

RESEARCH ARTICLE

High-resolution crystal structure of the *Borrelia burgdorferi* PlzA protein in complex with c-di-GMP: new insights into the interaction of c-di-GMP with the novel xPilZ domain

Avinash Singh^{1,†}, Jerilyn R. Izac^{2,†}, Edward J. A. Schuler², Dhara T. Patel², Christopher Davies^{1,‡,§} and Richard T. Marconi^{2,*,¶}

¹Department of Biochemistry and Molecular Biology, Medical University of South Carolina, Charleston, South Carolina 29425, USA and ²Department of Microbiology and Immunology, Virginia Commonwealth University Medical Center, 1112 East Clay Street, Room 101 McGuire Hall, Richmond, VA 23298-0678, USA

*Corresponding author: Department of Microbiology and Immunology, Virginia Commonwealth University Medical Center, 1112 East Clay Street, Room 101, McGuire Hall, Richmond, VA 23298-0678, USA. Tel: 18048283779; E-mail: Richard.marconi@vcuhealth.org

†Present Address: CSAB204, Department of Biochemistry & Molecular Biology, University of South Alabama, 5795 USA Drive North, Mobile, AL 36688, USA.

One sentence summary: Here, the authors present a high-resolution structure of the *Borrelia burgdorferi* PlzA protein complexed with c-di-GMP.

Editor: Hayley Newton

†These authors contributed equally to the work.

§Christopher Davies, <http://orcid.org/0000-0002-6687-4064>

¶Richard T. Marconi, <http://orcid.org/0000-0002-6035-7306>

ABSTRACT

In the tick-borne pathogens, *Borrelia burgdorferi* and *Borrelia hermsii*, c-di-GMP is produced by a single diguanylate cyclase (Rrp1). In these pathogens, the Plz proteins (PlzA, B and C) are the only c-di-GMP receptors identified to date and PlzA is the sole c-di-GMP receptor found in all *Borrelia* isolates. Bioinformatic analyses suggest that PlzA has a unique PilZN3-PilZ architecture with the relatively uncommon xPilZ domain. Here, we present the crystal structure of PlzA in complex with c-di-GMP (1.6 Å resolution). This is the first structure of a xPilZ domain in complex with c-di-GMP to be determined. PlzA has a two-domain structure, where each domain comprises topologically equivalent PilZ domains with minimal sequence identity but remarkable structural similarity. The c-di-GMP binding site is formed by the linker connecting the two domains. While the structure of apo PlzA could not be determined, previous fluorescence resonance energy transfer data suggest that apo and holo forms of the protein are structurally distinct. The information obtained from this study will facilitate ongoing efforts to identify the molecular mechanisms of PlzA-mediated regulation in ticks and mammals.

Keywords: borrelia; borrelia; c-di-GMP; PilZ; xPilZ; PlzA

Received: 25 April 2021; Accepted: 10 June 2021

© The Author(s) 2021. Published by Oxford University Press on behalf of the American Society for Nutrition. This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

INTRODUCTION

Lyme disease (LD) is the most common tick-transmitted infection in the northern hemisphere (Eisen *et al.* 2017; Sykes and Makiello 2017). The primary etiological agent in North America is *Borrelia burgdorferi* (Burgdorfer *et al.* 1982; Benach *et al.* 1983), while in Europe, *B. burgdorferi*, *Borrelia bavariensis*, *Borrelia garinii* and *Borrelia afzelii* are associated with disease (reviewed in O'Bier *et al.* 2020). Here, we collectively refer to *Borrelia* species associated with disease as the LD spirochetes. The Centers for Disease Control and Prevention estimates that there are more than 476 000 new cases of LD in humans each year in the USA alone (Kugeler *et al.* 2021).

The LD spirochetes persist in nature in an enzootic cycle involving *Ixodes* ticks and a diverse and expanding array of mammals and birds (Helble, McCarthy and Hu 2020). The LD spirochetes must be able to adapt to highly variable environmental conditions encountered during the enzootic cycle. The multiple regulatory pathways that are involved in these adaptive responses are complex (reviewed in Samuels *et al.* 2020). Both the LD spirochetes and the closely related tick-borne relapsing fever borrelia encode pathways for the synthesis and degradation of bis-(3'-5')-cyclic dimeric guanosine monophosphate (c-di-GMP), a ubiquitous bacterial second messenger molecule. *Borrelia* and *Borrelia* species encode a single diguanylate cyclase (Ryjenkov *et al.* 2005) designated as response regulatory protein 1 (Rrp1) (Rogers *et al.* 2009). The diguanylate cyclase activity of Rrp1 is phosphorylation dependent (Freedman *et al.* 2009) with Hpk1 being the presumed cognate histidine kinase (Caimano *et al.* 2011). Microarray and RNAseq analyses of *B. burgdorferi* Δ rrp1 deletion mutants revealed that Rrp1, and by extension c-di-GMP, is a positive regulator for the transcription of a large number of genes that encode functionally diverse proteins (Rogers *et al.* 2009; He *et al.* 2014; Caimano *et al.* 2015). C-di-GMP contributes to the regulation of chemotaxis (to N-acetyl glucosamine), motility, glycerol and chitobiose utilization, carbohydrate metabolism and immune evasion (Rogers *et al.* 2009; Pitzer *et al.* 2011; Sze *et al.* 2013; He *et al.* 2014; Caimano *et al.* 2015; Zhang *et al.* 2018). Rrp1 is essential for survival in ticks and, consistent with this, its transcription is upregulated in ticks (Freedman *et al.* 2009). By contrast, *B. burgdorferi* Δ rrp1 mutants retain their ability to infect mice (Caimano *et al.* 2011; Kostick *et al.* 2011; Sze *et al.* 2013), indicating that c-di-GMP is not required for survival in mammals.

