Characterization of cyclic di-GMP binding by the sole Borrelia burgdorferi and Borrelia hermsii PilZ domain-containing proteins

Katherine Louise Mallory
Virginia Commonwealth University

Follow this and additional works at: https://scholarscompass.vcu.edu/etd

Part of the Medicine and Health Sciences Commons

© The Author

Downloaded from
https://scholarscompass.vcu.edu/etd/532

This Thesis is brought to you for free and open access by the Graduate School at VCU Scholars Compass. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of VCU Scholars Compass. For more information, please contact libcompass@vcu.edu.
CHARACTERIZATION OF CYCLIC DI-GMP BINDING BY THE SOLE BORRELIA BURGDORFERI AND BORRELIA HERMSII PILZ DOMAIN-CONTAINING PROTEINS

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science at Virginia Commonwealth University

by

Katherine Louise Mallory
B.S. Virginia Commonwealth University, 2010

Director: Richard T. Marconi, Ph.D.
Professor, Department of Microbiology & Immunology
I would like to take this opportunity to show my appreciation of those individuals who have facilitated my personal and academic growth over these past several years. First and foremost, I would like to thank my advisor, Dr. Richard Marconi for his support and guidance. I would also like to thank my graduate committee, Dr. Todd Kitten, Dr. Kimberly Jefferson, Dr. Jessica Bell, and Dr. Jason Carlyon for investing their time and input into the development of this project. I would also like to thank individuals who contributed a great deal to the results obtained in this thesis in the form of both advice and aid: Dr. Jessica Kostick, Dr. Seth Deever, Dr. James Marrion, Dr. Christopher Earnhart, Dr. John McDowell, Daniel Miller, and Brittney Tegels. Lastly, I would like to thank all of the family and friends who have shared my laughs, dried my tears, and occasionally carried me through the thick and thin of this journey. I have no words for how fortunate I am to have known you all.
# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acknowledgments</td>
<td>ii</td>
</tr>
<tr>
<td>List of Tables</td>
<td>v</td>
</tr>
<tr>
<td>List of Figures</td>
<td>vi</td>
</tr>
<tr>
<td>Abbreviations</td>
<td>vii</td>
</tr>
<tr>
<td>Abstract</td>
<td>ix</td>
</tr>
<tr>
<td>Chapter</td>
<td></td>
</tr>
<tr>
<td>1 Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Borrelia genus</td>
<td>1</td>
</tr>
<tr>
<td>Lyme disease</td>
<td>5</td>
</tr>
<tr>
<td>Tick-borne relapsing fever</td>
<td>13</td>
</tr>
<tr>
<td>Cyclic di-GMP</td>
<td>17</td>
</tr>
<tr>
<td>C-di-GMP signaling in the Borrelia</td>
<td>23</td>
</tr>
<tr>
<td>Research objectives</td>
<td>25</td>
</tr>
<tr>
<td>2 Materials and Methods</td>
<td>28</td>
</tr>
<tr>
<td>Borrelia strains and cultivation</td>
<td>28</td>
</tr>
<tr>
<td>Plasmid construction</td>
<td>28</td>
</tr>
<tr>
<td>Protein expression and purification</td>
<td>30</td>
</tr>
<tr>
<td>FRET analysis</td>
<td>33</td>
</tr>
<tr>
<td>Size exclusion chromatography</td>
<td>33</td>
</tr>
<tr>
<td>Reverse phase chromatography</td>
<td>34</td>
</tr>
<tr>
<td>Isothermal titration calorimetry</td>
<td>34</td>
</tr>
<tr>
<td>3 Results</td>
<td>36</td>
</tr>
<tr>
<td>Recombinant protein purity</td>
<td>36</td>
</tr>
</tbody>
</table>
PlzA and PlzC undergo a FRET-detectable conformational change when c-di-GMP binds
.................................................................................................................. 36

PlzAR145D loses conformational change while PlzAR145K behaves similarly to PlzA ...................................................................................... 39

PlzA and PlzC exist as monomers both bound and unbound to c-di-GMP
.......................................................................................................................... 39

Thermodynamic characterization of the PlzA:c-di-GMP and PlzC:c-di-GMP interactions ........................................................................................................... 43

4 Discussion ......................................................................................................... 47
References .................................................................................................................. 59

Vita .............................................................................................................................. 72
List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-1</td>
<td>Case numbers and incidence rates in the United States per state</td>
<td>8</td>
</tr>
<tr>
<td>2-1</td>
<td>Strains and plasmids used in this study</td>
<td>29</td>
</tr>
<tr>
<td>2-2</td>
<td>Oligonucleotides used in this study</td>
<td>31</td>
</tr>
<tr>
<td>3-1</td>
<td>Thermodynamic profiles of PlzA, PlzC, PlzAR145D, and PlzAR145K</td>
<td>46</td>
</tr>
</tbody>
</table>
List of Figures

| Figure 1-1: | *Borrelia* periplasmic flagella | 2 |
| Figure 1-2: | Spirochete motility controlled by flagellar rotational direction | 4 |
| Figure 1-3: | 2011 case distribution of Lyme disease in the United States | 7 |
| Figure 1-4: | Enzootic cycle of the Lyme disease spirochete | 10 |
| Figure 1-5: | Tick-borne relapsing fever incidence in the United States | 15 |
| Figure 1-6: | Cyclic di-GMP structure | 18 |
| Figure 1-7: | Cyclic di-GMP conformations allow for multimerization | 19 |
| Figure 1-8: | Sequence homology between the *B. burgdorferi* and *B. hermsii* c-di-GMP metabolizing proteins | 24 |
| Figure 1-9: | Schematic of the c-di-GMP network of *B. burgdorferi* | 26 |
| Figure 2-1: | Schematic of FRET constructs and recombinant proteins | 32 |
| Figure 3-1: | SDS-PAGE of recombinant proteins purified in this study | 37 |
| Figure 3-2: | Fluorescence resonance energy transfer for conformational change detection | 38 |
| Figure 3-3: | C-di-GMP induced conformational change by PlzA and PlzC detectable by FRET | 40 |
| Figure 3-4: | FRET analyses of PlzA with a drastic and conservative substitution mutation of arginine 145 | 41 |
| Figure 3-5: | PlzA and PlzC exist as monomers in both the apo and holo form | 42 |
| Figure 3-6: | PlzA and PlzC bind c-di-GMP with a moderate affinity | 44 |
| Figure 4-1: | Sequence alignment of the *Borrelia* PilZ domain proteins with other characterized PilZ domain proteins | 54 |
Abbreviations

Δ Deletion
ABC ATP-binding cassette.
Arg Arginine
B. Borrelia
bp Base pair
BSK Barbour-Stoenner-Kelly medium
°C Degrees Celsius
CDC Centers for Disease Control and Prevention
C-di-GMP Bis-(3’-5’)-cyclic-dimeric guanosine monophosphate
CFP Cyan fluorescent protein
cm Centimeter
CO₂ Carbon dioxide
cp26 Circular plasmid 26
D Aspartic acid
DGC Diguanylate cyclase
DNA Deoxyribonucleic acid
EAL Glutamate-alanine-leucine
E. Escherichia
°F Degrees Fahrenheit
FRET Fluorescence resonance energy transfer
G. Gluconacetobacter
gDNA Genomic deoxyribonucleic acid
GGDEF Glycine-glycine-aspartate-glutamate-phenylalanine
glp Glycerol metabolism operon
GMP Guanosine monophosphate
GTP Guanosine-5'-triphosphate
HCl Hydrochloric acid
HD-GYP Histidine-aspartate-glycine-tyrosine-proline
His Histidine
hr Hour
Hpk1 Histidine sensor kinase 1
HPLC High-performance liquid chromatography
I. Ixodes
IPTG Isopropylthio-β-D-galactoside
ITC Isothermal titration calorimetry
K Lysine
kb  Kilobase
kDa  KiloDalton
KH$_2$PO$_4$  Monopotassium phosphate
μg  Microgram
μl  Microliter
μM  Micromolar
mg  Milligrams
mL  Milliliter
mM  Millimolar
min  Minute
mRNA  Messenger ribonucleic acid
M  Molar
NaCl  Sodium chloride
ng  Nanogram
Ni-NTA  Nickel-nitrilotriacetic acid
nm  Nanometer
nM  Nanomolar
O.  *Ornithodoros*
OspC  Outer surface protein C
*P.*  *Pseudomonas*
PAGE  Polyacrylamide gel electrophoresis
PBS  Phosphate buffered saline
PCR  Polymerase chain reaction
PDE  Phosphodiesterase
PdeA  Phosphodiesterase A
PdeB  Phosphodiesterase B
pGpG  5'  Phosphoguanylyl- (3' -> 5')- guanosine
PlzA  PilZ domain-containing protein A
PlzC  PilZ domain-containing protein C
R  Arginine
r-  Recombinant
RNA  Ribonucleic acid
rpm  Revolutions per minute
Rpo  RNA polymerase
Rrp1  Response regulator protein 1
*S.*  *Salmonella*
SDS  Sodium dodecyl sulfate
TBRF  Tick-borne relapsing fever
Tris  Tris(hydroxymethyl)aminomethane
*V.*  *Vibrio*
*vmp*  Variable major protein
x g  Gravitational force
YFP  Yellow fluorescent protein
Abstract

CHARACTERIZATION OF CYCLIC DI-GMP BINDING BY THE SOLE BORRELIA BURGDORFERI AND BORRELIA HERMSII PILZ DOMAIN-CONTAINING PROTEINS

By Katherine L. Mallory, M.S.

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science at Virginia Commonwealth University

Virginia Commonwealth University, 2013

Major Director: Richard T. Marconi, Ph.D.
Professor, Department of Microbiology & Immunology

Borrelia burgdorferi and Borrelia hermsii cause Lyme disease and relapsing fever, respectively. These spirochetes are maintained in an enzootic cycle, involving tick vectors and mammalian hosts. Differential gene expression is central in their survival in various environmental conditions. C-di-GMP has been demonstrated to be important in bacterial adaptation. Borrelia deletion mutant phenotypes have shown that c-di-GMP regulates motility, infectivity, and enzootic cycle progression. As the only known receptors encoded by Borrelia, PlzA and PlzC characterization is necessary in delineating c-di-GMP roles within the cell. In this study, biochemical, biophysical, and FRET methods demonstrated that these proteins exhibit a structural rearrangement when binding c-di-GMP likely significant to downstream activities. Substitution of a highly conserved residue within PlzA altered the structure and charge of the PilZ
domain, leading to abolished binding. PlzA and PlzC functionality studies are vital to discover mechanisms of c-di-GMP-mediated regulation of motility and host invasion by the Borrelia.
Chapter 1: Introduction

*Borrelia* genus

The genus *Borrelia* contains a group of enigmatic bacterial species which are defined by many unique properties. Despite being Gram-negative, the *Borrelia* lack the lipopolysaccharides characteristic of most Gram-negative species (134). They have developed highly fragmented genomes, consisting of a linear chromosome and multiple linear and circular plasmids (11, 13). In addition to fragmentation, *Borrelia* genomes have severely reduced metabolic and biosynthetic genetic information, making them highly reliant on their hosts for nutrients and basic metabolites (52, 99). These species have a distinct spiral morphology, the architecture of which is generated by the polar flagellar bundles (16). This spiral shape confers a very unique motility phenotype and contributes to infectivity (74, 130). Despite and because of these unusual properties, *Borrelia* species cause multiple infections across a range of mammalian, reptilian, avian, and arthropod hosts.

