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

Nerina Jusufovic[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

Present Address:
Christina R. Savage, National Cancer Institute, NIH, Bethesda, Maryland, 20892, USA

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 functional mechanism of PlzA was not previously known. We report that PlzA can bind DNA and RNA within the promoter and 5' UTR of the glycerol metabolism operon, *glpFKD*, and that nucleic acid binding requires c-di-GMP. In the presence of c-di-GMP, PlzA formed multimeric complexes with nucleic acids. PlzA contains two PilZ domains. Dissection of the domains demonstrated that the separated N-terminal domain bound nucleic acids in a c-di-GMP-independent manner. The C-terminal domain, which includes the c-di-GMP binding motif, did not bind nucleic acids under any tested conditions. Structural modeling suggests that c-di-GMP binding to the C-terminal domain stabilizes interactions between the two domains, facilitating nucleic acid binding through residues in the N-terminal domain.

INTRODUCTION

Cyclic bis-(3'→5')-dimeric guanosine monophosphate (c-di-GMP), is a ubiquitous bacterial cyclic di-nucleotide. In many species it is involved in the regulation of key processes such as transcription, motility, biofilm formation, and virulence (1–5). Signaling by c-di-GMP occurs in response to environmental changes and the impacts of c-di-GMP are mediated by regulating its synthesis and turnover. Two molecules of GTP are synthesized into c-di-GMP by diguanylate cyclases (DGC), while degradation is mediated by phosphodiesterases (PDE) (6–11). The abundance of DGCs and PDEs varies among bacteria, but at least one of each must be present to mediate c-di-GMP regulation. In some cases, c-di-GMP signaling is initiated by sensory histidine kinases that become stimulated by external signals and undergo autophosphorylation to set off a signaling cascade that leads to c-di-GMP synthesis (3, 4, 7). Responses to c-di-GMP are mediated by c-di-GMP binding receptors, which upon binding to c-di-GMP execute various effector functions (2, 4, 12). The first discovered c-di-GMP binding receptors were the PilZ domain-containing proteins (12–14). The PilZ domain consists of two conserved c-di-GMP binding motifs, RxxxR and [D/N]ZSXXG (12, 14–16). Given the widespread c-di-GMP signaling across Eubacteria, the PilZ domain containing proteins are one of the most well characterized but also diverse groups of c-di-GMP binding receptors (2–5, 13, 15, 16).

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 (7, 17–20). The borrelial Hk1/Rrp1 two-component system synthesizes c-di-GMP; activation of Hk1 occurs through an unknown signal (19, 21). Rrp1, the only known DGC in *B. burgdorferi*, is essential for acquisition and survival in the tick vector (17, 22). *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 (23–26). Synthesized c-di-GMP is bound by PlzA, the only universal c-di-GMP binding protein found in all isolates of Lyme disease *Borreliae* (27). PlzA binds c-di-GMP with high specificity and affinity and has been

implicated in regulating many *B. burgdorferi* processes, including alternative carbohydrate utilization, motility, and virulence (26, 28–31).

The glycerol metabolism operon, *glpFKD*, promotes *B. burgdorferi* survival in ticks (32). 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 utilize alternate carbon sources such as glycerol. Regulation of the *glpFKD* operon is controlled through the Hk1/Rrp1 system, and borrelial *plzA* mutants are defective in glycerol metabolism (17, 25, 31, 33). Despite the importance of c-di-GMP and PlzA for the survival and infectivity of *B. burgdorferi*, the molecular mechanism(s) of PlzA effector function was previously unknown.

A common mechanism of action for c-di-GMP effector proteins is to bind nucleic acids and function as transcription factors and/or nucleoid-associated proteins (34–40). 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 (41). Binding of c-di-GMP to the interdomain linker domain results in a conformational change; locking the PlzA domains into a rigid conformation (26, 29, 41). While no canonical DNA binding domain is evident in the PlzA amino acid sequence, we note that many of the borrelial nucleic acid binding proteins our lab has discovered and characterized possess novel binding motifs (31, 41–46). 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.

MATERIAL AND METHODS

Routine manipulations and molecular cloning

Oligonucleotides used in this study are listed in Supplementary Table S1 and were ordered from Integrated DNA Technologies (IDT). Genomic DNA (gDNA) was purified utilizing the E.Z.N.A Tissue 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 (47). 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. These mutations resulted in the amino-acid substitutions R145D and R149D within the c-di-GMP binding site of PlzA and have previously been shown to abolish c-di-GMP binding (18). 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 Supplementary Table S1.

Recombinant protein overexpression and purification

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 HisPur Ni-NTA resin (Thermo Scientific) per manufacturer protocol. Eluted protein solutions were then 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 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 Supplementary Table S1 (48, 49). DNA sequences used as templates for generating fluorescently tagged fragments >60 bp were either generated directly from PCR amplification of B31 genomic DNA, or from fragments cloned into the pCR 2.1 TA vector (Invitrogen) following TOPO TA protocols. PCR products amplified directly from gDNA were analyzed by agarose gel electrophoresis and gel-purified with Wizard® SV Gel and PCR Clean-Up System (Promega). Subsequently, target-specific or M13 F and R IRDye 800-labeled primers were used for PCR of the gel-purified or 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 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 50 nM or RNA probe at 50 or 100 nM for qualitative EMSAs. For quantitative EMSAs, the final probe concentrations were 10 nM. Some EMSAs were supplemented with either 3'5' c-di-GMP, 3'5' c-di-AMP, or 3'3' cGAMP (Sigma-Aldrich) to final concentrations of 100 µM, unless otherwise stated. Proteins were incubated with a cyclic-di-nucleotide for 5 min before the addition of the probe. Upon the addition of the probe, reactions were incubated for an additional 10 min. When appropriate, the non-specific competitor poly-dI-dC (Roche) was added to EMSA reactions prior to the labeled probe at a final concentration of 2.5 ng/µL and allowed to incubate with the protein and c-di-GMP for 5 min. Competition EMSAs were conducted similarly but with the unlabeled and specific competitor probe added after poly-dI-dC and allowed to incubate with protein and c-di-GMP prior to the addition of the labeled probe. For RNA EMSAs, RNA probes were heated above their T_m for 15 min before addition to the EMSA mix. When RNase contamination was present in recombinant proteins, RiboGuard RNase inhibitor (Lucigen) was added to RNA EMSA mixtures to a final concentration of 4 U/µL. For visualization during electrophoresis, EMSA loading dye (0.8 mg/mL Orange G, 15 mg/mL Ficoll 400) was added to each reaction. Novex TBE 6 or 10% gels (Invitrogen) were pre-run in 0.5x TBE buffer at 100V for a minimum of 30 min. The entire mix was then loaded onto the pre-run gels and resolved for 60-90 min at 100 V. Gels were run 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 for quantitative purposes unless otherwise stated. Lanes and bands were added manually and subsequently analyzed to generate the lane percentage totals for each band in a lane. Shifted bands were considered as the entire region above the free DNA in each lane. For competition EMSAs, the percentage of free DNA for each lane was normalized to the probe only control band percentage value. The band percentage values were used to determine the percent shifted of the respective probe relative to the free probe. Background and any ssDNA probe percentage that was shifted were subtracted from the percent totals. Nonlinear regression analysis was performed to calculate the apparent dissociation constant. Briefly, protein concentrations were plotted against the percent shifted per lane using the one-site specific binding setting in Prism GraphPad 9. Confidence was set at 95% with the analysis considering the mean of the replicate values from the triplicated EMSAs.

Protein structure modeling and bioinformatics

Protein structure prediction was performed via AlphaFold, while modeling of protein structures was carried out using ChimeraX 1.4 (50–52). Clustal Omega was used to align protein sequences, and the Clustal output files were then processed with Multiple Align Show for visualization of alignments (53, 54).

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 $PlzA_{WT}$, $PlzA_{RD\text{-}RD}$, $PlzA_{NTD}$, 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.

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 chromatograms for each sample were plotted as the relative abundance vs retention time in GraphPad Prism 9. 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.

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 (17, 25, 31, 33). Activation of *glpFKD* is critical for spirochete survival in between blood meals of the tick vector, which requires catabolism of glycerol (31–33). Given that requirement, we hypothesized that PlzA might bind the promoter of the *glpFKD* operon. To that end, we sought to determine if PlzA could bind DNA between *glpFKD* and the upstream locus. The transcriptional boundaries and promoter elements of the *glpFKD* operon were previously mapped (Figure 1A) (55, 56). A fluorescently labeled DNA fragment, encompassing the 410 bp intergenic region between the upstream gene, ORF BB_0239 (encoding a deoxyguanosine/deoxyadenosine kinase), and the start of *glpF* (BB_0240) was generated and is denoted as *glpFKD*(-219) (Figure 1A). EMSAs with purified recombinant PlzA revealed that PlzA bound to *glpFKD*(-219) DNA (Figure 2). Notably, two bands are visible in the probe-only control. When amplified with target-specific primers (Supplementary Table S1) from either genomic DNA or the TA cloned construct, the glpFKD(-219) PCR amplicon produces two bands despite gel purification of the appropriately sized band (Figure 2 lane 1). This suggests that native *glpFKD*(-219) forms secondary DNA structures.

To assess whether this interaction was specific, EMSAs with concentrations of PlzA within the range producing detectable shifts from the previous experiment were performed with the inclusion of the non-specific competitor poly-dI-dC (Figure 3A and B). The PlzA-*glpFKD*(-219) interaction was confirmed to be specific as poly-dI-dC did not detectably compete away shifted protein-DNA complexes. Note that secondary DNA structure is not observed in the probe-only lanes when *glpFKD*(-219) DNA is amplified from the pCR 2.1 vector using M13 F and R primers (Figure 3).

We next assessed the effects of c-di-GMP on DNA binding. Initially, c-di-GMP was supplemented in EMSA reactions to a final concentration of 100 μM. This concentration was determined empirically and was used to saturate all PlzA molecules present in the reaction to ensure complete binding throughout electrophoresis. Similarly, saturating concentrations have been used in other studies of c-di-GMP binding proteins that also bind DNA (35–38, 57, 58). The EMSA with *glpFKD*(-219), poly-dI-dC, and

increasing PlzA was repeated with supplemented c-di-GMP (Figure 3C). The addition of c-di-GMP enhanced PlzA binding affinity for *glpFKD*(-219) resulting in binding at lower concentrations of protein.

To further assess the requirement of c-di-GMP for PlzA DNA binding activity, a double mutant PlzA protein ($PlzA_{RD-RD}$) that cannot bind c-di-GMP was constructed by site-directed mutagenesis replacing the key arginine residues (R145 and R149) in the c-di-GMP binding domain with aspartic acids. The *glpFKD*(-219) EMSAs were repeated with $PlzA_{RD-RD}$ (Figure 3D-F). Binding to *glpFKD*(-219) was not observed under any tested conditions with $PlzA_{RD-RD}$.

As controls, competition EMSAs were performed with unlabeled DNA competitors derived from empty pCR 2.1 TA clones or those containing the *glpFKD*(-219) sequence (Supplementary Figure S1). The nonspecific DNA from the empty pCR 2.1 vector did not compete away $PlzA_{WT}$-*glpFKD*(-219) complexes (Supplementary Figure S1, lanes 3-5). Meanwhile, unlabeled *glpFKD*(-219) successfully competed away $PlzA_{WT}$-*glpFKD*(-219) complexes (Supplementary Figure S1, lanes 6-8). These results indicate that PlzA specifically bound the *glpFKD*-derived sequences in EMSAs.

Noting that c-di-GMP enhanced DNA binding by PlzA yet purified recombinant PlzA bound DNA without the addition of exogenous c-di-GMP, we investigated whether PlzA was purified with c-di-GMP already bound. Purified proteins were prepared as described above and assayed by liquid chromatography-tandem mass spectrometry to determine whether any cyclic di-nucleotides were present. 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 as 99.29 nM. A total of 31.7 µM of protein was sent for analysis indicating approximately 0.3% of the protein was bound with c-di-GMP. Prior studies determined that PlzA binds c-di-GMP with high affinity, with reported dissociation constants of 1.25 to 6.25 µM, so the finding of c-di-GMP in PlzA protein preparations was not unexpected (28, 29). Taken together, these data indicate that PlzA DNA binding to this region of the *glpFKD* operon is c-di-GMP dependent.

DNA binding affinity and specificity of PlzA

It was previously identified through *glpFKD-gfp* promoter fusions 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 (56). Given these findings and our data that PlzA can bind DNA, we focused further analysis of PlzA binding on a particular sequence within that region that was found to be important for transcription (56, 59). 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 1B). The *glpFKD*(-7/+35) probe was used in EMSAs with $PlzA_{WT}$ or $PlzA_{RD-RD}$, with or without c-di-GMP. $PlzA_{WT}$ bound to *glpFKD*(-7/+35), which was enhanced by increasing the concentration of c-di-GMP (Supplementary Figure S2A). The mutant $PlzA_{RD-RD}$ did not bind this smaller DNA under any tested conditions (Supplementary Figure S2B).

Next, we quantified the binding affinity of PlzA for *glpFKD*(-7/+35). To calculate the apparent dissociation constant $K_{d(app)}$ of the $PlzA_{WT}$- *glpFKD*(-7/+35*)* interaction, a gradient EMSA was performed

with increasing $PlzA_{WT}$ concentrations in the presence or absence of 100 μM c-di-GMP (Figure 5A and B respectively). Competitor poly-dI-dC was added to inhibit nonspecific binding interactions. The $K_{d(app)}$ was determined by non-linear regression analysis as described in the methods. With c-di-GMP, the $K_{d(app)}$ of $PlzA_{WT}$- *glpFKD*(-7/+35) was 2.45 μM (± 0.57 μM) (Figure 5C). Further, as protein concentration increased, a supershift was formed, suggesting a complex of multiple proteins and/or DNA molecules (Figure 5A). Thus, the stoichiometry of this complex(es) is being investigated, which will facilitate calculation of a true dissociation constant. The $K_{d(app)}$ without c-di-GMP was 738.9 μM (± 14232 μM), which is physiologically irrelevant (Figure 5C). The protein concentration range of 0.5-10 μM used in EMSAs did not cover the concentration needed to reach 50% of the free DNA shifted when c-di-GMP is not added (Figure 5B).

