Borrelia burgdorferi PlzA is a cyclic-di-GMP dependent DNA and RNA binding protein

Nerina Jusufovic[1], Andrew C. Krusenstjerna[1], Christina R. Savage[1], Timothy C. Saylor[1], Catherine A. Brissette[2], Wolfram R. Zückert[3], Paula J. Schlax[4], Md A. Motaleb[5], and Brian Stevenson[1,6,*]

[1] Microbiology, Immunology and Molecular Genetics, University of Kentucky College of Medicine, University of Kentucky, Lexington, Kentucky, 40526-0001, USA [2] Department of Biomedical Sciences, University of North Dakota, School of Medicine and Health Sciences, Grand Forks, ND 58203-9061, USA, [3] Department of Microbiology, Molecular Genetics and Immunology, University of Kansas School of Medicine, Kansas City, KS 66160, USA, [4] Department of Chemistry and Biochemistry, Bates College, Lewiston, ME, 04240-6030, USA, [5] Department of Microbiology and Immunology, Brody School of Medicine, East Carolina University, Greenville, NC 27834-435, USA, and [6] Department of Entomology, University of Kentucky, Lexington, Kentucky, 40526-0001, USA.

* To whom correspondence should be addressed. Tel: +1 859 257 9358; Fax: 859 257 8994; Email: brian.stevenson@uky.edu

ABSTRACT

The PilZ domain-containing protein, PlzA, is the only known cyclic di-GMP binding protein encoded by all Lyme disease spirochetes. PlzA has been implicated in the regulation of many borrelial processes, but the effector mechanism of PlzA was not previously known. Here we report that PlzA can bind DNA and RNA and that nucleic acid binding requires c-di-GMP, with the affinity of PlzA for nucleic acids increasing as concentrations of c-di-GMP were increased. A mutant PlzA that is incapable of binding c-di-GMP did not bind to any tested nucleic acids. We also determined that PlzA interacts predominantly with the major groove of DNA and that sequence length plays a role in DNA binding affinity. PlzA is a dual-domain protein with a PilZ-like N-terminal domain linked to a canonical C-terminal PilZ domain. Dissection of the domains demonstrated that the separated N-terminal domain bound nucleic acids independently of c-di-GMP. The C-terminal domain, which includes the c-di-GMP binding motifs, did not bind nucleic acids under any tested conditions. Our data are supported by computational docking, which predicts that c-di-GMP binding at the C-terminal domain stabilizes the overall protein structure and facilitates PlzA-DNA interactions via residues in the N-terminal domain. Based on our data, we propose that levels of c-di-GMP during the various stages of the enzootic life cycle direct PlzA binding to regulatory targets.

34 **INTRODUCTION**

35 Cyclic bis-(3'→5')-dimeric guanosine monophosphate (c-di-GMP) is a ubiquitous bacterial
36 cyclic di-nucleotide. In many bacterial species, it engages in the regulation of key processes
37 such as transcription, motility, biofilm formation, and virulence (Ryjenkov *et al.*, 2005; Hengge,
38 2009; Romling *et al.*, 2013; Jenal *et al.*, 2017; Valentini and Filloux, 2019). The Lyme disease
39 spirochete, *Borrelia burgdorferi* sensu lato, referred to as *Borrelia burgdorferi* henceforth, has
40 a c-di-GMP regulatory network consisting of a transmembrane sensor histidine kinase (Hk1),
41 a response regulator diguanylate cyclase (Rrp1), two phosphodiesterases (PdeA and PdeB),
42 and a single, chromosomally encoded PilZ domain-containing protein, PlzA (Galperin *et al.*,
43 2001; Freedman *et al.*, 2009; Rogers *et al.*, 2009; Sultan *et al.*, 2010; Caimano *et al.*, 2011;
44 Sultan *et al.*, 2011; Novak *et al.*, 2014).

45 The borrelial Hk1/Rrp1 two-component system synthesizes c-di-GMP; activation of Hk1
46 (BB_0420) occurs through an unknown signal (Caimano *et al.*, 2011; Bauer *et al.*, 2015). Rrp1
47 (BB_0419), the only identified diguanylate cyclase in *B. burgdorferi,* synthesizes c-di-GMP
48 from two GTP molecules and is essential for acquisition and survival in the tick vector (Rogers
49 *et al.*, 2009; Kostick *et al.*, 2011). Turnover of c-di-GMP is mediated by PdeA (BB_0363) and
50 PdeB (BB_0374), the only phosphodiesterases present in *B. burgdorferi*, which degrade c-di-
51 GMP to pGpG and GMP, respectively (Sultan *et al.*, 2010; Sultan *et al.*, 2011). *B. burgdorferi*
52 must delicately balance c-di-GMP signaling, as c-di-GMP is required for both maintenance in
53 the tick vector and successful transmission into the vertebrate host, but constitutive synthesis
54 is detrimental to spirochete survival during vertebrate infection (Sultan *et al.*, 2010; Sultan *et*
55 *al.*, 2011; Caimano *et al.*, 2015; Groshong *et al.*, 2021).

56 Synthesized c-di-GMP is bound by PlzA, the only universal c-di-GMP binding protein found in
57 all isolates of Lyme disease *Borreliae* (Kostick-Dunn *et al.*, 2018). Mutation of *plzA* has been
58 implicated in many *B. burgdorferi* phenotypes, including dysregulation of alternative
59 carbohydrate utilization, decreased motility, and virulence defects (Pitzer *et al.*, 2011; He *et*
60 *al.*, 2014; Mallory *et al.*, 2016; Zhang *et al.*, 2018; Groshong *et al.*, 2021). Binding of c-di-GMP

61 induces conformational changes in PlzA structure, which has led to the hypothesis that c-di-
62 GMP binding acts as a switching mechanism providing the holo and apo-forms of PlzA with
63 distinct functions (Mallory *et al.*, 2016; Groshong *et al.*, 2021; Grassmann *et al.*, 2023).
64 Specifically, c-di-GMP bound PlzA is thought to regulate genes required for survival in the tick
65 host, while apo-PlzA is speculated to play a role in regulation during vertebrate infection
66 (Groshong *et al.*, 2021; Grassmann *et al.*, 2023).

67 A common mechanism of action for c-di-GMP effector proteins is to bind nucleic acids (Wilksch
68 *et al.*, 2011; Tschowri *et al.*, 2014; Tan *et al.*, 2015; Wang *et al.*, 2016; Schäper *et al.*, 2017;
69 Hsieh *et al.*, 2018; Skotnicka *et al.*, 2020). It has been hypothesized that PlzA could bind DNA
70 and regulate genes through impacts on transcription, such as directly interacting with RNA
71 polymerase (RNAP), but this has yet to be validated experimentally (Groshong *et al.*, 2021;
72 Grassmann *et al.*, 2023). Conversely, others have argued that PlzA cannot bind DNA, as it
73 lacks an obvious, conventional DNA-binding motif (Caimano *et al.*, 2015; Zhang *et al.*, 2018).
74 Indeed, Zhang *et al.* concluded that PlzA does not bind DNA, as their electrophoretic mobility
75 shift assays (EMSA) did not identify binding to the *glp* promoter region (Zhang *et al.*, 2018).
76 We note that the raw data and methodologies, including protein and c-di-GMP concentrations,
77 were not shown in that report. After the submission of a preprint of our manuscript, another
78 group published that PlzA has RNA chaperone activity that is independent of c-di-GMP,
79 supporting the notion that holo- and apo-forms of PlzA have distinct functions (Van Gundy *et*
80 *al.*, 2023).

81 Given the multiple conflicting hypotheses regarding PlzA function, as well as the potential for
82 additional mechanisms between liganded and unliganded PlzA, we sought to independently
83 characterize PlzA nucleic acid binding function. To that end, we chose a well-established locus
84 PlzA is known to regulate, *glpFKD*, the glycerol catabolism operon. The *glpFKD* operon
85 promotes *B. burgdorferi* survival in ticks (Pappas *et al.*, 2011; Corona and Schwartz, 2015;
86 Helble *et al.*, 2021). After a tick has completed the digestion of its blood meal, it is hypothesized
87 that the midgut is devoid of glucose and many other nutrients. To survive in that environment,

88 Lyme disease spirochetes would need to utilize alternate carbon sources such as glycerol,
89 which is available in the tick host (Lee and Baust, 1987). The *glpFKD* operon is regulated
90 through the Hk1/Rrp1 pathway, and borrelial *plzA* mutants are defective in glycerol metabolism
91 (Rogers *et al.*, 2009; He *et al.*, 2011; Caimano *et al.*, 2015; Zhang *et al.*, 2018). Further, PlzA
92 was found to both positively and negatively affect the *glp* operon at higher and lower c-di-GMP
93 levels, respectively (Zhang *et al.*, 2018).

94 PlzA is classified as an xPilZ-PilZ domain protein, given its dual-domain structure. The binding
95 of c-di-GMP to the interdomain linker and C-terminal PilZ domain results in a conformational
96 change; locking the C-terminal PilZ and N-terminal xPilZ domains into rigid conformation,
97 which we hypothesized could permit DNA binding (Mallory *et al.*, 2016; Groshong *et al.*, 2021;
98 Singh *et al.*, 2021). While no canonical DNA binding domain is evident in the PlzA amino acid
99 sequence, our lab has discovered and characterized multiple nucleic acid binding proteins
100 possessing novel binding motifs (Babb *et al.*, 2006; Riley *et al.*, 2009; Jutras *et al.*, 2012;
101 Jutras *et al.*, 2013a; Jutras *et al.*, 2013b; Zhang *et al.*, 2018; Singh *et al.*, 2021). Here, we
102 report that PlzA can bind DNA and RNA directly downstream of the promoter and into the 5'
103 untranslated region (UTR) of the *glpFKD* operon and that binding to this region is c-di-GMP
104 dependent.

105 **RESULTS**

106 **PlzA binds *glpFKD* DNA and the interaction is c-di-GMP dependent**

107 Previous studies demonstrated that c-di-GMP and PlzA modulate expression of the glycerol
108 metabolism operon (Rogers *et al.*, 2009; He *et al.*, 2011; Caimano *et al.*, 2015; Zhang *et al.*,
109 2018). We hypothesized that PlzA might bind 5' of the *glpFKD* operon. To that end, we sought
110 to determine if PlzA could bind DNA between *glpFKD* and the upstream ORF BB_0239, which
111 encodes a deoxyguanosine/deoxyadenosine kinase. The transcriptional boundaries and
112 promoter elements of the *glpFKD* operon were previously mapped (Adams *et al.*, 2017; Grove
113 *et al.*, 2017). A fluorescently labeled DNA fragment was generated by PCR amplification from

114 a plasmid containing the 410 bp intergenic region between the upstream gene and the start
115 of *glpF* (BB_0240) and is denoted as *glpFKD*(-219) (Figure 1).

116 To assess if PlzA could bind this DNA, EMSAs with increasing concentrations of recombinant
117 PlzA were performed (Figure 2A). Binding was observed at higher micromolar concentrations,
118 as evidenced by the disappearance of free DNA. The experiment was repeated with added
119 non-specific competitor poly-dI-dC to a final concentration of 2.5 ng/μL. Poly-dI-dC did not
120 appreciably compete away the shifted protein-DNA complexes (Figure 2B).

121 Given that binding of c-di-GMP has been hypothesized to act as a switching mechanism for
122 PlzA function, we next investigated the effects of c-di-GMP on DNA binding. EMSA reaction
123 mixtures were supplemented with exogenous c-di-GMP to a final concentration of 100 μM.
124 This concentration was used to saturate all PlzA molecules present in the reaction to ensure
125 complete binding throughout electrophoresis. Saturating concentrations have similarly been
126 used in other studies with c-di-GMP binding proteins that also bind DNA (Wilksch *et al.*, 2011;
127 Yang *et al.*, 2013; Tan *et al.*, 2015; Zamorano-Sánchez *et al.*, 2015; Schäper *et al.*, 2017;
128 Skotnicka *et al.*, 2020). The EMSA with *glpFKD*(-219), poly-dI-dC, and increasing PlzA was
129 repeated with the supplemented c-di-GMP (Figure 2C). The addition of c-di-GMP enhanced
130 PlzA binding affinity for *glpFKD*(-219), resulting in binding at lower concentrations of protein,
131 and the formation of distinct protein-DNA complexes as indicated by the arrows. Further, as
132 PlzA concentration increased, the mobility of the shifted band decreased.

133 To further assess the requirement of c-di-GMP for PlzA DNA binding activity, a mutant PlzA
134 protein ($PlzA_{RD\text{-}RD}$) that cannot bind c-di-GMP was produced by site-directed mutagenesis.
135 This mutant replaced key arginine residues in the conserved c-di-GMP binding motif, RxxxR,
136 with aspartic acids (Ryjenkov *et al.*, 2006; Amikam and Galperin, 2006; Chou and Galperin,
137 2016). These mutations (R145D and R149D) have previously been shown to abolish c-di-
138 GMP binding (Freedman *et al.*, 2009; Mallory *et al.*, 2016; Singh *et al.*, 2021). The *glpFKD*(-

139 219) EMSAs were repeated with $PlzA_{RD-RD}$ (Figure 2D-F). Binding to *glpFKD*(-219) was not
140 observed under any tested conditions with $PlzA_{RD-RD}$.

141 To ensure binding was not due to the recognition of vector-derived sequences within the
142 probe, competition EMSAs were performed with unlabeled DNA competitors derived from
143 empty pCR 2.1 TA clones, or those containing the *glpFKD*(-219) sequence (Figure S1A). The
144 nonspecific DNA from the empty pCR 2.1 vector did not compete away $PlzA_{WT}$-*glpFKD*(-219)
145 complexes at a 20-fold higher concentration (Figure S1B, lanes 3-5). Meanwhile, as little as
146 6-fold molar excess of unlabeled *glpFKD*(-219) was able to compete with the probe, and a 10-
147 fold molar excess successfully competed away the $PlzA_{WT}$-*glpFKD*(-219) complexes (Figure
148 S1B, lanes 6-8). These results demonstrate PlzA preferentially bound the *glpFKD*-derived
149 sequences.

150 Noting that c-di-GMP enhanced DNA binding by PlzA yet free DNA decreased with purified
151 recombinant PlzA without adding exogenous c-di-GMP, we investigated the possibility that
152 PlzA was being purified with c-di-GMP already bound. Purified recombinant proteins were
153 prepared and assayed by liquid chromatography-tandem mass spectrometry. Extracted ion
154 chromatograms showed the presence of c-di-GMP in $PlzA_{WT}$ protein preparations but not in
155 $PlzA_{RD-RD}$ preparations (Figure 3). The concentration of c-di-GMP present in recombinant
156 $PlzA_{WT}$ was extrapolated from a c-di-GMP standard curve at 99.29 nM. A total of 31.7 μM of
157 protein had been analyzed, indicating that approximately 0.3% of the protein was bound with
158 c-di-GMP. Prior studies have reported dissociation constants of 1.25 to 6.25 μM of PlzA for c-
159 di-GMP, which are typical K_D values for PilZ receptors and c-di-GMP (Pitzer *et al.*, 2011;
160 Mallory *et al.*, 2016). The amount of co-purified c-di-GMP detected is lower than the K_D. This
161 is likely due to removal of some c-di-GMP during the protein purification process, as well as
162 *E. coli* maintaining low cellular concentrations of c-di-GMP (~40-80 nM) (Sarenko *et al.*, 2017;
163 Junkermeier and Hengge, 2023). Taken together, these data indicate that PlzA DNA binding
164 to this region of the *glpFKD* operon is specific and c-di-GMP dependent.

165 **c-di-GMP levels modulate PlzA DNA binding affinity**

166 It was previously identified with *glpFKD-gfp* promoter fusion studies that the minimal *glpFKD*
167 promoter sequence (-46 relative to the *glpFKD* transcription start site) with the full 195bp UTR
168 region was expressed at higher levels than the core promoter sequence alone (Grove *et al.*,
169 2017). Given these findings and our data that PlzA can bind DNA, we focused further analysis
170 of PlzA binding on a smaller sequence within that region that was found to be important for
171 regulation (Grove *et al.*, 2017; Savage *et al.*, 2018). A 42 bp DNA probe encompassing the -
172 7 to +35 sites, called *glpFKD*(-7/+35), was generated by annealing complementary
173 oligonucleotides with the forward oligonucleotide 5' conjugated to IRDye800 (Figure 4A). An
174 RNA probe conjugated to Alexa Fluor 488 was also made for subsequent RNA binding studies.

175 To determine how different c-di-GMP levels impact PlzA binding affinity, EMSAs with
176 increasing $PlzA_{WT}$ concentrations were performed at concentrations of added c-di-GMP
177 ranging from 0 to 100 μM in the reaction mixtures with the *glpFKD*(-7/+35) probe (Figure 4B).
178 Competitor poly-dI-dC was added at a final concentration of 2.5 ng/μL to reduce nonspecific
179 interactions. $PlzA_{WT}$ bound to *glpFKD*(-7/+35), and binding was enhanced by increasing the
180 concentration of c-di-GMP. In reactions with no added c-di-GMP, we observed that the amount
181 of free DNA decreased with increasing PlzA concentrations but did not detect a band
182 corresponding to the protein-DNA complex. The decrease in free DNA with no added c-di-
183 GMP is presumably due to the presence of co-purified c-di-GMP in recombinant PlzA. When
184 equimolar c-di-GMP and protein levels were tested, a discernable shift (formation of a new
185 band) was evident only at protein concentrations greater than 2.5 μM (purple arrow). As c-di-
186 GMP concentrations were further increased, shifts became discernable at lower protein
187 concentrations, with the strongest binding observed at c-di-GMP concentrations of 25, 50, and
188 100 μM. Further, as protein concentration was increased, we consistently observed
189 progressive super-shifting of complexes in a dose-dependent manner (Figure 4B, colored
190 arrows). At protein concentrations greater than or equal to 5 μM coupled with c-di-GMP
191 concentrations greater than 25 μM, complexes appeared to smear, suggesting the observed

192 super-shifting is due to complex dissociation. The stoichiometry of this complex(es) is under
193 investigation.

194 Next, we calculated the apparent binding affinity ($K_{D(app)}$) under these experimental conditions
195 for PlzA for *glpFKD*(-7/+35) at the various c-di-GMP concentrations (Figure 4C). The
196 calculations were performed under the assumption that a single molecule of PlzA binds to a
197 single molecule of DNA. As noted above, it is possible that the protein-nucleic acid ratio is not
198 1:1, which could affect the calculated values. The values described below are the $K_{D(app)}$ ±
199 SEM (standard error of the mean). Without any supplemented c-di-GMP, the extrapolated
200 $K_{D(app)}$ of $PlzA_{WT}$-*glpFKD*(-7/+35) was calculated as 7.86 (± 1.84 μM), however we again note
201 that some c-di-GMP is already present in the $PlzA_{WT}$ protein preparations (see above and
202 Figure 3). Binding affinity increased approximately 1.5-fold under equimolar c-di-GMP and
203 protein concentrations with a calculated $K_{D(app)}$ of 4.22 (± 0.43 μM). Under saturating c-di-GMP
204 conditions (exceeding the K_D of PlzA for c-di-GMP), binding affinity increased further.
205 Supplementing binding reactions with 10 μM c-di-GMP improved PlzA-DNA binding 2.5-fold
206 with a $K_{D(app)}$ of 2.81 (± 0.50 μM). An approximately 6- to 7-fold increase in binding affinity was
207 observed when the binding reaction was supplemented with either 25, 50, or 100 μM of c-di-
208 GMP resulting in calculated $K_{D(app)s}$ of 1.33 (± 0.12 μM), 1.25 (± 0.08 μM), and 1.09 μM (± 0.06
209 μM), respectively.

210 To determine if the calculated $K_{D(app)}$ values differed significantly, Welch's ANOVA followed by
211 a Dunnett's T3 multiple comparisons test was performed (Figure 4D). All calculated $K_{D(app)}$
212 values differed significantly ($p \leq 0.05$) from the 0 μM c-di-GMP $K_{D(app)}$ except for the equimolar
213 and 10 μM c-di-GMP conditions. Compared to the equimolar condition, the 25, 50, and 100
214 μM $K_{D(app)}$ values differed significantly ($p \leq 0.001$), while the 10 μM c-di-GMP $K_{D(app)}$ value did
215 not ($p > 0.05$). When comparing the 10, 25, 50, and 100 μM $K_{D(app)}$ values of $PlzA_{WT}$-*glpFKD*(-
216 7/+35) interaction, there were no significant differences ($p > 0.05$). The mutant $PlzA_{RD\text{-}RD}$ did
217 not bind *glpFKD*(-7/+35) under any tested conditions (Figure 4E). Overall, we conclude that
218 levels of c-di-GMP significantly impact PlzA-DNA binding.

219 **PlzA interacts predominantly with the major groove of DNA and sequence length**
220 **appears to play a role in DNA binding affinity**

221 We next sought to gain further information on the mode of binding between PlzA and dsDNA.
222 To explore these interactions, we performed binding assays with methyl green and
223 actinomycin D which bind the major and minor grooves of DNA, respectively (Kim and Nordén,
224 1993; Basu *et al.*, 2009; Liu *et al.*, 2016; Shi *et al.*, 2022). EMSAs were performed with
225 increasing concentrations of these drugs to determine if either could compete with $PlzA_{WT}$ for
226 binding to the *glpFKD*(-7/+35) probe (Figure 5). Methyl green was able to compete for binding
227 as evidenced by the increase in free DNA and by the disruption of PlzA-*glpFKD*(-7/+35)
228 complex formation at higher concentrations. Oppositely, competition with actinomycin D was
229 only evident at the highest concentrations tested. These results indicate that PlzA interacts
230 with dsDNA mainly through interactions with the major groove.