Most motile bacteria have external flagella or cilia to allow for translational motion (18). However, the *Borrelia* are unique in that their flagella are contained within the periplasmic space between the outer and cytoplasmic membrane depicted in Figure 1-1. These flagellar bundles are comprised of seven to eleven filaments insert into the cytoplasmic membrane at each end of the bacterium (75). These filaments form flat
**Figure 1-1. Borrelia periplasmic flagella.** The top panel is a depiction of a spirochete with an outer and cytoplasmic membrane, between which lies the periplasmic space. The unique, internal flagella wind around the protoplasmic cylinder, generating the spirochete morphology. The bottom panel shows the flagellum inserted into the cytoplasmic membrane by the flagellar motor.
ribbons and wind around the cell cylinder to generate the characteristic shape of these spirochetes (102). Each flagellum anchors into the cytoplasmic membrane by a motor complex that controls the rotational motion of the flagellum (75). Each of these motors is capable of clockwise and counter-clockwise rotation, which results in a “tumble” or “run” motility in most bacteria (18). However, due to the internal flagellar bundles at each pole, the *Borrelia* exhibit a “run,” “flex,” and “reverse” phenotype shown in Figure 1-2 (83). The “run” is when both flagellar motors are rotating in opposite directions and results in a corkscrew-like motility. When one bundle changes flagellar rotation, a spasming of the bacterium occurs. This is referred to as a “flex” and is a pause in translational motion, much like a “tumble” (18, 83). When the bundles are rotating asymmetrically once more, but in the opposite direction of the original “run,” it is a “reverse.” This motility pattern allows spirochetes to migrate through highly viscous environments such as connective tissue and contributes to pathogenicity (36, 68, 75, 130).

Two major branches exist within the *Borrelia* genus and these are defined by the diseases that the bacteria cause. One branch encompasses the Lyme disease-causing species, including *B. afzelii, B. americana, B. andersonii, B. bavariensis, B. bissettii, B. burgdorferi, B. californiensis, B. carolinensis, B. garinii, B. japonica, B. kurtenbachii, B. lusitaniae, B. sinica, B. tanukii, B. turdi, B. spielmanii, B. valaisiana, and B. yangtze* (112). The other branch contains the relapsing fever spirochetes, *B. hermsii, B. turicatae, B. duttonii, B. dugesi, B. crocidurae, B. parkeri, and B. recurrentis* (41, 47). All of the listed species are tick-borne pathogens with the exception of *B. recurrentis*, which is spread by the human body louse, *Pediculus humanus humanus* (47). The work
**Figure 1-2. Spirochete motility controlled by flagellar rotational direction.** The “run,” “flex,” and “reverse” motility of the spirochete is dependent on the rotational direction of the polar flagellar bundles. The top image shows the “run” phenotype where each pole is rotating in opposite directions. This allows translational motion to the right. The image below that is with the asymmetrical rotation of the flagella in the opposite direction of the “run” which results in a “reverse” and translation motion to the left. The bottom two images show the bundles rotating in the same directions, which presents as a “flex.”

contained within this thesis is concentrated on two species of *Borrelia* to represent the two branches within the genus, *B. burgdorferi* as a part of the Lyme disease branch and *B. hermsii* of the tick-borne relapsing fever branch. Both of these species are the most common cause of their respective diseases in North America.

**Lyme disease**

Lyme borreliosis is an emerging infectious disease that first made its debut in the mid-1970s near Lyme, CT (128). The disease was associated with severe arthritis and an expanding “bull’s eye” rash, previously described by Arvid Afzelius in 1909 (2). He observed it in relation to the bite of a hard-bodied tick and named this lesion the ‘erythema migrans’ rash. In addition to this dermatological manifestation, patients of the Lyme epidemic developed neurological, articular, and cardiac abnormalities (51, 85, 104, 127). These disease symptoms were associated with the bite of the *Ixodes scapularis* tick and, later, with the presence of a *B. burgdorferi* infection (27). Soon after, *B. burgdorferi* was isolated from the reservoir species, *Peromyscus leucupus* and the tick-mouse transmission confirmed (17, 24, 29). Since then, Lyme disease has been diagnosed in North America, Europe, and Asia with three major species causing the disease. Two of these (*B. garinii* and *B. afzelii*) are only present in Europe and Asia and spread by the *Ixodes ricinus* and *Ixodes persulcatus* ticks. In North America, *B. burgdorferi* is the primary causative agent and is spread by *I. scapularis* and *Ixodes pacificus* (28, 30, 64, 90).

In the United States, Lyme disease is the most common vector-borne disease and the sixth most common nationally notifiable disease (1, 34). In 2011, there were
33,097 confirmed and probable cases according to the CDC. This makes the national incidence rate 7.8 per 100,000 population (34). These statistics are considered to be under-representative of actual infection numbers due to severe under-reporting and misdiagnosis (33, 145). Additionally, 96% of the cases occurred in specific geographical regions, the Northeast, mid-Atlantic, and upper Midwest; this is shown in Figure 1-3 (34). The reported incidence of individual states per 100,000 was as high as 84.6 (Delaware) and was 28.7 between the 14 states that are considered to be part of the endemic regions: Connecticut, Delaware, Maine, Maryland, Massachusetts, Minnesota, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, Vermont, Virginia, and Wisconsin (individual state statistics can be found in Table 1-1) (34).

Lyme disease has a wide range of signs, symptoms, and manifestations. However, subclinical infections have been shown to occur through seroprevalence studies done in endemic regions (49). It is important to note that the human immune response to early infection by one strain of Lyme Borrelia is non-protective against other strains, resulting in reinfection (59, 86, 88, 122). The “hallmark” sign of early invasion is the erythema migrans rash, but only occurs in 60-85% of patients (85, 124). Other symptoms of acute infection can include fatigue, myalgia, arthralgia, headache, fever, and chills (85). However, an early diagnosis and subsequent treatment with oral antibiotics such as doxycycline or amoxicillin has proven very effective in the treatment of Lyme disease (125).

Due to non-specific signs of early infection, Lyme borreliosis is often allowed to progress untreated for long periods, which can lead to more serious sequelae. Later manifestations can include Lyme arthritis, neurological conditions such as
Figure 1-3. 2011 case distribution of Lyme disease in the United States. A map of the location of confirmed cases of Lyme disease in the United States that occurred in 2011. Each dot represents a single case randomly placed in the county of residence. The endemic region of Lyme disease primarily lies in the Mid-Atlantic, Northeast, and upper Midwest.

Table 1-1. Case numbers and incidence rates in the United States per state.

<table>
<thead>
<tr>
<th>State</th>
<th>Confirmed</th>
<th>Probable</th>
<th>Incidence*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alabama</td>
<td>9</td>
<td>15</td>
<td>0.2</td>
</tr>
<tr>
<td>Alaska</td>
<td>9</td>
<td>2</td>
<td>1.2</td>
</tr>
<tr>
<td>Arizona</td>
<td>8</td>
<td>7</td>
<td>0.1</td>
</tr>
<tr>
<td>Arkansas</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>California</td>
<td>79</td>
<td>13</td>
<td>0.2</td>
</tr>
<tr>
<td>Colorado</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Connecticut</td>
<td>2004</td>
<td>1035</td>
<td>56</td>
</tr>
<tr>
<td>Delaware</td>
<td>787</td>
<td>106</td>
<td>84.6</td>
</tr>
<tr>
<td>DC</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>Florida</td>
<td>78</td>
<td>37</td>
<td>0.4</td>
</tr>
<tr>
<td>Georgia</td>
<td>32</td>
<td>0</td>
<td>0.3</td>
</tr>
<tr>
<td>Hawaii</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Idaho</td>
<td>3</td>
<td>1</td>
<td>0.2</td>
</tr>
<tr>
<td>Illinois</td>
<td>194</td>
<td>0</td>
<td>1.5</td>
</tr>
<tr>
<td>Indiana</td>
<td>81</td>
<td>13</td>
<td>1.2</td>
</tr>
<tr>
<td>Iowa</td>
<td>72</td>
<td>28</td>
<td>2.4</td>
</tr>
<tr>
<td>Kansas</td>
<td>11</td>
<td>6</td>
<td>0.4</td>
</tr>
<tr>
<td>Kentucky</td>
<td>3</td>
<td>0</td>
<td>0.1</td>
</tr>
<tr>
<td>Louisiana</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Maine</td>
<td>801</td>
<td>205</td>
<td>60.3</td>
</tr>
<tr>
<td>Maryland</td>
<td>938</td>
<td>413</td>
<td>16.1</td>
</tr>
<tr>
<td>Massachusetts</td>
<td>1801</td>
<td>675</td>
<td>27.3</td>
</tr>
<tr>
<td>Michigan</td>
<td>89</td>
<td>15</td>
<td>0.9</td>
</tr>
<tr>
<td>Minnesota</td>
<td>1185</td>
<td>939</td>
<td>22.2</td>
</tr>
<tr>
<td>Mississippi</td>
<td>3</td>
<td>2</td>
<td>0.1</td>
</tr>
<tr>
<td>Missouri</td>
<td>5</td>
<td>3</td>
<td>0.1</td>
</tr>
<tr>
<td>Montana</td>
<td>9</td>
<td>2</td>
<td>0.9</td>
</tr>
<tr>
<td>Nebraska</td>
<td>7</td>
<td>4</td>
<td>0.4</td>
</tr>
<tr>
<td>Nevada</td>
<td>3</td>
<td>2</td>
<td>0.1</td>
</tr>
<tr>
<td>New Hampshire</td>
<td>887</td>
<td>412</td>
<td>67.3</td>
</tr>
<tr>
<td>New Jersey</td>
<td>3398</td>
<td>864</td>
<td>38.5</td>
</tr>
<tr>
<td>New Mexico</td>
<td>2</td>
<td>4</td>
<td>0.1</td>
</tr>
<tr>
<td>New York</td>
<td>3118</td>
<td>1372</td>
<td>16</td>
</tr>
<tr>
<td>North Carolina</td>
<td>18</td>
<td>70</td>
<td>0.2</td>
</tr>
<tr>
<td>North Dakota</td>
<td>22</td>
<td>5</td>
<td>3.2</td>
</tr>
<tr>
<td>Ohio</td>
<td>36</td>
<td>17</td>
<td>0.3</td>
</tr>
<tr>
<td>Oklahoma</td>
<td>2</td>
<td>0</td>
<td>0.1</td>
</tr>
<tr>
<td>Oregon</td>
<td>9</td>
<td>29</td>
<td>0.2</td>
</tr>
<tr>
<td>Pennsylvania</td>
<td>4739</td>
<td>623</td>
<td>37.2</td>
</tr>
<tr>
<td>Rhode Island</td>
<td>111</td>
<td>48</td>
<td>10.8</td>
</tr>
<tr>
<td>South Carolina</td>
<td>24</td>
<td>13</td>
<td>0.5</td>
</tr>
<tr>
<td>South Dakota</td>
<td>2</td>
<td>2</td>
<td>0.2</td>
</tr>
<tr>
<td>Tennessee</td>
<td>5</td>
<td>32</td>
<td>0.1</td>
</tr>
<tr>
<td>Texas</td>
<td>28</td>
<td>46</td>
<td>0.1</td>
</tr>
<tr>
<td>Utah</td>
<td>6</td>
<td>3</td>
<td>0.2</td>
</tr>
<tr>
<td>Vermont</td>
<td>476</td>
<td>147</td>
<td>76</td>
</tr>
<tr>
<td>Virginia</td>
<td>756</td>
<td>267</td>
<td>9.3</td>
</tr>
<tr>
<td>Washington</td>
<td>17</td>
<td>2</td>
<td>0.2</td>
</tr>
<tr>
<td>West Virginia</td>
<td>107</td>
<td>11</td>
<td>5.8</td>
</tr>
<tr>
<td>Wisconsin</td>
<td>2408</td>
<td>1241</td>
<td>42.2</td>
</tr>
<tr>
<td>Wyoming</td>
<td>1</td>
<td>1</td>
<td>0.2</td>
</tr>
</tbody>
</table>