To investigate the specificity of PlzA for *glpFKD*(-7/+35) DNA, competition EMSAs were performed with increasing molar excess of either unlabeled *glpFKD*(-7/+35) or unlabeled *flaB* DNAs as competitors against the labeled *glpFKD*(-7/+35) probe (Figure 6). The *flaB* competitor was chosen as it is a borrelial-derived sequence corresponding to a constitutively expressed gene that is a commonly used control in the field. Poly-dI-dC and c-di-GMP were added to each reaction at final concentrations of 2.5ng/μL and 100 μM respectively, while the PlzA protein concentration was held constant at 2.5 μM. Competition with 100-fold excess of unlabeled *glpFKD*(-7/+35) resulted in the loss of the highest order PlzA-DNA complexes but not the smaller complexes (Figure 6A Lane 3 complex B and A respectively). A 50-fold excess of unlabeled *glpFKD*(-7/+35) was sufficient to completely compete away this higher-order shift (Supplementary Figure S3 Lane 6). A 250-fold excess of unlabeled *glpFKD*(-7/+35) was required to detectably compete against complex A, but the overall decrease in percent shifted was like that of a 100-fold excess competitor and control based on densitometric analyses (Figure 6B). Ultimately, a 1000-fold excess of unlabeled *glpFKD*(-7/+35) was required for an approximately two-fold reduction in the $PlzA_{WT}$-*glpFKD*(-7/+35) complexes (Figure 6A Lane 6 and 6B). These results align with the higher calculated $K_{d(app)}$ for the PlzA-*glpFKD*(-7/+35) interaction as competition was not observed until the K_d of the interaction was surpassed (60).

Conversely, no competition was observed of complex A at 100 to 500-fold excess of unlabeled *flaB* competitor but was observed for complex B (Figure 6A lanes 8-10 and 6B). Some competition of complex A is observed at 1000-fold excess, but a 2000-fold excess of unlabeled *flaB* competitor was required to observe a two-fold reduction in the $PlzA_{WT}$-*glpFKD*(-7/+35) complexes (Figure 6A lanes 11 and 12 and 6B). This indicates that the affinity of $PlzA_{WT}$ for the *flaB* DNA is approximately two-fold less than for *glpFKD*(-7/+35) (Figure 6B).

Of note, as the concentration of unlabeled *glpFKD*(-7/+35) was increased, the single-stranded DNA species of the labeled probe disappeared (Supplementary Figure S3 lanes 3-9 and Figure 6 lanes 3-7). The single-stranded DNA species did not disappear when excess *flaB* competitor was added (Figure 6 lanes 8-12). This suggests that annealing of the labeled single-stranded DNA with the unlabeled complement strand could have occurred and increased the amount of labeled dsDNA probe in these lanes as compared to the protein and probe-only lanes. Thus, more labeled dsDNA was available, and the amount of unlabeled competitor required to decrease the observed shifts would be higher. Therefore, competition might occur at lower fold excesses of unlabeled *glpFKD*(-7/+35) than detected here.

PlzA also binds *glpFKD* RNA

Several borrelial proteins that were initially identified as having affinities for DNA have subsequently been found to also bind RNA (45, 59, 61). 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 1B). The *glpFKD*(-7/+35) counterpart RNA probe is designated *glpFKD*(UTR). EMSAs were performed with increasing concentrations of either $PlzA_{WT}$ or $PlzA_{RD\text{-}RD}$ with or without 100 µM c-di-GMP and labeled *glpFKD*(UTR) (Figure 7). $PlzA_{WT}$ bound RNA , which was enhanced by addition of c-di-GMP (Figure 7A). The binding pattern to RNA was comparable to DNA with increased protein concentration leading to larger complexes formed, evident as supershifts in the representative EMSAs. Those EMSA experiments were repeated with $PlzA_{RD\text{-}RD}$ to confirm if RNA binding is c-di-GMP dependent (Figure 7B). The mutant $PlzA_{RD\text{-}RD}$ protein did not bind *glpFKD*(UTR) RNA at any tested protein concentration, regardless of c-di-GMP addition. Collectively, our data indicate that PlzA binding activity is c-di-GMP dependent for both *glpFKD* DNA and RNA.

Dissection of PlzA reveals the nucleic acid binding domain

The solved structure of PlzA in complex with c-di-GMP has revealed a unique architecture consisting of a dual-domain protein connected through a linker domain. 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).

To determine which PlzA domain(s) binds nucleic acids, we performed EMSAs with increasing concentrations of $PlzA_{NTD}$ or $PlzA_{CTD}$ with or without 100 µM c-di-GMP and labeled *glpFKD*(-7/+35) DNA probe. $PlzA_{NTD}$ bound DNA independently of c-di-GMP, as supplementation with c-di-GMP did not affect binding. Higher order complexes were formed by the $PlzA_{NTD}$-*glpFKD*(-7/+35) interaction but not intermediary ones, in contrast to both types formed by $PlzA_{WT}$ (Figure 8B). This suggests that the C-terminus plays a role in protein-to-protein interactions required for the apparent multimerization of PlzA molecules during DNA binding.

$PlzA_{CTD}$ contains the motifs required for c-di-GMP binding, but the $PlzA_{CTD}$ domain did not bind to DNA with added c-di-GMP (Figure 8C). 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 9). Co-purified c-di-GMP was detected in $PlzA_{CTD}$ but not $PlzA_{NTD}$ by LC/MS-MS (Figure 9C and D respectively). 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. Therefore, the lack of DNA binding by $PlzA_{CTD}$ is not simply due to improper folding or loss of protein function. It was also determined that $PlzA_{NTD}$ could bind *glpFKD*(UTR) RNA with or without c-di-GMP, while $PlzA_{CTD}$ does not (Supplementary Figure S4A and B respectively).

PlzA DNA binding affinity is also enhanced by the cyclic di-nucleotide cGAMP

It has been shown previously that PlzA does not bind to other nucleotides like cGMP and GTP but binding to the di-nucleotides cyclic di-3',5'-adenosine monophosphate (c-di-AMP) and 3',3'-cyclic guanosine monophosphate-adenosine monophosphate (cGAMP) has not previously been explored (18, 28). Mass spectrometry analysis was performed to determine whether any of our recombinant proteins of interest contained c-di-AMP or cGAMP. C-di-AMP was not detected in any tested recombinant proteins (Figure 10A). An EMSA performed with *glpFKD*(-7/+35) and $PlzA_{WT}$ with increasing concentrations of c-di-AMP did not result in increased DNA binding affinity as observed with c-di-GMP (Figure 10B).

The cGAMP di-nucleotide was also not detected in the PlzA protein preparations (Figure 10C). Contrary to c-di-AMP, an effect was observed on PlzA DNA binding affinity when an EMSA was performed with increasing concentrations of exogenous 3'3' cGAMP (Figure 10D). Given that it required 25-fold more cGAMP than c-di-GMP to observe the same level of DNA binding, we conclude that PlzA has a greater affinity for c-di-GMP. *E. coli* is known to have several DGCs, and therefore produces copious amounts of c-di-GMP (62–65). Recently, a horizontally acquired cGAMP synthase was discovered in *E. coli* ECOR31, an animal commensal, and cGAMP was shown to have a regulatory role in biofilm formation (66). It is unclear, however, if common laboratory K-12 strains of *E. coli* possess cGAMP synthases (67). Therefore, cGAMP may not have been co-purified due to a lack of cGAMP production in lab strains of *E. coli* or due to greater affinity of PlzA for c-di-GMP. Direct binding experiments will need to be performed with c-di-AMP and cGAMP to determine the relative affinities of PlzA for these di-nucleotides.

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 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 to maintain vector competence (26, 28–32). Despite its singularity and importance to *B. burgdorferi*, little was previously known about the mechanistic functions of PlzA outside of binding c-di-GMP. Here we provide biochemical evidence that PlzA is a novel c-di-GMP-dependent nucleic acid binding protein that binds DNA and RNA.

While PlzA lacks a 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 (Figure 11A) (26, 29). Without bound c-di-GMP, PlzA has a structure that is flexible, potentially preventing the interactions necessary for DNA binding (Figure 11A left). This hypothesis is supported by the fact that crystallization of PlzA was only achieved with c-di-GMP bound (41). 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 (Figure 11A middle and right). Investigations into apo-PlzA are needed to identify whether it serves a function.

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 (Figure 11B left). In contrast, the C-terminal domain, $PlzA_{CTD}$, did not bind any tested nucleic acids although it still bound c-di-GMP. Noting that binding of c-di-GMP appears to stiffen interactions between the N and C-terminal domains, we suggest that flexibility of apo-PlzA results in the C-terminal domain blocking the ability of the N-terminal domain to bind DNA.

Although the N and C terminal PlzA domains are very similar, structural overlays of $PlzA_{NTD}$ and $PlzA_{CTD}$ reveal structural distinctions (Figure 11B right). The N-terminus of $PlzA_{NTD}$ contains an alpha-helix absent from the N-terminus of $PlzA_{CTD}$. Also, an alpha-helix is present at the beta-barrel pore of $PlzA_{NTD}$ but not in the $PlzA_{CTD}$. These differences are consistent with distinct roles in PlzA DNA binding dynamics. We are currently exploring the sequence motifs recognized by PlzA and the exact amino acid residues involved in nucleic acid binding.

Our studies revealed that increasing $PlzA_{WT}$ concentrations resulted in the formation of multimeric complexes when bound to *glpFKD*(-7/+35) DNA, indicating complexes of multiple proteins and/or DNA molecules. PlzA was found to be a monomer in solution regardless of c-di-GMP (29). While many DNA-binding proteins bind DNA as dimers, alternative binding models exist. It has been shown that some proteins are capable of binding sequentially as monomers and then dimerizing 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* (68–70). It is, therefore, possible that holo-PlzA could multimerize when bound to DNA. A dimerization or protein-protein interaction residue/motif/domain remains to be identified in PlzA. Given the absence of multimeric complexes in EMSAs with $PlzA_{NTD}$, $PlzA_{CTD}$ could be the domain involved in protein-protein interactions facilitating multimerization upon DNA binding.

Another possibility is that the observed EMSA supershifts could indicate bridging of DNA by PlzA. In competition EMSAs, complexes seemed to initially increase as more competitor was titrated which could also indicate bridging activity. Nucleoid-associated proteins (NAPs) are bacterial proteins that can bind DNA nonspecifically or with differential affinities and alter the architecture of the DNA molecule through mechanisms such as bridging, bending, and wrapping, all of which can impact transcription (71–74). Our results indicate that PlzA exhibits NAP-like properties. Recently, CdbA, a c-di-GMP binding protein of *Myxococcus xanthus,* was identified as being a NAP (37).

As of this writing, this is the first record of a c-di-GMP binding protein that can bind both DNA and RNA. More specifically, 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 (56, 59, 75). 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*, depending on c-di-GMP levels (31, 75). Although SpoVG can bind DNA and RNA of *glpFKD*, the transcriptional consequences of these interactions are unknown (59). How PlzA binding directly 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* (76).

To date, this is the first evidence of potential PlzA binding to a di-nucleotide other than c-di-GMP. The caveat of our results is that it is unknown if *B. burgdorferi* produces cGAMP or if exogenous cGAMP is available during the enzootic life cycle and can be imported into the cell. The bacterial di-nucleotide 3'3' cGAMP has been identified as an antiviral signaling molecule, much like its mammalian 2'3' linkage counterpart (77–81). A BLAST analysis of the *B. burgdorferi* B31 proteome did not reveal any plausible homologs to human cGAS synthase or to DncV, the first identified bacterial cGAMP synthase of *Vibrio cholerae* (78–85). Further, alignment of the two borrelial response regulators, Rrp1 and Rrp2, to DncV did not reveal homology or identify any key residues (data not shown). Recent studies have identified bacterial hybrid promiscuous (Hypr) diguanylate cyclases (DGCs) capable of producing both c-di-GMP and cGAMP (86–89). These studies identified a conserved residue within the GGDEF domain as key for di-nucleotide specificity. An aspartic acid residue at this position confers c-di-GMP only specificity, while a serine promotes di-nucleotide promiscuity. We aligned the GGDEF domain of the borrelial DGC, Rrp1, to a non-hybrid/canonical DGC called PleD of *Caulobacter vibroides* and to the Hypr DGCs Bd0367 and GacA of *Bdellovibrio bacteriovorus* and *Geobacter sulfurreducens* respectively (Supplementary Figure S5) (89). Rrp1, like PleD, has an aspartic acid residue (D199) at the conserved specificity site within the GGDEF domain. This infers that Rrp1 is not a Hypr DGC capable of both c-di-GMP and cGAMP production. More work will need to be done to determine if *B. burgdorferi* can produce or import cGAMP during the enzootic life cycle and what role, if any, cGAMP has in PlzA function and overall borrelial fitness.

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 discovered to also bind RNA, thus expanding the functional diversity of c-di-GMP binding proteins. It remains to be seen what additional targets in the *B. burgdorferi* genome/transcriptome fall under the umbrella of the PlzA regulon. The affinity and specificity of proteins for target nucleic acids are impacted by coeffectors (90). We observed that the addition of exogenous c-di-GMP increased DNA binding affinity of PlzA for both DNA and RNA. Cellular levels of c-di-GMP could alter the affinity and/or specificity of PlzA for given targets. Thus, targets regulated by PlzA could depend on the c-di-GMP levels available at different stages of the enzootic life cycle. 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 *Borrelia* regulates gene expression in response to environmental cues and help unravel the PlzA regulon.

