expression of the PilZ-
330 encoding genes is lower in *P. aeruginosa* PAO1.

331 Given that the methods used to assess gene expression differ between *P. aeruginosa*
332 PAO1 (microarrays) and *P. fluorescens* (Nanostring), we sought to use a common metric to
333 compare variation in gene expression between the two microbes. We calculated the coefficient
334 of variance for each data set by dividing the standard deviation by the mean of the expression of
335 each gene across conditions tested. Variability in the expression of these genes across conditions
336 was similar between the two organisms as judges by this metric, with both organisms exhibiting
337 lower variability in CDG gene than the median of non-CDG genes (**Figure S5C,D**). Conversely,

338 the variability of genes in the Pho regulon was higher than that of most non-CDG genes (**Figure**
339 **S5D**). Together these data suggest transcriptional regulation across multiple growth conditions is
340 not likely a major factor controlling the contribution of genes involved in the c-di-GMP network
341 in either *P. fluorescens* or *P. aeruginosa*.

342 To investigate potential transcriptional co-regulation common to both organisms, we
343 applied a simple pairwise Spearman correlation analysis of all genes in the compendium and
344 Nanostring datasets. In *P. fluorescens* we found that several genes seem to correlate strongly and
345 significantly with one another (**Figure S5E**), but with fewer genes being significantly and
346 strongly correlated with one another compared to *P. aeruginosa* (**Figure S5F**). These correlation
347 analyses suggest that while the magnitude of transcriptional changes are not large, groups of
348 CDG genes may be under related transcriptional regulation in both organisms.

349
350 **Evidence for nutrient-mediated, non-transcriptional control of proteins in the network.** We
351 next explored how much of the response of the network to the nutrients tested can be explained
352 by transcriptional regulation. That is, how often does a given mutant's defect when grown on
353 "nutrient X" correlate with a change in the transcription of that gene when grown on "nutrient
354 X"? To address this question, we assessed what portion of the conditions tested in both Biolog
355 and Nanostring were associated with changes in either biofilm formation, gene expression, both
356 or these phenotypes, or neither (**Figure 3B**).

357 In most cases that was no link between a change in biofilm formation in a particular
358 mutant and a change in expression of a c-di-GMP-related gene in that condition (**Figure 3B**).
359 However, we did observe instances of biofilm phenotypes that are associated with differences in
360 gene expression (**Figure 3B**), suggesting that gene expression may explain the impact of these

361 conditions. Data from the nutrients with both biofilm impact and gene expression difference are
362 displayed in **Figure 3C**. Interestingly, m-tartaric acid impacted both the biofilm formation and
363 expression of 13 different genes. m-Tartaric acid also induces one of the highest overall biofilm
364 tested, suggesting that m-tartaric acid may be an important signal or nutrient for *P. fluorescens*
365 and exert at least some degree of its impact on biofilm formation via transcriptional control. As
366 noted above, in a majority of circumstances, when biofilm formation occurs it can be attributed
367 to the chemical compound added in the absence of any detectable transcriptional change (**Figure**
368 **3B**). Given the large number of putative sensory domains present in many of the c-di-GMP-
369 related proteins in *P. fluorescens* Pf0-1, we suggest that nutrients may be impacting the c-di-
370 GMP network by controlling enzyme activity.

371

372 **Physical interaction is common for DGC and dual-domain proteins.** While patterns of
373 biofilm formation and transcriptional regulation emerged from the differing nutrient conditions
374 tested, these findings did not appear sufficient to explain how this larger network might be
375 ordered to prevent cross talk among its many enzymes and effectors. Furthermore, physical
376 interaction as a mechanism of signaling specificity has precedence in a number of bacterial c-di-
377 GMP systems including *E. coli*, *P. aeruginosa*, and *Xanthomonas axonopodis*. (13, 43, 44). In *P.*
378 *fluorescens*, physical interaction between the DGC GcbC and the effector LapD has previously
379 been shown to be required for specificity of GcbC signaling in biofilm formation (14). However,
380 it is largely unknown how common physical interaction is among enzymes and effectors in the
381 network, whether a given protein may have a single or many interaction partners, or if a given
382 type of c-di-GMP protein is more or less likely to form physical interactions than others.

383 To address these questions, each c-di-GMP related gene in *P. fluorescens* was cloned into
384 both “bait” and “prey” vectors of a bacterial two-hybrid (B2H) system to assess the capacity of
385 each protein for physical interaction. Each member of every protein type – GGDEF, EAL, dual-
386 domain, and PilZ – was tested against all members of the other types. Additionally, all dual-
387 domain proteins were also tested against all other dual-domain proteins as it is possible that such
388 proteins may act as DGCs, PDEs, both or as effectors. The summary of this analysis, which
389 represented close to 2000 bacterial two-hybrid assays, is presented in **Figure 4, Figure S6** and
390 **Table S9**.

391 Several dual-domain proteins showed a relatively high capacity for interaction with other
392 network members (**Figure 4A**). In particular LapD, Pfl01_0192, Pfl01_1252, Pfl01_2525,
393 Pfl01_2709, Pfl01_4086, and Pfl01_4487 interacted with 7 to 18 partners each. A subset of 15
394 proteins interacted with LapD (**Figure 4B**) and four additional dual-domain proteins interacted
395 with greater than 10 partners (**Table S9 and Figure S6**). Conversely, RapA and Pfl01_1887 are
396 not a part of these larger hubs, interacting with only two and one partners, respectively (**Figure**
397 **4A and Table S9**).

398 DGCs showed a highly variable propensity to interact (**Table S9**). 14 DGCs appear to
399 participate in physical interaction, having between one and eight interaction partners.
400 Interestingly, the four DGCs previously found to be necessary for biofilm formation under
401 K10T-1 media conditions have different interaction profiles. GcbA and GcbB failed to interact
402 with other proteins in this assay, while by contrast GcbC interacts with several dual-domain
403 proteins (including LapD) and with an EAL domain-containing protein. WspR has a single
404 interaction partner - Pfl01_1092. The remaining seven DGCs do not engage in detectable
405 physical interaction under our assay conditions.

406 Finally, little evidence was found to indicate that EAL and PilZ domain-containing
407 proteins interact with other members of the network (**Figure 4A**). Among the five EAL domain-
408 containing proteins, only Pfl01_2920 was found to interact with two DGCs and LapD. Likewise,
409 a single PilZ domain-containing protein, Pfl01_3860, interacts with the GGDEF domain-
410 containing protein Pfl01_3550 and the dual-domain protein Pfl01_0192. The Pfl01_3860 PilZ
411 domain-containing protein is part of a small node made up of a GGDEF domain and two dual-
412 domain proteins (**Fig. 4A**). Together, these data indicate that a number of the proteins in the
413 network have the potential to interact, and thereby potentially form signaling complexes.

414

415 **Relationship between interacting proteins and shared biofilm phenotypes.** Having shown
416 that some, but not all impacts of different conditions on biofilm formation may be explained by
417 transcriptional regulation, we next tried to determine how much of a role physical interaction
418 might play in contributing to signaling specificity in the network. To perform this analysis, we
419 assessed the conditions wherein both members of an interacting pair of proteins also displayed a
420 significant biofilm phenotype when mutated. We scored these biofilm phenotypes based on
421 whether they were in the same direction (i.e. both enhanced biofilm or both reduced biofilm) and
422 calculated the proportion of the total biofilm phenotypes each member of the interacting pairs
423 shared (**Figure 5A**). This analysis shows that most interaction partners impact biofilm in only
424 one or two of the same conditions, however some pairs do demonstrate a large overlap. For
425 example, the interaction pairs such as *wspR*-Pfl01_0192 and Pfl01_0692-Pfl01_1887, which only
426 impact biofilm formation in a few conditions, overlap in most of those conditions. Furthermore,
427 there are several instances wherein interacting pairs, when mutated, show opposite biofilm
428 phenotypes in a particular nutrient condition. Such a finding is likely physiologically relevant

429 given that the biofilm formation is impacted under a common growth condition. It is also not
430 surprising that overlapping conditions between interaction partners are generally consistent in
431 being all in the same direction or all in the opposite directions, as mutants generally only
432 impacted biofilm formation in one direction. However, it is worth noting that interaction partners
433 seem equally likely to act on biofilm formation in the same direction as they do in opposite
434 directions. These data suggest that there may be numerous examples of both antagonistic and
435 cooperative interactions in this network, often involving the same proteins, and possibly
436 modulated via physical interaction.

437
438 **Intersection between B2H and transcriptional control.** We next examined possible overlap
439 between proteins that interacted with one another and were coordinately expressed. We scored
440 the number of conditions where genes of an interacting pair changed in expression under the
441 same condition. A number of interacting pairs showed coordinated transcriptional changes under
442 one to seven conditions (**Figure 5B, Table S10**). Interesting, some genes that only showed
443 significantly different expression from the BMM in one or a small number of conditions (see
444 **Table S5**) did so under the same nutrient additions as genes encoding for their interaction
445 partner. For example, for the Pfl01_0192-Pfl01_4451 interacting pair, Pfl01_0192 shares all
446 seven conditions that produce differential expression from the BMM with the GGDEF domain-
447 containing Pfl01_4451. Together, these results suggest that in some cases, a combination of
448 transcriptional control and physical interactions may be utilized to modulate signalling
449 specificity.