231 To gain insights on the DNA sequence requirements for PlzA binding, the *glpFKD*(-7/+35)
232 sequence was changed at several nucleotides (Figure 6A). Mutant competitors 1 and 2
233 targeted the middle region of the sequence containing one of the few GC rich regions. Mutant
234 competitors 3-5 swapped an ATTA repeat to CGGC at the 5', 3' or both ends to introduce
235 greater GC content and potentially disrupt symmetry. Mutant competitors 7-9 were designed
236 to disrupt potential secondary structural elements. Competition EMSAs with the IRDye labeled
237 *glpFKD*(-7/+35) probe and the unlabeled mutagenized competitors were performed with
238 $PlzA_{WT}$ (Figure 6B-E). Despite introducing various mutations into the *glpFKD*(-7/+35)
239 sequence, competition still occurred, although some competed slightly better than others. For
240 example, mutant sequence 4, altering an ATTA to CGGC at the 5' end, disrupted complex
241 formation less than altering an ATTA at the 3' end or mutating both. Our results suggest that
242 the *glpFKD*(-7/+35) sequence may be a non-consensus or low affinity DNA binding site of
243 PlzA within the *glpFKD* operon or that PlzA binds in a sequence independent manner. This 42
244 bp sequence was chosen as it is crucial for *glpFKD* regulation and is known to be bound by
245 other proteins (Grove *et al.*, 2017; Savage *et al.*, 2018; Saylor *et al.*, 2023; Zhang *et al.*, 2024).

246 We hypothesize that there could be higher affinity binding sites upstream of this operon and
247 within the *B. burgdorferi* genome.

248 Since we could not identify an obvious PlzA sequence motif, we next investigated if PlzA
249 bound DNA in a length-dependent mechanism. The *glpFKD*(-7/+35) sequence was truncated
250 by nine bases at the 5' end and then an additional three nucleotides at a time, and competition
251 EMSAs were performed with these DNAs (Figure 7A). The truncated *glpFKD*(-7/+35)
252 competitors exhibited less competition than the unlabeled specific competitor (Comp 1/C1) at
253 the same concentration (Figure 7B). This could indicate that the nucleotides removed from the
254 5' end are necessary for PlzA recognition or that PlzA has a higher affinity for longer pieces of
255 DNA. Studies are ongoing to determine how PlzA recognizes nucleic acids and to find
256 additional PlzA binding sites, as further elaborated upon in the discussion.

257 **PlzA also binds *glp*-derived single-stranded nucleic acids**

258 Several borrelial proteins that were initially identified as having affinities for DNA have
259 subsequently been found to also bind RNA (Jutras *et al.*, 2013a; Savage *et al.*, 2018;
260 Krusenstjerna *et al.*, 2023). Given the role the 5' UTR of the *glpFKD* operon plays in its
261 regulation, we hypothesized PlzA could potentially also bind RNA. RNA probes used were
262 designed to correspond to only transcribed regions (Figure 4A). The *glpFKD*(-7/+35)
263 counterpart RNA probe is designated *glpFKD*(UTR). EMSAs were performed with increasing
264 concentrations of $PlzA_{WT}$ with 0 μM, 100 μM, or equimolar c-di-GMP levels and labeled
265 *glpFKD*(UTR) (Figure 8). $PlzA_{WT}$ bound the RNA, which was enhanced by addition of c-di-
266 GMP (Figure 8A). The binding pattern to the *glpFKD*(UTR) RNA was comparable to the
267 *glpFKD*(-7/+35) DNA, with increased protein and c-di-GMP concentrations leading to an
268 increased size of the PlzA-RNA shifts.

269 EMSAs were repeated with the $PlzA_{RD\text{-}RD}$ mutant protein at 0 or 100 μM c-di-GMP (Figure 8B).
270 The $PlzA_{RD\text{-}RD}$ protein did not bind *glpFKD*(UTR) RNA at any tested protein concentration,
271 regardless of whether c-di-GMP was included.

272 Given that PlzA can bind to both the DNA and RNA, we sought to determine if it could also
273 bind the cognate ssDNA corresponding to the coding strand. This is relevant as ssDNA is
274 exposed at open transcriptional complexes. The forward, IRDye labeled oligonucleotide used
275 to generate the *glpFKD*(-7/+35) dsDNA probe was used as a single-stranded DNA probe
276 (ss*glpF*) in EMSAs. No evident complexes were formed when exogenous c-di-GMP was
277 excluded from EMSA reactions, but free ssDNA did decrease at the highest tested PlzA
278 concentrations of 3.5-7.5 μM (Figure 8C right panel). Adding 100 μM c-di-GMP resulted in the
279 formation of PlzA-ssDNA complexes but at higher concentrations of protein than required to
280 achieve complex formation with dsDNA and RNA (Figure 8C left panel).

281 We next quantified the binding affinity of PlzA for the *glp* derived RNA and ssDNA sequences
282 through densitometry of the EMSA shifts, and the apparent $K_{D(app)s}$ calculated as described
283 above (Figure 4D). As with dsDNA, the binding affinity of PlzA for RNA was weakest when no
284 exogenous c-di-GMP was supplemented with a calculated $K_{D(app)}$ of 7.64 (± 1.85 μM). Binding
285 improved approximately 2-fold in equimolar c-di-GMP conditions with a $K_{D(app)}$ of 4.14 (± 0.74
286 μM). The $K_{D(app)}$ was 0.95 (± 0.11 μM) when the c-di-GMP concentration was 100 μM, an 8-
287 fold increase in PlzA binding affinity for the *glpFKD*(UTR) RNA as compared to when no c-di-
288 GMP was added. There was a significant difference between the $PlzA_{WT}$-*glpFKD*(UTR) RNA
289 $K_{D(app)}$ value calculated at 100 μM c-di-GMP and both the equimolar and 0 μM c-di-GMP values
290 ($p \leq 0.01$) (Figure 9A). No significant differences were detected between the $K_{D(app)}$ values
291 calculated at 0 μM and equimolar c-di-GMP levels ($p > 0.05$).

292 Although PlzA bound the ss*glpF* probe, it had a substantially lower affinity for ssDNA than
293 dsDNA or RNA (Figure 8D). The $K_{D(app)}$ of the $PlzA_{WT}$-ss*glpF* interaction was 8.74 (± 1.86 μM)
294 with 0 μM c-di-GMP and 4.42 (± 0.90 μM) with 100 μM c-di-GMP. These values were found to
295 be statistically significant from one another ($p < 0.05$) (Figure 9B). Only a 2-fold increase in
296 binding affinity of PlzA for the single-strand DNA was observed when the c-di-GMP
297 concentration was increased to 100 μM, whereas the same concentration of c-di-GMP
298 increased PlzA binding affinity 7-fold for dsDNA and 8-fold for RNA.

299 Lastly, we also compared the binding affinities of PlzA for all three tested nucleic acids (Figure
300 9C). The calculated $K_{D(app)}$ values of $PlzA_{WT}$ for each nucleic acid in the presence of 100 μM
301 c-di-GMP were compared with a Welch's ANOVA and Dunnett's T3 multiple comparisons test.
302 No significant difference ($p > 0.05$) was detected between the PlzA $K_{D(app)}$ values for dsDNA
303 (1.09 ± 0.06 μM) and RNA (0.95 ± 0.11 μM), indicating that PlzA binds the *glpFKD* dsDNA
304 and RNA with similar affinity. Both the dsDNA and RNA $K_{D(app)s}$ were significantly different from
305 the calculated ssDNA $K_{D(app)}$ (4.42 ± 0.90 μM) ($p \leq 0.01$), suggesting that PlzA prefers *glp*-
306 derived dsDNA and RNA over ssDNA. All the binding data presented throughout the
307 manuscript are summarized in Table 1. Collectively, our data indicate that PlzA binding activity
308 is c-di-GMP dependent for all tested *glpFKD* nucleic acids.

309 **Dissection of PlzA reveals the nucleic acid binding domain**

310 The first discovered c-di-GMP binding receptors were the PilZ domain-containing proteins
311 (Ryjenkov *et al.*, 2006; Amikam and Galperin, 2006; Galperin and Chou, 2020). The PilZ
312 domain consists of two conserved c-di-GMP binding motifs, RxxxR and [D/N]ZSXXG
313 (Ryjenkov *et al.*, 2006; Amikam and Galperin, 2006; Chou and Galperin, 2016; Cheang *et al.*,
314 2019). The recently solved crystal structure of *B. burgdorferi* PlzA revealed a unique dual-
315 domain topology consisting of an amino-terminal PilZ-like domain, called PilZN3, which is
316 connected to the carboxy-terminal PilZ domain via a linker domain that contains the RxxxR c-
317 di-GMP binding motif (Singh *et al.*, 2021). The first 141 amino acid residues make up the N-
318 terminus of PlzA ($PlzA_{NTD}$), while residues 142-261 make up the C-terminal ($PlzA_{CTD}$) PilZ
319 domain. Having determined that PlzA is a c-di-GMP-dependent DNA and RNA binding protein,
320 we next sought to determine which domain of PlzA binds nucleic acid. To generate
321 recombinant $PlzA_{NTD}$ and $PlzA_{CTD}$ domains, truncated *plzA* genes called *plzA*-NTD and *plzA*-
322 CTD were cloned to encode the residues described above, which correspond to each
323 respective domain (Figure 10A).

324 EMSAs were performed with labeled *glpFKD*(-7/+35) DNA probe and increasing
325 concentrations of $PlzA_{NTD}$ or $PlzA_{CTD,}$ with or without 100 μM c-di-GMP. $PlzA_{NTD}$ bound DNA

326 independently of c-di-GMP, as supplementation with c-di-GMP did not affect binding (Figure
327 10B). $PlzA_{CTD}$ contains the motifs required for c-di-GMP binding, but the $PlzA_{CTD}$ domain did
328 not bind to DNA with or without added c-di-GMP (Figure 10C). It was also determined that
329 $PlzA_{NTD}$ could bind *glpFKD*(UTR) RNA with or without c-di-GMP, while $PlzA_{CTD}$ does not
330 (Figure 10D and E respectively).

331 We also performed the LC/MS-MS analysis on $PlzA_{NTD}$ and $PlzA_{CTD}$ recombinant proteins to
332 determine if any di-nucleotides were present (Figure 11A). Co-purified c-di-GMP was detected
333 in $PlzA_{CTD}$ but not $PlzA_{NTD}$. The concentration of c-di-GMP present in recombinant $PlzA_{CTD}$ was
334 determined as 704.29 nM. A total of 66.7 μM of protein was sent for analysis indicating
335 approximately 1.1 % of $PlzA_{CTD}$ was bound with c-di-GMP. This indicates that $PlzA_{CTD}$ retained
336 the ability to bind c-di-GMP but cannot bind nucleic acid.

337 To confirm that the truncated proteins were stably folded, circular dichroism (CD) was
338 performed, and the CD spectra were obtained for $PlzA_{WT}$, $PlzA_{RD\text{-}RD}$, $PlzA_{NTD}$, and $PlzA_{CTD}$
339 (Figure 11B). The CD results show that the mutated/truncated proteins retained stable
340 structures matching that of $PlzA_{WT}$. Therefore, the lack of nucleic acid binding by $PlzA_{RD\text{-}RD}$
341 and $PlzA_{CTD}$ was not due to improper folding. Secondary structure analysis was performed via
342 the BeStSel server to further compare the recombinant proteins to the crystallized PlzA
343 structure (Micsonai *et al.*, 2015; Micsonai *et al.*, 2018; Micsonai *et al.*, 2022). All tested
344 recombinant proteins show some differences from the calculated secondary structures of the
345 7MIE PDB file. These deviations are primarily in the percentage of beta-sheet and turn
346 secondary structure present. We note that the His-Tag was not cleaved prior to CD analysis
347 of the recombinant proteins, and the resultant deviations could be due to the tag. The tag was
348 considered in all calculations for CD. Overall, we can deduce that protein folding is not the
349 cause of the lack of binding by $PlzA_{RD\text{-}RD}$ or $PlzA_{CTD}$.

350 **Computational docking of PlzA-nucleic acid interactions**

351 To gain a better understanding of the interactions between PlzA and DNA or RNA, we
352 performed a computational docking analysis with the HDOCK server (Huang and Zou, 2014;
353 Yan *et al.*, 2017a; Yan *et al.*, 2017b; Yan *et al.*, 2020). The *glpFKD*(-7/+35) DNA and
354 *glpFKD*(UTR) RNA sequences were input as ligands, while the PlzA crystal structure (PDB
355 ID: 7mie) was input as the receptor molecule. The modeling results indicate that interactions
356 between PlzA and *glpFKD*(-7/+35) DNA are facilitated primarily through the N-terminal domain
357 (Figure 12A). In the representative docking data, beta-strands six and seven of the N-terminus
358 are extended and protrude outward, intercalating a major groove of the DNA 5' to the protein
359 docking site (Figure 12A top panel). Polar and positively charged amino acids, including N110
360 and K111, are located at the loop region where beta-strands 6 and 7 converge, with the
361 potential to directly interact with bases of the DNA (Rooman *et al.*, 2002; Sathyapriya and
362 Vishveshwara, 2004). Interactions with a minor groove of the DNA are predicted to occur
363 through the loop region of beta-strands 2 and 3 and an unstructured linker region (Figure 12a
364 bottom panel). Several positively charged residues, including arginine (R32 and R38) and
365 lysine (K81), as well as an aromatic phenylalanine residue (F45), are predicted to directly
366 interact with the minor groove and backbone of the DNA. An additional lysine residue K87,
367 which is in the unstructured region, protrudes and interacts directly with bases of another major
368 groove of *glpFKD*(-7/+35). No residues of the C-terminal domain are predicted to interact
369 directly with the DNA; however, four residues are interfacing the DNA (K161, I213, D258, and
370 N261). Residue K161 hovers over the DNA backbone potentially interacting nonspecifically
371 via electrostatic interactions. We note that the docking analysis produced several potential
372 models that differed in the location of where residues interact with either the minor or major
373 groove. The model presented is the highest-ranking model. Overall, the computational docking
374 data supports our EMSA data showing that $PlzA_{NTD}$ binds DNA while $PlzA_{CTD}$ does not (Figures
375 10B and C) and that PlzA interacts predominately with the major groove of DNA (Figure 5).

376 The docking analysis was also performed with $PlzA_{WT}$ and *glpFKD*(UTR) RNA (Figure 12B).
377 Again, the N-terminal domain of PlzA facilitates most of the interactions. Similar regions of the

378 protein that interacted with dsDNA are implicated in the interaction with RNA including the
379 residues of the loop region of beta-strands 6 and 7 (K111 and Y113), the loop region of beta-
380 strands 2 and 3 (R38, Y42, and F45) and the unstructured linker region (R32, D85, K87, F92,
381 and F95). Unlike the docking model with dsDNA, the PlzA-RNA docking model indicates more
382 involvement of the C-terminal domain in RNA binding. Several residues of alpha-helixes 5
383 (D155, R158, K159, and K161) and 6 (D258 and Y259) in the C-terminal domain of PlzA, are
384 predicted to interact with RNA. Although our EMSA results show that $PlzA_{CTD}$ does not bind
385 any tested nucleic acids, the docking results indicate it could help stabilize $PlzA_{NTD}$-nucleic
386 acid interactions or could be involved in target recognition. Guided by the docking data, we
387 are in the process of generating site-direct PlzA mutant proteins that will explore residues
388 predicted to be involved in nucleic-acid binding.

389 Computational docking was attempted with PlzA and ss*glpF*, but the HDOCK server could not
390 complete the analysis. Given that PlzA bound dsDNA and RNA better than ssDNA, it is
391 interesting to note that the computational modeling predicts the RNA to fold into a double
392 stranded secondary structure. Combined with our EMSA data, this suggests that PlzA could
393 have a preference towards double stranded nucleic acids and/or secondary structures.

394 **DISCUSSION**

395 Cyclic-di-GMP has wide-ranging effects on many bacterial processes. The responses to c-di-
396 GMP are commonly mediated through c-di-GMP binding effector proteins. PlzA is the only
397 chromosomally and universally encoded c-di-GMP binding protein of the Lyme disease
398 spirochetes. It has been shown to be crucial for borrelial infection processes and in the
399 regulation of the metabolism of alternative carbon sources during colonization of vector ticks
400 (Pappas *et al.*, 2011; Pitzer *et al.*, 2011; He *et al.*, 2014; Mallory *et al.*, 2016; Zhang *et al.*,
401 2018; Groshong *et al.*, 2021). Despite its singularity and importance to *B. burgdorferi*, little was
402 previously known about the biochemical functions of PlzA outside of binding c-di-GMP. Here,
403 we provide evidence that PlzA is a novel c-di-GMP-dependent nucleic acid-binding protein.

404 While PlzA lacks an obvious canonical DNA binding motif, we hypothesized that PlzA might
405 be a nucleic acid binding protein given its role in the regulation of the *glpFKD* operon. Binding
406 was only observed by the wild-type PlzA protein but not the mutant, $PlzA_{RD\text{-}RD}$, which is
407 incapable of binding c-di-GMP. It has previously been shown that structural rearrangements
408 are induced upon the binding of c-di-GMP to PlzA (Mallory *et al.*, 2016; Groshong *et al.*, 2021).
409 Without bound c-di-GMP, The N- and C-terminal domains of PlzA can flex on an unstructured
410 linker, potentially inhibiting the stable interactions that are necessary for DNA and RNA
411 binding. This hypothesis is supported by the fact that crystallization of PlzA was only achieved
412 when saturated with c-di-GMP (Singh *et al.*, 2021). We surmise that binding of c-di-GMP to
413 the C-terminal PilZ domain results in a change of PlzA into a more rigid structure, which
414 permits nucleic acid binding through residues of the N-terminal domain.

415 We identified the N-terminal domain of PlzA as the nucleic acid binding domain and that
416 $PlzA_{NTD}$ bound DNA independently of c-di-GMP. In contrast, the C-terminal domain, $PlzA_{CTD}$,
417 did not bind any tested nucleic acids, although it still bound c-di-GMP. While $PlzA_{NTD}$ can
418 interact with nucleic acids, these interactions were not as strong as $PlzA_{WT}$. Binding of nucleic
419 acid by $PlzA_{NTD}$ resulted in smeared shifts that lacked defined bands. This suggests instability
420 of the $PlzA_{NTD}$-nucleic acid complexes resulting in dissociation of the complexes during
421 electrophoresis. Noting that binding of c-di-GMP appears to stiffen interactions between the
422 N- and C-terminal domains, we hypothesize the C-terminal domain may stabilize the
423 interactions of $PlzA_{NTD}$ with nucleic acids when c-di-GMP is bound. On the other hand, we
424 posit that the flexibility of apo-PlzA results in the C-terminal domain disrupting the ability of the
425 N-terminal domain to stably bind DNA. Although the N- and C- terminal PlzA domains are
426 similar, structural comparisons of $PlzA_{NTD}$ and $PlzA_{CTD}$ reveal distinct differences (Figure 10A).
427 The N-terminus of $PlzA_{NTD}$ contains an alpha-helix that is absent in the N-terminus of $PlzA_{CTD}$.
428 Additionally, an alpha-helix is present at the beta-barrel pore of $PlzA_{NTD}$ but not in the $PlzA_{CTD}$.
429 These differences suggest distinct roles in PlzA-nucleic acid binding dynamics. We are

430 currently exploring the sequence motifs recognized by PlzA and the exact amino acid residues
431 involved in nucleic acid binding.

432 Previous studies found that PlzA is a monomer in solution regardless of c-di-GMP (Mallory *et*
433 *al.*, 2016). While many DNA-binding proteins bind DNA as dimers, alternative binding models
434 exist (Kim and Little, 1992; Kohler *et al.*, 1999). It has been shown that some proteins are
435 capable of binding sequentially as monomers which then multimerize on the DNA via protein-
436 protein interactions, such as some members of the Leucine zipper and helix-loop-helix zipper
437 families, the LexA repressor of *E. coli*, and BpaB of *B. burgdorferi* (Kim and Little, 1992; Kohler
438 *et al.*, 1999; Burns *et al.*, 2010). Our studies revealed that increasing c-di-GMP and $PlzA_{WT}$
439 concentrations resulted in progressively slower migration of PlzA-*glpFKD*(-7/+35) DNA
440 complexes. The decreases in the migration of the complexes that are observed with increasing
441 protein concentration may indicate a stoichiometry that is not 1:1, or the observed smearing
442 may be the result of dynamic equilibrium of the complexes in the gels during electrophoresis
443 (Fried and Crothers, 1981; Fried, 1989). While increasing c-di-GMP enhances PlzA binding,
444 there could be a threshold where too much c-di-GMP is inhibitory and causes dissociation of
445 PlzA from nucleic acids. A dimerization or protein-protein interaction residue/motif/domain
446 remains to be identified in PlzA, and work is ongoing to address the DNA-binding mode and
447 stoichiometry of PlzA-nucleic acid complexes.