| U.S. TOTAL     | 24,364    | 8,733    | 7.8        |

*per 100,000 population

Table adapted from the 2012 CDC report
(http://www.cdc.gov/lyme/stats/chartstables/reportedcases_statelocality.html)
neuroborreliosis, cutaneous disorders, and cardiac abnormalities such as endocarditis (7, 51, 60, 129, 139). Beyond the potential for debilitating medical conditions, the financial burden of Lyme disease in the United States is severe. A study was done calculating the approximate cost per case; this came to approximately $8200 in a 2006 study (146). Despite the substantial financial burden that Lyme disease imposes, prevention is restricted to avoidance of exposure to high-risk environments and early recognition of acute symptoms. As the causative agent of Lyme disease in the United States, *B. burgdorferi* is a prominent pathogen with fascinating survival mechanisms in hosts spanning multiple phyla.

*B. burgdorferi* is a bacterial species that must survive by efficiently transmitting between two highly dissimilar host environments in a process known as the enzootic cycle. As a tick-borne pathogen, the bacteria are spread by the deer tick (*I. scapularis*), which has a two year life cycle depicted in Figure 1-4 (102). Eggs are laid in early spring and hatch in the summer (102). The larval tick that emerges is uninfected since *B. burgdorferi* are incapable of transovarial transmission (102). In late summer and early fall, the larvae feed on one of a plethora of mammalian, avian, and reptilian hosts (102). Transmission of *B. burgdorferi* to the tick vector often occurs at this stage by feeding on an infected member of the small mammal reservoir, commonly the white-footed mouse (102). The larvae then molt into the nymphal stage and lay dormant through the winter, through this time the bacteria remain in the mid-gut (102). These nymphal ticks take their second blood-meal in the spring or summer (102). It is this blood meal that causes the spirochetes to migrate into the salivary glands and effectively transmit to this secondary host, resulting in infection (73, 126). In the fall, nymphs moult into adults.
Figure 1-4. Enzootic cycle of the Lyme disease spirochete. The *B. burgdorferi* vector is primarily the *I. scapularis* tick, which has a complex, multi-stage life cycle. The cycle begins with eggs hatching in the summer. The larvae ticks that emerge are free of *Borrelia*. In the summer and fall, the ticks feed on their first host, which range widely in variety and include squirrels, mice, and birds. It is from this blood-meal that the ticks often obtain the bacteria. The fed tick drops off the host and remains dormant through the winter, moulting into the nymph life stage. The tick then feeds on a second host in the spring or summer and this is where transmission of the Lyme spirochete occurs. Infection of small mammal and avian hosts results in the continuity of the enzootic cycle. Fed nymphs drop off and moults into the adult stage. Adults feed and procreate on large mammals such as deer in the fall and winter, after which the adult females lay eggs in the spring. These eggs are uninfected since *B. burgdorferi* cannot transmit to the ova of the tick. It is significant to note that humans and canines are able to be infected, but they are dead-end hosts and do not contribute to the perpetuation of the bacteria in the environment. A hallmark symptom of human infection by *B. burgdorferi* is a “bulls-eye” or erythema migrans rash.

It is in this stage that the ticks feed and copulate on large mammals such as deer through the fall and spring (102). Deer are incompetent hosts for *B. burgdorferi*, but facilitate the spread of the disease by traversing great distances and expanding the range of potential hosts (102). In the spring, the enzootic cycle ends with the adult female laying eggs and dying (102). It is significant to note that while humans and dogs are capable of being infected, they are usually dead-end hosts and do not contribute to the perpetuation of the bacteria in the environment.

Because *B. burgdorferi* must survive within two disparate host environments, drastic transcriptional and translational changes must occur to effectively transmit and avoid each host’s defenses (50, 121). Despite the necessity for rapid, global transcriptional changes, the *B. burgdorferi* genome contains very few transcriptional regulators. For example, only three sigma factors (RpoD, RpoN, and RpoS) have been identified in the *B. burgdorferi* genome (52). RpoD (σ70) is the general, house-keeping sigma factor. RpoN (σ54) has been shown to transcribe multiple genes including RpoS, the other alternate sigma factor (26, 66). RpoS (σ38/σ8) aids in transcription of host-specific factors such as OspC, an outer surface protein only expressed in early mammalian infection (66). Synthesis of RpoS has multiple regulatory elements. Rrp2 (a two-component response regulator) is a σ54 enhancer-binding protein that requires phosphorylation to activate the transcription of *rpoS* (144). BosR is a DNA binding protein shown to directly interact with the *rpoS* promoter and allow transcription (91).

Along with transcriptional regulation, RpoS synthesis is also temperature-regulated at the translational level by a small RNA, DsrA, which is stabilized by the RNA chaperone Hfq (76, 77).
Regulatory elements are often suppressed or induced by environmental stimuli. A method that bacteria use to detect those stimuli is two-component systems. Most bacteria have many two-component regulatory systems, such as *Escherichia coli* with 30 to 60 (52). However, *B. burgdorferi* encodes five, only two of which are capable of global regulation (84, 144). These are Rrp1-Hpk1 and Rrp2-Hpk2 (107, 144). The Rrp2-Hpk2 system has been implicated in *rpoS* transcriptional regulation that initiates survival mechanisms specific to the mammalian host. Response regulators often contain DNA-binding domains, including Rrp2 (144). However, Rrp1 lacks a DNA-binding domain, but instead has a GGDEF domain and will be expounded upon later in this thesis (107, 116). As in other two-component systems, activation of the response regulator, Rrp1, requires phosphorylation by the cytoplasmic membrane-embedded histidine kinase, Hpk1 (32, 62, 116). In this manner, the sensory protein detects an external stimulus to initiate downstream functions and regulate cell activity (142). The scarcity of broad-spectrum regulatory pathways correlates with the overall minimization of the *B. burgdorferi* genome.

Small, fragmented, and condensed, the *B. burgdorferi* genome is comprised of a 0.91 Mb linear chromosome as well as a cohort of linear and circular plasmids that varies from strain to strain; the type strain B31 has 12 linear and 9 circular plasmids (105). The collective coding sequences of the chromosome occupy 93% of the genetic material with a great deal of gene overlap (52). This compacted chromosome expresses most of the housekeeping genes while others, many vital to bacterial survival in hosts, are expressed off of the bacterium’s arsenal of plasmids (e.g. OspC from cp26, 102). Despite its genetic dwarfism, *B. burgdorferi* has retained 54 chemotactic/motility genes;
this correlates with approximately 6% of the genome (52). This large percentage and the fact that non-motile B. burgdorferi are non-infectious exemplify the importance of motility for B. burgdorferi survival (74).

**Tick-borne relapsing fever**

Unlike the more recently discovered Lyme disease, relapsing fever has far more ancient roots in human history. Since the days of Hippocrates, a disease was described similar to what is known as relapsing fever today (41). Louse-borne relapsing fever (caused by B. recurrentis) is likely correlated with the yellow plague of Europe in 550 A.D. and was responsible for millions of cases during World War I and II (111). Tick-borne relapsing fever (TBRF) occurs throughout the world with the highest incidence rates in Africa (47). Unfortunately, most African countries do not have a reporting system for TBRF, so the actual infection rates of this disease are difficult to discern (78). However, infection studies in smaller populations have been conducted and provide insight into the severity of impact that TBRF can have.

The soft-bodied *Ornithodoros* tick is the vector for TBRF. Unlike the Lyme disease species, the TBRF spirochetes can use their vector as a reservoir species due to transovarial transmission and long-term survival (47, 123). *Ornithodoros* ticks live in protected niches of caves, mud-huts, trees, and cabins where small mammals frequent (118). Unlike the *Ixodes*, *Ornithodoros* ticks can live for many years and adult females reproduce after every blood-meal of that lifespan (8, 47, 79). Because of the life cycle and span of its tick vector, there are high percentages of TBRF-infected ticks in the environment (8, 41, 118).
Like Lyme disease, the endemic regions of TBRF are highly specific to the habitat and distribution of the vector species (93). For example, in the United States, 50% of TBRF cases have occurred in only 13 counties and this distribution is shown in Figure 1-5 (47). This endemic region correlates with the presence of the most common cause, \textit{B. hermsii} and its vector, \textit{O. hermsi} (45).

Another factor in TBRF epidemiology is the living conditions of the human population. For example, many African communities live in mud-huts, which are an ideal location for ticks to reside and feed on the human inhabitants at night (43, 92). Studies in Senegal revealed TBRF as the most common bacterial infection and the Tanzanian town of Mvumi reported TBRF to be the 7\textsuperscript{th} highest cause of death in children with a perinatal mortality rate of 436 per 1000 cases (42, 78, 136, 140). While cases in the United States are generally unsuspecting cabin vacationers, TBRF is a global health concern with severe medical outcomes.