450 **Discussion**

451 In this study, we investigated the c-di-GMP network of *P. fluorescens*. We selected this
452 model organism because it has one of the best-understood c-di-GMP circuits, from input to
453 output, with the control of localization of the cell-surface biofilm adhesion LapA as a key output
454 (21, 22). LapA appears to contribute to or be required for biofilm formation under every
455 condition we have examined to date, including the base minimal medium used in this study, and
456 this same medium supplemented with the 188 nutrients investigated here. Our analysis revealed
457 one simple, central conclusion: no one mode of regulation outlined above (nutrient input,
458 transcriptional regulation, protein-protein interaction) adequately describes the function of the
459 network. Instead, these modes of regulation can function in combination to contribute to tuning
460 of the network. We propose that this multimodal strategy of controlling c-di-GMP-mediated
461 outputs ultimately provides fidelity and specificity to the network.

462 We found that a majority of c-di-GMP related genes in *P. fluorescens* impact biofilm
463 formation when mutated and tested under a large array of conditions, indicating that context is
464 key for understanding the function of these large signaling networks. Further, this context
465 appears to shape the organization of the network in several ways. We found that transcriptional
466 regulation in *P. fluorescens* rarely results in over 5-fold changes in transcription across all tested
467 conditions. It was also rare for any given nutrient and any single gene that a transcriptional
468 change to be solely responsible for an observed biofilm defect. Despite this observation, we did
469 find that a small number of nutrients were responsible for the largest (albeit modest) change in
470 expression for many genes. For example, we identified several groups of genes whose expression
471 changed in the same nutrient conditions (i.e., m-tartaric acid, D-ribose, glycogen), possibly
472 suggesting co-regulation. Given this finding, we speculate that some genes in the c-di-GMP

473 network may be regulated as “suites”, and coordinated 2- to 5-fold changes to a single nutrient
474 across 8 to 12 genes may be an effective method of governing the network by controlling which
475 proteins are present under that given environmental condition. Why would these nutrients in
476 particular regulate suites of genes? Interestingly, m-tartaric acid is produced and metabolized by
477 fluorescent pseudomonads (45) and plays a critical role in solubilizing inorganic phosphate (46);
478 phosphate is a key regulator of biofilm formation by this microbe (47). The physiological
479 relevance of D-ribose and glycogen is less clear. Finally, given that many of the members of the
480 c-di-GMP network, including those encoded by these suites of genes, also have putative ligand
481 binding domains, it is quite possible that more than one mechanism regulates the function and/or
482 specificity of these network members.

483 Physical interaction was also found to play a role in the network. In comparison to our
484 transcriptional results, we found many discrete examples where two genes, when mutated and
485 shared an altered biofilm phenotype on the same nutrient source also physically interact with one
486 another. Of the 85 pairs of interacting proteins analyzed only 6 shared no biofilm phenotypes,
487 and more than a quarter (24) of the pairs shared a biofilm phenotype in 25 of more nutrient
488 conditions. The remaining pairs shared phenotypes between 1 and 25 nutrient conditions. We
489 interpret the relatively common frequency of interacting pairs sharing biofilm phenotypes to
490 indicate that physical interaction occurs among network members that can exert their influence
491 in similar environmental conditions. Such a combination of common nutrient inputs among
492 interacting proteins provides one of the strongest findings supporting the concept of multimodal
493 regulation. In the case of those interaction pairs which show little or no overlap in their mutant’s
494 biofilm phenotype, it possible that we have not yet identified the correct conditions for which the

495 interaction pair is relevant. Alternatively, some of these enzymes and effectors may interact but
496 are alternatively activated in differing environments.

497 The potential for physical interaction, especially among the DGCs and the LapD receptor
498 is potentially informative regarding signaling specificity. We reported previously that GcbC and
499 LapD interact (14), and a recent report by Cooley, O'Donnell and Sondermann (48) proposed
500 that LapD forms a dimer-of-dimers “basket”, with space for a DGC to nestle within the basket –
501 it is possible, for example, that all DGCs that interact with LapD have the potential to form this
502 signaling complex. Furthermore, given that many DGCs have associated ligand-binding
503 domains and show relatively few examples of ligand-mediated transcriptional control, we
504 propose that interaction of DGCs with the LapD receptor, coupled with ligand-mediated control
505 of DGC activity, could be a general mechanism of signaling specificity. Finally, the propensity
506 of members of the network to interact indicates the possibility that “local” signaling is a common
507 aspect of the network, and any models describing how the network functions cannot ignore this
508 feature (49).

509 We also identified a number of putative DGCs that worked ‘against type’, whereby their
510 absence actually resulted in an increase in biofilm formation in a number of conditions at varying
511 frequencies. Previous work has shown that the DGC GcbA in *P. aeruginosa* is partly responsible
512 for modulating a PDE linked to dispersal of biofilm (50), providing some precedence that DGCs
513 do not necessarily have to promote biofilm formation. However, we note here that two DGCs
514 (i.e., Pfl01_2049 and Pfl01_2297) that interact with LapD demonstrated this analogous
515 phenotype under some conditions, leading to the possibility that some DGCs may exert their
516 impact by binding and blocking effector proteins. Additionally, strains carrying mutations in
517 some DGCs were found to either enhance or reduce biofilm formation depending on the nutrient

518 tested, suggesting the mechanism of their impact on biofilm formation of some of these enzymes
519 may be different from what was reported for the *P. aeruginosa* GcbA enzyme. It is not
520 inconceivable that physical interaction may play a role in this phenomenon as well, with DGCs
521 blocking effectors until the correct nutrient/ligand activates the DGC, whereupon catalytic
522 activity is initiated or structural rearrangements cause the enzyme to change its interaction
523 profile.

524 While we have provided insight into the how a large c-di-GMP network is organized and
525 regulated, we note that there are also several hurdles to interpreting the results presented here.
526 One alternative explanation for the high percentage of c-di-GMP related proteins impacting
527 biofilm is that they are in fact responsible for different cellular process that indirectly have an
528 impact on biofilm formation. It is also likely that the nutrients tested here do not represent the
529 complete collection of environmental cues that these enzymes and effectors respond to in terms
530 of catalytic activation and transcriptional regulation. Indeed, the largest transcriptional change
531 observed was for *rapA*, the PDE up-regulated in a low phosphate condition, suggesting that
532 inputs other than organic compounds can impact the network. Further, we note that several
533 transcripts measured in our Nanostring experiments showed very low expression under most
534 conditions. Still, we find it compelling that a small subset of nutrients produced both the highest
535 and lowest transcriptional changes of a small number of genes, indicating that genes may be
536 regulated as suites in addition to being regulated individually by highly specific input signals.
537 Finally, we note that the bacterial two-hybrid only presents evidence of which proteins interact
538 with each other (or not) in a heterologous host under one growth condition (i.e., LB). That is, it is
539 conceivable that the ligand bound state of a DGC or PDE may not only activate/deactivate
540 enzyme activity, but it may also cause structural changes making the protein competent or

541 incompetent for interaction. Additionally, there may be further interactions we were not able to
542 observe in this system, either because the protein is unstable/unable to function outside of its
543 native host, or alternatively because the interaction was not observable under our assay
544 conditions.

545 Taken together, our results show that this large c-di-GMP network is a dynamic system
546 capable of responding in specific ways to a variety of inputs with individual members able to
547 take on different functions under differing circumstances. Importantly, we identified numerous
548 instances where combined modes of regulation are observed (nutrients inputs/protein-protein
549 interactions, transcription control/protein-protein interactions, nutrient inputs/transcriptional
550 control). Such multimodal regulation would allow the integration of multiple inputs to fine-tune
551 the output(s) of c-di-GMP-regulated processes, likely enhancing specificity of response by this
552 network. Future studies will be required to examine the detailed mechanisms of ligand
553 recognition, physical interaction among member proteins, and transcriptional regulation of genes
554 in the network.

555

556 **Materials and Methods**

557

558 **Strains and Media.** *P. fluorescens* Pfl01, *E. coli* S17 and *E. coli* BTH101 are used throughout
559 this study (51, 52). *P. fluorescens* and *E. coli* BTH101 were grown at 30°C on 1.5% agar LB
560 plates unless otherwise indicated. *E. coli* S17 was grown at 37°C. Media used in this study
561 includes K10T-1 as reported previously (52), and base minimal medium (BMM). BMM is
562 composed of 50 mM Tris HCL (pH 7.4), 7.56 mM (NH₄)₂SO₄, 0.15% glycerol, 1 mM K₂HP0₄,
563 and 0.6 mM MgSO₄. *P. fluorescens* strains harboring the pMQ72 expression plasmid were
564 grown overnight with 10 µg/ml gentamicin. Expression was induced during experimentation
565 with 0.2% arabinose. In-frame deletions were constructed as previously reported (53). Ligation
566 cloning of plasmids was conducted using standard laboratory techniques. All strains and
567 plasmids used in this study are listed in **Table S1**.

568

569 **Biofilm Assay.** *P. fluorescens* strains were struck out on LB plates overnight. In the case of
570 mutants coming from a previously generated mutant library, plates were supplemented with 30
571 µg/ml tetracycline (30). Single colonies were picked the following day to grow overnight in
572 liquid LB medium, and supplemented with 15 µg/ml tetracycline when appropriate (see **Table**
573 **S1**). Cells densities were measured and normalized to OD₆₀₀ = 0.716. Cells were mixed 1.5:100
574 with the BMM. 125 µl of the inoculated BMM was pipetted up and down in the Biolog PM1 and
575 PM2A nutrient plates; 100 µl of this mixture was then transferred from the Biolog plates to a
576 standard 96-well polystyrene plate (Costar) to perform the biofilm assays. Plates were covered
577 and placed in a humidified micro-chamber at 30°C for six hours, at which time the liquid in the

578 wells was discarded and the wells were stained with 125 μ l of 0.1% crystal violet (CV) at room
579 temperature for ~20 minutes, then rinsed 2-3 times with water. Wells were allowed to dry
580 overnight, and were de-stained the following day using 150 μ l of a solution of water, methanol,
581 and acetic acid (45:45:10 ratio by volume) for 20 minutes at room temperature. 125 μ l of the
582 solubilized CV solution was transferred to a flat-bottom 96 well plate and the OD was recorded
583 at 550 nm.

