bioRxiv preprint doi: https://doi.org/10.1101/2023.01.30.526351; this version posted September 4, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

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

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.30.526351; this version posted September 4, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

ABSTRACT

The PilZ domain-containing protein, PlzA, is the only 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. In studies with the promoter and 5' UTR of the glycerol catabolism operon, *glpFKD*, the affinity of PlzA for nucleic acids increased 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. PlzA is a dual-domain protein consisting of 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 dictate PlzA regulatory targets and functions.

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.30.526351; this version posted September 4, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

INTRODUCTION

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

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

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

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.30.526351; this version posted September 4, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

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

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

Given the multiple, conflicting hypotheses regarding PlzA function, as well as the potential for additional mechanisms between liganded and unliganded PlzA, we sought to further characterize PlzA nucleic acid binding function. We decided to examine a known regulatory target of PlzA, *glpFKD*, the glycerol catabolism operon. The *glpFKD* operon promotes *B. burgdorferi* survival in ticks (Pappas *et al.*, 2011; Corona and Schwartz, 2015; Helble *et al.*, 2021). After a tick has completed the digestion of its blood meal, it is hypothesized that the midgut is devoid of glucose and many other nutrients. To survive in that environment, Lyme disease spirochetes would need to utilize alternate carbon sources such as glycerol, which is

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.30.526351; this version posted September 4, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

available in the tick host (Lee and Baust, 1987). The *glpFKD* operon is regulated through the Hk1/Rrp1 pathway, and borrelial *plzA* mutants are defective in glycerol metabolism (Rogers *et al.*, 2009; He *et al.*, 2011; Caimano *et al.*, 2015; Zhang *et al.*, 2018). Further, PlzA was found to both positively and negatively affect the *glp* operon, at higher and lower c-di-GMP levels, respectively (Zhang *et al.*, 2018).

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

RESULTS

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

Previous studies demonstrated that c-di-GMP and PlzA modulate expression of the glycerol metabolism operon (Rogers *et al.*, 2009; He *et al.*, 2011; Caimano *et al.*, 2015; Zhang *et al.*, 2018). Activation of *glpFKD* is critical for spirochete survival in between blood meals of the tick vector, which requires catabolism of glycerol (He *et al.*, 2011; Pappas *et al.*, 2011; Zhang *et al.*, 2018). Given that requirement, we hypothesized that PlzA might bind 5' of the *glpFKD* operon. To that end, we sought to determine if PlzA could bind DNA between *glpFKD* and the upstream ORF BB_0239, which encodes a deoxyguanosine/deoxyadenosine kinase. The transcriptional boundaries and promoter elements of the *glpFKD* operon were previously

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.30.526351; this version posted September 4, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

mapped (Adams *et al.*, 2017; Grove *et al.*, 2017). A fluorescently labeled DNA fragment was generated by PCR amplification from a plasmid containing the 410 bp intergenic region between the upstream gene, and the start of *glpF* (BB_0240) and is denoted as *glpFKD*(-219) (Figure 1).

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

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

To further assess the requirement of c-di-GMP for PlzA DNA binding activity, a double mutant PlzA protein ($PlzA_{RD\text{-}RD}$) that cannot bind c-di-GMP was produced by site-directed mutagenesis which replaced key arginine residues in the conserved c-di-GMP binding motif, RxxxR, with aspartic acids (Ryjenkov *et al.*, 2006; Amikam and Galperin, 2006; Chou and Galperin, 2016). These mutations resulted in the amino-acid substitutions R145D and R149D

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.30.526351; this version posted September 4, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

within the c-di-GMP binding site of PlzA and have previously been shown to abolish c-di-GMP binding (Freedman *et al.*, 2009; Mallory *et al.*, 2016; Singh *et al.*, 2021). The *glpFKD*(-219) EMSAs were repeated with $PlzA_{RD-RD}$ (Figure 2D-F). Binding to *glpFKD*(-219) was not observed under any tested conditions with $PlzA_{RD-RD}$.

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

Noting that c-di-GMP enhanced DNA binding by PlzA yet free DNA decreased with purified recombinant PlzA without adding exogenous c-di-GMP, we investigated the possibility that PlzA was being purified with c-di-GMP already bound. Purified recombinant proteins were prepared and assayed by liquid chromatography-tandem mass spectrometry. Extracted ion chromatograms showed the presence of c-di-GMP in $PlzA_{WT}$ protein preparations but not in $PlzA_{RD-RD}$ preparations (Figure 4). The concentration of c-di-GMP present in recombinant $PlzA_{WT}$ was extrapolated from a c-di-GMP standard curve at 99.29 nM. A total of 31.7 μM of protein had been analyzed, indicating that approximately 0.3% of the protein was bound with c-di-GMP. Prior studies have reported dissociation constants of 1.25 to 6.25 μM of PlzA for c-di-GMP, which are typical K_D values for PilZ receptors and c-di-GMP (Pitzer *et al.*, 2011; Mallory *et al.*, 2016). The amount of co-purified c-di-GMP detected is lower than the K_D. This is likely due to removal of some c-di-GMP during the protein purification process, as well as *E. coli* maintaining low cellular concentrations of c-di-GMP (~40-80 nM) (Sarenko *et al.*, 2017; Junkermeier and Hengge, 2023). Taken together, these data indicate that PlzA DNA binding to this region of the *glpFKD* operon is specific and c-di-GMP dependent.

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.30.526351; this version posted September 4, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

c-di-GMP levels modulate PlzA DNA binding affinity

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

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

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.30.526351; this version posted September 4, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

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

Next, we calculated the apparent binding affinity ($K_{D(app)}$) under these experimental conditions for PlzA for *glpFKD*(-7/+35) at the various c-di-GMP concentrations (Figure 5C). The calculations were performed under the assumption that a single molecule of PlzA binds to a single molecule of DNA. As noted above, it is possible that the protein-nucleic acid ratio is not 1:1, which could affect the calculated values. The values described below are the $K_{D(app)}$ ± SEM (standard error of the mean). Without any supplemented c-di-GMP, the extrapolated $K_{D(app)}$ of $PlzA_{WT}$-*glpFKD*(-7/+35) was calculated as 7.86 (± 1.84 μM), however we again note that some c-di-GMP is already present in the $PlzA_{WT}$ protein preparations (see above and Figure 4). Binding affinity increased approximately 1.5-fold under equimolar c-di-GMP and protein concentrations with a calculated $K_{D(app)}$ of 4.22 (± 0.43 μM). Under saturating c-di-GMP conditions (exceeding the K_D of PlzA for c-di-GMP), binding affinity increased further. Supplementing binding reactions with 10 μM c-di-GMP improved PlzA-DNA binding 2.5-fold ($K_{D(app)}$ of 2.81 (± 0.50 μM). An approximately 6- to 7-fold increase in binding affinity was observed when the binding reaction was supplemented with either 25, 50, or 100 μM of c-di-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 μM), respectively.

To determine if the calculated $K_{D(app)}$ values differed significantly, Welch's ANOVA followed by a Dunnett's T3 multiple comparisons test was performed (Figure 5D). All calculated $K_{D(app)}$ values differed significantly ($p \leq 0.05$) from the 0 μM c-di-GMP $K_{D(app)}$ except for the equimolar and 10 μM c-di-GMP conditions. Compared to the equimolar condition, the 25, 50, and 100 μM $K_{D(app)}$ values differed significantly ($p \leq 0.001$), while the 10 μM c-di-GMP $K_{D(app)}$ value did not ($p > 0.05$). When comparing the 10, 25, 50, and 100 μM $K_{D(app)}$ values of $PlzA_{WT}$-*glpFKD*(-7/+35) interaction, there were no significant differences ($p > 0.05$). The mutant $PlzA_{RD\text{-}RD}$ did not bind *glpFKD*(-7/+35) under any tested conditions (Figure 5E). Overall, we conclude that levels of c-di-GMP significantly impact PlzA-DNA binding.

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.30.526351; this version posted September 4, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

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

Several borrelial proteins that were initially identified as having affinities for DNA have subsequently been found to also bind RNA (Jutras *et al.*, 2013a; Savage *et al.*, 2018; Krusenstjerna *et al.*, 2023). Given the role the 5' UTR of the *glpFKD* operon plays in its regulation, we hypothesized PlzA could potentially also bind RNA. RNA probes used were designed to correspond to only transcribed regions (Figure 5A). The *glpFKD*(-7/+35) counterpart RNA probe is designated *glpFKD*(UTR). EMSAs were performed with increasing concentrations of $PlzA_{WT}$ with 0 μM, 100 μM, or equimolar c-di-GMP levels and labeled *glpFKD*(UTR) (Figure 6A). We did not test additional c-di-GMP concentrations that were tested for impacts on DNA binding for the RNA, as no significant difference was found between 100 μM c-di-GMP and the other tested c-di-GMP concentrations. $PlzA_{WT}$ bound the RNA, which was enhanced by addition of c-di-GMP (Figure 6A). The binding pattern to the *glpFKD*(UTR) RNA was comparable to the *glpFKD*(-7/+35) DNA, with increased protein and c-di-GMP concentrations leading to an increase in the size of the PlzA-RNA shifts.

EMSAs were repeated with the $PlzA_{RD\text{-}RD}$ mutant protein at 0 or 100 μM c-di-GMP (Figure 6B). The $PlzA_{RD\text{-}RD}$ protein did not bind *glpFKD*(UTR) RNA at any tested protein concentration, regardless of whether c-di-GMP was included.

Given that PlzA can bind to both the DNA and RNA, we sought to determine if it could also bind the cognate ssDNA corresponding to the coding strand. This is relevant as ssDNA is exposed at open transcriptional complexes. The forward, IRDye labeled oligonucleotide used to generate the *glpFKD*(+7/-35) dsDNA probe was used as a single-stranded DNA probe (ss*glpF*) in EMSAs. No evident complexes were formed when exogenous c-di-GMP was excluded from EMSA reactions, but free ssDNA did decrease at the highest tested PlzA concentrations of 3.5-7.5 μM (Figure 6C right panel). Adding 100 μM c-di-GMP resulted in the formation of PlzA-ssDNA complexes but at higher concentrations of protein than required to achieve complex formation with dsDNA and RNA (Figure 6C left panel).

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.30.526351; this version posted September 4, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

We next quantified the binding affinity of PlzA for the *glp* derived RNA and ssDNA sequences through densitometry of the EMSA shifts, and the apparent $K_{D(app)s}$ calculated as described above (Figure 6D). As with dsDNA, the binding affinity of PlzA for RNA was weakest when no exogenous c-di-GMP was supplemented with a calculated $K_{D(app)}$ of 7.64 (± 1.85 μM). Binding improved approximately 2-fold in equimolar c-di-GMP conditions with a $K_{D(app)}$ of 4.14 (± 0.74 μM). The $K_{D(app)}$ was 0.95 (± 0.11 μM) when the c-di-GMP concentration was 100 μM, an 8-fold increase in PlzA binding affinity for the *glpFKD*(UTR) RNA as compared to when no c-di-GMP was added. There was a significant difference between the $PlzA_{WT}$-*glpFKD*(UTR) RNA $K_{D(app)}$ value calculated at 100 μM c-di-GMP and both the equimolar and 0 μM c-di-GMP values ($p \leq 0.01$) (Figure 7A). No significant differences were detected between the $K_{D(app)}$ values calculated at 0 μM and equimolar c-di-GMP levels ($p > 0.05$).

Although PlzA bound the ss*glpF* probe, it had a substantially lower affinity for ssDNA than dsDNA or RNA (Figure 6D). The $K_{D(app)}$ of the $PlzA_{WT}$-ss*glpF* interaction was 8.74 (± 1.86 μM) with 0 μM c-di-GMP and 4.42 (± 0.90 μM) with 100 μM c-di-GMP. These values were found to be statistically significant from one another ($p < 0.05$) (Figure 7B). Only a 2-fold increase in binding affinity of PlzA for the single-strand DNA was observed when the c-di-GMP concentration was increased to 100 μM, whereas the same concentration of c-di-GMP increased PlzA binding affinity 7-fold for dsDNA and 8-fold for RNA.