Considering the diverse array of cellular processes regulated by c-di-GMP in *B. burgdorferi* (Rogers *et al.* 2009; Kostick *et al.* 2011; He *et al.* 2014; Caimano *et al.* 2015), it is striking that most *Borrelia* strains encode only a single chromosomally encoded c-di-GMP binding protein, PlzA (Freedman *et al.* 2009). PlzA was annotated as a potential c-di-GMP receptor based on the presence of the PilZ signature motifs RXXXX and [D/N]hSXXG in its C-terminal domain. The PlzA:c-di-GMP binding interaction is strongly exothermic with the ΔH (-13.7 kcal mole⁻¹) and ΔS values (-22 cal mol K⁻¹), indicating an enthalpically driven interaction (Mallory *et al.* 2016). The K_d for the PlzA:c-di-GMP interaction is 6.25 ± 1.1 μM (Mallory *et al.* 2016), consistent with that reported for the c-di-GMP interaction with other receptors (Ryjenkov *et al.* 2006). The biological importance of PlzA was demonstrated by Pitzer *et al.* using allelic exchange replacement (Pitzer *et al.* 2011). *Borrelia burgdorferi* Δ plzA mutants are attenuated in their ability to infect mice and are unable to persist in ticks (Pitzer *et al.* 2011; Kostick-Dunn *et al.* 2018). A subset of *Borrelia* isolates carry a linear plasmid gene that encodes a second PilZ domain protein designated as PlzB (Freedman *et al.*

2009). Kostick *et al.* hypothesized that PlzA and PlzB are functionally distinct, as introduction of plzB into a *B. burgdorferi* Δ plzA deletion mutant did not restore wild-type phenotype (Kostick-Dunn *et al.* 2018). The biological significance and function of PlzB remains to be determined.

The molecular basis of PlzA-mediated regulation of cellular processes remains to be determined. PlzA possesses a divergent variant of the canonical PilZ domain referred to as an xPilZ domain (Galperin and Chou 2020). xPilZ domains have been identified in just a few bacteria, including *Xanthomonas campestris*, *Geobacter sulfurreducens*, *Escherichia coli* and *Klebsiella pneumoniae* (Galperin and Chou 2020). An up to date alignment of xPilZ domain amino acid sequences can be found in a recent review (Galperin and Chou 2020). Fluorescence resonance energy transfer (FRET) analyses using YFP-PlzA-CFP fusion proteins demonstrated that *B. burgdorferi* PlzA and *B. hermsii* PlzC undergo significant structural rearrangement upon c-di-GMP binding (Mallory *et al.* 2016). Structural transformation induced by c-di-GMP binding may serve as a toggle switch for c-di-GMP-dependent and c-di-GMP-independent functions in ticks and mammals, respectively (Kostick *et al.* 2011; He *et al.* 2014; Caimano *et al.* 2015; Mallory *et al.* 2016).

To gain insight into the molecular basis of PlzA function, we determined the crystal structure of PlzA complexed with c-di-GMP. This study is the first to determine the structure of a xPilZ domain-containing protein complexed with c-di-GMP. The data provided within provide insight into the molecular transitions that presumably allow PlzA to carry out environment-specific roles in *Borrelia* biology.

MATERIALS AND METHODS

Cloning, protein production and purification

Borrelia burgdorferi plzA and plzA site-directed mutant (S184A) genes were obtained by gene synthesis (codon optimized for *E. coli*; Genscript, Piscataway, N.J. USA) and cloned by the supplier into pET-45b(+). Both constructs express full-length PlzA (residues 1 to 261) with an N-terminal 2.4 kD fusion containing a hexahistidine tag. The plasmids were transformed into BL21/DE3 cells and cultivated in 2 liters of medium at 37°C, with shaking at 250 rpm to an OD_{600 nm} of 0.6. Isopropyl β-D-1-thiogalactopyranoside (1 mM) was added and cultivation continued for 6 h. The cells were recovered by centrifugation (5000 x g; 15 min; 4°C), frozen at -80°C, thawed and suspended in column buffer (50 mM NaH₂PO₄; 300 mM NaCl; 40 mM imidazole). The samples were sonicated and centrifuged to separate the soluble and insoluble fractions (15 500 x g; 30 min; 4°C). PlzA and PlzA_{S184A} were purified from the soluble fraction by nickel affinity chromatography using a Fast Protein Liquid Chromatography system (AKTA; Cytiva, Marlborough, MA, USA). In brief, the soluble fractions (in column buffer) were injected into a 50 mL Superloop (Cytiva) and run through a HisTrap FF column (Cytiva). Proteins were eluted with elution buffer (50 mM NaH₂PO₄; 300 mM NaCl; 500 mM imidazole) and fractions under the peak were collected, pooled and dialyzed into phosphate buffered saline (PBS; overnight) using Spectra/Por®1 (6-8 kDa cutoff) dialysis membranes (Spectrum Laboratories, Rancho Dominguez, CA, USA). Protein concentrations were determined using the BCA assay and protein purity was assessed by SDS-PAGE (AnyKDa Gels, Biorad, Hercules, CA, USA). PlzA protein containing selenomethionine (SeMet) in place of methionine was produced and purified commercially (Harker BIO, Buffalo, NY, USA) using the same expression vector detailed above on a fee-for-service basis.

Table 1. Data collection, refinement and phasing.

Data collection:	SAD	PlzA
Wavelength	0.9713	1.00
Space group	C2	P2 ₁
Cell dimensions:		
<i>a</i> , <i>b</i> , <i>c</i> (Å)	146.4, 36.3, 64.7	63.7, 36.9, 64.9
α , β , γ (°)	90.0, 105.6, 90.0	90.0, 113.9, 90.0
Resolution range (Å)	41.46–2.68	32.24–1.60
R _{merge} (%)	10.5 (33.1)	8.6 (57.3)
R _{pim}	0.024 (0.083)	0.040 (0.268)
Completeness (%)	99.5 (95.1)	97.7 (89.3)
Redundancy	19.7 (15.2)	5.7 (4.5)
/ $\langle\sigma I\rangle$	35.0 (5.0)	31.8 (2.2)
No. of unique reflections	9420 (427)	35 870 (1 619)
Refinement:		
Resolution (Å)	41.46–2.68	32.24–1.60
No. of non-hydrogen protein atoms		1881
No. of waters/ethylene glycols/PEG/chloride ions		122/2/1/1
No. of c-di-GMP atoms		92
R _{cryst} /R _{free} (%)		0.192/0.217
RMS deviations from ideal stereochemistry		
bond lengths (Å)		0.009
bond angles (°)		1.7
B factors:		
Mean B factor (main chain) (Å ²)		33.1
RMS deviation in main chain B factors (Å ²)		1.5
Mean B factor in side chains (Å ²)		37.3
RMS deviation in side chain B factors (Å ²)		2.5
Ramachandran plot:		
Residues in most favored region (%)		90.8
Residues in generously allowed region (%)		9.2
Residues in additional allowed region		0.0
Residues in disallowed region		0.0
PDB code		7MIE

