DISSERTATION

CHARACTERIZATION OF *MYCOBACTERIUM LEPRAE* DIGUANYLATE CYCLASES

Submitted by

Suwatchareeporn Rotcheewaphan

Department of Microbiology, Immunology, and Pathology

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Summer 2016

Doctoral Committee:

Advisor: John Belisle

Brad Borlee
Patrick Brennan
Mary Jackson
Jessica Prenni

ABSTRACT

CHARACTERIZATION OF *MYCOBACTERIUM LEPRAE* DIGUANYLATE CYCLASES

Mycobacterium leprae is the causative agent of leprosy, which is still a major health problem in several developing countries. Management of leprosy has been challenging because of the long incubation period of the disease and the development of a spectrum of clinical manifestations. Leprosy treatment is further complicated by the development of drug resistance. Knowledge of infection mechanisms and pathogenesis of leprosy is still limited. These fundamental gaps significantly limit the development of disease management, including treatment and prevention.

Although *M. leprae* is an obligate intracellular pathogen, this bacterium must possess mechanisms to adapt to different host defenses or cell types. The discovery of cyclic diguanylate monophosphate (c-di-GMP) and its potential roles in bacteria as a second messenger to regulate several cellular activities responding to environmental stimuli have stimulated an interest on c-di-GMP studies in *Mycobacterium* spp., especially *M. leprae* which has massive gene decay but still harbors several potential proteins functioning as diguanylate cyclases. The hypothesis of this study is that *M. leprae* has the ability to synthesize c-di-GMP. This study evaluated *M. leprae's* potential to synthesize c-di-GMP. Bioinformatics analyses were performed to identify proteins that are involved in c-di-GMP synthesis (diguanylate cyclase, DGC) and turnover (phosphodiesterase, PDE). Bioinformatics revealed that *M. leprae* harbors a putative DGC-PDE protein (ML1750c) and two putative DGC proteins (ML1419c and ML0397c). Interestingly, homologues of ML1419c

and ML0397c are not encoded by *Mycobacterium tuberculosis.* The *M. leprae* genes *ml1419c*, *ml0397c*, and *ml1750c* were cloned and expressed in *Pseudomonas aeruginosa* PAO1 and *Escherichia coli* BL21(DE3) pLysS. Assays for well-described phenotypes of c-di-GMP production (colony morphology, macromolecule synthesis, and biofilm formation) were performed with the recombinant clones. Direct measurement of c-di-GMP levels was accomplished by LC-MS. RNA was extracted from *M. leprae* infected mouse footpads, and expression of *ml1419c* and *ml0397c* was measured by droplet digital PCR. DGC proteins produced by *M. leprae* in armadillo tissue were also monitored with protein-specific polyclonal antibodies.

Phenotypic studies revealed that recombinant expression of *ml1419c* in *P. aeruginosa* altered colony morphology, motility, and biofilm formation, and the recombinant expression of *ml0397c* increased curli and cellulose production of *E. coli.* These phenotypes were consistent with increased DGC activity and c-di-GMP production. LC-MS analyses confirmed increased c-di-GMP production by ML1419c and ML0397c. *In vivo* gene expression studies revealed that *ml1419c*, *ml0397c*, and *ml1750c* are expressed by *M. leprae* during infection. Additionally, ML1419c and ML1750c proteins were clearly identified in whole cell sonicate of armadillo derived *M. leprae*. This study demonstrated that *M. leprae* has significant potential to produce c-di-GMP. ML1419c and ML0397c were confirmed as functional DGCs.

This study is significant because it provides evidence that *M. leprae* has the ability to produce c-di-GMP. Furthermore, these studies will pave the way for future research to characterize the biological roles of c-di-GMP in *M. leprae* and the pathogenesis of leprosy. Continued studies to elucidate the biological roles and the environmental signals for ML1419c, ML0397c, and ML1750c are being performed. These efforts are directed at defining the function of c-di-GMP in *M. leprae*. It is anticipated that these future efforts along with the data in this dissertation will shed

light on the signaling mechanisms that respond to environmental changes experienced by *M. leprae*.

ACKNOWLEDGEMENTS

There are several people who deserve acknowledgements for the successfulness of this work. First of all, I would like to sincerely thank Dr. Patrick Brennan who gave me the precious opportunity to study at Colorado State University and for all the suggestions during last five years. My time at Colorado State University is another challenging chapter in my life. I have received a lot of wonderful experiences here.

I gratefully thank my advisor, Dr. John Belisle for giving me the opportunity to be a part of his lab. I really appreciate his kindness and supports since I have been here. He has taught me not only knowledges, but also how to be a better critical thinker and scientific researcher.

I would like to thank my committee members Dr. Brad Borlee, Dr. Mary Jackson, and Dr. Jessica Prenni for all suggestions and guidance throughout my graduate career.

I would like to thank all of past and present Belisle, Slayden, and Borlee lab members for all helps and guidance during my time at CSU. This work cannot be accomplished without helps from those people.

I would like to thank Chulalongkorn University and Faculty of Medicine, Chulalongkorn University who provided financial support for pursuing my PhD.

I would like to thank my family that always supports me in every step of the way in my life.

TABLE OF CONTENTS

ABSTRACT ii

ACKNOWLEDGEMENTS v

LIST OF TABLES xi

LIST OF FIGURES xii

LIST OF KEYWORDS xvi

CHAPTER 1 INTRODUCTION TO *MYCOBACTERIUM* SPP. 1

1.1 General characteristics of genus *Mycobacterium* 1

1.2 Mycobacterial genome 2

1.3 Mycobacterial cell envelope 3

1.4 Mycobacterial adaptation to environmental changes 8

1.5 Leprosy and *Mycobacterium leprae* 9

1.5.1 Leprosy prevalence 10

1.5.2 Leprosy transmission 10

1.5.3 Leprosy clinical manifestations 11

1.5.4 Leprosy diagnosis 13

1.5.5 Leprosy treatment 15

1.5.6 *M. leprae* genome 16

1.5.7 Leprosy pathogenesis 19

1.5.8 Host immune response to *M. leprae* 21

1.5.9 Animal models for *M. leprae* study 23

REFERENCES 25

CHAPTER 2 INTRODUCTION TO CYCLIC NUCLEOTIDES 35

2.1 Bacterial signaling 35

2.1.1 Nucleotide-based second messengers 38

2.2 Cyclic diguanosine monophosphate (c-di-GMP) 45

2.2.1 Turnover mechanism 45

2.2.2 c-di-GMP signaling mechanism 48

2.2.3 Biological roles of c-di-GMP 50

2.2.4 c-di-GMP studies in mycobacteria 51
2.3 Purpose of the study and research hypothesis 52
2.4 Significance of the study 52
2.5 Scope of this study 53
REFERENCES 55
CHAPTER 3 C-DI-GMP DETECTION METHODOLOGY DEVELOPMENT 65
3.1 Introduction 65
3.2 Materials and Methods 68
3.2.1 Materials 68
3.2.2 LC-MS and LC-MS/MS method optimization 69
3.2.3 Calibration curves for quantitation of cyclic nucleotides and cyclic dinucleotides 69
3.2.4 Validation procedure 70
3.2.5 Cyclic nucleotide and cyclic dinucleotide extraction from *P. aeruginosa*, *E. coli*, and *M. tuberculosis* H37Ra 71
3.2.6 Statistical analyses 72
3.3 Results 73
3.3.1 LC-MS and LC-MS/MS method performance 73
3.3.2 LC-MS method calibration curve, LOD, and LOQ 77
3.3.3 Validation 79
3.3.4 Application of LC-MS and LC-MS/MS methods 82
3.4 Discussion 84
REFERENCES 86
CHAPTER 4 ML1419c AND ML0397c FUNCTION AS DIGUANYLATE CYCLASES TO PRODUCE C-DI-GMP 89
4.1 Introduction 89
4.2 Materials and methods 92
4.2.1 Bacterial strains, genomic DNA, and growth conditions 92
4.2.2 Construction of plasmids of *ml1419c*, *ml1750c*, and *ml0397c* for *P. aeruginosa* PAO1 92
4.2.3 Construction of plasmids of *ml1419c*, *ml1750c*, and *ml0397c* for *E. coli* 93
4.2.4 Site-directed mutagenesis of *ml1419c* and *ml0397c* 94

4.2.5 Protein and whole cell lysate isolation from *E. coli* 97
4.2.6 Polyclonal antibody test 98
4.2.7 Whole cell lysate isolation from *P. aeruginosa* PAO1 99
4.2.8 SDS-PAGE and Western blot analyses 99
4.2.9 *P. aeruginosa* phenotypic assays 100
4.2.10 *E. coli* phenotypic assays 102
4.2.11 Quantitative analysis of c-di-GMP by liquid chromatography-mass spectrometry (LC-MS) 102
4.2.12 Statistical analyses 103
4.2.13 Bioinformatics analyses 103
4.3 Results 104
4.3.1 Bioinformatics analyses and identification of c-di-GMP-related proteins in *M. leprae* 104
4.3.2 Conditional expression of *ml1750c*, *ml1419c*, and *ml0397c* in *E. coli* BL21(DE3) pLysS 107
4.3.3 Polyclonal antibody test 110
4.3.4 Conditional expression of *ml1750c*, *ml1419c*, and *ml0397c* in *P. aeruginosa* PAO1 111
4.3.5 Phenotypic studies of *P. aeruginosa* PAO1 and *E. coli* BL21(DE3) pLysS expressing *ml1750c*, *ml1419c*, and *ml0397c* 116
4.3.6 *ml1419c* alters *P. aeruginosa* colony morphology 120
4.3.7 *ml1419c* expression provides for quantifiable phenotypic differences associated with DGC activity 122
4.3.8 ML1419c and PAS-GGDEF ML0397c increased curli and cellulose formation in *E. coli* 125
4.3.9 Detection of c-di-GMP *in vivo* by LC-MS 128
4.4 Discussion 135
REFERENCES 140
CHAPTER 5 DIGUANYLATE CYCLASES OF *MYCOBACTERIUM LEPRAE* AND EXPRESSION *IN VIVO* 145
5.1 Introduction 145
5.2 Materials and methods 147
5.2.1 ddPCR primer design and optimization 147
5.2.2 ddPCR reaction optimization 148

5.2.3 *M. leprae* RNA purification 150
5.2.4 cDNA synthesis 150
5.2.5 ddPCR reaction 151
5.2.6 Detection of *M. leprae* proteins encoding DGCs in infected animal tissues 151
5.2.7 Statistical analyses 155
5.3 Results 155
5.3.1 Quantification of expression of genes encoding DGC proteins by ddPCR 155
5.3.2 Expression of *M. leprae ml1750c*, *dgcA*, and *dgcB* during infection 159
5.3.3 Detection of *M. leprae* ML1750c, DgcA, and DgcB in *M. leprae* subcellular fractions 162
5.3.4 Detection of *M. leprae* ML1750c, DgcA and DgcB by mass spectrometry 164
5.4 Discussion 166
REFERENCES 170
CHAPTER 6 EVALUATION OF *M. SMEGMATIS* AS A MODEL FOR DIGUANYLATE CYCLASE FUNCTION 173
6.1 Introduction 173
6.2 Materials and methods 174
6.2.1 Bacterial strains, genomic DNA, and growth conditions 174
6.2.2 Construction of *dgcA* and *dgcB* plasmids and recombinant *M. smegmatis* strains 174
6.2.3 qRT-PCR primer design and optimization for detection of *msmeg_2196* and *sigA* 175
6.2.4 *M. smegmatis* RNA purification 175
6.2.5 cDNA synthesis 176
6.2.6 qRT-PCR of *msmeg_2196* and *sigA* 177
6.2.7 *msmeg_2196* gene replacement by homologous recombination 179
6.2.8 *M. smegmatis* genomic DNA extraction 180
6.2.9 Confirmation of the Δ*msmeg_2196* mutant strain by PCR reaction and sequencing 181
6.2.10 Phenotypic assays of *M. smegmatis* strains 182
6.2.11 c-di-GMP detection in *M. smegmatis* by LC-MS analysis 184
6.2.12 Statistical analyses 184
6.3 Results 184

6.3.1 Construction of Δ*msmeg_2196* mutant 184
6.3.2 Phenotypic characterization of the *M. smegmatis* Δ*msmeg_2196* 186
6.3.3 *msmeg_2196* expression *in vitro* 193
6.3.4 Recombinant overexpression of *dgcA* and *dgcB* in *M. smegmatis* 195
6.4 Discussion 198
REFERENCES 201
CHAPTER 7 DISCUSSION AND CONCLUSIONS 204
REFERENCES 211
APPENDIX A 216
APPENDIX B 217
APPENDIX C 219
APPENDIX D 228
LIST OF ABBREVIATIONS 233

LIST OF TABLES

Table 3-1. Cyclic nucleotides/dinucleotides used in this study 68

Table 3-2. Composition of samples for standard curve and quality control 70

Table 3-3. Calibration curve data, LOD, and LOQ 78

Table 3-4. Intra- and inter-day accuracy and precision 80

Table 4-1. Primers used in this study 95

Table 4-2. Plasmids used in this study 96

Table 4-3. Conserved residues in proteins containing EAL domains 107

Table 4-4. Phenotypic studies in *P. aeruginosa* PAO1 and *E. coli* BL21(DE3) pLysS 120

Table 5-1. ddPCR primers used in this study 148

Table 5-2. Thermocycling conditions for ddPCR reaction 149

Table 5-3. Absolute quantification (copies/µl) of *sigA*, *dgcA*, *ml1750c*, and *dgcB* from two mouse footpads 161

Table 5-4. New *M. leprae* proteins identified in this study by LC-MS/MS 165

Table 5-5. ML1750c and DgcA protein identification by LC-MS/MS 166

Table 6-1. Temperature program used for real-time PCR by LightCycler 480 instrument 178

Table 6-2. Primers used in this study 181

LIST OF FIGURES

Figure 1-1. Mycobacterial cell envelope structure 8

Figure 1-2. Clinical manifestations of leprosy and immune responses 23

Figure 2-1. Bacterial signaling systems 37

Figure 2-2. Chemical structures, formulas, and monoisotopic masses of nucleotide-based second messengers 44

Figure 2-3. c-di-GMP turnover and signaling mechanism in bacteria 48

Figure 3-1. LC-MS chromatogram of cyclic nucleotide/dinucleotide standards 74

Figure 3-2. Precursors and product ions from MS/MS of cGMP, cAMP, c-di-AMP, c-di-GMP, cGAMP, and pGpG 76

Figure 3-3. Calibration curves of cyclic nucleotides/dinucleotides 78

Figure 3-4. c-di-GMP detection in *P. aeruginosa* PAO1 strains 83

Figure 3-5. 3’5’-cAMP and c-di-AMP detection in *M. tuberculosis* H37Ra 83

Figure 4-1. Bioinformatics analyses of putative DGC and PDE proteins of *M. leprae* 105

Figure 4-2. Alignment of conserved DGC domains of *M. leprae* proteins ML0397c, ML1419c and ML1750c 106

Figure 4-3. Comparison of the active residues and positions in conserved EAL domain containing proteins from *M. leprae* TN, *M. smegmatis* mc^2155, and *M. tuberculosis* H37Rv .. 107

Figure 4-4. Expression of *ml1419c*, *ml1750c*, and PAS-GGDEF *ml0397c* from *M. leprae* in *E. coli* 108

Figure 4-5. Expression of PAS-GGDEF *ml0397c* from *M. leprae* in *E. coli* BL21(DE3) pLysS 110

Figure 4-6. Anti-ML0397c, anti-ML1419c, and anti-ML1750c polyclonal serum test 111

Figure 4-7. Expression of *ml1750c*, *ml1419c*, and PAS-GGDEF *ml0397c* from *M. leprae* in *P. aeruginosa* PAO1..... 114

Figure 4-8. Expression of full length *ml0397c* from *M. leprae* in *P. aeruginosa* PAO1 115

Figure 4-9. Expression of *ml1419c* from *M. leprae* in *P. aeruginosa* PAO1 116

Figure 4-10. Recombinant expression of *ml1419c* and *ml1750c* in *P. aeruginosa* PAO1 alters biofilm formation 118

Figure 4-11. Colony morphology of *P. aeruginosa* PAO1 and recombinant strains expressing *M. leprae* DGC genes..... 119

Figure 4-12. Colony morphology of *P. aeruginosa* PAO1 and recombinant strains expressing *ml1419c* constructs..... 121

Figure 4-13. *P. aeruginosa* expressing *ml1419c* suppresses motility..... 123

Figure 4-14. Recombinant expression of *ml1419c* in *P. aeruginosa* PAO1 induces biofilm formation..... 125

Figure 4-15. Curli formation of recombinant *E. coli* BL21(DE3) pLysS strains 127

Figure 4-16. Cellulose formation of recombinant *E. coli* BL21(DE3) pLysS strains 128

Figure 4-17. Detection of c-di-GMP standard by LC-MS..... 130

Figure 4-18. c-di-GMP detection and relative quantification of c-di-GMP in recombinant *P. aeruginosa* strains 132

Figure 4-19. c-di-GMP detection by LC-MS and relative quantification of c-di-GMP in *P. aeruginosa* recombinant strains 132

Figure 4-20. c-di-GMP detection by LC-MS of c-di-GMP in *E. coli* recombinant strains 134

Figure 5-1. Annealing temperature optimization for *sigA*, *dgcA*, *ml1750c* and *dgcB* primer sets 156

Figure 5-2. Optimization of *M. leprae* cDNA amount for ddPCR reaction 158

Figure 5-3. ddPCR results of *sigA*, *dgcA*, *ml1750c*, and *dgcB* gene expression..... 160

Figure 5-4. Absolute quantification of *sigA*, *dgcA*, *ml1750c*, and *dgcB*..... 161

Figure 5-5. Detection of DgcB, DgcA, and ML1750c in *M. leprae* subcellular fractions 163

Figure 6-1. Construction of disrupted *msmeg_2196* by kanamycin resistance gene replacement 180

Figure 6-2. Electrophoresis gel of PCR products amplified from *M. smegmatis* mc^2155 and Δ*msmeg_2196*..... 186

Figure 6-3. Growth curves of *M. smegmatis* mc^2155 and Δ*msmeg_2196*..... 187

Figure 6-4. Colony morphology of *M. smegmatis* mc^2155 and Δ*msmeg_2196* 188

Figure 6-5. Colony morphology of *M. smegmatis* mc^2155 and Δ*msmeg_2196* after two and four weeks of incubation 189

Figure 6-6. Pellicle formation of *M. smegmatis* mc^2155 and Δ*msmeg_2196*..... 190

Figure 6-7. Biofilm formation of *M. smegmatis* mc^2155 and Δ*msmeg_2196* 191

Figure 6-8. Sliding motility of *M. smegmatis* mc^2155 and Δ*msmeg_2196* 192

Figure 6-9. Real-time PCR efficiency 194

Figure 6-10. Relative quantification of *msmeg_2196* expression..... 195

Figure 6-11. Colony morphology of *M. smegmatis* strains expressing *dgcA* and *dgcB* 196

Figure 6-12. Sliding motility of *M. smegmatis* strains expressing *dgcA* and PAS-GGDEF *dgcB* on 7H9 agar without a carbon source 197

Figure 6-13. Sliding motility of *M. smegmatis* strains expressing *dgcA* and *dgcB* on 7H9 agar with 2% glucose 198

LIST OF KEYWORDS

Mycobacterium leprae

Leprosy

c-di-GMP

Second messenger

GGDEF domain

EAL domain

Diguanylate cyclase

Phosphodiesterase

Colony morphology

Biofilm formation

Swimming motility

Twitching motility

Curli formation

Cellulose formation

Sliding motility

c-di-GMP detection

CHAPTER 1 INTRODUCTION TO *MYCOBACTERIUM* SPP.

1.1 General characteristics of genus *Mycobacterium*

Genus *Mycobacterium* is the only genus of the *Mycobacteriaceae* family belonging to the Order Actinomycetales (1). The *Mycobacterium* genus comprises more than a hundred species that include obligate pathogens, opportunistic pathogens, and commensals (1). Members of this genus have close relation in their 16s rRNA sequences (2, 3). There are a variety of remarkably pathogenic mycobacteria in this genus that cause diseases in human and animals, such as the *Mycobacterium tuberculosis* complex (4), *Mycobacterium leprae* (5), and *Mycobacterium ulcerans* (6). Additionally, this genus consists of many free-living environmental mycobacteria designated nontuberculous mycobacteria (NTM) or atypical mycobacteria such as *Mycobacterium avium* and *Mycobacterium xenopi* (7) which have various degrees of pathogenicity and virulence. The NTM can behave as opportunistic pathogens in immunocompromised individuals or patients with pre-existing lung disease. Furthermore, the NTM are usually found in environments such as rivers, lakes, and soil (8-10). NTM can be classified based on growth rate and phenotypes such as pigmentation (11). The pigment-forming mycobacteria can be classified as photochromogen such as *Mycobacterium gordonae* and *Mycobacterium kansasii* which require light to form pigment or scotochromogen such as *Mycobacterium gordonae* and *Mycobacterium scrofulaceum* that can form pigment in either a light or dark environment (7).

Most of the mycobacteria are aerobic, non-spore-forming, non-motile and produce curved or straight rod shaped cells with occasional branching. Moreover, aerial hyphae are normally absent. Cells of *Mycobacterium* spp. vary in size, ranging from 0.2 to 0.6 µM by 1.0 to 10 µM (1). Colony morphology also differs among species. For example, *M. tuberculosis* forms dry rough

colonies with nodular or wrinkled surfaces while *M. bovis* BCG produces a round colony with an irregular-edge (1). The optimal temperature for mycobacterial growth also differs based on the species. For example, *Mycobacterium chelonae* and *Mycobacterium abscessus* prefer growth at 30°C while *Mycobacterium phlei* and *Mycobacterium thermoresistibile* can grow in a broad temperature range, from 30°C to 52°C (7). Mycobacteria are not easily decolorized with acid-alcohol when stained with cabol-fuchsin, and are thus called acid fast bacilli (12).

The classification of *Mycobacterium* spp. based on growth rate can divide mycobacteria into two groups; fast-growing and slow-growing mycobacteria. The fast-growing mycobacteria such as *M. smegmatis*, *M. abscessus* and *M. chelonae* usually show visible colonies on solid medium within 7 days and have a doubling time between 2-6 h. In contrast, the slow-growing mycobacteria form visible colonies in more than 7 days because of a long doubling time that is more than 24 h (3, 7, 13, 14). For example, *M. tuberculosis* and *M. leprae* have doubling times of about 24 h and 13 days, respectively (15, 16). However, it should be noted that *M. leprae* cannot be grown *in vitro* but it can be viable in axenic medium such as 7H12 medium and has a stable metabolic state for a few weeks (17).

1.2 Mycobacterial genome

Mycobacterium spp. possess genomes with high G+C content, approximately 61 to 71 mol% (1). This number varies between *Mycobacterium* spp.. For example, *M. tuberculosis* has an overall G+C content of 65.6% (18) whereas *M. leprae* has only about 57.8% G+C content (19). The genome size also varies between *Mycobacterium* spp.. *M. leprae*, *M. smegmatis* mc^2155, and *M. tuberculosis* have genome sizes of 3,268,203 base pairs (bp), 6,988,209 bp and 4,411,529 bp, respectively (19, 20). For *M. tuberculosis*, the most common translation initiation codons are ATG

(about 61%) and GTG (about 35%). There are a few atypical initiation codons such as ATC used by the *infC* gene of *M. tuberculosis* (20).

The promoters and the transcription efficiency of ribosomal RNA (*rrn*) operons including the number of *rrn* operons determine the production of rRNA and regulate the number of ribosomes. The number of *rrn* operons of the *Mycobacterium* spp. is different between the slow-growing and fast-growing mycobacteria. Generally, slow grower mycobacteria harbor a single *rrn* operon per genome. In contrast, fast-growing mycobacteria possess two *rrn* operons (21, 22). The *rrn* of *M. tuberculosis* and *M. leprae* is situated about 1,500 and 1,300 kilobases (kb), respectively, from the chromosomal origin of replication (*oriC*). This arrangement may also be related to the slow growth of these mycobacteria since the *rrn* operon that is directly adjacent to the *oriC* results in increased rRNA production (18, 20). Consequently, *M. tuberculosis* has a doubling time of about 24 h (15) and *M. leprae* has a doubling time of about 13 days (16). However, the number of *rrn* per genome is not a good indicator of growth rate in mycobacteria (14).

1.3 Mycobacterial cell envelope

The cell envelope of bacteria is the first barrier that interacts with the environment. Bacteria have evolved cell envelopes that can help to protect them from a hostile environment, but that selectively allows the uptake of extracellular nutrients and elimination of waste products (23). Thus, the cell envelope must encompass machinery and transporters for secretion of proteins and the passage of small molecules or nutrients.

The structure of bacterial cell envelope is classified into two major groups; Gram-negative and Gram-positive bacteria (23). In general, the cell envelope is comprised of a phospholipid bilayer-inner membrane (plasma membrane), cell wall and capsule in some bacteria (23, 24).

Gram-negative bacterial cell envelope scecifically consists of an inner membrane and a thin peptidoglycan cell wall which is surrounded with an outer membrane containing lipopolysaccharides (LPS). In contrast, Gram-positive bacteria possess an inner membrane and a 30-100 nm thick cell wall of peptidoglycan with additional polysaccharides or polyol phosphate polymers (teichoic acid) without an outer membrane (23).

Most bacteria are classified as Gram-negative or Gram-positive; designations that are a reflection of their overall cell envelope structures. Mycobacteria are classified as Gram-positive bacteria based on phylogenic analysis of 16s RNA (25). However, the genome-based phylogenetic trees by Fu and Fu-Lui suggest that the *Mycobacterium* genus is more closely related to Gram-negative bacteria (26). The cell envelope of mycobacteria has characteristics of both Gram-positive and Gram-negative bacteria.

Similar to other bacteria, the mycobacterial envelope is comprised of a plasma membrane, cell wall and capsule (24). The plasma membrane of mycobacteria contains most derivatives of phosphatidic acid which are phosphatidylglycerol, diphosphatidylglycerol, phosphatidylethanolamine, and phosphatidylinositol mannosides (PIMs). The other components in plasma membrane are polyterpene-based products such as carotenoids and menaquinones. Additionally, the decaprenol and octahydroheptaprenol which are presumably involved in cell wall synthesis are found in the plasma membrane (27). However, some lipids, especially phosphatidylethanolamine and PIMs have also been purified with the cell wall fraction and the capsular surface of mycobacteria (28).

Structurally, mycobacterial cell wall can be divided into an inner layer and an outer layer that surround the cytoplasmic membrane (29). The inner layer of cell wall is composed of peptidoglycan (PG) or murein, arabinogalactan (AG), and mycolic acids which are covalently

linked together to form the cell wall skeleton. This structure is an essential core of mycobacterial cell wall and is chemically similar among each mycobacterium (24). This complex structure extends from the cytoplasmic membrane outward in layers, starting with PG as a basal layer, an intermediate layer composed of AG that is esterified with mycolic acids on the exterior (27, 29, 30).

Mycobacterial PG consists of alternating muramic acid (MurNAcyl) and N-acetylglucosamine (GlcNAc) residues that are linked to each other via β-1,4-glycosidic bonds (31, 32). The Mur residues can be either N-glycolylated (MurNGlyc) or N-acetylated (MurNAc). *M. leprae* cell wall contains only MurNAc since *namH* gene responsible for MurNGlyc biosynthesis is a pseudogene (33). The peptide side chains that are L-alanyl-D-isoglutaminyl-*meso*-diaminopimelyl-D-alanine are attached to MurNAc/ MurNGlyc residues to cross-link strands into a network (30). However, the tetrapeptide chains of *M. leprae* contains glycine instead of alanine (34). In addition, the isoglutamine and diaminopimelyl (DAP) can be amidated. PG crosslinking includes bonds between two DAP and between DAP and D-alanine (27, 35, 36).

The second component of mycobacterial cell wall is the AG. The AG is a polysaccharide that is a copolymer of arabinose and galactose with a structure of linear galactan arranged in a layer horizontal to the cytoplasmic membrane and substituted with two to three side chains of arabinose (31). Terminal arabinose residues are attached to mycolic acids by esterification (24). The AG is covalently bonded to the 6-postion of muramic acid residues in PG via phosphodiester bond (an α-L-Rha*p*-(1→3)-α-D-GlcNAc-(1→P) “linker region”) (37).

Lastly, the major component of the mycobacterial cell wall is the mycolic acids that are long-chain α-branched β-hydroxy fatty acids. The mycolic acids are constructed from two fatty acids. One is a saturated C_{20}–C_{26} fatty acid with an α-branched C_{60}–C_{90} meromycolate (38). There

are variations in mycolic acid in chain length and functional groups (double bonds, cyclopropane rings, or oxygen-containing constituents) among each of the *Mycobacterium* spp. (31). The mycolic acids provide an efficient barrier to reduce permeability for uptake of dangerous compounds leading to mycobacteria being intrinsically resistant to desiccation, therapeutic antibiotics, and chemical disinfectants (12, 24, 27).

The outer membrane or outer layer of mycobacteria is unlike Gram-negative bacteria. The lipid composing the outer leaflet is different in individual mycobacterium. In general, the outer leaflet is composed of non-covalently bound proteins, lipids and lipid-linked polysaccharides that include glycerophospholipids, glycopeptidolipids (GPLs), lipoarabinomannan (LAM), lipomannan (LM), phthiocerol dimycocerosate (PDIM), and phenolic glycolipid (PGL). These proteins and lipids can play roles as immunomodulators of the host immune system (27, 29). Specifically, *M. leprae* has phenolic glycolipid I (PGL-I) that is based on the backbone structure of PDIM (24, 39). Another major structures that may be present in the outer layer as well as the plasma membrane is the lipoglycan LAM. In slow-growing pathogenic mycobacteria such as *M. tuberculosis*, LAM is capped with mannose residues at the terminal β-Ara residue. Thus, it is called mannose-capped LAM (ManLAM) (40, 41). Fast-growing mycobacteria such as *M. smegmatis* have phosphoinositol-capped or phospho-*myo*-inositol-capped LAM (PILAM). In addition, some mycobacterial LAM is not capped (non-capped LAM or AraLAM), such as in *M. chelonae* (40).

In addition, the mycobacterial envelope contains proteins called porins that span the outer membrane of mycobacteria and mediate the influx of hydrophilic solutes (42) whereas lipophilic solutes can diffuse through the lipid bilayer of the cell envelope. The mycobacterial integral outer membrane proteins have been studied, and it has been shown that outer membrane porins (MspA) are involved in uptake of phosphate in *M. smegmatis* (43) while the OmpATb of *M. tuberculosis*

is involved in the uptake of serine at low pH condition and is also required for an adaptation to low pH and survival in macrophage and mice (44).

Mycobacterium spp. also possesses a capsule-like structure that surrounds the cell wall and plasma membrane. This structure is composed of polysaccharide, protein and lipid (45). The lipid constituents of the mycobacterial capsule are 2-3% of the capsule and include the PIM, PGL, GPL, dimycoloyl trahalose, and phosphatidylethanolamine (28). The polysaccharide components of the mycobacterial capsule are mainly glucan and arabinomannan while the protein components are secreted, cell envelope-associated or lipoproteins (24). Functionally, the capsule is involved in pathogenesis of mycobacteria, and facilitates the adhesion and penetration of mycobacteria into host cells. The capsule also contains enzymes that degrade host macromolecules facilitating mycobacterial replication in host cells and resistance to host defenses (45). Mycobacterial cell envelope structure is shown in Figure 1-1.

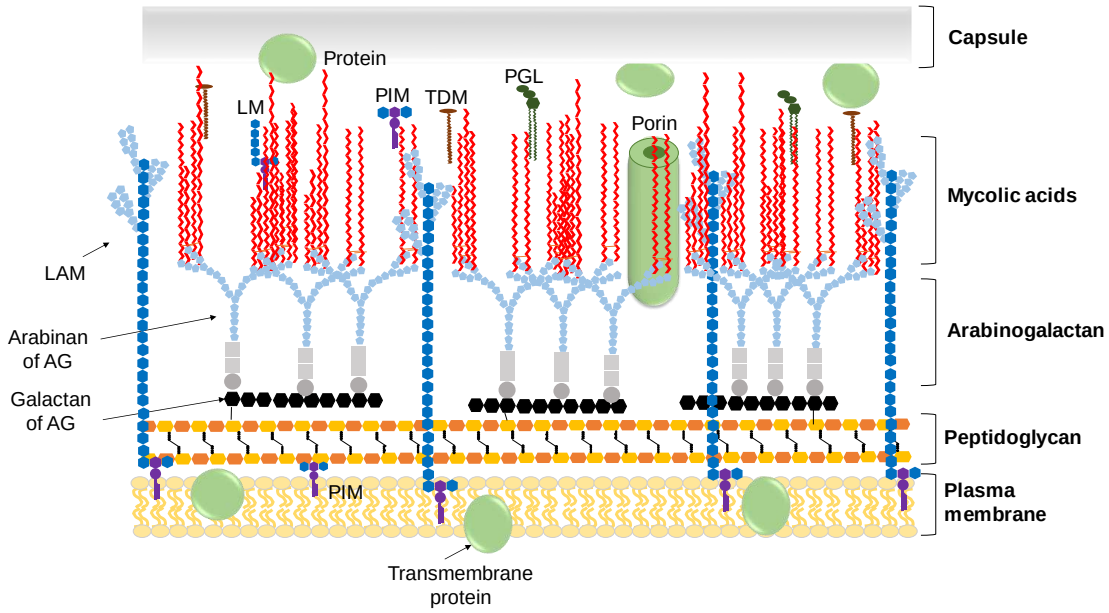


Figure 1-1. Mycobacterial cell envelope structure
(Drawn based on Patrick Brennan (1995) (27) and Mamadou Daffe and Philip Draper (1998) (24))

1.4 Mycobacterial adaptation to environmental changes

Members of the *Mycobacterium* genus have an ability to adapt to and survive in various conditions such as nutrient deprivation, hypoxia, several exogenous stress conditions and especially an intraphagosomal environment (14). When nutrient availability is altered, mycobacteria can utilize various carbon sources such as carbohydrates, fatty acids, lipids and carbon backbones of amino acids (46) and prefer amino acids, especially asparagine, glutamine and aspartate, as a nitrogen source (14). Nevertheless, mycobacteria can utilize ammonium in the absence of other nitrogen sources by using glutamine synthetase and glutamate synthase (47). Mycobacteria also are able to utilize phosphorous from inorganic phosphate for the biosynthesis of nucleic acids and phospholipids (48, 49). For cellular activities, mycobacteria require metals and trace elements such as potassium, iron, copper, and zinc (14).