Lastly, we also compared the binding affinities of PlzA for all three tested nucleic acids (Figure 7C). The calculated $K_{D(app)}$ values of $PlzA_{WT}$ for each nucleic acid in the presence of 100 μM c-di-GMP were compared with a Welch's ANOVA and Dunnett's T3 multiple comparisons test. No significant difference ($p > 0.05$) was detected between the PlzA $K_{D(app)}$ values for dsDNA (1.09 ± 0.06 μM) and RNA (0.95 ± 0.11 μM), indicating that PlzA binds the *glpFKD* dsDNA and RNA with similar affinity. Both the dsDNA and RNA $K_{D(app)s}$ were significantly different from the calculated ssDNA $K_{D(app)}$ (4.42 ± 0.90 μM) ($p \leq 0.01$), suggesting that PlzA prefers *glp*-derived dsDNA and RNA over ssDNA. All the binding data presented throughout the

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.30.526351; this version posted September 4, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

manuscript are summarized in Table 1. Collectively, our data indicate that PlzA binding activity is c-di-GMP dependent for all tested *glpFKD* nucleic acids.

Dissection of PlzA reveals the nucleic acid binding domain

The first discovered c-di-GMP binding receptors were the PilZ domain-containing proteins (Ryjenkov *et al.*, 2006; Amikam and Galperin, 2006; Galperin and Chou, 2020). The PilZ domain consists of two conserved c-di-GMP binding motifs, RxxxR and [D/N]ZSXXG (Ryjenkov *et al.*, 2006; Amikam and Galperin, 2006; Chou and Galperin, 2016; Cheang *et al.*, 2019). The recently solved crystal structure of *B. burgdorferi* PlzA revealed a unique dual-domain topology consisting of an amino-terminal PilZ-like domain, called PilZN3, which is connected to the carboxy-terminal PilZ domain via a linker domain that contains the RxxxR c-di-GMP binding motif (Singh *et al.*, 2021). The first 141 amino acid residues make up the N-terminus of PlzA ($PlzA_{NTD}$), while residues 142-261 make up the C-terminal ($PlzA_{CTD}$) PilZ domain. Having determined that PlzA is a c-di-GMP-dependent DNA and RNA binding protein, we next sought to determine which domain of PlzA binds nucleic acid. To generate recombinant $PlzA_{NTD}$ and $PlzA_{CTD}$ domains, truncated *plzA* genes called *plzA*-NTD and *plzA*-CTD were cloned to encode the residues described above, which correspond to each respective domain (Figure 8A).

EMSAs were performed with labeled *glpFKD*(-7/+35) DNA probe and increasing concentrations of $PlzA_{NTD}$ or $PlzA_{CTD,}$ with or without 100 μM c-di-GMP. $PlzA_{NTD}$ bound DNA independently of c-di-GMP, as supplementation with c-di-GMP did not affect binding (Figure 8B). $PlzA_{CTD}$ contains the motifs required for c-di-GMP binding, but the $PlzA_{CTD}$ domain did not bind to DNA with or without added c-di-GMP (Figure 8C). It was also determined that $PlzA_{NTD}$ could bind *glpFKD*(UTR) RNA with or without c-di-GMP, while $PlzA_{CTD}$ does not (Figure 8D and E respectively).

We also performed the LC/MS-MS analysis on $PlzA_{NTD}$ and $PlzA_{CTD}$ recombinant proteins to determine if any di-nucleotides were present (Figure 9A). Co-purified c-di-GMP was detected

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.30.526351; this version posted September 4, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

in $PlzA_{CTD}$ but not $PlzA_{NTD}$ by LC/MS-MS. The concentration of c-di-GMP present in recombinant $PlzA_{CTD}$ was determined as 704.29 nM. A total of 66.7 μM of protein was sent for analysis indicating approximately 1.1 % of $PlzA_{CTD}$ was bound with c-di-GMP. This indicates that $PlzA_{CTD}$ retained the ability to bind c-di-GMP but cannot bind nucleic acid.

To confirm that the truncated proteins were stable, circular dichroism (CD) was performed, and the CD spectra were obtained for $PlzA_{WT}$, $PlzA_{RD\text{-}RD}$, $PlzA_{NTD}$, and $PlzA_{CTD}$ in the binding buffer used for EMSAs (Figure 9B). The CD results show that the mutated proteins retained stable structures matching that of $PlzA_{WT}$. Therefore, the lack of DNA binding by $PlzA_{RD\text{-}RD}$ and $PlzA_{CTD}$ is not due to improper folding.

Computational docking of PlzA-nucleic acid interactions

To gain a better understanding of the interactions between PlzA and DNA or RNA, we performed a computational docking analysis with the HDOCK server (Huang and Zou, 2014; Yan *et al.*, 2017a; Yan *et al.*, 2017b; Yan *et al.*, 2020). The *glpFKD*(-7/+35) DNA and *glpFKD*(UTR) RNA sequences were input as ligands, while the PlzA crystal structure (PDB ID: 7mie) was input as the receptor molecule. The modeling results indicate that interactions between PlzA and *glpFKD*(-7/+35) DNA are facilitated primarily through the N-terminal domain (Figure 10a). Beta-strands six and seven of the N-terminus are extended and protrude outward, intercalating a major groove of the DNA 5' to the protein docking site (Figure 10a top panel). Polar and positively charged amino acids, including N110 and K111, are located at the loop region where beta-strands 6 and 7 converge, with the potential to directly interact with bases of the DNA (Rooman *et al.*, 2002; Sathyapriya and Vishveshwara, 2004). Interactions with a minor groove of the DNA are predicted to occur through the loop region of beta-strands 2 and 3 and an unstructured linker region (Figure 10a bottom panel). Several positively charged residues, including arginine (R32 and R38) and lysine (K81), as well as an aromatic phenylalanine residue (F45), are predicted to directly interact with the minor groove and backbone of the DNA. An additional lysine residue K87 located in the unstructured region,

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.30.526351; this version posted September 4, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

protrudes and interacts directly with bases of another major groove of *glpFKD*(-7/+35). No residues of the C-terminal domain are predicted to interact directly with the DNA; however, four residues are interfacing the DNA (K161, I213, D258, and N261). Residue K161 hovers over the DNA backbone potentially interacting nonspecifically via electrostatic interactions. The computational docking data supports our EMSA data showing that $PlzA_{NTD}$ binds DNA while $PlzA_{CTD}$ does not (Figures 8B and C).

The docking analysis was also performed with $PlzA_{WT}$ and *glpFKD*(UTR) RNA (Figure10B). Again, the N-terminal domain of PlzA facilitates most of the interactions. Similar regions of the protein that interacted with dsDNA are implicated in the interaction with RNA including the residues of the loop region of beta-strands 6 and 7 (K111 and Y113), the loop region of beta-strands 2 and 3 (R38, Y42, and F45) and the unstructured linker region (R32, D85, K87, F92, and F95). Unlike the docking model with dsDNA, the PlzA-RNA docking model indicates more involvement of the C-terminal domain in RNA binding. Several residues of alpha-helixes 5 (D155, R158, K159, and K161) and 6 (D258 and Y259) in the C-terminal domain of PlzA, are predicted to interact with RNA. Although our EMSA results show that $PlzA_{CTD}$ does not bind any tested nucleic acids, the docking results indicate it could help stabilize $PlzA_{NTD}$-nucleic acid interactions or could be involved in target recognition. Several PlzA residues predicted to be involved in nucleic-acid binding are currently being investigated by our lab.

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

DISCUSSION

Cyclic-di-GMP has wide-ranging effects on many bacterial processes. The responses to c-di-GMP are commonly mediated through c-di-GMP binding effector proteins. PlzA is the only

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.30.526351; this version posted September 4, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

chromosomally and universally encoded c-di-GMP binding protein of the Lyme disease spirochetes. It has been shown to be crucial for borrelial infection processes and in the regulation of the metabolism of alternative carbon sources during colonization of vector ticks (Pappas *et al.*, 2011; Pitzer *et al.*, 2011; He *et al.*, 2014; Mallory *et al.*, 2016; Zhang *et al.*, 2018; Groshong *et al.*, 2021). Despite its singularity and importance to *B. burgdorferi*, little was previously known about the biochemical functions of PlzA outside of binding c-di-GMP. Here, we provide evidence that PlzA is a novel c-di-GMP-dependent nucleic acid-binding protein.

While PlzA lacks an obvious canonical DNA binding motif, we hypothesized that PlzA might be a nucleic acid binding protein given its role in the regulation of the *glpFKD* operon. Binding was only observed by the wild-type PlzA protein but not the mutant, $PlzA_{RD\text{-}RD}$, which is incapable of binding c-di-GMP. It has previously been shown that structural rearrangements are induced upon the binding of c-di-GMP to PlzA (Mallory *et al.*, 2016; Groshong *et al.*, 2021). Without bound c-di-GMP, PlzA has a structure that is flexible, potentially preventing the interactions necessary for DNA binding. This hypothesis is supported by the fact that crystallization of PlzA was only achieved with c-di-GMP bound (Singh *et al.*, 2021). We surmise that binding of c-di-GMP to the C-terminal PilZ domain results in a conformational change of PlzA into a more rigid structure permitting nucleic acid binding through residues of the N-terminal domain.

We identified the N-terminal domain of PlzA as the nucleic acid binding domain and that $PlzA_{NTD}$ bound DNA independently of c-di-GMP. In contrast, the C-terminal domain, $PlzA_{CTD}$, did not bind any tested nucleic acids although it still bound c-di-GMP. While $PlzA_{NTD}$ can interact with nucleic acids, these interactions are not as strong as $PlzA_{WT}$. Binding of nucleic acid by $PlzA_{NTD}$ resulted in smeared shifts that lacked defined bands. This suggests instability of the $PlzA_{NTD}$-nucleic acid complexes resulting in dissociation of the complexes during electrophoresis. Noting that binding of c-di-GMP appears to stiffen interactions between the N- and C-terminal domains, we hypothesize the C-terminal domain stabilizes the interactions of $PlzA_{NTD}$ with nucleic acids when c-di-GMP bound. On the other hand, we posit that the

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.30.526351; this version posted September 4, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

flexibility of apo-PlzA results in the C-terminal domain disrupting the ability of the N-terminal domain to bind DNA. Although the N- and C- terminal PlzA domains are similar, structural comparisons of $PlzA_{NTD}$ and $PlzA_{CTD}$ reveal distinct differences (Figure 8). The N-terminus of $PlzA_{NTD}$ contains an alpha-helix that is absent in the N-terminus of $PlzA_{CTD}$. Additionally, an alpha-helix is present at the beta-barrel pore of $PlzA_{NTD}$ but not in the $PlzA_{CTD}$. These differences suggest distinct roles in PlzA-nucleic acid binding dynamics. We are currently exploring the sequence motifs recognized by PlzA and the exact amino acid residues involved in nucleic acid binding.

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

We also observed probe signal in the gel wells at high protein and c-di-GMP concentrations. This could be due to protein aggregation. Rather than simply being a negative consequence of environmental stress, recent studies indicate a role for protein aggregation in regulation.

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.30.526351; this version posted September 4, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Specifically, certain conditions can induce a protein to aggregate, and thus be sequestered, altering gene expression to permit bacterial adaptation to a specific environment (Schramm *et al.*, 2019). Whether high levels of c-di-GMP promote such processes remains to be tested, but c-di-GMP levels certainly alter PlzA binding affinity.

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

PlzA bound to a sequence adjacent to the promoter and extending 35 bp into the 5' UTR of *glpF*. This site has been shown to positively affect *glpFKD* transcript levels and is a site where two other borrelial regulators, SpoVG and BadR, also bind (Grove *et al.*, 2017; Savage *et al.*, 2018; Zhang *et al.*, 2022; Saylor *et al.*, 2023). BadR, the *Borrelia* Host Adaption Regulator, has recently been identified as a repressor of *glpFKD*, while PlzA has previously been shown to exert both positive and negative effects on *glpFKD*, at high and low levels of c-di-GMP respectively (Zhang *et al.*, 2018; Zhang *et al.*, 2022). Although SpoVG can bind the DNA and

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.30.526351; this version posted September 4, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

RNA of *glpFKD*, the transcriptional consequences of these interactions are unknown (Savage *et al.*, 2018; Saylor *et al.*, 2023). How PlzA binding affects *glpFKD* expression, and how these proteins work in tandem to regulate glycerol catabolism, is currently being investigated. Multi-layered regulation of glycerol catabolism is found in other bacteria indicating this is not a phenomenon unique to *Borrelia* (Bong *et al.*, 2019).