AVAILABILITY

Clustal Omega is a MSA tool used to align multiple, medium-large alignments (https://www.ebi.ac.uk/Tools/msa/clustalo/)

Multiple Align Show is protein alignment viewer available through the Sequence Manipulation Suite (https://www.bioinformatics.org/SMS/)

ACKNOWLEDGEMENT

We thank Andrew Krusenstjerna, Tatiana Castro-Padovani, and Jessamyn Moore for their support of the studies undertaken and for feedback on the manuscript. The LC-MS/MS to detect di-nucleotides was performed by the Mass Spectrometry and Metabolomics Core at Michigan State University. A sincere thank you to Tony Schilmiller at MSU for all his help with the mass spectrometry services. Figures 1 and 8A 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

1. 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.
2. Hengge,R. (2009) Principles of c-di-GMP signalling in bacteria. *Nat Rev Microbiol*, **7**, 263–273.
3. Jenal,U., Reinders,A. and Lori,C. (2017) Cyclic di-GMP: Second messenger extraordinaire. *Nat Rev Microbiol*, **15**, 271–284.
4. 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.
5. Valentini,M. and Filloux,A. (2019) Multiple roles of c-di-GMP signaling in bacterial pathogenesis. *Annu Rev Microbiol*, **73**, 387–406.
6. Ausmees,N., Mayer,R., Weinhouse,H., Volman,G., Amikam,D., Benziman,M. and Lindberg,M. (2001) Genetic data indicate that proteins containing the GGDEF domain possess diguanylate cyclase activity. *FEMS Microbiol Lett*, **204**, 163–167.
7. 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.
8. Schmidt,A.J., Ryjenkov,D.A. and Gomelsky,M. (2005) The ubiquitous protein domain EAL is a cyclic diguanylate-specific phosphodiesterase: Enzymatically active and inactive EAL domains. *J Bacteriol*, **187**, 4774–4781.
9. Christen,M., Christen,B., Folcher,M., Schauerte,A. and Jenal,U. (2005) Identification and characterization of a cyclic di-GMP-specific phosphodiesterase and its allosteric control by GTP. *Journal of Biological Chemistry*, **280**, 30829–30837.
10. Ryan,R.P., McCarthy,Y., Andrade,M., Farah,C.S., Armitage,J.P. and Dow,J.M. (2010) Cell-cell signal-dependent dynamic interactions between HD-GYP and GGDEF domain proteins mediate virulence in *Xanthomonas campestris*. *Proc Natl Acad Sci U S A*, **107**, 5989–5994.

11. Andrade,M.O., Alegria,M.C., Guzzo,C.R., Docena,C., Pareda Rosa,M.C., Ramos,C.H.I. and Farah,C.S. (2006) The HD-GYP domain of RpfG mediates a direct linkage between the Rpf quorum-sensing pathway and a subset of diguanylate cyclase proteins in the phytopathogen *Xanthomonas axonopodis* pv citri. *Mol Microbiol*, **62**, 537–551.
12. Amikam,D. and Galperin,M.Y. (2006) PilZ domain is part of the bacterial c-di-GMP binding protein. *Bioinformatics*, **22**, 3–6.
13. Galperin,M.Y. and Chou,S.H. (2020) Structural conservation and diversity of PilZ-related domains. *J Bacteriol*, **202**.
14. 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.
15. Chou,S.H. and Galperin,M.Y. (2016) Diversity of cyclic di-GMP-binding proteins and mechanisms. *J Bacteriol*, **198**, 32–46.
16. 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.
17. 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.
18. 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.
19. Caimano,M.J., Kenedy,M.R., Kairu,T., Desrosiers,D.C., Harman,M., Dunham-Ems,S., Akins,D.R., Pal,U. and Radolf,J.D. (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.
20. 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**.
21. 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.
22. 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.
23. 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.

24. 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.

25. 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 tick-to-mammal transmission. *Infect Immun*, **83**, 3043–3060.

26. 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.

27. 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.

28. 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.

29. Mallory,K.L., Miller,D.P., Oliver,L.D., Freedman,J.C., Kostick-Dunn,J.L., Carlyon,J.A., Marion,J.D., Bell,J.K. and Marconi,R.T. (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.

30. 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.

31. Zhang,J.J., Chen,T., Yang,Y., Du,J., Li,H., Troxell,B., He,M., Carrasco,S.E., Gomelsky,M. and Yang,X.F. (2018) Positive and Negative Regulation of Glycerol Utilization by the c-di-GMP Binding Protein PlzA in *Borrelia burgdorferi*. *J Bacteriol*, **200**.

32. 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**.

33. He,M., Ouyang,Z., Troxell,B., Xu,H., Moh,A., Piesman,J., Norgard,M. v., Gomelsky,M. and Yang,X.F. (2011) Cyclic di-gmp is essential for the survival of the lyme disease spirochete in ticks. *PLoS Pathog*, **7**.

34. Wang,F., He,Q., Su,K., Gao,F., Huang,Y., Lin,Z., Zhu,D. and Gu,L. (2016) The PilZ domain of MrkH represents a novel DNA binding motif. *Protein Cell*, **7**, 766–772.

35. Wilksch,J.J., Yang,J., Clements,A., Gabbe,J.L., Short,K.R., Cao,H., Cavaliere,R., James,C.E., Whitchurch,C.B., Schembri,M.A., *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.

36. Tan,J.W.H., Wilksch,J.J., Hocking,D.M., Wang,N., Srikhanta,Y.N., Tauschek,M., Lithgow,T., Robins-Browne,R.M., Yang,J. and Strugnell,R.A. (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.

37. 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**.

38. Schäper,S., Steinchen,W., Krol,E., Altegoer,F., Skotnicka,D., Søgaard-Andersen,L., Bange,G. and Becker,A. (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.

39. 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.

40. 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.

41. 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.

42. 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.

43. Riley,S.P., Bykowski,T., Cooley,A.E., Burns,L.H., Babb,K., Brissette,C.A., Bowman,A., Rotondi,M., Miller,C.M., Demoll,E., *et al.* (2009) *Borrelia burgdorferi* EbfC defines a newly-identified, widespread family of bacterial DNA-binding proteins. *Nucleic Acids Res*, **37**, 1973–1983.

44. 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.

45. Jutras,B.L., Chenail,A.M., Carroll,D.W., Miller,M.C., Zhu,H., Bowman,A. and Stevenson,B. (2013) 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.

46. Jutras,B.L., Chenail,A.M., Rowland,C.L., Carroll,D., Miller,M.C., Bykowski,T. and Stevenson,B. (2013) Eubacterial SpoVG Homologs Constitute a New Family of Site-Specific DNA-Binding Proteins. *PLoS One*, **8**, 66683.

47. 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.

48. Fraser,C.M., Casjens,S., Huang,W.M., Sutton,G.G., Clayton,R., Lathigra,R., White,O., Ketchum,K.A., Dodson,R., Hickey,E.K., *et al.* (1997) Genomic sequence of a Lyme disease spirochaete, *Borrelia burgdorferi*. *Nature*, **390**, 580–586.

49. Casjens,S., Palmer,N., van Vugt,R., Huang,W.M., Stevenson,B., Rosa,P., Lathigra,R., Sutton,G., Peterson,J., Dodson,R.J., *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.

50. Jumper,J., Evans,R., Pritzel,A., Green,T., Figurnov,M., Ronneberger,O., Tunyasuvunakool,K., Bates,R., Žídek,A., Potapenko,A., *et al.* (2021) Highly accurate protein structure prediction with AlphaFold. *Nature*, **596**, 583–589.

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

52. 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.

53. Sievers,F. and Higgins,D.G. (2018) Clustal Omega for making accurate alignments of many protein sequences. *Protein Science*, **27**, 135–145.

54. Madeira,F., Pearce,M., Tivey,A.R.N., Basutkar,P., Lee,J., Edbali,O., Madhusoodanan,N., Kolesnikov,A. and Lopez,R. (2022) Search and sequence analysis tools services from EMBL-EBI in 2022. *Nucleic Acids Res*, **50**, W276–W279.

55. 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.

56. Grove,A.P., Liveris,D., Iyer,R., Petzke,M., Rudman,J., Caimano,M.J., Radolf,J.D. and Schwartz,I. (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**.

57. 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.

58. Yang,J., Wilksch,J.J., Tan,J.W.H., Hocking,D.M., Webb,C.T., Lithgow,T., Robins-Browne,R.M. and Strugnell,R.A. (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**.

59. Savage,C.R., Jutras,B.L., Bestor,A., Tilly,K., Rosa,P.A., Tourand,Y., Stewart,P.E., Brissette,C.A. and Stevensona,B. (2018) *Borrelia burgdorferi* SpoVG DNA- and RNA-binding protein modulates the physiology of the Lyme disease spirochete. *J Bacteriol*, **200**.

60. Carey,M.F., Peterson,C.L. and Smale,S.T. (2012) Experimental strategies for the identification of DNA-binding proteins. *Cold Spring Harb Protoc*, **7**, 18–33.

61. Krusenstjerna,A.C., Arnold,W.K., Saylor,T.C., Tucker,J.S. and Stevenson,B. (2022) *Borrelia* burgdorferi DnaA and the nucleoid-associated protein EbfC coordinate expression of the *dnaX-ebfC* operon. *J Bacteriol*, **205**, e00396-22.

62. Hengge,R. (2010) Role of Cyclic Di-GMP in the Regulatory Networks of *Escherichia coli*. In Wolfe,A.J., Visick,K.L. (eds), *The Second Messenger Cyclic Di-GMP*. ASM Press, Washington, DC, pp. 230–252.

63. Povolotsky,T.L. and Hengge,R. (2012) 'Life-style' control networks in *Escherichia coli*: Signaling by the second messenger c-di-GMP. *J Biotechnol*, **160**, 10–16.

64. Povolotsky,T.L. and Hengge,R. (2016) Genome-based comparison of cyclic di-GMP signaling in pathogenic and commensal *Escherichia coli* strains. *J Bacteriol*, **198**, 111–126.

65. Hengge,R., Galperin,M.Y., Ghigo,J.M., Gomelsky,M., Green,J., Hughes,K.T., Jenal,U. and Landini,P. (2016) Systematic nomenclature for GGDEF and EAL domaincontaining cyclic di-GMP turnover proteins of *Escherichia coli*. *J Bacteriol*, **198**, 7–11.

66. Li,F., Cimdins,A., Rohde,M., Jänsch,L., Kaever,V., Nimtz,M. and Römling,U. (2019) DncV synthesizes cyclic GMP-AMP and regulates biofilm formation and motility in *Escherichia coli* ECOR31. *mBio*, **10**.

67. Li,F., Cao,L., Bähre,H., Kim,S.K., Schroeder,K., Jonas,K., Koonce,K., Mekonnen,S.A., Mohanty,S., Bai,F., *et al.* (2022) Patatin-like phospholipase CapV in Escherichia coli - morphological and physiological effects of one amino acid substitution. *NPJ Biofilms Microbiomes*, **8**.

68. 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.

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

70. Burns,L.H., Adams,C.A., Riley,S.P., Jutras,B.L., Bowman,A., Chenail,A.M., Cooley,A.E., Haselhorst,L.A., Moore,A.M., Babb,K., *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.

71. Drlica,K. and Rouviere-Yaniv,J. (1987) Histonelike proteins of bacteria. *Microbiol Rev*, **51**, 301–319.

72. Stavans,J. and Oppenheim,A. (2006) DNA-protein interactions and bacterial chromosome architecture. *Phys Biol*, **3**.

73. Dillon,S.C. and Dorman,C.J. (2010) Bacterial nucleoid-associated proteins, nucleoid structure and gene expression. *Nat Rev Microbiol*, **8**, 185–195.

74. Dorman,C.J. (2014) Function of nucleoid-associated proteins in chromosome structuring and transcriptional regulation. *J Mol Microbiol Biotechnol*, **24**, 316–331.

75. 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**

76. 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 glyr, crp, and sigf in *Mycobacterium smegmatis*. *J Bacteriol*, **201**.

77. Cohen,D., Melamed,S., Millman,A., Shulman,G., Oppenheimer-Shaanan,Y., Kacen,A., Doron,S., Amitai,G. and Sorek,R. (2019) Cyclic GMP–AMP signalling protects bacteria against viral infection. *Nature*, **574**, 691–695.

78. Wu,J., Sun,L., Chen,X., Du,F., Shi,H., Chen,C. and Chen,Z.J. (2013) Cyclic GMP-AMP is an endogenous second messenger in innate immune signaling by cytosolic DNA. *Science (1979)*, **339**, 826–830.

79. Sun,L., Wu,J., Du,F., Chen,X. and Chen,Z.J. (2013) Cyclic GMP–AMP synthase is a cytosolic DNA sensor that activates the type I interferon pathway. *Science (1979)*, **339**, 786–791.

80. Gao,P., Ascano,M., Wu,Y., Barchet,W., Gaffney,B.L., Zillinger,T., Serganov,A.A., Liu,Y., Jones,R.A., Hartmann,G., *et al.* (2013) Cyclic [G(2',5')pA(3',5')p] is the metazoan second messenger produced by DNA-activated cyclic GMP-AMP synthase. *Cell*, **153**, 1094–1107.

81. Kranzusch,P.J., Lee,A.S.Y., Wilson,S.C., Solovykh,M.S., Vance,R.E., Berger,J.M. and Doudna,J.A. (2014) Structure-guided reprogramming of human cgas dinucleotide linkage specificity. *Cell*, **158**, 1011–1021.

82. Zhu,D., Wang,L., Shang,G., Liu,X., Zhu,J., Lu,D., Wang,L., Kan,B., Zhang,J. ren and Xiang,Y. (2014) Structural Biochemistry of a *Vibrio cholerae* Dinucleotide Cyclase Reveals Cyclase Activity Regulation by Folates. *Mol Cell*, **55**, 931–937.