Mycobacteria also respond to environmental stresses such as heat shock (50) and iron starvation (51). Most studies have been performed in *M. tuberculosis*. As an obligate aerobic

organism, *M. tuberculosis* is able to sense oxygen tension, resulting in an ability to adapt to oxygen availability in the host. DosT/DosS/DosR system (52, 53) and WhiB3 (54) are well-described systems that sense oxygen in this mycobacterium. These mechanisms complicate treatment of tuberculosis and are associated with the development of dormancy. After invading the host cell, *M. tuberculosis* can enter a non-replicating state characterized by low metabolic activity and extreme tolerance to host defenses and drug treatment (55). This non-replicating state is, in part, a response to depletion of oxygen in the host cell (56). Patients with latent tuberculosis have no signs and symptoms and low numbers of infecting bacteria. However, if the host immune defense is not effective, the disease can be reactivated, resulting in an active infection (57). *M. tuberculosis* also adapts to host environments of low pH (acidification) (58), reactive oxygen and nitrogen species (ROS and RNS), hydrolytic enzymes, cationic antimicrobial peptides (CAMPs), and limited nutrients (57).

1.5 Leprosy and *Mycobacterium leprae*

Leprosy is a chronic infectious human disease that probably has its origin in Eastern Africa or the Near East. Comparative genomics have shown that Europeans or North Americans might have introduced leprosy into West African and America within the past 500 years (59). In 1873, the causative agent of leprosy was discovered by Armauer Hansen who identified a bacterium, now known as *Mycobacterium leprae*, to be present in skin lesions of leprosy patients. *M. leprae* was identified as the first pathogenic bacterium associated with human disease (7).

1.5.1 Leprosy prevalence

Leprosy is one of the most common diseases causing non-traumatic peripheral neuropathy around the world. With the introduction of multidrug therapy (MDT) by the World Health Organization (WHO), the number of leprosy patients has been dramatically reduced from millions of cases in the 1980s to 180,618 in 2013. Globally, the elimination of leprosy was achieved in the 2000s, with the global reduction of leprosy to prevalence rate of less than one case per 10,000 people. Approximately 16 million leprosy patients have been cured with MDT over the past 20 years (60). However, leprosy is still a disease that remains a public health concern in several low and middle income countries in Southeast Asia, Africa, and Latin America. Globally, approximately 200,000 new cases are reported per year, and 175,554 leprosy patients were registered for treatment in the first quarter of 2015. Although, the global prevalence rate is about 0.31 per 100,000 population, three countries contribute the majority of cases. The number of new leprosy cases in 2014 was 125,785, 31,064, and 17,025 for India, Brazil and Indonesia, respectively. There are several other countries that still had more than 1,000 new cases reported in 2014 such as Bangladesh, Ethiopia, Madagascar, Nepal, Myanmar, Nigeria, and the Philippines (61).

1.5.2 Leprosy transmission

The mode of transmission for leprosy is not definitely known. The principal mode of human-to-human *M. leprae* transmission is probably by aerosol spread of nasal secretions, especially by untreated lepromatous patients, and uptake through nasal or mucosa of the respiratory tract (62, 63). The study of leprosy transmission by PCR in skin and nasal secretions

of untreated multibacillary leprosy patients and household contacts has shown that both skin and nasal mucosa shed *M. leprae* into the environment and to contacts (64).

M. leprae can persist in soil, plants, water and in several species of organisms such as fish, amoeba, armadillo, and non-human primates (65, 66). The transmission of leprosy to humans from environmental or zoonotic reservoirs has been identified. Chimpanzees and sooty mangabeys may serve as zoonotic sources of *M. leprae* in some geographic areas (67). The armadillo, however, is more recognized as a major reservoir of *M. leprae* in the environment and is likely a source of *M. leprae* transmission to humans (68). Genome sequence analyses by Truman *et al*. has found that many leprosy patients were infected with the same strain of *M. leprae* present in wild armadillos observed in the southern United States (69). Recently, Wheat et al. revealed that *M. leprae* can survive in environmental free-living amoeba for up to eight months within amoebic cysts. In addition, *M. leprae* extracted from amoeba cysts was able to infect mice. Therefore, amoeba could be another reservoir of *M. leprae* in the environment and responsible for leprosy transmission to humans (70). In conclusion, the ability of *M. leprae* to survive in an environment outside the human body under various conditions could be responsible for new leprosy cases by indirect contact to those sources (71).

1.5.3 Leprosy clinical manifestations

Leprosy shows clinical manifestations in a continuous spectrum of disease. Patients can have paucibacillary to multibacillary disease that is associated with a few skin lesions to progressive widespread disease, respectively. Primarily, *M. leprae* invades Schwann cells in peripheral nerves in skin. The incubation period of leprosy is varied in each patient and the infected host can present clinical manifestations in ranges from three to ten years. The minimal incubation

period is a few weeks, reported in an infected young infant case, and the maximum incubation period is about 30 years (72, 73). Clinically, key manifestations of the disease involve the skin, nerve, eye, or systemic features. Consequently, the infection can lead to nerve damage and development of disabilities (74). The involvement of nerves results in inflammation with or without tenderness and sensory or motor loss in a lesion (74).

As described by Ridley and Jobling, leprosy is classified into two major groups and three intermediate forms of disease (75) based on histopathological and immunological criteria. The classifications are tuberculoid (TT), borderline tuberculoid (BT), mid-borderline (BB), borderline lepromatous (BL) and lepromatous polar leprosy (LL) (75) (Figure 1-2). The tuberculoid leprosy (TT) patient has a few lesions which can be large erythematous plaques or hypopigmented lesions with sharply raised outer edges. The TT lesion is a well-formed granuloma with differentiated macrophage, epithelioid and giant cells and a predominance of $CD4^+$ T cells with low or absent bacteria. The patient can lose sensation in the lesions, and a thickened peripheral nerve can be observed. On the other major pole of leprosy, the LL patient presents with diffuse cellular infiltration, and several distributed and symmetrical lesions with high bacterial load. The LL lesion is typically shiny and small. In addition, the lesions have a preponderance of $CD8^+$ T cells. The sensation of most lesions is not diminished and granuloma formation is absent (73, 75, 76). The LL patients have systemic clinical manifestations which involve nasal mucosa, bones, and testes. Additionally, LL leprosy patients can have eye damage from both nerve involvement and direct invasion by *M. leprae* (74). The clinical manifestations of the other three forms of leprosy; BT, BB, and BL have presentations of disease between TT and LL forms. For example, BT, BB and BL patients present with 10, 10 to 30, and more than 30 skin lesions, respectively. Loss of sensation in lesions is progressively diminished from BT, BB, to BL patients (73).

Patients with leprosy can also develop a reaction phase of disease that is characterized by an acute episode of inflammation. Leprosy reactions are classified into two types. Type I or reversal reaction (RR) occurs in BT, BB and BL patients, while Type II or erythema nodosum lepromasum (ENL) only occurs in BL and LL patients. The ENL is often recurrent and chronic (72, 77). The RR reaction is characterized by edema and erythematous of existing skin lesions. Furthermore, patients have neuritis that can lead to irreversible deformities. Type II reaction causes tender erythematous subcutaneous nodules with systemic symptoms such as leukocytosis, fever, malaise, enlarged lymph nodes and weight loss (77). The leprosy reaction is related to a shift in immunologic status of patients by chemotherapy, pregnancy, emotional and physical stress, or co-infection with other bacteria (72).

1.5.4 Leprosy diagnosis

The diagnosis of leprosy is based mainly on clinical manifestations, which include hypopigmented or reddish or copper-colored skin lesions with definite loss of sensation with or without thickened peripheral nerves. Skin lesions can be macules (flat), papules (raised), or nodules. BL and LL patients who are at high risk to develop disabilities may not have sensory loss. Therefore, other criteria should be included for diagnosis (74). A thickened nerve is always accompanied by other signs, including loss of sensation in skin and muscle weakness, as a result of damage to the nerve. Additionally, a positive acid-fast bacilli (AFB) detected on slit-skin smears or in skin biopsy specimens is also one of the diagnostic criteria (1, 78). The bacillary index or BI is used to quantify the density of *M. leprae* and assess bacterial load for classification and response to treatment (78).

The very first diagnostic test, the lepromin skin test, is an immunologic test, but does not differentiate between actual leprosy or exposure to *M. leprae*. This test measures an individual's cellular immune response to a mixture of *M. leprae* antigens. The lepromin test is also not specific to leprosy. False positive results can be developed by individuals without leprosy (74). A T cell immune response measured by interferon gamma (IFN-γ) production has been developed for leprosy diagnosis (79, 80).

A serologic test for routine laboratory diagnosis is not yet available. Polyclonal or monoclonal antibodies reactive to *M. leprae* phenolic glycolipid-I (PGL-I), which is a major lipid component in *M. leprae* cell wall, and other proteins have been used as a serological tool for leprosy diagnosis (81, 82). The development of diagnostic tests for leprosy seems to be difficult because of the diversity of cellular and humoral immune responses observed in leprosy patients and along the clinical spectrum (81, 83). For example, patients with paucibacillary leprosy or the TT form of disease present high levels of secreted IFN-γ *in vitro* with specific *M. leprae* antigen stimulation. Therefore, the tests using antibody against PGL-I may not be useful to these patients. In contrast, patients with multibacillary leprosy or the LL stage of disease do not produce IFN-γ *in vitro*. Therefore, the anti-PGL-I detection or bacterial DNA PCR is beneficial for diagnosis (84).

PCR assays have been developed to detect *M. leprae* and characterize its genotypes (85, 86). PCR assays targeting specific genes or repeat sequences in clinical specimens have high sensitivity and specificity for leprosy diagnosis and have been applied for many types of specimens including; skin smear, nasal smear, skin biopsies, blood, and nerve lesions (87-90). Examples of target genes are those encoding the proline-rich antigen, *M. leprae*-specific repetitive element (RLEP), Ag85B and 16s rRNA (84). PCR-based techniques could be applied for diagnosis in difficult cases such as patients with pure neural symptoms without skin lesions (91). Moreover,

PCR is useful for monitoring treatment outcome (84, 88) and the study of leprosy transmission and household contact surveillance (92, 93). Recently, the GenoType LepraeDR test, which is a reverse hybridization DNA strip test, has been developed as a diagnostic tool to detect antibiotic resistances in clinical specimens. This test was designed to detect mutations in *rpoB, gyrA*, and *folP1* genes which are associated to rifampicin, ofloxacin and dapsone resistance, respectively (85). Although, the PCR-based and reverse transcription PCR-based techniques have 100% specificity to *M. leprae*, their sensitivity varies from 34-80% in paucibacillary patients and greater than 90% in multibacillary patients (90). Currently, none of these diagnostic methods can be solely used for definitive diagnosis of leprosy. The combination of several approaches including clinical manifestations can be a definite approach for leprosy diagnosis.

1.5.5 Leprosy treatment

For therapeutic purpose, leprosy patients are classified based on the number of lesions, which are less or more than five for paucibacillary (PB) and multibacillary (MB) forms, respectively. WHO has provided multidrug therapy (MDT) to leprosy patients in all endemic countries since 1995. The MDT is a combination of three drugs (rifampicin, clofazimine and dapsone) for MB leprosy patients and two drugs (rifampicin and dapsone) for PB leprosy patients.

WHO guidelines for MB leprosy treatment recommend the use of rifampicin (600 mg, once a month), clofazimine (300 mg, once a month and 50 mg daily) and dapsone (100 mg daily) for one year. For PB patients, WHO recommends rifampicin (600 mg, once a month) and dapsone (100 mg daily) for 6 months. For patients with single skin lesion-paucibacillary leprosy, the standard regimen is a single dose of rifampicin (600 mg), ofloxacin (400 mg) and minocycline (100 mg) (94). Patients who are unable to have clofazimine or dapsone are provided second-line

drugs such as fluoroquinolones (ofloxacin, moxifloxacin or pefloxacin), minocycline, and macrolide (clarithromycin) (74). Leprosy bacilli resistant to dapsone have gradually appeared. Resistance to dapsone and rifampicin has been reported among patients who received rifampicin monotherapy or rifampicin with dapsone but relatively rare compared to rifampicin resistance in tueberculosis. Alternative regimens of anti-leprosy drugs should be incorporated for treatment (94). The treatment of leprosy with only one drug is not recommended since it can lead to drug resistance.

The main purpose of leprosy reaction treatment is to control acute inflammation, pain and eye and nerve damage (72). WHO recommends the use of corticosteroid, clofazimine and thalidomide for disease management. Prednisolone regiment which is a potent anti-inflammatory corticosteroid has been suggested for treatment of reversal reaction and neuritis in order to reduce nerve edema, prevent further disabilities and reverse initial nerve impairment. Nevertheless, MDT should be continued during reactional episodes. The control of fever and pain by adequate doses of analgesics is also recommended (94). Since thalidomide has teratogenic effects, WHO reinforces the use of prednisone and clofazimine for treatment of reactions (72).

1.5.6 *M. leprae* genome

The complete genome of *M. leprae* has been sequenced. The 3,268,203 bp length and G+C content of 57.8% (19) is much lower than *M. tuberculosis*, 4,411,532 bp and G+C content 65.61% (20). Only 49.5% of the *M. leprae* genome is predicted to possess coding capacity (1,604 open reading frames) (19). *M. tuberculosis*, in contrast, encodes 3,924 open reading frames with 91% of the genome possessing coding capacity (20). With the *M. leprae* genome, 1,116 reading frames (27% of the genome) are inactive or pseudogenes; a strong indication of reductive evolution (19).

Interestingly, pseudogenes of *M. leprae* are highly expressed and levels are likely altered during infection in the macrophage (95, 96). The low G+C content of *M. leprae* as compared to other mycobacteria may be a reflection of the low quantity of low repetitive DNA in the genome (18). The G+C content of active genes is about 60.1%, which is higher than the G+C content of pseudogenes, which is 56.5% (19).

Mycobacteria have the ability to adapt to and survive in an alteration of their extracellular environments and surrounding conditions by sensory perception. Sigma factors are proteins binding to a promoter and have binding specificity to the RNA polymerase. Most bacteria also encode several alternative sigma factors or extracytoplasmic function **(**ECF) sigma factors to control gene expression in response to various physiological and extracellular changes (97, 98). Each sigma factor regulates transcription of a different gene set (97). *M. leprae* contains nine sigma factor genes but only four functional sigma factors including *sigA* (19). In addition, *M. leprae* genome harbors fewer signal transduction systems than *M. tuberculosis* for response to environmental changes. Specifically, *M. tuberculosis* has 11 two-component regulatory systems and 14 eukaryotic type serine/threonine protein kinases (STPKs) that are involved in signaling networks (20). In contrast, *M. leprae* has only four STPKs (19, 98) and no two-component system has been identified. The reduced number of sensory systems likely reflect *M. leprae*'s narrow environmental niche (14). The loss of sigma factors and two-component systems was also hypothesized as a contribution to pseudogene formation in *M. leprae*, since several sets of genes that are activated by environmental stimuli cannot be expressed in this bacterium (98, 99).

The gene deletion and decay present in *M. leprae* has resulted in several defective in several important metabolic pathways such as sulphur acquisition and reduction, siderophore production, and oxidative stress response. In addition, *M. leprae* has lost most of the microaerophilic and

anaerobic respiratory chains such as formate dehydrogenase, nitrate and fumarate reductase (90). *M. leprae* has a major deficiency to acquire iron from the environment since it lacks the entire *mbt* operon responsible for encoding membrane and secreted forms of mycobactin (90). In addition, *M. leprae* lacks of many genes in catabolic metabolism such as those for lipolysis and acetate consumption (19). The lack of sufficient metacolic pathways leads to a well-known fact that *M. leprae* cannot be cultivated in artificial medium and requires several host factors to survive. Another interesting observation is that *M. leprae* has a large number of pseudogenes for proteins annotated as regulatory proteins (19). Thus, the set of functional genes present in the *M. leprae* genome provides insight into the necessary genes required for *in vivo* survival and pathogenesis in the host.

The genome sequence analysis of *M. leprae* strain TN, Br4923, Thai 53, and NHDP63 from India, Brazil, Thailand and United States, respectively, has shown that the four strains have 99.995% sequence identity and differences in 225 polymorphic sites. These polymorphisms are mainly SNPs and five pseudogenes (100). Recently, *Mycobacterium lepromatosis*, which causes diffuse lepromatous leprosy, has been proposed as a new species based on phylogenic analysis of the 16s rRNA, *rpoB*, and *hsp65* (101). The *M. lepromatosis* genome has 3,206,741 bp with a G+C content of 57.89% that is closely similar to *M. leprae* and has a high number of pseudogenes (1,334) and only 1,477 protein coding sequences. These protein-coding genes and pseudogenes share 93% nucleotide sequence homology and 82% identity to *M. leprae* (102, 103).

1.5.7 Leprosy pathogenesis

M. leprae cannot be cultured *in vitro* and is characterized by a degenerative genome that possesses a large number of pseudogenes (19). Nevertheless, this bacterium has the ability to adapt and survive within different environments of its human host. This includes infection of the upper respiratory tract, skin and peripheral nerves. *M. leprae* prefers areas of the human body with low temperatures, so it tends to locate near the surface of skin (5). Upon entry into the host, *M. leprae* replicates inside intracellular vesicles of macrophages, Schwann cells of peripheral nerves, and endothelial cells. The neuropathy observed in leprosy patients is a result of the infection of peripheral nerve by *M. leprae* in addition to host inflammatory and immunologic responses to the pathogen (90).

The entry of *M. leprae* into host cells starts with the binding of *M. leprae* to host cell receptors. For epithelial cells, *M. leprae* binds to fibronectin, β-integrin 6, and a 25 kDa glycoprotein. For macrophages, the terminal trisaccharide of PGL-1 on *M. leprae* binds to complement receptors CR1, CR4, and parts of CR3, facilitating phagocytosis by the classical complement pathway. In addition, the phagocytosis of *M. leprae* by monocyte-derived macrophages is also regulated by protein kinase (104, 105). This binding mechanism also involves fatty acid side-chains present on the PGL-1 molecule. Phagocytosis of *M. leprae* into macrophages is not associated with any oxidative bursts because of specific binding to the C3 complement (106).

For neural Schwann cells, a major target of infection by *M. leprae*, the possible entry route could be via the vascular endothelium with unknown mechanism (90). Once *M. leprae* accesses the endoneurial compartment, it binds to Schwann cells via several binding molecules leading to destruction of myelin and loss of axonal conductance. Consequently, infected patients lose sensory or motor function of the nerve (107). The conduction of action potentials is dependent on the

myelin sheath, which is a multi-laminar structure around axons and produced by oligodendrocytes and Schwann cells. Demyelinating neuropathy and axonal neuropathy caused by leprosy lead to a decrease in action-potential conduction velocity. Untreated infection can lead to chronic inflammation and fibrosis of nerves resulting in neuropathy and other disabilities (108). Both myelinated and non-myelinated Schwann cells are infected by *M. leprae.* Since *M. leprae* prefers the non-myelinating type of Schwann cell over the myelinated type for invasion and replication, the demyelination provides an environmental niche for *M. leprae* survival and may facilitate the progression of infection (90, 108).

It has been shown that *M. leprae* PGL-1 or LBP21 (laminin-binding 21 kDa protein) ligands on the bacterial surface can bind to the α-2 side chain of laminin-2, which is the major component of the basal lamina of Schwann cell-axon units, as well as the related α-dystroglycan receptor on Schwann cells (39, 109, 110). The α-dystroglycan is a component of the dystroglycan complex involved in the pathogenesis of muscular dystrophies. Specifically, *M. leprae* binds to α-dystroglycan only in the presence of the G domain on the α-2 side chain of laminin-2 and then causes demyelination (110). Furthermore, demyelination induced by *M. leprae* is caused by a direct ligation of bacterium to neuregulin receptor, ErbB2 and Erk1/2 activation, and subsequent mitogen-activated protein (MAP) kinase signaling and proliferation (111).

After binding to host cells, *M. leprae* is taken into host cells by phagocytosis and is encapsulated by a phagosome. The phagocytosis of *M. leprae* is mediated by the actin-dependent protein tyrosine kinase, calcium-dependent protein kinase, and phosphatidylinositol 3-kinase. After phagocytosis, *M. leprae* is killed through fusion with a phagolysosome. It is then digested by protease and oxidizing chemicals. *M. leprae* may be able to impede the lysosome fusion process with an unknown mechanism. Thus, it can survive and replicate inside the phagosome.

Macrophages or Schwann cells can introduce protease directly into the phagosome as a second line response against *M. leprae* (112). Infected Schwann cells are able to process and present antigen to T cells. If the bacterium can evade the destructive mechanisms and replicate inside infected cells without an effective immune response to the pathogen, *M. leprae* numbers will increase in the epineurium and endoneurium of Schwann cells. In addition, *M. leprae* may break out of cells and infect other cells. In contrast, if cellular immunity is effectively developed, a granuloma is developed in the epineurium and endoneurium of Schwann cells and stimulates perineurial fibrosis and thickening. The bacilli are then persistently maintained in the granuloma and may be relatively protected from immunologically mediated destruction (90).

1.5.8 Host immune response to *M. leprae*

The innate immune response is developed early after exposure to the infectious pathogen. Although innate immune response is not specific to a particular pathogen, this arm of immunity is crucial during the first critical hours to days after pathogen exposure to protect the host from an infection (113). Following exposure to *M. leprae*, dendritic cells (DCs), which are major antigen-presenting cells, may be the first cells to confront this pathogen (90). DCs are involved in the induction of cellular responses to intracellular pathogens like *M. leprae* by presenting antigen to T cells leading to the production of cytokines and chemokines (114). DCs have been found to be very effective at presenting *M. leprae* antigens such as PGL-I (115). Langerhans cells, which are a subset of DC cells, initiate immune responses in the epidermis of the skin. High numbers of Langerhans cells are found in TT patients' lesions, coexpressing high levels of CD1a and langerin as pattern recognition receptors (PRR) to recognize and present mycobacterial antigens to T cells. Specifically, Langerhans cells have high efficiency to present non-peptide mycobacterial antigens

to CD1a-restricted T cells. They also express langerin (CD207), which is a C-type PRR, as an important non-peptide antigen-uptake receptor for modulation of T cell responses (116). Furthermore, other PRRs involved in mycobacterial antigen presentation are dendritic cell-specific intercellular adhesion molecule-grabbing nonintegrin (DC-SIGN), which binds mannose-containing antigen such as mannose-capped lipoarabinomannan on pathogens (117). As for cytokine production, patients with TT leprosy have the dominant innate type I cytokines such as interleukin -12, -18 (IL-12, IL-18) and granulocyte-macrophage colony-stimulating factor (GM-CSF) whereas patients with LL leprosy have predominant type II cytokines IL-10 and IL-5 (118).

The adaptive immune response that is developed later is much more specific and effective in eliminating invading pathogens than the innate immune response. Two major classes of adaptive immune response are humoral and cellular-mediated immune response (113). The spectrum of clinical manifestations of leprosy is determined by the immunological response of the host to *M. leprae*. Importantly, *M. leprae* maintains its viability across a spectrum of disease pathology defined by two poles: 1) TT which is typified by a dominant Th1 immune response with vigorous cellular immune response to this mycobacterium resulting in a low number of bacilli, and 2) LL which presents with a high bacterial load and a non-protective, but robust Th2 immune response (5, 75). The TT pole of leprosy has a strong cellular immune response with T-cell proliferation and cytokine response to *M. leprae* antigen. The antibody response to *M. leprae* is absent or low. TT leprosy is infiltrated by CD4+ T cells or Th1-like T cells that produce high levels of Th1 cytokines IFN-γ, Tumor necrosis factor alpha (TNF-α), IL-2, and IL-15 (119, 120). In contrast, the LL pole of leprosy is characterized by an absence of specific cellular immune response and a high level of humoral immune response. LL patients present with a predominance of CD8+ T cells that

produce a type 2 or Th2 cytokine pattern such as IL-4 and IL-10. LL patients have a high titer of antibody to PGL-I and protein antigens specific for *M. leprae* (121).

Leprosy reactions are caused by the dynamic nature of the host immune response to *M. leprae*. Type I leprosy reaction or reversal reaction is a result of a spontaneous increase of CD4+ T cells. The reversal reaction is associated with the infiltration of IFN-γ and TNF-α secreting CD4+ T cells in skin lesions and nerves (120). Moreover, patients with reversal reaction usually have increased cytokine production by peripheral-blood lymphocytes (122). Another reaction is the type II reaction or ENL, which occurs only in BL and LL patients. ENL is a systemic inflammatory response with a high level of pro-inflammatory cytokines such as TNF-α, IL-1β and IL-6 (123, 124).

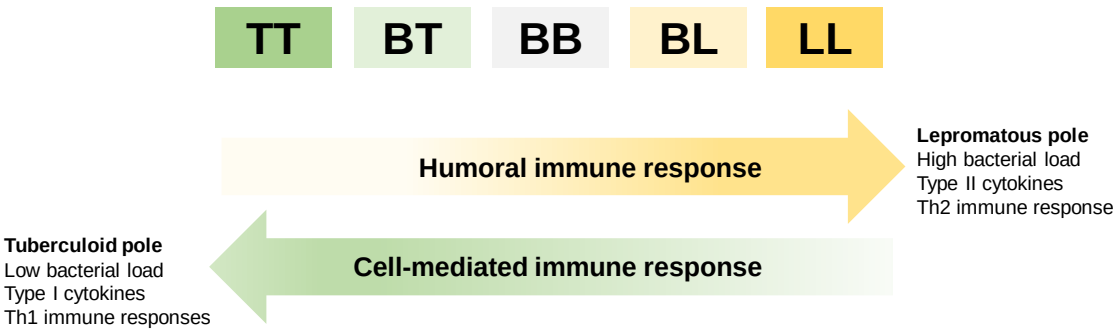


Figure 1-2. Clinical manifestations of leprosy and immune responses
(Drawn based on Ridley, D. S. & Jopling, W. H. (1966) (75))

1.5.9 Animal models for *M. leprae* study

M. leprae has been known to have a long doubling time (~14 days) and has not yet been successfully cultured *in vitro* (7, 16). *M. leprae* can be viable in axenic medium such as 7H12 medium and has a stable metabolic state for a few weeks (17). Therefore, the study of *M. leprae* requires animal models that can provide basic resources for genetic, metabolic and antigenic

studies. Nine-banded armadillo and atymic or gene knockout mice have been utilized as animal models (90, 125, 126). The armadillo is a natural reservoir of *M. leprae* in the environment and can cause zoonotic transmission to humans. Nine-banded armadillo (*Dasypus novemcinctus*) has been used as an experimental host for *M. leprae* study since it has a low core temperature, which facilitates *M. leprae* growth. In addition, it has a large body size and long life span, so it can provide a high number of bacilli. The nine-banded armadillo is also susceptible to several routes of infection, such as intravenous, intradermal, percutaneous and respiratory. The armadillo is an immunologically intact animal and can exhibit the full spectrum of histopathologic symptoms of leprosy including neurological involvement similar to humans (90, 125).

Shepard *et al.* demonstrated that a mouse can be used as an animal model for leprosy (80). *M. leprae* is injected and then multiplies in mouse footpads. The utilization of immunocompetent mice has revealed that the histopathological changes are minor, and disease or granuloma is rarely developed. In contrast, *M. leprae* is able to replicate up to 10^{10} or more bacilli per footpad in atymic *nu/nu* mice lacking a T-cell mediated immune response (90). The alternative mice used in leprosy research are knockout mice that lack genes related to cytokine production. Krahenbuhl *et al*. studied the use of knockout mice that have CGD (chronic granulomatous disease) and knockout mice lacking the ability to produce reactive oxygen intermediates (ROI) and reactive nitrogen intermediates (RNI). These mice could develop the immunological borderline areas of the leprosy spectrum whereas the normal and immunocompromised mice represented only two major poles of leprosy (126). Moreover, IFN-γ knockout mice also showed enlarged mouse footpads and had the borderline clinical features like BB and BL patients (127).

REFERENCES

1. Jorgensen JH, Pfaller MA, Carroll KC, Funke G, Landry ML, Richter SS, Warnock DW. 2015. Manual of Clinical Microbiology, 11th Edition doi:10.1128/9781555817381. American Society of Microbiology.

2. Tortoli E. 2006. The new mycobacteria: an update. *FEMS Immunology & Medical Microbiology* 48:159-178.

3. Stahl DA, Urbance JW. 1990. The Division between fast-growing and slow-growing species corresponds to natural relationships among the mycobacteria. *Journal of Bacteriology* 172:116-124.

4. Brosch R, Gordon SV, Marmiesse M, Brodin P, Buchrieser C, Eiglmeier K, Garnier T, Gutierrez C, Hewinson G, Kremer K, Parsons LM, Pym AS, Samper S, van Soolingen D, Cole ST. 2002. A new evolutionary scenario for the *Mycobacterium tuberculosis* complex. *Proc Natl Acad Sci U S A* 99:3684-3689.

5. Walker SL, Lockwood DN. 2007. Leprosy. *Clin Dermatol* 25:165-172.

6. Johnson PD, Stinear T, Small PL, Pluschke G, Merritt RW, Portaels F, Huygen K, Hayman JA, Asiedu K. 2005. Buruli ulcer (*M. ulcerans* infection): new insights, new hope for disease control. *PLoS Med* 2:e108.

7. Rastogi N, Legrand E, Sola C. 2001. The mycobacteria: an introduction to nomenclature and pathogenesis. *Rev Sci Tech* 20:21-54.

8. Cook JL. 2010. Nontuberculous mycobacteria: opportunistic environmental pathogens for predisposed hosts. *Br Med Bull* 96:45-59.

9. Falkinham JO, 3rd. 2009. Surrounded by mycobacteria: nontuberculous mycobacteria in the human environment. *J Appl Microbiol* 107:356-367.

10. Falkinham JO, 3rd. 1996. Epidemiology of infection by nontuberculous mycobacteria. *Clin Microbiol Rev* 9:177-215.

11. Runyon EH. 1959. Anonymous mycobacteria in pulmonary disease. *Med Clin North Am* 43:273-290.

12. Hett EC, Rubin EJ. 2008. Bacterial growth and cell division: a mycobacterial perspective. *Microbiol Mol Biol Rev* 72:126-156.

13. Rogall T, Wolters J, Flohr T, Bottger EC. 1990. Towards a phylogeny and definition of species at the molecular level within the genus *Mycobacterium*. *Int J Syst Bacteriol* 40:323-330.

14. Cook GM, Berney M, Gebhard S, Heinemann M, Cox RA, Danilchanka O, Niederweis M. 2009. Physiology of mycobacteria. *Adv Microb Physiol* 55:81-182, 318-189.

15. Ginsberg AM, Spigelman M. 2007. Challenges in tuberculosis drug research and development. *Nat Med* 13:290-294.

16. Shepard CC, McRae DH. 1965. *Mycobacterium leprae* in mice: minimal infectious dose, relationship between staining quality and infectivity, and effect of cortisone. *J Bacteriol* 89:365-372.

17. Truman RW, Krahenbuhl JL. 2001. Viable *M. leprae* as a research reagent. *Int J Lepr Other Mycobact Dis* 69:1-12.

18. Brosch R, Gordon SV, Pym A, Eiglmeier K, Garnier T, Cole ST. 2000. Comparative genomics of the mycobacteria. *Int J Med Microbiol* 290:143-152.

19. Cole ST, Eiglmeier K, Parkhill J, James KD, Thomson NR, Wheeler PR, Honore N, Garnier T, Churcher C, Harris D, Mungall K, Basham D, Brown D, Chillingworth T, Connor R, Davies RM, Devlin K, Duthoy S, Feltwell T, Fraser A, Hamlin N, Holroyd S, Hornsby T, Jagels K, Lacroix C, Maclean J, Moule S, Murphy L, Oliver K, Quail MA, Rajandream MA, Rutherford KM, Rutter S, Seeger K, Simon S, Simmonds M, Skelton J, Squares R, Squares S, Stevens K, Taylor K, Whitehead S, Woodward JR, Barrell BG. 2001. Massive gene decay in the leprosy bacillus. *Nature* 409:1007-1011.