During the preparation of this manuscript, Van Gundy et al. published a report identifying RNA chaperone activity by PlzA (Van Gundy *et al.*, 2023). The authors concluded that this activity is independent of c-di-GMP. This greatly deviates from our results that c-di-GMP is required for nucleic acid binding. This was acknowledged by those authors who suggested that differences in methodologies and conditions might have contributed to the differences in observed results. Some of these differences include their use of non-borrelial derived RNAs, a different PlzA site-mutant, equimolar concentrations of c-di-GMP, and filter binding assays used to calculate dissociation constants. Indeed, the differences in methodology and conditions between the two studies make it difficult to draw comparisons between calculated apparent K_D values, as they are only constant under the given conditions tested (Jarmoskaite *et al.*, 2020). However, some facets of the studies can still be compared. Their study utilized short, unstructured, non-borrelial derived RNAs, while the nucleic acids used in our study are borrelial derived, although examining just one locus PlzA is known to affect. Indeed, our data indicates that PlzA preferentially binds double-stranded/structured nucleic acids. No condition we tested yielded low nanomolar range binding. Regardless, the methodology yielded reproducible data that allowed for the identification of PlzA c-di-GMP-dependent nucleic acid-binding.

Given our finding that c-di-GMP levels greatly affect the affinity of PlzA for nucleic acids, we tested binding to both *glpFKD* DNA and RNA under equimolar conditions like those used by Van Gundy et al. In our hands, equimolar concentrations of c-di-GMP yielded significantly reduced nucleic acid binding affinities than those exceeding the calculated K_D of PlzA for c-di-GMP. Considering the higher K_D of the PlzA-c-di-GMP interaction (1.25-6.25 μM), under

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.30.526351; this version posted September 4, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

equimolar conditions reaching the K_D of PlzA-c-di-GMP depends on the protein concentration used. The protein concentrations used in the other study likely never reached a concentration that exceeded the K_D of PlzA for c-di-GMP to saturate half of the PlzA molecules, except at the higher protein ranges tested (1-2 μM). This could explain the observed weaker RNA binding and/or chaperone activity with holo-PlzA.

We chose to evaluate saturating c-di-GMP concentrations to evaluate the maximal nucleic acid binding activity of holo-PlzA and mitigate potential variability in the calculated apparent K_D values. It was also never confirmed if their PlzA protein preparations co-purified with c-di-GMP, which could contribute to observed binding in their apo-PlzA conditions. We also did not observe binding under any tested conditions with the $PlzA_{RD\text{-}RD}$ mutant, which cannot bind c-di-GMP. Lastly, the goals for our studies differed greatly as they sought to study RNA chaperone activity, while we set out to investigate PlzA nucleic acid binding at the *glpFKD* operon. Our two studies, however, do corroborate that PlzA can bind RNA and suggest that this function plays a role in the regulation of borrelial gene expression. We agree that further investigation is required to unravel the molecular mechanisms of PlzA's interactions with nucleic acids, and the roles that those interactions play in gene regulation in this important pathogen.

It is known that the affinity and specificity of proteins for target nucleic acids are impacted by co-effectors (Carey, 2022). We observed that adding exogenous c-di-GMP increased binding affinity of PlzA for *glp*-derived DNA and RNA. Studies in other bacteria with multiple PilZ domain proteins have suggested that the differential affinity of the various PilZ receptors for c-di-GMP sequentially regulates the processes that respond to c-di-GMP (Pultz *et al.*, 2012; Junkermeier and Hengge, 2023). Given that PlzA is the only PilZ domain protein in most *B. burgdorferi* isolates, an alternative to this could be that cellular levels of c-di-GMP alter the affinity and/or specificity of PlzA. We propose a model in which c-di-GMP availability alters the functionality of PlzA, and thus, the targets regulated by PlzA could depend on the c-di-GMP levels available at different stages of the enzootic life cycle (Figure 11).

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.30.526351; this version posted September 4, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

During vertebrate infection, no c-di-GMP is produced, favoring apo-PlzA functions such as RNA chaperone activity (Van Gundy *et al.*, 2023). When c-di-GMP is present, holo-PlzA can also anneal RNAs and potentially interact with RNAP (Grassmann *et al.*, 2023; Van Gundy *et al.*, 2023). At lower levels of c-di-GMP, PlzA nucleic-acid binding affinity is low for certain target genes such as *glpFKD*. As c-di-GMP levels increase, such as in a tick or during the blood meal, binding affinity for target nucleic acids is increased and potential multimerization of PlzA molecules on DNA could occur. Lastly, we hypothesize that an unknown signal might regulate PlzA activity at high concentrations resulting in the release of holo-PlzA from target DNAs. This hypothesis requires further investigation, as does overall regulation of c-di-GMP in *B. burgdorferi*.

In conclusion, PlzA is a c-di-GMP dependent nucleic acid binding protein. While other c-di-GMP binding receptors have been identified as DNA-binding proteins, PlzA is the first that has been demonstrated to also bind RNA, thus expanding the functional diversity of c-di-GMP binding proteins. Given that PlzA can bind to both DNA and RNA, it could have a role in coordinating transcriptional complexes, and it supports the notion that PlzA can interact with RNAP (Grassmann *et al.*, 2023). Our study examined PlzA nucleic-acid binding at a locus known to be affected by PlzA, the *glpFKD* catabolism operon, but sites/ligands of higher affinity could exist. It remains to be seen what additional targets in the *B. burgdorferi* genome/transcriptome fall under the umbrella of the PlzA regulon. Our work identifying PlzA as a novel nucleic acid-binding protein provides a basis for its functional mechanism. This will further inform our understanding of how *B. burgdorferi* regulates gene expression in response to environmental cues and help unravel the PlzA regulon.

MATERIAL AND METHODS

Routine manipulations and molecular cloning

Oligonucleotides used in this study are listed in Table 2 and were ordered from Integrated DNA Technologies (IDT). *B. burgdorferi* genomic DNA was purified utilizing the E.Z.N.A Tissue

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.30.526351; this version posted September 4, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

DNA kit (OMEGA). Standard and high-fidelity PCR were performed with 2X DreamTaq Green PCR Master Mix (Thermo Scientific) or Q5 High-Fidelity 2X Master Mix (NEB), respectively. Plasmids were isolated using QIAprep Spin Miniprep kits (Qiagen) according to the Miraprep protocol (Pronobis *et al.*, 2016). All plasmid constructs were transformed into chemically competent *E. coli* Top10 (Invitrogen) or DH5α (Invitrogen) strains for cloning and plasmid maintenance. Positive clones were identified by colony PCR screening with the appropriate primer sets, and DNA sequencing was performed by Eurofins Genomics LLC (Louisville, KY).

Plasmid construction

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

Generation of the site-directed PlzA R145D-R149D double mutant

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

Recombinant protein overexpression and purification

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.30.526351; this version posted September 4, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

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

For protein extraction, cell pellets were thawed on ice and resuspended in equilibration buffer (20mM sodium phosphate, 300 mM NaCl, 10 mM imidazole, pH 8.0) and lysed via sonication. Lysates were clarified by centrifugation at 23700 x *g* for 20 min at 4°C, and the supernatant was retained. When further clarification was required due to lysate viscosity, the supernatants were passed through a sterile 0.22 µM Millex-GS Syringe Filter Unit (MilliporeSigma). Protein purification was performed via column affinity chromatography using either HisPur Ni-NTA resin (Thermo Scientific) or the Magne-His purification system (Promega Corporation) for $PlzA_{RD\text{-}RD}$ per manufacturer protocol. Eluted protein solutions were dialyzed overnight into EMSA buffer (50 mM Tris-HCl [pH 7.5], 50 mM KCl, 1 mM dithiothreitol (DTT), 1 mM ethylenediaminetetraacetic acid (EDTA) [pH 8.0], 1 mM phenylmethanesulfonyl fluoride (PMSF), 10% glycerol (v/v), and 0.01% Tween-20). The dialyzed proteins were then concentrated using Amicon Ultra 3K ($PlzA_N$ and $PlzA_C$) or 10K ($PlzA_{WT}$ and $PlzA_{RD\text{-}RD}$) centrifugal filter units (MilliporeSigma). Aliquots of dialyzed proteins were assessed for purity by SDS-PAGE and Coomassie brilliant blue staining. Final recombinant protein concentrations were determined using Quick Start Bradford Protein Assays (Bio-Rad) with bovine serum

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.30.526351; this version posted September 4, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

albumin as the reference protein for standard curves. Purified protein aliquots were stored at -80°C.

Nucleic acid fragments design and generation

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

Smaller probes <60 bp long were produced by annealing complementary unlabeled and IRDye 800-labeled oligonucleotides at equimolar concentrations at 95°C for 5 min and were gradually cooled overnight at RT. All probes were aliquoted and stored at -20°C until further use. Unlabeled and 5' IRDye 800 (LI-COR Biosciences) labeled oligonucleotides were synthesized by IDT. Alexa Fluor 488 conjugated RNA probes corresponding to the transcribed regions of the respective DNA probes were also synthesized by IDT.

Electrophoretic mobility shift assays (EMSA)

EMSA reactions were performed in EMSA buffer at RT with a final concentration of dsDNA probe of 10 nM or RNA probe of either 10 nM or 50 nM. EMSA reaction mixtures were supplemented with 3'5' c-di-GMP (Sigma-Aldrich) at various concentrations ranging from 0 to 100 μM. c-di-GMP was not added to the gel or running buffer. Experiments using equimolar concentrations of c-di-GMP are relative to the final protein concentration in the reaction.

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.30.526351; this version posted September 4, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Proteins were co-incubated with c-di-GMP for 5 min before the addition of any nucleic acids. Prior to the addition of labeled probe, the non-specific competitor poly-dI-dC (Roche) was added to the EMSA reactions at a final concentration of 2.5 ng/µL, and the reaction mix was allowed to incubate another 5 min (Larouche *et al.*, 1996). Labeled probes were added last, and the reactions were subsequently incubated for an additional 10 min prior to loading onto gels. Competition EMSAs were conducted similarly but with the unlabeled, specific competitor added after poly-dI-dC and then allowed to incubate with the reaction prior to the addition of the labeled probe. Regarding RNA EMSAs, when RNase contamination was evident in recombinant protein preparations, RiboGuard (Lucigen) or SUPERase•In (Ambion) RNase inhibitor was added to the EMSA mixtures before the addition of RNA to a final concentration of 4 U/µL and 2 U/µL respectively to inhibit RNase activity. The final reaction volume was 10 µL prior to the addition of loading dye.

For visualization during electrophoresis, 2 µL of EMSA loading dye (0.8 mg/mL Orange G, 15 mg/mL Ficoll 400) was added to each reaction prior to gel loading. Novex TBE 6 or 10% gels (Invitrogen) were pre-run in 0.5x TBE running buffer (pH ~8.4 +/- 0.5), at 100V for a minimum of 30 min. Due to the complexes of the larger *glpFKD*(-219) probe and higher concentrations of PlzA not exiting the well, the pH of the running buffer was adjusted to a pH of 8.0 immediately prior to electrophoresis which allowed the complexes to enter the gel matrix. The pH was not adjusted for EMSAs using nucleic acid fragments <60 bp as complexes readily exited the well at the standard pH of 0.5x TBE running buffer. The entire reaction mixture was then loaded onto the pre-run gels and resolved at 100 V for 60-90 min at RT. EMSA images were acquired with a ChemiDoc Imaging System (Bio-Rad).