584

585 **Analysis of Biofilm Data.** For all analysis of experiments assaying biofilm formation using
586 nutrients from the Biolog plates, OD₅₅₀ values were adjusted in the following ways: First, a
587 baseline staining was established for each condition by growing a strain lacking the genes *gcbA*,
588 *gcbB*, *gcbC*, and *wspR* (referred to as Δ 4DGC) and subtracting each of these values from the
589 corresponding values in each condition from all other experiments. Second, the variability of WT
590 values for each conditions were assessed and conditions where WT exhibited a coefficient of
591 variance of greater than 0.35 were excluded from further analysis to reduce errors due to noisy
592 conditions. Third, one batch of data (Batch #11, **Table S2**) was higher than all others so that was
593 batch-normalized by multiplying all data from that day by the ratio between the WT median on
594 that day and the median WT value across all other days.

595 For all analyses except that for **Figure S2**, mutant values were compared to the mean WT
596 value for each condition and the difference between these values was expressed in terms of the
597 number of standard deviations of WT data in that condition. The portion of values in a normal
598 distribution that would be expected to be this many or more standard deviations from the mean
599 was expressed as a p value. All p values were adjusted using a Benjamini-Hochberg adjustment.

600 An example calculation of the P value is as follows: WT in carbon source X has mean =
601 0.3 and standard deviation = 0.1. Mutant A has a value of 0.496. Difference between WT mean
602 and value for mutant A = $0.496 - 0.3 = 0.196$. The number of WT standard deviations this value
603 differs from WT mean = $0.196 / 0.1 = 1.96$, and the portion of normally distributed population
604 one would expect to differ from the mean by 1.96 SDs or more = 0.05. Therefore the unadjusted
605 p value for Mutant A in condition X is 0.05.

606 For **Figure S2**, OD₅₅₀ values for *P. fluorescens* Pf01 and each mutant in each condition
607 were expressed as a ratio compared to the median OD₅₅₀ value of that mutant across all
608 conditions. This approach serves as a representation of the response to that condition relative to
609 an approximation of the baseline biofilm formation. A second ratio was then calculated by
610 comparing the ratio in each condition to the median of WT ratios in that condition. Using the
611 standard deviation of WT ratios for each condition as described above, we assigned a p value to
612 each value.

613

614 **Transcriptional Expression.** 11 replicates of 5 µl from overnight LB liquid cultures of the wild
615 type and mutant strains were spotted onto 1.5% agar plates containing either K10T-1 medium, or
616 the BMM alone or containing 2 mM of the desired nutrient sources. Exceptions were made for
617 nutrients that have previously been published as having biofilm phenotypes including arginine
618 (0.4%), citric acid (0.4%), and L-arabinose (0.2%), or where exact measurements were infeasible
619 (dextrin; soluble fraction – 0.2%). Cells were grown at 30°C for six hours, scraped from the plate
620 and resuspended in TE buffer, and processed using the RNAeasy extraction kit (Qiagen). 75 ng

621 of each RNA sample was added to the Nanostring nCounter kit, and the protocol provided by the
622 manufacture was followed without modification (Nanostring).

623
624 **Transcriptional Analysis.** Raw reads from the Nanostring system were normalized as a fraction
625 on the total number of reads measured for targeted genes, expressed as reads per 1000 transcripts
626 as described previously (34). A K10T-1 sample and a base medium sample was processed on
627 each cartridge, yielding six biological replicates. The six replicates of the BMM were used to test
628 significance in the difference between BMM and BMM plus other compounds. This analysis was
629 conducted in the same way as for the biofilm data analysis described above. Briefly, each
630 reading in the BMM plus a compound was compared to the mean value in base medium alone.
631 The standard deviation of that gene in the base medium was then used to estimate the portion of
632 values one would expect to find to be as different from the mean as the value being tested or
633 more extreme. This percentage was taken to be the p value. All p values were adjusted using the
634 Benjamini-Hochberg method. Of the 62 genes analyzed, 48 showed less than 2-fold variation
635 from the lowest to highest count between six replicate experiments when grown on K10T-1
636 (**Table S11**), and another 12 showed variation between 2 and 3-fold. Two genes showed greater
637 than 3-fold change; these genes were Pfl01_1252 and *gcbC*; both of these genes were found to
638 have low counts across all experiments (**Table S4, S11**).

639
640 **Bacterial Two-Hybrid Studies and Analysis.** BTH101 cells were electroporated with 50 ng of
641 the puT18 and pKNT25 plasmids containing the construct to be tested. Cells were recovered for
642 1 hr in 1ml LB at 37°C. 50 µl was then spread on selective agar containing 50 µg/ml
643 carbenicillin, 50 µg/ml kanamycin, 40 µg/ml 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside

644 (X-gal), and 0.5 mM Isopropyl β -D-1-thiogalactopyranoside (IPTG). Plates were incubated at
645 30°C for 20 hours, and colonies were examined for a blue color. A leucine zipper protein served
646 as a positive control, and the empty plasmids acted as the negative control. Each gene of interest
647 was fused into both the plasmids, and a positive result in either orientation was recorded as an
648 interaction.

649

650

651

652

653

654

655

656

657

658

659

660

661

662 **Acknowledgements.** The authors would like to thank Colleen Harty and T. Jarrod Smith for
663 their contributions in developing use of the Biolog plates, James Rudd for consulting on
664 hierarchical clustering methods in R, and Tom Hampton for advice on the statistical analysis of
665 the data.

666 **Author Contributions.** Project was conceptualized by GAO, KMD, and CSG. Data curation,
667 validation, and visualization was conducted by KMD and AJC. Formal Analysis was conducted
668 by KMD, AJC, JNT, and GD. Funding was acquired by GOA, CSG, and DAH. Investigation
669 was conducted by KMD, AJC, and TJG. Methodology was developed by KMD. Project was
670 administered, supervised, and resources provided by GAO. Software was contributed by AJC
671 and JNT. Writing the original manuscript was conducted by KMD, and review and editing was
672 conducted by KMD, AJC, GAO, and CSG.

673 **Competing Interests.** None.

674

675 **Funding.** This work was supported by NIGMS R01GM123609-06 (G.A.O.) and grant T32-
676 GM08704 (K.M.D.)

677

678 **Literature Cited**

- 679
- 680 1. **Kulasakara H, Lee V, Brencic A, Liberati N, Urbach J, Miyata S, Lee DG, Neely AN,**
681 **Hyodo M, Hayakawa Y, Ausubel FM, Lory S.** 2006. Analysis of *Pseudomonas*
682 *aeruginosa* diguanylate cyclases and phosphodiesterases reveals a role for bis-(3'-5')-
683 cyclic-GMP in virulence. Proc Natl Acad Sci U S A **103**:2839-2844.
- 684 2. **Tischler AD, Camilli A.** 2004. Cyclic diguanylate (c-di-GMP) regulates *Vibrio cholerae*
685 biofilm formation. Mol Microbiol **53**:857-869.
- 686 3. **Paul K, Nieto V, Carlquist WC, Blair DF, Harshey RM.** 2010. The c-di-GMP binding
687 protein YcgR controls flagellar motor direction and speed to affect chemotaxis by a
688 "backstop brake" mechanism. Molecular Cell **38**:128-139.
- 689 4. **Abel S, Bucher T, Nicollier M, Hug I, Kaever V, Abel Zur Wiesch P, Jenal U.** 2013. Bi-
690 modal distribution of the second messenger c-di-GMP controls cell fate and asymmetry
691 during the *caulobacter* cell cycle. PLoS Genet **9**:e1003744.
- 692 5. **Pultz IS, Christen M, Kulasekara HD, Kennard A, Kulasekara B, Miller SI.** 2012. The
693 response threshold of *Salmonella* PilZ domain proteins is determined by their binding
694 affinities for c-di-GMP. Mol Microbiol **86**:1424-1440.
- 695 6. **Christen M, Kulasekara HD, Christen B, Kulasekara BR, Hoffman LR, Miller SI.** 2010.
696 Asymmetrical distribution of the second messenger c-di-GMP upon bacterial cell
697 division. Science **328**:1295-1297.
- 698 7. **Sudarsan N, Lee ER, Weinberg Z, Moy RH, Kim JN, Link KH, Breaker RR.** 2008.
699 Riboswitches in eubacteria sense the second messenger cyclic di-GMP. Science **321**:411-
700 413.