c-di-GMP binding analyses

Radiolabeled c-di-GMP was enzymatically synthesized by incubating 3.3 μ M GTP[α -P³²] (PerkinElmer, Waltham, MA, USA) with purified recombinant *B. burgdorferi* Rrp1, as previously described (Ryjenkov *et al.* 2005) with minor modifications (Freedman *et al.* 2009). c-di-GMP binding assays were conducted as previously described (Freedman *et al.* 2009). In brief, recombinant PlzA (positive control), PlzA_{S184A} and Rrp1 (negative control) proteins (125 ng in PBS) were spotted onto nitrocellulose membranes and air-dried. The membranes were incubated in blocking buffer (PBS, 1% non-fat milk), washed with distilled H₂O and then incubated with either c-di-GMP[α -P³²] or dGTP[α -P³²] (negative control) in blocking buffer (2 h; room temperature). The membranes were washed three times with PBS, wrapped and exposed to X-ray film (90 min; -80°C) with intensifier screens.

Crystallization

For crystallization trials, purified His-tagged PlzA (5 mg ml⁻¹) with and without c-di-GMP (18 mM; Sigma, St. Louis, MO, USA) was incubated overnight at 4°C then centrifuged to remove any precipitate. Using a Gryphon liquid dispensing system (Art-Robbins, Sunnyvale, CA, USA), proteins were screened against the sparse matrix crystallization screens Core I, II, III, IV, JCSG + (Qiagen, Hilden, Germany) in 96-well INTELLI-PLATE 96 (Art-Robbins) by the vapor diffusion method in sitting drop configuration. Two hundred nL of protein was mixed with 200 nL of

well solution and plates were incubated at 18°C. PlzA coincubated with c-di-GMP formed plate-like crystals within a week in 0.1 M succinic acid (pH 7.0) and PEG 3350 (15%; w/v). Conditions were further optimized using 24-well Linbro plates (Jena Bioscience, Jena, Germany) by adjusting the concentration of precipitant and incubation temperature. The best crystals based on appearance and diffraction properties appeared after 7 days of incubation. Crystals of PlzA derivatized with selenomethionine were obtained under the same conditions.

Data collection and structure determination

Both native and SeMet crystals of PlzA were cryo-protected by quick passage through mother liquor (0.1 M succinic acid, pH 7.0) supplemented with PEG 3350 (15% w/v) and ethylene glycol (20% v/v) and flash-frozen in liquid nitrogen. Data were collected at the SER-CAT beamline at the Advanced Photon Source (Argonne, IL). For the SeMet crystals of PlzA, 2.7 Å resolution data were collected at a wavelength of 0.97913 Å on the 22-BM beamline, corresponding to the peak of the anomalous spectrum, as determined by a fluorescence scan ($f = -9.21$, $f' = 11.36$). The crystal-to-plate distance was 330 mm and 360 degrees of data were collected in 1° oscillations on a Rayonix MX300 detector with 2-s exposure per frame. Crystals occupied the C2 space group with cell dimensions $a = 146.4$, $b = 36.3$, $c = 64.7$ Å and $\beta = 105.6^\circ$. A 1.6 Å resolution dataset was then collected on the 22-ID beamline from native crystals of PlzA; 360° of data were