83. Ming,Z., Wang,W., Xie,Y., Ding,P., Chen,Y., Jin,D., Sun,Y., Xia,B., Yan,L. and Lou,Z. (2014) Crystal structure of the novel di-nucleotide cyclase from *Vibrio cholerae* (DncV) responsible for synthesizing a hybrid cyclic GMP-AMP. *Cell Res*, **24**, 1270–1273.

84. Whiteley,A.T., Eaglesham,J.B., de Oliveira Mann,C.C., Morehouse,B.R., Lowey,B., Nieminen,E.A., Danilchanka,O., King,D.S., Lee,A.S.Y., Mekalanos,J.J., *et al.* (2019) Bacterial cGAS-like enzymes synthesize diverse nucleotide signals. *Nature*, **567**, 194–199.

85. Davies,B.W., Bogard,R.W., Young,T.S. and Mekalanos,J.J. (2012) Coordinated regulation of accessory genetic elements produces cyclic di-nucleotides for *V. cholerae* virulence. *Cell*, **149**, 358–370.

86. Hallberg,Z.F., Chan,C.H., Wright,T.A., Kranzusch,P.J., Doxzen,K.W., Park,J.J., Bond,D.R. and Hammond,M.C. (2019) Structure and mechanism of a hypr GGDEF enzyme that activates cGAMP signaling to control extracellular metal respiration. *Elife*, **8**, 1–36.

87. Rangarajan,A.A. and Waters,C.M. (2022) Double take: A dual-functional Hypr GGDEF synthesizes both cyclic di-GMP and cyclic GMP-AMP to control predation in *Bdellovibrio bacteriovorus*. *PLoS Genet*, **18**, 1–5.

88. Hallberg,Z.F., Wang,X.C., Wright,T.A., Nan,B., Ad,O., Yeo,J. and Hammond,M.C. (2016) Hybrid promiscuous (Hypr) GGDEF enzymes produce cyclic AMP-GMP (3', 3'-cGAMP). *Proc Natl Acad Sci U S A*, **113**, 1790–1795.

89. Lowry,R.C., Hallberg,Z.F., Till,R., Simons,T.J., Nottingham,R., Want,F., Sockett,R.E., Hammond,M.C. and Lambert,C. (2022) Production of 3',3'-cGAMP by a *Bdellovibrio bacteriovorus* promiscuous GGDEF enzyme, Bd0367, regulates exit from prey by gliding motility. *PLoS Genet*, **18**, 1–24.

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

FIGURES LEGENDS

Figure 1. The *glpFKD* operon, 5' UTR, and designed probe schematic. (**A**) 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 (56). 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. (**B**) A schematic showing the region within the *glpFKD* promoter and 5' UTR corresponding to the probes used in DNA binding experiments. A DNA probe corresponding to a 42 bp sequence adjacent to the -10 site and proceeding 35 bp into the 5' UTR was designated *glpFKD*(-7/+35) was generated. An RNA probe was made corresponding to only nucleotides that are within the 5 'UTR was designated as *glpFKD*(UTR) RNA. Probes were conjugated to a fluorescent molecule for the detection of DNA binding in EMSA.

Figure 2. PlzA binds DNA of the *glpFKD* intergenic region. A representative EMSA performed with labeled probe *glpFKD*(-219), corresponding to the *glpFKD* intergenic region, and increasing concentrations of wild-type recombinant PlzA ($PlzA_{WT}$). Lane 1 is a probe-only control. All lanes contain 100 nM of probe. PlzA protein concentrations were as follows: Lane 2- 0.015 μM, Lane 3- 0.03 μM, Lane 4- 0.075 μM, Lane 5- 0.15 μM, Lane 6- 0.75 μM, Lane 7- 7.5 μM, and Lane 8- 15 μM. A shift (arrow) corresponding to a $PlzA_{WT}$-DNA complex was observed at micromolar concentrations of 7.5 μM and above. Two bands are visible in lane 1 despite gel purification of the probes, suggesting that it forms secondary DNA structures.

Figure 3. 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. 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 $PlzA_{RD\text{-}RD}$ 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 ($PlzA_{RD\text{-}RD}$).

Figure 4. Detection of c-di-GMP in recombinant $PlzA_{WT}$ and $PlzA_{RD\text{-}RD}$ by LC-MS/MS. Recombinant $PlzA_{WT}$ and $PlzA_{RD\text{-}RD}$ were sent off for mass spectrometry to detect c-di-GMP. Extracted ion

chromatograms of the LC-MS/MS analysis are displayed. (**A**) Chromatogram of the control c-di-GMP standard. (**B**) Chromatogram of the blank control. (**C**) Chromatogram of the recombinant $PlzA_{WT}$ sample. (**D**) Chromatogram of the recombinant $PlzA_{RD-RD}$ sample. 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.

Figure 5. Determination of the $K_{d(app)}$ of $PlzA_{WT}$ binding to *glpFKD*(-7/35+). (**A**) A gradient EMSA experiment performed with increasing protein concentrations of $PlzA_{WT}$ with 100 μM c-di-GMP and 10nM fluorescently labeled probe. Lane 1 is a probe-only control. Protein concentrations were as follows: Lane 2- 0.5 μM, Lane 3- 1 μM, Lane 4- 1.5 μM, Lane 5- 2 μM, Lane 6- 2.5 μM, Lane 7- 3.5 μM, Lane 8- 5 μM, Lane 9- 7.5 μM, and Lane 10- 10 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. (**B**) The same EMSA experiment was performed as in Figure 5A but without the addition of 100 μM of c-di-GMP. A representative gel is shown in both Figure 5A and B from triplicate EMSAs. (**C**) Triplicated EMSAs were quantitated by densitometry to determine the $K_{d(app)}$ of the $PlzA_{WT}$- *glpFKD*(-7/35+) interaction. The average of each replicate was analyzed by nonlinear regression using the one-site specific binding setting in GraphPad Prism. Errors bars represent one standard deviation (SD). The $K_{d(app)}$ for each condition is given with the respective standard error of the mean (SEM) and the coefficient of determination (R^2) values.

Figure 6. PlzA has higher affinity to *glpFKD*(-7/35+) than to *flaB* competitor DNA. (**A**) An EMSA performed with a constant $PlzA_{WT}$ protein concentration of 2.5 μM, a c-di-GMP concentration of 100 μM, and a labeled *glpFKD*(-7/35+) probe concentration of 10 nM. A final concentration of 2.5 ng/μL of the nonspecific competitor poly-dI-dC was added to each reaction. Lane 1 and lane 2 are probe only and probe + $PlzA_{WT}$ controls respectively. Unlabeled *glpFKD*(-7/35+) competitor was titrated into the reactions at increasing molar excess relative to labeled probe: Lane 3-1000 nM (100x), Lane 4- 2500 nM (250x), Lane 5- 5000 nM (500x), Lane 6- 10000 nM (1000x), and Lane 7- 20000 nM (2000x). In lanes 8-12, unlabeled *flaB* competitor was titrated into EMSA reactions at the same concentrations as unlabeled *glpFKD*(-7/35+) competitor in lanes 3-7. Complex A is the smaller order $PlzA_{WT}$- *glpFKD*(-7/35+) complex, while complex B is the larger order complex. A representative gel is shown from a triplicated EMSA experiment. (**B**) A bar graph depicting the percent free DNA as quantitated by densitometry. Bars represent the mean calculated from three EMSAs, while error bars represent the standard deviation. The mean values are shown within the bars. Competition with increasing molar excess of unlabeled *glpFKD*(-7/35+) resulted in higher percent free DNA than with unlabeled *flaB*. PlzA has an approximately two-fold higher affinity for *glpFKD*(-7/35+) than to *flaB*.

Figure 7. $PlzA_{WT}$ binds *glp* RNA and RNA binding is also c-di-GMP dependent. (**A**) A gradient EMSA was performed with increasing protein concentrations of $PlzA_{WT}$ with (lanes 1-8) or without (lanes 9-15) 100 μM c-di-GMP and 100 nM fluorescently labeled *glpFKD*(UTR) 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. (**B**) An identical EMSA experiment was performed as in Figure 7A but with $PlzA_{RD-RD}$. Riboguard was supplemented to the $PlzA_{RD-RD}$ EMSA mixtures to a final concentration of 4 U/µL. Representative gels are shown from triplicated ($PlzA_{WT}$) or duplicated ($PlzA_{RD-RD}$) EMSAs.

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. (**B**) A gradient EMSA was performed with increasing protein concentrations of $PlzA_{NTD}$ with (lanes 1-8) or without (lanes 9-15) 100 µM c-di-GMP and 50 nM fluorescently labeled *glpFKD*(-7/35+) probe. Lane 1 is a probe-only control. Protein concentrations were as follows: Lane 2- 0.75 µM, Lane 3- 1 µM, Lane 4- 1.5 µM, Lane 5- 2 µM, Lane 6- 3.5 µM, Lane 7- 5 µM, and Lane 8- 7.1 µM (**C**) An EMSA was performed with increasing protein concentrations of $PlzA_{CTD}$. Lane 1 is a probe-only control. Protein concentrations were as follows: Lane 2- 0.75 µM, Lane 3- 1 µM, Lane 4- 1.5 µM, Lane 5- 2 µM, Lane 6- 3.5 µM, Lane 7- 5 µM, and Lane 8- 7.5 µM. Representative gels are shown from triplicated ($PlzA_{NTD}$) or duplicated ($PlzA_{CTD}$) EMSAs.

Figure 9. Detection of c-di-GMP in recombinant $PlzA_{NTD}$ and $PlzA_{CTD}$ by LC-MS/MS. 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. (**A**) Chromatogram of the control c-di-GMP standard. (**B**) Chromatogram of the blank control. (**C**) Chromatogram of the recombinant $PlzA_{CTD}$ sample. (**D**) Chromatogram of the recombinant $PlzA_{NTD}$ sample. 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 control and blank chromatograms shown are the same as in Figure 4.

Figure 10. Detection of di-nucleotides by LC-MS/MS and impacts on DNA binding affinity. Recombinant proteins $PlzA_{WT}$, $PlzA_{CTD}$, $PlzA_{RD-RD}$, and $PlzA_{NTD}$ were sent off for mass spectrometry to detect the presence of cyclic di-nucleotides. (**A**) Extracted ion chromatograms of the LC-MS/MS analysis for c-di-AMP. None of the proteins co-purified c-di-AMP. (**B**) A representative gel showing an EMSA performed with increasing concentrations of c-di-AMP. 1, 3, or 6 µM of $PlzA_{WT}$ was incubated with *glpFKD*(-7/35+) DNA. Lane 1 is the probe-only control. Lanes 2, 5, 8, and 11= 1 µM, lanes 3, 6, 9, and 12= 3 µM, and lanes 4, 7, 10, and 13= 6 µM of protein. The c-di-AMP concentrations were as follows: Lanes 2-4= 0 µM, lanes 5-7= 5 µM, lanes 8-10= 50 µM, and Lane 11-13= 500 µM. Increased c-di-AMP had no effect on binding. (**C**) Ion chromatogram of the LC-MS/MS analysis for c-GAMP. None of the proteins co-purified c-GAMP. (**D**) A representative gel showing an EMSA performed with increasing concentrations of c-GAMP. The EMSA lanes and concentrations of protein and di-nucleotide are as in (**B**) but with c-GAMP rather than c-di-AMP. Arrows indicate protein-DNA complexes. An effect on DNA binding affinity between that of c-di-GMP and c-di-AMP was observed for c-GAMP. Representative gels are shown from duplicated EMSAs.

Figure 11. The protein structure of PlzA provides insights into DNA binding mechanism. (**A**) Proteins structures of apo-PlzA (AlphaFold) (left), holo-PlzA (PDB ID: 7mie) (middle), and a structural alignment of the holo-PlzA (cyan) and apo-PlzA (beige) forms. The binding of c-di-GMP to PlzA results in a conformational change providing rigidity to the structure of PlzA. (**B**) AlphaFold protein structures of $PlzA_{NTD}$ (left), $PlzA_{CTD}$ (right) and a structural alignment of $PlzA_{NTD}$ (red) and $PlzA_{CTD}$ (blue). Although the two domains are similar, $PlzA_{NTD}$ contains the putative DNA binding residues, while $PlzA_{CTD}$ harbors the c-di-GMP binding motifs. The structural overlay reveals a N-terminal alpha helix (α1) present in $PlzA_{NTD}$ that is absent from $PlzA_{CTD}$. All protein modeling was performed with ChimeraX 1.4 software.