- 701 8. **Hickman JW, Harwood CS.** 2008. Identification of FleQ from *Pseudomonas aeruginosa* as
702 a c-di-GMP-responsive transcription factor. *Mol Microbiol* **69**:376-389.
- 703 9. **Perez-Mendoza D, Coulthurst SJ, Sanjuan J, Salmond GP.** 2011. N-Acetylglucosamine-
704 dependent biofilm formation in *Pectobacterium atrosepticum* is cryptic and activated by
705 elevated c-di-GMP levels. *Microbiology* **157**:3340-3348.
- 706 10. **Hickman JW, Tifrea DF, Harwood CS.** 2005. A chemosensory system that regulates
707 biofilm formation through modulation of cyclic diguanylate levels. *Proc Natl Acad Sci U*
708 *S A* **102**:14422-14427.
- 709 11. **Krasteva PV, Fong JC, Shikuma NJ, Beyhan S, Navarro MV, Yildiz FH, Sondermann**
710 **H.** 2010. *Vibrio cholerae* VpsT regulates matrix production and motility by directly
711 sensing cyclic di-GMP. *Science* **327**:866-868.
- 712 12. **Bobrov AG, Kirillina O, Forman S, Mack D, Perry RD.** 2008. Insights into *Yersinia*
713 *pestis* biofilm development: topology and co-interaction of Hms inner membrane proteins
714 involved in exopolysaccharide production. *Environ Microbiol* **10**:1419-1432.
- 715 13. **Lindenberg S, Klauck G, Pesavento C, Klauck E, Hengge R.** 2013. The EAL domain
716 protein YciR acts as a trigger enzyme in a c-di-GMP signalling cascade in *E. coli* biofilm
717 control. *EMBO J* **32**:2001-2014.
- 718 14. **Dahlstrom KM, Giglio KM, Collins AJ, Sondermann H, O'Toole GA.** 2015.
719 Contribution of physical interactions to signaling specificity between a diguanylate
720 cyclase and its effector. *MBio* **6**:e01978-01915.
- 721 15. **Monds RD, Newell PD, Gross RH, O'Toole GA.** 2007. Phosphate-dependent modulation
722 of c-di-GMP levels regulates *Pseudomonas fluorescens* Pf0-1 biofilm formation by
723 controlling secretion of the adhesin LapA. *Mol Microbiol* **63**:656-679.

- 724 16. **Lee SH, Angelichio MJ, Mekalanos JJ, Camilli A.** 1998. Nucleotide sequence and
725 spatiotemporal expression of the *Vibrio cholerae* *vieSAB* genes during infection. *J*
726 *Bacteriol* **180**:2298-2305.
- 727 17. **Lim B, Beyhan S, Meir J, Yildiz FH.** 2006. Cyclic-diGMP signal transduction systems in
728 *Vibrio cholerae*: modulation of rugosity and biofilm formation. *Mol Microbiol* **60**:331-
729 348.
- 730 18. **Yildiz FH, Liu XS, Heydorn A, Schoolnik GK.** 2004. Molecular analysis of rugosity in a
731 *Vibrio cholerae* O1 El Tor phase variant. *Mol Microbiol* **53**:497-515.
- 732 19. **Tuckerman JR, Gonzalez G, Sousa EHS, Wan XH, Saito JA, Alam M, Gilles-Gonzalez**
733 **MA.** 2009. An oxygen-sensing diguanylate cyclase and phosphodiesterase couple for c-
734 di-GMP control. *Biochemistry* **48**:9764-9774.
- 735 20. **Zahringer F, Lacanna E, Jenal U, Schirmer T, Boehm A.** 2013. Structure and signaling
736 mechanism of a zinc-sensory diguanylate cyclase. *Structure* **21**:1149-1157.
- 737 21. **Dahlstrom KM, O'Toole GA.** 2017. A symphony of cyclases: Specificity in diguanylate
738 cyclase signaling. *Annu Rev Microbiol* doi:10.1146/annurev-micro-090816-093325.
- 739 22. **Boyd CD, O'Toole GA.** 2012. Second messenger regulation of biofilm formation:
740 breakthroughs in understanding c-di-GMP effector systems. *Annu Rev Cell Dev Biol*
741 **28**:439-462.
- 742 23. **Newell PD, Monds RD, O'Toole GA.** 2009. LapD is a bis-(3',5')-cyclic dimeric GMP-
743 binding protein that regulates surface attachment by *Pseudomonas fluorescens* Pf0-1.
744 *Proc Natl Acad Sci U S A* **106**:3461-3466.

- 745 24. **Newell PD, Boyd CD, Sondermann H, O'Toole GA.** 2011. A c-di-GMP effector system
746 controls cell adhesion by inside-out signaling and surface protein cleavage. *PLoS Biol*
747 **9**:e1000587.
- 748 25. **Liu N, Pak T, Boon EM.** 2010. Characterization of a diguanylate cyclase from *Shewanella*
749 *woodyi* with cyclase and phosphodiesterase activities. *Mol Biosyst* **6**:1561-1564.
- 750 26. **Navarro MV, De N, Bae N, Wang Q, Sondermann H.** 2009. Structural analysis of the
751 GGDEF-EAL domain-containing c-di-GMP receptor FimX. *Structure* **17**:1104-1116.
- 752 27. **Navarro MV, Newell PD, Krasteva PV, Chatterjee D, Madden DR, O'Toole GA,**
753 **Sondermann H.** 2011. Structural basis for c-di-GMP-mediated inside-out signaling
754 controlling periplasmic proteolysis. *PLoS Biol* **9**:e1000588.
- 755 28. **Ryjenkov DA, Simm R, Romling U, Gomelsky M.** 2006. The PilZ domain is a receptor for
756 the second messenger c-di-GMP: the PilZ domain protein YcgR controls motility in
757 enterobacteria. *J Biol Chem* **281**:30310-30314.
- 758 29. **Pratt JT, Tamayo R, Tischler AD, Camilli A.** 2007. PilZ domain proteins bind cyclic
759 diguanylate and regulate diverse processes in *Vibrio cholerae*. *J Biol Chem* **282**:12860-
760 12870.
- 761 30. **Newell PD, Yoshioka S, Hvorecny KL, Monds RD, O'Toole GA.** 2011. Systematic
762 analysis of diguanylate cyclases that promote biofilm formation by *Pseudomonas*
763 *fluorescens* Pf0-1. *J Bacteriol* **193**:4685-4698.
- 764 31. **Townsley L, Yildiz FH.** 2015. Temperature affects c-di-GMP signalling and biofilm
765 formation in *Vibrio cholerae*. *Environ Microbiol* doi:10.1111/1462-2920.12799.

- 766 32. **Ha DG, Richman ME, O'Toole GA.** 2014. Deletion mutant library for investigation of
767 functional outputs of cyclic diguanylate metabolism in *Pseudomonas aeruginosa* PA14.
768 Appl Environ Microbiol **80**:3384-3393.
- 769 33. **Spurbeck RR, Tarrien RJ, Mobley HL.** 2012. Enzymatically active and inactive
770 phosphodiesterases and diguanylate cyclases are involved in regulation of Motility or
771 sessility in *Escherichia coli* CFT073. MBio **3**.
- 772 34. **Hollomon JM, Grahl N, Willger SD, Koeppen K, Hogan DA.** 2016. Global role of cyclic
773 AMP signaling in pH-dependent responses in *Candida albicans*. mSphere **1**.
- 774 35. **Luo Y, Zhao K, Baker AE, Kuchma SL, Coggan KA, Wolfgang MC, Wong GC,**
775 **O'Toole GA.** 2015. A hierarchical cascade of second messengers regulates *Pseudomonas*
776 *aeruginosa* surface behaviors. MBio **6**.
- 777 36. **Guvener ZT, Harwood CS.** 2007. Subcellular location characteristics of the *Pseudomonas*
778 *aeruginosa* GGDEF protein, WspR, indicate that it produces cyclic-di-GMP in response
779 to growth on surfaces. Mol Microbiol **66**:1459-1473.
- 780 37. **Kuchma SL, Griffin EF, O'Toole GA.** 2012. Minor pilins of the type IV pilus system
781 participate in the negative regulation of swarming motility. J Bacteriol **194**:5388-5403.
- 782 38. **Petrova OE, Cherny KE, Sauer K.** 2014. The *Pseudomonas aeruginosa* diguanylate
783 cyclase GcbA, a homolog of *P. fluorescens* GcbA, promotes initial attachment to
784 surfaces, but not biofilm formation, via regulation of motility. J Bacteriol **196**:2827-2841.
- 785 39. **Baraquet C, Harwood CS.** 2013. Cyclic diguanosine monophosphate represses bacterial
786 flagella synthesis by interacting with the Walker A motif of the enhancer-binding protein
787 FleQ. Proc Natl Acad Sci U S A **110**:18478-18483.

- 788 40. **Lee ER, Baker JL, Weinberg Z, Sudarsan N, Breaker RR.** 2010. An allosteric self-
789 splicing ribozyme triggered by a bacterial second messenger. *Science* **329**:845-848.
- 790 41. **Dahlstrom KM, Giglio KM, Sondermann H, O'Toole GA.** 2016. The inhibitory site of a
791 diguanylate cyclase is a necessary element for interaction and signaling with an effector
792 protein. *J Bacteriol* **198**:1595-1603.
- 793 42. **Tan J, Doing G, Lewis KA, Price CE, Chen KM, Cady KC, Perchuk B, Laub MT,**
794 **Hogan DA, Greene CS.** 2017. Unsupervised extraction of stable expression signatures
795 from public compendia with an ensemble of neural networks. *Cell Syst* **5**:63-71 e66.
- 796 43. **Baker AE, Diepold A, Kuchma SL, Scott JE, Ha DG, Orazi G, Armitage JP, O'Toole**
797 **GA.** 2016. PilZ domain protein FlgZ mediates cyclic di-GMP-dependent swarming
798 motility control in *Pseudomonas aeruginosa*. *J Bacteriol* **198**:1837-1846.
- 799 44. **Andrade MO, Alegria MC, Guzzo CR, Docena C, Rosa MC, Ramos CH, Farah CS.**
800 2006. The HD-GYP domain of RpfG mediates a direct linkage between the Rpf quorum-
801 sensing pathway and a subset of diguanylate cyclase proteins in the phytopathogen
802 *Xanthomonas axonopodis* pv *citri*. *Mol Microbiol* **62**:537-551.
- 803 45. **Allen RH, Jakoby WB.** 1969. Tartaric acid metabolism. IX. Synthesis with tartrate
804 epoxidase. *J Biol Chem* **244**:2078-2084.
- 805 46. **Hurlbert RE, Jakoby WB.** 1965. Tartaric acid metabolism. I. Subunits of L(+)-tartaric acid
806 dehydrase. *J Biol Chem* **240**:2772-2777.
- 807 47. **Park KH, Lee CY, Son HJ.** 2009. Mechanism of insoluble phosphate solubilization by
808 *Pseudomonas fluorescens* RAF15 isolated from ginseng rhizosphere and its plant growth-
809 promoting activities. *Lett Appl Microbiol* **49**:222-228.