Disease progression of TBRF begins with exposure to the bacteria through a tick’s blood-meal that can last from a fifteen minutes to two hours (47, 79, 92). These blood-meals are brief, nocturnal, and painless, primarily occurring while the human host is asleep (47, 140). The transmission of the bacteria from the tick vector to the mammalian host is rapid and may take as little as 30 seconds, while \textit{B. burgdorferi} transmission requires 24 to 48 hours (47). After infection, there is an incubation period that lasts for several days before disease signs and symptoms begin (47). These are variable and can include headache, myalgia, nausea, arthralgia, and/or abdominal complaints (47). Some less transient manifestations include neuroborreliosis, splenomegaly, and myocarditis (4, 31, 141). However, the most common and name-
Figure 1-5. Tick-borne relapsing fever incidence in the United States. Above is the distribution of TBRF cases in North America. Highlighted are the counties in which TBRF patients had been diagnosed up to 2002.
sake symptom is a recurring, severe fever, often near or surpassing 104°F (45, 123). These febrile periods correlate with high levels of blood spirochetemia (47, 92). The first febrile period lasts an average of three days with approximately a week between the end of the initial febrile period and the start of the first relapse (47). As many as 13 relapses have been documented with spirochetemia reaching from 10^6 to 10^8 bacteria per mL of blood (42, 45). This clinical outcome is due to afebrile period sequestration of the bacteria in the internal organs and then reemergence resulting in the fever symptom (93). The molecular mechanism by which \textit{B. hermsii} causes this recurring fever is through antigenic variation (14).

\textit{B. hermsii} has been shown to have as many as 40 different serotypes that can be generated from a single spirochete (105). This is possible through several mechanisms of genetic recombination that result in one of many silent variable major protein (\textit{vmp}) genes translocating to the downstream region of the appropriate promoter, allowing transcription (12, 106). It is with the expression of a new Vmp that the dominant serotype of the \textit{B. hermsii} population changes, leading to immune evasion and increased spirochetemia (105).

Antigenic variation and other elaborate mammalian host-specific immune evasion mechanisms are often associated with efficient signal transduction systems to alter the \textit{B. hermsii} transcriptome upon host invasion. While this species is not quite as dependent on its mammalian host for survival as the Lyme disease spirochetes, \textit{B. hermsii} regularly infects small mammals such as tree squirrels, chipmunks, and ground squirrels (46). Therefore, this species must also harbor mechanisms of global transcriptional regulation specific to the tick-mammal and mammal-tick transmission.
Cyclic di-GMP

*Borrelia* undergo rapid, drastic changes in gene expression and biochemical processes that require efficient signal transduction systems. Bis-(3′-5′)-cyclic dimeric guanosine monophosphate (c-di-GMP) is a recently discovered bacterial secondary signaling molecule originally shown to regulate cellulose synthesis in *Gluconacetobacter xylinus* (110). Since then, it has been implicated in the regulation of many pathways among a plethora of bacterial species (5, 9, 67, 116).

All second messenger molecules function to relay signals, altering downstream cellular processes. These messengers are often small, making them energy efficient and capable of rapid diffusion throughout the cytoplasm. Structurally, c-di-GMP comprises two GMP molecules connected by two phosphodiester linkages and is depicted in Figure 1-6. These linkages result in a rigid, twelve-member ring, but flexibility in the glycosidic bonds between the ribose sugars and guanine bases allows for c-di-GMP to exist in two primary conformations (53, 117). The open conformation is depicted in Figure 1-7a, and the closed in Figure 1-7b. This difference in the glycosyl torsion angle is seen in other cyclic di-nucleotides, such as cyclic dimeric adenosine monophosphate (53). It is also significant to note that two monomers of c-di-GMP in the closed conformation are capable of associating due to base pair stacking. This results in an intercalated dimer, shown in Figure 1-7b (20, 48). C-di-GMP is capable of higher orders of oligomerization although it primarily exists in the monomer or dimer form at physiologically relevant concentrations (48, 117).
Figure 1-6. Cyclic di-GMP structure. Two GTP molecules are circularized by a diguanylate cyclase to form c-di-GMP. Two phosphodiester bonds generate a twelve member ring out of the ribose sugars and phosphate groups.
Figure 1-7. Cyclic di-GMP conformations allow for multimerization. The flexibility of the glycosidic bond between the ribose sugar and the guanine base allows for two primary conformations. The open conformation is shown in panel A while the closed conformation can be seen in panel B. In the closed conformation, two c-di-GMP molecules can also intercalate through base pair stacking of the guanines as well as hydrogen bond formation between the phosphate group of one molecule and the guanine base of the other.
Synthesis of c-di-GMP from two guanosine triphosphate (GTP) molecules is carried out by a group of proteins known as diguanylate cyclases (DGCs), which are typified by a GGDEF protein domain (9, 35, 63, 96, 135). It is at a conserved loop-forming motif (formed by the GGDEF sequence) where binding of the two GTPs occurs at a dimeric interface between two domains (95, 97). The GGDEF motif associates with one molecule of GTP by interacting with the phosphates as well as the guanine base and contributes to the high degree of binding specificity of these proteins (35). Breakdown is carried out by phosphodiesterases (PDEs), of which there are two groups. Some contain an EAL motif and these hydrolyze only one of the phosphodiester bonds of c-di-GMP, forming a pGpG molecule (21). Others contain an HD-GYP domain and catabolize c-di-GMP into its two GMP constituents (38, 57, 58, 114). It is through the activity of DGCs and PDEs that c-di-GMP levels and downstream effects are modulated within the cell.

C-di-GMP signaling regulates a variety of cellular functions among many bacterial species. As a matter of fact, c-di-GMP-metabolizing proteins have been discovered in most major phyla (109). The first discovered function of c-di-GMP was activation of exopolysaccharide biosynthesis in the form of cellulose production in *G. xylinus* (110). This has since expanded to include others, such as alginate production in *Pseudomonas aeruginosa* (81, 82). Other common regulatory functions of c-di-GMP include cell cycle progression, virulence factor production, and long-term persistence (23, 72, 80, 96, 138). However, perhaps the most defined functions of c-di-GMP are biofilm formation and motility (82). As a bacterial process known to confer properties that increase pathogenicity, biofilm formation has been a major area of focus in c-di-
GMP research (108). Among multiple species, including prominent pathogens such as *Yersinia pestis*, *Vibrio cholerae*, *Salmonella enterica*, and *P. aeruginosa*, increased levels of c-di-GMP induce biofilm formation while decreased levels result in an increase in planktonic cells and motility (22, 23, 40, 120, 137). These phenotypes are exhibited through c-di-GMP-induced regulation of gene expression and protein activity (37, 115, 143).

Cellular functions modulated by c-di-GMP can be altered at multiple stages within the biochemical pathways, including transcription, translation, and post-translational modifications (109). The transcription of messenger ribonucleic acid (mRNA) can be regulated by binding of transcription factors, enhancers, or suppressors (109). In this way, gene expression may be upregulated or downregulated depending on intracellular c-di-GMP levels (109). Once the target genes are transcribed, translational regulation can be carried out by c-di-GMP-activated proteins or riboswitches, which are secondary structures within mRNA that can directly bind c-di-GMP to promote or inhibit translation (109). Lastly, c-di-GMP can regulate protein levels and functionality through proteolysis, activation, or deactivation (44).

C-di-GMP-mediated regulation is highly reliant on the receptors/effectors to which this second messenger binds. Known receptors come in many forms and carry out an array of functions. Two types of riboswitches have been discovered to bind c-di-GMP and are present in many species, often with multiple incidences within a single genome (109). Other receptors can include degenerative GGDEF, EAL, or HD-GYP domain proteins that have lost their catalytic activity, but are still capable of binding c-di-GMP (109). The last group of common receptors contains a highly conserved domain
known as the PilZ domain (6). This domain confers c-di-GMP binding with a moderate to high affinity (6). Other receptors are less predictable and therefore difficult to find, but are gradually being discovered.

The PilZ domain was first discovered by bioinformatics analyses and postulated to bind c-di-GMP, which was quickly confirmed experimentally and named after the *P. aeruginosa* PilZ protein (3, 6, 115). These studies and more have shown that several residues are highly conserved and these form an RxxxR/DzSxxG motif where x represents any amino acid and z represents a hydrophobic amino acid. Crystallography and nuclear magnetic resonance studies have shown that these residues do, indeed, interact with c-di-GMP and that it can be in the monomeric or intercalated dimer conformation (15, 61, 69, 103, 119). While PilZ domain-containing proteins have been implicated in the regulation of many cellular processes, the best characterized have been shown to regulate motility.

PilZ protein-mediated regulation of motility can act through methods such as expression of motility genes or protein-protein interactions. An example of PilZ protein-regulated gene expression is *Klebsiella pneumoniae*’s MrkH (143). This transcription factor binds the promoter of a type III fimbriae gene cluster to facilitate biofilm formation in the presence of c-di-GMP (143). More commonly, however, PilZ proteins have been shown to interact with other proteins to alter function. The *S. enterica* serovar Typhimurium PilZ protein, YcgR, is the most thoroughly described c-di-GMP binding protein (115). It interacts with c-di-GMP to associate with the flagellar motor complex and inhibit motility (94, 115). PilZ domain proteins are an important part of a bacterium’s c-di-GMP regulatory network. The fact that almost every genome with a diguanylate
cyclase, including both *B. burgdorferi* and *B. hermsii*, also contains at least one PilZ domain protein lends evidence to that fact (109).

**C-di-GMP signaling in the *Borrelia***

GGDEF, EAL, HD-GYP, and PilZ domain containing proteins have been found in the genomes of most spirochete species, indicating the presence of c-di-GMP signaling systems (109). However, the surface of c-di-GMP functionality within the Spirochaete phylum has barely been scratched. *B. burgdorferi* and *B. hermsii* are particularly unique in that they each harbor only one diguanylate cyclase and two phosphodiesterases. Other species, such as *P. aeruginosa* have as many as 41 c-di-GMP metabolizing proteins (109). The simplicity and genetic malleability of these bacteria make them appealing systems in which to study c-di-GMP signaling.

The *B. burgdorferi* c-di-GMP network is slowly being unraveled. However, the *B. hermsii* system has largely been left untouched. A sequence comparison of the *B. burgdorferi* B31 and the *B. hermsii* DAH c-di-GMP-associated proteins reveals the high degree of homology between these two systems and can be found in Figure 1-8. The sequence identities of these proteins range from 65 to 76% and the sequence similarities from 82 to 89%. The BB0419 and BH0419 are the sole diguanylate cyclases in the respective species (Rrp1, 71, 107, 133). BB0363 and BH0363 refer to the EAL domain containing phosphodiesterases (PdeA, 107, 132). BB0374 and BH0374 are the HD-GYP phosphodiesterases (PdeB, 107, 131). BB0733 is the only known PilZ domain containing protein of *B. burgdorferi* (PlzA) and BH0733 is the *B. hermsii* counterpart (PlzC, 55, 98). Deletion mutants have been generated in each of the *B. burgdorferi*
Figure 1-8. Sequence homology between the *B. burgdorferi* and *B. hermsii* c-di-GMP metabolizing proteins. Each sequence alignment was generated in ClustalW. The asterisk (*) labels identical residues between the *B. burgdorferi* and *B. hermsii* species. The colon (:) labels residues with charge similarity. The period (.) labels residues with structural similarity. The BB0419 and BH0419 had 76% identity and 89% similarity. The GGDEF is highlighted. BB0363 and BH0363 had 69% identity and 84% similarity. The EAL domain is highlighted. BB0374 and BH0374 had 65% identity and 84% similarity. The HD-GYP domain is highlighted. BB0733 and BH0733 had 69% identity and 82% similarity. The PilZ domain motifs are highlighted.
genes listed above and their phenotypes described (55, 71, 98, 107, 131, 132, 133). A schematic of the *B. burgdorferi* c-di-GMP binding system can be found in Figure 1-9.