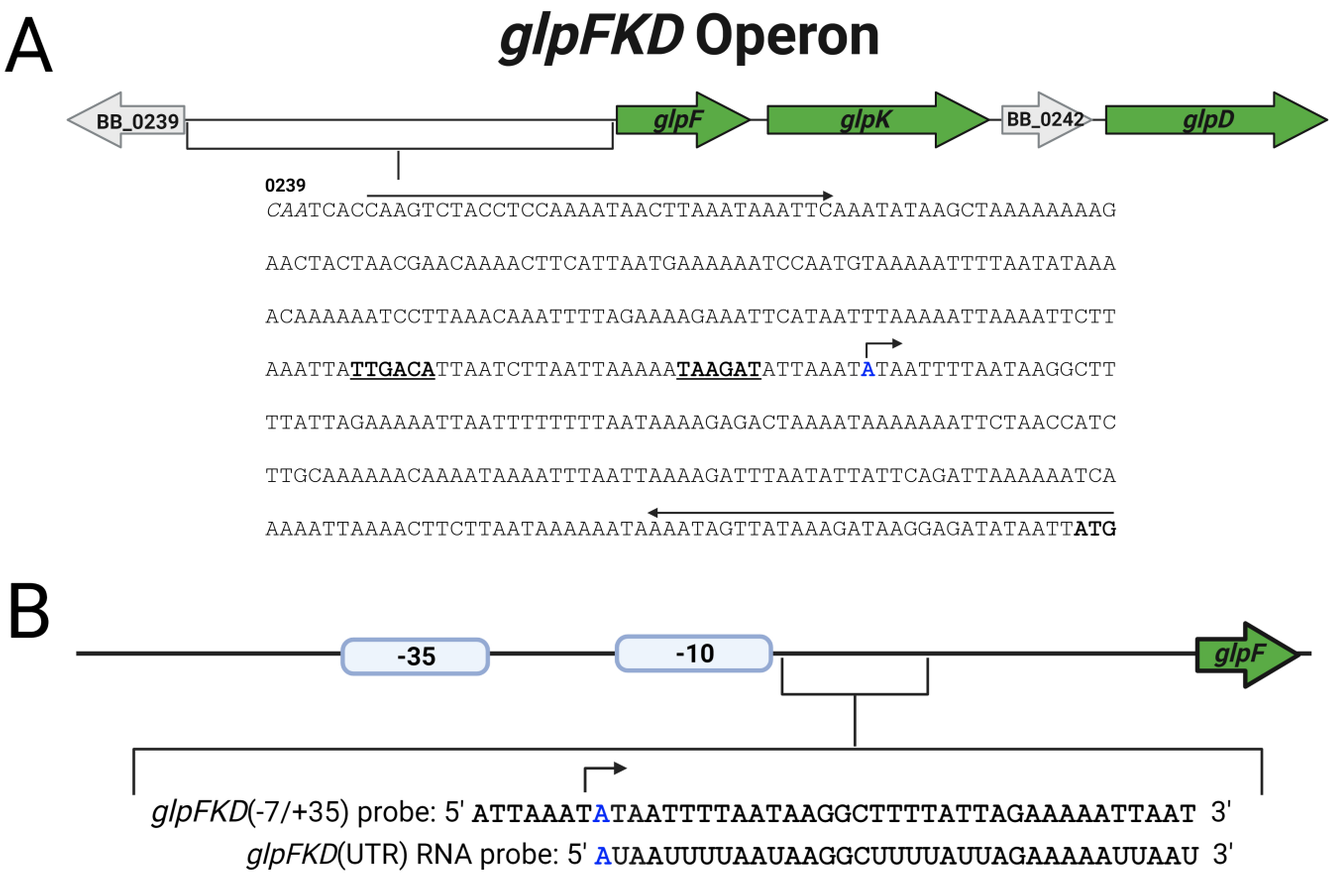


Figure 1. The glpFKD operon, 5' UTR, and designed probe schematic. (A) The glycerol catabolism gene operon (glpFKD) contains four genes: glpF (BB_0240, glycerol uptake facilitator protein), glpK (BB_0241, glycerol kinase), BB_0242 (hypothetical protein), and glpD (BB_0243, glycerol-3-phosphate dehydrogenase). The promoter and transcriptional boundaries were previously mapped (56). 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. (B) A schematic showing the region within the glpFKD promoter and 5' UTR corresponding to the probes used in DNA binding experiments. A DNA probe corresponding to a 42 bp sequence adjacent to the -10 site and proceeding 35 bp into the 5' UTR was designated glpFKD(-7/+35) was generated. An RNA probe was made corresponding to only nucleotides that are within the 5 'UTR was designated as glpFKD(UTR) RNA. Probes were conjugated to a fluorescent molecule for the detection of DNA binding in EMSA.

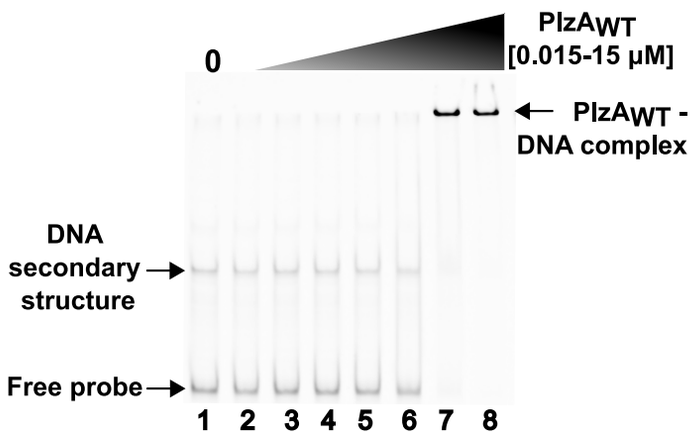


Figure 2. PlzA binds DNA of the glpFKD intergenic region. A representative EMSA performed with labeled probe glpFKD(-219), corresponding to the glpFKD intergenic region, and increasing concentrations of wild-type recombinant PlzA (PlzAWT). Lane 1 is a probe-only control. All lanes contain 100 nM of probe. PlzA protein concentrations were as follows: Lane 2- 0.015 µM, Lane 3- 0.03 µM, Lane 4- 0.075 µM, Lane 5- 0.15 µM, Lane 6- 0.75 µM, Lane 7- 7.5 µM, and Lane 8- 15 µM. A shift (arrow) corresponding to a PlzAWT-DNA complex was observed at micromolar concentrations of 7.5 µM and above. Two bands are visible in lane 1 despite gel purification of the probes, suggesting that it forms secondary DNA structures.

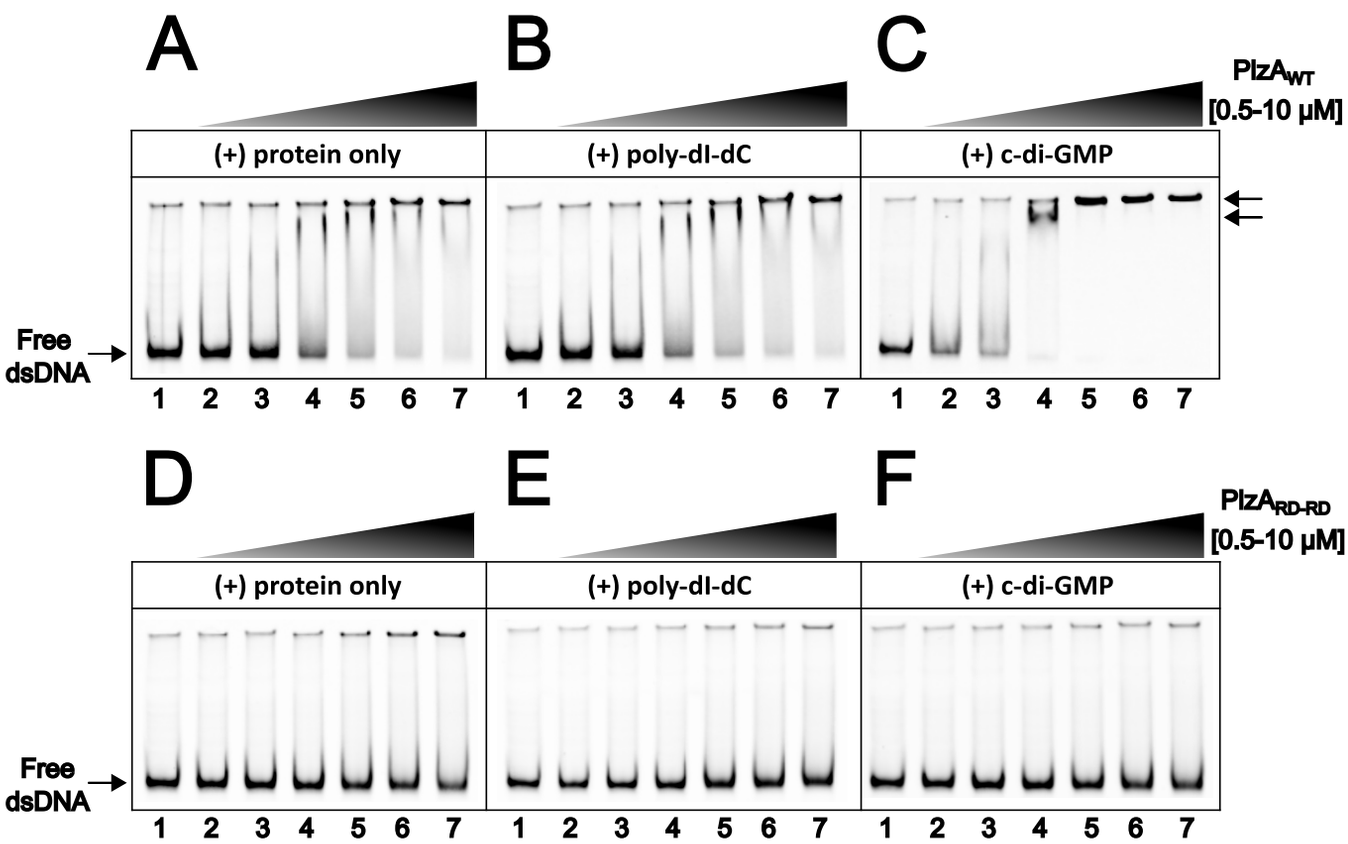


Figure 3. DNA binding by PlzA is c-di-GMP dependent. (A) EMSA with increasing concentrations of PlzAWT showing binding to glpFKD(-219) DNA. The mixture was incubated for 10 min before electrophoresis. (B) The same EMSA mixture as in (A) but with a final concentration of 2.5 ng/ µL of the nonspecific 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. 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 PlzARD-RD protein. All conditions were performed identically as with PlzAWT. Each panel is a representative gel from separate EMSAs which were performed in triplicate (PlzAWT) or duplicate (PlzARD-RD).

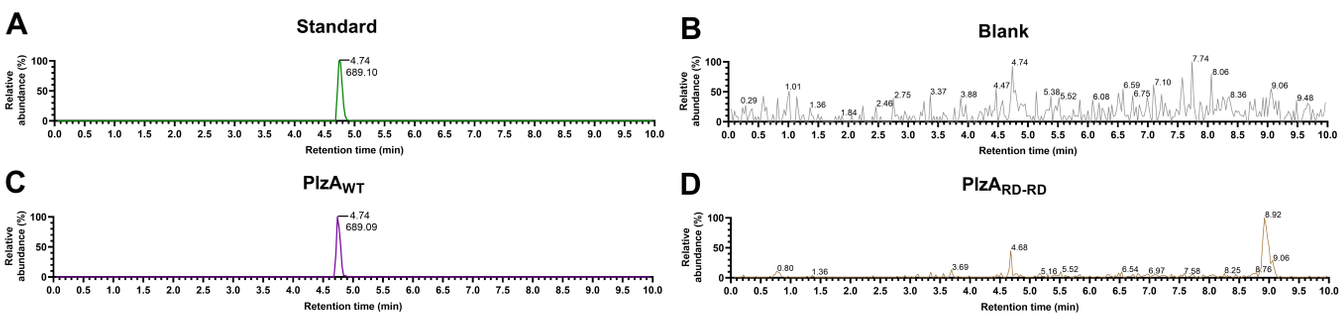
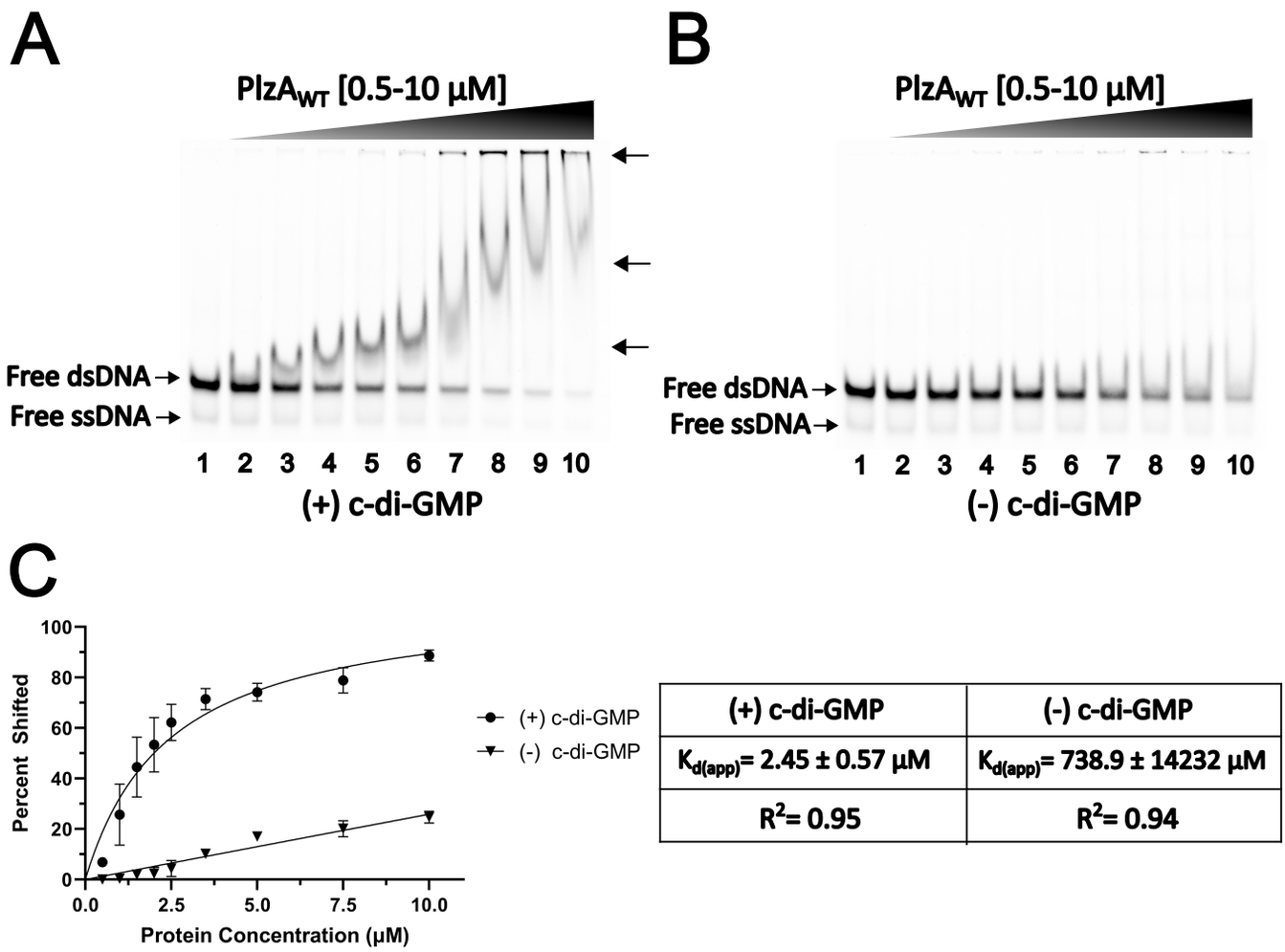


Figure 4. Detection of c-di-GMP in recombinant PlzAWT and PlzARD-RD by LC-MS/MS. Recombinant PlzAWT and PlzARD-RD were sent off for mass spectrometry to detect c-di-GMP. Extracted ion chromatograms of the LC-MS/MS analysis are displayed. (A) Chromatogram of the control c-di-GMP standard. (B) Chromatogram of the blank control. (C) Chromatogram of the recombinant PlzAWT sample. (D) Chromatogram of the recombinant PlzARD-RD sample. The LC-MS/MS analysis reveals PlzAWT co-purified with c-di-GMP, but PlzARD-RD did not. The concentration of c-di-GMP in recombinant PlzAWT was determined from a standard curve of c-di-GMP. PlzAWT co-purified with 99.29 nM of c-di-GMP and approximately 0.3% of the protein was bound with c-di-GMP.