Densitometry and statistics

Densitometry of EMSAs was performed using Image Lab 6.1 (Bio-Rad) software. Analyses were performed on triplicate EMSAs. Lanes as well as free and shifted bands were added manually, and intensity values were reported as band percentage. For dsDNA EMSAs, free

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.30.526351; this version posted September 4, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

dsDNA band percentages were normalized to the probe only control band percentage. The free DNA band percentage values were then used to determine the percent shifted of the respective probe relative to the free probe. For RNA and ssDNA EMSAs, lanes and bands were selected manually, and the shifted bands were analyzed to obtain the percent shifted values. Probe signal stuck in the gel wells, attributed to potential protein aggregation, were not considered shifts, and thus omitted from densitometric analysis.

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

Docking analysis and protein modeling

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

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.30.526351; this version posted September 4, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

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

Cyclic di-nucleotide LC-MS analysis

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

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.30.526351; this version posted September 4, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Data were processed using Waters Masslynx software. Extracted ion chromatograms were performed to look for the presence of c-di-GMP (m/z 689.09), c-di-AMP (m/z 657.09) and c-GAMP (673.09). The intensity values were converted to relative abundance as determined from the base peak values. The c-di-GMP concentration was determined from the peak area of the target compound against a standard curve of known c-di-GMP concentrations. These concentrations were calculated using the Targetlynx tool in the Waters Masslynx software. The percentage of protein bound with c-di-GMP was calculated from the total protein and calculated c-di-GMP concentrations in each sample.

Circular dichroism spectroscopy

The CD analysis was performed on $PlzA_{WT}$, $PlzA_{RD\text{-}RD}$, $PlzA_{NTD}$, and $PlzA_{CTD}$ in EMSA dialysis buffer (detailed above) at the University of Kentucky Center for Molecular Medicine Protein Core. Briefly, 400 µL of each protein solution at a concentration of 180-210 µg/mL was pipetted into a 1 mm quartz cuvette and loaded into a Jasco J-810 spectrophotometer. The samples were scanned at 0.1 nm increments and CD spectra collected for wavelengths 200-260 nM. EMSA buffer alone or supplemented with 50 µM c-di-GMP were also analyzed and used as controls. Data were visualized in GraphPad Prism 9.5.1.

AVAILABILITY

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

ACKNOWLEDGEMENT

We dedicate this manuscript to our friend and colleague Christina R. Savage, Ph.D., whose initial studies on PlzA were invaluable to this work. We also thank Jessamyn Moore and Tatiana Castro Padovani for their support of the studies undertaken. Circular Dichroism was performed at the University of Kentucky Center for Molecular Medicine Protein Core, and we thank Martin Chow for his assistance on these studies. The LC-MS/MS to detect di-nucleotides was performed by the Mass Spectrometry and Metabolomics Core at Michigan State

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.30.526351; this version posted September 4, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

University. A sincere thank you to Tony Schilmiller at MSU for all his help with the mass spectrometry services. Figures 1, 3a, 5a, 8a and 11 were created via BioRender.

FUNDING

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

CONFLICT OF INTEREST

None declared.

REFERENCES

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

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

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

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

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

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

Caimano, M.J., Dunham-Ems, S., Allard, A.M., Cassera, M.B., Kenedy, M., and Radolf, J.D. (2015) Cyclic di-GMP modulates gene expression in Lyme disease spirochetes at the tick-mammal interface to promote spirochete survival during the blood meal and

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.30.526351; this version posted September 4, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

tick-to-mammal transmission. *Infect Immun* **83**: 3043–3060 http://dx.doi.org/10.1128. Accessed May 1, 2021.

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

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

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

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

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

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

Fraser, C.M., Casjens, S., Huang, W.M., Sutton, G.G., Clayton, R., Lathigra, R., *et al.* (1997) Genomic sequence of a Lyme disease spirochaete, *Borrelia burgdorferi*. *Nature* **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.

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.30.526351; this version posted September 4, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

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**.

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

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

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

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

Jarmoskaite, I., Alsadhan, I., Vaidyanathan, P.P., and Herschlag, D. (2020) How to measure and evaluate binding affinities. *Elife* **9**: 1–34.

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

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.30.526351; this version posted September 4, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

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

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

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

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

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

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

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

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

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

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

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

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.30.526351; this version posted September 4, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

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

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

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

Pappas, C.J., Iyer, R., Petzke, M.M., Caimano, M.J., Radolf, J.D., and Schwartz, I. (2011) *Borrelia burgdorferi* requires glycerol for maximum fitness during the tick phase 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.

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.30.526351; this version posted September 4, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

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.

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

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

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

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

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

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

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

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.30.526351; this version posted September 4, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

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

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

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

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

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

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

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

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

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

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

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

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

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.30.526351; this version posted September 4, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

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

Zhang, J.J., Ranghunandanan, S., Wang, Q., Lou, Y., and Yang, X.F. (2022) Mechanism of repression of the glycerol utilization operon in *Borrelia burgdorferi*. *bioRxiv* **PREPRINT:NOT PEER REVIEWED**

FIGURES LEGENDS

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

FIGURE 2 DNA binding by PlzA is c-di-GMP dependent. (**a**) EMSA with increasing concentrations of $PlzA_{WT}$ showing binding to *glpFKD*(-219) DNA. The mixture was incubated for 10 min before electrophoresis. (**b**) The same EMSA mixture as in (**a**) but with a final concentration of 2.5 ng/μL of the nonspecific competitor poly-dI-dC. Protein was incubated with poly-dI-dC for 5 min prior to the addition of labeled *glpFKD*(-219) probe. (**c**) The same EMSA mixture as in (**b**) but supplemented with 100 μM c-di-GMP. Protein was incubated with c-di-GMP for 5 min prior to the addition of any nucleic acids. Arrows indicate PlzA- *glpFKD*(-219) complexes. Lane 1 in (**a-c**) is the probe only control, containing 50 nM of probe. Protein concentrations are as follows in each subsequent lane: Lane 2: 0.5 μM, Lane 3: 1 μM, Lane

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.30.526351; this version posted September 4, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

4: 2.5 μM, Lane 5: 5 μM, Lane 6: 7.5 μM, and Lane 7: 10 μM. (**d**), (**e**), (**f**) EMSA experiments as in (**a**), (**b**), and (**c**), respectively, but performed with the mutant $PlzA_{RD-RD}$ protein. All conditions were performed identically as with $PlzA_{WT}$. Each panel is a representative gel from separate EMSAs which were performed in triplicate ($PlzA_{WT}$) or duplicate ($PlzA_{RD-RD}$).

FIGURE 3 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.

FIGURE 4 Detection of c-di-GMP in recombinant $PlzA_{WT}$ and $PlzA_{RD-RD}$ by LC-MS/MS. Recombinant $PlzA_{WT}$ and $PlzA_{RD-RD}$ were sent off for mass spectrometry to detect c-di-GMP. Extracted ion chromatograms of the LC-MS/MS analysis for detection of c-di-GMP are displayed. Chromatograms depict the control c-di-GMP standard at 1 μM, recombinant $PlzA_{WT}$, and recombinant $PlzA_{RD-RD}$. The LC-MS/MS analysis reveals $PlzA_{WT}$ co-purified with c-di-GMP, but $PlzA_{RD-RD}$ did not. The concentration of c-di-GMP in recombinant $PlzA_{WT}$ was determined from a standard curve of c-di-GMP. $PlzA_{WT}$ co-purified with 99.29 nM of c-di-GMP

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.30.526351; this version posted September 4, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

and approximately 0.3% of the protein was bound with c-di-GMP. Black lines indicate chromatograms unrelated to this figure that were removed.

FIGURE 5 Levels of c-di-GMP alter DNA binding affinity of $PlzA_{WT}$. (**a**) A diagram of the generated DNA probe, designated as *glpFKD*(-7/+35), corresponding to a 42 bp sequence adjacent to the -10 site and proceeding 35 bp into the 5' UTR. An RNA probe was made corresponding to only nucleotides that are within the 5 'UTR which was designated as *glpFKD*(UTR) RNA. Probes were conjugated to a fluorescent molecule for the detection of binding in EMSAs. (**b**) EMSAs were performed with increasing protein concentrations of $PlzA_{WT,}$ and 10 nM fluorescently labeled probe at varying c-di-GMP concentrations (0, equimolar, 10, 25, 50, or 100 μM). For each EMSA shown, the indicated c-di-GMP concentration was held constant in all lanes. For each EMSA shown, lanes 1 and 7 are probe-only controls. Protein concentrations were as follows: Lane 2- 1 μM, Lane 3- 2.5 μM, Lane 4- 3.5 μM, Lane 5- 5 μM, Lane 6- 7.5 μM. A final concentration of 2.5 ng/μL of the nonspecific competitor poly-dI-dC was added to each lane. Arrows indicate higher-order complexes formed as protein concentration increases. A representative gel for each tested c-di-GMP concentration is shown from triplicate EMSAs. (**c**) Triplicate EMSAs were quantitated by densitometry to determine the $K_{d(app)}$ of the $PlzA_{WT}$-*glpFKD*(-7/35+) interaction by nonlinear regression analysis using the one-site specific binding setting in GraphPad Prism. Errors bars represent one standard deviation (SD). (**d**) Bar graph displaying the calculated $K_{d(app)}$ ± the standard error of the mean (SEM). The $K_{d(app)}$ values from three EMSAs were analyzed by a Welch ANOVA with Dunnett's T3 multiple comparisons test via GraphPad Prism. Adjusted P-values are indicated as follows: ns= $p > 0.05$, *= $p \leq 0.05$, and ***= $p \leq 0.001$. Comparisons between 10, 25, 50, and 100 μM c-di-GMP are not indicated on the graph as they were not significant. (**e**) A similar EMSA experiment was performed as in Figure 5B but with $PlzA_{RD\text{-}RD}$ and either 100 (lanes 1-6) or 0 (lanes 7-12) μM of c-di-GMP supplemented. No binding is observed.

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.30.526351; this version posted September 4, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 6 $PlzA_{WT}$ also binds single stranded RNA and DNA. (**a**) Gradient EMSAs were performed with increasing protein concentrations of $PlzA_{WT}$ without c-di-GMP (left panel), equimolar (middle panel), or 100 μM c-di-GMP (right panel) and 10 nM fluorescently labeled *glpFKD*(UTR) RNA probe. Lane 1 is the probe-only control in each EMSA. Lane protein concentrations were as follows: Lane 2- 1 μM, Lane 3- 2.5 μM, Lane 4- 3.5 μM, Lane 5- 5 μM, Lane 6- 7.5 μM. (**b**) A similar EMSA experiment was performed with $PlzA_{RD\text{-}RD}$ with (lanes 1-8) or without (lanes 9-15) 100 μM c-di-GMP and 50 nM of *glpFKD*(UTR) RNA probe. Lane 1 is a probe-only control. Protein concentrations were as follows: Lanes 2 and 9- 1 μM, Lanes 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, Lanes 7 and 14- 5 μM, Lanes 8 and 15- 10 μM. Riboguard was supplemented in the $PlzA_{RD\text{-}RD}$ EMSA mixtures to a final concentration of 4 U/μL. (**c**) Gradient EMSAs were performed with increasing protein concentrations of $PlzA_{WT}$ with 100 μM c-di-GMP (left panel) or 0 μM c-di-GMP (right panel) with 10 nM fluorescently labeled ss*glpF* probe. Lanes 1 and 7 are the probe-only control in each EMSA. Lane protein concentrations were as follows: Lanes 2 and 8- 1 μM, Lanes 3 and 9- 2.5 μM, Lanes 4 and 10- 3.5 μM, Lanes 5 and 11- 5 μM, Lanes 6 and 12-7.5 μM. Representative gels are shown from triplicate ($PlzA_{WT}$) or duplicate ($PlzA_{RD\text{-}RD}$) EMSA experiments. (**d**) Triplicate EMSAs from (**a**) and (**c**) were quantitated by densitometry to determine the $K_{d(app)}$ of $PlzA_{WT}$ for the *glpFKD* RNA or ssDNA interactions by nonlinear regression analysis using the one-site specific binding setting in GraphPad Prism. Errors bars represent one standard deviation (SD).