Interference in the c-di-GMP signaling system of the *Borrelia* results in multiple phenotypes. Deletion of the sole diguanylate cyclase gene, *rrp1*, resulted in an inability to colonize ticks (62, 71, 133). A deletion mutant of *hpk1* had a very similar phenotype because the diguanylate cyclase activity of Rrp1 is dependent on phosphorylation by Hpk1 (32, 62, 116). Alternatively, deletion of one of the phosphodiesterase genes (*pdeA*) abolished mouse infectivity, but retained tick colonization comparable to wild-type (132). However, deletion of the other phosphodiesterase (*pdeB*) resulted in normal murine infectivity and reduced, but not abolished tick survival (131). Lastly, a deletion of the only c-di-GMP binding molecule found in *B. burgdorferi*, *plzA*, led to reduced invasion of both the mouse and tick host (98). Most of these mutants also had alterations in motility and chemotaxis (71, 98, 131, 132, 133). The only consistent phenotype in all mutants was an inability to transfer between the mammalian host and tick vector to complete the enzootic cycle, which is paramount for the perpetuation of *B. burgdorferi* in the environment.

**Research objectives**

PlzA is the only confirmed c-di-GMP binding molecule produced by *B. burgdorferi* (55). As such, it is the only possible known c-di-GMP effector. Therefore, characterization of PlzA activity and its interaction with c-di-GMP could elucidate the molecular mechanisms of c-di-GMP regulation in *B. burgdorferi*. 
The B. burgdorferi c-di-GMP network has been characterized as to enzymatic activity if not downstream functionality. Hpk1 is embedded in the cytoplasmic membrane and theoretically detects an external stimuli to induce autophosphorylation of histidine residues. This, in turn, activates the diguanylate cyclase Rrp1 to synthesize c-di-GMP from two GTP molecules. Degredation of c-di-GMP is carried out by PdeA through linearization of the secondary messenger into the pGpG form or by PdeB which breaks it down into two GMP molecules. The last B. burgdorferi c-di-GMP interacting protein is PlzA, which contains a functional PilZ domain that confers c-di-GMP binding.

Figure 1-9. Schematic of the c-di-GMP network of B. burgdorferi. The B. burgdorferi c-di-GMP network has been characterized as to enzymatic activity if not downstream functionality. Hpk1 is embedded in the cytoplasmic membrane and theoretically detects an external stimuli to induce autophosphorylation of histidine residues. This, in turn, activates the diguanylate cyclase Rrp1 to synthesize c-di-GMP from two GTP molecules. Degredation of c-di-GMP is carried out by PdeA through linearization of the secondary messenger into the pGpG form or by PdeB which breaks it down into two GMP molecules. The last B. burgdorferi c-di-GMP interacting protein is PlzA, which contains a functional PilZ domain that confers c-di-GMP binding.
Due to its sequence homology to PlzA, the study of the *B. hermsii* PilZ domain protein, PlzC, would be very appealing as one of the first steps in delineating the c-di-GMP regulatory network of the TBRF spirochete. While nothing has been published on c-di-GMP signaling in the relapsing fever bacterium, preliminary data from our lab implies that c-di-GMP signaling has an effect on the motility, chemotaxis, and disease of *B. hermsii* (54). Unfortunately, altered phenotypes in enzootic cycle progression cannot be assessed at this time since there are no uninfected *O. hermsi* ticks commercially available.

Through biochemical and biophysical methods such as amino acid substitution, fluorescence resonance energy transfer, isothermal titration calorimetry, and size exclusion chromatography, our lab has shed light on the c-di-GMP interaction with the only PilZ domain containing proteins of *B. burgdorferi* and *B. hermsii*, PlzA and PlzC respectively.
Chapter 2: Materials and Methods

*Borrelia* strains and cultivation

*B. burgdorferi* B31-5A4 cells were cultivated in Barbour-Stoenner-Kelly (BSK-H) media supplemented with 6% rabbit serum (10). *B. hermsii* DAH cells were cultivated in BSK-H media supplemented with 12% rabbit serum. All strains were maintained under humidified conditions with 5% CO$_2$ and at 25, 33, or 37°C. These isolates are listed in Table 2-1.

Plasmid construction

Strains and plasmids used in this study are listed in Table 2-1. All fluorescent resonance energy transfer (FRET) expression plasmids were derived from pFRET12AA plasmid (89). The construct constitutes a flanking fluorescence resonance energy transfer (FRET) pair (YFP and CFP) on either side of a multiple cloning site and encodes twelve amino acids. The *mYPet* (YFP) and *mCyPet* (CFP) genes contain mutations that prevent oligomerization that could interfere with signal. A positive control FRET construct, MC0118 (kindly provided by Dr. Samuel Miller) contains the *Salmonella enterica* serovar Typhimurium gene encoding the PilZ domain protein YcgR and exhibits a characterized conformational change upon binding of c-di-GMP using FRET [Christen, 2010 #2452]. A SpeI site was introduced at the 5’ end and a KpnI site
Table 2-1. Strains and plasmids used in this study.

<table>
<thead>
<tr>
<th>Strain or Plasmid</th>
<th>Description</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>E. coli NovaBlue</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E. coli BL21 (DE3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B. hermsii DAH</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B. burgdorferi B31</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pFRET12aa</td>
<td>pET15b::mYPet_12AA_mCyPet</td>
<td>(89)</td>
</tr>
<tr>
<td>MC0118</td>
<td>pET15b::mYPet_ycgR_mCyPet</td>
<td>(39)</td>
</tr>
<tr>
<td>pFRETplzC</td>
<td>pET15b::mYPet_plzC_mCyPet</td>
<td>this work</td>
</tr>
<tr>
<td>pFRETplzA</td>
<td>pET15b::mYPet_plzA_mCyPet</td>
<td>this work</td>
</tr>
<tr>
<td>pFRETplzAR145D</td>
<td>pET15b::mYPet_plzAR145D_mCyPet</td>
<td>this work</td>
</tr>
<tr>
<td>pFRETplzAR145K</td>
<td>pET15b::mYPet_plzAR145K_mCyPet</td>
<td>this work</td>
</tr>
</tbody>
</table>
at the 3' end of all PCR products. *plzA* (BB0733) was amplified from genomic DNA of the *B. burgdorferi* strain B31. *plzC* (BH0733) was amplified from the DAH strain of *B. hermsii*. Using splice overlap exchange techniques, the *plzA* arginine residue at position 145 was substituted with an aspartic acid (AGA to GAT) and a lysine (AGA to AAA). All primers used in this study can be found in Table 2-2. Upon sequence confirmation of each construct, the plasmids were transformed into BL21 (DE3) cells. Schematics of the pFRETplzA plasmid and the pFRETplzC plasmid can be found in Figure 2-1.

For the purpose of His-tagging PlzA, PlzAR145D, PlzAR145K, and PlzC, *plzA* was PCR amplified from B31 gDNA and *plzC* from DAH gDNA. *plzAR145D* and *plzAR145K* were amplified from their respective FRET constructs using the *plzA* specific primers. PCR products were then inserted into pET45b(+, Novagen) using restriction sites indicated by the primer name in Table 2-2. Plasmids were sequence verified and transformed into BL21 (DE3) cells for protein expression.

**Protein expression and purification.**

All proteins in the study were expressed in Lysogeny Broth supplemented with 100 µg/mL ampicillin. Cultures of the transformed BL21 (DE3) cells were grown to log phase at 37°C then chilled on ice for 15 min. Protein expression was induced with isopropyl β-D-1-thiogalactopyranoside (IPTG). FRET proteins were induced with IPTG at a 20 µM concentration for 24 hours at 18°C. His-tagged proteins were induced with IPTG at a concentration of 200 µM for 18 hours at room temperature. Cells were recovered by centrifugation then lysed with lysozyme (1 mg/mL) and sonication.
<table>
<thead>
<tr>
<th>Primer name</th>
<th>Primer sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>PlzA Spe I 5' F</td>
<td>ATAGTGACTAGTATGCTTTTTATCTAGAAAAATAAGAGATTATGGGGCTAAGTGATAGG</td>
</tr>
<tr>
<td>PlzA Kpn I 3' R</td>
<td>ATAGTTGGTACATTGAAGATAATCATGGATCAACATAGTATCATCAAGTGTTATTTTATC</td>
</tr>
<tr>
<td>PlzA R145D SOE 5' F</td>
<td>GGGCAAAATCAGGAGTATTCTAGGAGGGATTATTATCAGTAAAGATTCTATTAGAAAGC</td>
</tr>
<tr>
<td>PlzA R145D SOE 3' R</td>
<td>CCTCTCATGAATATAGCTTGTAGCTGCTTTGCCCCAAGGTTTAAATCAAGAGAGC</td>
</tr>
<tr>
<td>PlzA R145K SOE 5' F</td>
<td>GGGCAAAATCAGAAAAATTCTAGGAGGGATTATTATCAGTAAAGATTCTATTAGAAAGC</td>
</tr>
<tr>
<td>PlzA R145K SOE 3' R</td>
<td>CCTCTCATGAATTTTTGCTTGGTGGCCGAAAGGTTTAAATCAAGAGAGC</td>
</tr>
<tr>
<td>PlzC Spe I 5' F</td>
<td>ATAGTGACTAGTATGCTTTTTGCTCAGAAGAGTTATGGAAGTTAAGTATAAGG</td>
</tr>
<tr>
<td>PlzC Kpn I 3' R</td>
<td>ATAGTGGGTACATTAAATATCCATGGGATTACATAGTATAAGCAAGGGAAGTC</td>
</tr>
<tr>
<td>PlzA Age I 5' F</td>
<td>ATAGTCACCCTGTTCTTTATCTAGGAAAAATAAGGAGATATGG</td>
</tr>
<tr>
<td>PlzA Xho I 3' R</td>
<td>CTATCACTCGAGTTAAATGAAAAATATCATGGATCAACTAGTATACTC</td>
</tr>
<tr>
<td>PlzC Xho I 5' F</td>
<td>TATCACTCGAGTTTTTTGCTCAGAAGAGTTATGAAGTTAAGTATAAGG</td>
</tr>
<tr>
<td>PlzC Xho I 3' R</td>
<td>TATCACTTTAAAATATCATGGGATTACATAGTATAAGCAAGGGAAGTC</td>
</tr>
</tbody>
</table>
Figure 2-1. Schematic of FRET constructs and recombinant proteins. Panel A is a depiction of the FRET protein expression plasmids used in this study. The left was used for expression of YFP-PlzA-CFP, YFP-PlzAR145D-CFP, and YFP-PlzAR145K-CFP. The right was used for expression of YFP-PlzC-CFP. Panel B is a depiction of all of the FRET proteins used in this study, including the YFP-CFP negative control and the YFP-YcgR-CFP positive control.
Soluble proteins were isolated by high speed centrifugation and purified using Ni-NTA affinity chromatography with His-trap columns (GE Healthcare) on an AKTA purification system (GE Healthcare). The FRET proteins were dialyzed into PBS and the His-tagged proteins into a buffer containing 150 mM NaCl and 10 mM Tris (pH 8.0) using Spectra/Por dialysis membranes (Spectrum Laboratories). Proteins were quantified using a bicinchoninic acid protein assay kit (Pierce) or 280 nm wavelength absorption, then visualized via Coomasie-stained sodium dodecyl sulfate polyacrylamide gel electrophoresis (BioRad) to assess purity.