(+) c-di-GMP	(-) c-di-GMP
$K_{d(app)}$= 2.45 ± 0.57 µM	$K_{d(app)}$= 738.9 ± 14232 µM
R^2= 0.95	R^2= 0.94

Figure 5. Determination of the Kd(app) of PlzAWT binding to glpFKD(-7/35+). (A) A gradient EMSA experiment performed with increasing protein concentrations of PlzAWT with 100 µM c-di-GMP and 10nM fluorescently labeled probe. Lane 1 is a probe-only control. Protein concentrations were as follows: Lane 2- 0.5 µM, Lane 3- 1 µM, Lane 4- 1.5 µM, Lane 5- 2 µM, Lane 6- 2.5 µM, Lane 7- 3.5 µM, Lane 8- 5 µM, Lane 9- 7.5 µM, and Lane 10- 10 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. (B) The same EMSA experiment was performed as in Figure 5A but without the addition of 100 µM of c-di-GMP. A representative gel is shown in both Figure 5A and B from triplicate EMSAs. (C) Triplicated EMSAs were quantitated by densitometry to determine the Kd(app) of the PlzAWT- glpFKD(-7/35+) interaction. The average of each replicate was analyzed by nonlinear regression using the one-site specific binding setting in GraphPad Prism. Errors bars represent one standard deviation (SD). The Kd(app) for each condition is given with the respective standard error of the mean (SEM) and the coefficient of determination (R2) values.

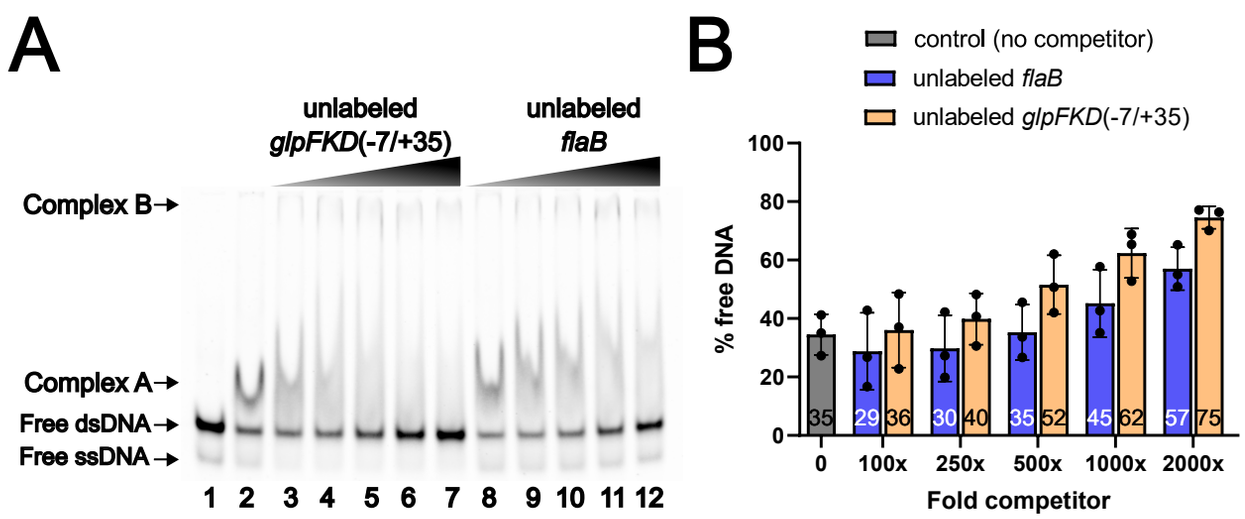


Figure 6. PlzA has higher affinity to glpFKD(-7/35+) than to flaB competitor DNA. (A) An EMSA performed with a constant PlzAWT protein concentration of 2.5 µM, a c-di-GMP concentration of 100 µM, and a labeled glpFKD(-7/35+) probe concentration of 10 nM. A final concentration of 2.5 ng/µL of the nonspecific competitor poly-dI-dC was added to each reaction. Lane 1 and lane 2 are probe only and probe + PlzAWT controls respectively. Unlabeled glpFKD(-7/35+) competitor was titrated into the reactions at increasing molar excess relative to labeled probe: Lane 3-1000 nM (100x), Lane 4- 2500 nM (250x), Lane 5- 5000 nM (500x), Lane 6- 10000 nM (1000x), and Lane 7- 20000 nM (2000x). In lanes 8-12, unlabeled flaB competitor was titrated into EMSA reactions at the same concentrations as unlabeled glpFKD(-7/35+) competitor in lanes 3-7. Complex A is the smaller order PlzAWT- glpFKD(-7/35+) complex, while complex B is the larger order complex.
A representative gel is shown from a triplicated EMSA experiment. (B) A bar graph depicting the percent free DNA as quantitated by densitometry. Bars represent the mean calculated from three EMSAs, while error bars represent the standard deviation. The mean values are shown within the bars. Competition with increasing molar excess of unlabeled glpFKD(-7/35+) resulted in higher percent free DNA than with unlabeled flaB. PlzA has an approximately two-fold higher affinity for glpFKD(-7/35+) than to flaB.

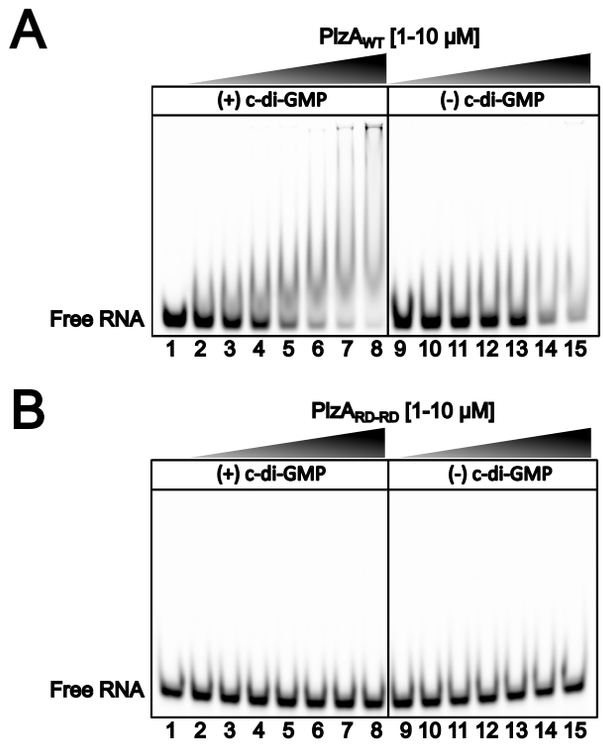


Figure 7. PlzAWT binds glp RNA and RNA binding is also c-di-GMP dependent. (A) A gradient EMSA was performed with increasing protein concentrations of PlzAWT with (lanes 1-8) or without (lanes 9-15) 100 µM c-di-GMP and 100 nM fluorescently labeled glpFKD(UTR) probe. Lane 1 is a probe-only control. Protein 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. (B) An identical EMSA experiment was performed as in Figure 7A but with PlzARD-RD. Riboguard was supplemented to the PlzARD-RD EMSA mixtures to a final concentration of 4 U/µL. Representative gels are shown from triplicated (PlzAWT) or duplicated (PlzARD-RD) EMSAs.

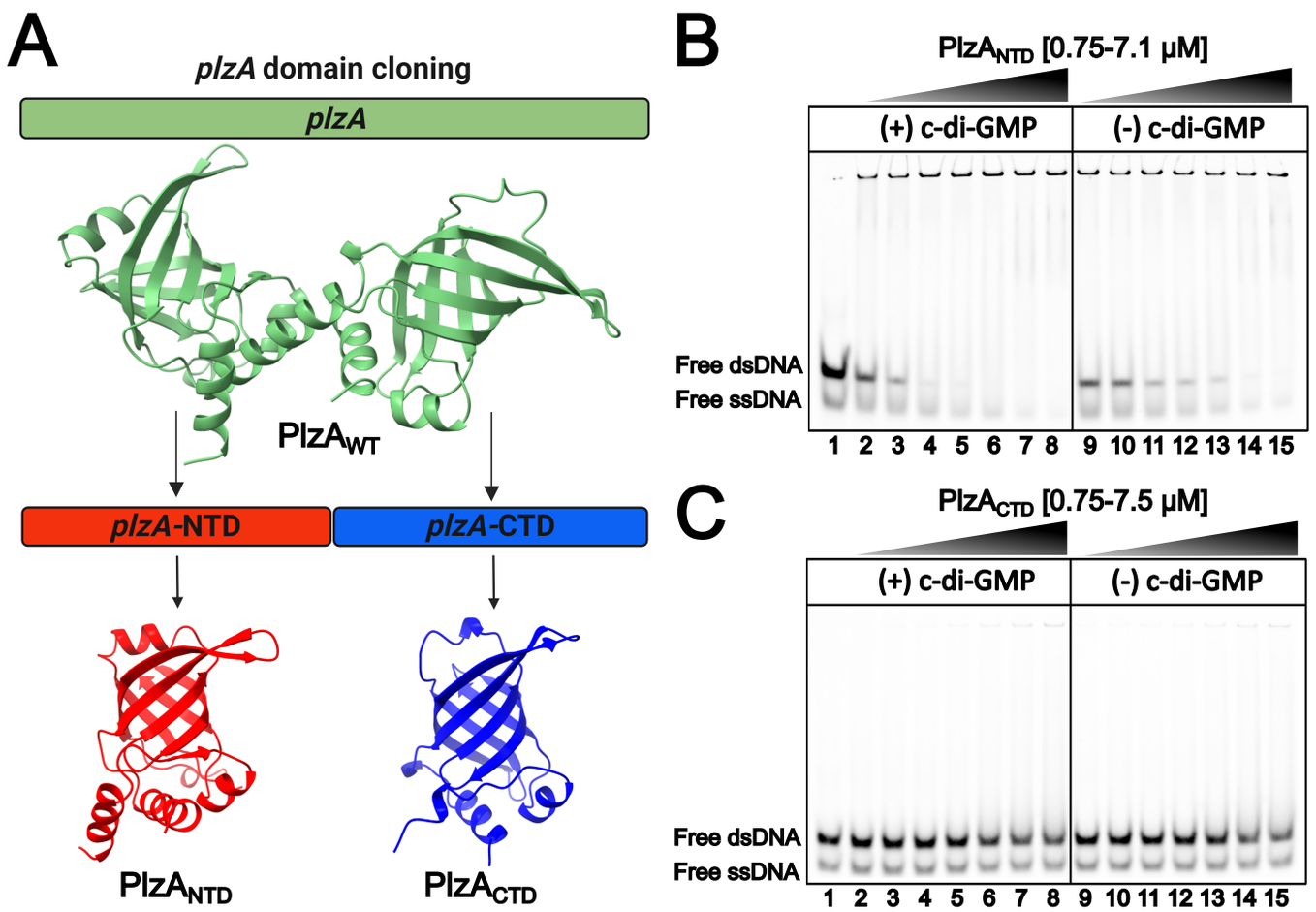


Figure 8. The N-terminal domain of PlzA is the nucleic acid binding domain. (A) The cloning strategy used to produce recombinant N (PlzANTD) and C (PlzACTD) terminal PlzA domains. The plzA gene portions corresponding to the N-terminus, plzA-NTD, and the C-terminus, plzA-CTD, were generated and cloned into an expression vector to produce recombinant PlzANTD and PlzACTD proteins. (B) A gradient EMSA was performed with increasing protein concentrations of PlzANTD with (lanes 1-8) or without (lanes 9-15) 100 µM c-di-GMP and 50 nM fluorescently labeled glpFKD(-7/35+) probe. Lane 1 is a probe-only control. Protein concentrations were as follows: Lane 2- 0.75 µM, Lane 3- 1 µM, Lane 4- 1.5 µM, Lane 5- 2 µM, Lane 6- 3.5 µM, Lane 7- 5 µM, and Lane 8- 7.1 µM (C) An EMSA was performed with increasing protein concentrations of PlzACTD. Lane 1 is a probe-only control. Protein concentrations were as follows: Lane 2- 0.75 µM, Lane 3- 1 µM, Lane 4- 1.5 µM, Lane 5- 2 µM, Lane 6- 3.5 µM, Lane 7- 5 µM, and Lane 8- 7.5 µM. Representative gels are shown from triplicated (PlzANTD) or duplicated (PlzACTD) EMSAs.