FIGURE 7 c-di-GMP significantly impacts PlzA-nucleic acid binding and PlzA prefers dsDNA and RNA over ssDNA. (**a**) The $K_{d(app)}$ values from triplicate RNA EMSAs at each tested c-di-GMP concentration in Figure 6A were analyzed by a Welch ANOVA with Dunnett's T3 multiple comparisons test. (**b**) The $K_{d(app)}$ values from triplicate ssDNA EMSAs at the tested c-di-GMP concentrations in Figure 6C were analyzed by a Welch's t-test. (**c**) The $K_{d(app)}$ values from EMSAs with PlzA and each tested nucleic acid at 100 μM c-di-GMP were compared via a Welch ANOVA with Dunnett's T3 multiple comparisons test. Bar graphs display the calculated

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.30.526351; this version posted September 4, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

$K_{d(app)s}$ ± the standard error of the mean (SEM). Adjusted P-values are indicated as follows: ns= p >0.05, *= p ≤0.05, and **= p ≤0.01.

FIGURE 8 The N-terminal domain of PlzA is the nucleic acid binding domain. (**a**) The cloning strategy used to produce recombinant N ($PlzA_{NTD}$) and C ($PlzA_{CTD}$) terminal PlzA domains. The *plzA* gene portions corresponding to the N-terminus, *plzA*-NTD, and the C-terminus, *plzA*-CTD, were generated and cloned into an expression vector to produce recombinant $PlzA_{NTD}$ and $PlzA_{CTD}$ proteins. Gradient EMSAs were performed with increasing protein concentrations of $PlzA_{NTD}$ (**b, d**) or $PlzA_{CTD}$ (**c, e**) with (lanes 1-6) or without (lanes 7-12) 100 μM c-di-GMP and 10 nM fluorescently labeled *glpFKD*(-7/35+) dsDNA (**b, c**) or *glpFKD*(UTR) RNA (**d, e**) probe. Lanes 1 and 7 are probe-only controls. Protein concentrations were as follows: Lanes 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 6 and 12- 7.5 μM. SUPERase•In RNase inhibitor was supplemented to a final concentration of 2 U/μL in RNA EMSA mixtures. The asterisk indicates $PlzA_{NTD}$-dsDNA shifts. Representative gels are shown from triplicate ($PlzA_{NTD}$) or duplicate ($PlzA_{CTD}$) EMSAs. N= 2 or 3 independent experiments.

FIGURE 9 Detection of c-di-GMP in recombinant $PlzA_{NTD}$ and $PlzA_{CTD}$ by LC-MS/MS and circular dichroism of proteins. (**a**) Recombinant $PlzA_{NTD}$ and $PlzA_{CTD}$ were sent off for mass spectrometry to detect the presence of c-di-GMP. Extracted ion chromatograms of the LC-MS/MS analysis are displayed. Chromatograms depict the control c-di-GMP standard at 1 μM, recombinant $PlzA_{CTD}$, and recombinant $PlzA_{NTD}$. The LC-MS/MS analysis revealed that $PlzA_{CTD}$ co-purified with c-di-GMP, but $PlzA_{NTD}$ did not. The concentration of c-di-GMP in recombinant $PlzA_{CTD}$ was determined from a standard curve of c-di-GMP. $PlzA_{CTD}$ co-purified with 704.29 nM of c-di-GMP and approximately 1.1 % of the protein was bound with c-di-GMP. The c-di-GMP standard chromatogram shown is the same as in Figure 4. (**b**) Circular dichroism spectra of recombinant $PlzA_{WT}$, $PlzA_{RD\text{-}RD}$, $PlzA_{NTD}$, and $PlzA_{CTD}$ measured in EMSA buffer (see methods). EMSA buffer alone or plus 50 μM c-di-GMP were also measured as controls.

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.30.526351; this version posted September 4, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 10 Computational docking analysis of PlzA with *glp* DNA and RNA. Computational docking was performed using the HDOCK server with PlzA and either (**a**) the 42 bp *glpFKD*(-7/+35) dsDNA or (**b**) the *glpFKD*(UTR) RNA sequence. In (**a**) only 26 of 42 bp of the *glpFKD*(-7/+35) sequence is shown. The dashed lines represent where bases were left out of the figure to zoom in on the protein-DNA interactions. The dsDNA model was linear, and removal of the bases has no impact on the interpretation of the data. The top predicted model from each computation is shown. The HDOCK docking and confidence scores for each model were -210.60 and 0.77 for PlzA-*glp* DNA, and -280.71 and 0.93 for PlzA-*glp* RNA, respectively. The protein is rotated 180° from the top to the bottom panels. Select amino acid residues that interface the nucleic acids are labeled.

FIGURE 11 Model depicting the modulation of PlzA function by levels of c-di-GMP. Apo and holo-PlzA are hypothesized to have distinct functions in the vertebrate and tick hosts respectively. Recent studies identified PlzA RNA chaperone activity that is c-di-GMP independent (Van Gundy *et al.*, 2023). In environments with no c-di-GMP, such as the vertebrate host, apo-PlzA engages in RNA chaperone activity to modulate gene expression required for survival in the vertebrate host. The phosphodiesterases PdeA and PdeB are likely to be highly active upon transmission and into vertebrate infection, lowering c-di-GMP levels and facilitating apo-PlzA functions (Sultan *et al.*, 2010; Pitzer *et al.*, 2011; Sultan *et al.*, 2011; Novak *et al.*, 2014; Groshong *et al.*, 2021). Phosphodiesterase activity is likely to occur at all stages to regulate c-di-GMP levels and indirectly, PlzA function. When Hk1 is activated (signal unknown), Rrp1 is phosphorylated and produces c-di-GMP. Holo-PlzA can also engage in RNA chaperone activity by annealing RNAs, and potentially interact with RNAP (Grassmann *et al.*, 2023; Van Gundy *et al.*, 2023). We propose that levels of c-di-GMP impact the specific genes regulated by PlzA through altering nucleic acid binding affinity. At lower c-di-GMP levels, few holo-PlzA monomers exist resulting in weak PlzA binding affinity for certain target genes such as *glpFKD*. As levels of c-di-GMP increase, such as in the tick host or during a blood meal, more holo-PlzA molecules bind target genes with higher affinity, and PlzA could

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.30.526351; this version posted September 4, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

potentially multimerize on DNA. At high protein and c-di-GMP concentrations, we hypothesize release of DNA by holo-PlzA via a yet unknown signal.

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.30.526351; this version posted September 4, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

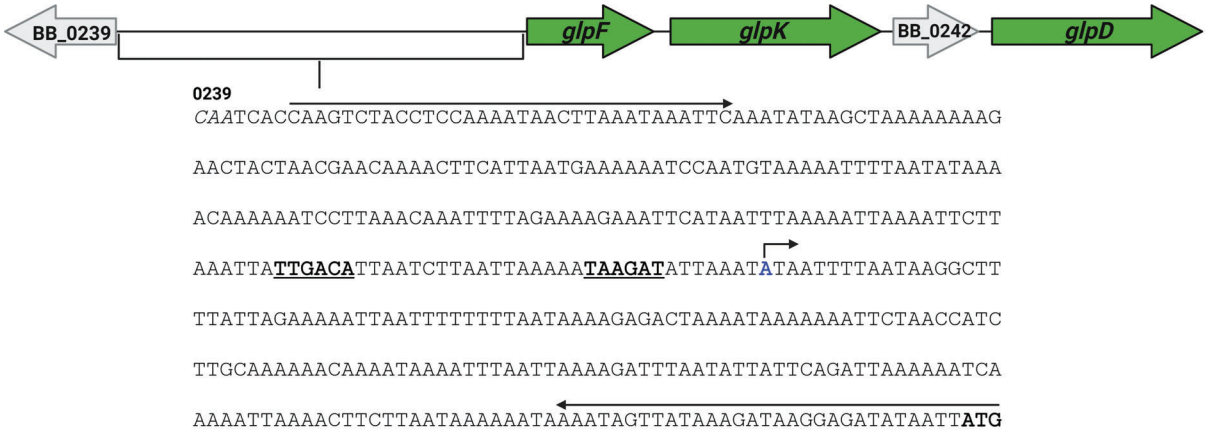

FIGURE 1 The *glpFKD* operon, 5' UTR, and designed probe schematic. The glycerol catabolism gene operon (*glpFKD*) contains four genes: *glpF* (BB_0240, glycerol uptake facilitator protein), *glpK* (BB_0241, glycerol kinase), BB_0242 (hypothetical protein), and *glpD* (BB_0243, glycerol-3-phosphate dehydrogenase).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

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.30.526351; this version posted September 4, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

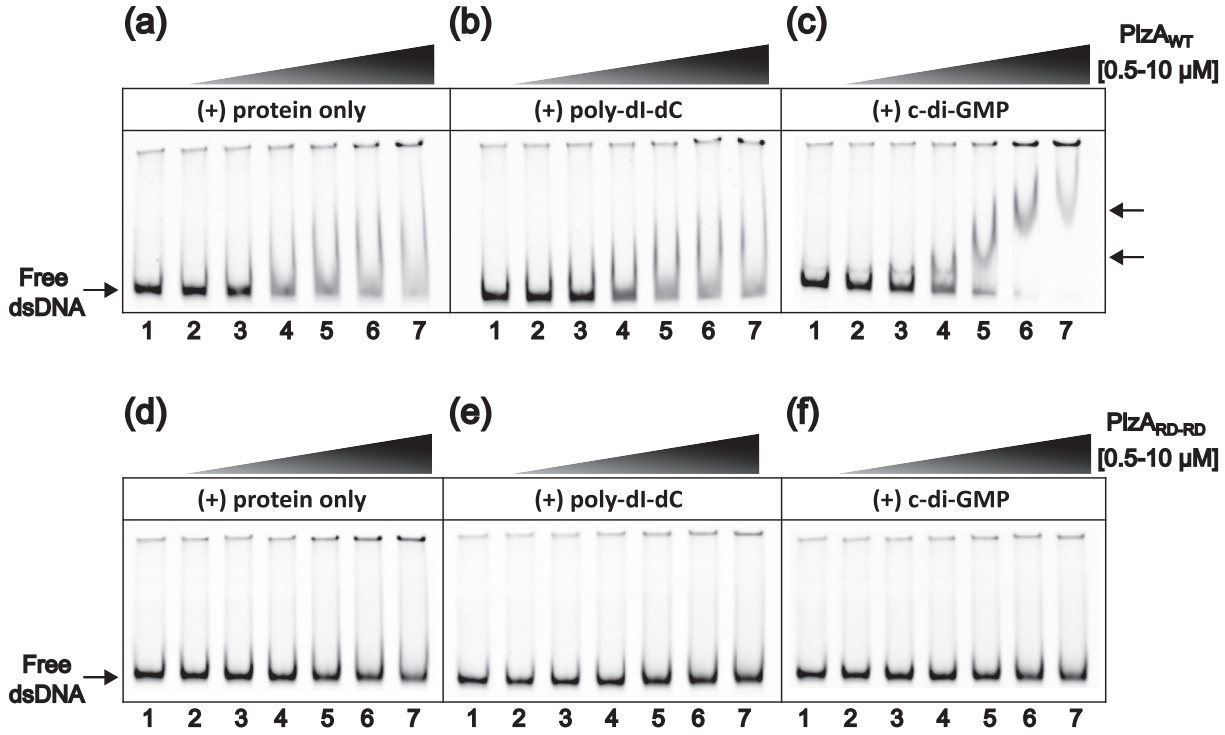

FIGURE 2 DNA binding by PlzA is c-di-GMP dependent. (**a**) EMSA with increasing concentrations of PlzA$_{WT}$ showing binding to *glpFKD*(-219) DNA. The mixture was incubated for 10 min before electrophoresis. (**b**) The same EMSA mixture as in (a) but with a final concentration of 2.5 ng/μL of the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

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.30.526351; this version posted September 4, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

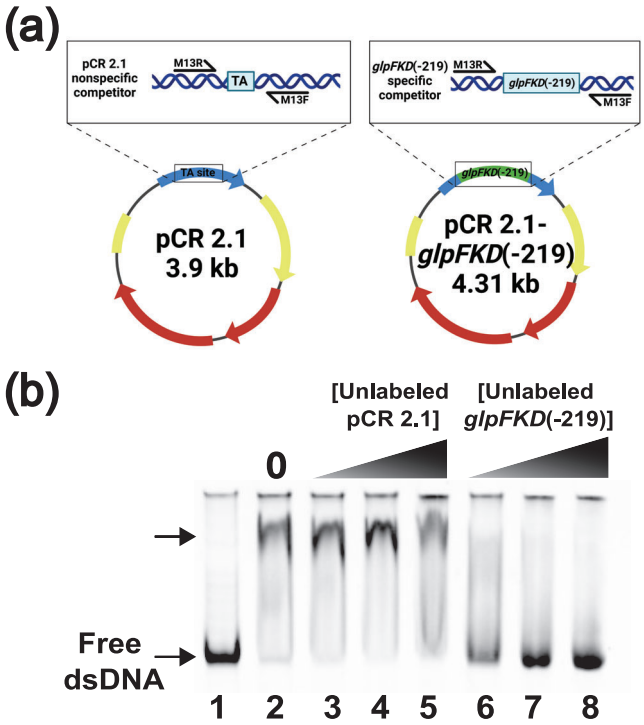

FIGURE 3 PlzA recognizes *glpFKD*(-219) and not pCR 2.1 sequences. (**a**) A schematic of the unlabeled DNA competitors derived from the empty pCR 2.1 vector backbone or the vector containing the *glpFKD*(-219) sequence. A labeled M13R and an unlabeled M13F primer were used to amplify the target DNAs 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.