**FRET analysis**

Assays were performed using a *PHERAstar FS microplate reader* (BMG) with a YPET CYPET 425 480 530 filter. Excitation was induced with a light wavelength of 425 nm and fluorescence intensity was detected at 480 nm (CFP emission) and 530 nm (YFP emission). Fluorescence intensity was measured every 0.2 seconds at each wavelength over 30 seconds. Recombinant FRET proteins (500 nM) were assayed in triplicate in low volume 384-well black plates (ProxiPlate). 5 µL of c-di-GMP (Biolog) was injected at the 5 second time point for a final concentration of 25 µM in 25 µL.

**Size exclusion chromatography**

His-tagged PlzA and PlzC (40 µM) were incubated with and without c-di-GMP (40 µM) in 550 µL of a buffer containing 150 mM NaCl and 10 mM Tris (pH 8.0) on ice for 1 hour. A control of c-di-GMP alone was also incubated. 500 µL of each reaction was separated on a Superdex-75 column (GE Healthcare) using an AKTA purification
system (GE Healthcare) and elution monitored at 280 nm. 250 µL fractions were collected for later analyses.

**Reverse phase chromatography**

Three of the protein and c-di-GMP elution fractions were consolidated into 750 µL samples. These samples were boiled for 10 min to denature the protein, then centrifuged at 14,000 rpm for 15 min to remove the precipitated protein. The aqueous portions were filtered using Amicon Ultracel filters with a molecular weight cut off of 3 kDa (Millipore). 500 µL of each sample was injected into the 15 x 4.6 cm Supelcosil LC-18-T column (Supelco) using an AKTA purification system (GE healthcare). The reverse phase gradient program used a buffer on line A containing 100 mM KH$_2$PO$_4$ and 4 mM tetrabutyl ammonium hydrogen sulfate (pH 5.9) and the buffer on line B was 75% buffer A and 25% methanol. The step gradient ran at a flow rate of 0.5 mL min$^{-1}$ for 18 mL. The program is as follows with the time point and percentage of buffer B: 0 min, 0%; 7.0 min, 30%; 14.0 min, 60%; 19.6 min, 100%; 30.8 min, 50%; and 32.2 min, 0%. Absorbance was measured at the 254 nm wavelength for the detection of nucleic acids.

**Isothermal titration calorimetry**

ITC was performed using a VP-ITC microcalorimeter (MicroCal) at 25°C. His-tagged proteins and c-di-GMP were suspended in a buffer containing 150 mM NaCl and 10 mM Tris-HCl (pH 8.0). PlzA and the two substitution mutant proteins were assayed at a concentration of 50 µM in the cell with 1 mM c-di-GMP titrating in the cell from the syringe. 20 µM of PlzC was assayed in the cell with 0.2 mM c-di-GMP in the syringe.
For all reactions, the initial c-di-GMP injection was a 0.5 µL volume over a duration of 1 s with a spacing of 120 s. 23 more injections of 1.5 µL followed for a 3 s duration and 120 s spacing. Data was analyzed using Origin software assuming a single binding site (MicroCal).
Chapter 3: Results

Recombinant protein purity

FRET recombinant proteins (YFP-YcgR-CFP, YFP-PlzC-CFP, YFP-PlzA-CFP, and subsequent PlzA mutants) were shown at the expected size of ~85 kDa while the negative control (YFP-CFP) migrated at a molecular weight of 58 kDa (Figure 3-1a). His-tagged PlzA, PlzAR145D, and PlzAR145K migrated to the expected molecular weight of 30 kDa and PlzC, to 35 kDa (Figure 3-1b).

PlzA and PlzC undergo a FRET-detectable conformational change when c-di-GMP binds

FRET is a technique that utilizes fluorophores that share the same wavelength range of emission by a donor and excitation by an acceptor. Therefore, the relative proximity of a donor to an acceptor molecule can be measured based on the amount of transference. If a FRET pair flanks a ligand receptor that exhibits a conformational change upon ligand binding, the changes in the distance and/or orientation of the two fluorophores due to ligand binding can be detected through fluctuations in the emission ratio. The specifics of this technique can be seen in Figure 3-2.

In vitro kinetics of the YFP-PlzA-CFP and YFP-PlzC-CFP proteins displayed a change in the emission ratio (480 nm/530 nm) upon addition of c-di-GMP that is
Figure 3-1. SDS-PAGE of recombinant proteins purified in this study. 1 µg of purified recombinant protein was run on a 15% Tris-HCl polyacrylamide gel. The gel was stained with Coomassie Brilliant Blue. Panel A consists of the His-tagged FRET recombinant proteins. Panel B consists of the His-tagged recombinant proteins. All proteins migrated to the expected molecular weight, shown by the ladder and arrows on the left.
Figure 3-2. Fluorescence resonance energy transfer for conformational change detection. CFP is excited by a laser beam with a wavelength of 425 nm, which can in turn excite YFP with a 480 nm emission, which results in a secondary emission at 530 nm. The efficiency of this energy transfer is dependent on the distance between these two fluorophores. Therefore, the relative proximity/orientation of the YFP and CFP can be measured based on the energy transfer between them. By introducing a spacer protein between the FRET pair, a structural rearrangement due to ligand binding can be detected by a change in the emission ratio of YFP and CFP.
depicted in Figure 3-3. This is indicative of a c-di-GMP-induced conformational change in PlzA and PlzC. YFP-YcgR-CFP displayed a similar decrease in FRET efficiency after c-di-GMP injection that is comparable to previous studies (39). In vitro kinetics of the YFP-CFP protein confirmed that the fluorophores do not independently undergo a c-di-GMP-induced conformation change in since there is no change in the emission ratio when c-di-GMP is injected.

**PlzAR\textsubscript{145D} loses conformational change while PlzAR\textsubscript{145K} behaves similarly to PlzA**

The YFP-PlzAR\textsubscript{145D}-CFP protein exhibited no change in the emission ratio after c-di-GMP injection. Therefore, there was a loss of c-di-GMP-induced conformational change by PlzA when the arginine at position 145 was mutated into an aspartic acid, shown in Figure 3-4. However, the YFP-PlzAR\textsubscript{145K}-CFP protein did have a change in the emission ratio, therefore, a structural rearrangement was detected when binding c-di-GMP and mutating Arg145 into a lysine did not disrupt the c-di-GMP-induced conformational change by PlzA.

**PlzA and PlzC exist as monomers both bound and unbound to c-di-GMP**

Using size exclusion chromatography, our lab analyzed the c-di-GMP-bound and unbound state of PlzA and PlzC for oligomerization. PlzA and PlzC alone eluted off of the column at the monomeric molecular weight. Additionally, after being incubated with c-di-GMP, each protein eluted at a molecular weight comparable to the unbound protein, indicating that the monomeric state was maintained (Figure 3-5).
Figure 3-3. C-di-GMP induced conformational change by PlzA and PlzC detectable by FRET. Excitation of FRET proteins was induced with a 425 nm laser beam and fluorescence intensity was detected at 480 nm (donor emission) and 530 nm (acceptor emission). Fluorescence intensity was measured at each emission wavelength for 30 s. The raw fluorescence intensity measurements over time are displayed in the bottom graph of each panel. C-di-GMP was injected into the wells at the 5 s time point. Therefore, a change in the emission ratio at 5 s is indicative of a c-di-GMP-induced conformational change. The emission ratio of each protein over time can be found in the top graph of each panel. YFP-CFP acted as a negative control with no change in the emission ratio. YFP-YcgR-CFP acted as a positive control, with a change in the emission ratio at the 5 s time point. Both YFP-PlzA-CFP and YFP-PlzC-CFP exhibited a change in the emission ratio upon addition of c-di-GMP.
Figure 3-4. FRET analyses of PlzA with a drastic and conservative substitution mutation of arginine 145. PlzA substitution mutants of position 145 were introduced into the FRET construct to characterize the PlzA c-di-GMP binding domain. FRET assays were conducted similarly to Figure 3-3. The YFP-PlzA-CFP graphs on the left are for comparison. The R145D substitution mutant protein resulted in an abolishment of the change in the emission ratio upon c-di-GMP binding. However, the R145K substitution mutant retained the structural rearrangement exhibited by the YFP-PlzA-CFP protein.
Figure 3-5. PlzA and PlzC exist as monomers in both the apo and holo form. Size exclusion chromatography results of proteins with and without c-di-GMP on an Superdex-75 column, detection was with 280 nm absorbance. Specific fractions were collected and run on a reverse phase column. Those fractions are denoted by the brackets and the reverse phase results of those fractions indicated by the arrow. C-di-GMP detection was with 254 nm absorbance. The PlzA and PlzC alone fractions (panel A and C) saw no detection of c-di-GMP. They each eluted at the monomeric molecular weight based on a standard curve. Incubation of PlzA and PlzC with an equivalent concentration of c-di-GMP was carried out, then run on a gel filtration column. Again, PlzA and PlzC eluted at the monomer molecular weight (panel B and D). The protein and c-di-GMP elution fractions were then injected onto a reverse phase column to detect for c-di-GMP and confirm that the proteins were bound to c-di-GMP. To verify the c-di-GMP retention volume, an equal concentration and volume of c-di-GMP was run on the gel filtration column and the elution fraction run on the reverse phase column. The c-di-GMP retention volume off of the gel filtration column and the elution volume off of the reverse phase column were consistent with the PlzA:c-di-GMP and PlzC:c-di-GMP runs (panel E).
To confirm that c-di-GMP was present or absent in the appropriate eluates, elution fractions from the size exclusion chromatography were loaded onto a reverse phase column to detect the di-nucleotide. A peak at the 18 mL retention volume in the protein and c-di-GMP elution fractions of both the PlzA-c-di-GMP and PlzC-c-di-GMP run confirmed that c-di-GMP is bound by these proteins. The absence of this peak in the protein fraction of the PlzA and PlzC alone runs served as negative controls.