- 810 48. **Cooley RB, O'Donnell JP, Sondermann H.** 2016. Coincidence detection and bi-directional
811 transmembrane signaling control a bacterial second messenger receptor. *Elife* **5**.
- 812 49. **Yan J, Deforet M, Boyle KE, Rahman R, Liang R, Okegbe C, Dietrich LEP, Qiu W,**
813 **Xavier JB.** 2017. Bow-tie signaling in c-di-GMP: Machine learning in a simple
814 biochemical network. *PLoS Comput Biol* **13**:e1005677.
- 815 50. **Petrova OE, Cherny KE, Sauer K.** 2015. The diguanylate cyclase GcbA facilitates
816 *Pseudomonas aeruginosa* biofilm dispersion by activating BdlA. *J Bacteriol* **197**:174-
817 187.
- 818 51. **Simon R, Priefer U, Puhler A.** 1983. A broad host range mobilization system for invivo
819 genetic-engineering - Transposon mutagenesis in gram-negative bacteria. *Bio-*
820 *Technology* **1**:784-791.
- 821 52. **Monds RD, Newell PD, Schwartzman JA, O'Toole GA.** 2006. Conservation of the Pho
822 regulon in *Pseudomonas fluorescens* Pf0-1. *Appl Environ Microbiol* **72**:1910-1924.
- 823 53. **Shanks RM, Caiazza NC, Hinsa SM, Toutain CM, O'Toole GA.** 2006. *Saccharomyces*
824 *cerevisiae*-based molecular tool kit for manipulation of genes from gram-negative
825 bacteria. *Appl Environ Microbiol* **72**:5027-5036.

826

827

828

829

830

831 **Figure Legends**

832

833 **Figure 1. Domain-level depiction of c-di-GMP-related proteins in *P. fluorescens*.** All
834 proteins in *P. fluorescens* with a predicted EAL, GGDEF, PilZ, or other known c-di-GMP-
835 binding domain are depicted. Domains illustrated include those predicted by Pfam and SMART.
836 In cases where these two databases predicted different domain architecture over the same amino
837 acid span, the Pfam prediction was used and Pfam domain naming conventions are used in all
838 cases. If a gene/protein name has been assigned/reported, it is also shown here. All recurrent
839 domains are listed in the key, with the remainder of domains labeled on each protein. Models are
840 not drawn to scale, although large amino acid stretches without a domain depicted represent
841 space in the protein where no known domains are predicted. **A.** GGDEF domain-containing
842 proteins. **B.** Dual domain-containing proteins. **C.** EAL, PilZ, and FleQ domain-containing
843 proteins. The domain key is shown boxed.

844

845 **Figure 2. A majority of c-di-GMP-related genes show a defect in biofilm development when**
846 **mutated.** **A.** Data from biofilm formation assays conducted under all 188 tested nutrients for
847 each gene indicated are displayed as a boxplot. Each box is color-coded based on the domain
848 associated with the indicated gene (see legend on right side). Mutants were compared to the wild
849 type using a Wilcoxon signed rank test and p values were adjusted using a Bonferroni correction.
850 ***p < 0.001. **B.** Shown is a bar graph of each of the mutants versus the number of nutrient
851 conditions wherein the mutant shows a biofilm phenotype, defined as a value significantly
852 different than the mean WT value using the test described in the Materials and Methods. The
853 height of the bar above and below the origin indicates the number of conditions that each mutant

854 had a significantly enhanced and reduced biofilm, respectively. The bars are color coded to
855 match the domain of the protein encoded by the mutated gene with the legend shown on the right
856 side of the panel.

857
858 **Figure 3. Analysis of the expression of genes in the network.** **A.** Shown is the \log_2 expression
859 ratio for each gene in the network across the 50 nutrients tested. The ratio was calculated with
860 dividing the number of counts per 1000 normalized mRNA transcripts on BMM plus the
861 indicated nutrient by the number of mRNA transcripts in the BMM. A positive value indicates
862 an increase in expression in response to the nutrient, while a negative value indicates a decrease
863 in expression. All ratios were transformed to their \log_2 values then plotted. Each red dot
864 indicates a condition in which expression was not significantly different than the expression on
865 BMM alone, while blue dots indicate conditions in which expression was significantly different
866 from BMM (determined as described in the Materials and Methods). A version of this plot with
867 the individual dots assigned to nutrients is shown in **Figure S4A.** **B.** Shown is a plot relating the
868 number of nutrient conditions for which transcription of the indicated gene was changed AND a
869 biofilm defect in a strain carrying a mutation in that same gene was observed in that same
870 nutrient. Data are shown for all the genes in the network. **C.** Conditions where a gene with both a
871 significant change in transcription and the mutation of that gene had a significant impact on
872 biofilm formation are plotted and color coded by nutrient.

873
874 **Figure 4. Physical interaction map of c-di-GMP-related proteins in *P. fluorescens*.** **A.**
875 BTH101 *E. coli* cells were grown on LB medium supplemented with X-gal for 20 hours and
876 colonies were assessed for blue color. Protein pairs of each domain type were tested for physical

877 interaction in the bacterial two-hybrid system against each member of every other domain type,
878 with dual-domain proteins also being tested against each other. Each gene was tested on both the
879 pUT18 and pKNT25 vectors; a positive result was recorded if either the T18- or T25-fused
880 protein produced blue colonies with another member of the network. Proteins that failed to
881 interact are shown at bottom. FleQ was not tested in this analysis. The interaction map was
882 generated by Cytoscape v 3.5.1. **B.** LapD serves as a node of interaction among several proteins
883 shown here. The legend is shown on the bottom, right of the figure: GGDEF-containing proteins
884 are represented by green circles, EAL-containing proteins by orange diamonds, dual-domain
885 proteins by blue rectangles, and PilZ-containing proteins by purple triangles.

886

887 **Figure 5. Assessing the protein-protein interactions in the network.** **A.** Shown is a plot of
888 interacting pairs of proteins in the network versus the number of nutrient conditions wherein
889 each pair shared a biofilm phenotype. Orange coloration indicates the number of conditions
890 where both interactors affected biofilm formation in the same way, while blue color indicated the
891 number of conditions that interactors affected biofilm formation in opposite directions. **B.** The
892 stacked box plot shows the relationship between interacting pairs of proteins in the network
893 versus a change in expression of both genes in the interacting pair in the same nutrient. Orange
894 coloration indicates the number of conditions where both interactors showed a change in
895 transcription in the same direction, while blue color indicated the number of conditions that
896 interactors showed a change in transcription in the opposite direction. Note the difference in the
897 scale bar between panels A and B.

898

899

900 **Supplemental Figure Legends**

901

902 **Figure S1. Biofilm phenotypes vary with nutrients.** Shown is a histogram of the quantified
903 biofilm for the WT as a function of the nutrient added to BMM. The histogram is arranged in
904 order from lowest biofilm formed to highest biofilm formed. The red bar indicated the biofilm
905 formed in BMM.

906

907 **Figure S2. Heat map of biofilm phenotypes for each mutant with each nutrient.** In this
908 heatmap, red indicates a higher ratio than WT and blue is lower ratio than WT. That is, cells that
909 are red indicate conditions where a mutant biofilm either increased more or decreased less than
910 WT in a condition relative to the median biofilm formed by that mutant. Black indicates values
911 not significantly different from WT.

912

913 **Figure S3. Heat map of gene expression for each gene in the network on 50 different**
914 **nutrients.** Expression of genes when grown on the indicated nutrient-supplemented BBM in
915 liquid medium. Only values which are significantly different from expression on BMM are
916 colored (see Materials and Methods for details of statistical tests). Red cells indicate conditions
917 where genes were expressed at significantly higher than in BMM, while blue cells indicate
918 conditions where genes were expressed at significantly lower levels compared to BMM.