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.30.526351; this version posted September 4, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

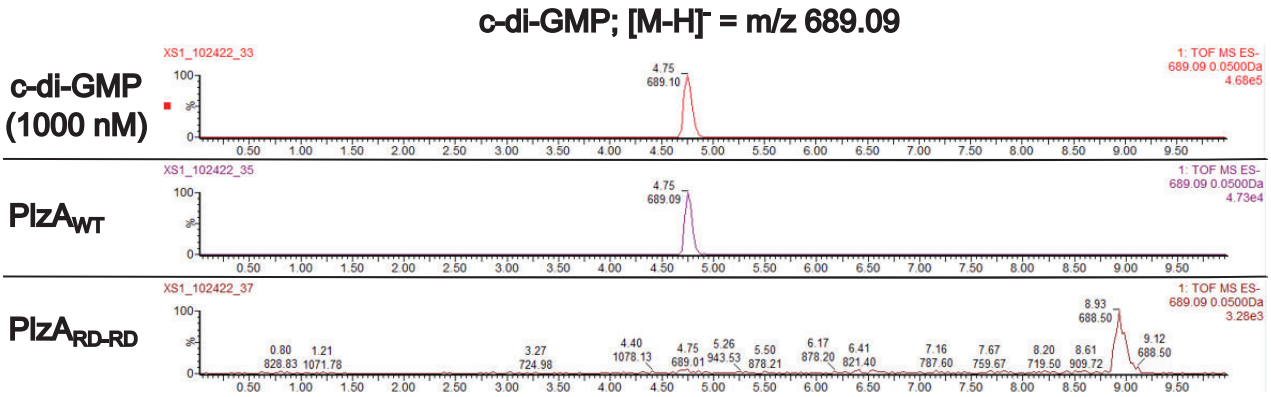

FIGURE 4 Detection of c-di-GMP in recombinant PlzA$_{WT}$ and PlzA$_{RD-RD}$ by LC-MS/MS. Recombinant PlzA$_{WT}$ and PlzA$_{RD-RD}$ were sent off for mass spectrometry to detect c-di-GMP. Extracted ion chromatograms of the LC-MS/MS analysis for detection of c-di-GMP are displayed. Chromatograms depict the control c-di-GMP standard at 1 μM, recombinant PlzA$_{WT}$, and recombinant PlzA$_{RD-RD}$. The LC-MS/MS analysis reveals PlzA$_{WT}$ co-purified with c-di-GMP, but PlzA$_{RD-RD}$ did not. The concentration of c-di-GMP in recombinant PlzA$_{WT}$ was determined from a standard curve of c-di-GMP. PlzA$_{WT}$ co-purified with 99.29 nM of c-di-GMP and approximately 0.3% of the protein was bound with c-di-GMP. Black lines indicate chromatograms unrelated to this figure that were removed.

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.30.526351; this version posted September 4, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

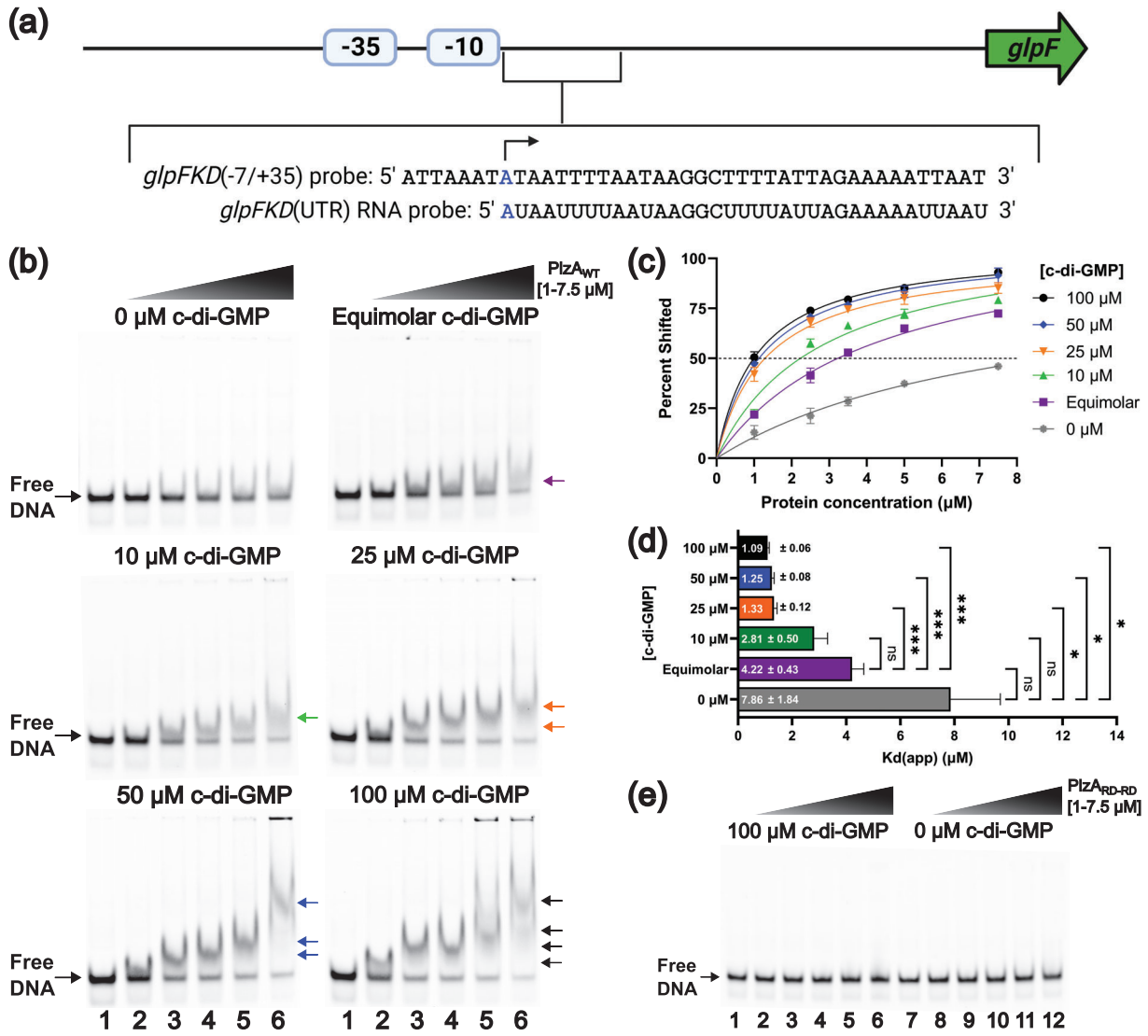

FIGURE 5 Levels of c-di-GMP alter DNA binding affinity of $PlzA_{WT}$. (a) A diagram of the generated DNA probe, designated as *glpFKD(*-7/+35), corresponding to a 42 bp sequence adjacent to the -10 site and proceeding 35 bp into the 5' UTR. An RNA probe was made corresponding to only nucleotides that are within the 5 'UTR which was designated as *glpFKD*(UTR) RNA. Probes were conjugated to a fluorescent molecule for the detection of binding in EMSAs. (b) EMSAs were performed with increasing protein concentrations of $PlzA_{WT}$, and 10 nM fluorescently labeled probe at varying c-di-GMP concentrations (0, equimolar, 10, 25, 50, or 100 μM). For each EMSA shown, the indicated c-di-GMP concentration was held constant in all lanes. For each EMSA shown, lanes 1 and 7 are probe-only controls. Protein concentrations were as follows: Lane 2- 1 μM, Lane 3- 2.5 μM, Lane 4- 3.5 μM, Lane 5- 5 μM, Lane 6- 7.5 μM. A final concentration of 2.5 ng/μL of the nonspecific competitor poly-dI-dC was added to each lane. Arrows indicate higher-order complexes formed as protein concentration increases. A representative gel for each tested c-di-GMP concentration is shown from triplicated EMSAs. (c) Triplicate EMSAs were quantitated by densitometry to determine the $K_{d(app)}$ of the $PlzA_{WT}$-*glpFKD*(-7/35+) interaction by nonlinear regression analysis using the one-site specific binding setting in GraphPad Prism. Errors bars represent one standard deviation (SD). (d) Bar graph displaying the calculated $K_{d(app)}$ ± the standard error of the mean (SEM). The Kd(app) values from three EMSAs were analyzed by a Welch ANOVA with Dunnett's T3 multiple comparisons test via GraphPad Prism. Adjusted P-values are indicated as follows: ns= p >0.05, *= p ≤0.05, and ***= p ≤0.001. Comparisons between 10, 25, 50, and 100 μM c-di-GMP are not indicated on the graph as they were not significant. (e) A similar EMSA experiment was performed as in Figure 5B but with $PlzA_{RD-RD}$ and either 100 (lanes 1-6) or 0 (lanes 7-12) μM of c-di-GMP supplemented. No binding is observed.