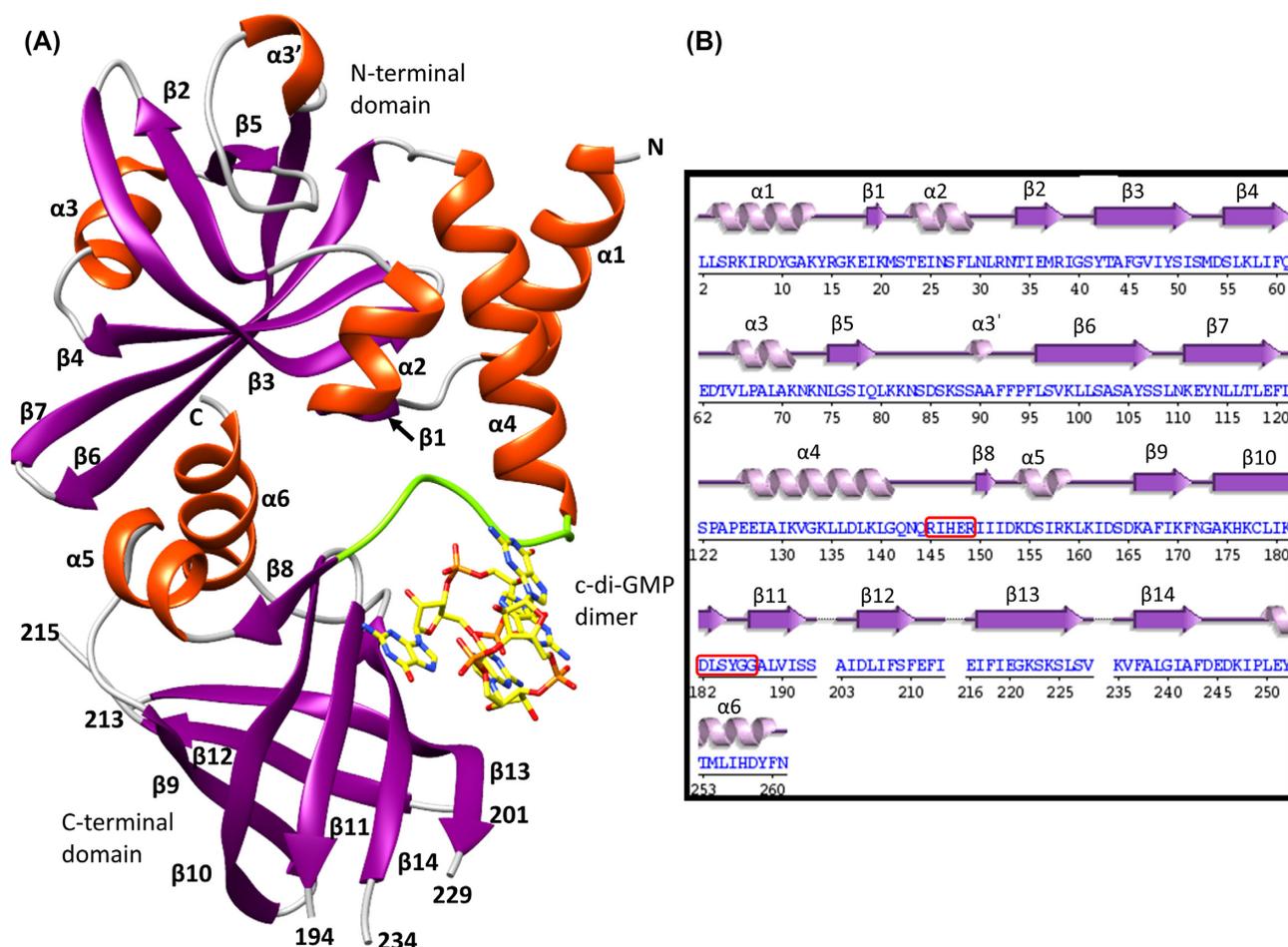


Figure 1. Crystal structure of *B. burgdorferi* PlzA in complex with c-di-GMP dimer. (A) The structure of PlzA shown in ribbon format, where helices and β strands are colored in orange and purple, respectively, and are numbered. The linker joining the N- and C-terminal domains is shown in green. The intercalated c-di-GMP dimer is shown in stick format. (B) Secondary structure assignment of PlzA, where the conserved sequence motifs associated with c-di-GMP binding are boxed in red.

collected in 0.25° oscillations on an EIGER 16M detector with an exposure of 0.2 s per frame. These crystals occupied the $P2_1$ space group with cell dimensions $a = 63.7$, $b = 36.9$, $c = 64.9$ Å and $\beta = 113.9^\circ$. All data were indexed, integrated and scaled using HKL2000 (Otwinowski and Minor 1997).

The structure of SeMet-containing PlzA was solved by single wavelength anomalous diffraction (SAD) phasing using Autosol in PHENIX (Terwilliger et al. 2009). PlzA contains five methionine residues, including one at the N-terminus (Terwilliger et al. 2008). The model was improved by several rounds of refinement and rebuilding with the graphics program, COOT (Emsley et al. 2010) and then used as a search model to solve the structure of native PlzA by molecular replacement using PHASER (McCoy et al. 2007). Two molecules of c-di-GMP were modeled into the $|Fo| - |Fc|$ difference electron density map and the model was refined further using COOT (Emsley et al. 2010) and REFMAC (Murshudov, Vagin and Dodson 1997). The final model has an R_{factor} and R_{free} of 19.2% and 21.7%, respectively (Table 1). The stereochemistry of the model was analyzed with PROCHECK (Laskowski et al. 1996). The entire N domain could be modeled into the electron density map, whereas density was absent or weak for amino acids 194–202, 214–215 and 229–235 in the C-domain, comprising surface loops. The model also contains two c-di-GMP molecules as an intercalated dimer. An example of the final $2|Fo| - |Fc|$ electron density is shown in Fig. S2. Examination of the crystal-packing interactions of the structure shows

that the outermost c-di-GMP (CDG2) molecule is contacted by a symmetry-related molecule (Fig. S3). Specifically, residues from helix $\alpha 1$ in a symmetry-related molecule form hydrogen bonds with the phosphodiester backbone of the c-di-GMP molecule. Since these occur as a consequence of crystal formation, it is unlikely that such interactions have biological relevance. Atomic coordinates and structure factors have been deposited in the Protein Data Bank with accession code 7MIE.

RESULTS AND DISCUSSION

PlzA structure description

The crystal structure of PlzA complexed with c-di-GMP was solved at 1.6 Å resolution in space group $P2_1$. An initial structure was solved at 2.7 Å resolution by SAD from crystals of PlzA containing SeMet obtained in space group C2 and this was used to generate the native structure using crystals obtained in the $P2_1$ space group. Statistics for the data collection and refinement of the native model are shown in Table 1. Attempts to grow crystals of PlzA in the absence of c-di-GMP were unsuccessful, presumably because the protein exhibits flexibility without bound ligand.

PlzA complexed to c-di-GMP has a bilobed structure comprising N- and C-terminal domains separated by an eight-amino acid linker (Fig. 1A). The binding site for c-di-GMP comprises