919

920 **Figure S4. Summary of expression analysis. A.** Shown is the \log_2 expression ratio for each
921 gene in the network across the 50 nutrients tested. The ratio was calculated with dividing the
922 number of mRNA transcripts on BMM plus the indicated nutrient by the number of mRNA

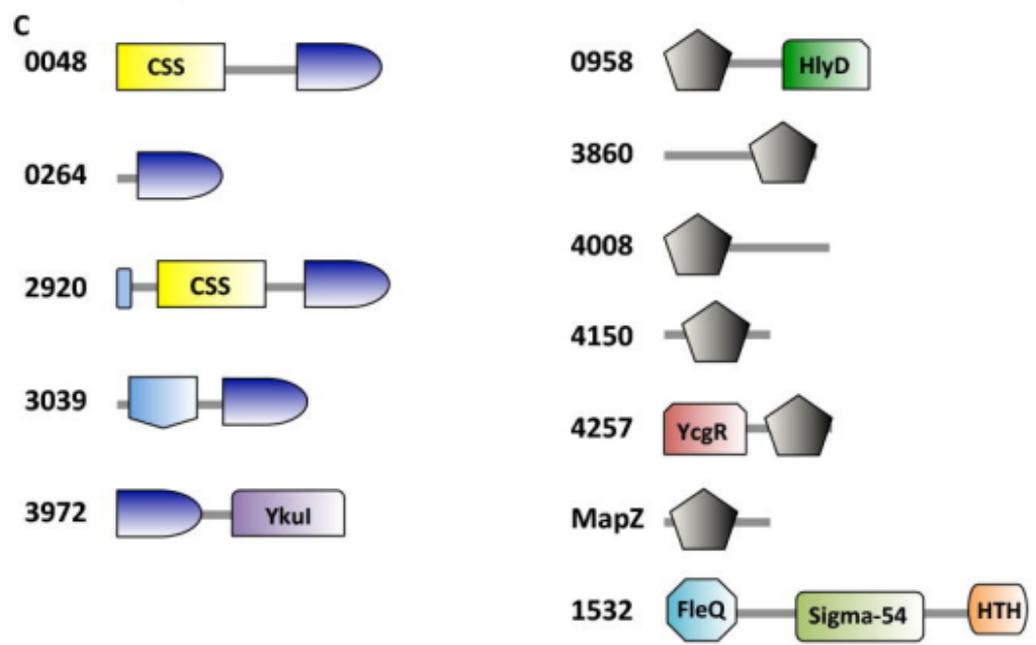
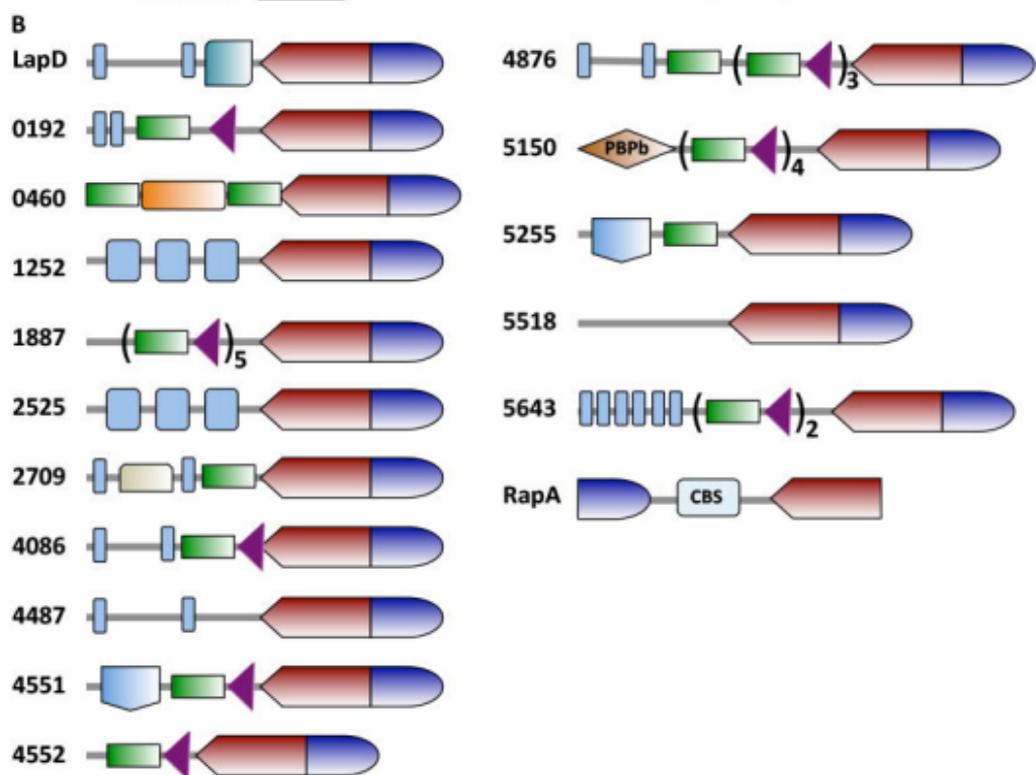
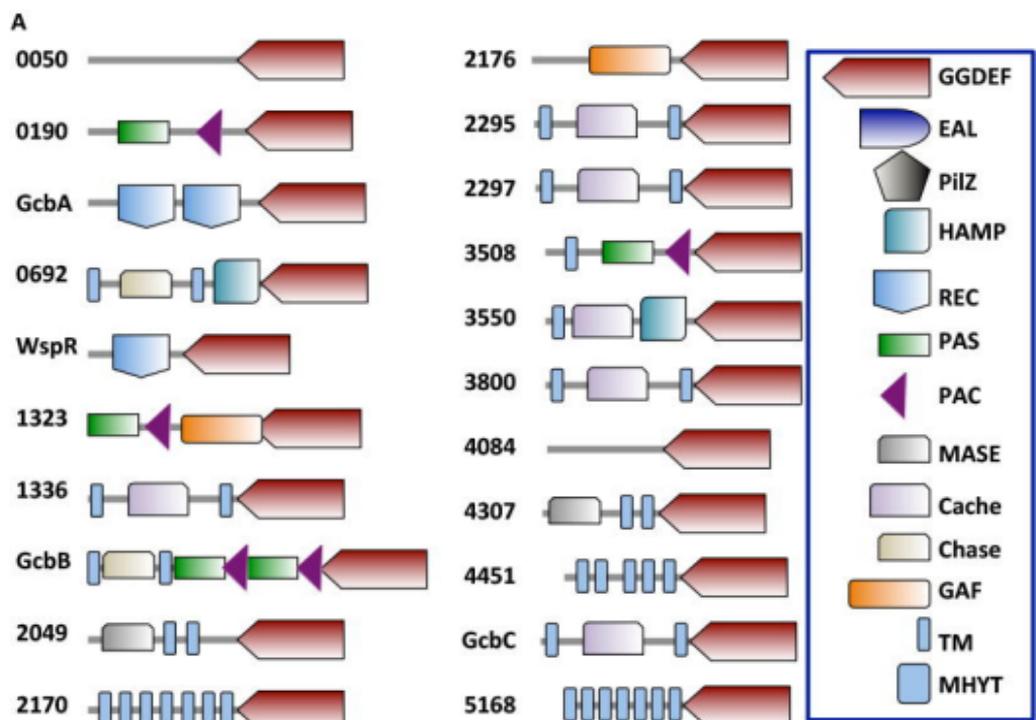
923 transcripts in the BMM. A positive value indicates an increase in expression in response to the
924 nutrient compared to BMM, while a negative value indicates a decrease in expression compared
925 to BMM. All ratios were transformed to their \log_2 values then plotted. Each colored dot
926 indicated the ratio of a different carbon source, with the key to the colors shown on the right. **B.**
927 \log_2 transformed ratio of expression of genes when cells were grown on a solid medium divided
928 by their expression in the same medium in broth. The carbon source added to the liquid or solid
929 BMM is indicated. * $p < 0.05$ Wilcoxon signed rank test.