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.30.526351; this version posted September 4, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

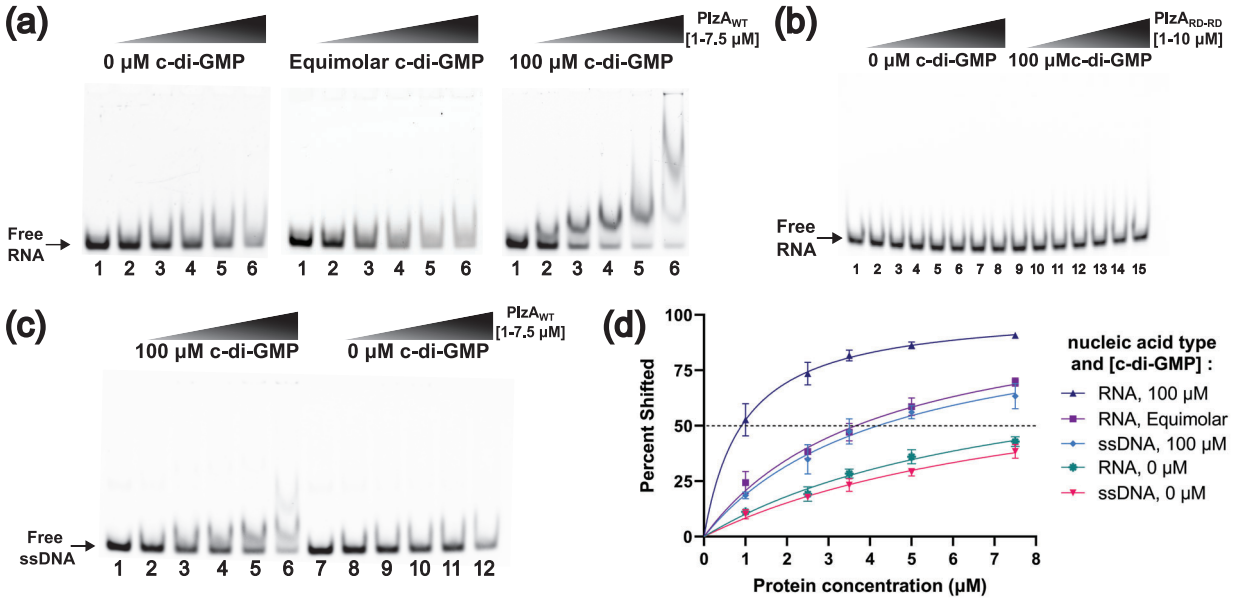

FIGURE 6 PlzA$_{WT}$ also binds single stranded RNA and DNA. (**a**) Gradient EMSAs were performed with increasing protein concentrations of PlzA$_{WT}$ without c-di-GMP (left panel), equimolar (middle panel), or 100 μM c-di-GMP (right panel) and 10 nM fluorescently labeled *glpFKD*(UTR) RNA probe. Lane 1 is the probe-only control in each EMSA. Lane protein concentrations were as follows: Lane 2- 1 μM, Lane 3- 2.5 μM, Lane 4- 3.5 μM, Lane 5- 5 μM, Lane 6- 7.5 μM. (**b**) A similar EMSA experiment was performed with PlzA$_{RD-RD}$ with (lanes 1-8) or without (lanes 9-15) 100 μM c-di-GMP and 50 nM of *glpFKD*(UTR) RNA probe. Lane 1 is a probe-only control. Protein concentrations were as follows: Lanes 2 and 9- 1 μM, Lanes 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, Lanes 7 and 14- 5 μM, Lanes 8 and 15- 10 μM. Riboguard was supplemented in the PlzA$_{RD-RD}$ EMSA mixtures to a final concentration of 4 U/μL. (**c**) Gradient EMSAs were performed with increasing protein concentrations of PlzA$_{WT}$ with 100 μM c-di-GMP (left panel) or 0 μM c-di-GMP (right panel) with 10 nM fluorescently labeled ss*glpF* probe. Lanes 1 and 7 are the probe-only control in each EMSA. Lane protein concentrations were as follows: Lanes 2 and 8- 1 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- 7.5 μM. Representative gels are shown from triplicated (PlzA$_{WT}$) or duplicated (PlzA$_{RD-RD)}$ EMSA experiments. (**d**) Triplicated EMSAs from (**a**) and (**c**) were quantitated by densitometry to determine the Kd(app) of PlzA$_{WT}$ for the *glpFKD* RNA or ssDNA interactions by nonlinear regression analysis using the one-site specific binding setting in GraphPad Prism. Errors bars represent one standard deviation (SD).

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.30.526351; this version posted September 4, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

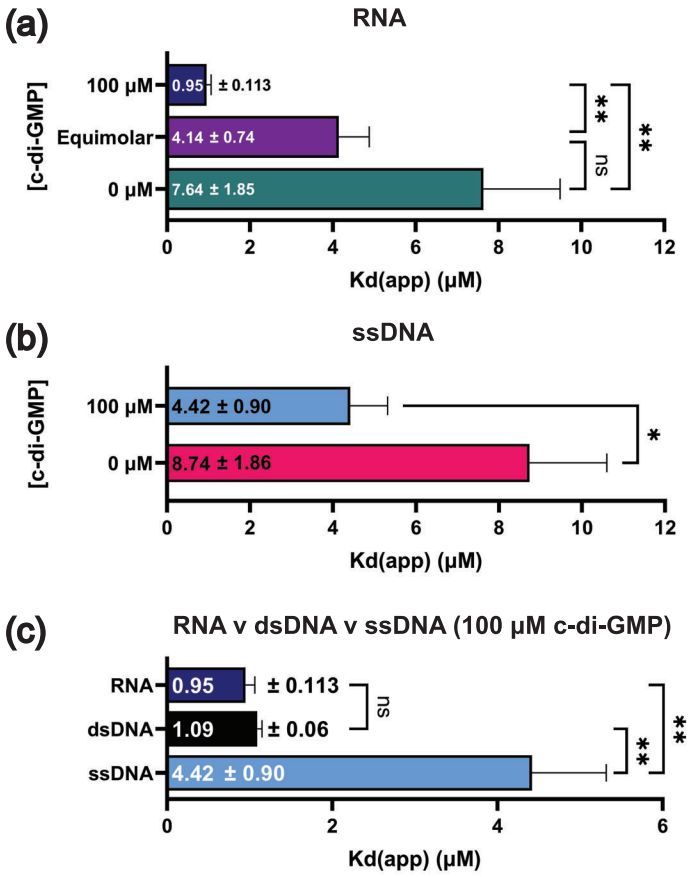

FIGURE 7 c-di-GMP significantly impacts PlzA-nucleic acid binding and PlzA prefers dsDNA and RNA over ssDNA. (**a**) The Kd(app) values from triplicate RNA EMSAs at each tested c-di-GMP concentration in Figure 6A were analyzed by a Welch ANOVA with Dunnett's T3 multiple comparisons test. (**b**) The Kd(app) values from triplicate ssDNA EMSAs at the tested c-di-GMP concentrations in Figure 6C were analyzed by a Welch's t-test. (c) The Kd(app) values from EMSAs with PlzA and each tested nucleic acid at 100 μM c-di-GMP were compared via a Welch ANOVA with Dunnett's T3 multiple comparisons test. Bar graphs display the calculated Kd(app)s ± the standard error of the mean (SEM). Adjusted P-values are indicated as follows: ns= p >0.05, *= p ≤0.05, and **= p ≤0.01.

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.30.526351; this version posted September 4, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

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

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.30.526351; this version posted September 4, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

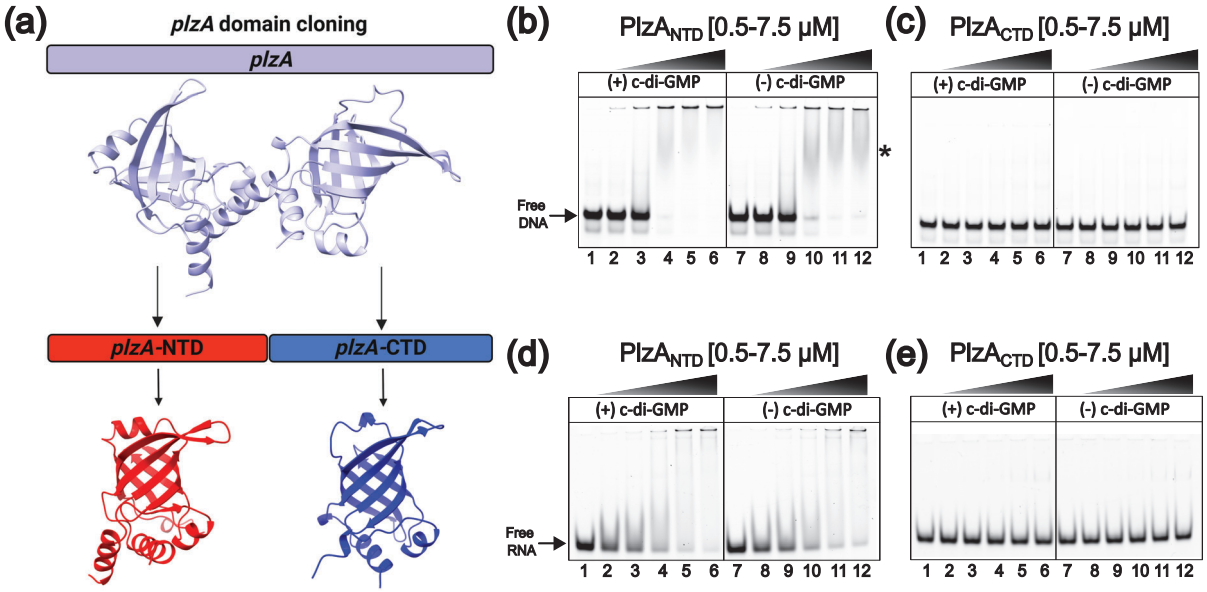

FIGURE 8 The N-terminal domain of PlzA is the nucleic acid binding domain. (**a**) The cloning strategy used to produce recombinant N (PlzA$_{NTD}$) and C (PlzA$_{CTD}$) terminal PlzA domains. The plzA gene portions corresponding to the N-terminus, plzA-NTD, and the C-terminus, plzA-CTD, were generated and cloned into an expression vector to produce recombinant PlzA$_{NTD}$ and PlzA$_{CTD}$ proteins. Gradient EMSAs were performed with increasing protein concentrations of PlzA$_{NTD}$ (**b**, **d**) or PlzA$_{CTD}$ (**c**, **e**) with (lanes 1-6) or without (lanes 7-12) 100 μM c-di-GMP and 10 nM fluorescently labeled *glpFKD*(-7/35+) dsDNA (b, c) or *glpFKD*(UTR) RNA (**d**, **e**) probe. Lanes 1 and 7 are probe-only controls. Protein concentrations were as follows: Lanes 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 6 and 12- 7.5 μM. SUPERase•In RNase inhibitor was supplemented to a final concentration of 2 U/μL in RNA EMSA mixtures. Representative gels are shown from triplicated (PlzA$_{NTD}$) or duplicated (PlzA$_{CTD}$) EMSAs.

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.30.526351; this version posted September 4, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