20. Cole ST, Brosch R, Parkhill J, Garnier T, Churcher C, Harris D, Gordon SV, Eiglmeier K, Gas S, Barry CE, 3rd, Tekaia F, Badcock K, Basham D, Brown D, Chillingworth T, Connor R, Davies R, Devlin K, Feltwell T, Gentles S, Hamlin N, Holroyd S, Hornsby T, Jagels K, Krogh A, McLean J, Moule S, Murphy L, Oliver K, Osborne J, Quail MA, Rajandream MA, Rogers J, Rutter S, Seeger K, Skelton J, Squares R, Squares S, Sulston JE, Taylor K, Whitehead S, Barrell BG. 1998. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* 393:537-544.

21. Ji Y-e, Colston MJ, Cox RA. 1994. The ribosomal RNA (*rrn*) operons of fast-growing mycobacteria: primary and secondary structures and their relation to rrn operons of pathogenic slow-growers. *Microbiology* 140:2829-2840.

22. Gonzalez-y-Merchand JA, Colston MJ, Cox RA. 1996. The rRNA operons of *Mycobacterium smegmatis* and *Mycobacterium tuberculosis*: Comparison of promoter elements and of neighbouring upstream genes. *Microbiology* 142:667-674.

23. Silhavy TJ, Kahne D, Walker S. 2010. The bacterial cell envelope. *Cold Spring Harb Perspect Biol* 2:a000414.

24. Daffe M, Draper P. 1998. The envelope layers of mycobacteria with reference to their pathogenicity. *Adv Microb Physiol* 39:131-203.

25. Pitulle C, Dorsch M, Kazda J, Wolters J, Stackebrandt E. 1992. Phylogeny of rapidly growing members of the genus *Mycobacterium*. *Int J Syst Bacteriol* 42:337-343.

26. Fu LM, Fu-Liu CS. 2002. Is *Mycobacterium tuberculosis* a closer relative to Gram-positive or Gram-negative bacterial pathogens? *Tuberculosis (Edinb)* 82:85-90.

27. Brennan PJ, Nikaido H. 1995. The envelope of mycobacteria. *Annu Rev Biochem* 64:29-63.

28. Ortalo-Magne A, Lemassu A, Laneelle MA, Bardou F, Silve G, Gounon P, Marchal G, Daffe M. 1996. Identification of the surface-exposed lipids on the cell envelopes of *Mycobacterium tuberculosis* and other mycobacterial species. *J Bacteriol* 178:456-461.

29. Brennan PJ. 2003. Structure, function, and biogenesis of the cell wall of *Mycobacterium tuberculosis*. *Tuberculosis (Edinb)* 83:91-97.

30. Lederer E, Adam A, Ciorbaru R, Petit JF, Wietzerbin J. 1975. Cell walls of mycobacteria and related organisms; chemistry and immunostimulant properties. *Mol Cell Biochem* 7:87-104.

31. Dmitriev BA, Ehlers S, Rietschel ET, Brennan PJ. 2000. Molecular mechanics of the mycobacterial cell wall: from horizontal layers to vertical scaffolds. *Int J Med Microbiol* 290:251-258.

32. Mahapatra S, Scherman H, Brennan PJ, Crick DC. 2005. N Glycolylation of the nucleotide precursors of peptidoglycan biosynthesis of *Mycobacterium* spp. is altered by drug treatment. *Journal of Bacteriology* 187:2341-2347.

33. Mahapatra S, Crick DC, McNeil MR, Brennan PJ. 2008. Unique structural features of the peptidoglycan of *Mycobacterium leprae*. *Journal of Bacteriology* 190:655-661.

34. Draper P, Kandler O, Darbre A. 1987. Peptidoglycan and arabinogalactan of *Mycobacterium leprae*. *J Gen Microbiol* 133:1187-1194.

35. Mahapatra S, Yagi T, Belisle JT, Espinosa BJ, Hill PJ, McNeil MR, Brennan PJ, Crick DC. 2005. Mycobacterial lipid II is composed of a complex mixture of modified muramyl and peptide moieties linked to decaprenyl phosphate. *Journal of Bacteriology* 187:2747-2757.

36. Kumar P, Arora K, Lloyd JR, Lee IY, Nair V, Fischer E, Boshoff HI, Barry CE, 3rd. 2012. Meropenem inhibits D,D-carboxypeptidase activity in *Mycobacterium tuberculosis*. *Mol Microbiol* 86:367-381.

37. McNeil M, Daffe M, Brennan PJ. 1990. Evidence for the nature of the link between the arabinogalactan and peptidoglycan of mycobacterial cell walls. *Journal of Biological Chemistry* 265:18200-18206.

38. Takayama K, Wang C, Besra GS. 2005. Pathway to synthesis and processing of mycolic acids in *Mycobacterium tuberculosis*. *Clin Microbiol Rev* 18:81-101.

39. Ng V, Zanazzi G, Timpl R, Talts JF, Salzer JL, Brennan PJ, Rambukkana A. 2000. Role of the cell wall phenolic glycolipid-1 in the peripheral nerve predilection of *Mycobacterium leprae*. *Cell* 103:511-524.

40. Nigou J, Gilleron M, Puzo G. 2003. Lipoarabinomannans: from structure to biosynthesis. *Biochimie* 85:153-166.

41. Chatterjee D, Lowell K, Rivoire B, McNeil MR, Brennan PJ. 1992. Lipoarabinomannan of *Mycobacterium tuberculosis.* Capping with mannosyl residues in some strains. *J Biol Chem* 267:6234-6239.

42. Niederweis M. 2003. Mycobacterial porins-new channel proteins in unique outer membranes. *Mol Microbiol* 49:1167-1177.

43. Wolschendorf F, Mahfoud M, Niederweis M. 2007. Porins are required for uptake of phosphates by *Mycobacterium smegmatis*. *J Bacteriol* 189:2435-2442.

44. Raynaud C, Papavinasasundaram KG, Speight RA, Springer B, Sander P, Bottger EC, Colston MJ, Draper P. 2002. The functions of OmpATb, a pore-forming protein of *Mycobacterium tuberculosis*. *Mol Microbiol* 46:191-201.

45. Daffé M, Etienne G. 1999. The capsule of *Mycobacterium tuberculosis* and its implications for pathogenicity. *Tubercle and Lung Disease* 79:153-169.

46. de Carvalho LP, Fischer SM, Marrero J, Nathan C, Ehrt S, Rhee KY. 2010. Metabolomics of *Mycobacterium tuberculosis* reveals compartmentalized co-catabolism of carbon substrates. *Chem Biol* 17:1122-1131.

47. Gouzy A, Poquet Y, Neyrolles O. 2014. Nitrogen metabolism in *Mycobacterium tuberculosis* physiology and virulence. *Nat Rev Microbiol* 12:729-737.

48. Peirs P, Lefevre P, Boarbi S, Wang XM, Denis O, Braibant M, Pethe K, Locht C, Huygen K, Content J. 2005. *Mycobacterium tuberculosis* with disruption in genes encoding the phosphate binding proteins PstS1 and PstS2 is deficient in phosphate uptake and demonstrates reduced *in vivo* virulence. *Infection and Immunity* 73:1898-1902.

49. Rifat D, Bishai WR, Karakousis PC. 2009. Phosphate depletion: a novel trigger for *Mycobacterium tuberculosis* persistence. *J Infect Dis* 200:1126-1135.

50. Stewart GR, Wernisch L, Stabler R, Mangan JA, Hinds J, Laing KG, Young DB, Butcher PD. 2002. Dissection of the heat-shock response in *Mycobacterium tuberculosis* using mutants and microarrays. *Microbiology* 148:3129-3138.

51. Rodriguez GM, Smith I. 2003. Mechanisms of iron regulation in mycobacteria: role in physiology and virulence. *Mol Microbiol* 47:1485-1494.

52. Park HD, Guinn KM, Harrell MI, Liao R, Voskuil MI, Tompa M, Schoolnik GK, Sherman DR. 2003. Rv3133c/*dosR* is a transcription factor that mediates the hypoxic response of *Mycobacterium tuberculosis*. *Molecular Microbiology* 48:833-843.

53. Rustad TR, Harrell MI, Liao R, Sherman DR. 2008. The enduring hypoxic response of *Mycobacterium tuberculosis*. *PLoS One* 3:e1502.

54. Banaiee N, Jacobs WR, Jr., Ernst JD. 2006. Regulation of *Mycobacterium tuberculosis whiB3* in the mouse lung and macrophages. *Infect Immun* 74:6449-6457.

55. Wayne LG, Hayes LG. 1996. An *in vitro* model for sequential study of shiftdown of *Mycobacterium tuberculosis* through two stages of nonreplicating persistence. *Infection and Immunity* 64:2062-2069.

56. Wayne LG, Sohaskey CD. 2001. Nonreplicating persistence of *Mycobacterium tuberculosis*. *Annu Rev Microbiol* 55:139-163.

57. Gengenbacher M, Kaufmann SHE. 2012. *Mycobacterium tuberculosis*: success through dormancy. *FEMS Microbiology Reviews* 36:514-532.

58. Amer AO, Swanson MS. 2002. A phagosome of one's own: a microbial guide to life in the macrophage. *Curr Opin Microbiol* 5:56-61.

59. Monot M, Honoré N, Garnier T, Araoz R, Coppée J-Y, Lacroix C, Sow S, Spencer JS, Truman RW, Williams DL, Gelber R, Virmond M, Flageul B, Cho S-N, Ji B, Paniz-Mondolfi A, Convit J, Young S, Fine PE, Rasolofo V, Brennan PJ, Cole ST. 2005. On the origin of leprosy. *Science* 308:1040-1042.

60. WHO. 2015. Leprosy Fact sheet. http://www.who.int/mediacentre/factsheets/fs101/en/. Accessed

61. WHO. 2015. Weekly Epidemiological Record (WER) 36:461-476.

62. Rees RJ, McDougall AC. 1977. Airborne infection with *Mycobacterium leprae* in mice. *J Med Microbiol* 10:63-68.

63. Chehl S, Job CK, Hastings RC. 1985. Transmission of leprosy in nude mice. *Am J Trop Med Hyg* 34:1161-1166.

64. Job CK, Jayakumar J, Kearney M, Gillis TP. 2008. Transmission of leprosy: A study of skin and nasal secretions of household contacts of leprosy patients using PCR. *American Journal of Tropical Medicine and Hygiene* 78:518-521.

65. Truman R, Fine PE. 2010. 'Environmental' sources of *Mycobacterium leprae*: issues and evidence. *Lepr Rev* 81:89-95.

66. Loughry WJ, Truman RW, McDonough CM, Tilak MK, Garnier S, Delsuc F. 2009. Is leprosy spreading among nine-banded armadillos in the southeastern United States? *J Wildl Dis* 45:144-152.

67. Meyers WM, Gormus BJ, Walsh GP, Baskin GB, Hubbard GB. 1991. Naturally acquired and experimental leprosy in nonhuman primates. *Am J Trop Med Hyg* 44:24-27.

68. Truman R. 2005. Leprosy in wild armadillos. *Lepr Rev* 76:198-208.

69. Truman RW, Singh P, Sharma R, Busso P, Rougemont J, Paniz-Mondolfi A, Kapopoulou A, Brisse S, Scollard DM, Gillis TP, Cole ST. 2011. Probable zoonotic leprosy in the southern United States. *N Engl J Med* 364:1626-1633.

70. Wheat WH, Casali AL, Thomas V, Spencer JS, Lahiri R, Williams DL, McDonnell GE, Gonzalez-Juarrero M, Brennan PJ, Jackson M. 2014. Long-term survival and virulence of *Mycobacterium leprae* in amoebal cysts. *PLoS Negl Trop Dis* 8:e3405.

71. Desikan KV, Sreevatsa. 1995. Extended studies on the viability of *Mycobacterium leprae* outside the human body. *Leprosy review* 66:287-295.

72. Pinheiro RO, de Souza Salles J, Sarno EN, Sampaio EP. 2011. *Mycobacterium leprae*-host-cell interactions and genetic determinants in leprosy: an overview. *Future Microbiol* 6:217-230.

73. Bhat RM, Prakash C. 2012. Leprosy: an overview of pathophysiology. *Interdiscip Perspect Infect Dis* 2012:181089.

74. Britton WJ, Lockwood DN. 2004. Leprosy. *Lancet* 363:1209-1219.

75. Ridley DS, Jopling WH. 1966. Classification of leprosy according to immunity. A five-group system. *Int J Lepr Other Mycobact Dis* 34:255-273.

76. Van Voorhis WC, Kaplan G, Sarno EN, Horwitz MA, Steinman RM, Levis WR, Nogueira N, Hair LS, Gattass CR, Arrick BA, Cohn ZA. 1982. The cutaneous infiltrates of leprosy: cellular characteristics and the predominant T-cell phenotypes. *N Engl J Med* 307:1593-1597.

77. WHO. WHO Guidelines for the management of severe erythema nodosum leprosum (ENL) reactions.

78. Walker SL, Lockwood DN. 2006. The clinical and immunological features of leprosy. *Br Med Bull* 77-78:103-121.

79. Spencer JS, Dockrell HM, Kim HJ, Marques MA, Williams DL, Martins MV, Martins ML, Lima MC, Sarno EN, Pereira GM, Matos H, Fonseca LS, Sampaio EP, Ottenhoff TH, Geluk A, Cho SN, Stoker NG, Cole ST, Brennan PJ, Pessolani MC. 2005. Identification of specific proteins and peptides in *Mycobacterium leprae* suitable for the selective diagnosis of leprosy. *J Immunol* 175:7930-7938.

80. Shepard CC. 1960. The experimental disease that follows the injection of human leprosy bacilli into foot-pads of mice. *J Exp Med* 112:445-454.

81. Spencer JS, Kim HJ, Wheat WH, Chatterjee D, Balagon MV, Cellona RV, Tan EV, Gelber R, Saunderson P, Duthie MS, Reece ST, Burman W, Belknap R, Mac Kenzie WR, Geluk A, Oskam L, Dockrell HM, Brennan PJ. 2011. Analysis of antibody responses to *Mycobacterium leprae* phenolic glycolipid I, lipoarabinomannan, and recombinant proteins to define disease subtype-specific antigenic profiles in leprosy. *Clin Vaccine Immunol* 18:260-267.

82. Duthie MS, Truman RW, Goto W, O'Donnell J, Hay MN, Spencer JS, Carter D, Reed SG. 2011. Insight toward early diagnosis of leprosy through analysis of the developing antibody responses of *Mycobacterium leprae*-infected armadillos. *Clin Vaccine Immunol* 18:254-259.

83. Roche PW, Britton WJ, Failbus SS, Ludwig H, Theuvenet WJ, Adiga RB. 1990. Heterogeneity of serological responses in paucibacillary leprosy-differential responses to protein and carbohydrate antigens and correlation with clinical parameters. *Int J Lepr Other Mycobact Dis* 58:319-327.

84. Martinez AN, Talhari C, Moraes MO, Talhari S. 2014. PCR-based techniques for leprosy diagnosis: From the laboratory to the clinic. *PLoS Negl Trop Dis* 8:e2655.

85. Cambau E, Chauffour-Nevejans A, Tejmar-Kolar L, Matsuoka M, Jarlier V. 2012. Detection of antibiotic resistance in leprosy using GenoType LepraeDR, a novel ready-to-use molecular test. *PLoS Negl Trop Dis* 6:e1739.

86. Sakamuri RM, Kimura M, Li W, Kim HC, Lee H, Kiran MD, Black WCt, Balagon M, Gelber R, Cho SN, Brennan PJ, Vissa V. 2009. Population-based molecular epidemiology of leprosy in Cebu, Philippines. *J Clin Microbiol* 47:2844-2854.

87. Williams DL, Gillis TP, Booth RJ, Looker D, Watson JD. 1990. The use of a specific DNA probe and polymerase chain reaction for the detection of *Mycobacterium leprae*. *J Infect Dis* 162:193-200.

88. Kampirapap K, Singtham N, Klatser PR, Wiriyawipart S. 1998. DNA amplification for detection of leprosy and assessment of efficacy of leprosy chemotherapy. *Int J Lepr Other Mycobact Dis* 66:16-21.

89. Martinez AN, Ribeiro-Alves M, Sarno EN, Moraes MO. 2011. Evaluation of qPCR-based assays for leprosy diagnosis directly in clinical specimens. *PLoS Negl Trop Dis* 5:e1354.

90. Scollard DM, Adams LB, Gillis TP, Krahenbuhl JL, Truman RW, Williams DL. 2006. The continuing challenges of leprosy. *Clinical Microbiology Reviews* 19:338-381.

91. Bezerra Da Cunha FM, Werneck MC, Scola RH, Werneck LC. 2006. Pure neural leprosy: diagnostic value of the polymerase chain reaction. *Muscle Nerve* 33:409-414.

92. de Wit MY, Douglas JT, McFadden J, Klatser PR. 1993. Polymerase chain reaction for detection of *Mycobacterium leprae* in nasal swab specimens. *J Clin Microbiol* 31:502-506.

93. Almeida EC, Martinez AN, Maniero VC, Sales AM, Duppre NC, Sarno EN, Santos AR, Moraes MO. 2004. Detection of *Mycobacterium leprae* DNA by polymerase chain reaction in the blood and nasal secretion of Brazilian household contacts. *Mem Inst Oswaldo Cruz* 99:509-511.

94. WHO. 1994. Chemotherapy of leprosy.

95. Suzuki K, Nakata N, Bang PD, Ishii N, Makino M. 2006. High-level expression of pseudogenes in *Mycobacterium leprae*. *FEMS Microbiol Lett* 259:208-214.

96. Williams DL, Slayden RA, Amin A, Martinez AN, Pittman TL, Mira A, Mitra A, Nagaraja V, Morrison NE, Moraes M, Gillis TP. 2009. Implications of high level pseudogene transcription in *Mycobacterium leprae*. *BMC Genomics* 10:397.

97. Missiakas D, Raina S. 1998. The extracytoplasmic function sigma factors: role and regulation. *Mol Microbiol* 28:1059-1066.

98. Tyagi JS, Saini DK. 2004. Did the loss of two-component systems initiate pseudogene accumulation in *Mycobacterium leprae*? *Microbiology* 150:4-7.

99. Madan Babu M. 2003. Did the loss of sigma factors initiate pseudogene accumulation in *M. leprae*? *Trends Microbiol* 11:59-61.

100. Monot M, Honore N, Garnier T, Zidane N, Sherafi D, Paniz-Mondolfi A, Matsuoka M, Taylor GM, Donoghue HD, Bouwman A, Mays S, Watson C, Lockwood D, Khamesipour A, Dowlati Y, Jianping S, Rea TH, Vera-Cabrera L, Stefani MM, Banu S, Macdonald M, Sapkota BR, Spencer JS, Thomas J, Harshman K, Singh P, Busso P, Gattiker A, Rougemont J, Brennan PJ, Cole ST. 2009. Comparative genomic and phylogeographic analysis of *Mycobacterium leprae*. *Nat Genet* 41:1282-1289.

101. Han XY, Seo YH, Sizer KC, Schoberle T, May GS, Spencer JS, Li W, Nair RG. 2008. A new *Mycobacterium* species causing diffuse lepromatous leprosy. *Am J Clin Pathol* 130:856-864.

102. Singh P, Benjak A, Schuenemann VJ, Herbig A, Avanzi C, Busso P, Nieselt K, Krause J, Vera-Cabrera L, Cole ST. 2015. Insight into the evolution and origin of leprosy bacilli from the genome sequence of *Mycobacterium lepromatosis*. *Proc Natl Acad Sci U S A* 112:4459-4464.

103. Han XY, Mistry NA, Thompson EJ, Tang HL, Khanna K, Zhang L. 2015. Draft genome sequence of new leprosy agent *Mycobacterium lepromatosis*. *Genome Announc* 3:e00513-00515.

104. Schlesinger LS, Horwitz MA. 1991. Phagocytosis of *Mycobacterium leprae* by human monocyte-derived macrophages is mediated by complement receptors CR1 (CD35), CR3 (CD11b/CD18), and CR4 (CD11c/CD18) and IFN-gamma activation inhibits complement receptor function and phagocytosis of this bacterium. *J Immunol* 147:1983-1994.

105. Prabhakaran K, Harris EB, Randhawa B. 2000. Regulation by protein kinase of phagocytosis of *Mycobacterium leprae* by macrophages. *J Med Microbiol* 49:339-342.

106. Schlesinger LS, Horwitz MA. 1991. Phenolic glycolipid-1 of *Mycobacterium leprae* binds complement component C3 in serum and mediates phagocytosis by human monocytes. *J Exp Med* 174:1031-1038.

107. Rambukkana A, Zanazzi G, Tapinos N, Salzer JL. 2002. Contact-dependent demyelination by *Mycobacterium leprae* in the absence of immune cells. *Science* 296:927-931.

108. Rambukkana A. 2004. *Mycobacterium leprae*-induced demyelination: a model for early nerve degeneration. *Current Opinion in Immunology* 16:511-518.

109. Marques MA, Ant nio VL, Sarno EN, Brennan PJ, Pessolani MC. 2001. Binding of α2-laminins by pathogenic and non-pathogenic mycobacteria and adherence to Schwann cells. *J Med Microbiol* 50:23-28.

110. Rambukkana A, Yamada H, Zanazzi G, Mathus T, Salzer JL, Yurchenco PD, Campbell KP, Fischetti VA. 1998. Role of alpha-dystroglycan as a Schwann cell receptor for *Mycobacterium leprae*. *Science* 282:2076-2079.

111. Tapinos N, Ohnishi M, Rambukkana A. 2006. ErbB2 receptor tyrosine kinase signaling mediates early demyelination induced by leprosy bacilli. *Nat Med* 12:961-966.

112. Steinhoff U, Golecki JR, Kazda J, Kaufmann SH. 1989. Evidence for phagosome-lysosome fusion in *Mycobacterium leprae*-infected murine Schwann cells. *Infect Immun* 57:1008-1010.

113. Medzhitov R, Janeway Jr CA. 1997. Innate immunity: impact on the adaptive immune response. *Current Opinion in Immunology* 9:4-9.

114. Demangel C, Britton WJ. 2000. Interaction of dendritic cells with mycobacteria: where the action starts. *Immunol Cell Biol* 78:318-324.

115. Hashimoto K, Maeda Y, Kimura H, Suzuki K, Masuda A, Matsuoka M, Makino M. 2002. *Mycobacterium leprae* infection in monocyte-derived dendritic cells and its influence on antigen-presenting function. *Infect Immun* 70:5167-5176.

116. Hunger RE, Sieling PA, Ochoa MT, Sugaya M, Burdick AE, Rea TH, Brennan PJ, Belisle JT, Blauvelt A, Porcelli SA, Modlin RL. 2004. Langerhans cells utilize CD1a and langerin to efficiently present nonpeptide antigens to T cells. *J Clin Invest* 113:701-708.

117. Tailleux L, Schwartz O, Herrmann JL, Pivert E, Jackson M, Amara A, Legres L, Dreher D, Nicod LP, Gluckman JC, Lagrange PH, Gicquel B, Neyrolles O. 2003. DC-SIGN is the major *Mycobacterium tuberculosis* receptor on human dendritic cells. *J Exp Med* 197:121-127.

118. Montoya D, Modlin RL. 2010. Learning from leprosy: insight into the human innate immune response. *Adv Immunol* 105:1-24.

119. Jullien D, Sieling PA, Uyemura K, Mar ND, Rea TH, Modlin RL. 1997. IL-15, an immunomodulator of T cell responses in intracellular infection. *J Immunol* 158:800-806.

120. Little D, Khanolkar-Young S, Coulthart A, Suneetha S, Lockwood DN. 2001. Immunohistochemical analysis of cellular infiltrate and gamma interferon, interleukin-12, and inducible nitric oxide synthase expression in leprosy type 1 (reversal) reactions before and during prednisolone treatment. *Infect Immun* 69:3413-3417.

121. Cho SN, Cellona RV, Villahermosa LG, Fajardo TT, Jr., Balagon MV, Abalos RM, Tan EV, Walsh GP, Kim JD, Brennan PJ. 2001. Detection of phenolic glycolipid I of *Mycobacterium leprae* in sera from leprosy patients before and after start of multidrug therapy. *Clin Diagn Lab Immunol* 8:138-142.

122. Manandhar R, Shrestha N, Butlin CR, Roche PW. 2002. High levels of inflammatory cytokines are associated with poor clinical response to steroid treatment and recurrent episodes of type 1 reactions in leprosy. *Clin Exp Immunol* 128:333-338.

123. Sarno EN, Grau GE, Vieira LM, Nery JA. 1991. Serum levels of tumour necrosis factor-alpha and interleukin-1 beta during leprosy reactional states. *Clin Exp Immunol* 84:103-108.

124. Khanolkar-Young S, Rayment N, Brickell PM, Katz DR, Vinayakumar S, Colston MJ, Lockwood DN. 1995. Tumour necrosis factor-alpha (TNF-alpha) synthesis is associated with the skin and peripheral nerve pathology of leprosy reversal reactions. *Clin Exp Immunol* 99:196-202.

125. Balamayooran G, Pena M, Sharma R, Truman RW. 2015. The armadillo as an animal model and reservoir host for *Mycobacterium leprae*. *Clin Dermatol* 33:108-115.

126. Krahenbuhl J, Adams LB. 2000. Exploitation of gene knockout mice models to study the pathogenesis of leprosy. *Lepr Rev* 71 Suppl:S170-175.

127. Adams LB, Scollard DM, Ray NA, Cooper AM, Frank AA, Orme IM, Krahenbuhl JL. 2002. The study of *Mycobacterium leprae* infection in interferon-gamma gene-disrupted mice as a model to explore the immunopathologic spectrum of leprosy. *J Infect Dis* 185 Suppl 1:S1-8.

CHAPTER 2 INTRODUCTION TO CYCLIC NUCLEOTIDES

2.1 Bacterial signaling

Alterations in extracellular environment and intracellular physiological status cause bacteria to adapt and survive in a given situation. Bacteria sense changes in extracellular and/or intracellular parameters through the signal transduction machinery. Once a change is detected, bacteria transmit signals to cellular components that mediate an appropriate alteration in bacterial physiology, metabolism, and cell behaviors (1).

Bacterial signaling can be classified into extracellular, transmembrane and intracellular processes (2). The main extracellular signaling is cell-cell communication or quorum sensing (QS). This type of intercellular communication can occur in the same or different species of bacteria (3). Quorum sensing is cell-cell communication used by bacteria to regulate their gene expression via the production, detection, and response to chemical signal molecules called autoinducers that accumulate extracellularly and intracellularly in a cell-density dependent manner. The detection of accumulated signal molecules occurs at a minimal threshold stimulatory level (3) initiating signal transduction cascades. Bacteria utilize quorum sensing to control several cellular activities such as biofilm formation, virulence factor secretion, sporulation and conjugation (4-7). Two major types of autoinducers include acyl homoserine lactones (AHLs) used by Gram-negative bacteria (8), and modified oligopeptides used by Gram-positive bacteria (9) (Figure 2-1). Some bacteria such as *Vibrio harveyi* and *Streptococcus pyogenes* can utilize a particular autoincuder designated as AI-2 for quorum sensing (10, 11). The AI-2 has been revealed as 3(2H) class of furanones (12).

Another type of bacterial signaling mechanism occurres via transmembrane. Generally, transmembrane signaling in bacteria is a two-component system that transduces extracellular signal into the intracellular compartment. The two-component system is comprised by a histidine sensor kinase receptors, which respond to particular signals and then phosphorylates a cognate response regulators. The integrated-membrane histidine receptor is a sensor for extracellular stimuli and transfers the information across the membrane. A response regulator then transmits the signal originating from the kinase to regulate transcription of target genes or function of target proteins (13). The histidine sensor kinase autophosphorylates at a conserved histidine residue and relays the phosphoryl group to a conserved aspartate residues in the response regulator. As a result, the phosphorylated response regulator mediates the output responses. Response regulators are classified into several groups based on effector domains, structures, and functions. They can bind to DNA or RNA, have enzymatic activity, or be involved in protein-protein interactions (14).

In addition, bacteria are able to encode sigma factors called extra-cytoplasmic sigma factors (ECF) that respond to specific environmental stimuli or a changing environment. Sigma factors locate in outer membrane or periplasm of bacterial cell wall. They are often cotranscribed with anti-sigma factors that bind and inhibit their cognate sigma factors. Once the sensor of anti-sigma factor receives environmental stimuli, the sigma factor is released and binds to RNA polymerase (15). These sigma factors are involved in the recruitment of RNA polymerase to the appropriate response genes and regulate the expression of proteins (16) (Figure 2-1).

Intracellular signaling involves in biochemical reactions that are mediated by enzymatic activity or secondary messengers resulting in an alteration of gene expression or protein activity. Examples of bacterial intracellular signaling machinery include the response regulator of the two-component system, transcriptional regulators (13), and small molecules functioning as secondary

messengers such as cyclic adenosine 3’5’-monophosphate (cAMP), cyclic diguanosine monophosphate (c-di-GMP), and ppGpp (17). Moreover, eukaryotic-like serine/threonine kinases and phosphatases have been identified in bacteria. Phosphorylation is a crucial mechanism to control protein activity that consequently regulates cellular functions. After sensing stimuli, the phosphorylation of specific amino acid residues (most commonly serine, threonine, tyrosine, histidine and aspartate) is able to control protein activities by activating conformational changes in the active site or regulating protein-protein interactions (18). Bacterial signaling systems are shown in Figure 2-1.

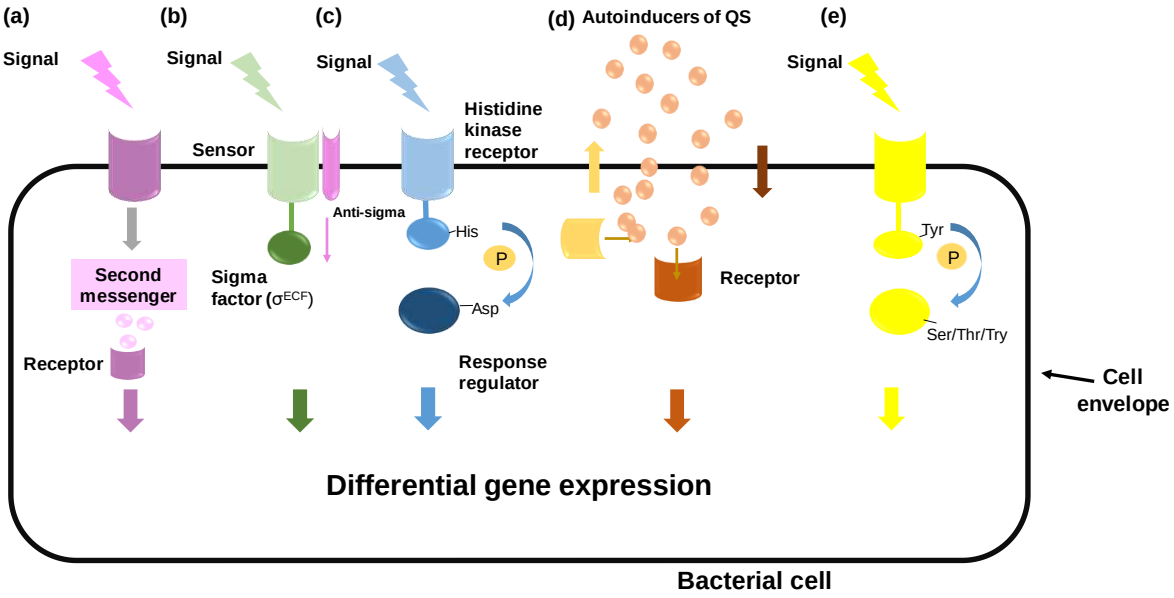


Figure 2-1. Bacterial signaling systems
Several bacterial signaling systems sense and respond to environmental stimuli and consequently regulate expression of genes; (a) nucleotide-based second messenger, (b) extracytoplasmic sigma factors, (c) two-component system, (d) quorum sensing, and (e) eukaryotic-like serine/threonine protein kinase.

2.1.1 Nucleotide-based second messengers

Signal transduction mediates bacterial recognition and response to environmental changes. Second messengers are molecules that transmit the signals received by receptors on the cell surface to intracellular effector-protein targets (19). In addition, the second messenger can serve as an amplifier to greatly enhance the signal strength and trigger the alteration of biochemical activities within the cells. Second messenger systems are able to integrate multiple sensory inputs, and offer flexibility in recognition and response.

2.1.1.1 Nucleotide-based second messengers in eukaryotic cells

In eukaryotic cells, especially human, the second messenger can be classified into three major groups: cyclic nucleotide-based second messengers, inositol trisphosphate (IP3) and diacylglycerol (DAG), and calcium ion ($Ca2^{+}$) (20). The signaling compound that is shared between eukaryotes and prokaryotes is cyclic nucleotides-based second messenger.