448 We also observed probe signal in the gel wells at high protein and c-di-GMP concentrations.
449 This could be due to protein aggregation. Rather than simply being a negative consequence
450 of environmental stress, recent studies indicate a role for protein aggregation in regulation.
451 Specifically, certain conditions can induce a protein to aggregate, and thus be sequestered,
452 altering gene expression to permit bacterial adaptation to a specific environment (Schramm *et*
453 *al.*, 2019). Whether high levels of c-di-GMP promote such processes remains to be tested, but
454 c-di-GMP levels certainly alter PlzA binding affinity.

455 Our study sought to identify if PlzA can bind nucleic acids and what role c-di-GMP plays in
456 that function. The data are conclusive that PlzA indeed binds DNA, both double and single-
457 stranded, and RNA, with a preference for dsDNA and RNA. PlzA also displayed higher affinity
458 for certain nucleic acid types over others, and PlzA-*glp* DNA binding was competed away by
459 specific competitors but not poly-dI-dC. Further, the nucleic acid binding at the *glpFKD* locus
460 by PlzA is dependent on c-di-GMP. A caveat of the data is that the calculated dissociation
461 constants for PlzA-nucleic acids are lower than typically observed for protein-nucleic acid
462 interactions. The higher dissociation constants could be attributed to our method of choice,
463 EMSA, as some potential dissociation and protein aggregation were observed. Attempts were
464 made with various alterations aimed to improve the stability of the complexes during
465 electrophoresis, such as altering the gel percentage, pH of the running buffer, salt/ion
466 concentrations of the reaction buffer, and addition of detergents to the reaction (Hellman and
467 Fried, 2007). Altering the pH of the running buffer to 8.0 decreased the aggregation in the
468 wells but smearing still occurred at higher protein and c-di-GMP concentrations (data not
469 shown). The experimental conditions presented and reported here yielded the most consistent
470 results.

471 PlzA bound to a sequence adjacent to the promoter and extending 35 bp into the 5' UTR of
472 *glpF*. This site has been shown to positively affect *glpFKD* transcript levels and is a site where
473 two other borrelial regulators, SpoVG and BadR, also bind (Grove *et al.*, 2017; Savage *et al.*,
474 2018; Saylor *et al.*, 2023; Zhang *et al.*, 2024). BadR, the *Borrelia* Host Adaption Regulator,
475 has recently been identified as a repressor of *glpFKD*, while PlzA has previously been shown
476 to exert both positive and negative effects on *glpFKD*, at high and low levels of c-di-GMP
477 respectively (Zhang *et al.*, 2018; Zhang *et al.*, 2024). Although SpoVG can bind the DNA and
478 RNA of *glpFKD*, the transcriptional consequences of these interactions are unknown (Savage
479 *et al.*, 2018; Saylor *et al.*, 2023). How these proteins work in tandem to regulate glycerol
480 catabolism is currently being investigated. Multi-layered regulation of glycerol catabolism is

481 found in other bacteria, indicating this is not a phenomenon unique to *Borrelia* (Bong *et al.*,
482 2019).

483 Our data show that PlzA is a DNA and RNA binding protein that is c-di-GMP dependent.
484 Obvious questions that remain are what sequence motif(s) PlzA recognizes and what
485 additional binding sites may exist in the *B. burgdorferi* genome. Our studies were unable to
486 identify a sequence motif in the *glpFKD* intergenic region but were able to identify that PlzA
487 interacts mainly with the major groove of DNA. We also saw a preference of PlzA for longer
488 sequences and suggest PlzA may recognize gene targets in a sequence independent manner.
489 The lesser affinity for shorter sequences could indicate that PlzA might recognize structured
490 nucleic acids, as shorter DNA is less likely to form structured conformations. Given that poly-
491 dI-dC did not abolish binding, PlzA likely exhibits preferential affinity to certain sequences.
492 Whether the preferential affinity is due to a sequence motif or structural motifs, remains to be
493 determined. What is clear, from multiple studies, is that the cellular effects exerted by PlzA
494 are specific.

495 A logical step to address these questions is to perform chromatin immunoprecipitation with
496 sequencing (ChIP-Seq). Given that PlzA binding is reliant on an effector molecule, c-di-GMP,
497 there are some caveats in performing this assay at this stage. First, the activating stimulus for
498 the Hk1/Rrp1 pathway is currently unknown. As a result, it is unknown how levels of c-di-GMP
499 fluctuate in culture medium and across growth phases, and therefore how much c-di-GMP is
500 produced in a bacterial cell at a given time. This can affect the ability to obtain consistent ChIP
501 results. The tendency of protein to occupy each site will likely vary with culture conditions.
502 Therefore, optimal experimental conditions must be determined before obtaining and drawing
503 conclusions about PlzA binding sites derived from ChIP experiments.

504 Next, as mentioned above, two other proteins have been found to bind the same site within
505 *glpFKD*, SpoVG and BadR(Savage *et al.*, 2018; Saylor *et al.*, 2023; Zhang *et al.*, 2024).
506 Competition between these proteins could also impact the ChIP data. Additionally, the

507 molecular consequence of nucleic acid binding remains to be determined. We are actively
508 working to address these questions by characterizing several PlzA site-direct mutant proteins
509 incapable of binding nucleic acids to determine impacts on *B. burgdorferi* physiology and gene
510 regulation. Such work will also aid in identifying targets of PlzA regulation to better guide the
511 identification of additional binding sites.

512 During the preparation of this manuscript, Van Gundy et al. published a report identifying RNA
513 chaperone activity by PlzA (Van Gundy *et al.*, 2023). The authors concluded that this activity
514 is independent of c-di-GMP, which deviates from our results showing that c-di-GMP is required
515 for nucleic acid binding. Those authors acknowledged this and suggested that differences in
516 methodologies and conditions might have contributed to the differences in observed results.
517 Some of these differences include using non-borrelial derived RNAs, a different PlzA site-
518 mutant, equimolar concentrations of c-di-GMP, and filter binding assays used to calculate
519 dissociation constants.

520 We chose to evaluate saturating c-di-GMP concentrations to evaluate the maximal nucleic
521 acid binding activity of holo-PlzA and mitigate potential variability in the calculated apparent
522 K_D values. Van Gundy et al. did not assess whether their PlzA protein preparations co-purified
523 with c-di-GMP, which could have contributed to their observed binding by apo-PlzA. Unlike
524 their study, we did not observe binding by the $PlzA_{RD\text{-}RD}$ mutant, which cannot bind c-di-GMP,
525 under any tested condition. Given the stabilizing effect c-di-GMP binding has on the structure
526 of PlzA (only the holo form was able to be crystallized), the ability of their mutant and apo-
527 PlzA to bind RNA merits further mechanistic investigation. Lastly, the goals for our studies
528 differed greatly as they sought to study RNA chaperone activity, while we set out to investigate
529 the ability of PlzA to bind nucleic acids. Our two studies, however, corroborate that PlzA can
530 bind RNA and suggest that this function plays a role in regulating borrelial gene expression.
531 We agree that further investigation is required to unravel the molecular mechanisms of PlzA
532 interactions with nucleic acids, and the roles that those interactions play in gene regulation in
533 this important pathogen.

534 It is known that the affinity and specificity of proteins for target nucleic acids are impacted by
535 co-effectors (Carey, 2022). We observed that adding exogenous c-di-GMP increased binding
536 affinity of PlzA for *glp*-derived DNA and RNA. The localization of c-di-GMP and associated
537 proteins in *B. burgdorferi* is currently unknown. Whether the c-di-GMP signaling network is
538 global or localized will be key to understanding how c-di-GMP induces PlzA effector functions.
539 This is actively being investigated by our laboratory. Studies in other bacteria with multiple
540 PilZ domain proteins have suggested that the differential affinity of the various PilZ receptors
541 for c-di-GMP sequentially regulates the processes that respond to c-di-GMP (Pultz *et al.*, 2012;
542 Junkermeier and Hengge, 2023). Given that PlzA is the only PilZ domain protein in most *B.*
543 *burgdorferi* isolates, an alternative to this could be that cellular levels of c-di-GMP alter the
544 affinity and/or specificity of PlzA.

545 We propose a model in which c-di-GMP availability alters the functionality of PlzA, and thus,
546 the targets regulated by PlzA could depend on the c-di-GMP levels available at different stages
547 of the enzootic life cycle (Figure 13). During vertebrate infection, no c-di-GMP is produced,
548 favoring apo-PlzA functions such as RNA chaperone activity (Van Gundy *et al.*, 2023). When
549 c-di-GMP is present, holo-PlzA can also anneal RNAs and potentially interact with RNAP
550 (Grassmann *et al.*, 2023; Van Gundy *et al.*, 2023). At lower levels of c-di-GMP, PlzA nucleic-
551 acid binding affinity is low for target loci. As c-di-GMP levels increase further, binding affinity
552 for target nucleic acids is increased and potential multimerization of PlzA molecules on DNA
553 could occur.

554 In conclusion, PlzA is a c-di-GMP-dependent nucleic acid-binding protein. While other c-di-
555 GMP binding receptors have been identified as DNA-binding proteins, PlzA is the first that has
556 been demonstrated to also bind RNA, thus expanding the functional diversity of c-di-GMP
557 binding proteins. Given that PlzA can bind to both DNA and RNA, it could have a role in
558 coordinating transcriptional complexes, and it supports the notion that PlzA can interact with
559 RNAP (Grassmann *et al.*, 2023). Our study examined PlzA nucleic acid binding at a locus
560 known to be affected by PlzA, the *glpFKD* catabolism operon, but sites of higher affinity could

561 exist. Our work identifying PlzA as a novel nucleic acid-binding protein provides a basis for its
562 functional mechanism. This will further inform our understanding of how *B. burgdorferi*
563 regulates gene expression in response to environmental cues and help unravel the PlzA
564 regulon.

565 **MATERIAL AND METHODS**

566 **Routine manipulations and molecular cloning**

567 Oligonucleotides used in this study are listed in Table 2 and were ordered from Integrated
568 DNA Technologies (IDT). *B. burgdorferi* genomic DNA was purified utilizing the E.Z.N.A Tissue
569 DNA kit (OMEGA). Standard and high-fidelity PCR were performed with 2X DreamTaq Green
570 PCR Master Mix (Thermo Scientific) or Q5 High-Fidelity 2X Master Mix (NEB), respectively.
571 Plasmids were isolated using QIAprep Spin Miniprep kits (Qiagen) according to the Miraprep
572 protocol (Pronobis *et al.*, 2016). All plasmid constructs were transformed into chemically
573 competent *E. coli* Top10 (Invitrogen) or DH5α (Invitrogen) strains for cloning and plasmid
574 maintenance. Positive clones were identified by colony PCR screening with the appropriate
575 primer sets, and DNA sequencing was performed by Eurofins Genomics LLC (Louisville, KY).

576 **Plasmid construction**

577 The wild-type (WT) *plzA* gene from *B. burgdorferi* strain B31, was cloned between the
578 NdeI/XhoI sites of pET28a(+) and thus fused in-frame with an amino-terminal 6xHis-tag.
579 Truncated *plzA* genes, named *plzA*-NTD and *plzA*-CTD encoding the N-terminal (residues 1-
580 141) and C-terminal (residues 142-261) domains of PlzA, respectively, were generated via
581 gene synthesis and cloned to produce 6xHis-tagged proteins by Genscript (Piscataway, NJ).
582 Briefly, the synthesized truncations were cloned into the pET28b(+) vector with *plzA*-NTD
583 cloned between the BamHI/NotI sites producing an amino-terminal 6xHis-tag fusion, while
584 *plzA* CTD was cloned between the NcoI/XhoI sites producing a carboxy-terminal 6xHis-tag
585 fusion.

586 **Generation of the site-directed PlzA R145D-R149D mutant**

587 Site-directed mutagenesis by overlap extension PCR was used to generate a PlzA mutant that
588 cannot bind c-di-GMP. Two-point mutations were introduced into a pTrcHis TOPO (Invitrogen)
589 construct containing the *plzA* gene. The successful introduction of the mutations was
590 confirmed by DNA sequencing. The site-directed PlzA double mutant is denoted as $PlzA_{RD\text{-}RD}$.
591 The oligonucleotide primers used for site-directed mutagenesis are listed in Table 2.

592 **Recombinant protein overexpression and purification**

593 For recombinant protein expression, *E. coli* Rosetta 2(DE3)(pLysS) (Novagen) chemically
594 competent cells were used for $PlzA_{WT}$, $PlzA_{RD\text{-}RD}$, and N-terminal PlzA domain ($PlzA_{NTD}$), while
595 One Shot BL21 Star (DE3) cells were used for the C-terminal PlzA domain ($PlzA_{CTD}$).
596 Overnight bacterial cultures were diluted 1:100 into Super broth (tryptone 32 g, yeast extract
597 20 g, NaCl 5 g per liter) supplemented with the appropriate antibiotics (kanamycin-50 µg/ml,
598 carbenicillin-100 µg/ml, and/or chloramphenicol-30 µg/ml). The cultures were allowed to grow
599 to an OD_{600} of 0.5-1.0 and protein expression was subsequently induced by the addition of
600 0.25 mM ($PlzA_{NTD}$ only) or 0.5 mM isopropyl-β-D-thiogalactopyranoside (IPTG). Induced
601 cultures were incubated at either 37°C for 3-4 h ($PlzA_{WT}$ and $PlzA_{CTD}$), 29°C for 4-6 h ($PlzA_{WT}$),
602 or overnight at room temperature ($PlzA_{RD\text{-}RD}$ and $PlzA_{NTD}$). The cells were harvested by
603 centrifugation at 5400 x *g* for 30 min at 4°C, and the cell pellets were frozen at -80°C until
604 protein purification.

605 For protein extraction, cell pellets were thawed on ice and resuspended in equilibration buffer
606 (20 mM sodium phosphate, 300 mM NaCl, 10 mM imidazole, pH 8.0) and lysed via sonication.
607 Lysates were clarified by centrifugation at 23700 x *g* for 20 min at 4°C, and the supernatant
608 was retained. When further clarification was required due to lysate viscosity, the supernatants
609 were passed through a sterile 0.22 µM Millex-GS Syringe Filter Unit (MilliporeSigma). Protein
610 purification was performed via column affinity chromatography using either HisPur Ni-NTA
611 resin (Thermo Scientific) or the Magne-His purification system (Promega Corporation) for

612 $PlzA_{RD-RD}$ per manufacturer protocol. Eluted protein solutions were dialyzed overnight into
613 either EMSA buffer (50 mM Tris-HCl [pH 7.5], 50 mM KCl, 1 mM dithiothreitol (DTT), 1 mM
614 ethylenediaminetetraacetic acid (EDTA) [pH 8.0], 1 mM phenylmethanesulfonyl fluoride
615 (PMSF), 10% glycerol (v/v), and 0.01% Tween-20) or 20 mM sodium phosphate, pH 7.2. The
616 dialyzed proteins were then concentrated using Amicon Ultra 3K ($PlzA_N$ and $PlzA_C$) or 10K
617 ($PlzA_{WT}$ and $PlzA_{RD-RD}$) centrifugal filter units (MilliporeSigma). Aliquots of dialyzed proteins
618 were assessed for purity by SDS-PAGE and Coomassie brilliant blue staining. Final
619 recombinant protein concentrations were determined using Quick Start Bradford Protein
620 Assays (Bio-Rad) with bovine serum albumin as the reference protein for standard curves.
621 Purified protein aliquots were stored at -80°C.

622 **Nucleic acid fragments design and generation**

623 All probe sequences were derived from the genomic sequence of *B. burgdorferi* type strain
624 B31, and the corresponding primers are reported in Table 2 (Fraser *et al.*, 1997; Casjens *et*
625 *al.*, 2000). Large and small DNA substrates were used in this study. Larger fluorescently
626 tagged fragments >60 bp were generated by PCR amplification of the target sequences from
627 B31 genomic DNA and subsequent cloning into the pCR 2.1 TA vector (Invitrogen) following
628 TOPO TA protocols. Subsequently, target-specific or M13 F and R IRDye 800-labeled primers
629 were used for PCR of the plasmid templates to produce labeled DNA fragments for binding
630 studies. All PCR-generated probes were treated with exonuclease I (NEB) according to the
631 manufacturer's protocol to remove single-stranded DNA, followed by ethanol precipitation.
632 The DNA was resuspended in molecular-grade water (Ambion) and quantified by Nanodrop
633 UV spectrometry.

634 Smaller probes <60 bp long were produced by annealing complementary unlabeled and IRDye
635 800-labeled oligonucleotides at equimolar concentrations at 95°C for 5 min and were gradually
636 cooled overnight at RT. The same annealing procedure was done for unlabeled competitors.
637 All probes and competitors were aliquoted and stored at -20°C until further use. Unlabeled

638 and 5' IRDye 800 (LI-COR Biosciences) labeled oligonucleotides were synthesized by IDT.
639 Alexa Fluor 488 conjugated RNA probes corresponding to the transcribed regions of the
640 respective DNA probes were also synthesized by IDT.

641 **Electrophoretic mobility shift assays (EMSA)**

642 EMSA reactions were performed in EMSA buffer at RT with a final concentration of dsDNA
643 probe of 10 nM or RNA probe of either 10 nM or 50 nM. EMSA reaction mixtures were
644 supplemented with 3'5' c-di-GMP (Sigma-Aldrich) at various concentrations ranging from 0 to
645 100 μM. c-di-GMP was not added to the gel or running buffer. Experiments using equimolar
646 concentrations of c-di-GMP are relative to the final protein concentration in the reaction.
647 Proteins were co-incubated with c-di-GMP for 5 min before the addition of any nucleic acids.
648 Prior to the addition of labeled probe, the non-specific competitor poly-dI-dC (Roche) was
649 added to the EMSA reactions at a final concentration of 2.5 ng/μL, and the reaction mix was
650 allowed to incubate another 5 min (Larouche *et al.*, 1996). Labeled probes were added last,
651 and the reactions were subsequently incubated for an additional 10 min prior to loading onto
652 gels. Competition EMSAs were conducted similarly but with the unlabeled, specific competitor
653 added after poly-dI-dC and then allowed to incubate with the reaction prior to the addition of
654 the labeled probe. Regarding RNA EMSAs, when RNase contamination was evident in
655 recombinant protein preparations, RiboGuard (Lucigen) or SUPERase•In (Ambion) RNase
656 inhibitor was added to the EMSA mixtures before the addition of RNA to a final concentration
657 of 4 U/μL and 2 U/μL respectively to inhibit RNase activity. The final reaction volume was 10
658 μL prior to the addition of loading dye.

659 For visualization during electrophoresis, 2 μL of EMSA loading dye (0.8 mg/mL Orange G, 15
660 mg/mL Ficoll 400) was added to each reaction prior to gel loading. Novex TBE 6 or 10% gels
661 (Invitrogen) were pre-run in 0.5x TBE running buffer (pH ~8.4 +/- 0.5), at 100V for a minimum
662 of 30 min. Due to the complexes of the larger *glpFKD*(-219) probe and higher concentrations
663 of PlzA not exiting the well, the pH of the running buffer was adjusted to a pH of 8.0

664 immediately prior to electrophoresis which allowed the complexes to enter the gel matrix. The
665 pH was not adjusted for EMSAs using nucleic acid fragments <60 bp as complexes readily
666 exited the well at the standard pH of 0.5x TBE running buffer. The entire reaction mixture was
667 then loaded onto the pre-run gels and resolved at 100 V for 60-90 min at RT. EMSA images
668 were acquired with a ChemiDoc Imaging System (Bio-Rad).

669 **Densitometry and statistics**

670 Densitometry of EMSAs was performed using Image Lab 6.1 (Bio-Rad) software. Analyses
671 were performed on triplicate EMSAs. Lanes as well as free and shifted bands were added
672 manually, and intensity values were reported as band percentage. For dsDNA EMSAs, free
673 dsDNA band percentages were normalized to the probe only control band percentage. The
674 free DNA band percentage values were then used to determine the percent shifted of the
675 respective probe relative to the free probe. For RNA and ssDNA EMSAs, lanes and bands
676 were selected manually, and the shifted bands were analyzed to obtain the percent shifted
677 values. Background and any ssDNA probe percentage that was shifted were subtracted from
678 the percent totals. For competition EMSAs, the band intensity values of each band in a lane
679 were determined manually, and the free probe band intensity values determined. Any probe
680 signal stuck in the gel wells, attributed to potential protein aggregation, were not considered
681 shifts, and thus omitted from densitometric analysis.

682 To calculate the apparent dissociation constant (K_D), nonlinear regression analysis was
683 performed to determine the best-fit values for each experimental condition tested. Briefly,
684 protein concentrations were plotted against the shifted band percentage values obtained from
685 densitometry using the one-site specific binding setting in Prism GraphPad 9.5.1. Confidence
686 was set at 95% with the analysis considering each replicate value from triplicate EMSAs
687 performed for the analysis. Either a Welch's t-test or Welch ANOVA followed by a Dunnett's
688 T3 multiple comparisons test, was used to compare fits.