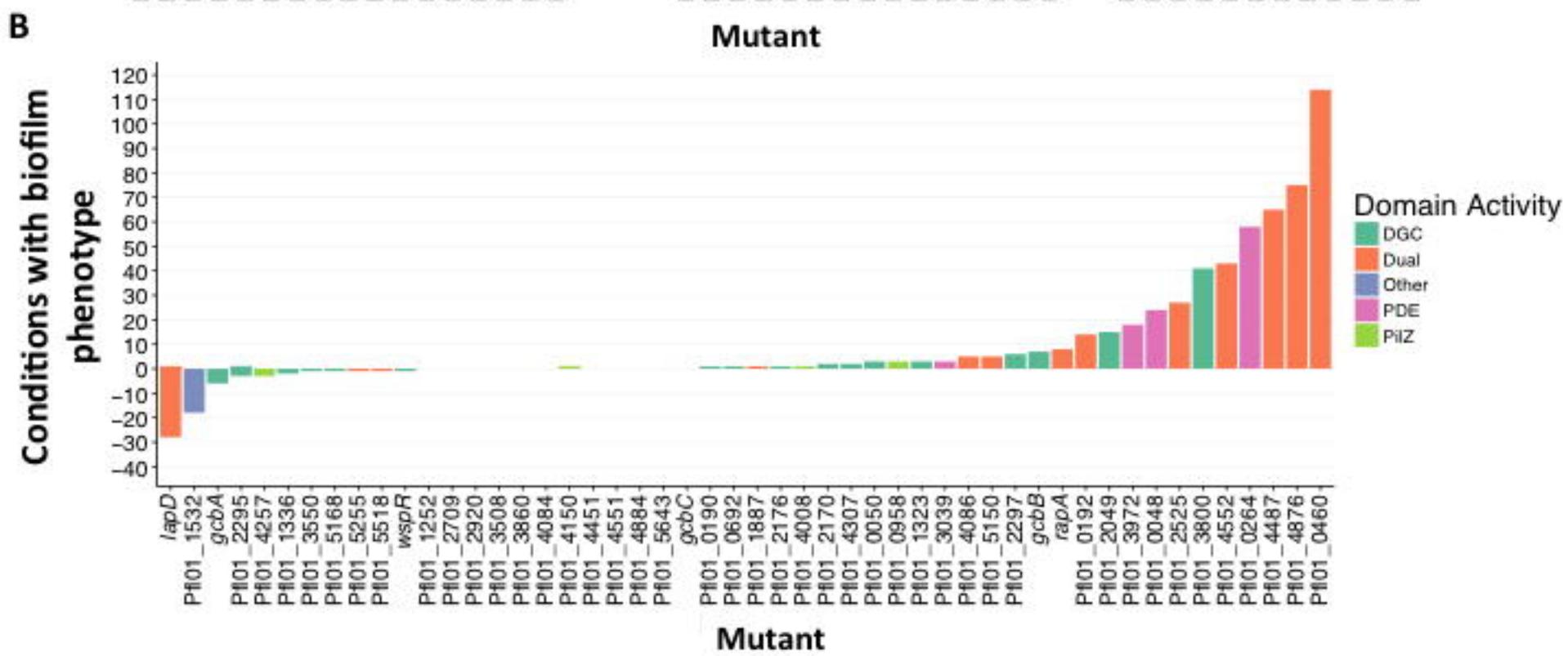
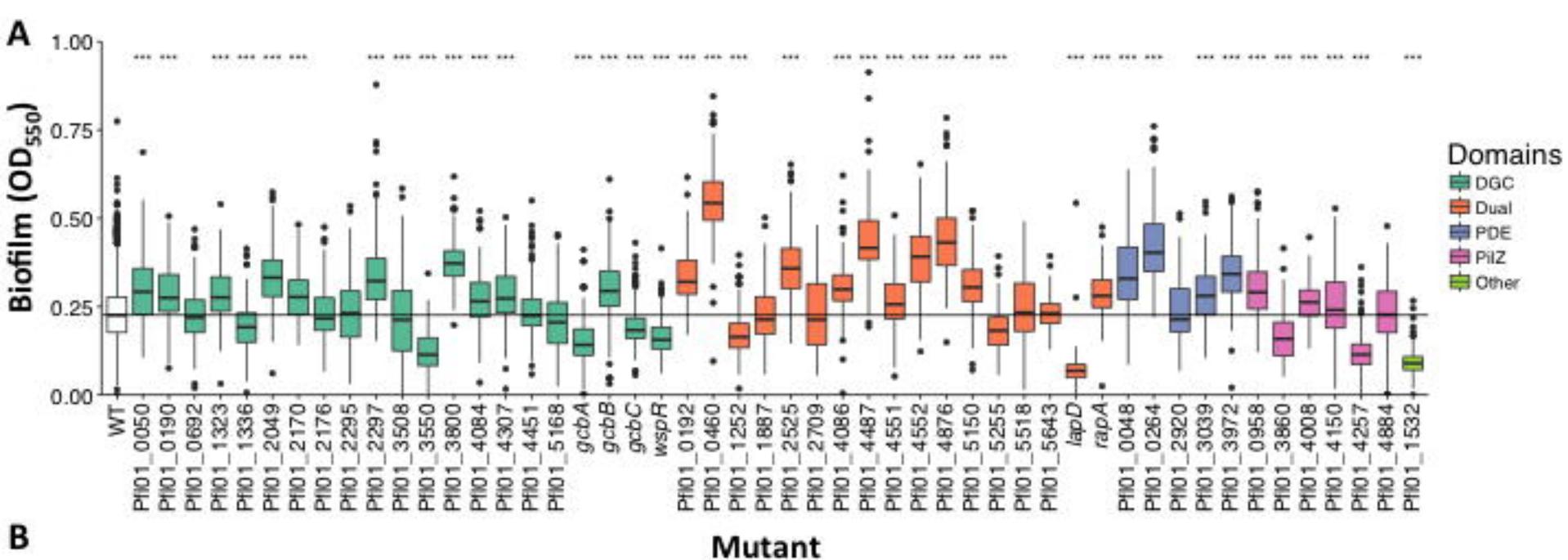
930
931 **Figure S5. Comparing characteristics of c-di-GMP gene expression between organisms.** *P.*
932 *fluorescens* Pf01 Nanostring expression data were acquired as described in the Materials and
933 Methods. *P. aeruginosa* PAO1 expression data were retrieved from the NCBI database of
934 published microarray analyses using the tools developed by Tan *et al.* (42). The *P. aeruginosa*
935 PAO1 compendium data were normalized by assigning the highest value in each experiment the
936 value 1, and the lowest value the value 0. All other values in the experiment were scaled linearly
937 between those values using the following formula: for each value x : $z_i = (x_i - \min(x)) / (\max(x) -$
938 $\min(x))$ where z_i is the normalized value of x_i . **A-D.** Genes were grouped according by predicted
939 proteins domains into the groups GGDEF, EAL, Dual, HD-GYP, and PilZ for c-di-GMP-related
940 genes. Genes not encoding one of those domains was termed “non-cdg”. In addition, genes in the
941 Pho regulon (42) were included as an example of the characteristics of genes under
942 transcriptional regulation. **A-B.** The median normalized expression of each gene was calculated
943 and the distribution of those values is plotted as a boxplot with genes grouped according to the
944 characterization of the domains they encode. Panel **A** shows data from *P. fluorescens* and panel **B**
945 from *P. aeruginosa* PAO1. **C-D.** The coefficient of variance of the normalized expression of

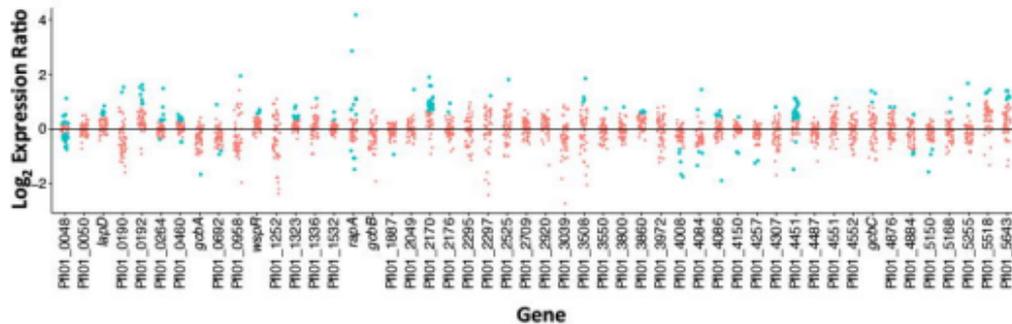
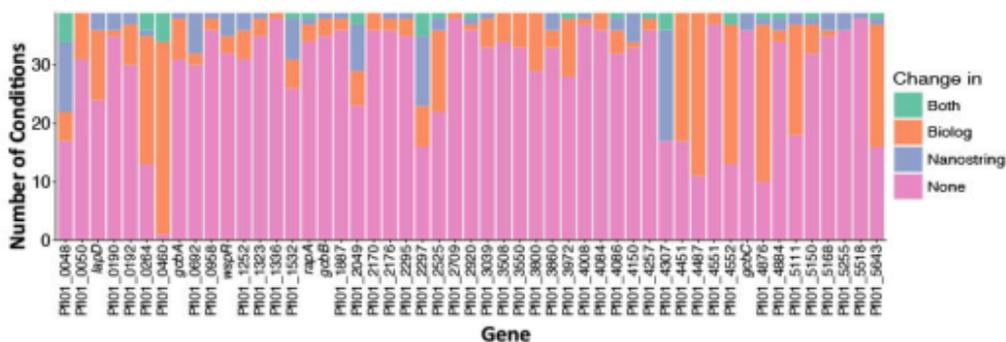
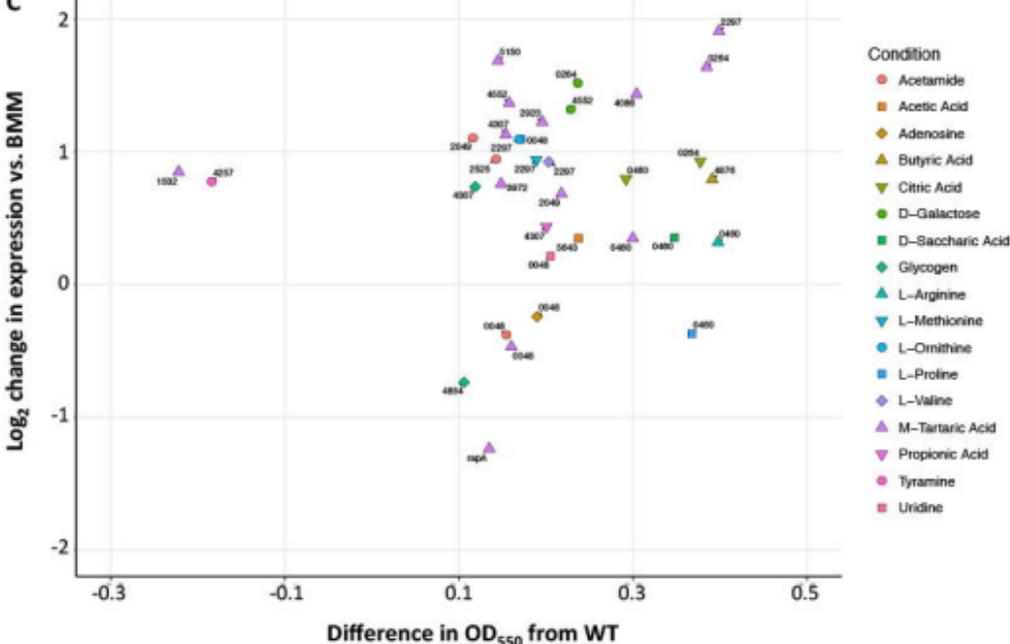
946 each gene was calculated by dividing the standard deviation by the mean, and the distribution of
947 those values is plotted as a boxplot with genes grouped according to the characterization of the
948 domain-containing proteins they encode. Panel **C** shows data from *P. fluorescens* and panel **D**
949 from *P. aeruginosa* PAO1. **E-F**. Spearman correlation values for each pair of c-di-GMP-related
950 gene in each organism were calculated and plotted using the R package “corrplot”. Significance
951 of these correlations was calculated using the “corrttest” function in this package and p values
952 were adjusted using a Bonferroni correction. Correlations with adjusted p values below 0.05
953 were not plotted. Panel **E** shows data from *P. fluorescens* and panel **E** from *P. aeruginosa* PAO1.
954