Cyclic adenosine 3’5’ monophosphate (cAMP) and cyclic guanosine 3’5’ monophosphate (cGMP) are the important second messengers in higher eukaryotes. They regulate important functions such as muscle contraction, vision, sleep, and memory by mediating effects of light, hormones, nitric oxide, and other signals (21). cAMP and cGMP are synthesized by adenylyl and guanylyl cyclases, respectively, and degraded by phosphodiesterase (PDE) specific to cAMP or cGMP. These enzymes are part of the class III adenylate/guanylate cyclase family. The substrates for cAMP and cGMP synthesis are ATP and GTP, respectively. The specificity of these cyclases for ATP or GTP is determined by a few residues in the substrate-binding pocket (22).

Recently, the discovery of a eukaryotic cyclic dinucleotide second messenger, cyclic GMP-AMP hybrid molecule (cGAMP), has brought new insight in the second messenger field. The 2’3’-

cGAMP is synthesized in response to the presence of microbial or self DNA in the cytosol of mammalian cells. Moreover, the STING (stimulation of interferon gene)-stimulating signals such as c-di-GMP, cyclic dimeric adenosine monophosphate (c-di-AMP), and DNA can mediate the synthesis of cGAMP. The presence of these microbial compounds or DNA in mammalian cell cytoplasm is detected by the DNA sensor cGAMP synthase (cGAS), which catalyzes the production of cGAMP to activate an innate immune response characterized by the STING-synthesizing type I interferon and inflammatory cytokines (23, 24). cGAMP has also been discovered in bacteria. *V. cholerae* is able to produce a hybrid 3'3'-cGAMP molecule from GTP and ATP by using dinucleotide cyclase DncV. The 3'3'-cGAMP is responsible for bacterial virulence including intestinal colonization, regulation of chemotaxis, and a phenotype associated with hyperinfectivity (25). In addition, the environmental bacterium *Geobacter sulfurreducens* is able to produce cGAMP as well. GEMM-Ib class riboswitches (Genes for the Environment, Membranes, and Motility) that regulate genes related to extracellular electron transfer have been identified as the receptor of cGAMP in *G. sulfurreducens* (26).

Although c-di-GMP is mainly synthesized and found in prokaryotes as a signaling molecule, this compound is reported to be synthesized by eukaryotes as well. The c-di-GMP can modulate cellular activity in eukaryotes, including immunologic responses as well as physiological cellular activity. Interestingly, whole genome analysis reveals that *Dictyostelium discoideum* amoeba harbors a prokaryotic gene encoding diguanylate cyclase to synthesize c-di-GMP in its genome. Disruption of diguanylate cyclase leads to a defect in stalk formation preventing development of a fruiting body and decreasing sessility. Moreover, depletion of c-di-GMP can be replaced by secreted c-di-GMP of wild-type cells and by c-di-GMP from a distantly related dictyostelid (27).

2.1.1.2 Nucleotide-based second messengers in bacteria

In bacteria, two common cyclic nucleotide-based second messengers are cAMP and guanosine-3,5-bis(pyrophosphate) or guanosine tetraphosphate (ppGpp). cAMP is a universal second messenger that is used by a variety of organisms including mammals, fungi, protozoa and bacteria (28). cAMP is synthesized from ATP by adenylate cyclases (Class III nucleotide cyclases) and is degraded by PDE (19). cAMP mediates regulatory effects through binding and allosteric interaction with a transcription factor of the cAMP-receptor protein (CRP) family or catabolite gene-activator protein (CAP). In *Escherichia coli*, the cAMP-CRP complex mediates glucose response or catabolite regulation protein regulating catabolic operons for the utilization of alternative carbon sources (29). The cAMP-CRP complex also plays a role in controlling other cellular processes such as biofilm-regulation, type III-secretion, and virulence gene-expression in many bacteria (28).

cAMP is a signaling molecule that is present in mycobacteria. Genome analysis revealed that *M. tuberculosis*, *M. avium* and *M. leprae* contain 16, 12 and four adenylyl cyclases that are classified as class III nucleotide cyclases (30, 31). *M. tuberculosis* is known to have ten active adenylyl cyclases of the existing 16 adenylyl cyclases, but only one PDE (Rv0805) to degrade cAMP (31, 32). Importantly, the homolog of *M. tuberculosis* Rv0805 is present in the genome of *M. leprae* (ML2210), which has extensive gene loss and the presence of numerous pseudogenes (32). The activity of *M. tuberculosis* adenylyl cyclases is directly influenced by pH, carbon dioxide, fatty acids, hypoxia and starvation (19, 33). However, cAMP signaling in the *M. tuberculosis* complex is not relevant to the classical catabolite repression function found in *E. coli* (19). One of the predicted and active adenylyl cyclases, Rv0386, has been shown to produce and secrete cAMP within macrophages during infection. In addition, cAMP synthesized by Rv0386 is

involved in bacterial virulence. The deletion of Rv0386 in *M. tuberculosis* results in decreased TNF-α production in infected mice, and a reduction in bacterial survival (34).

Other important second messengers in bacteria are ppGpp and (p)ppGpp (guanosine pentaphosphate) or magic spot. ppGpp and pppGpp are synthesized by the addition of a pyrophosphate moiety to the 3′ position of GDP and GTP, respectively, by a ribosome-associated protein. There are several classes of enzymes that control the synthesis and degradation of (p)ppGpp. The first class is a RelA-SpoT homologue (RSH) that has both synthase and hydrolase domains. This bifunctional protein plays roles in synthesizing (p)ppGpp from ATP and either GTP or GDP. In addition, it hydrolyzes (p)ppGpp to pyrophosphate and GTP or GDP. Furthermore, the small alarmone synthetases (SAS) and hydrolases (SAH) are capable of synthesizing and degrading (p)ppGpp, respectively (35). Low levels of charged tRNAs results in the production of (p)ppGpp. (p)ppGpp plays many important roles in bacterial stringent response, a stress responses to amino acid starvation, heat shock, and other environmental changes (36, 37). (p)ppGpp can regulate cell replication, metabolism, transcription, and translation through direct interaction with its targets, providing a mechanism by which bacteria are able to survive and adapt to challenging stresses (37). For example, in Gram-negative bacteria such as *E. coli*, (p)ppGpp mediates cellular response to starvation by binding to RNA polymerase to repress genes encoding rRNA, and tRNA while also activating amino acid biosynthesis and transport genes (38, 39). In Gram-positive bacteria such as *Bacillus subtilis*, (p)ppGpp negatively regulates cellular GTP levels during starvation in order to resist nutritional stress and to maintain GTP homeostasis (40, 41).

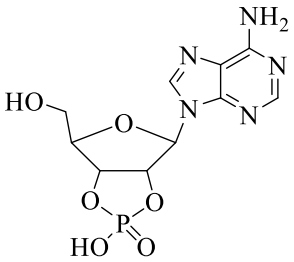
Recent studies have elucidated that 3’5’-cGMP also acts as a bacterial second messenger. A study in *Rhodospirillum centenum* has revealed that 3′5′-cGMP is used as a messenger to regulate desiccation-resistant cyst development (42). Interestingly, An *et al*. (43) demonstrated a

complex cascade where c-di-GMP synthesis is regulated by cGMP in *Xanthomonas campestris*, a causative agent of black rot disease of cruciferous plants. Through analysis of a transposon mutant library An *et al*. (43) discovered that the *XC_0250* gene encodes a class III nucleotide cyclase that has a guanylate cyclase activity. This gene is involved in virulence, transport, motility, biofilm formation, stress tolerance, drug resistance, detoxification and signal transduction of the bacterium. Moreover, this study has shown the relationship of cGMP and c-di-GMP. The cGMP synthesized by XC_0250 interacts with the cNMP-binding domain of the diguanylate cyclase XC_0249 to stimulate c-di-GMP synthesis (43).

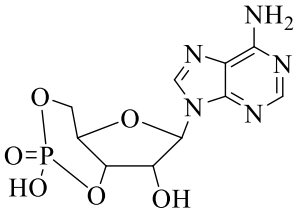
In addition to cyclic nucleotides, cyclic dinucleotides also play important roles as second messengers to regulate several activities in bacteria. Two well described cyclic dinucleotides in bacteria are c-di-AMP and c-di-GMP. c-di-AMP was first described in the Gram-positive bacterium, *B. subtilis*, as a monitor of chromosome integrity (44). c-di-AMP is synthesized from two molecules of ATP by diadenylate cyclase (DAC) and depleted to phosphoadenylyl adenosine (pApA) by PDE (45). The first evidence of c-di-AMP secretion by bacteria was from the intracellular bacterial pathogen, *Listeria monocytogenes*. In this bacterium, the c-di-AMP was secreted by multidrug efflux pumps (MDRs) and detected in the cytosol of host cells. Immunologically, the c-di-AMP is recognized by the innate immune system activating host type I IFN production, which is IFN-β (46). Furthermore, c-di-AMP was detected in cellular extracts from several bacteria such as *Staphylococcus aureus*, *Streptococcus pyogenes* and *Mycobacterium* spp. (45). In *M. tuberculosis*, the production of c-di-AMP has been recognized to be mediated by Rv3586 (DacA) diadenylate cyclase protein (47), and depletion of c-di-AMP occurs when it is enzymatically linearized by Rv2837 PDE into pApA, which is further hydrolyzed to 5'-AMP (48). Currently, the c-di-AMP receptor has only been identified in *M. smegmatis*. Ms5346 or DarR, a

TetR transcriptional factor was discovered to be a c-di-AMP receptor (49). The c-di-AMP was revealed to be involved in *M. tuberculosis* virulence. High levels of c-di-AMP results in reduced bacterial virulence in mice, enhanced IFN-β production, and increased macrophage autophagy (50, 51). Chemical structures, formulas, and monoisotopic masses of cyclic nucleotides/dinucleotides are shown in Figure 2-2.

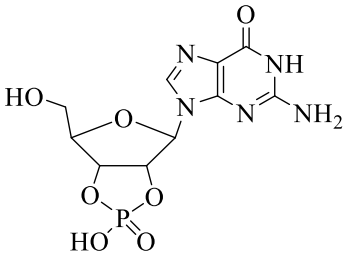
2'3'-cAMP ($C_{10}H_{12}N_5O_6P$, 329.0545 Da)



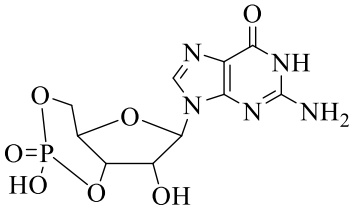
3'5'-cAMP ($C_{10}H_{12}N_5O_6P$, 329.0545 Da)



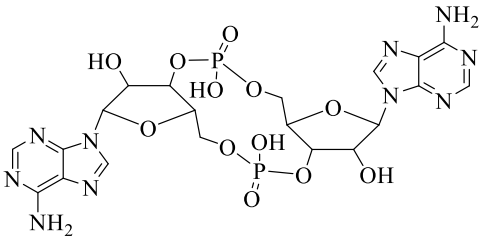
2'3'-cGMP ($C_{10}H_{12}N_5O_7P$, 345.0474 Da)



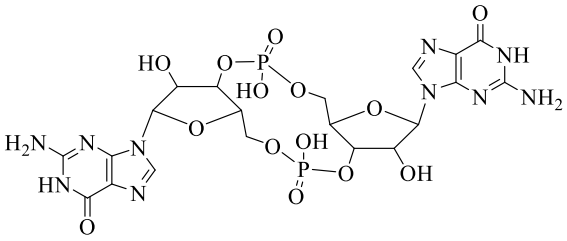
3'5'-cGMP ($C_{10}H_{12}N_5O_7P$, 345.0474 Da)



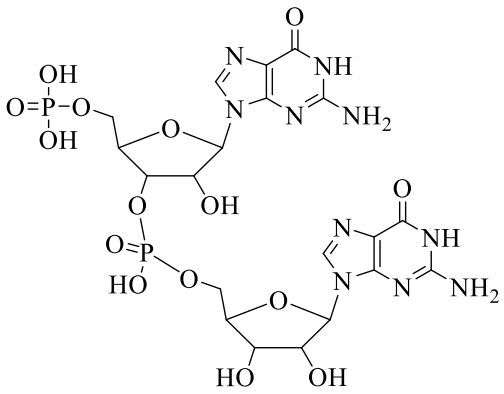
c-di-AMP ($C_{20}H_{24}N_{10}O_{12}P_2$, 658.105 Da)



c-di-GMP ($C_{20}H_{24}N_{10}O_{14}P_2$, 690.0949 Da)



pGpG ($C_{20}H_{26}N_{10}O_{15}P_2$, 708.1054 Da)



cGAMP ($C_{20}H_{24}N_{10}O_{13}P_2$, 674.1 Da)

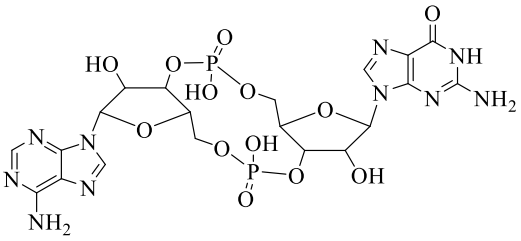


Figure 2-2. Chemical structures, formulas, and monoisotopic masses of nucleotide-based second messengers

Chemical structures were generated by ChemDraw Professional software (PerkinElmer Informatics, Waltham, MA). The monoisotopic masses were calculated using Agilent MassHunter softwarer (Agilent, Santa Clara, CA).

2.2 Cyclic diguanosine monophosphate (c-di-GMP)

Cyclic dimeric (3'5') guanosine monophosphate or the c-di-GMP molecule is a well-described second messenger in bacteria that was first described and isolated from *Acitobacter xylinum* which has been renamed as *Gluconacetobacter xylinus* in the early 1980s by Ross *et al*. in the laboratory of Moshe Benzimen (52). The initial discovery showed c-di-GMP functions to regulate the production of cellulose (a homopolymer of D-glucose residues linked in β-1,4-glucosidic bonds) by targeting membrane-bound cellulose synthase (53). Chemical analyses revealed that the cellulose synthase activator is composed of two moieties of GMP linked by a 3'5' phosphodiester bond with a molecular mass of 690 Da. The structure of this activator was identified as bis(3'5')-cyclic diguanylic acid, which is susceptible to 3'5'-phosphodiesterase hydrolysis. In addition, Ross *et al.* revealed that c-di-GMP is synthesized from two molecules of GTP by diguanylate cyclase enzyme (DGC), which is promoted by polyethylene-glycol (PEG) and degraded to 5'-GMP by PDE that is inhibited by Ca^{2+}. One molecule of pyrophosphate (PPi) is released during the synthesis reaction for every molecule of GMP residue incorporated from GTP into c-di-GMP (52).

2.2.1 Turnover mechanism

The formation of c-di-GMP is catalyzed via DGC, an enzyme that utilizes GTP as the substrate and possesses a conserved GG(D/E)EF motif (glycine-glycine-aspartate/glutamate-glutamate-phenylalanine) that is part of the active site domain (54, 55). Approximately 50% of proteins with characterized DGC activity (56) are regulated via allosteric control and the binding

of c-di-GMP to a conserved inhibitory site (I-site) motif (RxxD) directly upstream of the GGDEF motif (55, 57). Initially, the synthesis of c-di-GMP by a functional DGC involves the binding of two molecules of GTP to the two glycines of the GGDEF motif forming the functional homodimer of the GGDEF domain-containing protein, a typical mechanism of type III nucleotide cyclases (54). Moreover, the fourth residue (glutamic acid) is involved in metal ion coordination while the third amino acid of the motif (aspartic acid or glutamic acid) is essential for catalysis and also metal coordination. Specifically, the formation of a phosphodiester bond between the two GMP molecules requires two Mg^{2+} or Mn^{2+}cations. In addition, the formation of c-di-GMP from GTPs is a two-step reaction that has a diguanosine-tetraphosphate (5′-pppGpG) as a reaction intermediate (54, 55).

The degradation of c-di-GMP is also a two-step pathway. PDE proteins that possess EAL (glutamate-alanine-leucine) domains are responsible for the depletion of c-di-GMP and conversion to a linear nucleotide 5′phosphoguanylyl 3′5′ guanosine (pGpG) (58), and proteins containing HD-GYP (histidine-aspartate-glycine-tyrosine-proline) domains are able to deplete c-di-GMP to two molecules of guanosine monophosphate (GMP) (59). The PDE activity of the EAL domain is dependent on Mg^{2+} or Mn^{2+} ions that are involved in coordination with the glutamate residue (E) of the EAL motif. The PDE activity of EAL is strongly inhibited by Ca^{2+} and Zn^{2+} ions (58). HD-GYP domain contains a HHExxDGxxGYP motif, which is not related to the EAL domain (60). Crystal structure analysis of HD-GYP protein, which is a subset of HD family proteins, has shown that the GYP motif is not involved in PDE activity, but the tyrosine (Y) residue is involved as a binding site of bound c-di-GMP. The PDE activity of HD-GYP protein requires catalytic metals (Fe^{2+}) involved in trinuclear iron binding sites (61). Recently, oligoribonuclease (Orn) with an ability to degrade RNA oligomers to mononucleotides has also been identified to participate in the

conversion of pGpG to GMP in *Pseudomonas aeruginosa*. The activity of Orn is dependent on Mn^{2+} but it is not sensitive to Ca^{2+} (62). The activities of GGDEF and EAL domains can also be incorporated into a single bi-functional protein or separated on individual proteins.

Interestingly, a recent study revealed that one of the GGDEF domain-containing proteins from *G. sulfurreducens* functions as a dinucleotide cyclase to synthesize 3'3'-cGAMP by forming an asymmetric dimer to bind ATP and GTP. The enzymatic activity of this dinucleotide cyclase is dependent on ATP to GTP ratios (63). Another exceptional paradigm for c-di-GMP turnover is in the c-di-GMP degradation. He *et al*. has shown that c-di-AMP PDE, Rv2837, from *M. tuberculosis* harbors the ability to degrade c-di-AMP and c-di-GMP, with degradation of c-di-GMP occuring at a lower rate than that of c-di-AMP (64). Therefore, the signaling mechanisms of c-di-GMP in bacteria are complicated and involve a variety of unanticipated mechanisms and machineries. The turnover mechanism and signaling of c-di-GMP are shown in Figure 2-3.

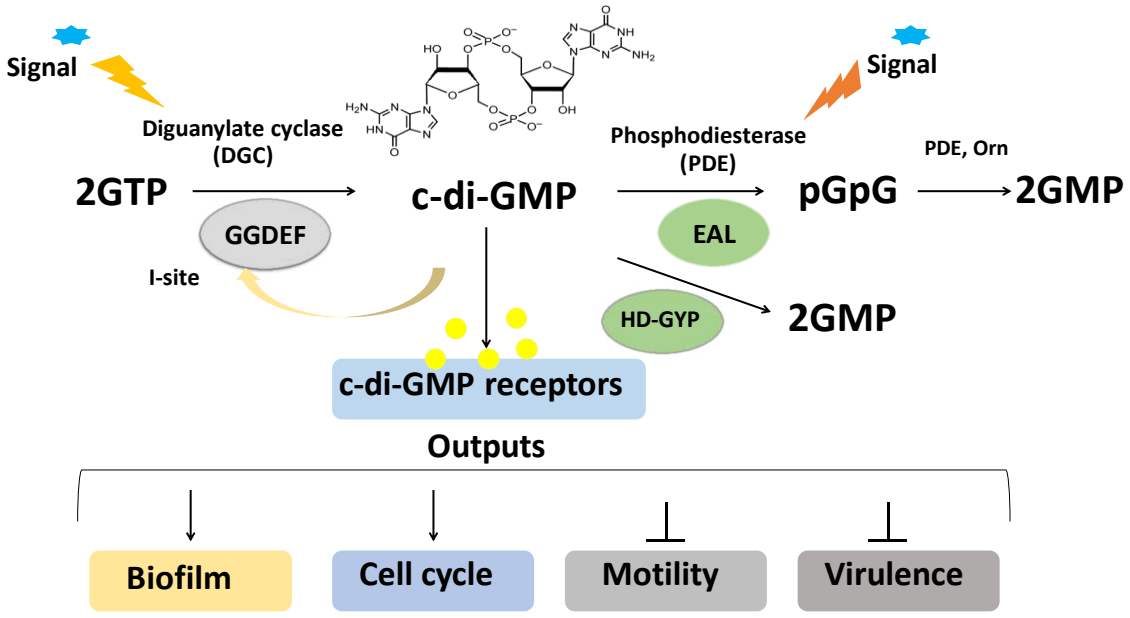


Figure 2-3. c-di-GMP turnover and signaling mechanism in bacteria
(Drawn based on Romling *et al*. (56) and Regine Hengge (2009) (65))

2.2.2 c-di-GMP signaling mechanism

The mechanism of c-di-GMP signaling in bacterial cells starts when the sensory input domains linked to GGDEF, EAL, or HD-GYP domain perceive environmental signals, which activates the DGC or PDE to synthesize or break down c-di-GMP, respectively. More importantly, sensory input domains typically located in the N-terminus of proteins containing DGC or PDE domains receive environmental signals that initiate the DGC or PDE activity (56, 66). The potential combinatorial diversity of various sensory input domains coupled to individual DGC or PDE encoding proteins as well as bi-functional proteins allows organisms such as *Pseudomonas* spp. to utilize c-di-GMP in the regulation of multiple activities such as biofilm formation, motility and virulence (67-69). The most common cytoplasmic proteins coupled with sensory domains are GGDEF, EAL and HD-GYP proteins with PAS, GAF, and receiver REC domains (56). They monitor cytoplasmic or extracellular levels of specific ligands and respond by producing or

degrading c-di-GMP depending on their localization and domain components. Nevertheless, only a small number of signals triggering those sensory domains have been identified.

PAS (Per-ARNT-Sim) domain got its name because it was commonly found in the PERIOD circadian protein (Per), the aryl hydrocarbon receptor nuclear translocator (ARNT), and the fly developmental regulator single-minded (SIM) (70). The GAF domain received its name because it was originally identified in cGMP-specific PDEs, adenylyl and guanylyl cyclases, phytochromes, FhlA and NifA transcription factors (71). Several ligands that bind to PAS and GAF domains to modulate DGC and PDE activity have been identified. PAS and GAF domains can sense various kinds of environmental signals such as light (72), heme (73), FMN (flavin mononucleotide) and other chromophores (73), FAD (flavin adenosine dinucleotide) (74), O_2 (75, 76), NO, CO, and redox potential (77). The GGDEF and/or EAL or HD-GYP domains with REC domains are the response regulators of the two-component signal system responsible for phosphorylation. This system modulates c-di-GMP levels in response to extracellular or intracellular signals perceived by the sensor histidine kinases (56). In addition to PAS, GAF and REC domains, other sensory domains involved in the c-di-GMP signaling network have been described. These sensory domains are the LOV (light-oxygen-voltage) domain which detects blue light (78) and globin domains involved in O_2 sensing (79). Furthermore, similar to quorum sensing, the c-di-GMP moiety has been shown to regulate several bacterial activities. Therefore, regulatory connections between the quorum sensing system and c-di-GMP signaling have been determined. Population density has been shown to be one of the environmental cues that modulates by c-di-GMP signaling (80-82).

After c-di-GMP synthesis, c-di-GMP interacts with several classes of receptors. These cellular effectors can be transcriptional factors (83, 84), proteins in the PilZ domain family with a

conserved motif, RxxR-D/NxSxxG (85), degenerate GGDEF or EAL domains (86, 87), riboswitches (26) and the I-site (54, 57). Finally, the receptors interact with a downstream target to affect cellular functions (56, 66).

2.2.3 Biological roles of c-di-GMP

c-di-GMP has been described in a broad number of bacterial species and is associated with regulation of a variety of bacterial activities including survival (88) virulence (68), cell differentiation (89), and biofilm formation (67, 90, 91). Pathogen-produced c-di-GMP has also been found to stimulate the host innate immune response (92). Most information about c-di-GMP signaling and biological roles has been well-described in Gram-negative bacteria such as *P. aeruginosa*, *Caulobacter crescentus*, *Salmonella enterica*, *E. coli*, and *Vibrio cholera* (68, 69, 90, 93-96). In Gram-positive bacteria, c-di-GMP has been characterized in only a few bacteria including *Clostridium difficile*, *B. subtilis*, and *S. aureus*. In *B. subtilis*, c-di-GMP plays a role in bacterial swarming motility, and the PilZ receptor has been identified as a c-di-GMP receptor (97, 98). *C. difficile* is another Gram-positive bacterium for which c-di-GMP has been characterized. c-di-GMP has been revealed to be involved in the regulation of type IV pili which leads to reduced *C. difficile* flagellar motility. This leads to cells being aggregated in liquid culture because swimming motility is reduced (99, 100). c-di-GMP also inhibits toxin production that is a key virulence factor in *C. difficile* (101).

Although c-di-GMP has been described as an intracellular signaling molecule, previous studies have shown that extracellular c-di-GMP could affect cellular activities of bacteria. Karaolis *et al.* showed that exogenous c-di-GMP could inhibit *S. aureus* intercellular adhesive interactions and biofilm formation (102). However, the mechanisms of action have not been elucidated. A

possible mechanism is that c-di-GMP binds to a surface receptor, which subsequently modulates gene expression. Alternatively, c-di-GMP may be able to enter *S. aureus* and directly trigger gene expression. Similar to c-di-AMP, c-di-GMP can act as an immunomodulator, stimulating secretion of cytokines and chemokines and expression of chemokine receptor by induction of dendritic cell maturation (103, 104). Specifically, c-di-GMP can function as a pathogen-associated molecular pattern (PAMP), stimulating the innate immune response to eliminate the pathogen. c-di-AMP and c-di-GMP are recognized by the helicase DDX41, a pattern recognition receptor (PRR) (105) which signals via the STING sensor and activates the transcription factor ,nuclear factor kappa B, (NF-κB) and IFN- regulatory factor 3 (IRF3) (46, 106, 107) to trigger a host type I IFN production (46, 104).

2.2.4 c-di-GMP studies in mycobacteria

In mycobacteria, the production and regulatory activity of c-di-GMP has been investigated in *M. tuberculosis,* the causative agent of tuberculosis (108), *Mycobacterium bovis* (109) and *M. smegmatis,* a fast growing mycobacterium used as a model organism (110). *M. tuberculosis* encodes a single bi-functional DGC-PDE protein (Rv1354c) (111). Disruption of Rv1354c PDE activity decreases pathogenicity and dormancy in *M. tuberculosis* (112). Studies of c-di-GMP PDE activity of Rv1357c in the closely related *M. bovis* BCG Pasteur 1173P2 demonstrated that c-di-GMP was associated with the regulation of lipid production and pellicle growth, and promoted resistance to nitrosative stress (109). Protein interaction studies also suggested that c-di-GMP production in *M. tuberculosis* is involved in regulation of rhamnose biosynthesis; a key sugar in the formation of the mycobacterial cell wall (113). Moreover, studies of the *M. smegmatis* homolog of *rv1354c* (*msmeg_2196*) demonstrated that both GGDEF and EAL domains are active and that

c-di-GMP is involved in colony morphology and long-term survival during nutrient starvation (88, 114, 115).

2.3 Purpose of the study and research hypothesis

The discovery of c-di-GMP and its potential roles in bacteria as a second messenger to regulate cellular activities has stimulated an interest for c-di-GMP studies in the *Mycobacterium* spp., especially *M. leprae* which has massive gene decay, but harbors several potential proteins functioning as diguanylate cyclases. The hypothesis of this study is that *M. leprae* has the ability to synthesize c-di-GMP by its potential DGC proteins. The rationale for this study is that after completion of the proposed research, the potential of *M. leprae* to synthesize c-di-GMP as a second messenger during infection will be revealed. Additionally, the presence of transcriptional expression of the genes encoding GGDEF-containing protein *in vivo* and the impacts of c-di-GMP on cellular activities in surrogate hosts or model organisms will be characterized.

2.4 Significance of the study

This study is significant because it generated the first demonstration that c-di-GMP exists in *M. leprae*. *M. leprae* is the causative agent of leprosy, which is still a major health problem in several developing countries. Management of leprosy has been challenging because of the long incubation period of disease. This is a major hurdle in disease diagnosis. The development of drug resistance additionally increases the challenges in management of leprosy. Leprosy treatment is further complicated by the development of a spectrum of clinical manifestations. After invading the host cell, *M. leprae* can cause various clinical manifestations depending on the host immune response (116). Knowledge of pathogenesis of leprosy is still limited. These fundamental gaps

significantly limit the development of disease management including treatment and prevention. We have began to effectively address these knowledge gaps through a better understanding of *M. leprae* intracellular communication by the second messenger, c-di-GMP.

From bioinformatics analyses, *M. leprae* harbors four genes, *ml1750c*, *ml1419c*, *ml0397c*, and *ml1752c* which are predicted to be responsible for c-di-GMP turnover. The successful completion of this study has provided preliminary knowledge that *M. leprae* has the ability to produce c-di-GMP. This study will pave the way for future research to characterize biological roles of c-di-GMP in *M. leprae* that may be involved in the pathogenesis of leprosy. This study has also shed light on the signaling mechanisms responding to environmental changes experienced by *M. leprae*.

2.5 Scope of this study

The scope of this study is based on genetic and biochemical approaches investigating c-di-GMP production by *M. leprae* DGC proteins, and the presence of those proteins during infection. Since *M. leprae* cannot be grown *in vitro*, genetic manipulation of *M. leprae* genes must be performed in model organisms. Elucidation of biological roles of c-di-GMP produced by *M. leprae* in model organisms might be useful to imply c-di-GMP biological roles in *M. leprae*. The genetic approach focuses on the identification of potential proteins that could function as DGCs to produce c-di-GMP using bioinformatics analyses. Functional activities of those proteins are determined by gene replacement or site-directed mutagenesis strategies to generate inactive DGC proteins. Additionally, expression of genes encoding DGCs during infection is studied by reverse transcription polymerase chain reaction. The production of protein in clinical samples from infected animals will be determined by specific antibody to those DGC proteins and will be confirmed by mass spectrometry. The biochemical approach is used to identify c-di-GMP levels

in bacterial cell lysate and cell cultures by a high performance method such as lipid chromatography mass spectrometry.

REFERENCES

1. Camilli A, Bassler BL. 2006. Bacterial small-molecule signaling pathways. *Science* 311:1113-1116.

2. Anonymous. 2010. Front Matter, p I-XXV, Bacterial Signaling doi:10.1002/9783527629237.fmatter. Wiley-VCH Verlag GmbH & Co. KGaA.

3. Waters CM, Bassler BL. 2005. Quorum sensing: cell-to-cell communication in bacteria. *Annu Rev Cell Dev Biol* 21:319-346.

4. Miller MB, Skorupski K, Lenz DH, Taylor RK, Bassler BL. 2002. Parallel quorum sensing systems converge to regulate virulence in *Vibrio cholerae*. *Cell* 110:303-314.

5. Hammer BK, Bassler BL. 2003. Quorum sensing controls biofilm formation in *Vibrio cholerae*. *Mol Microbiol* 50:101-104.

6. Magnuson R, Solomon J, Grossman AD. 1994. Biochemical and genetic characterization of a competence pheromone from *B. subtilis*. *Cell* 77:207-216.

7. Ji G, Beavis RC, Novick RP. 1995. Cell density control of staphylococcal virulence mediated by an octapeptide pheromone. *Proc Natl Acad Sci U S A* 92:12055-12059.

8. Fuqua C, Parsek MR, Greenberg EP. 2001. Regulation of gene expression by cell-to-cell communication: acyl-homoserine lactone quorum sensing. *Annu Rev Genet* 35:439-468.

9. Kleerebezem M, Quadri LEN, Kuipers OP, De Vos WM. 1997. Quorum sensing by peptide pheromones and two-component signal-transduction systems in Gram-positive bacteria. *Molecular Microbiology* 24:895-904.

10. Schauder S, Bassler BL. 2001. The languages of bacteria. *Genes Dev* 15:1468-1480.

11. Lyon WR, Madden JC, Levin JC, Stein JL, Caparon MG. 2001. Mutation of luxS affects growth and virulence factor expression in Streptococcus pyogenes. *Mol Microbiol* 42:145-157.

12. Schauder S, Shokat K, Surette MG, Bassler BL. 2001. The LuxS family of bacterial autoinducers: biosynthesis of a novel quorum-sensing signal molecule. *Mol Microbiol* 41:463-476.

13. Mitrophanov AY, Groisman EA. 2008. Signal integration in bacterial two-component regulatory systems. *Genes & Development* 22:2601-2611.

14. Gao R, Mack TR, Stock AM. 2007. Bacterial response regulators: versatile regulatory strategies from common domains. *Trends Biochem Sci* 32:225-234.

15. Helmann JD. 2002. The extracytoplasmic function (ECF) sigma factors. *Adv Microb Physiol* 46:47-110.

16. Missiakas D, Raina S. 1998. The extracytoplasmic function sigma factors: role and regulation. *Mol Microbiol* 28:1059-1066.

17. Pesavento C, Hengge R. 2009. Bacterial nucleotide-based second messengers. *Curr Opin Microbiol* 12:170-176.

18. Pereira SFF, Goss L, Dworkin J. 2011. Eukaryote-like serine/threonine kinases and phosphatases in bacteria. *Microbiology and Molecular Biology Reviews : MMBR* 75:192-212.