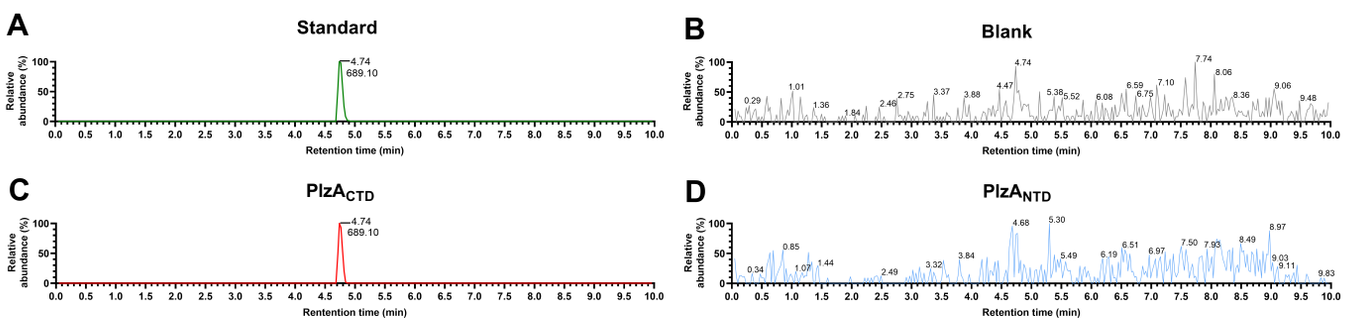


Figure 9. Detection of c-di-GMP in recombinant PlzANTD and PlzACTD by LC-MS/MS. Recombinant PlzANTD and PlzACTD were sent off for mass spectrometry to detect the presence of c-di-GMP. Extracted ion chromatograms of the LC-MS/MS analysis are displayed. (A) Chromatogram of the control c-di-GMP standard. (B) Chromatogram of the blank control. (C) Chromatogram of the recombinant PlzACTD sample. (D) Chromatogram of the recombinant PlzANTD sample. The LC-MS/MS analysis revealed that PlzACTD co-purified with c-di-GMP, but PlzANTD did not. The concentration of c-di-GMP in recombinant PlzACTD was determined from a standard curve of c-di-GMP. PlzACTD 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 control and blank chromatograms shown are the same as in Figure 4.

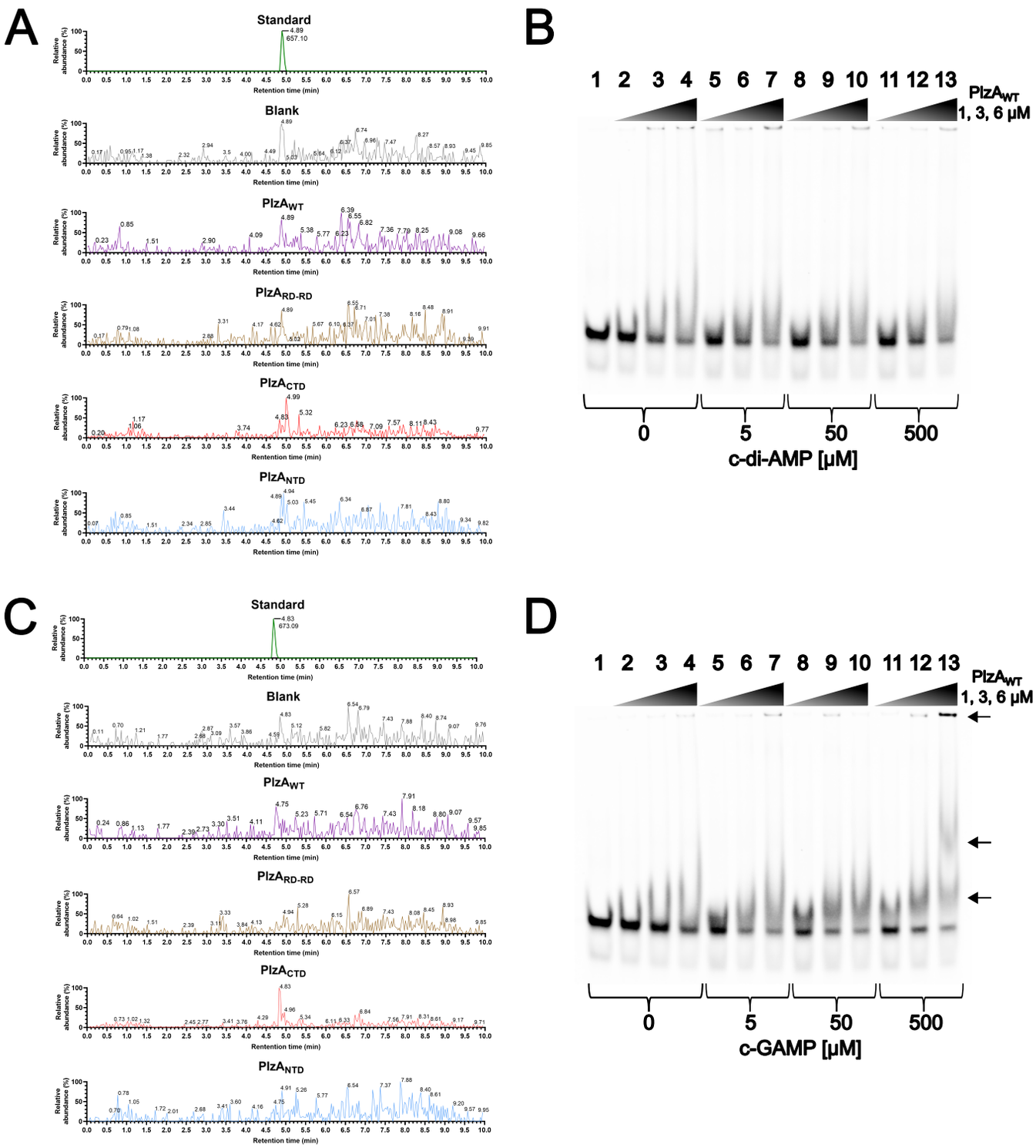


Figure 10. Detection of di-nucleotides by LC-MS/MS and impacts on DNA binding affinity. Recombinant proteins PlzAWT, PlzACTD, PlzARD-RD, and PlzANTD were sent off for mass spectrometry to detect the presence of cyclic di-nucleotides. (A) Extracted ion chromatograms of the LC-MS/MS analysis for c-di-AMP. None of the proteins co-purified c-di-AMP. (B) A representative gel showing an EMSA performed with increasing concentrations of c-di-AMP. 1, 3, or 6 µM of PlzAWT was incubated with glpFKD(-7/35+) DNA. Lane 1 is the probe-only control. Lanes 2, 5, 8, and 11= 1 µM, lanes 3, 6, 9, and 12= 3 µM, and lanes 4, 7, 10, and 13= 6 µM of protein. The c-di-AMP concentrations were as follows: Lanes 2-4= 0 µM, lanes 5-7= 5 µM, lanes 8-10= 50 µM, and Lane 11-13= 500 µM. Increased c-di-AMP had no effect on binding. (C) Ion chromatogram of the LC-MS/MS analysis for c-GAMP. None of the proteins co-purified c-GAMP. (D) A representative gel showing an EMSA performed with increasing concentrations of c-GAMP. The EMSA lanes and concentrations of protein and di-nucleotide are as in (B) but with c-GAMP rather than c-di-AMP. Arrows indicate protein-DNA complexes. An effect on DNA binding affinity between that of c-di-GMP and c-di-AMP was observed for c-GAMP. Representative gels are shown from duplicated EMSAs.

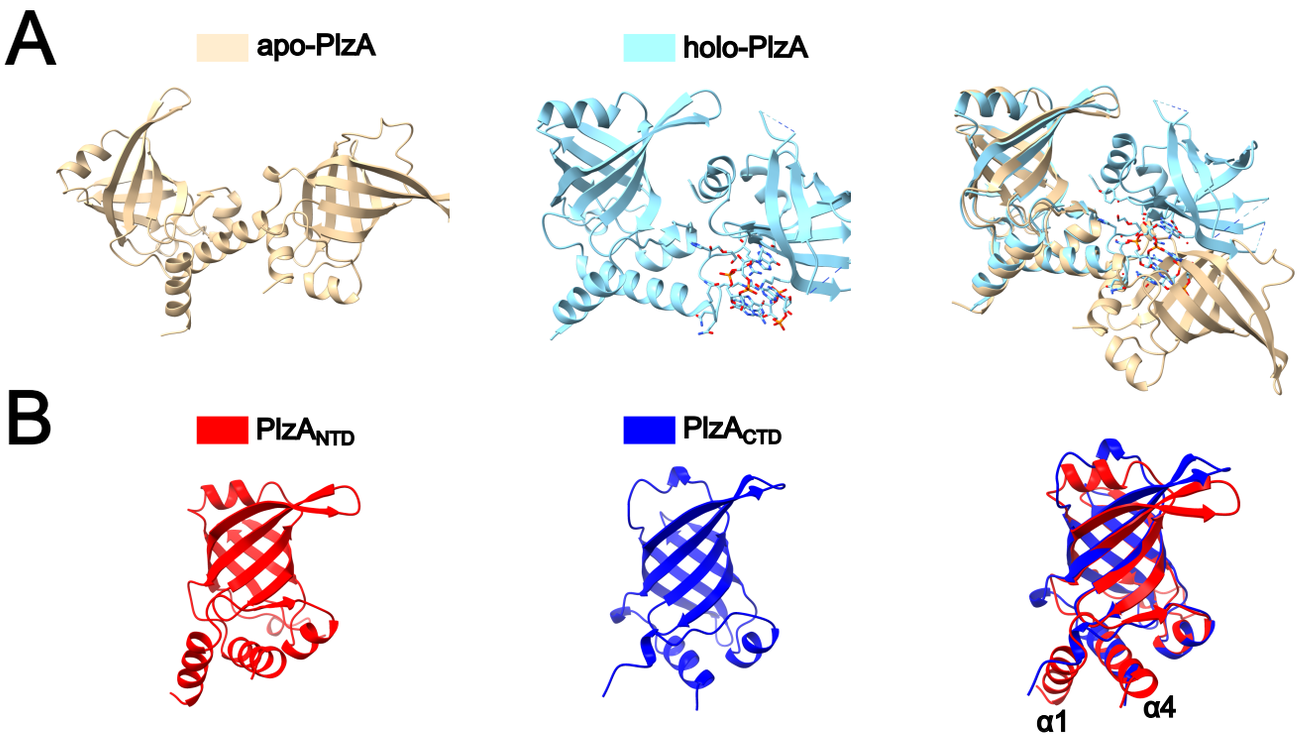


Figure 11. The protein structure of PlzA provides insights into DNA binding mechanism. (A) Proteins structures of apo-PlzA (AlphaFold) (left), holo-PlzA (PDB ID: 7mie) (middle), and a structural alignment of the holo-PlzA (cyan) and apo-PlzA (beige) forms. The binding of c-di-GMP to PlzA results in a conformational change providing rigidity to the structure of PlzA. (B) AlphaFold protein structures of PlzANTD (left), PlzACTD (right) and a structural alignment of PlzANTD (red) and PlzACTD (blue). Although the two domains are similar, PlzANTD contains the putative DNA binding residues, while PlzACTD harbors the c-di-GMP binding motifs. The structural overlay reveals a N-terminal alpha helix (α1) present in PlzANTD that is absent from PlzACTD. All protein modeling was performed with ChimeraX 1.4 software.