689 **Major and Minor Groove Binding Assay**

690 The IRDye 800 labeled *glpFKD*(-7/+35) DNA probe was used in the major and minor groove
691 binding assay. The reactions were assembled similarly to the EMSA reactions described
692 above. Briefly, 2.5 μM of WT PlzA was incubated with 100 μM c-di-GMP in EMSA buffer for 5
693 minutes. The *glpFKD*(-7/+35) DNA probe (10 nM) was then added to the reaction mixture
694 which was allowed to incubate for an additional 10 minutes at room temperature. After the
695 probe incubation, 0.25-250 μM of either methyl green (major groove binder) or actinomycin D
696 (minor groove binder) were added to the reaction mixture and allowed to incubate for 5
697 minutes at room temperature. EMSA loading dye (0.8 mg/mL Orange G, 15 mg/mL Ficoll 400)
698 was added to each reaction, and the reaction mixtures were subsequently loaded onto pre-
699 run 10% Novex TBE gels. Gels were resolved and images acquired as described above for
700 EMSAs.

701 **Docking analysis and protein modeling**

702 Computational modeling of holo-PlzA (PDB ID: 7mie) in complex with the nucleic acids of
703 interest was performed using the HDOCK server: http://hdock.phys.hust.edu.cn/. Briefly, the
704 PDB and chain ID of PlzA was entered for the input receptor molecule to retrieve the crystal
705 structure of the protein for the analysis. For each analysis, the corresponding nucleic acid
706 sequence of interest was then pasted as the input ligand molecule, and the type of nucleic
707 acid selected from the drop-down menu. The input nucleic acid sequences are listed in Table
708 2. All analyses were performed with default parameters. The top model for each computation
709 is shown. Docking and confidence scores are provided in the corresponding figure legends.
710 The more negative the docking score, the more probable the binding. A confidence score
711 above 0.7 indicates the molecules would be highly likely to bind.

712 Protein structure prediction for $PlzA_{NTD}$, $PlzA_{CTD}$, and $PlzA_{RD\text{-}RD}$ was performed via AlphaFold.
713 All output PDB files generated from the various analyses were viewed with ChimeraX 1.4
714 (Goddard *et al.*, 2018; Jumper *et al.*, 2021; Pettersen *et al.*, 2021).

715 **Cyclic di-nucleotide LC-MS analysis**

716 Detection of cyclic di-nucleotides in recombinant protein purifications was conducted by liquid
717 chromatography-tandem mass spectrometry (LC-MS/MS) at the Michigan State University
718 Mass Spectrometry core. Briefly, recombinant proteins were purified and quantified as
719 mentioned previously and subsequently dialyzed into EMSA buffer without Tween-20. For LC-
720 MS/MS analysis, approximately 1 mg/mL of $PlzA_{WT}$, $PlzA_{RD-RD,}$ $PlzA_{NTD}$, and $PlzA_{CTD}$ were
721 aliquoted and stored at -80°C until shipment to the core facility. Nucleotides bound to purified
722 protein samples (1mg/mL) were analyzed by precipitating protein with 3 volumes of acetonitrile
723 followed by centrifugation to pellet precipitated protein. The supernatant was transferred to a
724 new tube and evaporated to dryness. The sample was reconstituted in mobile phase solvent
725 A (10 mM tributylamine + 15 mM acetic acid in water/methanol, 97:3 v/v). LC-MS analysis was
726 performed on a Waters Xevo G2-XS Quadrupole-Time-of-Flight (QTof) mass spectrometer
727 interfaced with a Waters Acquity UPLC system. 10 µL of sample was injected onto a Waters
728 Acquity BEH-C18 UPLC column (2.1x50 mm) and compounds were separated using the
729 following gradient: initial conditions were 100% mobile phase A, hold at 100% A until 1 min,
730 ramp to 99% B at 7 min (mobile phase B: methanol), hold at 99% B until 8 min, return to 100%
731 A at 8.01 min and hold until 10 min. The flow rate through the column was 0.3 ml/min and the
732 column temperature was held at 40°C. Compounds were ionized by electrospray ionization
733 operated in negative mode. Capillary voltage was 2.0 kV, cone voltage was 35, source
734 temperature was 100C and desolvation temperature was 350C. The cone gas and desolvation
735 gas flows were 50 L/hr and 600 L/hr respectively. A TOF MS scan method with targeted
736 enhancement of m/z 689 was used with 0.5 second scan time. Lockmass correction was
737 performed using leucine enkephalin as the reference compound.

738 Data were processed using Waters Masslynx software. Extracted ion chromatograms were
739 performed to look for the presence of c-di-GMP (m/z 689.09), c-di-AMP (m/z 657.09) and c-
740 GAMP (673.09). The intensity values were converted to relative abundance as determined
741 from the base peak values. The c-di-GMP concentration was determined from the peak area
742 of the target compound against a standard curve of known c-di-GMP concentrations. These

743 concentrations were calculated using the Targetlynx tool in the Waters Masslynx software.
744 The percentage of protein bound with c-di-GMP was calculated from the total protein and
745 calculated c-di-GMP concentrations in each sample.

746 **Circular dichroism spectroscopy**

747 The CD analysis was performed at the University of Kentucky Center for Molecular Medicine
748 Protein Core. Briefly, proteins were either buffer exchanged or diluted into 10 mM sodium
749 phosphate buffer, pH 7.0. A 400 µL solution of each protein at a concentration of 100 µg/mL
750 was pipetted into a 1 mm quartz cuvette and loaded into a Jasco J-810 spectrophotometer.
751 The samples were scanned at 0.1 nm increments and CD spectra collected for wavelengths
752 185-260 nM with the temperature held constant at 25°C. A buffer only sample was used as a
753 control. Two independent recombinant protein preparations were analyzed for each protein
754 type. Spectral data was deconvoluted using CDtoolX. Deconvolution consisted of buffer
755 spectra subtraction and the conversion of the CD spectra to molar residue ellipticity to
756 normalize the data for sample concentration and sequence length. Data were visualized in
757 GraphPad Prism 10.1.2.

758 Secondary structure prediction was performed using the BeStSel webserver using the single
759 spectrum analysis function (Micsonai *et al.*, 2015; Micsonai *et al.*, 2018; Micsonai *et al.*, 2022).
760 Mean residue ellipticity values from the deconvoluted CD data were input into the server for
761 analysis. The PDB file, 7MIE, was used to determine the secondary structure of the
762 crystallized PlzA protein for comparisons to the generated recombinant proteins.

763 **AVAILABILITY**

764 HDOCK (http://hdock.phys.hust.edu.cn/)

765 CDtoolX (https://cdtools.cryst.bbk.ac.uk/)

766 BeStSel (https://bestsel.elte.hu/index.php)

767 **ACKNOWLEDGEMENT**

768 We dedicate this manuscript to our friend and colleague Christina R. Savage, Ph.D., whose
769 initial studies on PlzA were invaluable to this work. We also thank Jessamyn Moore and
770 Tatiana Castro Padovani for their support of the studies undertaken. Circular Dichroism was
771 performed at the University of Kentucky Center for Molecular Medicine, and we thank Martin
772 Chow for his assistance on these studies. The LC-MS/MS to detect di-nucleotides was
773 performed by the Mass Spectrometry and Metabolomics Core at Michigan State University. A
774 sincere thank you to Tony Schilmiller at MSU for all his help with the mass spectrometry
775 services. Figures 1, 4a, 6a, 7a, 10a, and 13 were created via BioRender. Lastly, we thank Dr.
776 József Kardos at Eötvös Loránd University for providing the secondary structure analysis for
777 the 7MIE PDB file via email.

778 **FUNDING**

779 Funding for open access charge: US National Institutes of Health (grant R01 AI144126-3 to
780 B.S.).

781 **CONFLICT OF INTEREST**

782 None declared.

783 **REFERENCES**

784 Adams, P.P., Avile, C.F., Popitsch, N., Bilusic, I., Schroeder, R., Lybecker, M., and
785 Jewett, M.W. (2017) In vivo expression technology and 5′ end mapping of the *Borrelia*
786 *burgdorferi* transcriptome identify novel RNAs expressed during mammalian infection.
787 *Nucleic Acids Res* **45**: 775–792.

788 Amikam, D., and Galperin, M.Y. (2006) PilZ domain is part of the bacterial c-di-GMP
789 binding protein. *Bioinformatics* **22**: 3–6
790 https://academic.oup.com/bioinformatics/article/22/1/3/218448. Accessed February 18,
791 2022.

792 Babb, K., Bykowski, T., Riley, S.P., Miller, M.C., Demoll, E., and Stevenson, B. (2006)
793 *Borrelia burgdorferi* EbfC, a novel, chromosomally encoded protein, binds specific DNA
794 sequences adjacent to erp loci on the spirochete's resident cp32 prophages. *J Bacteriol*
795 **188**: 4331–4339 https://journals.asm.org/journal/jb. Accessed April 8, 2022.

796 Basu, D., Khare, G., Singh, S., Tyagi, A., Khosla, S., and Mande, S.C. (2009) A novel
797 nucleoid-associated protein of Mycobacterium tuberculosis is a sequence homolog of
798 GroEL. *Nucleic Acids Res* **37**: 4944–4954.

799 Bauer, W.J., Luthra, A., Zhu, G., Radolf, J.D., Malkowski, M.G., and Caimano, M.J.
800 (2015) Structural characterization and modeling of the *Borrelia burgdorferi* hybrid
801 histidine kinase Hk1 periplasmic sensor: A system for sensing small molecules
802 associated with tick feeding. *J Struct Biol* **192**: 48–58.

803 Bong, H.J., Ko, E.M., Song, S.Y., Ko, I.J., and Oh, J. Il (2019) Tripartite regulation of
804 the *glpFKD* operon involved in glycerol catabolism by gylr, crp, and sigf in
805 *Mycobacterium smegmatis*. *J Bacteriol* **201**.

806 Burns, L.H., Adams, C.A., Riley, S.P., Jutras, B.L., Bowman, A., Chenail, A.M., *et al.*
807 (2010) BpaB, a novel protein encoded by the Lyme disease spirochete's cp32
808 prophages, binds to *erp* Operator 2 DNA. *Nucleic Acids Res* **38**: 5443–5455.

809 Caimano, M.J., Dunham-Ems, S., Allard, A.M., Cassera, M.B., Kenedy, M., and Radolf,
810 J.D. (2015) Cyclic di-GMP modulates gene expression in Lyme disease spirochetes at
811 the tick-mammal interface to promote spirochete survival during the blood meal and
812 tick-to-mammal transmission. *Infect Immun* **83**: 3043–3060 http://dx.doi.org/10.1128.
813 Accessed May 1, 2021.

814 Caimano, M.J., Kenedy, M.R., Kairu, T., Desrosiers, D.C., Harman, M., Dunham-Ems,
815 S., *et al.* (2011) The hybrid histidine kinase Hk1 is part of a two-component system that
816 is essential for survival of *Borrelia burgdorferi* in feeding *Ixodes scapularis* ticks. *Infect*
817 *Immun* **79**: 3117–3130 http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml. Accessed
818 February 13, 2022.

819 Carey, J. (2022) Affinity, Specificity, and Cooperativity of DNA Binding by Bacterial
820 Gene Regulatory Proteins. *Int J Mol Sci* **23**: 1–16.

821 Casjens, S., Palmer, N., Vugt, R. Van, Huang, W.M., Stevenson, B., Rosa, P., *et al.*
822 (2000) A bacterial genome in flux: The twelve linear and nine circular
823 extrachromosomal DNAs in an infectious isolate of the Lyme disease spirochete
824 *Borrelia burgdorferi*. *Mol Microbiol* **35**: 490–516.

825 Cheang, Q.W., Xin, L., Chea, R.Y.F., and Liang, Z.X. (2019) Emerging paradigms for
826 PilZ domain-mediated C-di-GMP signaling. *Biochem Soc Trans* **47**: 381–388.

827 Chou, S.H., and Galperin, M.Y. (2016) Diversity of cyclic di-GMP-binding proteins and
828 mechanisms. *J Bacteriol* **198**: 32–46 http://ncbi.nlm.nih.gov/Complete_Genomes/c-di-
829 GMP.html. Accessed February 13, 2022.

830 Corona, A., and Schwartz, I. (2015) *Borrelia burgdorferi* : Carbon Metabolism and the
831 Tick-Mammal Enzootic Cycle. *Microbiol Spectr* **3**.

832 Fraser, C.M., Casjens, S., Huang, W.M., Sutton, G.G., Clayton, R., Lathigra, R., *et al.*
833 (1997) Genomic sequence of a Lyme disease spirochaete, *Borrelia burgdorferi*. *Nature*
834 **390**: 580–586.

Freedman, J.C., Rogers, E.A., Kostick, J.L., Zhang, H., Iyer, R., Schwartz, I., and Marconi, R.T. (2009) Identification and molecular characterization of a cyclic-di-GMP effector protein, PlzA (BB0733): Additional evidence for the existence of a functional cyclic-di-GMP regulatory network in the Lyme disease spirochete, *Borrelia burgdorferi*. *FEMS Immunol Med Microbiol* **58**: 285–294 https://academic.oup.com/femspd/article/58/2/285/597665. Accessed April 26, 2021.

Fried, M., and Crothers, D.M. (1981) Equilibria and kinetics of lac repressor-operator interactions by polyacrylamide gel electrophoresis. *Nucleic Acids Res* **9**: 1981.

Fried, M.G. (1989) Measurement of protein-DNA interaction parameters by electrophoresis mobility shift assay. *Electrophoresis* **10**: 366–376.

Galperin, M.Y., and Chou, S.H. (2020) Structural conservation and diversity of PilZ-related domains. *J Bacteriol* **202**.

Galperin, M.Y., Nikolskaya, A.N., and Koonin, E. V. (2001) Novel domains of the prokaryotic two-component signal transduction systems. *FEMS Microbiol Lett* **203**: 11–21 https://academic.oup.com/femsle/article/203/1/11/478433. Accessed March 17, 2022.

Goddard, T.D., Huang, C.C., Meng, E.C., Pettersen, E.F., Couch, G.S., Morris, J.H., and Ferrin, T.E. (2018) UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Science* **27**: 14–25.

Grassmann, A.A., Tokarz, R., Golino, C., McLain, M.A., Groshong, A.M., Radolf, J.D., and Caimano, M.J. (2023) BosR and PlzA reciprocally regulate RpoS function to sustain *Borrelia burgdorferi* in ticks and mammals. *Journal of Clinical Investigation* http://www.jci.org/articles/view/166710.

Groshong, A.M., Grassmann, A.A., Luthra, A., Mclain, M.A., Provatas, A.A., Radolf, J.D., and Caimano, M.J. (2021) PlzA is a bifunctional c-di-GMP biosensor that promotes tick and mammalian host- adaptation of *Borrelia burgdorferi*. *PLoS Pathog* **17**: 1–31.

Grove, A.P., Liveris, D., Iyer, R., Petzke, M., Rudman, J., Caimano, M.J., *et al.* (2017) Two Distinct Mechanisms govern RpoS-Mediated Repression of Tick-Phase Genes during Mammalian Host Adaptation by *Borrelia burgdorferi*, the Lyme Disease Spirochete. *mBio* **8** https://journals.asm.org/doi/full/10.1128/mBio.01204-17. Accessed June 18, 2022.

He, M., Ouyang, Z., Troxell, B., Xu, H., Moh, A., Piesman, J., *et al.* (2011) Cyclic di-gmp is essential for the survival of the lyme disease spirochete in ticks. *PLoS Pathog* **7**.

He, M., Zhang, J.J., Ye, M., Lou, Y., and Yang, X.F. (2014) Cyclic Di-GMP receptor PlzA controls virulence gene expression through RpoS in *Borrelia burgdorferi*. *Infect Immun* **82**: 445–452 http://iai.asm.org/. Accessed April 26, 2021.

Helble, J.D., McCarthy, J.E., and Hu, L.T. (2021) Interactions between *Borrelia burgdorferi* and its hosts across the enzootic cycle. *Parasite Immunol* **43**.

874 Hellman, L.M., and Fried, M.G. (2007) Electrophoretic mobility shift assay (EMSA) for
875 detecting protein-nucleic acid interactions. *Nat Protoc* **2**: 1849–1861.

876 Hengge, R. (2009) Principles of c-di-GMP signalling in bacteria. *Nat Rev Microbiol* **7**:
877 263–273.

878 Hsieh, M.L., Hinton, D.M., and Waters, C.M. (2018) VpsR and cyclic di-GMP together
879 drive transcription initiation to activate biofilm formation in *Vibrio cholerae*. *Nucleic*
880 *Acids Res* **46**: 8876–8887.

881 Huang, S.Y., and Zou, X. (2014) A knowledge-based scoring function for protein-RNA
882 interactions derived from a statistical mechanics-based iterative method. *Nucleic Acids*
883 *Res* **42**.

884 Jenal, U., Reinders, A., and Lori, C. (2017) Cyclic di-GMP: Second messenger
885 extraordinaire. *Nat Rev Microbiol* **15**: 271–284
886 http://dx.doi.org/10.1038/nrmicro.2016.190.

887 Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., *et al.*
888 (2021) Highly accurate protein structure prediction with AlphaFold. *Nature* **596**: 583–
889 589 http://dx.doi.org/10.1038/s41586-021-03819-2.

890 Junkermeier, E.H., and Hengge, R. (2023) Local signaling enhances output specificity
891 of bacterial c-di-GMP signaling networks. *microLife* **4**.

892 Jutras, B.L., Bowman, A., Brissette, C.A., Adams, C.A., Verma, A., Chenail, A.M., and
893 Stevenson, B. (2012) EbfC (YbaB) is a new type of bacterial nucleoid-associated
894 protein and a global regulator of gene expression in the Lyme disease spirochete. *J*
895 *Bacteriol* **194**: 3395–3406.

896 Jutras, B.L., Chenail, A.M., Carroll, D.W., Miller, M.C., Zhu, H., Bowman, A., and
897 Stevenson, B. (2013a) Bpur, the Lyme disease spirochete's PUR domain protein:
898 Identification as a transcriptional modulator and characterization of nucleic acid
899 interactions. *Journal of Biological Chemistry* **288**: 26220–26234 http://www.jbc.org/.
900 Accessed October 5, 2020.

901 Jutras, B.L., Chenail, A.M., Rowland, C.L., Carroll, D., Miller, M.C., Bykowski, T., and
902 Stevenson, B. (2013b) Eubacterial SpoVG Homologs Constitute a New Family of Site-
903 Specific DNA-Binding Proteins. *PLoS One* **8**: 66683 www.plosone.org. Accessed April
904 26, 2021.

905 Kim, B., and Little, J.W. (1992) Dimerization of a specific DNA-binding protein on the
906 DNA. *Science (1979)* **255**: 203–206.

907 Kim, S.K., and Nordén, B. (1993) Methyl green. A DNA major-groove binding drug.
908 *FEBS Lett* **315**: 61–64.

909 Kohler, J.J., Metallo, S.J., Schneider, T.L., and Schepartz, A. (1999) DNA specificity
910 enhanced by sequential binding of protein monomers. *Proc Natl Acad Sci U S A* **96**:
911 11735–11739.

912 Kostick, J.L., Szkotnicki, L.T., Rogers, E.A., Bocci, P., Raffaelli, N., and Marconi, R.T.
913 (2011) The diguanylate cyclase, Rrp1, regulates critical steps in the enzootic cycle of
914 the Lyme disease spirochetes. *Mol Microbiol* **81**: 219–231
915 https://onlinelibrary.wiley.com/doi/full/10.1111/j.1365-2958.2011.07687.x. Accessed
916 April 15, 2022.

917 Kostick-Dunn, J.L., Izac, J.R., Freedman, J.C., Szkotnicki, L.T., Oliver, L.D., and
918 Marconi, R.T. (2018) The *Borrelia burgdorferi* c-di-GMP Binding Receptors, PlzA and
919 PlzB, Are Functionally Distinct. *Front Cell Infect Microbiol* **8**: 213 www.frontiersin.org.
920 Accessed April 26, 2021.

921 Krusenstjerna, A.C., Arnold, W.K., Saylor, T.C., Tucker, J.S., and Stevenson, B. (2023)
922 *Borrelia* burgdorferi DnaA and the nucleoid-associated protein EbfC coordinate
923 expression of the *dnaX-ebfC* operon. *J Bacteriol* **205**: e00396-22
924 https://doi.org/10.1128/jb.00396-22.

925 Larouche, K., Bergeron, M.-J., Leclerc, S., and Guerin, S. (1996) Optimization of
926 competitor poly(dI-dC).poly(dI-dC) levels is advised in DNA-protein interaction studies
927 involving enriched nuclear proteins. *Biotechniques* **20**: 439–444.

928 Lee, R.E., and Baust, J.G. (1987) Cold-Hardiness in the Antarctic Tick, *Ixodes uriae*.
929 *Physiol Zool* **60**: 499–506 https://about.jstor.org/terms.

930 Liu, Y., Wang, H., Cui, T., Zhou, X., Jia, Y., Zhang, H., and He, Z.G. (2016) NapM, a
931 new nucleoid-associated protein, broadly regulates gene expression and affects
932 mycobacterial resistance to anti-tuberculosis drugs. *Mol Microbiol* **101**: 167–181.

933 Mallory, K.L., Miller, D.P., Oliver, L.D., Freedman, J.C., Kostick-Dunn, J.L., Carlyon,
934 J.A., *et al.* (2016) Cyclic-di-GMP binding induces structural rearrangements in the PlzA
935 and PlzC proteins of the Lyme disease and relapsing fever spirochetes: A possible
936 switch mechanism for c-di-GMP-mediated effector functions. *Pathog Dis* **74**: 1–8.