19. Knapp GS, Mcdonough KA. 2014. Cyclic AMP signaling in mycobacteria, Molecular Genetics of Mycobacteria, Second Edition doi:doi:10.1128/microbiolspec.MGM2-0011-2013. American Society of Microbiology.

20. Lodish H BA, Zipursky SL, et al. 2000. Second messengers, Molecular Cell Biology 4th edition. W. H. Freeman, New York.

21. Beavo JA, Brunton LL. 2002. Cyclic nucleotide research - still expanding after half a century. *Nat Rev Mol Cell Biol* 3:710-718.

22. Sunahara RK, Beuve A, Tesmer JJ, Sprang SR, Garbers DL, Gilman AG. 1998. Exchange of substrate and inhibitor specificities between adenylyl and guanylyl cyclases. *J Biol Chem* 273:16332-16338.

23. Zhang X, Shi H, Wu J, Zhang X, Sun L, Chen C, Chen ZJ. 2013. Cyclic GMP-AMP containing mixed phosphodiester linkages is an endogenous high-affinity ligand for STING. *Mol Cell* 51:226-235.

24. Civril F, Deimling T, de Oliveira Mann CC, Ablasser A, Moldt M, Witte G, Hornung V, Hopfner KP. 2013. Structural mechanism of cytosolic DNA sensing by cGAS. *Nature* 498:332-337.

25. Davies Bryan W, Bogard Ryan W, Young Travis S, Mekalanos John J. 2012. Coordinated regulation of accessory genetic elements produces cyclic di-nucleotides for *V. cholerae* virulence. *Cell* 149:358-370.

26. Kellenberger CA, Wilson SC, Hickey SF, Gonzalez TL, Su Y, Hallberg ZF, Brewer TF, Iavarone AT, Carlson HK, Hsieh YF, Hammond MC. 2015. GEMM-I riboswitches from *Geobacter* sense the bacterial second messenger cyclic AMP-GMP. *Proc Natl Acad Sci U S A* 112:5383-5388.

27. Chen Z-h, Schaap P. 2012. The prokaryote messenger c-di-GMP triggers stalk cell differentiation in *Dictyostelium*. *Nature* 488:680-683.

28. McDonough KA, Rodriguez A. 2012. The myriad roles of cyclic AMP in microbial pathogens: from signal to sword. *Nat Rev Micro* 10:27-38.

29. Harman JG. 2001. Allosteric regulation of the cAMP receptor protein. *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology* 1547:1-17.

30. Shenoy AR, Sivakumar K, Krupa A, Srinivasan N, Visweswariah SS. 2004. A survey of nucleotide cyclases in actinobacteria: unique domain organization and expansion of the class III cyclase family in *Mycobacterium tuberculosis*. *Comp Funct Genomics* 5:17-38.

31. Shenoy AR, Visweswariah SS. 2006. Mycobacterial adenylyl cyclases: Biochemical diversity and structural plasticity. *FEBS Letters* 580:3344-3352.

32. Shenoy AR, Sreenath N, Podobnik M, Kovacevic M, Visweswariah SS. 2005. The Rv0805 gene from *Mycobacterium tuberculosis* encodes a 3',5'-cyclic nucleotide phosphodiesterase: biochemical and mutational analysis. *Biochemistry* 44:15695-15704.

33. Bai G, Knapp GS, McDonough KA. 2011. Cyclic AMP signalling in mycobacteria: redirecting the conversation with a common currency. *Cell Microbiol* 13:349-358.

34. Agarwal N, Lamichhane G, Gupta R, Nolan S, Bishai WR. 2009. Cyclic AMP intoxication of macrophages by a *Mycobacterium tuberculosis* adenylate cyclase. *Nature* 460:98-102.

35. Atkinson GC, Tenson T, Hauryliuk V. 2011. The RelA/SpoT Homolog (RSH) Superfamily: Distribution and functional evolution of ppGpp synthetases and hydrolases across the tree of life. *PLoS ONE* 6:e23479.

36. Hauryliuk V, Atkinson GC, Murakami KS, Tenson T, Gerdes K. 2015. Recent functional insights into the role of (p)ppGpp in bacterial physiology. *Nat Rev Micro* 13:298-309.

37. Liu K, Bittner AN, Wang JD. 2015. Diversity in (p)ppGpp metabolism and effectors. *Curr Opin Microbiol* 24:72-79.

38. Reiness G, Yang HL, Zubay G, Cashel M. 1975. Effects of guanosine tetraphosphate on cell-free synthesis of *Escherichia coli* ribosomal RNA and other gene products. *Proceedings of the National Academy of Sciences of the United States of America* 72:2881-2885.

39. Paul BJ, Berkmen MB, Gourse RL. 2005. DksA potentiates direct activation of amino acid promoters by ppGpp. *Proc Natl Acad Sci U S A* 102:7823-7828.

40. Bittner AN, Kriel A, Wang JD. 2014. Lowering GTP level increases survival of amino acid starvation but slows growth rate for *Bacillus subtilis* cells lacking (p)ppGpp. *J Bacteriol* 196:2067-2076.

41. Kriel A, Bittner Alycia N, Kim Sok H, Liu K, Tehranchi Ashley K, Zou Winnie Y, Rendon S, Chen R, Tu Benjamin P, Wang Jue D. 2012. Direct regulation of GTP

homeostasis by (p)ppGpp: A critical component of viability and stress resistance. *Molecular Cell* 48:231-241.

42. Roychowdhury S, Dong Q, Bauer CE. 2015. DNA-binding properties of a cGMP-binding CRP homologue that controls development of metabolically dormant cysts of *Rhodospirillum centenum*. *Microbiology* 161:2256-2264.

43. An SQ, Chin KH, Febrer M, McCarthy Y, Yang JG, Liu CL, Swarbreck D, Rogers J, Maxwell Dow J, Chou SH, Ryan RP. 2013. A cyclic GMP-dependent signalling pathway regulates bacterial phytopathogenesis. *Embo j* 32:2430-2438.

44. Witte G, Hartung S, Büttner K, Hopfner K-P. 2008. Structural biochemistry of a bacterial checkpoint protein reveals diadenylate cyclase activity regulated by DNA recombination intermediates. *Molecular Cell* 30:167-178.

45. Corrigan RM, Grundling A. 2013. Cyclic di-AMP: another second messenger enters the fray. *Nat Rev Micro* 11:513-524.

46. Woodward JJ, Iavarone AT, Portnoy DA. 2010. c-di-AMP secreted by intracellular *Listeria monocytogenes* activates a host type I interferon response. *Science* 328:1703-1705.

47. Bai Y, Yang J, Zhou X, Ding X, Eisele LE, Bai G. 2012. *Mycobacterium tuberculosis* Rv3586 (DacA) is a diadenylate cyclase that converts ATP or ADP into c-di-AMP. *PLoS One* 7:e35206.

48. Manikandan K, Sabareesh V, Singh N, Saigal K, Mechold U, Sinha KM. 2014. Two-step synthesis and hydrolysis of cyclic di-AMP in *Mycobacterium tuberculosis*. *PLoS One* 9:e86096.

49. Zhang L, Li W, He ZG. 2013. DarR, a TetR-like transcriptional factor, is a cyclic di-AMP-responsive repressor in *Mycobacterium smegmatis*. *J Biol Chem* 288:3085-3096.

50. Yang J, Bai Y, Zhang Y, Gabrielle VD, Jin L, Bai G. 2014. Deletion of the cyclic di-AMP phosphodiesterase gene (cnpB) in *Mycobacterium tuberculosis* leads to reduced virulence in a mouse model of infection. *Mol Microbiol* 93:65-79.

51. Dey B, Dey RJ, Cheung LS, Pokkali S, Guo H, Lee JH, Bishai WR. 2015. A bacterial cyclic dinucleotide activates the cytosolic surveillance pathway and mediates innate resistance to tuberculosis. *Nat Med* 21:401-406.

52. Ross P, Weinhouse H, Aloni Y, Michaeli D, Weinberger-Ohana P, Mayer R, Braun S, de Vroom E, van der Marel GA, van Boom JH, Benziman M. 1987. Regulation of cellulose synthesis in *Acetobacter xylinum* by cyclic diguanylic acid. *Nature* 325:279-281.

53. Ross P, Mayer R, Benziman M. 1991. Cellulose biosynthesis and function in bacteria. *Microbiol Rev* 55:35-58.

54. Chan C, Paul R, Samoray D, Amiot NC, Giese B, Jenal U, Schirmer T. 2004. Structural basis of activity and allosteric control of diguanylate cyclase. *Proc Natl Acad Sci U S A* 101:17084-17089.

55. Wassmann P, Chan C, Paul R, Beck A, Heerklotz H, Jenal U, Schirmer T. 2007. Structure of $BeF3^-$ -modified response regulator PleD: implications for diguanylate cyclase activation, catalysis, and feedback inhibition. *Structure* 15:915-927.

56. Romling U, Galperin MY, Gomelsky M. 2013. Cyclic di-GMP: the first 25 years of a universal bacterial second messenger. *Microbiol Mol Biol Rev* 77:1-52.

57. Christen B, Christen M, Paul R, Schmid F, Folcher M, Jenoe P, Meuwly M, Jenal U. 2006. Allosteric control of cyclic di-GMP signaling. *J Biol Chem* 281:32015-32024.

58. Schmidt AJ, Ryjenkov DA, Gomelsky M. 2005. The ubiquitous protein domain EAL is a cyclic diguanylate-specific phosphodiesterase: Enzymatically active and inactive EAL domains. *Journal of Bacteriology* 187:4774-4781.

59. Ryan RP, Fouhy Y, Lucey JF, Crossman LC, Spiro S, He YW, Zhang LH, Heeb S, Camara M, Williams P, Dow JM. 2006. Cell-cell signaling in *Xanthomonas campestris* involves an HD-GYP domain protein that functions in cyclic di-GMP turnover. *Proc Natl Acad Sci U S A* 103:6712-6717.

60. Dow JM, Fouhy Y, Lucey JF, Ryan RP. 2006. The HD-GYP domain, cyclic di-GMP signaling, and bacterial virulence to plants. *Mol Plant Microbe Interact* 19:1378-1384.

61. Bellini D, Caly DL, McCarthy Y, Bumann M, An SQ, Dow JM, Ryan RP, Walsh MA. 2014. Crystal structure of an HD-GYP domain cyclic-di-GMP phosphodiesterase reveals an enzyme with a novel trinuclear catalytic iron centre. *Mol Microbiol* 91:26-38.

62. Cohen D, Mechold U, Nevenzal H, Yarmiyhu Y, Randall TE, Bay DC, Rich JD, Parsek MR, Kaever V, Harrison JJ, Banin E. 2015. Oligoribonuclease is a central feature of cyclic diguanylate signaling in *Pseudomonas aeruginosa*. *Proc Natl Acad Sci U S A* 112:11359-11364.

63. Hallberg ZF, Wang XC, Wright TA, Nan B, Ad O, Yeo J, Hammond MC. 2016. Hybrid promiscuous (Hypr) GGDEF enzymes produce cyclic AMP-GMP (3', 3'-cGAMP). *Proc Natl Acad Sci U S A* 113:1790-1795.

64. He Q, Wang F, Liu S, Zhu D, Cong H, Gao F, Li B, Wang H, Lin Z, Liao J, Gu L. 2016. Structural and Biochemical Insight into the Mechanism of Rv2837c from *Mycobacterium tuberculosis* as a c-di-NMP Phosphodiesterase. *J Biol Chem* 291:3668-3681.

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

66. Sondermann H, Shikuma NJ, Yildiz FH. 2012. You've come a long way: c-di-GMP signaling. *Curr Opin Microbiol* 15:140-146.

67. Hickman JW, Tifrea DF, Harwood CS. 2005. A chemosensory system that regulates biofilm formation through modulation of cyclic diguanylate levels. *Proc Natl Acad Sci U S A* 102:14422-14427.

68. Kulasekara H, Lee V, Brencic A, Liberati N, Urbach J, Miyata S, Lee DG, Neely AN, Hyodo M, Hayakawa Y, Ausubel FM, Lory S. 2006. Analysis of *Pseudomonas aeruginosa* diguanylate cyclases and phosphodiesterases reveals a role for bis-(3'-5')-cyclic-GMP in virulence. *Proc Natl Acad Sci U S A* 103:2839-2844.

69. Merritt JH, Brothers KM, Kuchma SL, O'Toole GA. 2007. SadC reciprocally influences biofilm formation and swarming motility via modulation of exopolysaccharide production and flagellar function. *J Bacteriol* 189:8154-8164.

70. Huang ZJ, Edery I, Rosbash M. 1993. PAS is a dimerization domain common to *Drosophila* period and several transcription factors. *Nature* 364:259-262.

71. Ho YS, Burden LM, Hurley JH. 2000. Structure of the GAF domain, a ubiquitous signaling motif and a new class of cyclic GMP receptor. *Embo j* 19:5288-5299.

72. Savakis P, De Causmaecker S, Angerer V, Ruppert U, Anders K, Essen LO, Wilde A. 2012. Light-induced alteration of c-di-GMP level controls motility of *Synechocystis* sp. PCC 6803. *Mol Microbiol* 85:239-251.

73. Sardiwal S, Kendall SL, Movahedzadeh F, Rison SC, Stoker NG, Djordjevic S. 2005. A GAF domain in the hypoxia/NO-inducible *Mycobacterium tuberculosis* DosS protein binds haem. *J Mol Biol* 353:929-936.

74. Qi Y, Rao F, Luo Z, Liang ZX. 2009. A flavin cofactor-binding PAS domain regulates c-di-GMP synthesis in AxDGC2 from *Acetobacter xylinum*. *Biochemistry* 48:10275-10285.

75. Chang AL, Tuckerman JR, Gonzalez G, Mayer R, Weinhouse H, Volman G, Amikam D, Benziman M, Gilles-Gonzalez MA. 2001. Phosphodiesterase A1, a regulator of cellulose synthesis in *Acetobacter xylinum*, is a heme-based sensor. *Biochemistry* 40:3420-3426.

76. Tuckerman JR, Gonzalez G, Sousa EH, Wan X, Saito JA, Alam M, Gilles-Gonzalez MA. 2009. An oxygen-sensing diguanylate cyclase and phosphodiesterase couple for c-di-GMP control. *Biochemistry* 48:9764-9774.

77. Henry JT, Crosson S. 2011. Ligand-binding PAS domains in a genomic, cellular, and structural context. *Annu Rev Microbiol* 65:261-286.

78. Cao Z, Livoti E, Losi A, Gartner W. 2010. A blue light-inducible phosphodiesterase activity in the cyanobacterium *Synechococcus elongatus*. *Photochem Photobiol* 86:606-611.

79. Wan X, Tuckerman JR, Saito JA, Freitas TA, Newhouse JS, Denery JR, Galperin MY, Gonzalez G, Gilles-Gonzalez MA, Alam M. 2009. Globins synthesize the second

messenger bis-(3'-5')-cyclic diguanosine monophosphate in bacteria. *J Mol Biol* 388:262-270.

80. Ryan RP, McCarthy Y, Andrade M, Farah CS, Armitage JP, Dow JM. 2010. Cell-cell signal-dependent dynamic interactions between HD-GYP and GGDEF domain proteins mediate virulence in *Xanthomonas campestris*. *Proceedings of the National Academy of Sciences of the United States of America* 107:5989-5994.

81. Srivastava D, Harris RC, Waters CM. 2011. Integration of cyclic di-GMP and quorum sensing in the control of *vpsT* and *aphA* in *Vibrio cholerae*. *J Bacteriol* 193:6331-6341.

82. Trimble MJ, McCarter LL. 2011. Bis-(3'-5')-cyclic dimeric GMP-linked quorum sensing controls swarming in *Vibrio parahaemolyticus*. *Proc Natl Acad Sci U S A* 108:18079-18084.

83. Hickman JW, Harwood CS. 2008. Identification of FleQ from *Pseudomonas aeruginosa* as a c-di-GMP-responsive transcription factor. *Mol Microbiol* 69:376-389.

84. Krasteva PV, Fong JC, Shikuma NJ, Beyhan S, Navarro MV, Yildiz FH, Sondermann H. 2010. *Vibrio cholerae* VpsT regulates matrix production and motility by directly sensing cyclic di-GMP. *Science* 327:866-868.

85. Amikam D, Galperin MY. 2006. PilZ domain is part of the bacterial c-di-GMP binding protein. *Bioinformatics* 22:3-6.

86. Jenal U, Malone J. 2006. Mechanisms of cyclic-di-GMP signaling in bacteria. *Annu Rev Genet* 40:385-407.

87. Navarro MV, De N, Bae N, Wang Q, Sondermann H. 2009. Structural analysis of the GGDEF-EAL domain-containing c-di-GMP receptor FimX. *Structure* 17:1104-1116.

88. Bharati BK, Sharma IM, Kasetty S, Kumar M, Mukherjee R, Chatterji D. 2012. A full-length bifunctional protein involved in c-di-GMP turnover is required for long-term survival under nutrient starvation in *Mycobacterium smegmatis*. *Microbiology* 158:1415-1427.

89. Abel S, Chien P, Wassmann P, Schirmer T, Kaever V, Laub MT, Baker TA, Jenal U. 2011. Regulatory cohesion of cell cycle and cell differentiation through interlinked phosphorylation and second messenger networks. *Mol Cell* 43:550-560.

90. Irie Y, Borlee BR, O'Connor JR, Hill PJ, Harwood CS, Wozniak DJ, Parsek MR. 2012. Self-produced exopolysaccharide is a signal that stimulates biofilm formation in *Pseudomonas aeruginosa*. *Proc Natl Acad Sci U S A* 109:20632-20636.

91. Borlee BR, Goldman AD, Murakami K, Samudrala R, Wozniak DJ, Parsek MR. 2010. *Pseudomonas aeruginosa* uses a cyclic-di-GMP-regulated adhesin to reinforce the biofilm extracellular matrix. *Mol Microbiol* 75:827-842.

92. Karaolis DKR, Means TK, Yang D, Takahashi M, Yoshimura T, Muraille E, Philpott D, Schroeder JT, Hyodo M, Hayakawa Y, Talbot BG, Brouillette E, Malouin F. 2007. Bacterial c-di-GMP is an immunostimulatory molecule. *The Journal of Immunology* 178:2171-2181.

93. Abel S, Chien P, Wassmann P, Schirmer T, Kaever V, Laub MT, Baker TA, Jenal U. 2011. Regulatory cohesion of cell cycle and cell differentiation through inter linked phosphorylation and second messenger networks. *Molecular Cell* 43:550-560.

94. Simm R, Morr M, Kader A, Nimtz M, Romling U. 2004. GGDEF and EAL domains inversely regulate cyclic di-GMP levels and transition from sessility to motility. *Mol Microbiol* 53:1123-1134.

95. Paul K, Nieto V, Carlquist WC, Blair DF, Harshey RM. 2010. The c-di-GMP binding protein YcgR controls flagellar motor direction and speed to affect chemotaxis by a "backstop brake" mechanism. *Mol Cell* 38:128-139.

96. Tischler AD, Camilli A. 2004. Cyclic diguanylate (c-di-GMP) regulates *Vibrio cholerae* biofilm formation. *Mol Microbiol* 53:857-869.

97. Gao X, Mukherjee S, Matthews PM, Hammad LA, Kearns DB, Dann CE, 3rd. 2013. Functional characterization of core components of the *Bacillus subtilis* cyclic-di-GMP signaling pathway. *J Bacteriol* 195:4782-4792.

98. Chen Y, Chai Y, Guo JH, Losick R. 2012. Evidence for cyclic di-GMP-mediated signaling in *Bacillus subtilis*. *J Bacteriol* 194:5080-5090.

99. Purcell EB, McKee RW, McBride SM, Waters CM, Tamayo R. 2012. Cyclic diguanylate inversely regulates motility and aggregation in *Clostridium difficile*. *J Bacteriol* 194:3307-3316.

100. Bordeleau E, Purcell EB, Lafontaine DA, Fortier LC, Tamayo R, Burrus V. 2015. Cyclic di-GMP riboswitch-regulated type IV pili contribute to aggregation of *Clostridium difficile*. *J Bacteriol* 197:819-832.

101. McKee RW, Mangalea MR, Purcell EB, Borchardt EK, Tamayo R. 2013. The second messenger cyclic di-GMP regulates *Clostridium difficile* toxin production by controlling expression of sigD. *J Bacteriol* 195:5174-5185.

102. Karaolis DK, Rashid MH, Chythanya R, Luo W, Hyodo M, Hayakawa Y. 2005. c-di-GMP (3'-5'-cyclic diguanylic acid) inhibits *Staphylococcus aureus* cell-cell interactions and biofilm formation. *Antimicrob Agents Chemother* 49:1029-1038.

103. Karaolis DK, Means TK, Yang D, Takahashi M, Yoshimura T, Muraille E, Philpott D, Schroeder JT, Hyodo M, Hayakawa Y, Talbot BG, Brouillette E, Malouin F. 2007. Bacterial c-di-GMP is an immunostimulatory molecule. *J Immunol* 178:2171-2181.

104. McWhirter SM, Barbalat R, Monroe KM, Fontana MF, Hyodo M, Joncker NT, Ishii KJ, Akira S, Colonna M, Chen ZJJ, Fitzgerald KA, Hayakawa Y, Vance RE. 2009. A host type I interferon response is induced by cytosolic sensing of the bacterial second messenger cyclic-di-GMP. *Journal of Experimental Medicine* 206:1899-1911.

105. Parvatiyar K, Zhang Z, Teles RM, Ouyang S, Jiang Y, Iyer SS, Zaver SA, Schenk M, Zeng S, Zhong W, Liu ZJ, Modlin RL, Liu YJ, Cheng G. 2012. The helicase DDX41 recognizes the bacterial secondary messengers cyclic di-GMP and cyclic di-AMP to activate a type I interferon immune response. *Nat Immunol* 13:1155-1161.

106. Burdette DL, Monroe KM, Sotelo-Troha K, Iwig JS, Eckert B, Hyodo M, Hayakawa Y, Vance RE. 2011. STING is a direct innate immune sensor of cyclic di-GMP. *Nature* 478:515-U111.

107. Sauer JD, Sotelo-Troha K, von Moltke J, Monroe KM, Rae CS, Brubaker SW, Hyodo M, Hayakawa Y, Woodward JJ, Portnoy DA, Vance RE. 2011. The N-ethyl-N-nitrosourea-induced Goldenticket mouse mutant reveals an essential function of Sting in the *in vivo* interferon response to Listeria monocytogenes and cyclic dinucleotides. *Infect Immun* 79:688-694.

108. Gengenbacher M, Kaufmann SHE. 2012. *Mycobacterium tuberculosis*: success through dormancy. *FEMS Microbiology Reviews* 36:514-532.

109. Flores-Valdez MA, de Jesus Aceves-Sanchez M, Pedroza-Roldan C, Vega-Dominguez PJ, Prado-Montes de Oca E, Bravo-Madrigal J, Laval F, Daffe M, Koestler B, Waters CM. 2015. The cyclic di-GMP phosphodiesterase gene Rv1357c/BCG1419c affects BCG pellicle production and *in vivo* maintenance. *IUBMB Life* 67:129-138.

110. Reyrat J-M, Kahn D. 2001. *Mycobacterium smegmatis*: an absurd model for tuberculosis? *Trends in Microbiology* 9:472-473.

111. Gupta K, Kumar P, Chatterji D. 2010. Identification, activity and disulfide connectivity of c-di-GMP regulating proteins in *Mycobacterium tuberculosis*. *PLoS One* 5:(11): e15072. doi:15010.11371/journal.pone.0015072.

112. Hong Y, Zhou X, Fang H, Yu D, Li C, Sun B. 2013. Cyclic di-GMP mediates *Mycobacterium tuberculosis* dormancy and pathogenecity. *Tuberculosis (Edinb)* 93:625-634.

113. Deng J, Bi L, Zhou L, Guo SJ, Fleming J, Jiang HW, Zhou Y, Gu J, Zhong Q, Wang ZX, Liu Z, Deng RP, Gao J, Chen T, Li W, Wang JF, Wang X, Li H, Ge F, Zhu G, Zhang HN, Gu J, Wu FL, Zhang Z, Wang D, Hang H, Li Y, Cheng L, He X, Tao SC, Zhang XE. 2014. *Mycobacterium tuberculosis* proteome microarray for global studies of protein function and immunogenicity. *Cell Rep* 9:2317-2329.

114. Sharma IM, Prakash S, Dhanaraman T, Chatterji D. 2014. Characterization of a dual-active enzyme, DcpA, involved in cyclic diguanosine monophosphate turnover in *Mycobacterium smegmatis*. *Microbiology* 160:2304-2318.

115. Gupta KR, Kasetty S, Chatterji D. 2015. Novel functions of (p)ppGpp and cyclic di-GMP in mycobacterial physiology revealed by phenotype microarray analysis of wild-type and isogenic strains of *Mycobacterium smegmatis*. *Appl Environ Microbiol* 81:2571-2578.

116. Ridley DS, Jopling WH. 1966. Classification of leprosy according to immunity. A five-group system. *Int J Lepr Other Mycobact Dis* 34:255-273.

CHAPTER 3 C-DI-GMP DETECTION METHODOLOGY DEVELOPMENT

3.1 Introduction

Quantification and detection of c-di-GMP is crucial for assessing the activity of c-di-GMP metabolizing proteins for both *in vivo* and *in vitro* experiments. Several analytical techniques have been developed to detect c-di-GMP in enzymatic assays, particularly for the levels of c-di-GMP found in bacterial cells during *in vivo* studies. These techniques have both advantages and drawbacks for detecting c-di-GMP. Fluorescent protein based assays like fluorescence resonance energy transfer (FRET) have been used as biosensors to monitor c-di-GMP levels in live bacterial cells, and can be visualized by microscopy (1, 2). Fluorescent intercalator based methods have also been applied for c-di-GMP detection. Thiazole orange (TO, 1-methyl-4-[(3-methyl-2(3H)-benzothiazolylidene)-methyl] quinolinium *p*-tosylate) is a fluorescent intercalator that can bind nucleic acids, DNA and RNA. It has been applied to detect and bind to c-di-GMP based on the ability of c-di-GMP to form G-quadruplexes when the aromatic thiazole orange is present. Changes in the absorption and fluorescence spectra of thiazole orange upon interaction with c-di-GMP occur in a concentration-dependent manner. However, this method is not sensitive enough to detect c-di-GMP at concentrations lower than 5 μM (3). Another approach used to determine c-di-GMP concentration is intercalator-mediated peroxidase formation. Similar to thiazole orange, the c-di-GMP can form G-quadruplexes in the presence of proflavine which enhances peroxidation of hemin. The c-di-GMP-proflavine-hemin complex is able to oxidize the colorless compound 2,2'-azino-bis(3-ethylbezothiazo-line-6-sulfonic acid) to form colored products. Although, this method is a simple detection system which does not require any instruments and can be visualized by eye, the limit of detection is approximately 1 μM (4).

An aptamer strategy has also been developed to detect c-di-GMP. In this strategy, a class I riboswitch domain or aptamer is used as the c-di-GMP sensing region, and is coupled with spinach RNA as a fluorescent reporter or ribozyme responsible for RNA cleavage. This method can detect c-di-GMP in a nanomolar range (5, 6). Recently, a label-free electrochemical biosensor coupled with a self-assembled riboswitch has been created to detect c-di-GMP, and has a limit of detection as low as 50 nM. This method is based on electrochemical impedance spectroscopy (EIS) which is a quantitative method with low destruction to biological samples (7).

Reversed-phase high-performance liquid chromatography (HPLC) has been used to separate c-di-GMP from bacterial extracts and detect c-di-GMP by UV absorption at 254 nm (8, 9). However, this method has low sensitivity and can provide poor peak separation, particularly when this method is applied to detect c-di-GMP in cellular extracts with a high concentration of matrix and other nucleotides that may have similar mass to c-di-GMP. Therefore, low levels of c-di-GMP can be detected using a mass spectrometry method that provides adequate sensitivity. Currently, the separation of c-di-GMP from crude cell extracts by HPLC combined with mass spectrometry for c-di-GMP quantification is considered the gold standard (10). Matrix-assisted laser desorption/ionization-time of flight mass spectrometry or MALDI-TOF can be used to detect c-di-GMP in bacterial samples or in sample fractions after HPLC separation (8, 11, 12). MALDI-TOF analysis is a technique that ionizes molecules that are embedded in a matrix. Analytes from the sample are cocrystallized with a saturated matrix solution. The co-crystals are irradiated by laser energy that induces ionization of the matrix that then desorbs and transfers ions to an electric field where they are detected. The time of flight to a detector in an electric field is used to determine the molecular weight of the substance (mass-to-charge ratio [*m/z*]), which is longer for larger molecules than for smaller ones. MALDI-TOF is considered to be a soft ionization technique

because it generates minimal or no fragmentation of ions (13, 14). Nevertheless, MALDI-TOF has some drawbacks. First, it requires an additional step for chromatographic work-up of c-di-GMP. Second, uneven matrix crystallization and unequal spreading of matrix on the template can occur. Consequently, reliability and reproducibility is a major limitation of this method (12).

High performance liquid chromatography-coupled mass spectrometry and tandem-mass spectrometry (LC-MS and LC-MS/MS) has been a widely used as a methodology to detect c-di-GMP in biological samples. The separation of ionized samples based on their mass-to-charge ratio (*m/z*) after partitioning by LC leads to the high sensitivity and specificity of the LC-MS approach (15). Specifically, the LC-MS/MS provides a better sensitivity and selectivity since the isolated ion from the first MS is further fragmented into product ions. The concentration of intracellular c-di-GMP is important for determination of the c-di-GMP regulatory network. Therefore, reliable and sensitive quantitation methods to determine *in vivo* c-di-GMP levels and/or *in vitro* experiments assessing enzymatic activity of DGC are required to get a better understanding of the c-di-GMP regulatory network and the physiological output from the c-di-GMP pool in bacterial cells. Of particular note, bacterial cells or cultures contain numerous nucleotides and compounds other than c-di-GMP that could interfere with c-di-GMP detection or co-elute with c-di-GMP. In this study, LC-MS and LC-MS/MS methods have been developed and optimized to detect c-di-GMP in biological samples from different bacterial species. The method can also be applied to determine other important nucleotides in bacterial cells and eukaryotes including; c-di-AMP, cGAMP, pGpG, 2’3’-cAMP, 3’5’-cAMP, 2’3’-cGMP, and 3’5’-cGMP. These optimized techniques will likely provide invaluable methods for acquiring high yields of c-di-GMP and other cyclic nucleotides from bacteria, and may be applied for detection of cyclic nucleotides in clinical specimens.

3.2 Materials and Methods

3.2.1 Materials

Cyclic nucleotides purchased from manufacturers were used as standards for method optimization. Cyclic diguanosine monophosphate (c-di-GMP), cyclic diadenosine monophosphate (c-di-AMP), cyclic GMP-AMP (cGAMP), and 5'- phosphoguanylyl- (3'5')- guanosine (pGpG), guanosine- 3'5'-cyclic monophosphate (3'5'-cGMP), and guanosine- 2'3'-cyclic monophosphate (2'3'-cGMP) were purchased from BIOLOG Life Science Institute (Germany). Adenosine 3'5'-cyclic monophosphate (3'5'-cAMP) and adenosine 2'3'-cyclic monophosphate (2'3'-cAMP) were purchased from Sigma Aldrich (St. Louis, MO). Isotopic adenosine ([1'-^{13}C]adenosine) used as an internal standard (IS) was purchased from Omicron Biochemicals (South Bend, IN). All compounds were stored at -20ºC upon arrival. They were suspended in LC-MS grade water to a final concentration of 1 mM or 10 mM and stored at -80ºC as stock solutions until use. The formula and molecular weight of the compounds are shown in Table 3-1. LC-MS grade water and methanol were obtained from Honeywell Burdick and Jackson (Morris Plains, NJ). Acetic acid and ammonium acetate were purchased from Sigma Aldrich.

Table 3-1. Cyclic nucleotides/dinucleotides used in this study

Name	Empirical Formula	Molecular weight (Da)
c-di-GMP	$C_{20}H_{24}N_{10}O_{14}P_2$	690.0949
c-di-AMP	$C_{20}H_{24}N_{10}O_{12}P_2$	658.105
cGMP	$C_{10}H_{12}N_5O_7P$	345.0474
cAMP	$C_{10}H_{12}N_5O_6P$	329.0545
cGAMP	$C_{20}H_{24}N_{10}O_{13}P_2$	674.1
pGpG	$C_{20}H_{26}N_{10}O_{15}P_2$	708.1054
[^{13}C]adenosine	$^{13}CC_9H_{13}N_5O_4$	268.1001

3.2.2 LC-MS and LC-MS/MS method optimization

LC-MS and LC-MS/MS experiments were performed on an Agilent 1200 HPLC system coupled to an Agilent 6520 quadrupole time-of-flight (Q-TOF) mass spectrometer (Santa Clara, CA). Standard compounds were resolved on a reverse phase Atlantis T3, C18 column (3 µm particle size, 2.1 x 150 mm, Waters (Millford, MA)) at a flow rate of 350 µl/min at 30ºC. The gradient consisted of 100% solvent A (0.1% acetic acid, 10 mM ammonium acetate) for 1 min followed by a 1 min linear gradient to 10% solvent B (methanol) (16), a 1.4 min linear gradient to 20% buffer B and a 1 min linear gradient to 100% buffer B and held for 2 min. The eluted compound from LC was diverted to waste for 2 min before being diverted to the mass spectrometry. The Q-TOF mass spectrometer was operated in positive ion mode at 2 GHz extended dynamic range, a *m/z* range of 100 to 1700 and at a scan rate of one spectra per second. MS data were collected in profile and centroid mode. Electrospray ionization (ESI) source parameters were 2.5 kv, 350°C gas temperature, a drying gas flow rate of 11 L/min, and a nebulizer flow rate of 45 psi. Confirmation of cyclic nucleotides/ dinucleotides peaks was performed by LC-MS/MS.