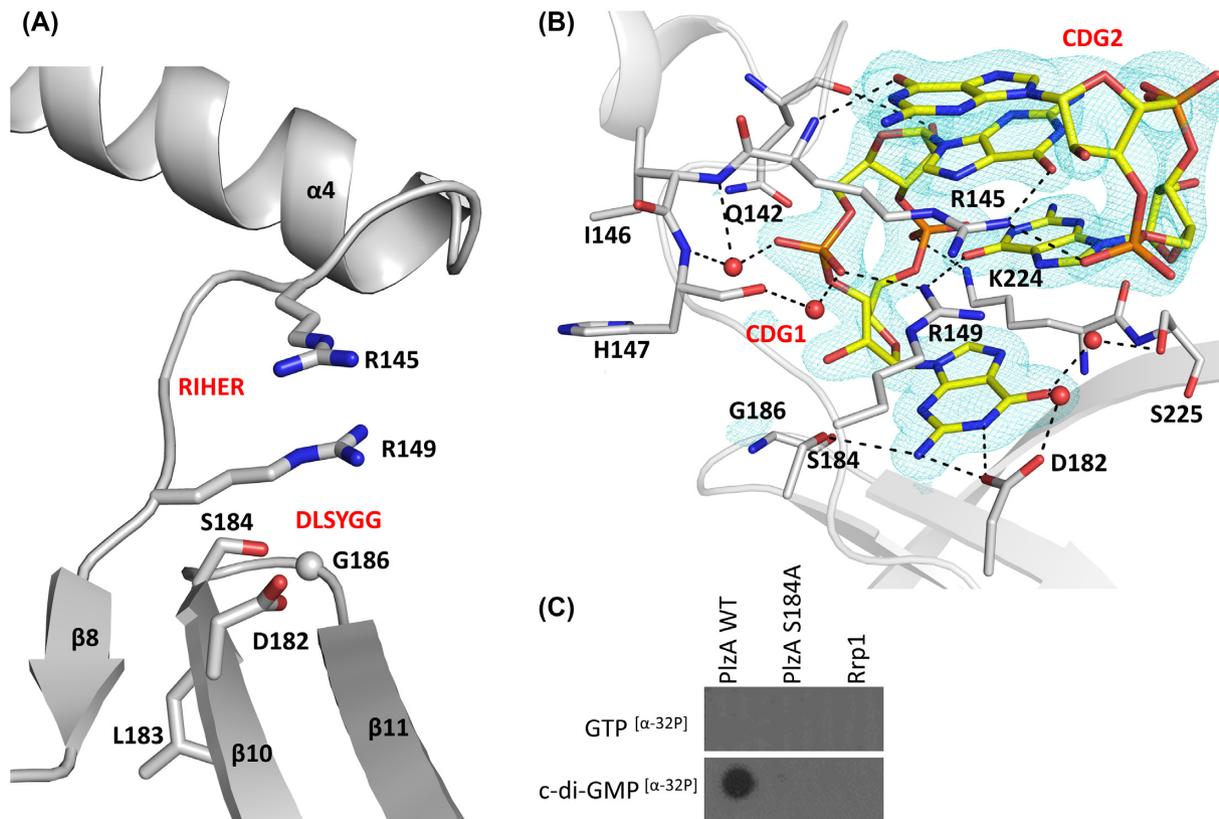


Figure 2. The c-di-GMP binding site of PlzA. **(A)** c-di-GMP-binding residues of the two conserved motifs located on the $\alpha 4$ - $\beta 8$ and $\beta 10$ - $\beta 11$ loops. **(B)** The two molecules of c-di-GMP (labeled CDG1 and 2) bound to PlzA colored with yellow bonds with the corresponding 2[Fo]-F_c electron density map (blue) contoured at 1σ . Residues of PlzA in contact with the ligand are shown with gray bonds and potential hydrogen bonds are shown as dashed lines. Red spheres represent water molecules. **(C)** Recombinant PlzA with a S184A site-directed mutation was tested for c-di-GMP. Binding was assessed using an immunoblot overlay approach as detailed in the text. Wild-type PlzA served as a positive control for c-di-GMP binding and the diguanylate cyclase, Rrp1, served as a negative control.

one face of the β -barrel in the PilZ domain (C-terminal domain), the interdomain linker and the C-terminal end of $\alpha 4$ from the N-terminal domain. The N-terminal domain of PlzA consists of seven β strands ($\beta 1$ - $\beta 7$), four alpha helices ($\alpha 1$ - $\alpha 4$) and a 3_{10} helix ($\alpha 3$). $\beta 1$ is a short β strand between $\alpha 1$ and $\alpha 2$. $\beta 2$ - $\beta 7$ forms a central six-stranded antiparallel β -barrel, with helix $\alpha 3$ linking strands $\beta 4$ and $\beta 5$, and the 3_{10} helix ($\alpha 3$) linking $\beta 5$ and $\beta 6$. The domain ends with a relatively long α helix, $\alpha 4$. The C-terminal domain of the protein is comprised of seven beta strands ($\beta 8$ - $\beta 14$) and two alpha helices ($\alpha 5$ and $\alpha 6$). Similar to the N-terminal domain, the C-terminal domain also comprises a six-stranded barrel. It begins with a short antiparallel β strand ($\beta 8$) and an alpha helix ($\alpha 5$) before forming the β barrel. In contrast to the N-terminal domain, all β strands in the barrel are joined by loops without intervening helices, but it is also notable that two loops ($\beta 11$ - $\beta 12$ and $\beta 13$ - $\beta 14$) are disordered. The *B. burgdorferi* PlzA sequence and its structural elements are shown in Fig. 1B. The fold of the C-terminal domain corresponds to the well-known canonical PilZ domain that is found in many bacterial proteins (Galperin and Chou 2020).

c-di-GMP-binding site

The binding site for c-di-GMP is formed by two conserved sequence motifs, RXXXR and D/NhSXXG (where X is a variable residue and h is a hydrophobic residue) (Amikam and Galperin 2006). In PlzA, these sequences correspond to ¹⁴⁵RIHER and ¹⁸²DLSYGG (Fig. 1B) (Mallory et al. 2016). The RIHER motif

comprises the linker region between the two domains and the DLSYGG motif forms the loop between $\beta 10$ and $\beta 11$ (Fig. 2A). Two molecules of c-di-GMP (designated as CDG1 and CDG2) are observed in the structure as an intercalated dimer (Fig. 2B). The binding of an intercalated c-di-GMP dimer has been reported for the structure of other PilZ-containing proteins, e.g. PP4397 (Ko et al. 2010). Within the c-di-GMP dimer, the four guanine moieties form an internal stack, surrounded by ribosyls and phosphoryls on each side (Fig. 2B). The two Arg residues from the RIHER motif form several hydrogen bonds within this arrangement (Fig. 2B). The guanidinium group of Arg145 interacts with the guanyl oxygen of CDG1 and the O2 phosphoryl of CDG2, and that of Arg149 similarly interacts with the guanyl oxygen of CDG2 and the O2 phosphoryl of CDG1. The two Arg residues also form stacking interactions between guanyl groups of c-di-GMP. Other residues of PlzA also mediate contacts with the c-di-GMP dimer. Asp182 and Ser184 from the $\beta 10$ - $\beta 11$ loop form hydrogen bonds with the guanyl group of CDG1, Lys224 contacts O₂P and the main chain carbonyl oxygen of Gln142 forms a weak hydrogen bond with the 2 hydroxyl group of the ribose sugar of CDG1. There are also interactions involving ordered water molecules. Previous site-directed mutation studies of PlzA confirmed the involvement of Arg145 and Arg149 in c-di-GMP binding (Freedman et al. 2009; Mallory et al. 2016). In this report, we assessed the direct involvement of Ser184 in binding through site-directed mutagenesis. Ser184 was substituted with Ala and the mutated protein was assessed for c-di-GMP binding using an immunoblot overlay assay approach (Fig. 2C) (Freedman et al.