937 Micsonai, A., Moussong, É., Wien, F., Boros, E., Vadászi, H., Murvai, N., *et al.* (2022)
938 BeStSel: Webserver for secondary structure and fold prediction for protein CD
939 spectroscopy. *Nucleic Acids Res* **50**: W90–W98.

940 Micsonai, A., Wien, F., Bulyáki, É., Kun, J., Moussong, É., Lee, Y.H., *et al.* (2018)
941 BeStSel: A web server for accurate protein secondary structure prediction and fold
942 recognition from the circular dichroism spectra. *Nucleic Acids Res* **46**: W315–W322.

943 Micsonai, A., Wien, F., Kernya, L., Lee, Y.H., Goto, Y., Réfrégiers, M., and Kardos, J.
944 (2015) Accurate secondary structure prediction and fold recognition for circular
945 dichroism spectroscopy. *Proc Natl Acad Sci U S A* **112**: E3095–E3103.

946 Novak, E.A., Sultan, S.Z., and Motaleb, M.A. (2014) The cyclic-di-GMP signaling
947 pathway in the Lyme disease spirochete, *Borrelia burgdorferi*. *Front Cell Infect Microbiol*
948 **4** www.frontiersin.org. Accessed April 26, 2021.

949 Pappas, C.J., Iyer, R., Petzke, M.M., Caimano, M.J., Radolf, J.D., and Schwartz, I.
950 (2011) *Borrelia burgdorferi* requires glycerol for maximum fitness during the tick phase
951 of the enzootic cycle. *PLoS Pathog* **7**.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Meng, E.C., Couch, G.S., Croll, T.I., *et al.* (2021) UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Science* **30**: 70–82.

Pitzer, J.E., Syed, S.Z., Hayakawa, Y., Hobbs, G., Miller, M.R., and Motaleb, M.A. (2011) Analysis of the *Borrelia burgdorferi* Cyclic-di-GMP-Binding Protein PlzA Reveals a Role in Motility and Virulence. *Infect Immun* **79**: 1815–1825 http://iai.asm.org/. Accessed April 26, 2021.

Pronobis, M.I., Deuitch, N., and Peifer, M. (2016) The Miraprep: A protocol that uses a Miniprep kit and provides Maxiprep yields. *PLoS One* **11**: 1–12.

Pultz, I.S., Christen, M., Kulasekara, H.D., Kennard, A., Kulasekara, B., and Miller, S.I. (2012) The response threshold of Salmonella PilZ domain proteins is determined by their binding affinities for c-di-GMP. *Mol Microbiol* **86**: 1424–1440.

Riley, S.P., Bykowski, T., Cooley, A.E., Burns, L.H., Babb, K., Brissette, C.A., *et al.* (2009) *Borrelia burgdorferi* EbfC defines a newly-identified, widespread family of bacterial DNA-binding proteins. *Nucleic Acids Res* **37**: 1973–1983 http://www.rcsb.org/. Accessed April 8, 2022.

Rogers, E.A., Terekhova, D., Zhang, H.-M.M., Hovis, K.M., Schwartz, I., and Marconi, R.T. (2009) Rrp1, a cyclic-di-GMP-producing response regulator, is an important regulator of *Borrelia burgdorferi* core cellular functions. *Mol Microbiol* **71**: 1551–1573 http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2843504&tool=pmcentrez&rendertype=abstract. Accessed February 18, 2022.

Romling, U., Galperin, M.Y., and Gomelsky, M. (2013) Cyclic di-GMP: the First 25 Years of a Universal Bacterial Second Messenger. *Microbiology and Molecular Biology Reviews* **77**: 1–52.

Rooman, M., Lié Vin, J., Buisine, E., and Wintjens, R. (2002) Cation-p/H-bond Stair Motifs at Protein-DNA Interfaces. *J Mol Biol* **319**: 67–76 http://www.idealibrary.com.

Ryjenkov, D.A., Simm, R., Römling, U., and Gomelsky, M. (2006) The PilZ domain is a receptor for the second messenger c-di-GMP: The PilZ domain protein YcgR controls motility in enterobacteria. *Journal of Biological Chemistry* **281**: 30310–30314.

Ryjenkov, D.A., Tarutina, M., Moskvin, O. V., and Gomelsky, M. (2005) Cyclic diguanylate is a ubiquitous signaling molecule in bacteria: Insights into biochemistry of the GGDEF protein domain. *J Bacteriol* **187**: 1792–1798 https://journals.asm.org/doi/abs/10.1128/JB.187.5.1792-1798.2005. Accessed March 17, 2022.

Sarenko, O., Klauck, G., Wilke, F.M., Pfiffer, V., Richter, A.M., Herbst, S., *et al.* (2017) More than enzymes that make or break cyclic Di-GMP—local signaling in the interactome of GGDEF/EAL domain proteins of Escherichia coli. *mBio* **8**.

Sathyapriya, R., and Vishveshwara, S. (2004) Interaction of DNA with clusters of amino acids in proteins. *Nucleic Acids Res* **32**: 4109–4118.

991 Savage, C.R., Jutras, B.L., Bestor, A., Tilly, K., Rosa, P.A., Tourand, Y., *et al.* (2018)
992 *Borrelia burgdorferi* SpoVG DNA- and RNA-binding protein modulates the physiology of
993 the Lyme disease spirochete. *J Bacteriol* **200** http://jb.asm.org/. Accessed April 26,
994 2021.

995 Saylor, T.C., Savage, C.R., Krusenstjerna, A.C., Jusufovic, N., Zückert, W.R., Brissette,
996 C.A., *et al.* (2023) Quantitative analyses of interactions between SpoVG and RNA/DNA.
997 *Biochem Biophys Res Commun* **654**: 40–46.

998 Schäper, S., Steinchen, W., Krol, E., Altegoer, F., Skotnicka, D., Søgaard-Andersen, L.,
999 *et al.* (2017) AraC-like transcriptional activator CuxR binds c-di-GMP by a PilZ-like
1000 mechanism to regulate extracellular polysaccharide production. *Proc Natl Acad Sci U S*
1001 *A* **114**: E4822–E4831.

1002 Schramm, F.D., Schroeder, K., and Jonas, K. (2019) Protein aggregation in bacteria.
1003 *FEMS Microbiol Rev* **44**: 54–72.

1004 Shi, W.T., Zhang, B., Li, M.L., Liu, K.H., Jiao, J., and Tian, C.F. (2022) The convergent
1005 xenogeneic silencer MucR predisposes α-proteobacteria to integrate AT-rich symbiosis
1006 genes. *Nucleic Acids Res* **50**: 8580–8598.

1007 Singh, A., Izac, J.R., Schuler, E.J.A., Patel, D.T., Davies, C., and Marconi, R.T. (2021)
1008 High-resolution crystal structure of the *Borreliella burgdorferi* PlzA protein in complex
1009 with c-di-GMP: new insights into the interaction of c-di-GMP with the novel xPilZ
1010 domain. *Pathog Dis* **79**: 1–9.

1011 Skotnicka, D., Steinchen, W., Szadkowski, D., Cadby, I.T., Lovering, A.L., Bange, G.,
1012 and Søgaard-Andersen, L. (2020) CdbA is a DNA-binding protein and c-di-GMP
1013 receptor important for nucleoid organization and segregation in *Myxococcus xanthus*.
1014 *Nat Commun* **11** https://doi.org/10.1038/s41467-020-15628-8. Accessed February 13,
1015 2022.

1016 Sultan, S.Z., Pitzer, J.E., Boquoi, T., Hobbs, G., Miller, M.R., and Motaleb, M.A. (2011)
1017 Analysis of the HD-GYP domain cyclic dimeric gmp phosphodiesterase reveals a role in
1018 motility and the enzootic life cycle of *Borrelia burgdorferi*. *Infect Immun* **79**: 3273–3283.

1019 Sultan, S.Z., Pitzer, J.E., Miller, M.R., and Motaleb, M.A. (2010) Analysis of a *Borrelia*
1020 *burgdorferi* phosphodiesterase demonstrates a role for cyclic-di-guanosine
1021 monophosphate in motility and virulence. *Mol Microbiol* **77**: 128–142.

1022 Tan, J.W.H., Wilksch, J.J., Hocking, D.M., Wang, N., Srikhanta, Y.N., Tauschek, M., *et*
1023 *al.* (2015) Positive Autoregulation of *mrkHI* by the Cyclic Di-GMP-Dependent MrkH
1024 Protein in the Biofilm Regulatory Circuit of *Klebsiella pneumoniae*. *J Bacteriol* **197**:
1025 1659–1667.

1026 Tschowri, N., Schumacher, M.A., Schlimpert, S., Chinnam, N.B., Findlay, K.C.,
1027 Brennan, R.G., and Buttner, M.J. (2014) Tetrameric c-di-GMP mediates effective
1028 transcription factor dimerization to control *Streptomyces* development. *Cell* **158**: 1136–
1029 1147 http://dx.doi.org/10.1016/j.cell.2014.07.022.

1030 Valentini, M., and Filloux, A. (2019) Multiple roles of c-di-GMP signaling in bacterial
1031 pathogenesis. *Annu Rev Microbiol* **73**: 387–406.

1032 Van Gundy, T., Patel, D., Bowler, B.E., Rothfuss, M.T., Hall, A.J., Davies, C., *et al.*
1033 (2023) c-di-GMP regulates activity of the PlzA RNA chaperone from the Lyme disease
1034 spirochete. *Mol Microbiol* **119**: 711–727.

1035 Wang, F., He, Q., Su, K., Gao, F., Huang, Y., Lin, Z., *et al.* (2016) The PilZ domain of
1036 MrkH represents a novel DNA binding motif. *Protein Cell* **7**: 766–772
1037 /pmc/articles/PMC5055493/. Accessed February 13, 2022.

1038 Wilksch, J.J., Yang, J., Clements, A., Gabbe, J.L., Short, K.R., Cao, H., *et al.* (2011)
1039 MrkH, a Novel c-di-GMP-Dependent Transcriptional Activator, Controls *Klebsiella*
1040 *pneumoniae* Biofilm Formation by Regulating Type 3 Fimbriae Expression. *PLoS*
1041 *Pathog* **7**: e1002204
1042 https://journals.plos.org/plospathogens/article?id=10.1371/journal.ppat.1002204.
1043 Accessed February 14, 2022.

1044 Yan, Y., Tao, H., He, J., and Huang, S.Y. (2020) The HDOCK server for integrated
1045 protein–protein docking. *Nat Protoc* **15**: 1829–1852.

1046 Yan, Y., Wen, Z., Wang, X., and Huang, S.Y. (2017a) Addressing recent docking
1047 challenges: A hybrid strategy to integrate template-based and free protein-protein
1048 docking. *Proteins: Structure, Function and Bioinformatics* **85**: 497–512.

1049 Yan, Y., Zhang, D., Zhou, P., Li, B., and Huang, S.Y. (2017b) HDOCK: A web server for
1050 protein-protein and protein-DNA/RNA docking based on a hybrid strategy. *Nucleic*
1051 *Acids Res* **45**: W365–W373.

1052 Yang, J., Wilksch, J.J., Tan, J.W.H., Hocking, D.M., Webb, C.T., Lithgow, T., *et al.*
1053 (2013) Transcriptional activation of the *mrkA* promoter of the *Klebsiella pneumoniae*
1054 type 3 fimbrial operon by the c-di-GMP-dependent MrkH protein. *PLoS One* **8**.

1055 Zamorano-Sánchez, D., Fong, J.C.N., Kilic, S., Erill, I., and Yildiz, F.H. (2015)
1056 Identification and characterization of VpsR and VpsT binding sites in *Vibrio cholerae*. *J*
1057 *Bacteriol* **197**: 1221–1235 http://dx.doi.org/10.1128. Accessed May 2, 2021.

1058 Zhang, J.J., Chen, T., Yang, Y., Du, J., Li, H., Troxell, B., *et al.* (2018) Positive and
1059 Negative Regulation of Glycerol Utilization by the c-di-GMP Binding Protein PlzA in
1060 *Borrelia burgdorferi*. *J Bacteriol* **200** https://doi.org/10.1128/JB. Accessed April 14,
1061 2021.

1062 Zhang, J.-J., Raghunandanan, S., Wang, Q., Priya, R., Alanazi, F., Lou, Y., and Yang,
1063 X.F. (2024) BadR directly represses the expression of the glycerol utilization operon in
1064 the Lyme disease pathogen. *J Bacteriol* https://journals.asm.org/doi/10.1128/jb.00340-
1065 23.

1066

1067 **FIGURES LEGENDS**

1068 **FIGURE 1** The *glpFKD* operon, 5' UTR, and designed probe schematic. The glycerol
1069 catabolism gene operon (*glpFKD*) contains four genes: *glpF* (BB_0240, glycerol uptake
1070 facilitator protein), *glpK* (BB_0241, glycerol kinase), BB_0242 (hypothetical protein), and *glpD*
1071 (BB_0243, glycerol-3-phosphate dehydrogenase). The promoter and transcriptional
1072 boundaries were previously mapped (Grove *et al.*, 2017). The -35 and -10 sites are bolded
1073 and underlined. The transcriptional start site nucleotide is highlighted in blue and topped with
1074 an arrow indicating the direction of transcription. The start of the *glpF* ORF is marked by the
1075 bolded ATG, while the start of the upstream BB_0239 ORF is labeled, and its nucleotides are
1076 italicized. The long arrows indicate the location of primers used to generate a large probe
1077 incorporating the entire intergenic region which is denoted as *glpFKD*(-219) for initial DNA
1078 binding studies.

1079 **FIGURE 2** DNA binding by PlzA is c-di-GMP dependent. (**a**) EMSA with increasing
1080 concentrations of $PlzA_{WT}$ showing binding to *glpFKD*(-219) DNA. The mixture was incubated
1081 for 10 min before electrophoresis. (**b**) The same EMSA mixture as in (**a**) but with a final
1082 concentration of 2.5 ng/μL of the nonspecific competitor poly-dI-dC. Protein was incubated
1083 with poly-dI-dC for 5 min prior to the addition of labeled *glpFKD*(-219) probe. (**c**) The same
1084 EMSA mixture as in (**b**) but supplemented with 100 μM c-di-GMP. Protein was incubated with
1085 c-di-GMP for 5 min prior to the addition of any nucleic acids. Arrows indicate PlzA- *glpFKD*(-
1086 219) complexes. Lane 1 in (**a-c**) is the probe only control, containing 50 nM of probe. Protein
1087 concentrations are as follows in each subsequent lane: Lane 2: 0.5 μM, Lane 3: 1 μM, Lane
1088 4: 2.5 μM, Lane 5: 5 μM, Lane 6: 7.5 μM, and Lane 7: 10 μM. (**d**), (**e**), (**f**) EMSA experiments
1089 as in (**a**), (**b**), and (**c**), respectively, but performed with the mutant $PlzA_{RD\text{-}RD}$ protein. All
1090 conditions were performed identically as with $PlzA_{WT}$. Each panel is a representative gel from
1091 separate EMSAs which were performed in triplicate ($PlzA_{WT}$) or duplicate ($PlzA_{RD\text{-}RD}$).

1092 **FIGURE 3** Detection of c-di-GMP in recombinant $PlzA_{WT}$ and $PlzA_{RD\text{-}RD}$ by LC-MS/MS.
1093 Recombinant $PlzA_{WT}$ and $PlzA_{RD\text{-}RD}$ were sent off for mass spectrometry to detect c-di-GMP.
1094 Extracted ion chromatograms of the LC-MS/MS analysis for detection of c-di-GMP are

1095 displayed. Chromatograms depict the control c-di-GMP standard at 1 μM, recombinant
1096 $PlzA_{WT}$, and recombinant $PlzA_{RD-RD}$. The LC-MS/MS analysis reveals $PlzA_{WT}$ co-purified with
1097 c-di-GMP, but $PlzA_{RD-RD}$ did not. The concentration of c-di-GMP in recombinant $PlzA_{WT}$ was
1098 determined from a standard curve of c-di-GMP. $PlzA_{WT}$ co-purified with 99.29 nM of c-di-GMP
1099 and approximately 0.3% of the protein was bound with c-di-GMP. Black lines indicate
1100 chromatograms unrelated to this figure that were removed.

1101 **FIGURE 4** Levels of c-di-GMP alter DNA binding affinity of $PlzA_{WT}$. (**a**) A diagram of the
1102 generated DNA probe, designated as *glpFKD*(-7/+35), corresponding to a 42 bp sequence
1103 adjacent to the -10 site and proceeding 35 bp into the 5' UTR. An RNA probe was made
1104 corresponding to only nucleotides that are within the 5 'UTR which was designated as
1105 *glpFKD*(UTR) RNA. Probes were conjugated to a fluorescent molecule for the detection of
1106 binding in EMSAs. (**b**) EMSAs were performed with increasing protein concentrations of
1107 $PlzA_{WT,}$ and 10 nM fluorescently labeled probe at varying c-di-GMP concentrations (0,
1108 equimolar, 10, 25, 50, or 100 μM). For each EMSA shown, the indicated c-di-GMP
1109 concentration was held constant in all lanes. For each EMSA shown, lanes 1 and 7 are probe-
1110 only controls. Protein concentrations were as follows: Lane 2- 1 μM, Lane 3- 2.5 μM, Lane 4-
1111 3.5 μM, Lane 5- 5 μM, Lane 6- 7.5 μM. A final concentration of 2.5 ng/μL of the nonspecific
1112 competitor poly-dI-dC was added to each lane. Arrows indicate higher-order complexes
1113 formed as protein concentration increases. A representative gel for each tested c-di-GMP
1114 concentration is shown from triplicate EMSAs. (**c**) Triplicate EMSAs were quantitated by
1115 densitometry to determine the $K_{d(app)}$ of the $PlzA_{WT}$-*glpFKD*(-7/35+) interaction by nonlinear
1116 regression analysis using the one-site specific binding setting in GraphPad Prism. Errors bars
1117 represent one standard deviation (SD). (**d**) Bar graph displaying the calculated $K_{d(app)}$ ± the
1118 standard error of the mean (SEM). The $K_{d(app)}$ values from three EMSAs were analyzed by a
1119 Welch ANOVA with Dunnett's T3 multiple comparisons test via GraphPad Prism. Adjusted P-
1120 values are indicated as follows: ns= $p > 0.05$, *= $p \leq 0.05$, and ***= $p \leq 0.001$. Comparisons
1121 between 10, 25, 50, and 100 μM c-di-GMP are not indicated on the graph as they were not

1122 significant. (**e**) A similar EMSA experiment was performed as in Figure 4B but with $PlzA_{RD-RD}$
1123 and either 100 (lanes 1-6) or 0 (lanes 7-12) μM of c-di-GMP supplemented. No binding is
1124 observed.

1125 **FIGURE 5** Major and Minor Groove Binding Assay. Competition EMSAs were performed with
1126 a constant $PlzA_{WT}$ concentration of 3.5 μM, 10 nM fluorescently labeled *glpFKD*(-7/35+) probe,
1127 and a constant c-di-GMP concentration of 100 μM. Lane 1 is the probe-only control, while lane
1128 2 is the probe + protein only control. Lanes 3-7 contain increasing concentrations of methyl
1129 green: Lane 3- 0.5 μM, Lane 4- 5 μM, Lane 5- 50 μM, Lane 6- 3 100 μM, and Lane 7- 200 μM.
1130 Lanes 8-12 contain increasing concentrations of Actinomycin D: Lane 8- 0.5 μM, Lane 9- 5
1131 μM, Lane 10- 50 μM, Lane 11- 100 μM and Lane 12- 200 μM. A representative gel is shown
1132 from three independently performed experiments.

1133 **FIGURE 6** Competition EMSAs with mutagenized *glpFKD*(-7/35+) competitors. (**a**) A diagram
1134 of the generated unlabeled, mutagenized *glpFKD*(-7/+35) competitors. Comp 1 (C1) is the
1135 unlabeled, specific *glpFKD*(-7/+35) competitor. Introduced mutations are indicated in red for
1136 each mutant competitor. The transcriptional start site is the nucleotide labeled as blue.
1137 Competition EMSAs were performed with the various mutant competitors and the *glpFKD*(-
1138 7/35+) probe. (**b**) $PlzA_{WT}$, c-di-GMP, poly-dI-dC, and *glpFKD*(-7/35+) probe concentrations
1139 were held constant at 2.5 μM, 100 μM, 2.5 ng/μL, and 10 nM respectively. Lane 1 is the probe-
1140 only control, while lane 2 is the probe + protein only control. Lanes 3-5: 250, 500, & 1000x
1141 molar excess (relative to the probe) of comp 1 (unlabeled, *glpFKD*(-7/35+). Lanes 6-8: Lanes
1142 3-5: 250, 500, & 1000x mut 1 competitor. Lanes 6-8: Lanes 9-11: 250, 500, & 1000x mut 2
1143 competitor. (**c**) Concentrations for protein, effector, nonspecific competitor, and probe as in
1144 (**a**). Lane 1 is the probe-only control, while lane 2 is the probe + protein only control. Lane 3 is
1145 the specific competitor control with 1000x of comp 1. Lanes 4-6: 250, 500, & 1000x mut 4
1146 competitor. Lanes 7-9: 250, 500, & 1000x mut 3 competitor. Lanes 10-12: 250, 500, & 1000x
1147 mut 5 competitor. (**d**) $PlzA_{WT}$, c-di-GMP, poly-dI-dC, and *glpFKD*(-7/35+) probe concentrations
1148 were held constant at 3.5 μM, 100 μM, 2.5 ng/μL, and 10 nM respectively. Lane 1 is the probe-

1149 only control, while lane 2 is the probe + protein only control. Lanes 3-5: 250, 500, & 1000x
1150 molar excess (relative to the probe) of comp 1 (unlabeled, *glpFKD*(-7/35+). Lanes 6-8: Lanes
1151 3-5: 250, 500, & 1000x mut 6 competitor. The black line denotes irrelevant lanes that were
1152 removed from the same gel. (**e**) $PlzA_{WT}$, c-di-GMP, poly-dI-dC, and *glpFKD*(-7/35+) probe
1153 concentrations were held constant at 2.0 μM, 100 μM, 2.5 ng/μL, and 10 nM respectively. Lane
1154 1 is the probe-only control, while lane 2 is the probe + protein only control. Lane 3 is the
1155 specific competitor control with 1000x of comp 1. Lanes 4-6: 250, 500, & 1000x mut 7
1156 competitor. Lanes 7-9: 250, 500, & 1000x mut 8 competitor. Lanes 10-12: 250, 500, & 1000x
1157 mut 9 competitor.