**Thermodynamic characterization of the PlzA:c-di-GMP and PlzC:c-di-GMP interactions**

ITC methods were performed to measure the energy usage in maintaining a set temperature as the c-di-GMP ligand was titrated into a solution containing the *Borrelia* c-di-GMP receptors. In this way, the amount of heat produced or absorbed by the reaction could be detected and a sigmoidal titration curve obtained. From this curve, the stoichiometry and affinity of the receptor for the ligand could be calculated. In addition, by detecting the energy necessary to return the sample temperature to 25°C, the net enthalpy change can be measured from which the net change in entropy can be calculated using the Gibbs free energy equation. Thus, the c-di-GMP interactions with recombinant PlzA and PlzC were characterized. Raw kinetic measurements of power compensation and titration curves can be found in Figure 3-6. Based on the data, the dissociation constant of PlzA:c-di-GMP was calculated to be 6.25±1.10 µM and PlzC:c-di-GMP was 5.92±0.963 µM. The binding stoichiometries of both associations were ~1:1. These interactions were exothermic and enthalpy-driven, exhibiting high net changes in enthalpy and low net changes in entropy. Drastic net entropic changes...
Figure 3-6. PlzA and PlzC bind c-di-GMP with moderate affinity. Isothermal titration calorimetry was conducted to characterize the interaction of PlzA, PlzAR145D, PlzAR145K, and PlzC with c-di-GMP at 25°C. The top panels depict the raw isotherm titration calorimetry data. The bottom panels show the binding isotherm data with calculated, dilution-corrected heat release for each reaction with a fit curve using a single site interaction model. Thermodynamic profiles of these reactions can be found in Table 3-1.
would suggest hydrophobic interactions because of a reduced organization of the water molecules that interact with the protein surface. However, the PlzA:c-di-GMP and PlzC:c-di-GMP ITC data indicate that the heat release in these reactions is due to hydrogen bond formation, not hydrophobic interactions. The complete thermodynamic profile of PlzA:c-di-GMP and PlzC:c-di-GMP can be found in Table 3-1.

For the purpose of the PlzA and PlzC PilZ domain characterization, our lab mutated Arg145 into an aspartic acid and lysine residue. The PlzAR145D and PlzAR145K interactions with c-di-GMP were measured under the same conditions as PlzA and supported the FRET analyses. PlzAR145D exhibited no significant binding events with c-di-GMP and PlzAR145K binding was comparable to PlzA.
Table 3-1. Thermodynamic profiles of PlzA, PlzC, PlzA<sub>R145D</sub>, and PlzA<sub>R145K</sub>.

<table>
<thead>
<tr>
<th></th>
<th>Syringe: c-di-GMP</th>
<th>Cell: PlzA</th>
<th>Cell: PlzA&lt;sub&gt;R145K&lt;/sub&gt;</th>
<th>Cell: PlzC</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>K&lt;sub&gt;d&lt;/sub&gt; (µM)</strong></td>
<td>6.25±1.10</td>
<td>1.18±0.460</td>
<td>5.92±0.963</td>
<td></td>
</tr>
<tr>
<td><strong>N (syringe/cell)</strong></td>
<td>0.91±0.032</td>
<td>0.693±0.026</td>
<td>1.12±0.034</td>
<td></td>
</tr>
<tr>
<td><strong>ΔH (kcal mol&lt;sup&gt;-1&lt;/sup&gt;)</strong></td>
<td>-13.7±0.673</td>
<td>-11.6±0.599</td>
<td>-13.7±0.749</td>
<td></td>
</tr>
<tr>
<td><strong>ΔS (cal mol&lt;sup&gt;-1&lt;/sup&gt;K&lt;sup&gt;-1&lt;/sup&gt;)</strong></td>
<td>-22.1</td>
<td>-9.95</td>
<td>-22.0</td>
<td></td>
</tr>
</tbody>
</table>
Chapter 4: Discussion

C-di-GMP is a ubiquitous signaling molecule utilized by many bacterial species in the regulation of cellular processes (109). Since its discovery 25 years ago, this di-nucleotid e has been implicated in the modulation of many factors contributing to host invasion (109). However, it is not c-di-GMP itself that carries out these effector functions, but the receptors to which it binds (109). Therefore, the discovery and characterization of these c-di-GMP effectors is highly desirable.

Many c-di-GMP receptors contain other recognizable domains, such as the helix-turn-helix of a DNA binding domain or a membrane fusion domain for ABC transporters (81, 109, 143). While PlzA and PlzC each have C-terminal PilZ domains, their sequences contain no other clues as to their purpose within the cell since they harbor no significant sequence homology to known protein domains. Therefore, insights into their potential function and activity must be delineated by examining c-di-GMP effects within the Borrelia.

Multiple groups have worked to divulge c-di-GMP functionality in the Lyme spirochete. Rogers et al. (107) was the first to generate an rrp1 deletion mutant (Δrrp1) and eliminate the source of c-di-GMP synthesis. Through transcriptome analysis of Δrrp1, it was found that in the absence of the diguanylate cyclase, 140 genes had altered expression, which correlates with ~10% of the genome. Several of these genes
are involved in motility, chemotaxis, and metabolism, implying that c-di-GMP may function to regulate these processes.

One of the many interesting genes within the Rrp1 regulon is *rpoS*, one of the alternate sigma factors of *B. burgdorferi*. Specifically, this transcriptional modulation appears to be through the oxidative stress response regulator, BosR (133). Decreased protein expression of both BosR and RpoS occurred in Δ*rrp1* (133). How BosR is activated to promote RpoS transcription is unknown, but appears to be through a c-di-GMP effector.

Many other studies have found that both Rrp1 and Hpk1 are upregulated during tick feeding (55, 87, 107). This implies that the Rrp1-Hpk1 system may be important in the expression of genes significant for spirochete survival in the tick and was confirmed experimentally with tick infection studies (62, 71). Thus, it became readily apparent that c-di-GMP has a part to play in *B. burgdorferi* survival within the tick.

One pathway by which c-di-GMP appears to affect tick invasion is through activation of the glycerol metabolic pathway (62). An operon of glycerol transport and metabolism genes (*glp*) was reported to be a part of the Rrp1 regulon through multiple microarray studies (62, 107). Since glucose is an ample source of nutrients in the mammalian host, Δ*rrp1* would be unaffected during murine infection. However, glycerol is likely an important carbon source for the spirochete in the tick vector. A deletion mutant was generated in the *glp* operon to assess if this pathway is the sole cause of the altered Δ*rrp1* phenotype (62). The *glp* mutant was capable of survival within the tick,
albeit at a reduced level (62). This indicates that there is at least one other c-di-GMP-regulated factor that contributes to *B. burgdorferi* proliferation in the tick.

While the absence of c-di-GMP appears to only significantly affect colonization of the tick vector, uncontrolled production has a negative effect on both tick and mammalian invasion by *B. burgdorferi* (71, 131, 132). The *pdeA* deletion mutant had increased intracellular c-di-GMP accumulation and was unable to infect mice even at very high bacterial loads (132). Additionally, an Rrp1 overexpression strain was unable to infect either the murine or arthropod host (71). This phenotype within the *Borrelia* is consistent with the rest of the bacterial world where high levels of c-di-GMP correlate with decreased infectivity (109).

Since c-di-GMP signaling in the *Borrelia* appears to be very important in host invasion and enzootic cycle progression, the discovery of the responsible effector molecule(s) is vital to delineating the mechanism of c-di-GMP activity. PlzA and PlzC are the only confirmed c-di-GMP binding molecules within the *Borrelia*. Like *rrp1*, *plzA* transcript levels are upregulated during tick feeding (55). This indicates that factors of the bloodmeal induce the transcription of *plzA* even though it is not within the Rrp1 regulon (107). Multiple groups have generated *plzA* mutant strains and each had decreased infectivity in mice by needle inoculation as well as decreased colonization of the tick by immersion (70, 98). Despite invasion of both the vector and mammalian host, albeit mitigated, natural transmission of the spirochete from one host to the other did not occur (98). Therefore, a *plzA* mutant is incapable of completing the enzootic cycle, which correlates with the phenotypes of other *B. burgdorferi* strains with mutated c-di-GMP-associated genes. In addition to a *plzA* mutant, Pitzer et al. created a *plzA*
deletion in the \textit{pdeA} mutant background (98). This strain had detectably higher intracellular c-di-GMP levels, but the phenotype of the \textit{pdeA} and \textit{pdeA-plzA} deletion mutants were identical. Therefore, the non-pathogenic phenotype of the \textit{pdeA} mutant is not because of altered PlzA activity due to bolstered c-di-GMP levels in the spirochete.

Along with mitigated infectivity, the \textit{plzA} deletion mutant displayed decreased motility speed (70). This phenotype could be accounted for through altered motility gene expression or protein-protein interactions with the flagellar motor. Changes in the transcription levels of \textit{fliG1}, \textit{cheY3}, and \textit{fliI} by the \textit{plzA} mutant implicate the former, but does not exclude the latter (70). Unlike most other bacterial species, the \textit{B. burgdorferi} genome encodes two motor switch FliG proteins (FliG1 and FliG2, 52). FliG2 appears to function as a traditional motor switch protein, localizing to each periplasmic flagellar bundle (74). Conversely, FliG1 localizes to one of the poles and, interestingly, a \textit{fliG1} deletion mutant has only one functional flagellar bundle (74). Analysis of the \textit{fliG1} mutant in viscous media revealed that each bundle comprises normal flagellar filaments, but only one managed to gyrate, which resulted in severely decreased translational motion (74). This presented a unique opportunity to test the necessity of motility in \textit{B. burgdorferi} infectivity without interfering with the architecture that the flagella provide (74). Murine infection studies supported that a \textit{fliG1} mutant is non-infectious, reaffirming the significance of motility in spirochetal pathogenicity (25, 74).

While both motors of the \textit{plzA} mutant do manage to gyrate in viscous media, the translational speed and flexation per second were severely reduced (70). It is tempting to speculate that this phenotype is because of altered FliG1 activity due to a lack of
PlzA regulation. However, more studies must be conducted on the functionality and potential interaction of PlzA with FliG1.

In an effort to shed light on the activity and biochemistry of PlzA and the *B. hermsii* homologue, PlzC, this study sought to explore their interaction with c-di-GMP. Often, ligand binding induces a conformational change in the receptor and results in an alteration in function. A method of detecting this is FRET where the protein in question is flanked by a FRET pair. Changes in the relative distance or orientation of the protein termini upon ligand binding can be detected by measuring the emission ratio of the FRET pair.