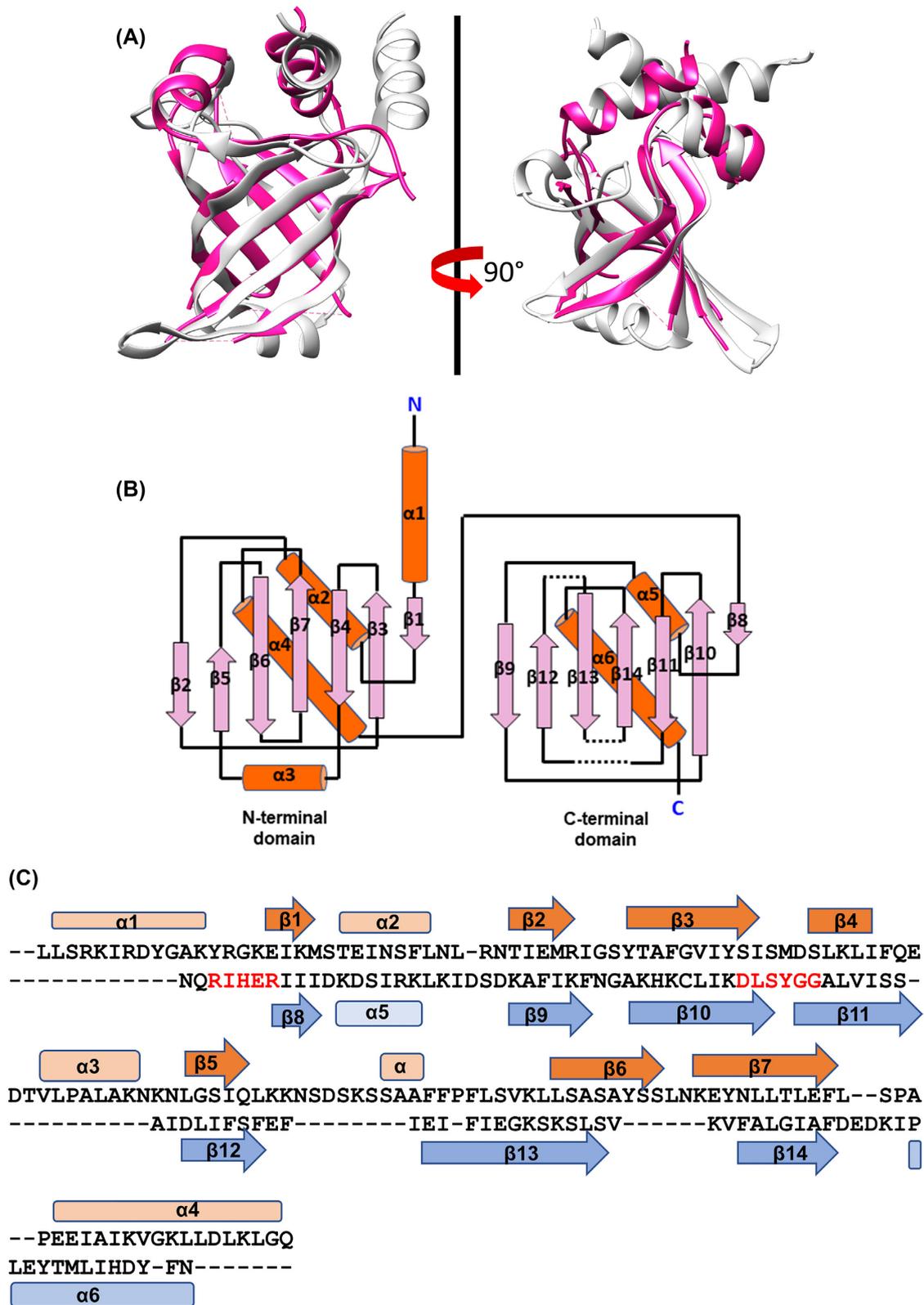


Figure 3. The N- and C-terminal domains of PilZ are structurally equivalent. (A) Superimposition of the N- and C-terminal domains of PilZa, where the N-terminal domain is gray and the C-terminal domain is pink. (B) A topology diagram generated from the superimposition showing the structural relationship between the two domains, where β strands are shown as pink arrows, α helices are orange cylinders and connecting loops are black lines. (C) Alignment of the N- and C-terminal sequences with corresponding secondary structure elements above (N-terminal domain, orange) and below (C-terminal domain, blue).

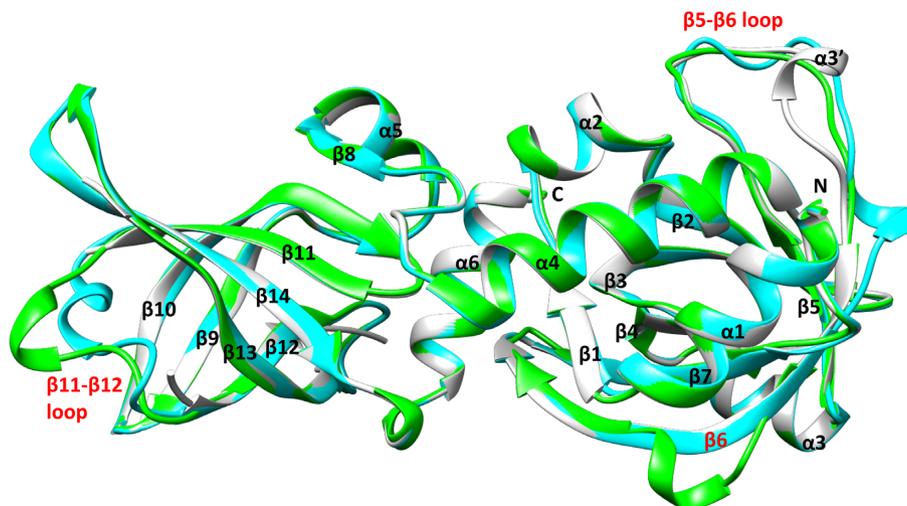


Figure 4. Structural homology between *B. burgdorferi* PlzA and PlzB, and *B. hermsii* PlzC. Homology models for PlzB and PlzC were generated based on the PlzA structure. PlzA, PlzB and PlzC are shown in gray, cyan and green, respectively. Structural elements are indicated and loops where structural differences are noted are labeled in red.