1158 **FIGURE 7** Competition EMSAs with truncated *glpFKD*(-7/35+) competitors. (**a**) A diagram of
1159 the generated unlabeled, truncated *glpFKD*(-7/+35) competitors. Comp 1 (C1) is the
1160 unlabeled, specific *glpFKD*(-7/+35) competitor. Competitors 2, 3, and 4 are unlabeled
1161 *glpFKD*(-7/+35) truncated at the 5' end by either 9, 12, or 15 bases respectively. The
1162 transcriptional start site is the nucleotide labeled as blue. (**b**) $PlzA_{WT}$, c-di-GMP, poly-dI-dC,
1163 and *glpFKD*(-7/35+) probe concentrations were held constant at 2.0 μM, 100 μM, 2.5 ng/μL,
1164 and 10 nM respectively. Lane 1 is the probe-only control, while lane 2 is the probe + protein
1165 only control. Lane 3 is the specific competitor control with 1000x of comp 1. Lanes 4-6: 250,
1166 500, & 1000x comp 2 competitor. Lanes 7-9: 250, 500, & 1000x comp 3 competitor. Lanes 10-
1167 12: 250, 500, & 1000x comp 4 competitor. A representative gel is shown from two
1168 independently performed experiments.

1169 **FIGURE 8** $PlzA_{WT}$ also binds single stranded RNA and DNA. (**a**) Gradient EMSAs were
1170 performed with increasing protein concentrations of $PlzA_{WT}$ without c-di-GMP (left panel),
1171 equimolar (middle panel), or 100 μM c-di-GMP (right panel) and 10 nM fluorescently labeled
1172 *glpFKD*(UTR) RNA probe. Lane 1 is the probe-only control in each EMSA. Lane protein
1173 concentrations were as follows: Lane 2- 1 μM, Lane 3- 2.5 μM, Lane 4- 3.5 μM, Lane 5- 5 μM,
1174 Lane 6- 7.5 μM. (**b**) A similar EMSA experiment was performed with $PlzA_{RD\text{-}RD}$ with (lanes 1-
1175 8) or without (lanes 9-15) 100 μM c-di-GMP and 50 nM of *glpFKD*(UTR) RNA probe. Lane 1

1176 is a probe-only control. Protein concentrations were as follows: Lanes 2 and 9- 1 μM, Lanes
1177 3 and 10- 1.25 μM, Lanes 4 and 11- 2 μM, Lanes 5 and 12- 2.5 μM, Lanes 6 and 13- 3.3 μM,
1178 Lanes 7 and 14- 5 μM, Lanes 8 and 15- 10 μM. Riboguard was supplemented in the $PlzA_{RD-}$
1179 $_{RD}$ EMSA mixtures to a final concentration of 4 U/μL. (**c**) Gradient EMSAs were performed with
1180 increasing protein concentrations of $PlzA_{WT}$ with 100 μM c-di-GMP (left panel) or 0 μM c-di-
1181 GMP (right panel) with 10 nM fluorescently labeled ss*glpF* probe. Lanes 1 and 7 are the probe-
1182 only control in each EMSA. Lane protein concentrations were as follows: Lanes 2 and 8- 1
1183 μM, Lanes 3 and 9- 2.5 μM, Lanes 4 and 10- 3.5 μM, Lanes 5 and 11- 5 μM, Lanes 6 and 12-
1184 7.5 μM. Representative gels are shown from triplicate ($PlzA_{WT}$) or duplicate ($PlzA_{RD-RD}$) EMSA
1185 experiments. (**d**) Triplicate EMSAs from (**a**) and (**c**) were quantitated by densitometry to
1186 determine the $K_{d(app)}$ of $PlzA_{WT}$ for the *glpFKD* RNA or ssDNA interactions by nonlinear
1187 regression analysis using the one-site specific binding setting in GraphPad Prism. Errors bars
1188 represent one standard deviation (SD).

1189 **FIGURE 9** c-di-GMP significantly impacts PlzA-nucleic acid binding and PlzA prefers dsDNA
1190 and RNA over ssDNA. (**a**) The $K_{d(app)}$ values from triplicate RNA EMSAs at each tested c-di-
1191 GMP concentration in Figure 5A were analyzed by a Welch ANOVA with Dunnett's T3 multiple
1192 comparisons test. (**b**) The $K_{d(app)}$ values from triplicate ssDNA EMSAs at the tested c-di-GMP
1193 concentrations in Figure 5C were analyzed by a Welch's t-test. (**c**) The $K_{d(app)}$ values from
1194 EMSAs with PlzA and each tested nucleic acid at 100 μM c-di-GMP were compared via a
1195 Welch ANOVA with Dunnett's T3 multiple comparisons test. Bar graphs display the calculated
1196 $K_{d(app)s}$ ± the standard error of the mean (SEM). Adjusted P-values are indicated as follows:
1197 ns= $p > 0.05$, *= $p \leq 0.05$, and **= $p \leq 0.01$.

1198 **FIGURE 10** The N-terminal domain of PlzA is the nucleic acid binding domain. (**a**) The cloning
1199 strategy used to produce recombinant N ($PlzA_{NTD}$) and C ($PlzA_{CTD}$) terminal PlzA domains.
1200 The *plzA* gene portions corresponding to the N-terminus, *plzA*-NTD, and the C-terminus, *plzA*-
1201 CTD, were generated and cloned into an expression vector to produce recombinant $PlzA_{NTD}$
1202 and $PlzA_{CTD}$ proteins. Gradient EMSAs were performed with increasing protein concentrations

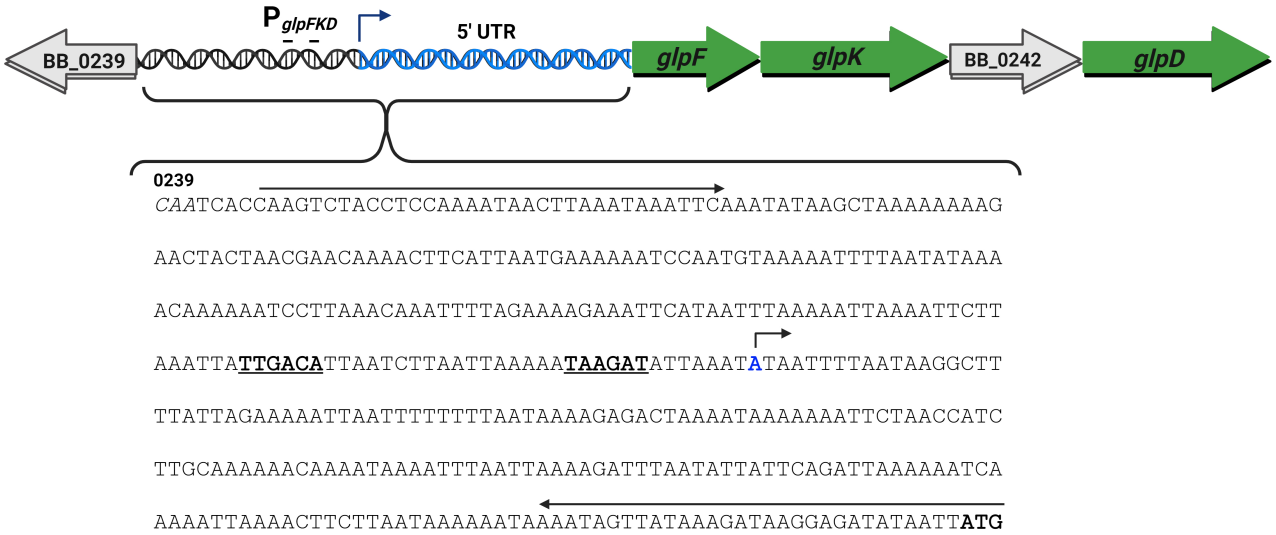
1203 of $PlzA_{NTD}$ (**b, d**) or $PlzA_{CTD}$ (**c, e**) with (lanes 1-6) or without (lanes 7-12) 100 μM c-di-GMP
1204 and 10 nM fluorescently labeled *glpFKD*(-7/35+) dsDNA (**b, c**) or *glpFKD*(UTR) RNA (**d, e**)
1205 probe. Lanes 1 and 7 are probe-only controls. Protein concentrations were as follows: Lanes
1206 2 and 8- 0.5 μM, Lanes 3 and 9- 1 μM, Lanes 4 and 10- 2.5 μM, Lanes 5 and 11- 5 μM, Lanes
1207 6 and 12- 7.5 μM. SUPERase•In RNase inhibitor was supplemented to a final concentration
1208 of 2 U/μL in RNA EMSA mixtures. The asterisk indicates $PlzA_{NTD}$-dsDNA shifts.
1209 Representative gels are shown from triplicate ($PlzA_{NTD}$) or duplicate ($PlzA_{CTD}$) EMSAs. N= 2
1210 or 3 independent experiments.

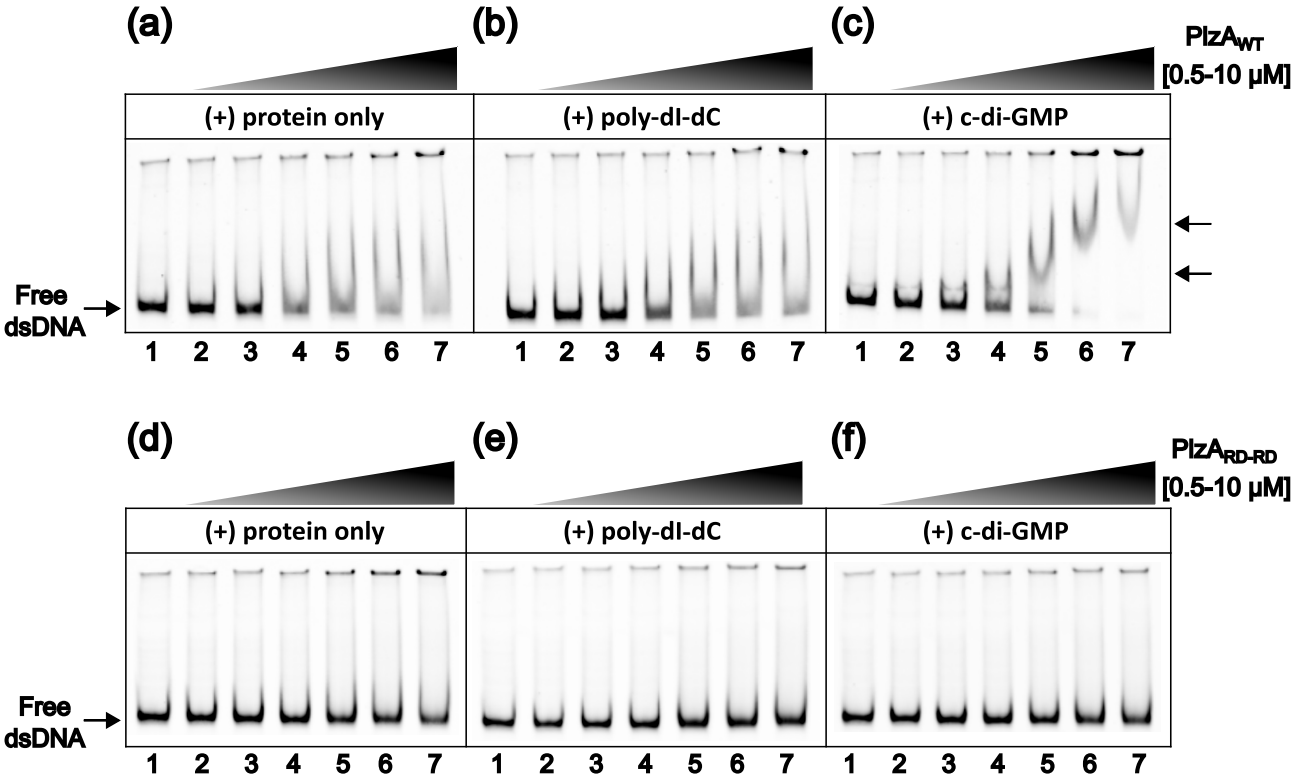
1211 **FIGURE 11** Detection of c-di-GMP in recombinant $PlzA_{NTD}$ and $PlzA_{CTD}$ by LC-MS/MS and
1212 circular dichroism of proteins. (**a**) Recombinant $PlzA_{NTD}$ and $PlzA_{CTD}$ were sent off for mass
1213 spectrometry to detect the presence of c-di-GMP. Extracted ion chromatograms of the LC-
1214 MS/MS analysis are displayed. Chromatograms depict the control c-di-GMP standard at 1 μM,
1215 recombinant $PlzA_{CTD}$, and recombinant $PlzA_{NTD}$. The LC-MS/MS analysis revealed that
1216 $PlzA_{CTD}$ co-purified with c-di-GMP, but $PlzA_{NTD}$ did not. The concentration of c-di-GMP in
1217 recombinant $PlzA_{CTD}$ was determined from a standard curve of c-di-GMP. $PlzA_{CTD}$ co-purified
1218 with 704.29 nM of c-di-GMP and approximately 1.1 % of the protein was bound with c-di-GMP.
1219 The c-di-GMP standard chromatogram shown is the same as in Figure 3. (**b**) Circular
1220 dichroism spectra of recombinant $PlzA_{WT}$, $PlzA_{RD\text{-}RD}$, $PlzA_{NTD}$, and $PlzA_{CTD}$ (see methods), and
1221 a summary table of the secondary structure analysis via BeStSel of recombinant proteins
1222 versus the PlzA PDB file 7MIE. CD spectra were derived from averaging the CD results from
1223 two different recombinant protein preparations for each protein tested. A buffer only control
1224 was performed and subtracted from the averaged measured recombinant protein CD spectra.

1225 **FIGURE 12** Computational docking analysis of PlzA with *glp* DNA and RNA. Computational
1226 docking was performed using the HDOCK server with PlzA and either (**a**) the 42 bp *glpFKD*(-
1227 7/+35) dsDNA or (**b**) the *glpFKD*(UTR) RNA sequence. In (**a**) only 26 of 42 bp of the *glpFKD*(-
1228 7/+35) sequence is shown. The dashed lines represent where bases were left out of the figure
1229 to zoom in on the protein-DNA interactions. The dsDNA model was linear, and removal of the

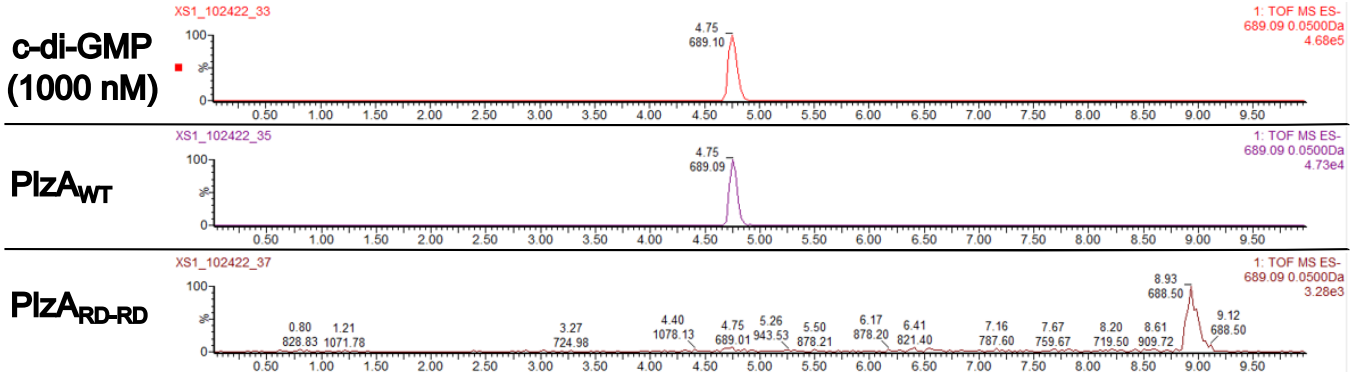
1230 bases has no impact on the interpretation of the data. The top predicted model from each
1231 computation is shown. The HDOCK docking and confidence scores for each model were -
1232 210.60 and 0.77 for PlzA-*glp* DNA, and -280.71 and 0.93 for PlzA-*glp* RNA, respectively. The
1233 protein is rotated 180° from the top to the bottom panels. Select amino acid residues that
1234 interface the nucleic acids are labeled.

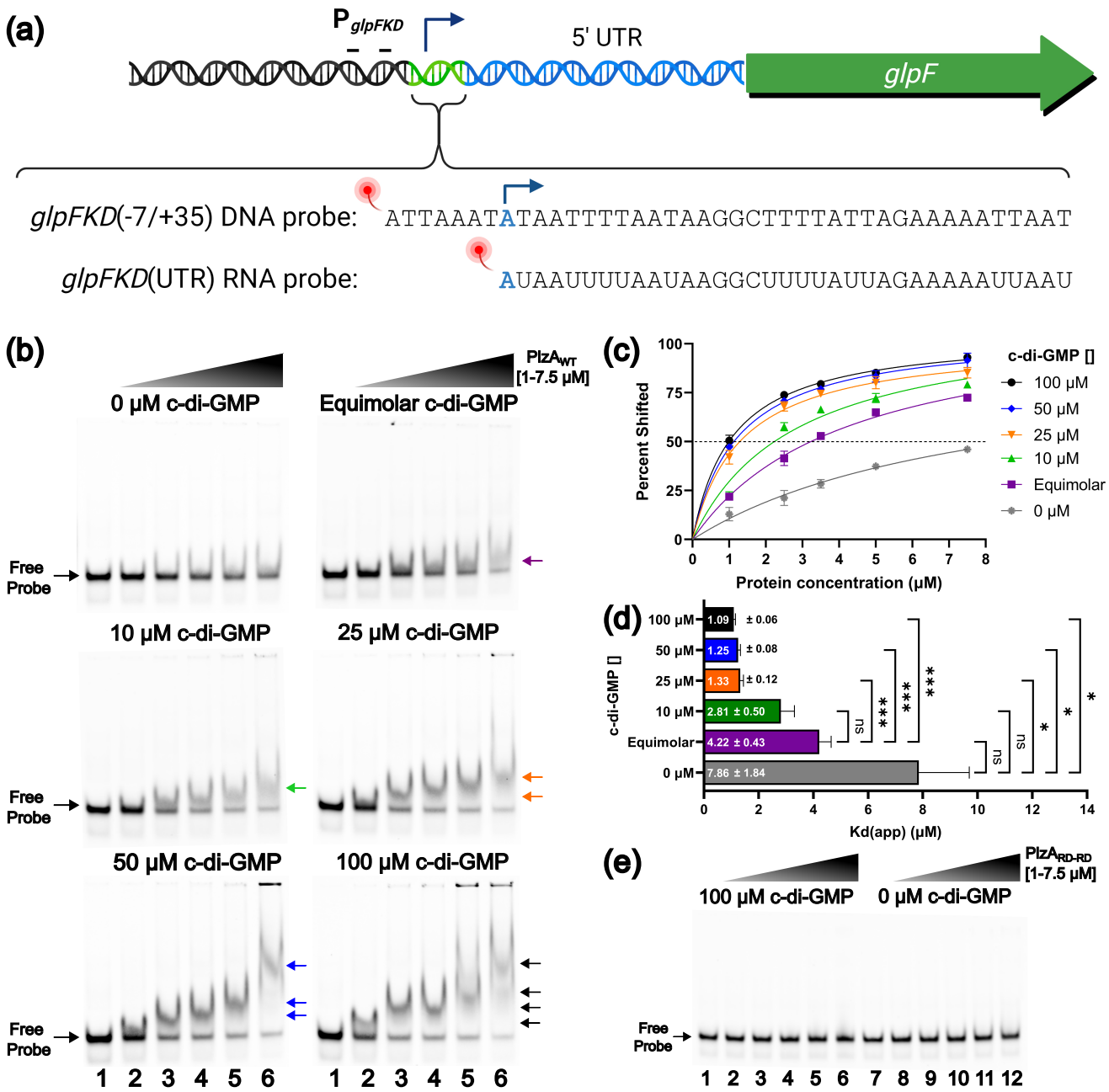
1235 **FIGURE 13** Model depicting the modulation of PlzA function by levels of c-di-GMP. Apo and
1236 holo-PlzA are hypothesized to have distinct functions in the vertebrate and tick hosts
1237 respectively. Recent studies identified PlzA RNA chaperone activity that is c-di-GMP
1238 independent (Van Gundy *et al.*, 2023). In environments with no c-di-GMP, such as the
1239 vertebrate host, apo-PlzA engages in RNA chaperone activity to modulate gene expression
1240 required for survival in the vertebrate host. The phosphodiesterases PdeA and PdeB are likely
1241 to be highly active upon transmission and into vertebrate infection, lowering c-di-GMP levels
1242 and facilitating apo-PlzA functions (Sultan *et al.*, 2010; Pitzer *et al.*, 2011; Sultan *et al.*, 2011;
1243 Novak *et al.*, 2014; Groshong *et al.*, 2021). Phosphodiesterase activity is likely to occur at all
1244 stages to regulate c-di-GMP levels and indirectly, PlzA function. When Hk1 is activated (signal
1245 unknown), Rrp1 is phosphorylated and produces c-di-GMP. Holo-PlzA can also engage in
1246 RNA chaperone activity by annealing RNAs, and potentially interact with RNAP (Grassmann
1247 *et al.*, 2023; Van Gundy *et al.*, 2023). We propose that levels of c-di-GMP impact the specific
1248 genes regulated by PlzA through altering nucleic acid binding affinity. At lower c-di-GMP
1249 levels, few holo-PlzA monomers exist resulting in weak PlzA binding affinity for certain target
1250 loci. As levels of c-di-GMP increase, more holo-PlzA molecules bind target genes with higher
1251 affinity, and PlzA could potentially multimerize on DNA.