3.2.3 Calibration curves for quantitation of cyclic nucleotides and cyclic dinucleotides

Stock solutions of eight compounds and the internal standard, [^{13}C]adenosine, were utilized to construct calibration curves. Concentrations of each compound were prepared as dilutions within the range of 25 nM – 1 µM with a 100 nM or 200 nM [^{13}C]adenosine internal standard in LC-MS grade water. Quality control samples (QC) of each compound were prepared at three concentrations in LC-MS grade water, 40 nM, 200 nM and 800 nM with 200 nM [^{13}C]adenosine. All samples for standard curve construction were aliquoted prior to the study and were stored at -80ºC to avoid a refreeze-thaw cycle that could cause compound degradation. The volume of

samples injected into LC-MS was 10 µl. The composition of standard curve samples with 200 nM [^{13}C]adenosine internal standard are shown in Table 3-2.

Table 3-2. Composition of samples for standard curve and quality control

Calibration curve concentrations	**Compound volume and concentration**	**10 µM [^{13}C]adenosine]***	**LC-MS grade water**	**Final volume**
25 nM	50 µl of 250 nM	100 µl	400 µl	500 µl
50 nM	50 µl of 500 nM	100 µl	400 µl	500 µl
100 nM	50 µl of 1 µM	100 µl	400 µl	500 µl
250 nM	50 µl of 2.5 µM	100 µl	400 µl	500 µl
500 nM	50 µl of 5 µM	100 µl	400 µl	500 µl
1000 nM	50 µl of 10 µM	100 µl	400 µl	500 µl
QC 40 nM	50 µl of 400 nM	100 µl	400 µl	500 µl
QC 200 nM	50 µl of 2 µM	100 µl	400 µl	500 µl
QC 800 nM	50 µl of 8 µM	100 µl	400 µl	500 µl

* Final concentration of IS is 200 nM. QC is quality control.

3.2.4 Validation procedure

Calibration curves were generated by plotting the peak area ratio of an internal standard against the concentration (nM) of each compound. Each calibration curve was obtained from six calibration standard samples analyzed in triplicate in three separate analytical runs. Calibration curves were constructed with Agilent MassHunter Quantitative Analysis software using a linear regression model without a weighting factor. The linearity, inter-, and intra-day accuracy and precision were examined by analysis of three QC sample concentrations for each compound. Each OQ concentration was analyzed using three analytical runs performed in replicates of five. The limit of detection (LOD) and limit of quantification (LOQ) were determined at the concentration of each compound that had a signal to noise ratio (SNR) of at least three and 10, respectively, in the mass chromatogram.

3.2.5 Cyclic nucleotide and cyclic dinucleotide extraction from *P. aeruginosa*, *E. coli*, and *M. tuberculosis* H37Ra

P. aeruginosa PAO1

tpbB (PA1120) gene expression in *P. aeruginosa* PAO1 was accomplished using the arabinose-inducible vector pJN105 containing the P_{BAD} promoter (17). Plasmids were provided by Dr. Brad Borlee, Colorado State University. The *tpbB* is a well-characterized diguanylate cyclase known to synthesize c-di-GMP in *P. aeruginosa* (9, 18). *P. aeruginosa* expressing *tpbB* and a vector control were used for method validation. Overnight cultures of *P. aeruginosa* PAO1 strains grown in LB medium containing 100 µg/ml gentamicin were diluted with VBMM (Vogel-Bonner Minimal Medium) (1:100) containing 100 µg/ml gentamicin (Gold Biotechnology, St. Louis, MO) with or without 0.2% L-arabinose (Gold Biotechnology). Cells collected from aliquots (2 ml) of *P. aeruginosa* culture grown to mid log-phase (OD_{600}~0.6-0.7) were extracted with 100 µl of 0.6 M (final concentration) perchloric acid (Sigma Aldrich) (19) spiked with 100 nM [^{13}C]adenosine. The precipitate of the perchloric acid extraction was used to determine protein concentration by the BCA assay and sample normalization (20), and extracts were neutralized with 20 µl of 2.5 M potassium bicarbonate (Sigma Aldrich). Neutralized supernatants were stored at -80°C. The experiment was performed with three technical replicates for each of three independent biological replicates.

M. tuberculosis H37Ra

The protocol for c-di-GMP and other cyclic nucleotide/dinucleotide extraction was modified from a previous study (21). The initial culture of *M. tuberculosis* H37Ra was grown in Middlebrook 7H9 medium (Becton Dickinson, Franklin Lakes, NJ) supplemented with 0.2%

glycerol, 10% OADC, and 0.05% Tween 80 at 37ºC with shaking at 100 rpm to log phase. The log-phase *M. tuberculosis* culture was diluted with 7H9 medium supplemented with 0.2% glycerol and 10% OADC to OD_{600} 0.05 and grown at 37ºC with shaking at 100 rpm. *M. tuberculosis* cells were collected from 30 ml of cultures grown for one week and three weeks by spinning at 3,000 rpm, 4ºC for 10 min. The extraction was performed using 500 µl to 1 ml of 0.6 M perchloric acid (aapproximately 1 ml of 0.6 M perchloric acid per 1 g of bacterial cells) spiked with [^{13}C]adenosine. The precipitate of the perchloric acid extraction was used for determination of protein concentration by the BCA assay (ThermoFisher Scientific, Grand Island, NY) and sample normalization (20), and extracts were neutralized with 20 µl of 2.5 M potassium bicarbonate per 100 µl of 0.6 M perchloric acid solution used. Neutralized supernatants were stored at -80°C and lyophilized (Freezone by LABCONCO, Kansas City, MO). The lyophilized samples were resuspended with 200 µl LC-MS grade water that provided a final concentration of [^{13}C]adenosine at 200 nM. The experiment was performed with three technical replicates for each of two independent biological replicates. Workflow for cyclic nucleotides/dinucleotides extraction is shown in Appendix A.

3.2.6 Statistical analyses

The *p* values of the biological sample experiment were calculated by one-way or two-way analysis of variance (ANOVA) followed by multiple comparison using GraphPad Prism version 6.0 (GraphPad Software, Inc., San Diego, CA). Data were expressed as mean values ± SD. A *p* value < 0.05 was considered to be statistically significant.

3.3 Results

3.3.1 LC-MS and LC-MS/MS method performance

Cyclic nucleotides/dinucleotides play important roles in bacterial and eukaryotic cells. The development of a method that simultaneously detects these compounds would provide an advantage for the study of the level of these compounds in those organisms. The presence of purine or pyrimidine rings could result in weak basic characteristics in these compounds. These cyclic nucleotides/dinucleotides are polar compounds, making reverse phase chromatography a suitable technique for their separation. A reverse phase C18 column was used for this study. Solvent A of the mobile phase, which contained both acetic acid and ammonium acetate in water can promote the ionization of compounds having both basic and acid characteristics. ESI is valuable for analysis of these polar compounds. Initially, method optimization of practical liquid chromatographic conditions for simultaneous analyses of multiple cyclic nucleotide/dinucleotide compounds in a short chromatographic window was performed. The optimized gradient separated all eight target compounds within 8 min (Figure 3-1), including the isomers of cAMP and cGMP. The compound peak can be clearly identified by the retention time and its distinct *m/z* transitions performed by tandem mass spectrometry, LC-MS/MS. The precursors and product ions of each compound, and corresponding collision energy (CE) used for fragmentation are shown in Figure 3-2.

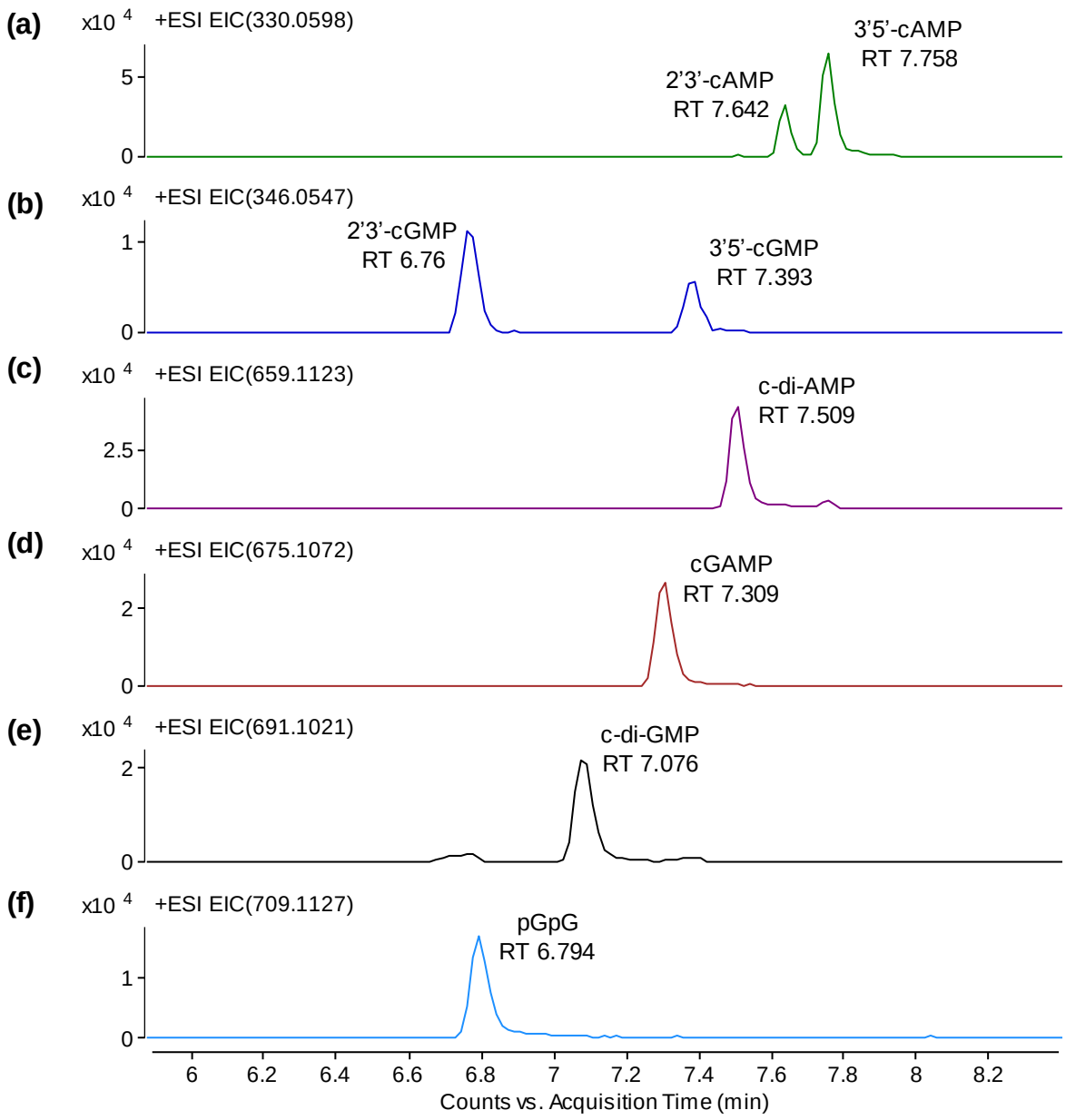
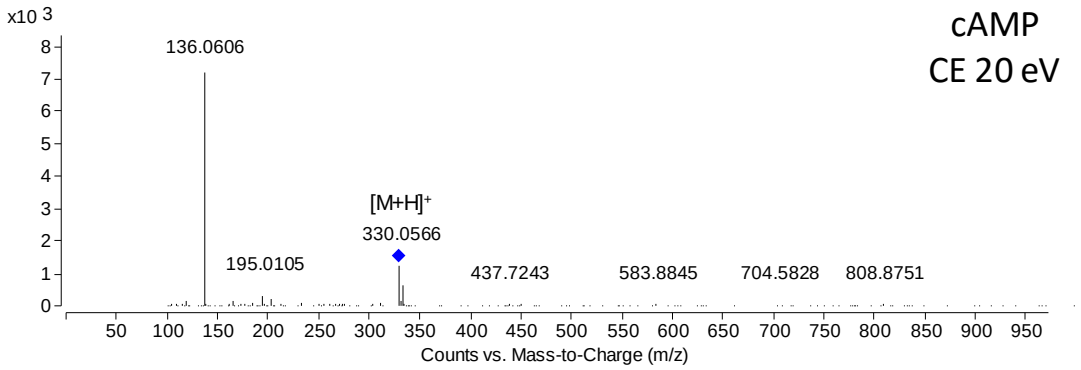
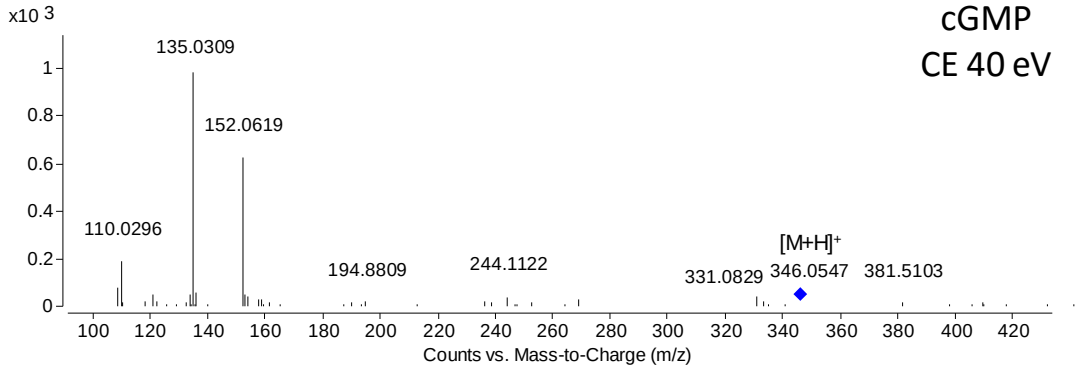


Figure 3-1. LC-MS chromatogram of cyclic nucleotide/dinucleotide standards
The representative chromatograms and retention times of 2'3'-cAMP (a), 3'5'-cAMP (a), 2'3'-cGMP (b), 3'5'-cGMP (b), c-di-AMP (c), cGAMP (d), c-di-GMP (e) and pGpG (f) analyzed in one analytical run.

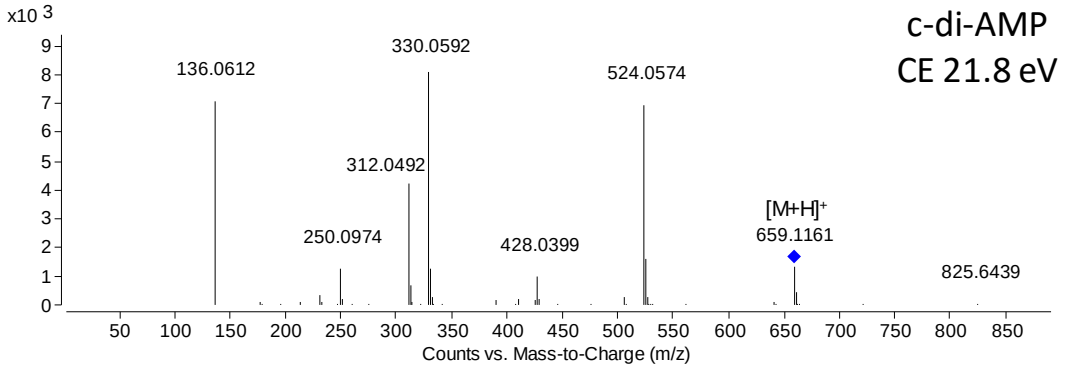
(a)


x10 3
136.0606
[M+H]+
330.0566
195.0105
437.7243
583.8845
704.5828
808.8751
cAMP
CE 20 eV
Counts vs. Mass-to-Charge (m/z)

(b)


x10 3
135.0309
152.0619
110.0296
194.8809
244.1122
[M+H]+
331.0829
346.0547
381.5103
cGMP
CE 40 eV
Counts vs. Mass-to-Charge (m/z)

(c)


x10 3
136.0612
330.0592
524.0574
312.0492
250.0974
428.0399
[M+H]+
659.1161
825.6439
c-di-AMP
CE 21.8 eV
Counts vs. Mass-to-Charge (m/z)

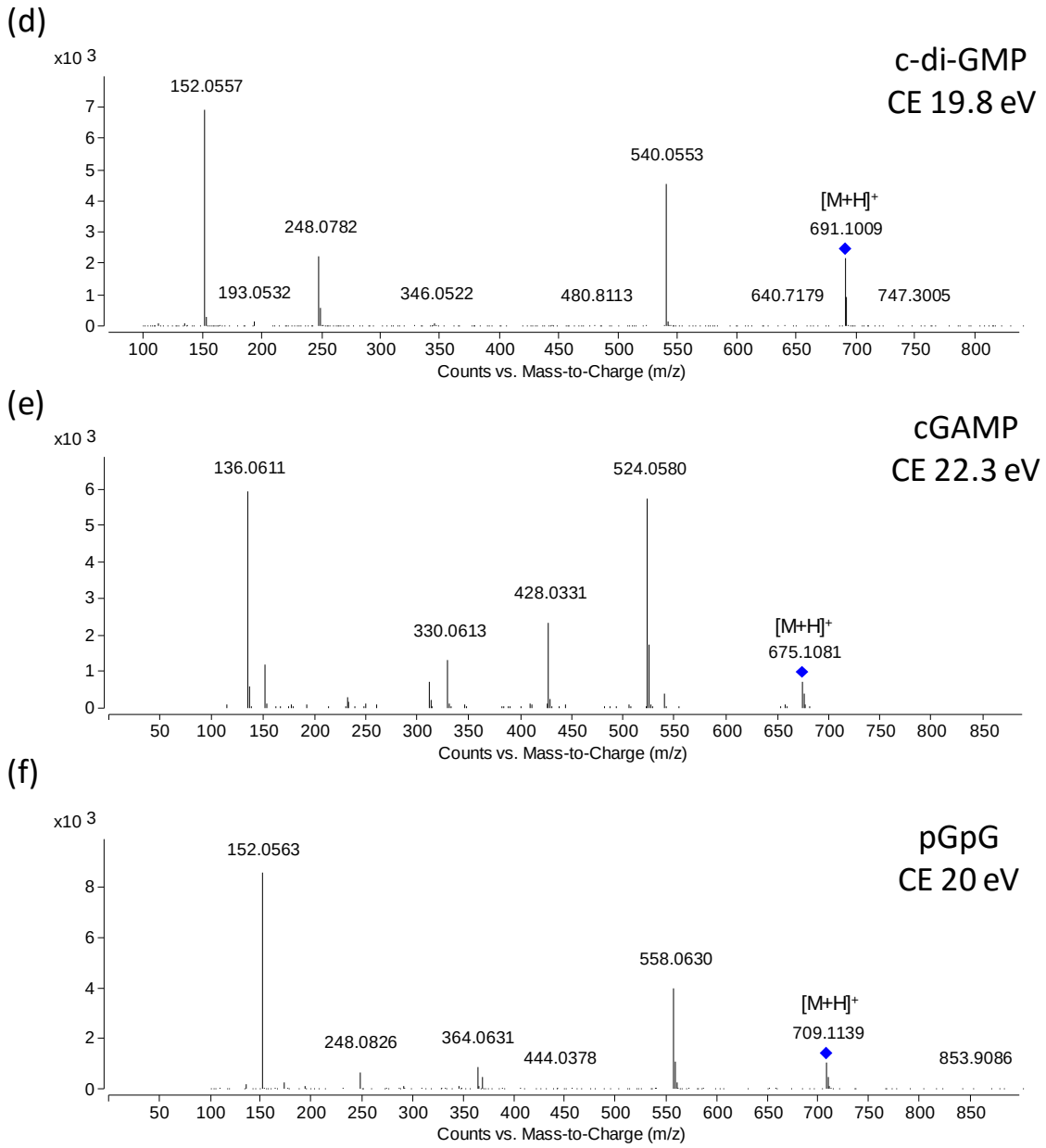


Figure 3-2. Precursors and product ions from MS/MS of cGMP, cAMP, c-di-AMP, c-di-GMP, cGAMP, and pGpG

3.3.2 LC-MS method calibration curve, LOD, and LOQ

The calibration curves for 2'3'-cAMP, 3'5'-cAMP, 2'3'-cGMP, 3'5'-cGMP, c-di-AMP, c-di-GMP, cGAMP, and pGpG were generated from dilutions of standard samples spiked with an internal standard. [^{13}C]adenosine was chosen since it is absent in bacteria and has relative chemical structure to target compounds. The peak area ratio of each compound and an internal standard was calculated and plotted against the compound concentration using linear regression analysis. The calibration curve of each compound (Figure 3-3) was obtained from six concentration samples analyzed in replicates of three. This method can only detect 2'3'-cGMP at a concentration equal to or higher than 50 nM. Therefore, the calibration curve of 2'3'-cGMP was generated from five concentration samples ranging from 50 nM-1 µM. The LOD and LOQ of all compounds are shown in Table 3-3. For the calibration curves, the linear correlation (R^2) for all compounds was greater than 0.99 with intercepts close to zero as shown in Table 3-3.

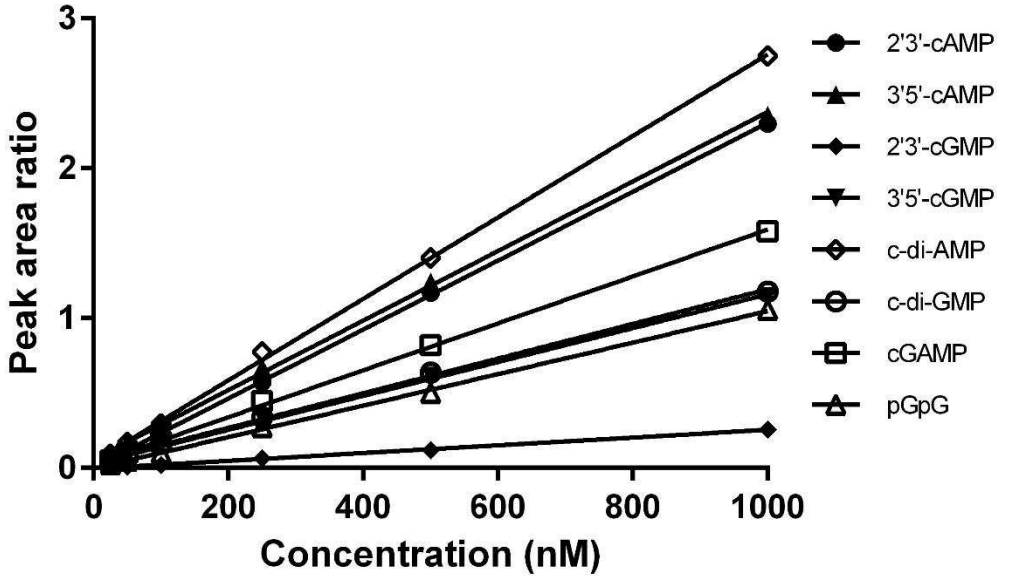


Figure 3-3. Calibration curves of cyclic nucleotides/dinucleotides
Calibration curves for quantitative determination of 2'3'-cAMP, 3'5'-cAMP, 3'5'-cGMP, c-di-GMP, c-di-AMP, cGAMP, and pGpG. Each data point is the average of triplicate analytical run.

Table 3-3. Calibration curve data, LOD, and LOQ

Compound	Slope, a*	Intercept, b*	Correlation coefficient, R^2	LOD	LOQ
2'3'-cAMP	0.0099	-0.0342	0.9983	50 fmole	200 fmole
3'5'-cAMP	0.0101	0.0914	0.9979	50 fmole	200 fmole
2'3'-cGMP	0.001	-0.00746	0.9976	1 pmole	2 pmole
3'5'-cGMP	0.0036	0.0326	0.9977	200 fmole	1 pmole
c-di-AMP	0.0049	0.065	0.9985	100 fmole	500 fmole
c-di-GMP	0.0022	0.0394	0.9989	500 fmole	1 pmole
cGAMP	0.0031	0.0454	0.9988	200 fmole	1 pmole

*Values are the mean from three calibration curves, slope (a) and intercept (b) refer to the regression equation, y = ax+b. These data from the calibration curves generated from concentrations 25 nM-1 µM except 2'3'-cGMP calibration curve generated from 50 nM-1 µM.

3.3.3 Validation

QC samples were prepared at three concentrations representing low, intermediate, and high concentrations in the calibration curve, and these were used for determining the accuracy and precision of the method. The intra-day accuracy and precision were determined from the variability of replicates of QC samples (n=5) analyzed within the same day while the inter-day accuracy and precision were calculated from analytical runs performed on three different days (22). Specifically, accuracy was determined as the percent difference between the mean concentration of the QC samples that could be analyzed as intra- or inter-day analytical run per analytical run and the nominal concentration. The coefficient of variation (CV) or precision was determined by percent of the ratio of the standard deviation and the mean concentration of the replicates of QC in the same analytical run (22). The intra- and inter-day accuracy of all compounds ranged between 80-120% and the precision ranged between 0.3-11% as shown in Table 3-4. These results indicate that the method is highly reproducible.

Table 3-4. Intra- and inter-day accuracy and precision

		QC 1: 40 nM			QC 2: 200 nM			QC 3: 800 nM		
Compound		Observed concentration (nM)	Accuracy (%)	Precision (%)	Observed concentration (nM)	Accuracy (%)	Precision (%)	Observed concentration (nM)	Accuracy (%)	Precision (%)
2'3' cAMP	intraday 1	45.77 ± 2.48	114.43	5.42	192.17 ± 9.13	96.09	4.75	744.96 ± 12.84	93.12	1.72
	intraday 2	44.62 ± 1.37	111.56	3.07	198.53 ± 6.96	99.27	3.51	811.42 ± 10.13	101.43	1.25
	intraday 3	43.03 ± 0.86	107.58	2.00	206.75 ± 11.35	103.38	5.49	799.38 ± 5.66	99.92	0.71
	interday	44.47 ± 1.37	111.19	3.10	199.15 ± 7.3	99.58	3.67	785.26 ± 35.41	98.16	4.51
3'5' cAMP	intraday 1	36.02 ± 1.46	91.57	5.67	229.32 ± 3.17	114.73	1.39	864.27 ± 5.23	107.48	1.41
	intraday 2	39.68 ± 1.75	95.66	10.15	207.83 ± 8.9	103.32	4.53	801.09 ± 2.61	100.5	0.94
	intraday 3	42.72 ± 2.49	100.97	6.10	202.03 ± 12.02	102.65	7.04	800.12 ± 12.06	101.41	3.67
	interday	39.47 ± 3.35	97.76	7.64	213.06 ± 14.38	106.90	6.35	821.82 ± 36.76	103.13	3.68
2'3' cGMP	intraday 1	-	-	-	199.39 ± 2.3	99.69	1.15	783.07 ± 13.99	97.89	1.79
	intraday 2	-	-	-	197.76 ± 3.43	98.88	1.73	833.53 ± 2.89	104.19	0.35
	intraday 3	-	-	-	213.63 ± 24.91	106.81	11.66	786.35 ± 6.81	98.29	0.87
	interday	-	-	-	203.59 ± 8.73	101.8	4.29	800.98 ± 28.23	100.12	3.52
3'5' cGMP	intraday 1	38.85 ± 1.89	97.13	4.86	218.58 ± 4.94	109.62	2.37	810.62 ± 6.66	101.33	2.11
	intraday 2	39.86 ± 1.58	99.64	3.96	215.62 ± 7.9	108.44	3.91	829.58 ± 8.97	103.7	1.64
	intraday 3	35.3 ± 0.31	88.24	0.89	216.9 ± 3.2	108.7	1.58	821.48 ± 2.16	102.68	0.68
	interday	38 ± 2.4	95	6.31	217.03 ± 1.49	108.92	0.57	820.51 ± 15.02	102.57	1.16
c-di-AMP	intraday 1	39.69 ± 0.88	99.21	2.21	204.66 ± 4.98	102.33	2.43	762.75 ± 21.04	95.34	2.76
	intraday 2	40.63 ± 0.68	101.57	1.67	205.06 ± 1.66	102.53	0.81	798.98 ± 6.23	99.87	0.78
	intraday 3	39.55 ± 1.22	98.88	3.09	221.84 ± 2.19	110.92	0.9944	775.14 ± 1.15	96.89	0.15
	interday	39.96 ± 0.59	99.89	1.47	210.52 ± 9.81	105.26	4.66	778.95 ± 18.41	97.37	2.36
c-di-GMP	intraday 1	39.12 ± 0.65	97.79	1.67	220.6 ± 4.52	110.30	2.05	793.26 ± 6.05	99.16	0.76
	intraday 2	40.16 ± 2.05	100.40	5.10	209.47 ± 6.51	104.74	3.11	849.16 ± 6.78	106.14	0.8
	intraday 3	32.21 ± 2.71	80.53	8.44	233.56 ± 4.76	116.78	2.04	820.34 ± 17.64	102.54	2.15
	interday	37.16 ± 4.32	92.91	11.62	221.21 ± 12.06	110.60	5.45	820.91 ± 27.95	102.61	3.41

cGAMP	intraday 1	38.78 ± 2.35	96.95	6.05	220.34 ± 9.68	110.17	4.39	792.24 ± 6.4	99.03	0.81
	intraday 2	40.07 ± 2.06	100.17	5.14	223.21 ± 5.67	111.6	2.54	845.14 ± 19.67	105.64	2.33
	intraday 3	40.91 ± 0.85	102.28	2.10	218.18 ± 3.6	109.09	1.65	789.32 ± 5.63	98.66	0.71
	interday	39.92 ± 1.07	99.80	2.69	220.58 ± 2.5	110.29	1.14	808.9 ± 31.41	101.11	3.88
pGpG	intraday 1	45.38 ± 3.01	113.45	6.64	205.2 ± 10.74	102.60	5.24	768.63 ± 14.68	96.08	1.91
	intraday 2	42.94 ± 0.87	107.36	2.02	193.72 ± 8.37	96.86	4.32	758.27 ± 14.74	94.78	1.94
	intraday 3	42.47 ± 3.68	106.16	8.65	202.51 ± 9.74	101.25	4.81	731.4 ± 9.7	91.43	1.33
	interday	43.6 ± 1.56	108.99	3.59	200.47 ± 6	100.24	2.99	752.77 ± 19.22	94.1	2.56

* Accuracy (%) represents the difference between the measured and the true concentrations of the QC samples. Precision or coefficient of variation (CV) (%) is the average standard variation of the measurements (n=5 for intra-day assay and n=3 for inter-day assay).

3.3.4 Application of LC-MS and LC-MS/MS methods

Application of the optimized method was conducted to detect cyclic nucleotides/dinucleotides in bacteria. Perchloric acid was used to extract and stabilize those compounds spiked with the [^{13}C]adenosine internal standard, which had the same concentration as in the calibration curve. Extracts from *P. aeruginosa* PAO1 with an arabinose-inducible pJN105 vector control and constitutively expressing *tpbB* encoding diguanylate cyclase to produce c-di-GMP were analyzed by optimized LC-MS method. Arabinose was utilized to induce high-level expression of *tpbB.* The experiment was performed in three independent bacterial cultures with technical replicates for each culture. The optimized method can detect changes in c-di-GMP level for extracts from the *P. aeruginosa* PAO1 pJN105 vector control and *P. aeruginosa* expressing *tpbB* in the presence or absence of arabinose at levels of 16.59, 325.72, and 711.53 nmole per mg protein, respectively (Figure 3-4). The method was also applied to detect 3’5’-cAMP and c-di-AMP in *M. tuberculosis* H37Ra, and changes in 3’5’-cAMP and c-di-AMP levels along with growth time were determined. The experiment was performed in two independent bacterial cultures with three technical replicates for each culture. Extracts from one-week (log-phase) and three-week (late stationary phase) cultures were analyzed by the LC-MS instrument. The level of 3’5’-cAMP was significantly decreased in three-week cultures (0.44 pmole/mg protein) as compared to one-week cultures (19.82 pmole/mg protein), while c-di-AMP levels were increased in three-week cultures (50.47 pmole/mg protein) as compared to one-week cultures (2.49 pmole/mg protein) (Figure 3-5). This result correlates to previous studies that have shown increased c-di-AMP level in the late stationary phase of growth phase compared to the early log phase (23).