2009). Replacement of S184 with Ala abolished c-di-GMP binding. The positive (PlzA) and negative (Rrp1) controls behaved as anticipated. PlzA bound c-di-GMP whereas Rrp1 did not. Binding specificity was verified by incubating the recombinant proteins with dGTP [α - P^{32}]. No detectable binding was observed.

Domain arrangement of PlzA

The presence of two domains in PlzA with a connecting linker suggests that the protein can occupy different domain arrangements in solution. In support of this, we could only obtain crystals in the presence of c-di-GMP, which suggests that the apo protein exhibits more flexibility compared with the ligand-bound form. It appears that binding of c-di-GMP stabilizes a specific domain arrangement of the protein observed in the crystal structure, whereas the apo protein can occupy multiple conformations. This, in turn, could be related to function, where binding by c-di-GMP leads to a closed conformation of PlzA with specificity for one set of binding targets, and the apo protein has a different set of targets. This is consistent with an earlier FRET-based study that demonstrated a domain rearrangement of PlzA as a result of c-di-GMP binding (Mallory et al. 2016).

PlzA contains two PilZ domains

A notable observation from the crystal structure is that the N- and C-terminal domains of PlzA are highly similar in structure. Using the SSM algorithm in the COOT program (Krissinel and Henrick 2004), 91 common residues comprising the core β -barrel regions of the molecule can be superimposed with an root-mean-square deviation (RMSD) of 1.9 Å (Fig. 3A). The similarity arises because both domains comprise six-stranded β -barrels with the same topology (Fig. 3B). The main differences between the two domains are the presence of α 1 in the N-terminal domain that lacks a counterpart at the N terminus of the C-terminal domain and, as noted, the absence of helices in the connecting loops of the C-terminal domain. Smaller connections make the C-terminal domain smaller overall than the N-terminal domain (116 vs 145 residues, respectively). Their structural similarity means that both domains of PlzA can be considered PilZ domains and this feature appears to be unique

among PilZ domain-containing proteins (Galperin and Chou 2020). Despite this close structural overlap, alignment of the N- and C-terminal domain sequences based on the superimposition revealed no significant sequence identity between the two domains (Fig. 3C).

The close structural similarity between domains raises the question of whether there is any indication of a c-di-GMP binding site in regions of the N-terminal domain equivalent to the two conserved motifs that form the c-di-GMP-binding site in the C-terminal domain. The superimposition shows that the linker domain containing the RIHER motif is topologically equivalent to the loop following the α 1 helix at the beginning of the N-terminal domain. However, while several arginines are present in this region (e.g. arginines 8 and 15), none conform to the RXXXR motif. Similarly, there is no counterpart for the [D/N]hSXXG motif in the connecting loop between β 3 and β 4, which is topologically equivalent to β 10- β 11 in the C-terminal domain (Fig. 3C). It is therefore unclear whether the structural overlap between domains has arisen through gene duplication or through a convergent process driven by the stability of β -barrels.

Homology modeling of *Borrelia burgdorferi* PlzB and *Borrelia hermsii* PlzC

PlzA shares 61% amino acid sequence identity with PlzB from *B. burgdorferi* and 66% with PlzC from *B. hermsii*. Based on this sequence conservation, homology models of PlzB and PlzC were generated based on the structure of PlzA using SWISS-MODEL (Waterhouse et al. 2018). As expected, the structures all adopt the same fold, but there are some differences in loop regions (Fig. 4). In the N-terminal domain, the connecting loop between β 5 and β 6 is structured differently in all three proteins and β 6 appears to deviate from β secondary structure in PlzC. There is relatively less variance in the C-terminal domain and the only parts that differ correspond to disordered regions of PlzA (the β 11- β 12 and β 13- β 14 loops). Of these, the β 11- β 12 loop is predicted to adopt a helical structure in both PlzB and PlzC. The c-di-GMP binding site is highly conserved in the three proteins. Overall, the homology modeling suggests that both PlzB and PlzC adopt essentially the same structure as PlzA and bind c-di-GMP in a similar way.

While PlzA, PlzB and PlzC are likely to interact with c-di-GMP in a similar manner, it is noteworthy that the Plz proteins appear to mediate different functions, since complementation of a *plzA* deletion mutant with *plzB* did not restore wild-type phenotype (Kostick-Dunn et al. 2018).

In conclusion, this study is the first to determine a high-resolution structure of an xPilz domain containing protein complexed with c-di-GMP. The structural information gleaned from these analyses will facilitate efforts to define the functional determinants of PlzA and other xPilz domain containing proteins. The results of this study support the hypothesis that the *apo* and *holo* forms of PlzA are structurally distinct and have c-di-GMP-dependent and c-di-GMP-independent functions in the tick and mammalian environments, respectively (Pitzer et al. 2011; He et al. 2014; Mallory et al. 2016; Kostick-Dunn et al. 2018; Zhang et al. 2018). In future analyses, we will investigate the downstream effector mechanisms of the Plz proteins and assess the potential of specific domains to interact with ligands induced in the tick and mammalian environments.

SUPPLEMENTARY DATA

Supplementary data are available at [FEMSPD](https://www.femspd.com) online.