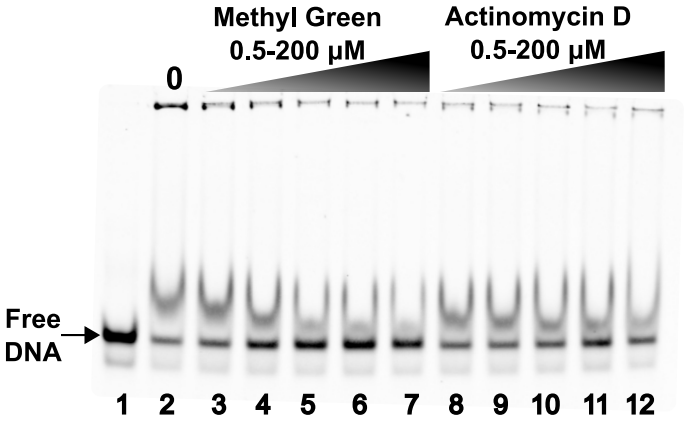

P glpFKD
5' UTR
BB_0239
glpF
glpK
BB_0242
glpD
0239
CAATCACCAAGTCTACCTCCAAAATAACTTAAATAAATTCAAATATAAGCTAAAAAAAAG
AACTACTAACGAACAAAACTTCATTAATGAAAAAATCCAATGTAAAAATTTTAATATAAA
ACAAAAAATCCTTAAACAAATTTTAGAAAAGAAATTCATAATTTAAAAATTAAAATTCTT
AAATTATTGACATTAATCTTAATTAAAAATAAGATATTAAATATAATTTTAATAAGGCTT
TTATTAGAAAAATTAATTTTTTTTAATAAAAGAGACTAAAATAAAAAAATTCTAACCATC
TTGCAAAAAACAAAATAAAATTTAATTAAAAGATTTAATATTATTCAGATTAAAAAATCA
AAAATTAAAACTTCTTAATAAAAAATAAAATAGTTATAAAGATAAGGAGATATAATTATG

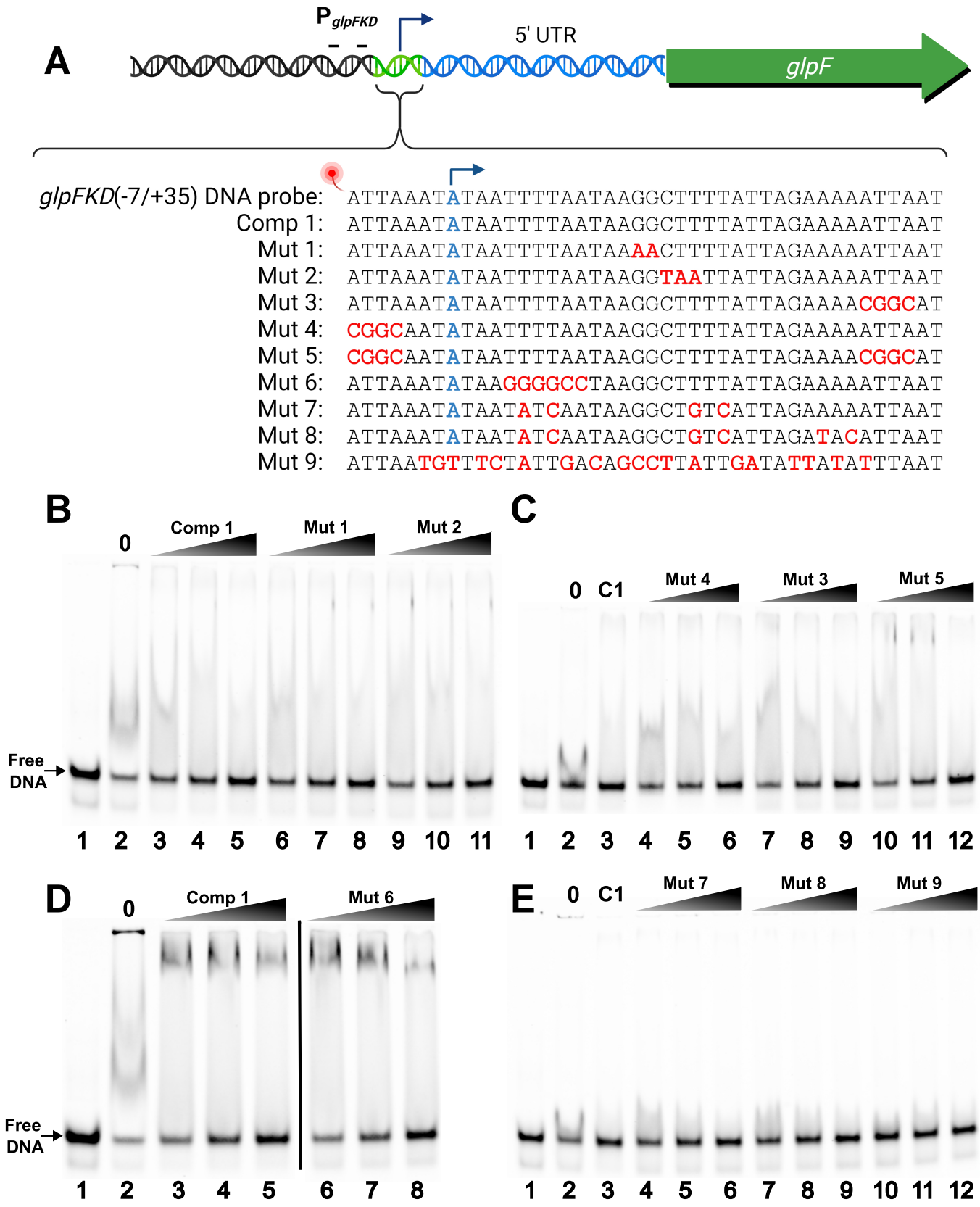

(a)
(b)
(c)
PlzA$_{WT}$
[0.5-10 μM]
(+) protein only
(+) poly-dI-dC
(+) c-di-GMP
Free
dsDNA
1 2 3 4 5 6 7
1 2 3 4 5 6 7
1 2 3 4 5 6 7
(d)
(e)
(f)
PlzA$_{RD-RD}$
[0.5-10 μM]
(+) protein only
(+) poly-dI-dC
(+) c-di-GMP
Free
dsDNA
1 2 3 4 5 6 7
1 2 3 4 5 6 7
1 2 3 4 5 6 7

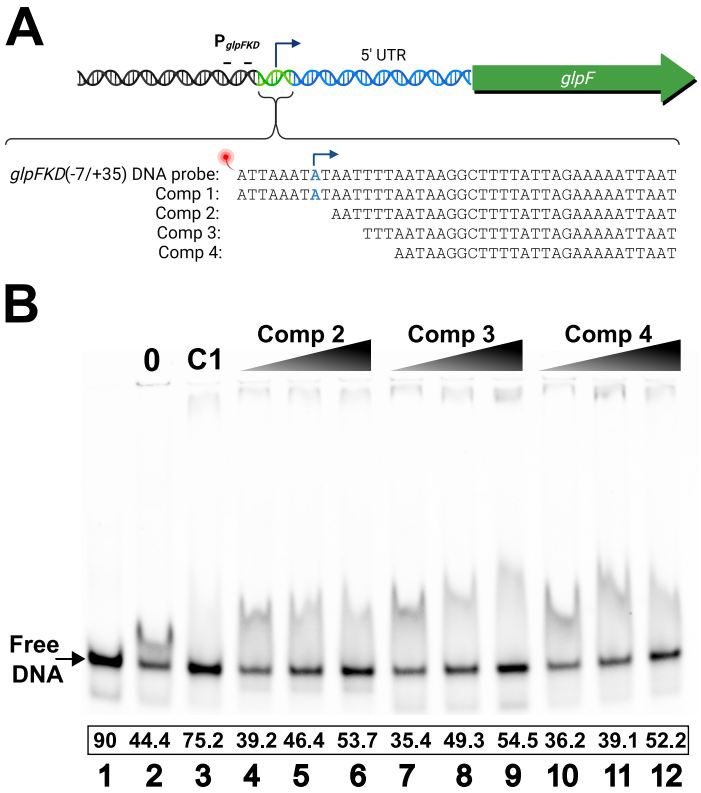
c-di-GMP; $[M-H]^-$ = m/z 689.09

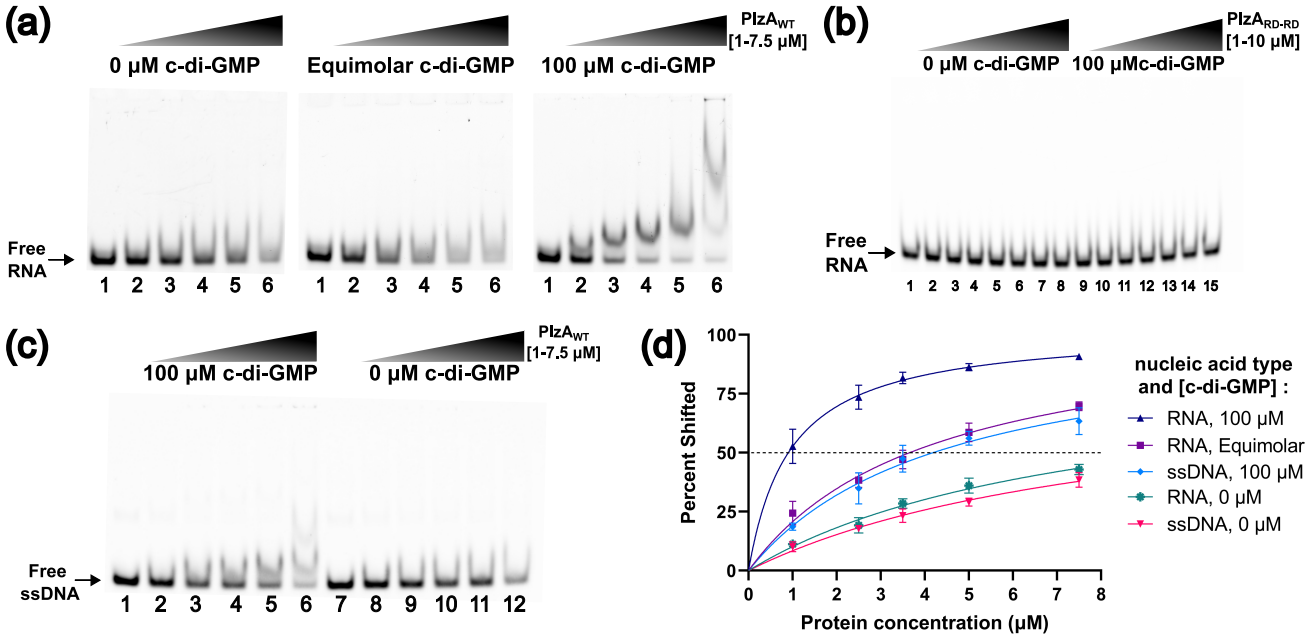


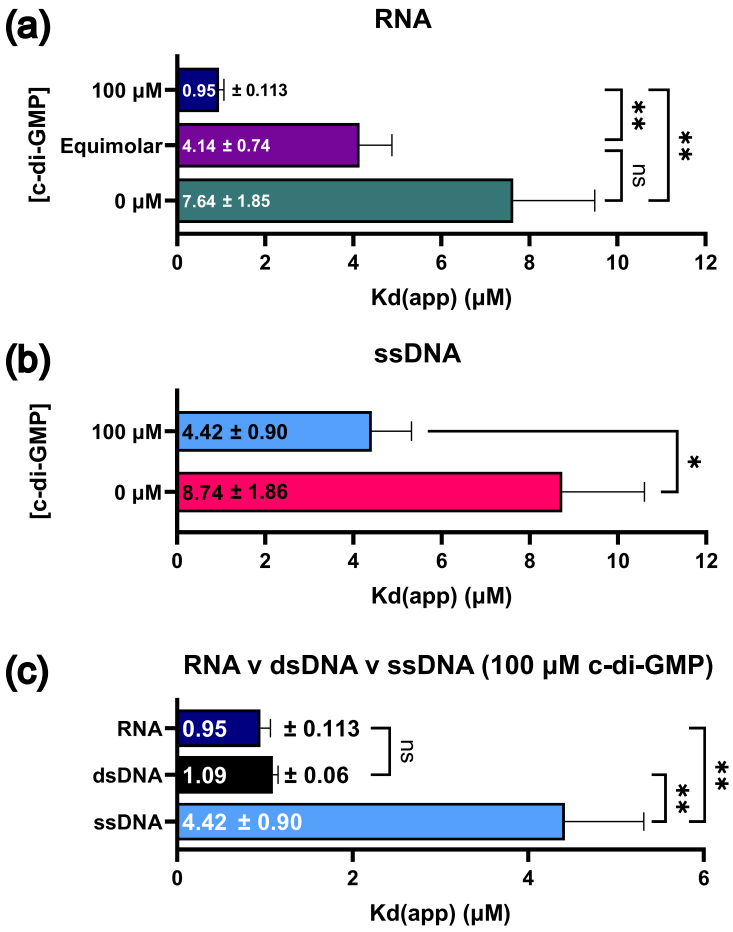

(a)
P glpFKD
5' UTR
glpF
glpFKD(-7/+35) DNA probe: ATTAAATATAATTTTAATAAGGCTTTTATTAGAAAAATTAAT
glpFKD(UTR) RNA probe: AUAAUUUUAAUAAGGCUUUUAUUAGAAAAAUUAAU
(b)
PlzA WT [1-7.5 µM]
0 µM c-di-GMP
Equimolar c-di-GMP
10 µM c-di-GMP
25 µM c-di-GMP
50 µM c-di-GMP
100 µM c-di-GMP
Free Probe
1 2 3 4 5 6
1 2 3 4 5 6
(c)
Percent Shifted
Protein concentration (µM)
c-di-GMP []
100 µM
50 µM
25 µM
10 µM
Equimolar
0 µM
(d)
c-di-GMP []
100 µM 1.09 ± 0.06
50 µM 1.25 ± 0.08
25 µM 1.33 ± 0.12
10 µM 2.81 ± 0.50
Equimolar 4.22 ± 0.43
0 µM 7.86 ± 1.84
Kd(app) (µM)
(e)
PlzA RD-RD [1-7.5 µM]
100 µM c-di-GMP
0 µM c-di-GMP
Free Probe
1 2 3 4 5 6 7 8 9 10 11 12

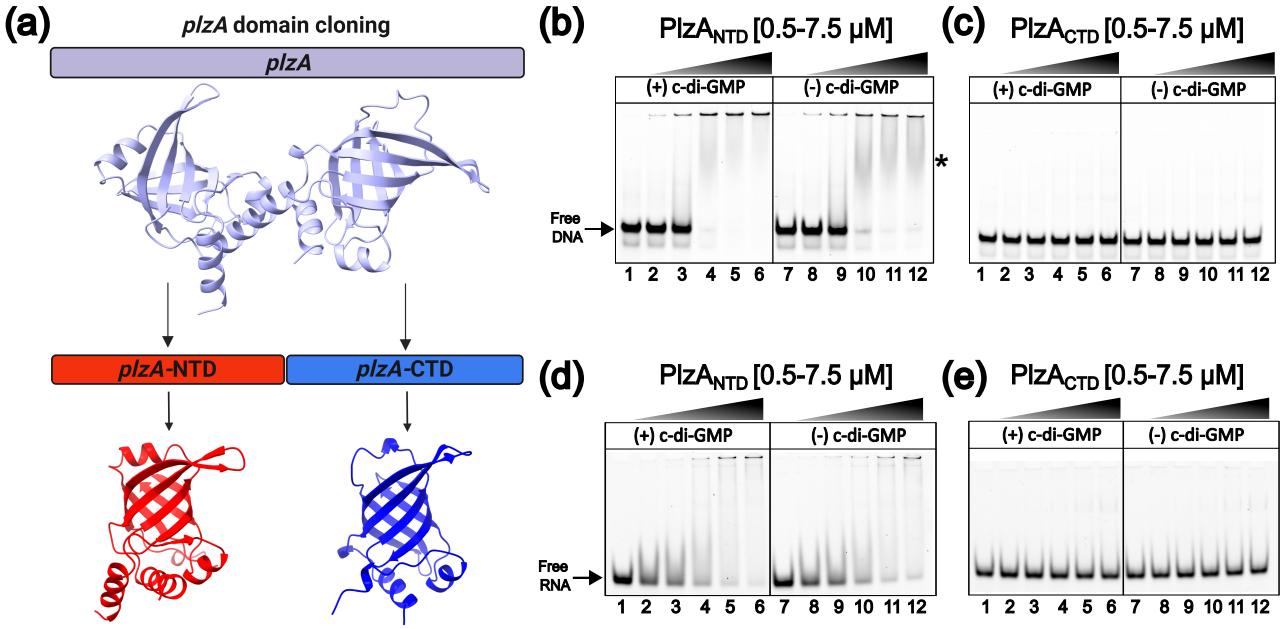

Methyl Green
0.5-200 μM
Actinomycin D
0.5-200 μM
0
Free
DNA
1 2 3 4 5 6 7 8 9 10 11 12


A
P glpFKD
5' UTR
glpF
glpFKD(-7/+35) DNA probe: ATTAAATATAATTTTAATAAGGCTTTTATTAGAAAAATTAAT
Comp 1: ATTAAATATAATTTTAATAAGGCTTTTATTAGAAAAATTAAT
Mut 1: ATTAAATATAATTTTAATAAAACTTTTATTAGAAAAATTAAT
Mut 2: ATTAAATATAATTTTAATAAGGTAATTATTAGAAAAATTAAT
Mut 3: ATTAAATATAATTTTAATAAGGCTTTTATTAGAAAACGGCAT
Mut 4: CGGCAATATAATTTTAATAAGGCTTTTATTAGAAAAATTAAT
Mut 5: CGGCAATATAATTTTAATAAGGCTTTTATTAGAAAACGGCAT
Mut 6: ATTAAATATAAGGGGCCTAAGGCTTTTATTAGAAAAATTAAT
Mut 7: ATTAAATATAATATCAATAAGGCTGTCATTAGAAAAATTAAT
Mut 8: ATTAAATATAATATCAATAAGGCTGTCATTAGATACATTAAT
Mut 9: ATTAATGTTTCTATTGACAGCCTTATTGATATTATATTTAAT
B
0
Comp 1
Mut 1
Mut 2
Free DNA
1 2 3 4 5 6 7 8 9 10 11
C
0 C1
Mut 4
Mut 3
Mut 5
1 2 3 4 5 6 7 8 9 10 11 12
D
0
Comp 1
Mut 6
Free DNA
1 2 3 4 5 6 7 8
E
0 C1
Mut 7
Mut 8
Mut 9
1 2 3 4 5 6 7 8 9 10 11 12


A
P glpFKD
5' UTR
glpF
glpFKD(-7/+35) DNA probe: ATTAAATATAATTTTAATAAGGCTTTTATTAGAAAAATTAAT
Comp 1: ATTAAATATAATTTTAATAAGGCTTTTATTAGAAAAATTAAT
Comp 2: AATTTTAATAAGGCTTTTATTAGAAAAATTAAT
Comp 3: TTTAATAAGGCTTTTATTAGAAAAATTAAT
Comp 4: AATAAGGCTTTTATTAGAAAAATTAAT
B
0
C1
Comp 2
Comp 3
Comp 4
Free DNA
90 44.4 75.2 39.2 46.4 53.7 35.4 49.3 54.5 36.2 39.1 52.2
1 2 3 4 5 6 7 8 9 10 11 12