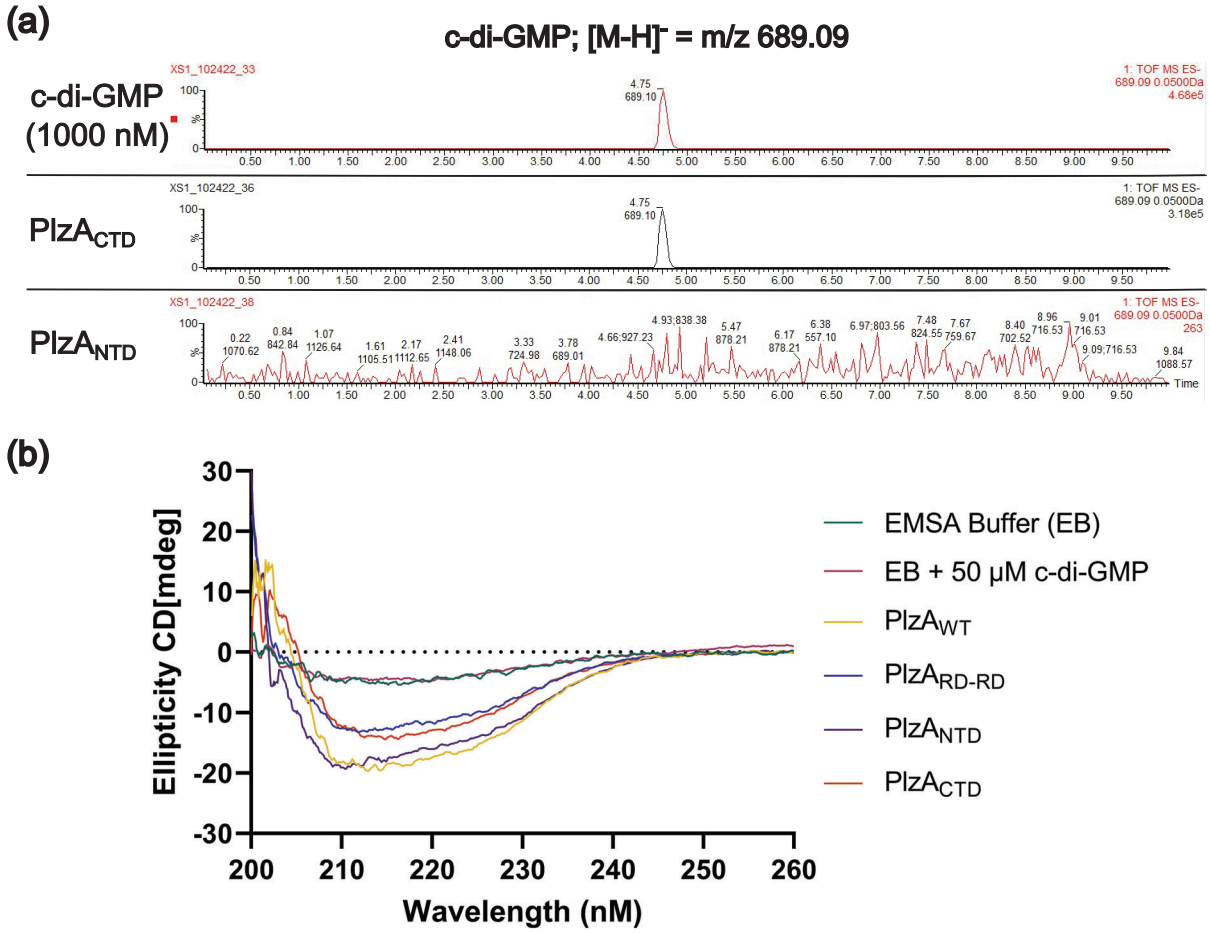

FIGURE 9 Detection of c-di-GMP in recombinant $PlzA_{NTD}$ and $PlzA_{CTD}$ by LC-MS/MS and circular dichroism of proteins. (**a**) Recombinant $PlzA_{NTD}$ and $PlzA_{CTD}$ were sent off for mass spectrometry to detect the presence of c-di-GMP. Extracted ion chromatograms of the LC-MS/MS analysis are displayed. Chromatograms depict the control c-di-GMP standard at 1 μM, recombinant $PlzA_{CTD}$, and recombinant $PlzA_{NTD}$. The LC-MS/MS analysis revealed that $PlzA_{CTD}$ co-purified with c-di-GMP, but $PlzA_{NTD}$ did not. The concentration of c-di-GMP in recombinant $PlzA_{CTD}$ was determined from a standard curve of c-di-GMP. $PlzA_{CTD}$ co-purified with 704.29 nM of c-di-GMP and approximately 1.1 % of the protein was bound with c-di-GMP. The c-di-GMP standard chromatogram shown is the same as in Figure 4. (**b**) Circular dichroism spectra of recombinant $PlzA_{WT}$, $PlzA_{RD-RD}$, $PlzA_{NTD}$, and $PlzA_{CTD}$ measured in EMSA buffer (see methods). EMSA buffer alone or plus 50 μM c-di-GMP were also measured as controls.

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.30.526351; this version posted September 4, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

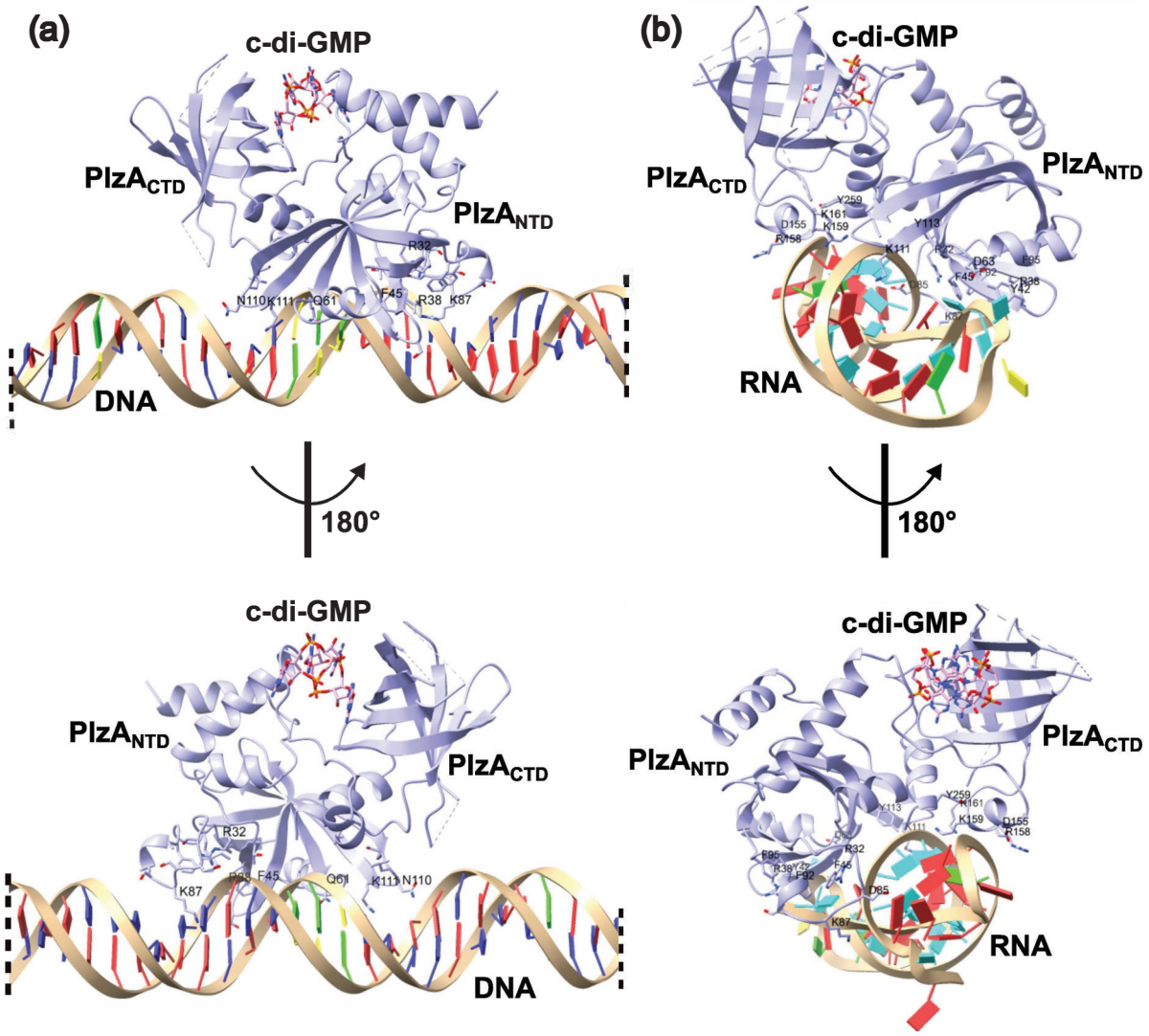

FIGURE 10 Computational docking analysis of PlzA with *glp* DNA and RNA. Computational docking was performed using the HDOCK server with PlzA and either (a) the 42 bp *glpFKD*(-7/+35) dsDNA or (b) the *glpFKD*(UTR) RNA sequence. In (a) only 26 of 42 bp of the *glpFKD*(-7/+35) sequence is shown. The dashed lines represent where bases were left out of the figure to zoom in on the protein-DNA interactions. The dsDNA model was linear, and removal of the bases has no impact on the interpretation of the data. The top predicted model from each computation is shown. The HDOCK docking and confidence scores for each model were -210.60 and 0.77 for PlzA-*glp* DNA and -280.71 and 0.93 for PlzA-*glp* RNA, respectively. The protein is rotated 180° from the top to the bottom panels. Select amino acid residues that interface the nucleic acids are labeled.

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.30.526351; this version posted September 4, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

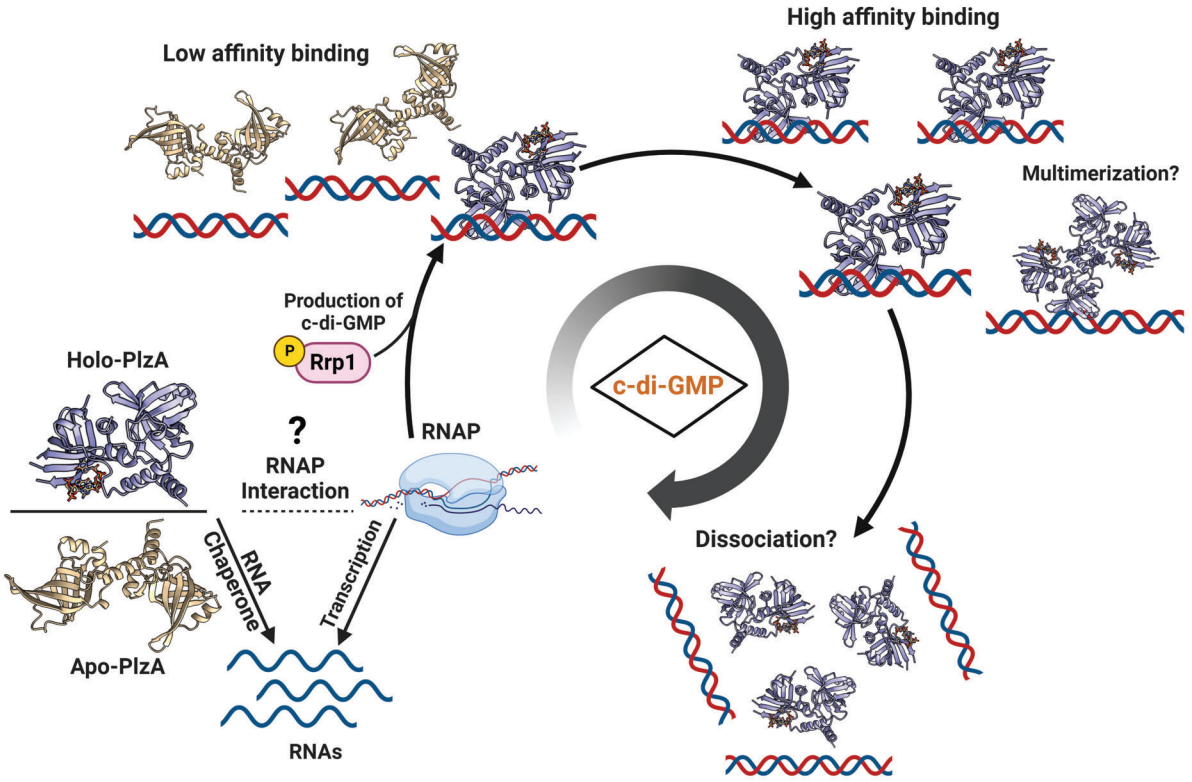

FIGURE 11 Model depicting the modulation of PlzA function by levels of c-di-GMP. Apo and holo-PlzA are hypothesized to have distinct functions in the vertebrate and tick hosts respectively. Recent studies identified PlzA RNA chaperone activity that is c-di-GMP independent (Van Gundy et al., 2023). In environments with no c-di-GMP, such as the vertebrate host, apo-PlzA engages in RNA chaperone activity to modulate gene expression required for survival in the vertebrate host. The phosphodiesterases PdeA and PdeB are likely to be highly active upon transmission and into vertebrate infection, lowering c-di-GMP levels and facilitating apo-PlzA functions (Sultan et al., 2010; Pitzer et al., 2011; Sultan et al., 2011; Novak et al., 2014; Groshong et al., 2021). Phosphodiesterase activity is likely to occur at all stages to regulate c-di-GMP levels and indirectly, PlzA function. When Hk1 is activated (signal unknown), Rrp1 is phosphorylated and produces c-di-GMP. Holo-PlzA can also engage in RNA chaperone activity by annealing RNAs, and potentially interact with RNAP (Grassmann et al., 2023; Van Gundy et al., 2023). We propose that levels of c-di-GMP impact the specific genes regulated by PlzA through altering nucleic acid binding affinity. At lower c-di-GMP levels, few holo-PlzA monomers exist resulting in weak PlzA binding affinity for certain target genes such as *glpFKD*. As levels of c-di-GMP increase, such as in the tick host or during a blood meal, more holo-PlzA molecules bind target genes with higher affinity, and PlzA could potentially multimerize on DNA. At high protein and c-di-GMP concentrations, we hypothesize release of DNA by holo-PlzA via a yet unknown signal.

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.30.526351; this version posted September 4, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

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.