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

by

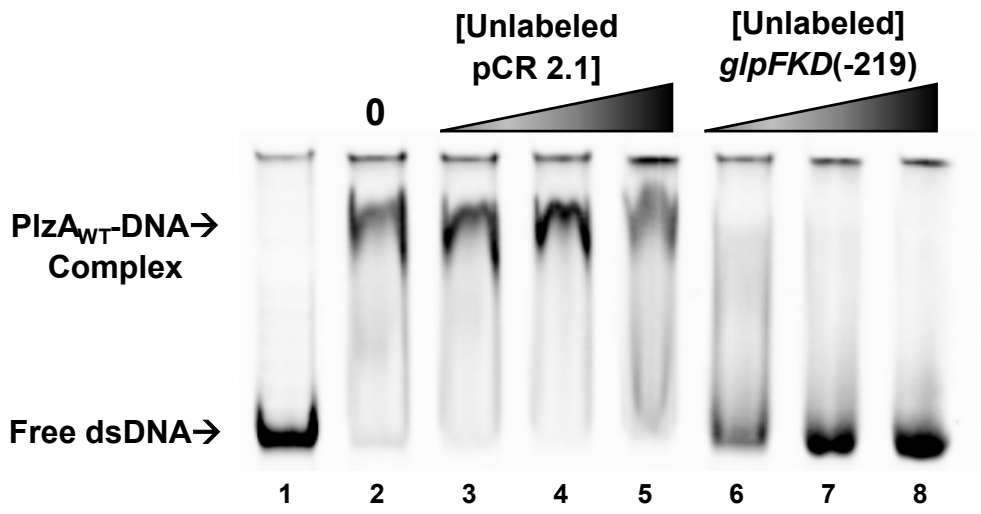
Nerina Jusufovic, Christina R. Savage, Timothy C. Saylor, Catherine A. Brissette, Wolfram R. Zückert, Paula J. Schlax, Md A. Motaleb, and Brian Stevenson

Supplementary Data File

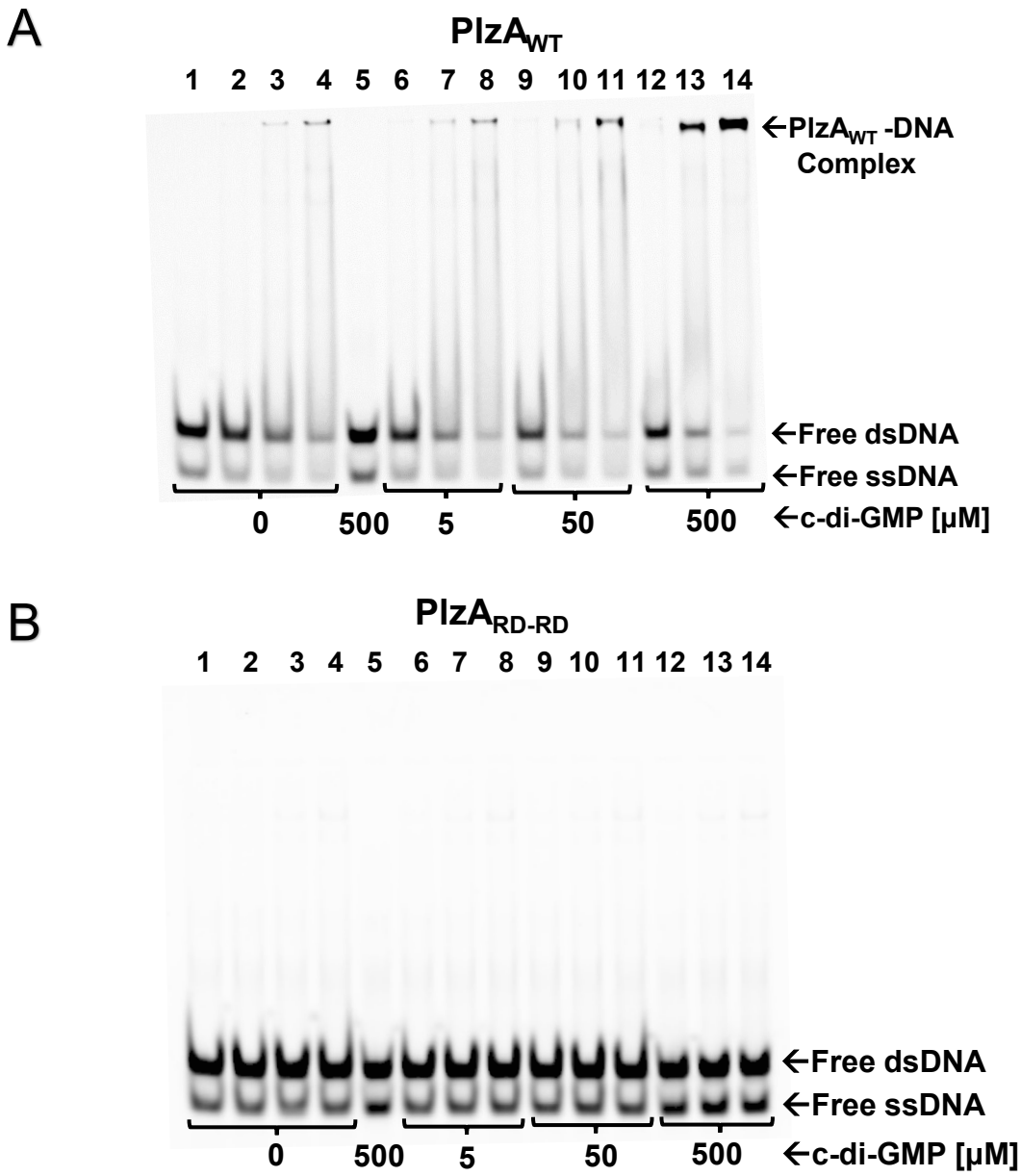
Supplementary Table S1. Oligonucleotides used in this study.

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 and Primers		
glpFKD(-219)	*glp*UP F: CAAGTCTACCTCCAAAATAACTTAAATAAATTC *glp*UP R 5'IRD800: CATAATTATATCTCCTTATCTTTATAACTATTTTAT M13 F: GTAAAACGACGGCCAG M13 R 5'IRD800CW: CAGGAAACAGCTATGAC	410 654
glpFKD(-7/+35)	*glpF* F IRD800: ATTAAATATAATTTTAATAAGGCTTTTATTAGAAAAATTAAT *glpF* R: ATTAATTTTTCTAATAAAAGCCTTATTAAAATTATATTTAAT	42
glpFKD(UTR) RNA	*glpF* RNA 5' Alexa488: AUAAUUUUAAUAAGGCUUUUAUUAGAAAAAUUAAU	35
flaB 1	*flaB* 1F: AACAGGCAAAAGGATTTGCCAAAGTCAGAAATT *flaB* 1R: AATTTCTGACTTTGGCAAATCCTTTTGCCTGTT	33

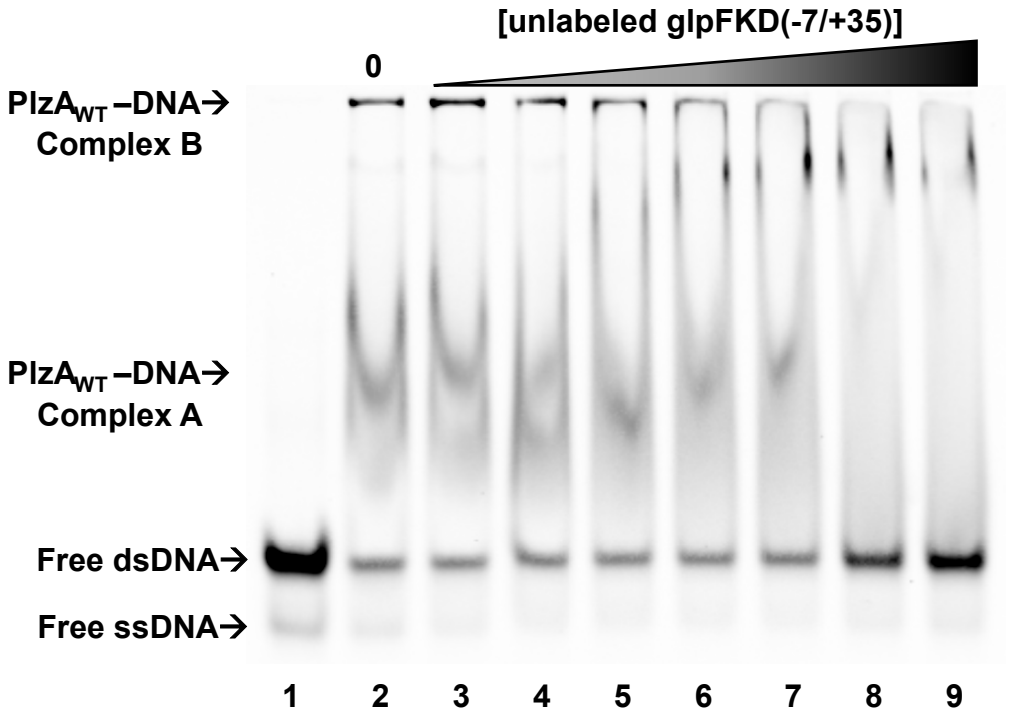
[a] Bolded and underlined nucleotides indicate restriction enzyme sites.



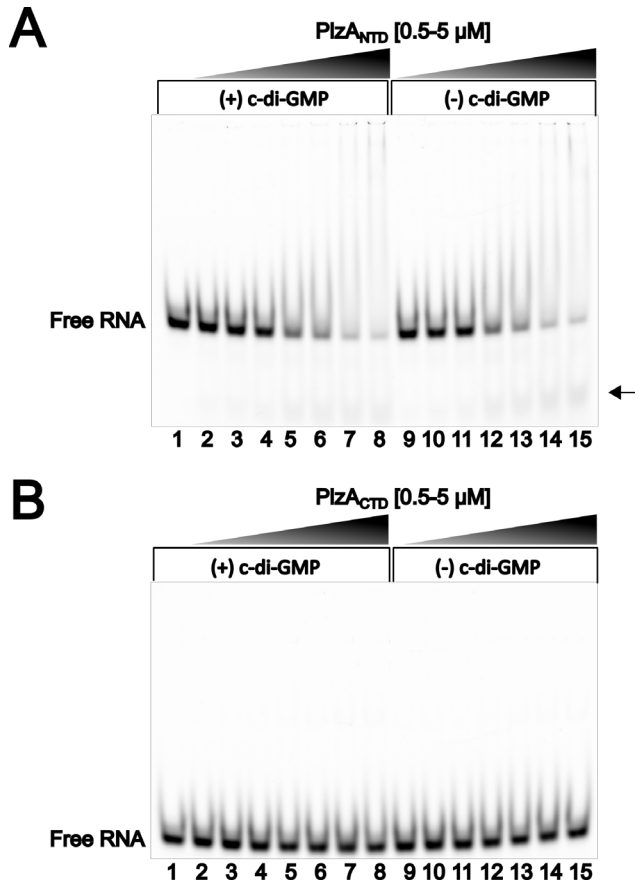
Supplementary Figure S1. PlzA recognizes *glpFKD*(-219) and not pCR 2.1 sequences. As a control, competitive EMSAs were performed with unlabeled DNA competitors derived from empty pCR 2.1 or the vector containing the *glpFKD*(-219) sequence. Lane 1 and lane 2 are probe only and probe + PlzA$_{WT}$ controls, respectively. All lanes contained a c-di-GMP concentration of 100 µM, a final poly-dI-dC concentration of 2.5ng/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. 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 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.



Supplementary Figure S2. Binding to *glpFKD*(-7/35+) is also c-di-GMP dependent. (**A**) Representative EMSA of 1, 3, or 6 μM of PlzA$_{WT}$ incubated with *glpFKD*(-7/35+) DNA at varying c-di-GMP concentrations. Increased c-di-GMP concentrations increase PlzA$_{WT}$ binding affinity for DNA. (**B**) Representative EMSA of 1, 3, or 6 μM of PlzA$_{RD-RD}$ incubated with *glpFKD*(-7/35+) DNA at varying c-di-GMP concentrations. PlzA$_{RD-RD}$ does not bind *glpFKD*(-7/35+) and c-di-GMP shows no effect on DNA binding. For both representative gels, lanes 2, 6, 9, and 12= 1 μM, lanes 3, 7, 10, and 13= 3 μM, and lanes 4, 8, 11, and 14= 6 μM of the respective protein, while lanes 2-4= 0 μM, lanes 6-8= 5 μM, lanes 9-11= 50 μM, and lanes 12-14= 500 μM of c-di-GMP. Lanes 1 and 5 are probe-only controls with 0 or 500 μM of c-di-GMP respectively.



Supplementary Figure S3. Gradient competition EMSA with unlabeled *glpFKD*(-7/35+) competitor. EMSA was performed with a constant PlzA$_{WT}$ protein concentration of 3.5 μM, a c-di-GMP concentration of 100 μM, and a labeled *glpFKD*(-7/35+) probe concentration of 10 nM. A final concentration of 2.5 ng/μL of the nonspecific competitor poly-dI-dC was added to each reaction. Lane 1 and lane 2 are probe only and probe + PlzA$_{WT}$ controls respectively. Unlabeled *glpFKD*(-7/35+) competitor was titrated into the reactions at increasing molar excess relative to labeled probe: Lane 3-50 nM (5x), Lane 4- 100 nM (10x), Lane 5- 250 nM (25x), Lane 6- 500 nM (50x), Lane 7- 1000 nM (100x), Lane 8- 5000 nM (500x), and Lane 9- 10000 nM (1000x). Complex A is the smaller order PlzA$_{WT}$- *glpFKD*(-7/35+) complex, while complex B is the larger order complex. A representative gel is shown from a triplicated EMSA.



Supplementary Figure S4. The N-terminal domain of PlzA also binds RNA. (**A**) A gradient EMSA was performed with increasing protein concentrations of $PlzA_{NTD}$ with (lanes 1-8) or without (lanes 9-15) 100 µM c-di-GMP and 50 nM fluorescently labeled *glpFKD*(UTR) probe. The arrow indicates the degradation of RNA despite the addition of RNase inhibitor. (**B**) An identical EMSA to that in (**A**) was performed but with increasing protein concentrations of $PlzA_{CTD}$. For both gels, lane 1 is a probe-only control. Protein concentrations in the gels were as follows: Lane 2- 0.5 µM, Lane 3- 0.75 µM, Lane 4- 1 µM, Lane 5- 2 µM, Lane 6- 2.5 µM, Lane 7- 3.5 µM, and Lane 8- 5 µM. Riboguard was supplemented to a final concentration of 4 U/µL in all EMSA mixtures. Representative gels are shown from triplicated ($PlzA_{NTD}$) or duplicated ($PlzA_{CTD}$) EMSAs.

Supplementary Figure S5. Protein alignment of Rrp1 to canonical and Hypr DGC GGDEF domains. The amino acid sequences corresponding to the GGDEF domains of the following proteins were aligned using Clustal Omega: *Borrelia burgdorferi* Rrp1 (BB0419), *Caulobacter vibroides* PleD (CB15), *Bdellovibrio bacteriovorus* Bd0367, and *Geobacter sulfurreducens* GacA (GSU1658). The selectivity residues of each protein are outlined in blue. The two Hypr DGCs, Bd0367 and GacA, have serine residues at the selectivity residue, while the canonical DGC PleD has an aspartic acid residue. Residue D199 of Rrp1 corresponds to the selectivity residue indicating Rrp1 is not a Hypr DGC. The GGDEF domain (green) and I-site residues (red) are highlighted for reference. Note that Rrp1 does not harbor the canonical I-site residues as regulatory activity is mediated through phosphorylation via Hk1, the sensory histidine kinase (1). Figure adapted from Lowry et al. 2022 (89) (CC BY 4.0).