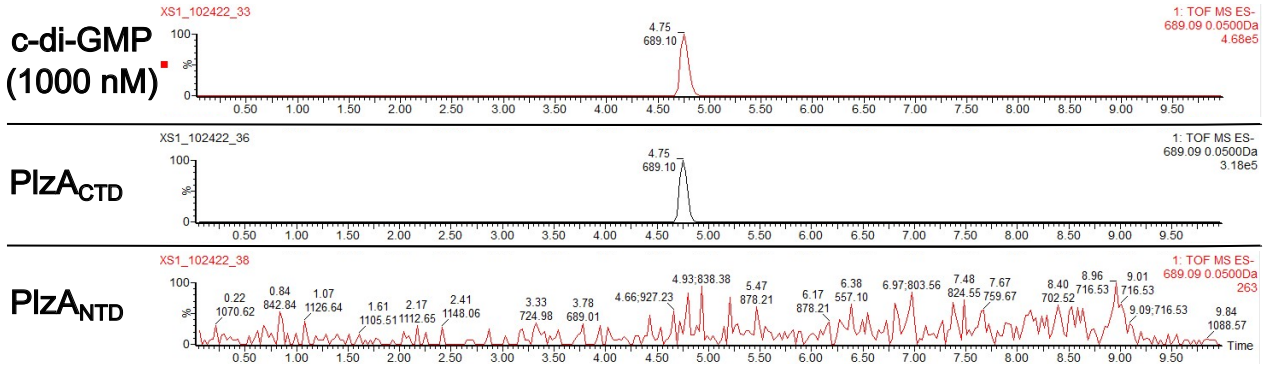

(a)
0 μM c-di-GMP
Equimolar c-di-GMP
100 μM c-di-GMP
PlzA$_{WT}$ [1-7.5 μM]
Free RNA
1 2 3 4 5 6
1 2 3 4 5 6
1 2 3 4 5 6
(b)
0 μM c-di-GMP
100 μMc-di-GMP
PlzA$_{RD-RD}$ [1-10 μM]
Free RNA
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
(c)
100 μM c-di-GMP
0 μM c-di-GMP
PlzA$_{WT}$ [1-7.5 μM]
Free ssDNA
1 2 3 4 5 6 7 8 9 10 11 12
(d)
Percent Shifted
100
75
50
25
0
0 1 2 3 4 5 6 7 8
Protein concentration (μM)
nucleic acid type and [c-di-GMP] :
RNA, 100 μM
RNA, Equimolar
ssDNA, 100 μM
RNA, 0 μM
ssDNA, 0 μM


(a)
RNA
[c-di-GMP]
100 μM
Equimolar
0 μM
0.95
± 0.113
4.14 ± 0.74
7.64 ± 1.85
**
ns
**
0
2
4
6
8
10
12
Kd(app) (μM)
(b)
ssDNA
[c-di-GMP]
100 μM
0 μM
4.42 ± 0.90
8.74 ± 1.86
*
0
2
4
6
8
10
12
Kd(app) (μM)
(c)
RNA v dsDNA v ssDNA (100 μM c-di-GMP)
RNA
dsDNA
ssDNA
0.95
± 0.113
1.09
± 0.06
4.42 ± 0.90
ns
**
**
0
2
4
6
Kd(app) (μM)

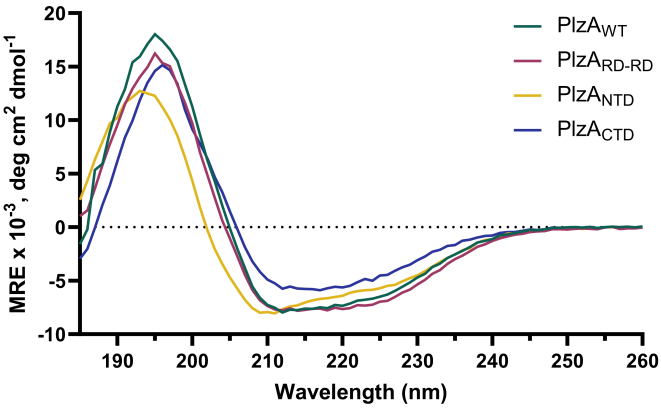

(a)
plzA domain cloning
plzA
plzA-NTD
plzA-CTD
(b)
PlzA$_{NTD}$ [0.5-7.5 μM]
(+) c-di-GMP
(-) c-di-GMP
Free DNA
*
1 2 3 4 5 6 7 8 9 10 11 12
(c)
PlzA$_{CTD}$ [0.5-7.5 μM]
(+) c-di-GMP
(-) c-di-GMP
1 2 3 4 5 6 7 8 9 10 11 12
(d)
PlzA$_{NTD}$ [0.5-7.5 μM]
(+) c-di-GMP
(-) c-di-GMP
Free RNA
1 2 3 4 5 6 7 8 9 10 11 12
(e)
PlzA$_{CTD}$ [0.5-7.5 μM]
(+) c-di-GMP
(-) c-di-GMP
1 2 3 4 5 6 7 8 9 10 11 12

(a)

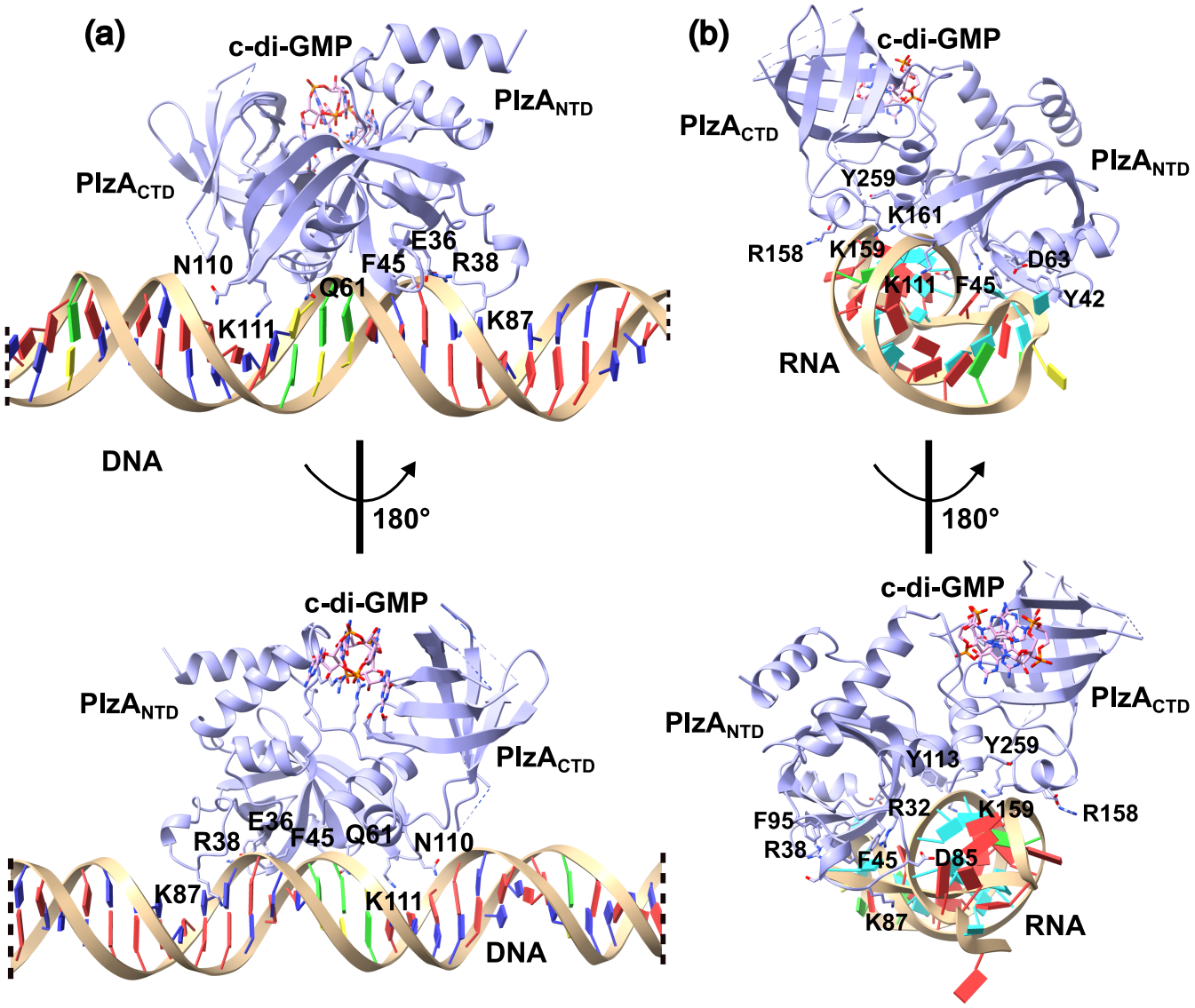
c-di-GMP; [M-H]⁻ = m/z 689.09



(b)



	PDB 7MIE	PlzA$_{WT}$	PlzA$_{RD-RD}$	PlzA$_{NTD}$	PlzA$_{CTD}$
Helix	20.5	22.1	22.7	19.7	19.1
Sheet	35.2	29.9	26.1	26.9	27
Turn	6.3	10.8	12.6	13.7	12.7
Other	38	37.2	38.5	39.6	41.2


(a)
c-di-GMP
PlzA$_{NTD}$
PlzA$_{CTD}$
N110
E36
F45
R38
Q61
K111
K87
DNA
180°
c-di-GMP
PlzA$_{NTD}$
PlzA$_{CTD}$
E36
R38
F45
Q61
N110
K87
K111
DNA
(b)
c-di-GMP
PlzA$_{CTD}$
PlzA$_{NTD}$
Y259
K161
R158
K159
D63
K111
F45
Y42
RNA
180°
c-di-GMP
PlzA$_{CTD}$
PlzA$_{NTD}$
Y113
Y259
R32
K159
R158
F95
R38
F45
D85
K87
RNA

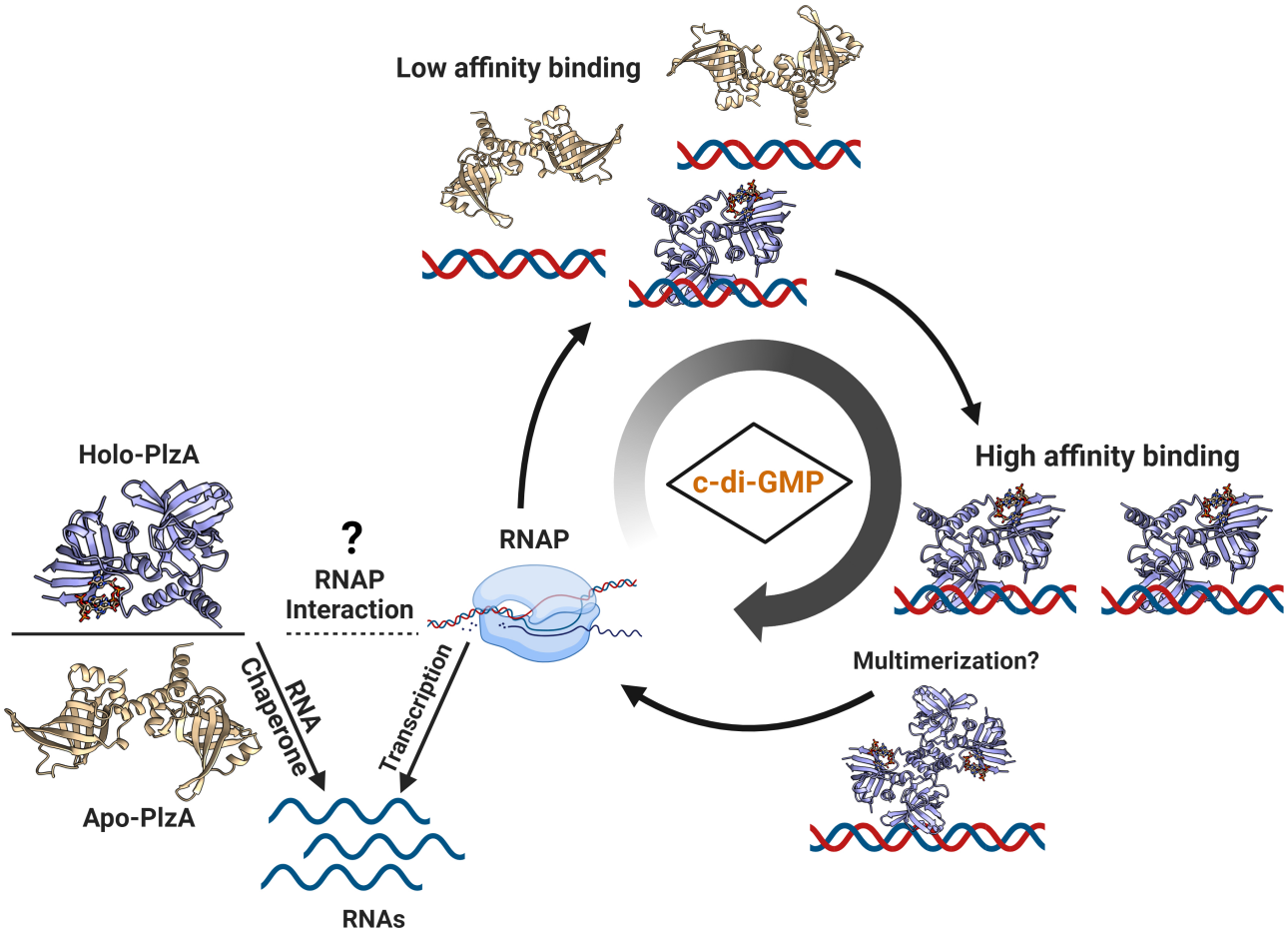

Low affinity binding
High affinity binding
c-di-GMP
Holo-PlzA
Apo-PlzA
?
RNAP
Interaction
RNA
Chaperone
Transcription
RNAs
Multimerization?

Table 1. Binding interaction summary of $PlzA_{WT}$ and the tested nucleic acids at various c-di-GMP concentrations.

***glpFKD*(-7/+35) dsDNA**			
c-di-GMP (μM)	**$K_{d(app)}$ (μM), 95% C.I. [LL, UL]**	**B_{max} (% Shifted), 95% C.I. [LL, UL]**	**Adj. R-squared**
100	1.09, [0.97, 1.23]	105.20, [102.20, 108.40]	0.986
50	1.25, [1.08, 1.45]	105.50, [101.50, 109.90]	0.981
25	1.33, [1.10, 1.59]	101.60, [96.58, 107.10]	0.971
10	2.81, [1.98, 4.02]	113.00, [98.79, 132.30]	0.940
Equimolar	4.22, [3.40, 5.26]	115.40, [104.40, 129.10]	0.983
0	7.86, [4.85, 14.06]	94.25, [71.72, 138.80]	0.950
***glpFKD*(UTR) RNA**			
c-di-GMP (μM)	**$K_{d(app)}$ (μM), 95% C.I. [LL, UL]**	**B_{max} (% Shifted), 95% C.I. [LL, UL]**	**Adj. R-squared**
100	0.95, [0.73, 1.22]	102.6, [96.76, 109.20]	0.930
Equimolar	4.14, [2.79, 6.30]	106.7, [89.50, 132.60]	0.944
0	7.64, [4.71, 13.66]	87.73, [66.88, 128.80]	0.950
***glpF* ssDNA**			
c-di-GMP (μM)	**$K_{d(app)}$ (μM), 95% C.I. [LL, UL]**	**Bmax (% Shifted), 95% C.I. [LL, UL]**	**Adj. R-squared**
100	4.42, [2.94, 6.84]	102.80, [85.32, 129.90]	0.920
0	8.74, [5.67, 14.70]	82.04, [63.53, 116.90]	0.950

Table 2. Oligonucleotides used in this study.

Target	Name and Sequence (5' → 3')	Size (bp)
Cloning and Sequencing Primers		
plzA	*plzA* F: GGCAGC**CATATG**CTTTTATCTAGAAAAATAAGAGATTATG *plzA* R: GTGGTG**CTCGAG**TTAATTGAAATAATCATGGATCAAC	786
MCS pET28a(+):*plzA*	T7 F: TAATACGACTCACTATAGGG T7 R: TAGTTATTGCTCAGCGGTGG	1026
pCR 2.1 TA clones	M13 F: GTAAAACGACGGCCAG M13 R: CAGGAAACAGCTATGAC	N/A
plzA-NTD	*plzA* 1F: ATGCTTTTATCTAGAAAAATAAGAGATTATG plzA N R: TTACTGATTTTGCCCAAGCTTTAAATC	435
plzA-CTD	*plzA* C F: ATGGATTTAAAGCTTGGGCAA plzA 1R: ATTGAAATAATCATGGATCAACATAGTATAC	378
Site-directed Mutagenesis Primers		
plzA R145-149D	*plzA*-mut F: AATCAGGACATTCATGAGGACATTATTATCGATAAAGATTCTATTAG plzA-mut R: GTCCTCATGAATGTCCTGATTTTGCCCAAGCTTTAAATCAAGAAGCTTTCC	786
EMSA Probes, Competitors, and Primers		
glpFKD(-219)	M13 F: GTAAAACGACGGCCAG M13 R 5'IRD800CW or unlabeled: CAGGAAACAGCTATGAC	654
glpFKD(-7/+35)	*glpF* F IRD800: ATTAAATATAATTTTAATAAGGCTTTTATTAGAAAAATTAAT *glpF* R: ATTAATTTTTCTAATAAAAGCCTTATTAAAATTATATTTAAT	42
ss*glpF*	*glpF* F IRD800: ATTAAATATAATTTTAATAAGGCTTTTATTAGAAAAATTAAT	42
glpFKD(UTR) RNA	*glpF* RNA 5' Alexa488: AUAAUUUUAAUAAGGCUUUUAUUAGAAAAAUUAAU	35
pCR 2.1 competitor	M13 F: GTAAAACGACGGCCAG M13 R: CAGGAAACAGCTATGAC	210

[a] Bolded and underlined nucleotides indicate restriction enzyme sites.

Borrelia burgdorferi PlzA is a cyclic-di-GMP dependent DNA and RNA binding protein

by

Nerina Jusufovic, Andrew C. Krusenstjerna, Christina R. Savage, Timothy C. Saylor, Catherine A. Brissette, Wolfram R. Zückert, Paula J. Schlax, Md A. Motaleb, and Brian Stevenson

Supplementary Information

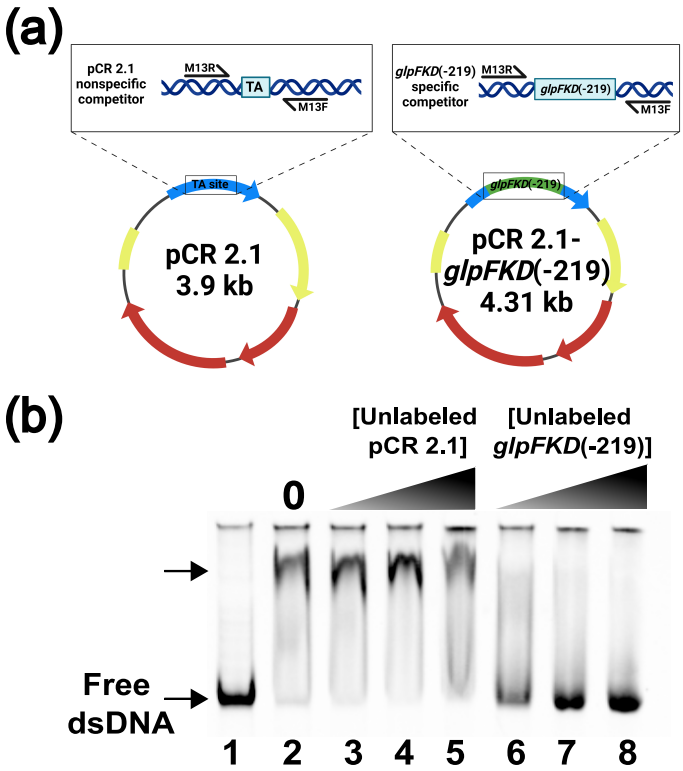


FIGURE S1 PlzA recognizes *glpFKD*(-219) and not pCR 2.1 sequences. (**a**) A schematic of the unlabeled DNA competitors derived from the empty pCR 2.1 vector backbone or the vector containing the *glpFKD*(-219) sequence. A labeled M13R and an unlabeled M13F primer were used to amplify the target DNAs from the respective plasmids to generate competitors. These competitors were then used as a control in competition EMSAs performed with labeled *glpFKD*(-219) DNA probe. (**b**) Lane 1 and lane 2 are probe only and probe + $PlzA_{WT}$ controls respectively. All lanes contained a c-di-GMP concentration of 100 μM, a final poly-dI-dC concentration of 2.5 ng/μL, and a labeled *glpFKD*(-219) probe concentration of 50 nM. Lanes 2-8 contained a constant $PlzA_{WT}$ protein concentration of 1.5 μM. The unlabeled pCR 2.1 competitor was titrated into the reactions at increasing molar excess relative to labeled probe: Lane 3-300 nM (6x), lane 4- 500 nM (10x), and lane 5- 1000 nM (20x). Unlabeled *glpFKD(*-219) competitor was titrated into the reactions at increasing molar excess relative to labeled probe: Lane 6-300 nM (6x), lane 7- 500 nM (10x), and lane 8- 600 nM (12x). The free arrow indicates PlzA-*glpFKD*(-219) complexes. The pCR 2.1 competitor could not compete away the PlzA-*glpFKD*(-219) complex at any tested concentration. Unlabeled *glpFKD*(-219) competitor was able to compete at all tested concentrations as low as 6x molar excess.