ACKNOWLEDGEMENTS

Use of the Advanced Photon Source was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract No. W-31-109-ENG-38. Data were collected at Southeast Regional Collaborative Access Team (SER-CAT) 22-ID beam line at the Advanced Photon Source, Argonne National Laboratory. Supporting institutions may be found at www.ser-cat.org/members.html. The X-ray crystallography facility used for this work is supported by the Medical University of South Carolina's Research Resource Facilities program. The liquid handling robot used for crystallization was obtained via an NIH Shared Instrumentation Award (S10 RR027139).

Conflict of interest. None of the authors have a real or perceived conflict of interest.

REFERENCES

- Amikam D, Galperin MY. PilZ domain is part of the bacterial c-di-GMP binding protein. *Bioinformatics* 2006;**22**:3–6.
- Benach JL, Bosler EM, Hanrahan JP et al. Spirochetes isolated from the blood of two patients with Lyme disease. *N Engl J Med* 1983;**308**:740–2.
- Burgdorfer W, Barbour AG, Hayes SF et al. Lyme disease—a tick-borne spirochetosis? *Science* 1982;**216**:1317–9.
- Caimano MJ, Dunham-Ems S, Allard AM et al. 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* 2015;**83**:3043–60.
- Caimano MJ, Kenedy MR, Kairu T et al. 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* 2011;**79**:3117–30.
- Eisen RJ, Kugeler KJ, Eisen L et al. Tick-borne zoonoses in the United States: persistent and emerging threats to human health. *ILAR J* 2017;**58**:319–35.
- Emsley P, Lohkamp B, Scott WG et al. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 2010;**66**:486–501.
- Freedman JC, Rogers EA, Kostick JL et al. 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* 2009;**58**:285–94.
- Galperin MY, Chou SH. Structural conservation and diversity of pilz-related domains. *J Bacteriol* 2020;**202**:e00664–19.
- He M, Zhang JJ, Ye M et al. Cyclic Di-GMP receptor PlzA controls virulence gene expression through RpoS in *Borrelia burgdorferi*. *Infect Immun* 2014;**82**:445–52.
- Helble JD, McCarthy JE, Hu LT. Interactions between *Borrelia burgdorferi* and its hosts across the enzootic cycle. *Parasite Immunol* 2020;**43**:e12816.
- Ko J, Ryu KS, Kim H et al. Structure of PP4397 reveals the molecular basis for different c-di-GMP binding modes by Pilz domain proteins. *J Mol Biol* 2010;**398**:97–110.
- Kostick JL, Szkotnicki LT, Rogers EA et al. The diguanylate cyclase, Rrp1, regulates critical steps in the enzootic cycle of the Lyme disease spirochetes. *Mol Microbiol* 2011;**81**:219–31.
- Kostick-Dunn JL, Izac JR, Freedman JC et al. The *Borrelia burgdorferi* c-di-GMP binding receptors, PlzA and PlzB, are functionally distinct. *Front Cell Infect Microbiol* 2018;**8**:213.
- Krissinel E, Henrick K. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr D Biol Crystallogr* 2004;**60**:2256–68.
- Kugeler KJ, Schwartz AM, Delorey MJ et al. Estimating the frequency of Lyme disease diagnoses, United States, 2010–2018. *Emerg Infect Dis* 2021;**27**:616–9.
- Laskowski RA, Rullmannn JA, MacArthur MW et al. AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J Biomol NMR* 1996;**8**:477–86.
- Mallory KL, Miller DP, Oliver LD, Jr. et al. 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* 2016;**74**:ftw105.
- McCoy AJ, Grosse-Kunstleve RW, Adams PD et al. Phaser crystallographic software. *J Appl Crystallogr* 2007;**40**:658–74.
- Murshudov GN, Vagin AA, Dodson EJ. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* 1997;**53**:240–55.
- O'Bier NS, Hatke AL, Camire AC et al. Human and veterinary vaccines for Lyme disease. *Curr Issues Mol Biol* 2020;**42**:191–222.
- Otwinowski Z, Minor W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* 1997;**276**:307–26.
- Pitzer JE, Sultan SZ, Hayakawa Y et al. Analysis of the *Borrelia burgdorferi* cyclic-di-GMP binding protein PlzA reveals a role in motility and virulence. *Infect Immun* 2011;**79**:1815–25.
- Rogers EA, Terekhova D, Zhang HM et al. Rrp1, a cyclic-di-GMP-producing response regulator, is an important regulator of *Borrelia burgdorferi* core cellular functions. *Mol Microbiol* 2009;**71**:1551–73.
- Ryjenkov DA, Simm R, Romling U et al. The PilZ domain is a receptor for the second messenger c-di-GMP: the PilZ domain protein YcgR controls motility in enterobacteria. *J Biol Chem* 2006;**281**:30310–4.
- Ryjenkov DA, Tarutina M, Moskvina OV et al. Cyclic diguanylate is a ubiquitous signaling molecule in bacteria: insights into biochemistry of the GGDEF protein domain. *J Bacteriol* 2005;**187**:1792–8.
- Samuels DS, Lybecker MC, Yang XF et al. Gene regulation and transcriptomics. *Curr Issues Mol Biol* 2020;**42**:223–66.

- Sykes RA, Makiello P. An estimate of Lyme borreliosis incidence in Western Europe. *J Public Health* 2017;**39**:74–81.
- Sze CW, Smith A, Choi YH et al. Study of the response regulator Rrp1 reveals its regulatory role in chitobiose utilization and virulence of *Borrelia burgdorferi*. *Infect Immun* 2013;**81**:1775–87.
- Terwilliger TC, Adams PD, Read RJ et al. Decision-making in structure solution using Bayesian estimates of map quality: the PHENIX AutoSol wizard. *Acta Crystallogr D Biol Crystallogr* 2009;**65**:582–601.
- Terwilliger TC, Grosse-Kunstleve RW, Afonine PV et al. Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. *Acta Crystallogr D Biol Crystallogr* 2008;**64**:61–9.
- Waterhouse A, Bertoni M, Bienert S et al. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res* 2018;**46**:W296–303.
- Zhang JJ, Chen T, Yang Y et al. Positive and negative regulation of glycerol utilization by the c-di-GMP binding protein PlzA in *Borrelia burgdorferi*. *J Bacteriol* 2018;**200**:e00243–18.