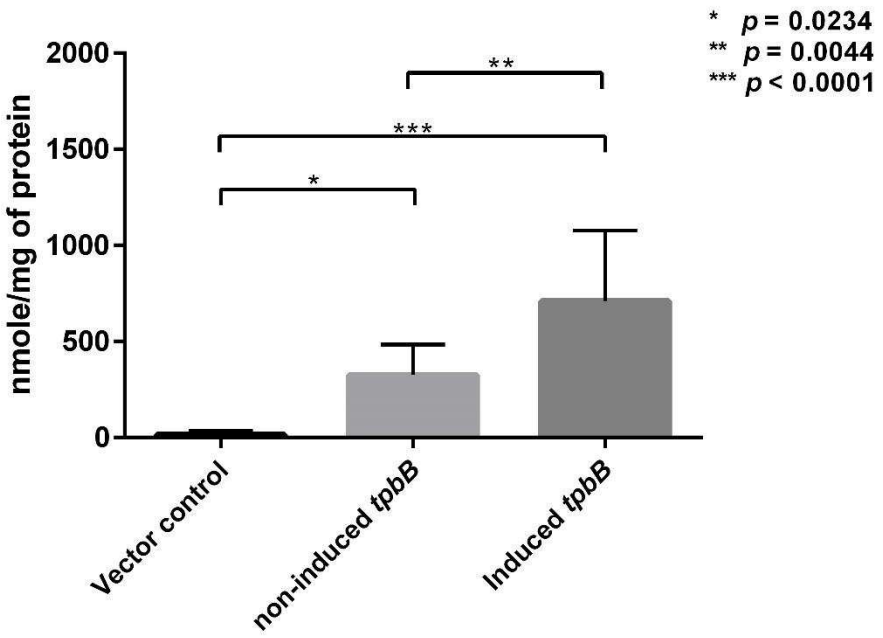


Figure 3-4. c-di-GMP detection in *P. aeruginosa* PAO1 strains

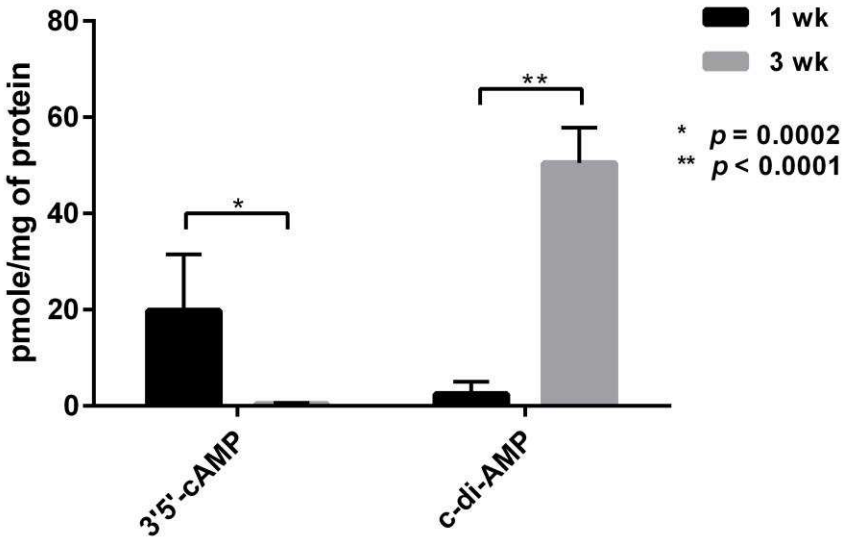


Figure 3-5. 3’5’-cAMP and c-di-AMP detection in *M. tuberculosis* H37Ra

3.4 Discussion

Cyclic nucleotides and cyclic dinucleotides regulate several biological and physiological roles in various organisms including bacteria and eukaryotic cells. To obtain a better understanding of the regulatory mechanisms and signaling of these compounds in those organisms, a highly selective, specific, reliable and sensitive detection method for *in vitro* and *in vivo* experiments is necessary. LC-MS method development for cyclic nucleotides/dinucleotides detection in biological samples would provide advantages for identification of these compounds. This is particularly true when physiological or biological factors that can have an effect on the level of these compounds in bacteria must be investigated. LC-MS and LC-MS/MS are sensitive, specific and fast techniques. Positive-ion mode ESI-MS was used in this study since it provides more sensitivity than the negative ion mode for c-di-GMP. The ESI has been used to detect nucleotides and cyclic nucleotides in previous studies and provides high sensitivity (16, 24-27). Separation and detection of compounds in the samples was accomplished by liquid chromatography coupled with mass spectrometry. Separation of several nucleotides or cyclic nucleotides/dinucleotides is complicated because of their relatively similar structures (28). The flow rate of the mobile phase also has an effect on the chromatographic separation of all compounds. A flow rate of 350 µl/ml was selected since it provided the best separation of all compounds, especially for isomers like 2'3'- and 3'5'-cAMP and 2'3'- and 3'5'-cGMP which have the same *m/z* and *m/z* fragments for MS/MS. However, a higher flow rate may provide less sensitivity than lower flow rates. Therefore, modification of the method can be performed depending on the objective of the study.

Although the method optimization was mainly performed based on the c-di-GMP compound, the method developed in this study can simultaneously detect several compounds in a short time in the same analytical run. Specifically, it can differentiate the 2'3'- and 3'5'- isomer

compounds that have the same *m/z* and ion fragments for MS/MS. In addition, the method is reproducible and sensitive with high precision (coefficient of variance <15%) and accuracy (80-120%). The concentration of internal standard can be adjusted depending on the concentration of compound in biological samples. LC-MS/MS could be applied when confirmation of the compound is needed. However, detection of 2'3'-cGMP by this method lacked the high sensitivity achieved with other compounds.

This method can be applied to complex biological samples such as extracts from bacteria that have matrix effects. Specifically, the alteration of compound levels can be distinguished by this method as shown in the detection of c-di-GMP in different conditions in *P. aeruginosa* PAO1, *E. coli* (data not shown) and detection of 3'5'-cAMP and c-di-AMP in various growth phases of *M. tuberculosis* H37Ra. Additionally, *Mycobacterium* spp. have a unique cell wall structure, and the technique used to extract and detect cyclic nucleotides/dinucleotides from *M. tuberculosis* in this study may be useful for application to other mycobacteria species. The inability to detect c-di-GMP in *M. tuberculosis* results from the low level of c-di-GMP in this bacterium that has only one GGDEF-containing protein (Rv1354c) (29, 30) or rapid degradation of c-di-GMP. The 2'3' cGAMP is synthesized in response to the presence of microbial or self DNA in the cytosol of mammalian cells (31). Therefore, further investigation to analyze extracts from eukaryotic cells is also required to assess utility for cGAMP detection. In summary, this method is fast, sensitive and has high reproducibility for detection and quantitation of cyclic nucleotides and dinucleotides which can be applied in biological samples.

REFERENCES

1. Christen M, Kulasekara HD, Christen B, Kulasekara BR, Hoffman LR, Miller SI. 2010. Asymmetrical distribution of the second messenger c-di-GMP upon bacterial cell division. *Science* 328:1295-1297.

2. Ho CL, Chong KS, Oppong JA, Chuah ML, Tan SM, Liang ZX. 2013. Visualizing the perturbation of cellular cyclic di-GMP levels in bacterial cells. *J Am Chem Soc* 135:566-569.

3. Nakayama S, Kelsey I, Wang J, Roelofs K, Stefane B, Luo Y, Lee VT, Sintim HO. 2011. Thiazole orange-induced c-di-GMP quadruplex formation facilitates a simple fluorescent detection of this ubiquitous biofilm regulating molecule. *J Am Chem Soc* 133:4856-4864.

4. Nakayama S, Roelofs K, Lee VT, Sintim HO. 2012. A c-di-GMP-proflavine-hemin supramolecular complex has peroxidase activity-implication for a simple colorimetric detection. *Mol Biosyst* 8:726-729.

5. Nakayama S, Luo Y, Zhou J, Dayie TK, Sintim HO. 2012. Nanomolar fluorescent detection of c-di-GMP using a modular aptamer strategy. *Chem Commun (Camb)* 48:9059-9061.

6. Gu H, Furukawa K, Breaker RR. 2012. Engineered allosteric ribozymes that sense the bacterial second messenger cyclic diguanosyl 5'-monophosphate. *Anal Chem* 84:4935-4941.

7. Xie Q, Zhao F, Liu H, Shan Y, Liu F. 2015. A label-free and self-assembled electrochemical biosensor for highly sensitive detection of cyclic diguanylate monophosphate (c-di-GMP) based on RNA riboswitch. *Anal Chim Acta* 882:22-26.

8. Schmidt AJ, Ryjenkov DA, Gomelsky M. 2005. The ubiquitous protein domain EAL is a cyclic diguanylate-specific phosphodiesterase: Enzymatically active and inactive EAL domains. *Journal of Bacteriology* 187:4774-4781.

9. Kulasekara H, Lee V, Brencic A, Liberati N, Urbach J, Miyata S, Lee DG, Neely AN, Hyodo M, Hayakawa Y, Ausubel FM, Lory S. 2006. Analysis of *Pseudomonas aeruginosa* diguanylate cyclases and phosphodiesterases reveals a role for bis-(3'-5')-cyclic-GMP in virulence. *Proc Natl Acad Sci U S A* 103:2839-2844.

10. Romling U, Galperin MY, Gomelsky M. 2013. Cyclic di-GMP: the first 25 years of a universal bacterial second messenger. *Microbiol Mol Biol Rev* 77:1-52.

11. Simm R, Morr M, Kader A, Nimtz M, Romling U. 2004. GGDEF and EAL domains inversely regulate cyclic di-GMP levels and transition from sessility to motility. *Mol Microbiol* 53:1123-1134.

12. Simm R, Morr M, Remminghorst U, Andersson M, Romling U. 2009. Quantitative determination of cyclic diguanosine monophosphate concentrations in nucleotide extracts of bacteria by matrix-assisted laser desorption/ionization-time-of-flight mass spectrometry. *Anal Biochem* 386:53-58.

13. Marvin LF, Roberts MA, Fay LB. 2003. Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry in clinical chemistry. *Clin Chim Acta* 337:11-21.

14. Jurinke C, Oeth P, Boom D. MALDI-TOF mass spectrometry. *Molecular Biotechnology* 26:147-163.

15. Pitt JJ. 2009. Principles and applications of liquid chromatography-mass spectrometry in clinical biochemistry. *Clin Biochem Rev* 30:19-34.

16. Burhenne H, Kaever V. 2013. Quantification of cyclic dinucleotides by reversed-phase LC-MS/MS, p 27-37. *In* Gehring C (ed), Cyclic nucleotide signaling in plants: methods and protocols doi:10.1007/978-1-62703-441-8_3. Humana Press, Totowa, NJ.

17. Newman JR, Fuqua C. 1999. Broad-host-range expression vectors that carry the L-arabinose-inducible *Escherichia coli araBAD* promoter and the *araC* regulator. *Gene* 227:197-203.

18. Ueda A, Wood TK. 2009. Connecting quorum sensing, c-di-GMP, pel polysaccharide, and biofilm formation in *Pseudomonas aeruginosa* through tyrosine phosphatase TpbA (PA3885). *PLoS Pathog* 5:(6): e1000483. doi:1000410.1001371/journal.ppat.1000483.

19. Hickman JW, Harwood CS. 2008. Identification of FleQ from *Pseudomonas aeruginosa* as a c-di-GMP-responsive transcription factor. *Mol Microbiol* 69:376-389.

20. Irie Y, Parsek MR. 2014. LC/MS/MS-based quantitative assay for the secondary messenger molecule, c-di-GMP. *Methods Mol Biol* 1149:271-279.

21. Weinhouse H, Sapir S, Amikam D, Shilo Y, Volman G, Ohana P, Benziman M. 1997. c-di-GMP-binding protein, a new factor regulating cellulose synthesis in *Acetobacter xylinum*. *FEBS Lett* 416:207-211.

22. Islam NM, Yoo HH, Lee MW, Dong MS, Park YI, Jeong HS, Kim DH. 2008. Simultaneous quantitation of five flavonoid glycosides in Herba *Epimedii* by high-performance liquid chromatography-tandem mass spectrometry. *Phytochem Anal* 19:71-77.

23. Dey B, Dey RJ, Cheung LS, Pokkali S, Guo H, Lee J-H, Bishai WR. 2015. A bacterial cyclic dinucleotide activates the cytosolic surveillance pathway and mediates innate resistance to tuberculosis. *Nature medicine* 21:401-406.

24. Lorenzetti R, Lilla S, Donato JL, de Nucci G. 2007. Simultaneous quantification of GMP, AMP, cyclic GMP and cyclic AMP by liquid chromatography coupled to tandem mass spectrometry. *J Chromatogr B Analyt Technol Biomed Life Sci* 859:37-41.

25. Cai Z, Song F, Yang MS. 2002. Capillary liquid chromatographic-high-resolution mass spectrometric analysis of ribonucleotides. *J Chromatogr A* 976:135-143.

26. Spangler C, Bohm A, Jenal U, Seifert R, Kaever V. 2010. A liquid chromatography-coupled tandem mass spectrometry method for quantitation of cyclic di-guanosine monophosphate. *J Microbiol Methods* 81:226-231.

27. Jia X, Fontaine BM, Strobel F, Weinert EE. 2014. A facile and sensitive method for quantification of cyclic nucleotide monophosphates in mammalian organs: basal levels of eight cNMPs and identification of 2',3'-cIMP. *Biomolecules* 4:1070-1092.

28. Kehr J, Chavko M. 1986. Separation of nucleotides by reversed-phase high-performance liquid chromatography: Advantages and limitations. *Fresenius' Zeitschrift für analytische Chemie* 325:466-469.

29. Hong Y, Zhou X, Fang H, Yu D, Li C, Sun B. 2013. Cyclic di-GMP mediates *Mycobacterium tuberculosis* dormancy and pathogenecity. *Tuberculosis (Edinb)* 93:625-634.

30. Gupta K, Kumar P, Chatterji D. 2010. Identification, activity and disulfide connectivity of c-di-GMP regulating proteins in *Mycobacterium tuberculosis*. *PLoS One* 5:(11): e15072. doi:15010.11371/journal.pone.0015072.

31. Zhang X, Shi H, Wu J, Zhang X, Sun L, Chen C, Chen ZJ. 2013. Cyclic GMP-AMP containing mixed phosphodiester linkages is an endogenous high-affinity ligand for STING. *Mol Cell* 51:226-235.

CHAPTER 4 ML1419C AND ML0397C FUNCTION AS DIGUANYLATE CYCLASES TO PRODUCE C-DI-GMP[1]

4.1 Introduction

Generally, bacteria contain more than one protein containing GGDEF and/or EAL, HD-GYP domains. The DGC proteins that synthesize c-di-GMP in various bacteria can be differently arranged as a single GGDEF protein, or in tandem such as hybrid GGDEF-EAL proteins, and hybrid GGDEF-HD-GYP domains. GGDEF and EAL domains are often observed on the same protein as parts of multidomain proteins. Those proteins may have both DGC and PDE functions or either DGC or PDE formation. Significantly, the GGDEF-EAL proteins are conserved in various bacteria. As approximately 1/3 of all GGDEF domains and 2/3 of all EAL domains are conserved as hybrid GGDEF-EAL proteins (1). Proteins containing GGDEF, EAL or HD-GYP domains are typically linked with other protein domains such as sensory domains and transmembrane domains (2, 3). Although several bacteria have more than one protein with GGDEF, EAL or HD-GYP domains, not all these proteins are active. They can be degenerate proteins that function in other cellular mechanisms such as c-di-GMP receptors or protein-protein interactions mediated events. Additionally, when multiple DGC or PDE proteins are present in the same bacteria, they should not be active at the same time and location, and if they are present at the same time and under same condition, they likely regulate different target outputs (1, 4).

---

[1] Parts of this chapter are in a manuscript submitted to Microbiology: Suwatchareeporn Rotcheewaphan, John T. Belisle, Kristofor J. Webb, Hee-Jin Kim, John S. Spencer, and Bradley R. Borlee. Diguanylate cyclase activity of the *Mycobacterium leprae* T cell antigen ML1419c.

c-di-GMP has been well-described in several bacteria including *P. aeruginosa* and *E. coli*. *P. aeruginosa* is a model organism for studying c-di-GMP impact on biofilm formation (5). This organism is an opportunistic pathogen of patients with cystic fibrosis (6), septicemia and urinary tract infections in catheterized patients (7). The genome of *P. aeruginosa* PAO1 encodes 41 proteins predicted to metabolize intracellular c-di-GMP, including 17 GGDEF-containing proteins, five EAL domain-containing proteins, 16 proteins that contain composite GGDEF-EAL domains, and three proteins with HD-GYP domains (1, 8-10). For HD-GYP containing proteins, only two proteins (PA4108 and PA4781) harbor HD-GYP domains and another protein (PA2572) contains a domain with variant residues (YN-GYP) (9). In general, high levels of intracellular c-di-GMP in *P. aeruginosa* promote exopolysaccharide (EPS) production, and decrease flagella- and pilus-mediated motility. In addition, c-di-GMP increases the production of biofilm matrix components in *P. aeruginosa* including the Pel and Psl polysaccharides and a biofilm-associated adhesin whereas low intracellular c-di-GMP increases motility and inhibits EPS production (2, 4, 5, 11, 12). Two HD-GYP proteins of *P. aeruignosa*, PA4108 and PA4781, have been shown to regulate biofilm formation, swarming motility, virulence factor synthesis, and virulence of *P. aeruginosa* in a manner consistent with c-di-GMP PDE activity (9).

E. coli is another bacterium in which c-di-GMP plays several biological roles. *E. coli* is one of the most widely-used systems for recombinant heterogeneous gene expression (13). Similar to *P. aeruginosa*, *E. coli* K-12 contains more than one proteins with GGDEF and/or EAL domains. This bacterium has 12 proteins with a GGDEF domain, seven proteins with a GGDEF-EAL domain and 10 proteins with an EAL domain. In *E. coli*, increased c-di-GMP levels leads to decreased bacterial motility, increased cellulose and curli fimbriae formation, and reduced virulence (14). Based on these phenotypic properties associated with c-di-GMP, the current studies

used *P. aeruginosa* and *E. coli* as model organisms to study recombinant *M. leprae* proteins related to c-di-GMP.

Bioinformatics analyses of the annotated *M. leprae* genome revealed three putative DGC proteins, all possessing a conserved GGDEF domain. One of the putative DGC coding sequences of *M. leprae*, ML1419c, was previously shown to be expressed when *M. leprae* was experimentally infected in the mouse footpad (15). Immunological studies also revealed a strong antigen-specific T cell response to peptides of ML1419c in paucibacillary patients and in the household contacts of multibacillary patients (16). This demonstrate that *M. leprae* produces ML1419c in the early stages of leprosy; however, the physiological function of this protein was unknown. Another putative DGC protein is ML0397c. Previous studies have been shown that the gene encoding ML0397c is also expressed during infection in animal models (15). Similar to *ml1419c*, it may be involved in leprosy disease but the physiological function needed to be investigated (15, 16).

The inability to culture *M. leprae in vitro* or subject it to genetic manipulation are major impediments in characterizing the physiological function of *M. leprae* proteins. To overcome these limitations, we chose to express *ml1419c* and *ml0397c* in an alternative heterologous host, *P. aeruginosa* PAO1 and *E. coli*, in which the phenotypes associated with modulating c-di-GMP levels are well documented (2, 5, 8, 12, 17-22). This study demonstrated the enzymatic function of ML1419c and ML0397c and provides strong evidence that *M. leprae* is capable of producing the second messenger c-di-GMP.

4.2 Materials and methods

4.2.1 Bacterial strains, genomic DNA, and growth conditions

P. aeruginosa PAO1 and recombinant strains were grown at 37°C in either Lennox LB medium or VBMM (Vogel-Bonner Minimal Medium) with L-arabinose (Gold Biotechnology) for inducible expression experiments. *E. coli* strain BL21(DE3) and strain BL21(DE3) pLysS (Invitrogen, Grand Island, NY) were grown in Miller LB medium at 37°C with isopropyl β-D-thiogalactoside (IPTG) for inducible expression experiments. Gentamicin (100 µg/ml) (Gold Biotechnology) was used for selection of recombinant *P. aeruginosa* strains. Kanamycin (50 µg/ml) (Sigma Aldrich), gentamicin (10 µg/ml) and chloramphenicol (34 µg/ml) (Sigma Aldrich) were used for selection of recombinant *E. coli* strains. The following reagent was obtained through Biodefense and Emerging Infections (BEI) Research Resources Repository (Manassas, VA), NIAID, NIH: Genomic DNA from *M. leprae*, Strain NHDP, NR-19350 and Strain Thai-53, NR-19352.

4.2.2 Construction of plasmids of *ml1419c*, *ml1750c*, and *ml0397c* for *P. aeruginosa* PAO1

Heterologous gene expression in *P. aeruginosa* PAO1 was accomplished by using the arabinose-inducible vector pJN105 containing P_{BAD} promoter (23). *M. leprae ml1419c* and *ml1750c* were amplified by PCR from NHDP63 genomic DNA of *M. leprae* with Q5 high fidelity DNA polymerase (New England BioLabs, Ipswich, MA) using *ml1419c* pJN105 forward and *ml1419c* pJN105 reverse primers; and *ml1750c* pJN105 forward and *ml1750c* pJN105 reverse primers, respectively (Table 4-1). *M. leprae* full length *ml0397c* and *ml0397c* containing only PAS and GGDEF domains (PAS-GGDEF *ml0397c*) were amplified by PCR from Thai 53 genomic DNA of *M. leprae* with Q5 high fidelity DNA polymerase using *ml0397c* pJN105 forward primer,

and FL-*ml0397c* pJN105 reverse primer or PG-*ml0397c* pJN105 reverse primer, respectively (Figure 4-1). Underlined sequences in *ml1419c*, FL-*ml0397c* and PG-*ml0397c* primers represent *Eco*RI and *Spe*I sites in the forward and reverse primers, respectively. Underlined sequences in *ml1750c* primers represent *Spe*I and *Sac*I sites in the forward and reverse primers, respectively. Bolded sequence represents inclusion of an optimized ribosome binding site (24). The 1692-bp *ml1419c*, 1809-bp full length *ml0397c*, and 900-bp PAS-GGDEF *ml0397c* fragments were cloned into pJN105 (23) using the *Eco*RI and *Spe*I sites to generate pMRLB105, pMRLB110 and pMRLB111, respectively. The 1872-bp *ml1750c* gene fragment was cloned into pJN105 using *Spe*I and *Sac*I sites to generate pMRLB112. Plasmids were sent for DNA sequencing at Proteomics and Metabolomics facility (PMF) at Colorado State University. All primers and plasmids used in this study are shown in Table 4-1 and Table 4-2, respectively.

4.2.3 Construction of plasmids of *ml1419c*, *ml1750c*, and *ml0397c* for *E. coli*

Gene expression of *ml1419c*, *ml1750c*, and *ml0397c* in *E. coli* BL21(DE3) or *E. coli* BL21(DE3) pLysS was achieved by using the IPTG inducible vector, pET28a. *M. leprae ml1419c*, full length *ml0397c* and PAS-GGDEF *ml0397c* were amplified by PCR from *M. leprae* genomic DNA with Q5 high fidelity DNA polymerase using *ml1419c* pET28a forward and *ml1419c* pET28a reverse primers; and *ml0397c* pET28a forward primer and FL-*ml0397c* pET28a reverse or PG-*ml0397c* pET28a reverse primers, respectively (Table 4-1). The *ml750c* was amplified by PCR from *M. leprae* genomic DNA with GC-RICH PCR System (Roche, Indianapolis, IN) using *ml1750c* pET28a forward and *ml1750c* pET28a reverse primers. Underlined sequences represent *Nde*I and *Hin*dIII sites in the forward and reverse primers, respectively. The 1692-bp *ml1419c*, 1872-bp *ml1750c*, 1809-bp full length *ml0397c* and 900-bp PAS-GGDEF *ml0397c* fragments were

cloned into pET28a using the *Nde*I and *Hin*dIII sites to generate pMRLB109, pMRLB113, pMRLB114 and pMRLB115, respectively, with N-terminal His_6 tag.

To express PAS-GGDEF *ml0397c* in *E. coli* for protein purification, PCR amplification was performed using *M. leprae* genomic DNA and the PG-*ml0397c* pET28a forward-2 and PG-*ml0397c* pET28a reverse-2 primers. Underlined sequences represent *Nde*I and *Hin*dIII sites in the forward and reverse primers, respectively. The 927-bp PAS-GGDEF *ml0397c* fragment was cloned into the expression vector pET28a using the *Nde*I and *Hin*dIII sites to generate pMRLB116. All plasmids were sent for DNA sequencing at PMF at Colorado State University. All primers and plasmids used in this study are shown in Table 4-1 and Table 4-2, respectively.

4.2.4 Site-directed mutagenesis of *ml1419c* and *ml0397c*

Site-directed mutagenesis of *ml1419c* was accomplished with the QuikChange Lightning Site-Directed Mutagenesis kit (Agilent Technologies) following the manufacturer's recommendations. The *ml1419c*$_{\Delta GGDEF}$ forward and *ml1419c*$_{\Delta GGDEF}$ reverse primers (Table 4-1) were used to generate *ml1419c* sequences encoding proteins with a deletion of the $_{472}$GGDEF$_{476}$ motif (ML1419c$_{\Delta GGDEF}$). The *ml1419c*$_{\Delta GGDEF}$ construct was digested with *Eco*RI and *Spe*I endonucleases and cloned into the pJN105 to generate pMRLB108. The plasmid pJN1120 that expresses *tpbB* (PA1120), a well-characterized diguanylate cyclase from *P. aeruginosa*, was used as a positive control (8, 25) and pJN105 was used as a negative control.

Site-directed mutagenesis of *ml0397c* was accomplished with the QuikChange Lightning Site-Directed Mutagenesis kit following the manufacturer's recommendations. The PG-*ml0397c*$_{\Delta GGDEF}$ forward and PG-*ml0397c*$_{\Delta GGDEF}$ reverse primers (Table 4-1) were used to generate *ml0397c* sequences encoding proteins with a deletion of the $_{219}$GGDEF$_{223}$ motif (PAS-GGDEF

ML0397c$_{\Delta GGDEF}$). The PAS-GGDEF *ml0397c*$_{\Delta GGDEF}$ construct was digested with *Nde*I and *Hin*dIII sites endonucleases and cloned into the pET28a to generate pMRLB117. All constructs were confirmed by nucleotide sequencing. All primers and plasmids used in this study are shown in Table 4-1 and Table 4-2, respectively.

Table 4-1. Primers used in this study

Primers	Sequences
ml1419c pJN105 forward	5′-GGAATTC**GAGGAGGATATT**CGTGTTGGAGACGGTGCGTAG-3′
ml1419c pJN105 reverse	5′-GGACTAGTTCAGCTAGGTTGTTGGTTGAACGTG-3′
ml1750c pJN105 forward	5'-GGACTAGT**GAGGAGGATATT**CATGACCCGGTCCCTGGAC-3'
ml1750c pJN105 reverse	5'-GCGAGCTCTTACGATCGGTGTATCTGGTGTTC-3'
*ml0397*c pJN105 forward	5'-GGAATTC**GAGGAGGATATT**CGTGGATCACACGACGAAGTG-3'
FL-*ml0397c* pJN105 reverse	5'-GGACTAGT TCATTTGATAATGCCGATCTTGCG-3'
PG-*ml0397c* pJN105 reverse	5'-GGACTAGTTCACGTGACGGCACATTGTTTCTTG-3'
ml1750c pET28a forward	5'-AAGCATATGACCCGGTCCCTGGACGAA-3'
ml1750c pET28a reverse	5'-GTTAAGCTTTTACGATCGGTGTATCTGGT-3'
ml1419c pET28a forward	5′-AAACATATGTTGGAGACGGTGCGTAGCG-3′
ml1419c pET28a reverse	5′-TTAAGCTTGCTAGGTTGTTGGTTGAACG-3′
ml0397c pET28a forward	5’-AAGCATATGGTGGATCACACGACGAAGTG-3’
FL-*ml0397c* pET28a reverse	5’-GGGAAGCTTTCATTTGATAATGCCGATC-3’
PG-*ml0397c* pET28a reverse	5’-GAGAAGCTTTCAGACGGCACATTGTTTCTTG-3’
PG-*ml0397c* pET28a forward-2	5'-AAGCATATGGCCGACATTACATCAGAGG-3'
PG-*ml0397c* pET28a reverse-2	5'-GAGAAGCTTTCAGACGGCACATTGTTTCTTG-3'
ml1419c$_{\Delta GGDEF}$ forward	5'-GTGGTGGGTAGGTTCGTCGCTCTGATCCTG-3'
ml1419c$_{\Delta GGDEF}$ reverse	5'-CAGGATCAGAGCGACGAACCTACCCACCAC-3'
PG-*ml0397c*$_{\Delta GGDEF}$ forward	5'-TCGTCGCGCGGCTGATCGTACTACTCCG-3'
PG-*ml0397c*$_{\Delta GGDEF}$ reverse	5'-CGGAGTAGTACGATCAGCCGCGCGACGA-3'

Table 4-2. Plasmids used in this study

Plasmids	Description	Reference or source
pJN105	*P. aeruginosa* arabinose inducible expression vector, Gm^R, *araC-*P_{BAD}	(23)
pJN1120	pJN105 containing *tpbB* (PA1120)	(8, 25)
pMRLB105	pJN105 containing *ml1419c*	This study
pMRLB108	pJN105 containing *ml1419c*$_{\Delta GGDEF}$	This study
pMRLB109	pET28a (Kan^R, IPTG inducible, *T7*) containing *ml1419c*	This study
pMRLB110	pJN105 containing full length *ml0397c*	This study
pMRLB111	pJN105 containing PAS-GGDEF *ml0397c*	This study
pMRLB112	pJN105 containing *ml1750c*	This study
pMRLB113	pET28a containing *ml1750c*	This study
pMRLB114	pET28a containing full length *ml0397c*	This study
pMRLB115	pET28a containing PAS-GGDEF *ml0397c* (900 bp)	This study
pMRLB116	pET28a containing PAS-GGDEF *ml0397c* (927 bp)	This study
pMRLB117	pET28a containing PAS-GGDEF *ml0397c*$_{\Delta GGDEF}$	This study

4.2.5 Protein and whole cell lysate isolation from *E. coli*

Recombinant ML1419c was purified from *E. coli* BL21(DE3). *E. coli* transformed with pMRLB109 was cultured to an OD_{600} ~0.4-0.6 and induced with 0.5 mM IPTG (EMD Millipore) for 3 h. Cells were lysed in lysis buffer [PBS (pH 7.4) with 1X proteinase inhibitor (Roche), 50 µg/ml DNase I (Sigma Aldrich) and 10 µg/ml RNase A (Sigma Aldrich)] using an ultrasonic processor (Vibra-cell VC750) with an amplitude setting of 30% with six 20 sec pulses and a 59 sec pause between the pulse cycles. Protein inclusion bodies were collected by centrifugation and suspended in binding buffer [50 mM Tris-HCl (pH 8.0), 300 mM NaCl, 10 mM imidazole and 8 M urea]. Protein purification was achieved by immobilized metal affinity chromatography (IMAC) with Ni-NTA agarose resin (Qiagen, Valencia, CA). The purified protein was eluted with binding buffer containing 150 mM imidazole, and dialyzed against 50 mM Tris-HCl (pH 8.0), 150 mM NaCl with a gradual reduction of urea. Protein concentrations were determined using the bicinchoninic acid (BCA) assay (26). The purified recombinant ML1419c was provided to Lampire Biological Laboratories (Pipersville, PA) for the production of rabbit anti-ML1419c polyclonal serum.

E. coli BL21(DE3) transformed with pMRLB113 (*ml1750c*) was cultured to an OD_{600} ~0.4-0.6 and induced with 0.5 mM IPTG (EMD Millipore) for 3 h. Cells were lysed using an ultrasonic processor (Vibra-cell VC750) with an amplitude setting of 30% with six 20 sec pulses and 59 sec pauses between the pulse cycles in lysis buffer [PBS (pH 7.4) with 1X proteinase inhibitor, 10 µg/ml DNase I and 10 µg/ml RNase A, 300 µg/ml lysozyme]. The*E. coli* expressing pMRLB116 (PAS-GGDEF *ml0397c*) was cultured to OD600 ~0.4-0.6 and induced with 1 mM IPTG for 2 h and cells were lysed in lysis buffer [PBS (pH 7.4) with 1X proteinase inhibitor, 50 µg/ml DNase I and 10 µg/ml RNase A, 300 µg/ml lysozyme]. Protein inclusion bodies were

collected by centrifugation and suspended in binding buffer [100 mM NaH_2PO_4, 10 mM Tris-HCl (pH 8.0), 10 mM imidazole and 8 M urea]. Protein purification was achieved by IMAC with Ni-NTA agarose resin. The purified protein was eluted with binding buffer containing 150 mM, 250 mM, and 500 mM imidazole, and dialyzed against 20 mM Tris-HCl (pH 8.0), 500 mM NaCl with a gradual reduction of urea. Protein concentrations were determined using the BCA assay (26). The purified recombinant ML1750c and PAS-GGDEF ML0397c were provided to Lampire Biological Laboratories for the production of rabbit anti-ML1750c and anti-ML0397c polyclonal serum.