Through FRET analyses, this study divulged that the only known c-di-GMP receptors encoded by the Lyme spirochete, *B. burgdorferi*, and the TBRF spirochete, *B. hermsii*, exhibit a structural rearrangement when binding c-di-GMP likely significant to downstream functions. Both of the PilZ proteins had a decrease in FRET efficiency upon addition of c-di-GMP; this could be due to an increase in the distance between the amino and carboxyl termini of the protein or simply due to a change in the orientation. However, c-di-GMP-induced oligomerization or monomerization of the FRET proteins could also result in interference and affect the analysis. Therefore, the potential for PlzA and PlzC multimerization had to be assessed.

The *Pseudomonas putida* PilZ domain protein PP4397 has been shown to transition from a dimer to a monomer based on c-di-GMP binding (119). When bound by PP4397, c-di-GMP occludes the dimeric interface, preventing dimerization (69). Therefore, a change in state induced by c-di-GMP binding does occur in other PilZ domain proteins. However, size exclusion chromatography of PlzA and PlzC bound and
unbound to c-di-GMP upheld that these proteins exist as monomers in the apo and holo state. Therefore, the change in FRET efficiency exhibited by PlzA and PlzC is due to intramolecular rearrangements upon ligand binding and not intermolecular interactions.

Cyclic di-GMP is a freely diffusible molecule and bacteria lack an apparent means of physical sequestration in the form of membrane-bound organelles. Therefore, varying receptor affinities, intracellular localization, and differential expression levels have a key role to play in where and when c-di-GMP affects cell processes (65). A study was conducted in *S. enterica* serovar Typhimurium in which the disparate affinities of the only two PilZ domain proteins were found to be significant in governing two different cellular processes based on intracellular c-di-GMP concentrations (101). YcgR has been implicated in c-di-GMP-induced negative regulation of flagellar rotation through protein-protein interactions. This is in order to initiate the transition from motility into sessility. YcgR is also a c-di-GMP receptor with one of the highest known PilZ domain affinity for c-di-GMP (~190 nM, 37, 101). Alternately, the only other PilZ domain protein, BcsA, has a comparatively modest affinity of 8.2 µM. Upon c-di-GMP binding, BcsA positively regulates cellulose production, a function that facilitates biofilm formation. Therefore, it was hypothesized by Putz et al. that there is a c-di-GMP concentration threshold for each of these proteins that must be surpassed before the c-di-GMP-dependent function can be performed (101). In the case of this study in *Salmonella* Typhimurium, the YcgR threshold is lower so that as c-di-GMP is synthesized motility is depressed before the BcsA threshold is reached and cellulose production initiated (101).
The binding affinities of PlzA and PlzC for c-di-GMP were calculated using titration methods. They share a similar affinity, which is not surprising considering the high degree of sequence homology between the two. In this study, PlzA’s affinity for c-di-GMP was calculated to be 6.25 µM and PlzC’s was 5.92 µM. Comparing these affinities with the *Salmonella* PilZ proteins revealed that they are closer to the affinity of BcsA than YcgR. However, the phenotype of a plzA mutant more closely correlates with YcgR activity in that it affects bacterial motility. Since YcgR interacts with the motor switch protein, FliG and FliM, it is appealing to consider that PlzA may act in a similar manner.

The *Borrelia* c-di-GMP receptors, PlzA and PlzC, harbor multiple residues that are highly conserved among PilZ domain proteins. A sequence comparison of PlzA and PlzC to other PilZ domains can be seen in Figure 4-1. Substitution mutations and structural determination of other PilZ domains have shown that the two arginines of the RxxxR motif directly interact with c-di-GMP, primarily through hydrogen bond formation and non-covalent interactions (69). Several other residues also seem to affect c-di-GMP binding, these can include those in the DzSxxG motif or other residues more distal in sequence from the PilZ domain (37, 56, 69, 81, 113, 115, 119). However, unlike the RxxxR, these seem to vary a great deal.

In order to further characterize the PilZ domain of PlzA, our lab generated recombinant proteins with substitutions of the arginine at position 145, which is the first in PlzA’s 145RxxxR motif. The first was a conservative mutation from an arginine to a lysine. Both of these amino acids are basic with the two highest isoelectric points of the amino acids. Additionally, their structures are alike in that their side chains are similar in
Figure 4-1. Sequence alignment of the *Borrelia* PilZ domain proteins with other characterized PilZ domain proteins. Alignment was generated by ClustalW 2.0 and ESPRIPT. Identical amino acids are highlighted and similar amino acids are surrounded by a box. High conservation of the RxxxR/DzSxxG PilZ motif between all proteins, including PlzA and PlzC is shown. VCA0042 is PlzD of *V. cholera*. DgrA and DgrB are encoded by *Caulobacter crescentus*. PA4608 is a *P. aeruginosa* PilZ domain protein. PP4397 is produced by *P. putida*. CesA is from *G. xylinum*. MrkH is encoded by *Azospirillium brasilense*. YcgR is a c-di-GMP binding protein associated with motility in *Salmonella Typhimurium*. 
length and end in a positively charged nitrogen-based group, although arginine comprises a guanidinium group and lysine, an amino group. The retention of c-di-GMP binding and structural rearrangement capabilities must be attributed to the high degree of similarity between these two amino acids. The second was a drastic mutation from an arginine to an aspartic acid. The difference in the isoelectric points of these two amino acids is the most extreme of any pair. This mutation ensured that the localized surface charge of the 145RxxxR motif would be significantly altered. Results obtained in this study have shown that this mutation did, indeed, not only abolish the conformational change, but resulted in no detectable binding of c-di-GMP by PlzA. Other mutations of this particular residue within other PilZ domain proteins from an arginine to an alanine have also led to abolished c-di-GMP binding (100, 119). The fact that the PlzA<sub>R145K</sub> substitution recombinant protein was still capable c-di-GMP binding and conformational change suggests that the interaction is not necessarily primary sequence dependent. However, the abolishment of binding in the PlzA<sub>R145D</sub> protein and retention of binding in the PlzA<sub>R145K</sub> protein indicates that a positively charged at position 145 is very important in the PlzA-c-di-GMP interaction.

Many PilZ domain proteins contain another arginine residue just upstream of the RxxxR motif, making it RRxxxR. Titration studies of these proteins with c-di-GMP have revealed that every characterized protein that harbors this additional arginine residue binds a c-di-GMP intercalated dimer, including YcgR (37, 56, 69, 113, 115, 119). Conversely, if the PilZ domain protein retains a residue other than an arginine at that position, the protein binds the monomeric form of c-di-GMP, including PlzD of V. cholera (15, 69, 81). Examination of the PlzA and PlzC sequence shows that neither encodes
an arginine at this position. The titration analyses in this study revealed that both of these bind c-di-GMP with an ~1:1 ratio. An interesting study mutated the residue just upstream of the RxxxR motif to encode an arginine in PlzD of *V. cholera* and investigated the binding stoichiometry of this substitution mutant protein to find that the ratio had shifted to a 1:2 (69). Considering the conservative metabolic nature of the *Borrelia* with genetic and metabolic materials, it is not surprising that these bacteria should evolve to encode a c-di-GMP receptor that only binds one molecule of its ligand instead of two.

While PlzA is the only known c-di-GMP receptor encoded by the Lyme spirochete, in comparing the phenotypes of the *rrp1* and *plzA* deletion mutants, it is apparent that there may be other c-di-GMP effectors produced by the Lyme spirochete. Otherwise, the two mutants would exhibit more comparable phenotypes. Another speculation is that PlzA may have functionality independent of c-di-GMP binding. Phenotypic characterization of a *B. burgdorferi* mutant encoding a non-c-di-GMP binding PlzA, such as PlzAR145D, would be crucial in elucidating the role of c-di-GMP binding by PlzA *in vivo*.

Most research efforts in delineating the c-di-GMP functionality in spirochetal regulatory networks has been concentrated in *B. burgdorferi* and, more recently, *B. hermsii*. However, it is significant to note that there appear to be c-di-GMP networks in all of the *Borrelia* with sequenced genomes (109). Analyses of the *B. bissettii*, *B. afzelii*, *B. valaisiana*, *B. garinii*, *B. spielmanii*, *B. turicatae*, *B. duttoni*, *B. recurrentis*, and *B. crocidurae* genomes have revealed that these Lyme and relapsing fever spirochetes harbor genes that share a great deal of homology with the *B. burgdorferi* and *B. hermsii*
c-di-GMP systems, including PlzA and PlzC. Additionally, the activity of a PilZ domain protein encoded by another pathogenic spirochete, *Treponema denticola*, was investigated very recently (19). The results of this study revealed that a deletion mutant of the c-di-GMP receptor TDE0214 exhibited altered motility and reduced biofilm formation (19). Research in c-di-GMP signaling by the spirochete is expanding to further delineate its potential roles in biochemical processes and pathogenesis.

In summary, the c-di-GMP regulatory networks of *B. burgdorferi* and *B. hermsii* have significant involvement in many processes, including motility, metabolism, and survival. While the modulation of c-di-GMP levels within the cell through diguanylate cyclases and phosphodiesterases is very important for governing downstream effectors, c-di-GMP receptors are what perform the mediated functions of this signaling molecule. As the only known c-di-GMP binding factors within the *Borrelia*, the interactions of PlzA and PlzC with c-di-GMP are vital in determining the mechanism of c-di-GMP-mediated host invasion by the Lyme and relapsing fever spirochetes. In other effectors, the mechanism that c-di-GMP alters protein activity literally hinges around the loop-forming RxxxR motif into which di-nucleotide nestles. In this study, the structural rearrangement, monomeric state, binding stoichiometry, and ligand affinity of PlzA and PlzC were determined. These properties must contribute to their function within the cell, which has been shown to be significant in motility, infectivity, and enzootic cycle progression.

Future endeavors in characterizing the c-di-GMP regulatory networks of the *Borrelia* would include detection of c-di-GMP synthesis *in vivo* during different stages of the enzootic cycle. The use of a biosensor with a detectable output would be imperative. Studies in *Caulobacter crescentus* have used a FRET biosensor in this manner (39).
Therefore, mutagenesis of the Lyme and relapsing fever spirochetes to express the FRET proteins described in this study would further delineate how c-di-GMP functions within these bacterial species.
References


97. **Pei, J., and N. V. Grishin.** 2001. GGDEF domain is homologous to adenylyl cyclase. Proteins **42:**210-6.


Vita

Publications


Manuscripts in preparation


Presentations

November 2009, poster presentation, “A strain of the Lyme spirochete, Borrelia burgdorferi expressing a significant truncation of the critical virulence factor Outer surface protein C (OspC) retains infectivity, potentially due to compensatory expression of other surface proteins,” Katherine L. Mallory, Christopher G. Earnhart, and Richard T. Marconi. Virginia Branch American Society for Microbiology, Richmond, VA.