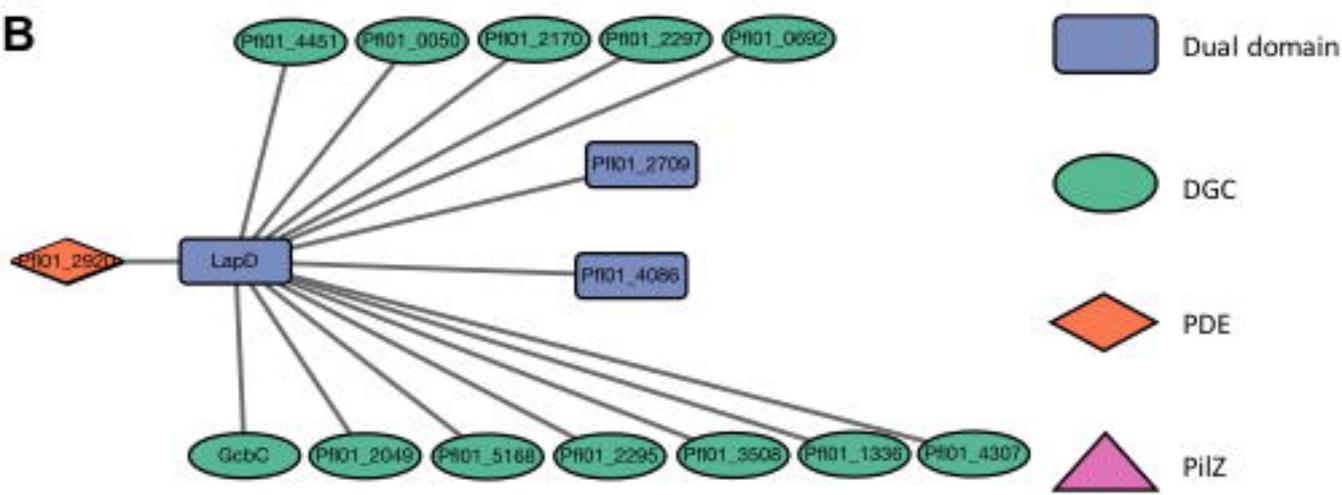
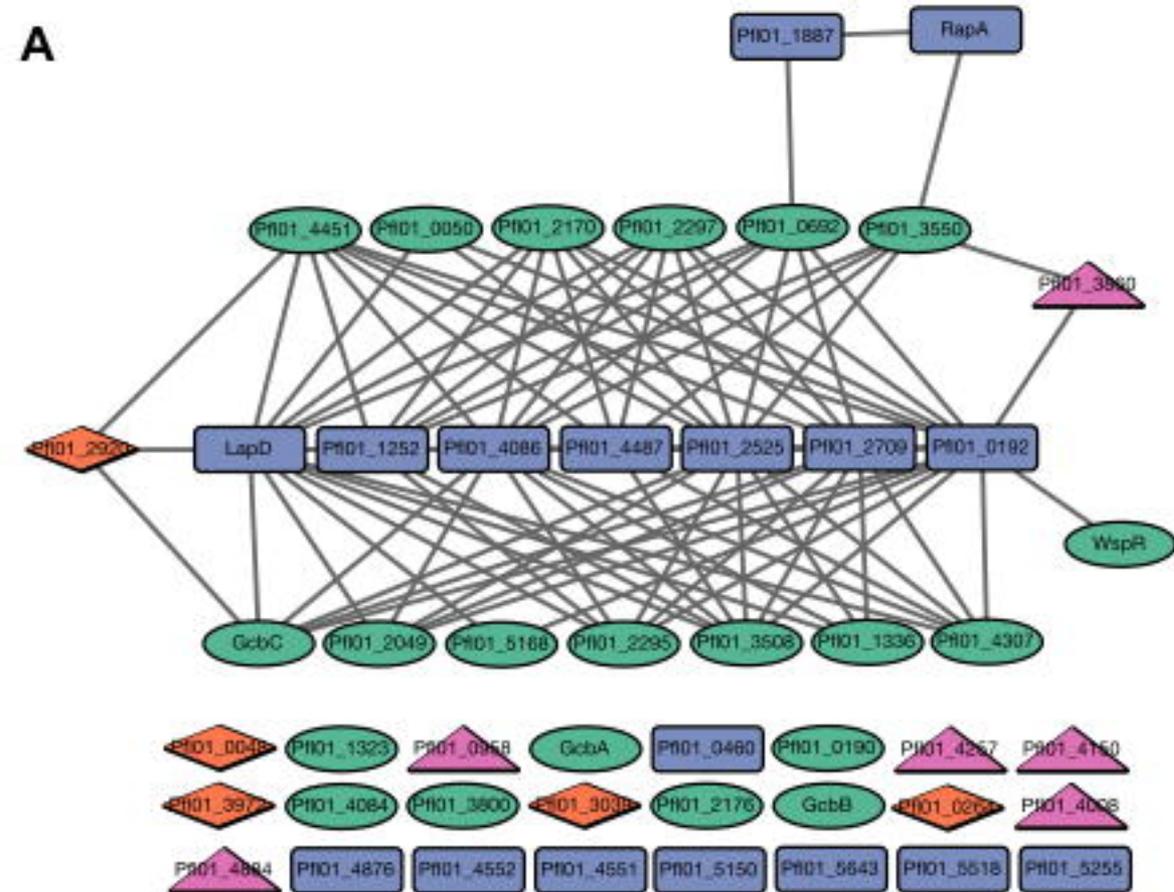
955 **Figure S6. Schematic of protein-protein interaction sub-networks.** **A.** Shown is the key for
956 each domain type shown in panels B-E. **B-E.** Small interaction networks showing proteins that
957 interact with a large number of other proteins. Such “hub” proteins are Pfl0_0192 (panel B),
958 Pfl0_2709 (panel C), Pfl0_2525 (panel D) and Pfl0_4068 (panel E). A summary of all such
959 interactions is presented in **Table S10** and **Figure 4**.

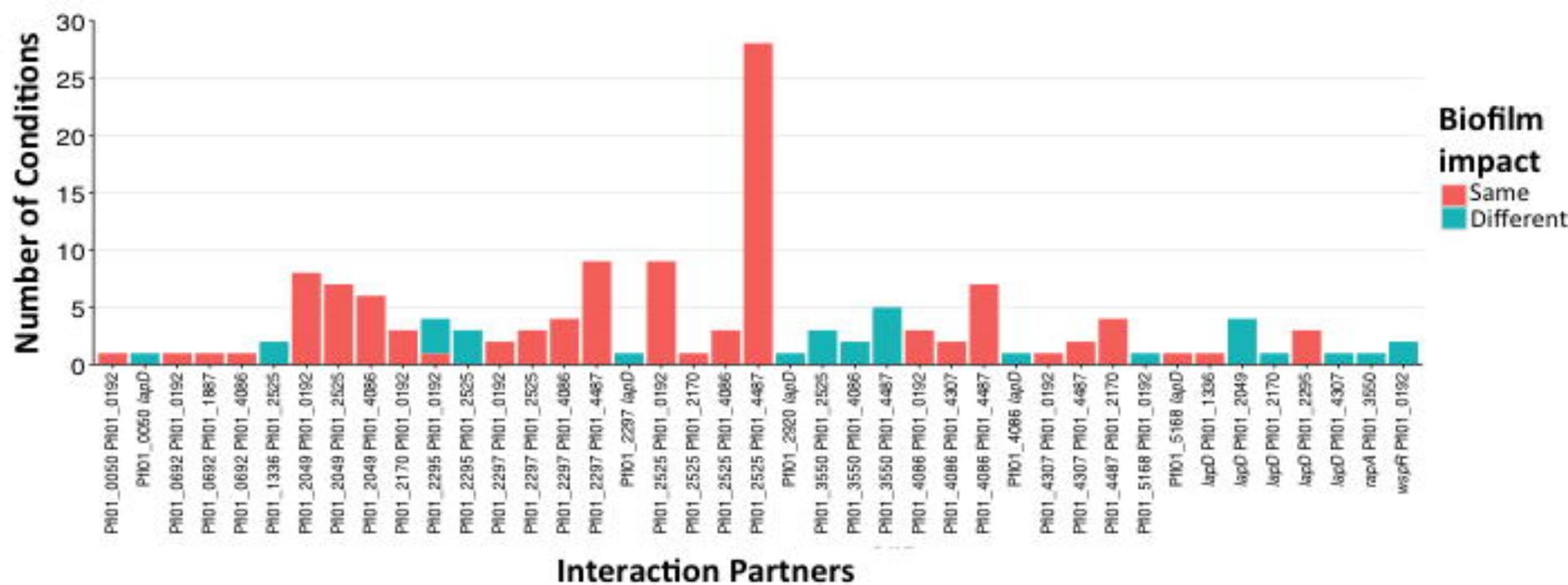
960





A**B****C**



A**B**