For phenotypic studies and c-di-GMP extraction, *E. coli* BL21(DE3) pLysS was transformed with pMRLB109, pMRLB113, pMRLB114, pMRLB115, pMRLB117 and pET28a vector by chemical transformation following the manufacturer's recommendations. For whole cell lysate, overnight cultures of *E. coli* strains were diluted (1:100) into fresh Miller LB broth containing kanamycin (50 µg/ml), chloramphenicol (34 µg/ml) and 0.5 mM IPTG and grown to log-phase (OD_{600} ~0.6-0.7). Cells were collected by centrifugation and suspended in lysis buffer described for recombinant *ml1419c* expressing and lysed by using an ultrasonic processor (Microson XL2000) with three 30 sec pulses and two min pause between the pulse cycles. Total protein concentrations were determined using the BCA assay.

4.2.6 Polyclonal antibody test

Purified ML1750c, ML1419c, ML0397c (2.5 µg) was resolved on a Nupage 4-12% Bis-Tris Zoom protein gel, IPG well (Invitrogen). Proteins were transferred to PVDF membranes (EMD Millipore) that were cut into 10-12 slices. Each slice of protein membrane was incubated in serial dilutions of primary antibodies which are anti-ML1750c, anti-ML1419c and anti-

ML0397c polyclonal antibodies from rabbit serum at day 57 after antigen injection (Term-bleed). The polyclonal serum was diluted from 1:1,000 to 1:200,000. The second antibody is anti-rabbit IgG conjugate to alkaline phosphatase (1:10,000) (Sigma Aldrich).

4.2.7 Whole cell lysate isolation from *P. aeruginosa* PAO1

P. aeruginosa PAO1 was transformed with pMRLB105, pMRLB108, pMRLB110, pMRLB111, pMRLB112, pJN1120 or pJN105 by electroporation (27). Overnight cultures of *P. aeruginosa* strains were diluted (1:100) into fresh Lennox LB broth containing 100 µg/ml gentamicin and 0.2% L-arabinose and grown to log-phase (OD_{600} ~0.6-0.7). Cells were collected by centrifugation and suspended in lysis buffer and lysed as described for ML1419c purification from *E. coli*. Total protein concentrations were determined using the BCA assay.

4.2.8 SDS-PAGE and Western blot analyses

ML1419c, ML1750c, and ML0397c production in *P. aeruginosa* PAO1 were initially screened by Western blot analysis. Purified recombinant ML1419c (125 ng), ML1750c (250 ng) and ML0397c (250 ng) and *P. aeruginosa* whole cell lysates (30 µg) were resolved under denaturing conditions on NuPAGE 4-12% bis-tris polyacrylamide gels (Invitrogen). Proteins were visualized by staining with Coomassie G-250 stain (Invitrogen). For western blots, proteins were electro-transferred to nitrocellulose membrane and probed with mouse anti-ML1419c polyclonal serum (1:5,000) (provided by Dr. Spencer), rabbit anti-ML1750c polyclonal serum (1:100,000), and rabbit anti-ML0397c polyclonal serum (1:100,000). The secondary antibody was anti-mouse or anti-rabbit IgG conjugated to alkaline phosphatase (1:20,000) (Sigma Aldrich). Reactive

proteins were visualized by incubation with BCIP/NBT (5-Bromo-4-chloro-3-indolyl phosphate/Nitro blue tetrazolium) tablets (Sigma Aldrich).

For the *ml1419c* site-directed mutagenesis study, purified recombinant ML1419c (500 ng or 25 ng) and *P. aeruginosa* whole cell lysates (5 µg) were resolved under denaturing conditions on NuPAGE 4-12% bis-tris polyacrylamide gels. Proteins were visualized by staining with Coomassie G-250 stain. For Western blots, proteins were electrotransferred to PVDF membranes and probed with rabbit anti-ML1419c polyclonal serum (1:200,000). The secondary antibody was anti-rabbit IgG conjugated to HRP (1:20,000) (Promega, Madison, WI). Chemiluminescence was used to visualize reactive proteins by incubating membrane with luminol-based enhanced chemiluminescence HRP substrate (SuperSignal West Dura Extended Duration Substrate) (Thermo Scientific).

E. coli whole cell lysates (2.5 µg and 5 µg) were resolved on NuPAGE 4-12% bis-tris polyacrylamide gels gel for Western blots and visualized with Coomassie G-250 stain as described for *P. aeruginosa* whole cell lysates. For western blots, proteins were electrotransferred to nitrocellulose membrane (Bio-Rad, Hercules, CA) and probed with mouse anti-polyhistidine monoclonal antibody (1:3,000) (Sigma Aldrich). The secondary antibody was anti-mouse IgG conjugated to alkaline phosphatase (1:20,000) (Sigma Aldrich). Reactive proteins were visualized by incubation with BCIP/NBT tablets.

4.2.9 *P. aeruginosa* phenotypic assays

Colony morphology and dye binding - *P. aeruginosa* strains were grown on VBMM agar containing 40 µg/ml Congo red (Sigma Aldrich), 15 µg/ml Coomassie Brilliant Blue (Sigma Aldrich), 1% L-arabinose and 100 µg/ml gentamicin at 30°C for 48 h modified from previous

study (17). The colony morphology was observed under a Leica MZ9.5 stereomicroscope (Leica Microsystems, Buffalo Grove, IL).

Motility assays - Swimming and twitching motility were assessed as previously described (28, 29) with minor modifications. Swimming motility was assayed by stab inoculating 1 µl of a *P. aeruginosa* overnight culture into low-viscosity Lennox LB agar (0.3% Bacto agar) containing with 100 µg/ml gentamicin and 0.2% L-arabinose. The diameters of the swimming zone were measured after growth at 37°C for 24 h. Twitching motility was assessed by stab inoculating a colony of *P. aeruginosa* through Lennox LB agar containing 100 µg/ml gentamicin and 0.2% L-arabinose and cultured at 37°C for 48 h. The migration of bacteria attached on the polystyrene plate surface was visualized by staining with 0.1% crystal violet and the diameters were measured.

Biofilm formation - Biofilm formation was assessed as previously described (30) with modifications. The log phase *P. aeruginosa* cultures (OD_{600}~0.6-0.7) were adjusted to an OD_{600} of ~0.1 in VBMM containing 100 µg/ml gentamicin and 0.2% L-arabinose. Aliquots (150 µl) of diluted cultures were added to 96-well polystyrene plates (Nunc™ Microwell™ 96-well microplates (#243656), Thermo Scientific) in replicates of six and incubated at 37°C in a sealed bag for 24 h. The 96-well plates were washed twice with water, stained with 0.1% crystal violet, for 10 min, and washed twice with water. Bound crystal violet was solubilized with 30% acetic acid and the absorbance was measured at 590 nm.

4.2.10 *E. coli* phenotypic assays

Curli formation - The assay was performed as described previously with modification (21). The curli formation assay was assessed by inoculating 1 µl of overnight *E. coli* cultures onto YESCA agar (0.1% yeast extract, 1% casamino acids, 2% Difco agar) containing 50 µg/ml Congo red dye, 50 µg/ml kanamycin, 34 µg/ml chloramphenicol and various concentrations of IPTG (0 mM, 0.05 mM, 0.1 mM, 0.5 mM, or 1.0 mM). The curli formation assay plates were incubated at 30ºC for 3 days and then at room temperature for 24 h. The red color of colony indicates of the production of curli.

Cellulose formation - The assay was performed as described previously with modification (21). Cellulose production was detected by inoculating 1 µl of overnight *E. coli* cultures onto low salt LB agar (0.5 g NaCl/liter) containing 0.02% Calcofluor white stain (Fluorescent Brightener 28) (Sigma Aldrich), 50 µg/ml kanamycin, 34 µg/ml chloramphenicol and various concentrations of IPTG (0 mM, 0.05 mM, 0.1 mM, 0.5 mM or 1 mM). The cellulose production was determined under UV light in Gel Doc XR imager (Bio-Rad) after incubation at 30ºC for 3 days. White fluorescence of the colonies indicated cellulose production

4.2.11 Quantitative analysis of c-di-GMP by liquid chromatography-mass spectrometry (LC-MS)

Overnight cultures of *P. aeruginosa* were diluted with VBMM (1:100) containing 100 µg/ml gentamicin and 0.2% L-arabinose. Overnight cultures of *E. coli* BL21(DE3) pLysS strains were diluted with LB (1:100) containing 50 µg/ml kanamycin, 34 µg/ml chloramphenicol and 0.5 mM IPTG. An aliquot (2 ml) of each *P. aeruginosa* and *E. coli* culture grown to an OD_{600}~0.6-

0.7 was spiked with 100 nM or 200 nM [^{13}C] adenosine (Omicron Biochemicals) and extracted with 100 µl of 0.6 M (final concentration) perchloric acid (Sigma Aldrich) (17). In addition, the *E. coli* cell pellets were washed twice with water before the extraction to discard the LB medium. The precipitate of the perchloric acid extraction was used for determination of protein concentration by the BCA assay and sample normalization (31), and the extracts were neutralized with 20 µl of 2.5 M potassium bicarbonate. The neutralized supernatants were stored at -80°C. Extracts (10 µl) of c-di-GMP were analyzed by LC-MS as described by the method in Chapter 3 (3.2.2).

4.2.12 Statistical analyses

P values were calculated by one-way analysis of variance (ANOVA) followed by Tukey comparison using GraphPad Prism version 6.0 (GraphPad Software). Data were expressed as mean values ± SD. The p value < 0.05 was considered as statistically significant.

4.2.13 Bioinformatics analyses

The annotated proteins containing either GGDEF, EAL, or HD-GYP domains of *M. leprae* strain TN were identified from the NCBI Conserved Domain Database (CDD) (http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml) (32) using the domain accession numbers: GGDEF, cd01949; EAL, cd01948; and HD-GYP, cd00077 and SMART database (http://smart.embl-heidelberg.de/) (33) using the domain accession numbers: GGDEF, SM00267; EAL, SM00052; and HD-GYP, SM00471. *M. leprae* proteins containing GGDEF, EAL or HD-GYP motifs that were identified in both databases were used for further analyses. ML1419c and ML0397c heme-binding sites in sensory domains were characterized by the NCBI CDD.

Alignment of *M. leprae* strain TN protein sequences possessing GGDEF and/or EAL motifs was performed with the T-Coffee alignment tool (34). Evaluation of putative transmembrane domains was performed with TMHMM program (http://www.cbs.dtu.dk/services/TMHMM/) (35).

4.3 Results

4.3.1 Bioinformatics analyses and identification of c-di-GMP-related proteins in *M. leprae*

Identification of putative DGC proteins (GGDEF) or c-di-GMP PDE proteins (EAL or HD-GYP) encoded in the genome of *M. leprae* strain TN was achieved by query of the NCBI CDD and the SMART databases. Bioinformatics analyses revealed that *M. leprae* harbors one putative hybrid DGC-PDE protein (ML1750c), two putative DGC proteins (ML1419c and ML0397c) and one putative PDE protein with an EAL domain (ML1752c) (Figure 4-1). No proteins with a HD-GYP motif were identified. Results from these bioinformatics analyses of *M. leprae* strain TN agree with the large scale census of c-di-GMP-related proteins (1).

The hybrid DGC-PDE protein ML1750c possesses an N-terminal GAF sensor domain as well as GGDEF and EAL domains (Figure 4-1 (a)). This protein is homologous to *M. smegmatis* MSMEG_2196 with 64.39% identity and *M. tuberculosis* Rv1354c with 62.30% identity, both of which were experimentally defined as possessing active GGDEF and EAL domains (36-38). Unexpectedly, *M. leprae* with its reduced genome encoded two additional DGC proteins, ML1419c and ML0397c (Figure 4-1 (b), (c)), that are not encoded by *M. smegmatis* or *M. tuberculosis*. ML1419c possesses three sequential PAS signaling domains N-terminal to the GGDEF domain. Two ML1419c heme-binding sites in PAS domains were identified based on the data from NCBI CDD. ML0397c harbors a single N-terminal PAS sensor domain with heme-binding site linked to a GGDEF domain and ten C-terminal transmembrane domains as predicted

by analysis with the TMHMM program. A single predicted PDE protein (ML1752c) is homologous to Rv1357c of *M. tuberculosis* (1, 36), but it has no homology to proteins encoded by *M. smegmatis* (Figure 4-1 (d)).

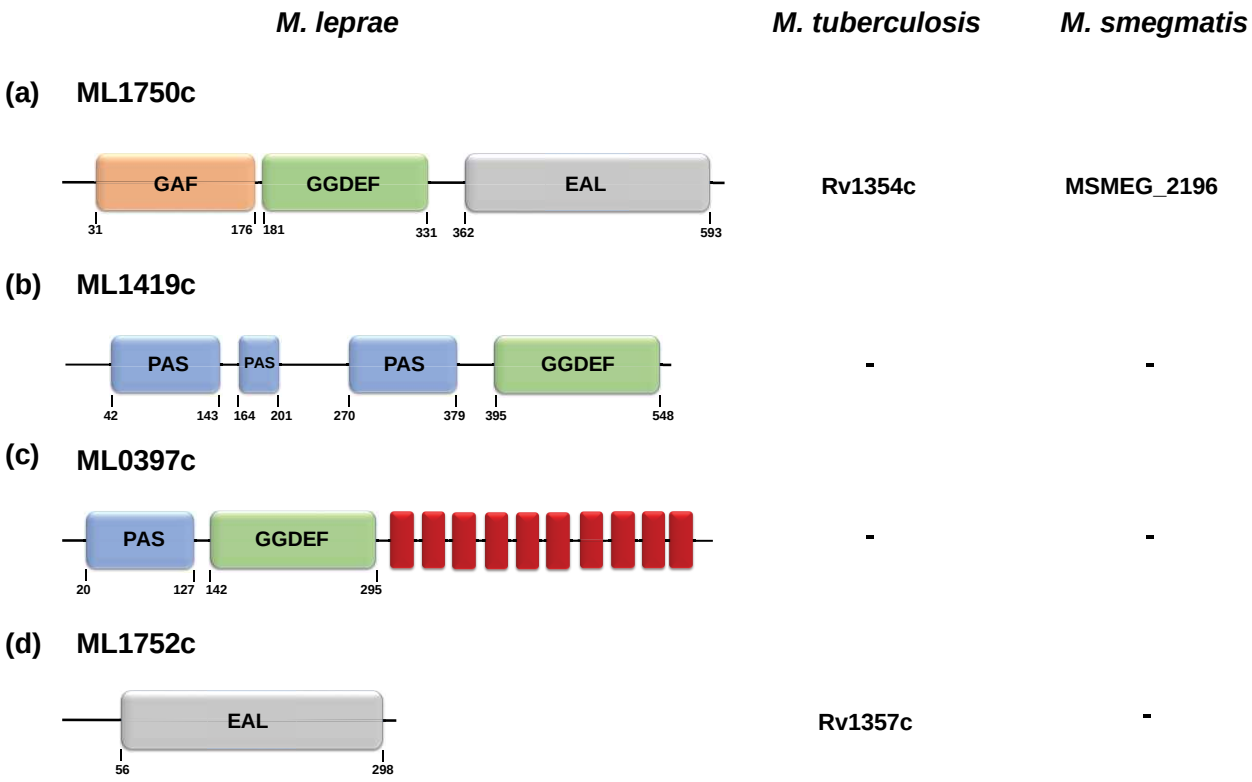


Figure 4-1. Bioinformatics analyses of putative DGC and PDE proteins of *M. leprae*
(a) ML1750c (623 aa) (gi|15827936|NP_302199) is a hybrid protein containing both GGDEF and EAL motifs, and a N-terminal GAF sensory domain. (b) ML1419c (563 aa) (gi|15827746|NP_302009) contains a GGDEF motif and three consecutive PAS sensory domains upstream to GGDEF domain. (c) ML0397c (602 aa) (gi|15827122|NP_301385) possesses a GGDEF motif, a N-terminal PAS sensor domain, and 10 transmembrane α-helices (red rectangles). (d) ML1752c (302 aa) (gi|15827938|NP_302201) has a single EAL motif and lacks a sensory domain. Homologs of ML1750c are produced in both *M. tuberculosis* and *M. smegmatis* and a homolog of ML1752c is identified in *M. tuberculosis*. Numbers indicate amino acid positions as reported by CDD NCBI.

Multiple alignments of the *M. leprae* proteins containing GGDEF and EAL domains (Figure 4-2 and Figure 4-3) were performed with T-Coffee (34). This demonstrated conservation of key amino acids in putative active sites. The three predicted DGCs of *M. leprae* (ML1750c, ML1419c and ML0397c) all contain the A-site sequence, RxGGDEF (39). Thus, these proteins are expected to act as functional DGCs producing c-di-GMP. In addition, ML1419c and ML0397c possess an I-site, RxxD motif (40) that is located directly up-stream of the A-site. Multiple alignment analyses for the proteins with predicted PDE activity and EAL domains revealed the conserved residues of the EAL active site are present in ML1750c and ML1752c. This conservation included appropriately spaced residues of E, N, E, E, D, K, and E, except for the last E residue of ML1752c which is replaced with a K residue (1) (Table 4-3). These *in silico* data indicate that *M. leprae* has a greater capacity than *M. tuberculosis* or *M. smegmatis* for c-di-GMP production.

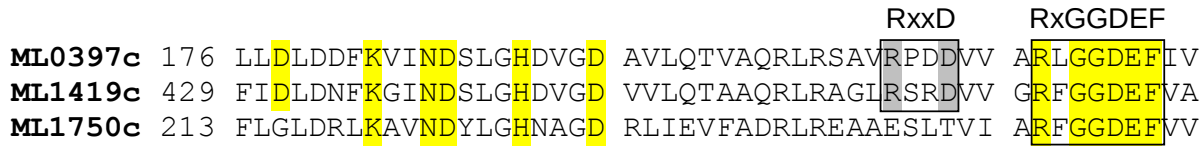


Figure 4-2. Alignment of conserved DGC domains of *M. leprae* proteins ML0397c, ML1419c and ML1750c
The conserved I-site, RxxD motif of ML0397c and ML1419c is highlighted in gray. The conserved A-site, RxGGDEF motif, is present in all proteins. Conserved amino acids involved in enzymatic activity are highlighted in yellow.

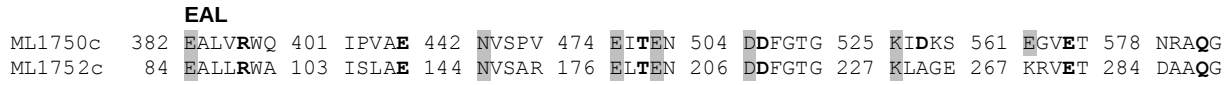


Figure 4-3. Comparison of the active residues and positions in conserved EAL domain containing proteins from *M. leprae* TN, *M. smegmatis* mc^{2}155, and *M. tuberculosis* H37Rv
Alignment of protein sequence containing EAL domains of *M. leprae* TN (ML1750c and ML1752c). The amino acids that contribute to PDE activity are highlighted with gray. The other known conserved amino acids surrounding the active site are in bold text (1).

Table 4-3. Conserved residues in proteins containing EAL domains

Proteins	**Active site residues**						
ML1750c	E (382)	N (442)	E (474)	E (477)	D (504)	K (525)	E (561)
ML1752c	E (84)	N (144)	E (176)	E (179)	D (206)	K (227)	K (267)
MSMEG_2196	E (384)	N (444)	E (476)	E (479)	D (506)	K (527)	E (563)
Rv1354c	E (389)	N (449)	E (481)	E (484)	D (511)	K (532)	E (568)
Rv1357c	E (89)	N (149)	E (181)	E (184)	D (211)	K (232)	K (272)

4.3.2 Conditional expression of *ml1750c*, *ml1419c*, and *ml0397c* in *E. coli* BL21(DE3) pLysS

ml1750c, *ml1419c*, full length *ml0397c*, and PAS-GGDEF *ml0397c* were conditionally expressed in BL21(DE3) pLysS *E. coli* under the control of the IPTG induction. Recombinant protein production was assessed by SDS-PAGE and Western blot of whole cell lysates of *E. coli* strains (Figure 4-4). Proteins of the expected size; ML1750c (67.4 kDa), ML1419c (60.9 kDa), and PAS-GGDEF ML0397c (34 kDa) that reacted with anti-histidine monoclonal antibody were observed in the whole cell lysates of *E. coli* BL21(DE3) pLysS expressing *ml1750c*, *ml1419c*, and PAS-GGDEF *ml0397c*, respectively when grown under the presence of IPTG. The expression of full length of *ml0397c* was not accomplished in various induction conditions, which are different IPTG concentrations or at low temperature. Only proteins with lower molecular weight than expected size were observed (Figure 4-4). With IPTG induction, the *E. coli* expressing full length of *ml0397c* had reduced growth compared to PAS-GGDEF *ml0397c* (data not shown).

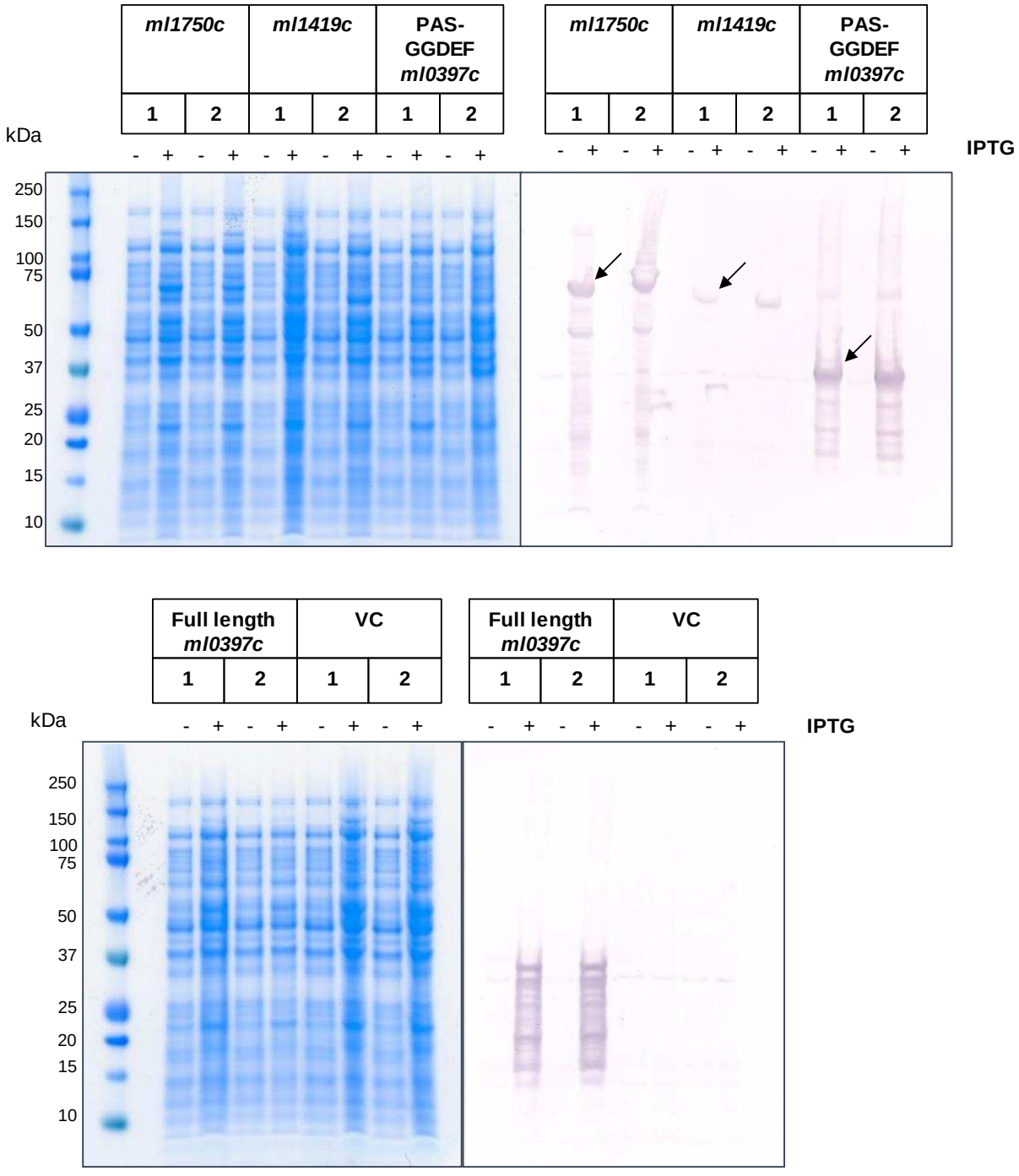


Figure 4-4. Expression of *ml1419c*, *ml1750c*, and PAS-GGDEF *ml0397c* from *M. leprae* in *E. coli*

Two whole cell lysates from *E. coli* BL21(DE3) pLysS expressing *ml1750c*, *ml1419c*, PAS-GGDEF *ml0397c*, full length *ml0397c* and pET28a vector control (VC) grown under the presence (+) or absence (-) of IPTG were analyzed by SDS-PAGE with Coomassie blue staining (left panel) and Western blot (right panel) using anti-histidine monoclonal antibody. Black arrows indicate proteins with expected size; ML1750c (67.4 kDa), ML1419c (60.9 kDa), PAS-GGDEF ML0397c (34 kDa).

In addition, PAS-GGDEF *ml0397c* and a mutated construct of this gene (PAS-GGDEF *ml0397c*$_{\Delta GGDEF}$) was conditionally expressed in *E. coli* BL21(DE3) pLysS under the control of the IPTG induction (Figure 4-5). A protein of expected size, 34 kDa of PAS-GGDEF ML0397c, that reacted with anti-histidine monoclonal antibody was observed in the whole cell lysates of *E. coli* expressing PAS-GGDEF *ml0397c*, or PAS-GGDEF *ml0397c*$_{\Delta GGDEF}$ when grown in the presence of IPTG (Figure 4-5). In absence of IPTG, no recombinant protein production was observed for all strains. Additionally, no products reactive to the anti-histidine monoclonal antibody were observed for *E. coli* containing the pET28a vector control regardless of the presence or absence of IPTG (Figure 4-4 and Figure 4-5).

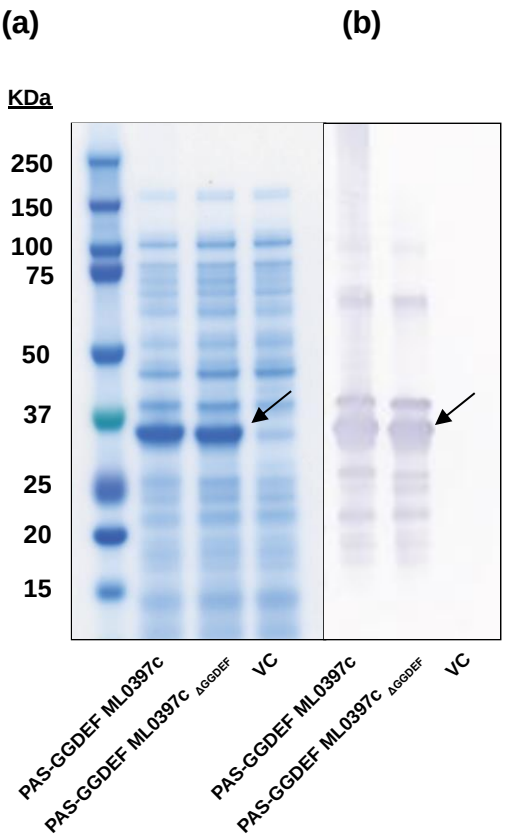


Figure 4-5. Expression of PAS-GGDEF *ml0397c* from *M. leprae* in *E. coli* BL21(DE3) pLysS Recombinant *E. coli* strains containing PAS-GGDEF *ml0397c*, PAS-GGDEF *ml0397c*$_{\Delta GGDEF}$, or the pET28a vector control (VC) were grown under the presence of 0.5 mM IPTG. Whole cell lysates, 5 µg and 2.5 µg, of the recombinant strains were analyzed by SDS-PAGE with Coomassie blue staining (a) and Western blot using anti-histidine monoclonal antibody (b), respectively. Black arrow indicates the protein with size 34 kDa.

4.3.3 Polyclonal antibody test

The sensitivity of anti-ML1750c, anti-ML1419c and anti-ML0397c polyclonal sera generated in rabbit by using purified proteins from *E. coli* as antigens was determined before detection of ML1750c, ML1419c and ML0397c in *M. leprae* subcellular fractions (Chapter 5) or *P. aeruginosa* whole cell lystae. The recombinant ML1750c, ML1419c, and ML0397c proteins purified from *E. coli* were used as substrates for polyclonal antibody detection. The polyclonal antibodies diluted to titer of 1:100,000 or 1:200,000 could efficiently detect ML1750c, Ml1419c,

and ML0397c (Figure 4-6). The ML1750c, ML1419c, and ML0397c were not detected by D0 rabbit serum (data not shown).

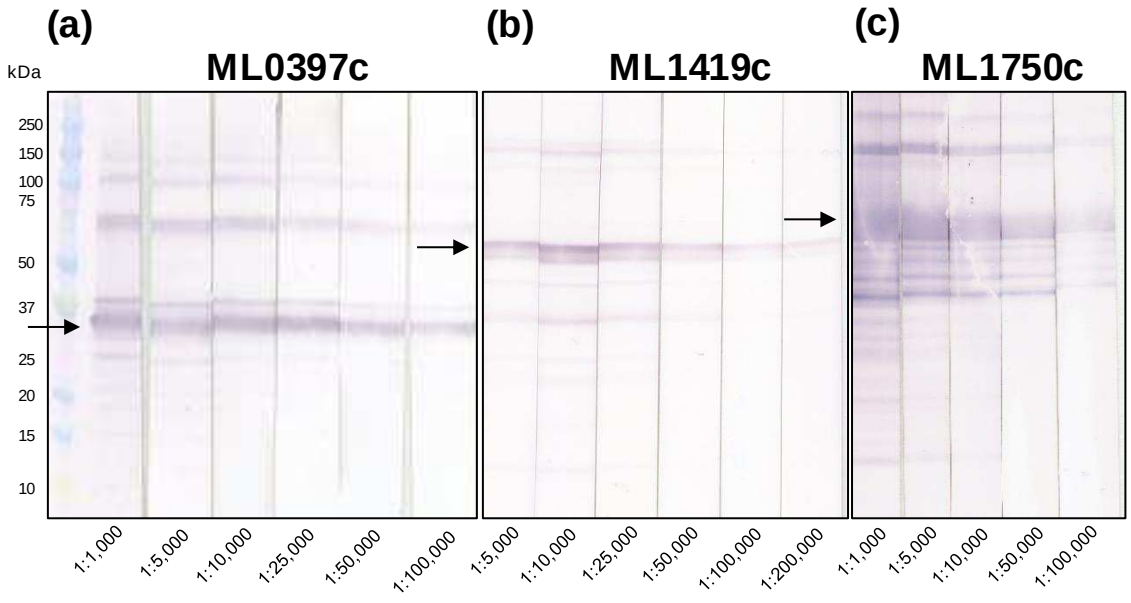


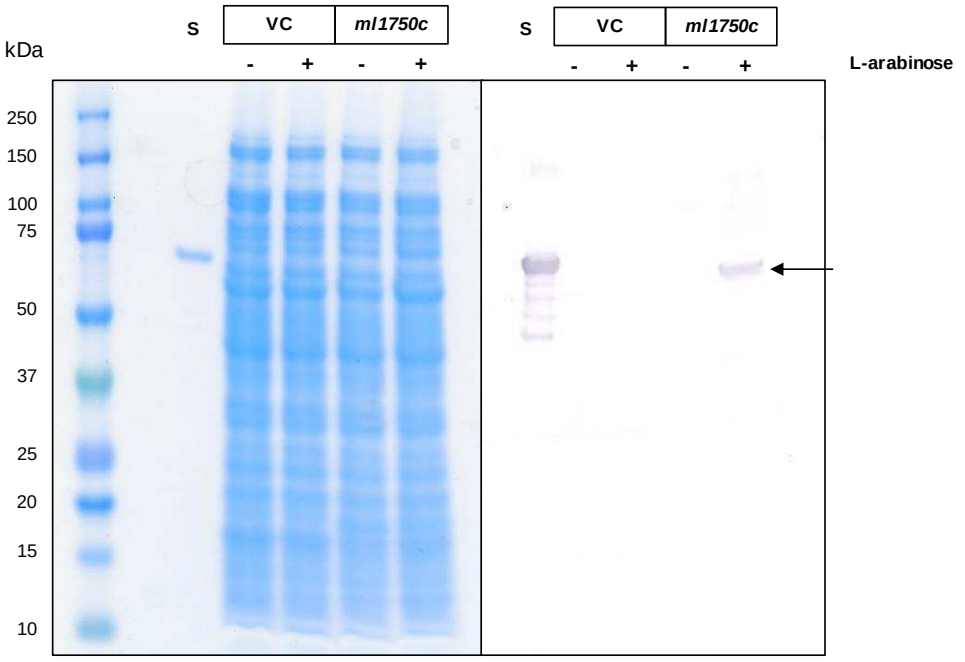
Figure 4-6. Anti-ML0397c, anti-ML1419c, and anti-ML1750c polyclonal serum test
The anti-ML0397c (a), anti-ML1419c (b), and anti-ML1750c (c) polyclonal sera was applied to monitor recombinant ML0397c proteins without transmembrane domain (34 kDa), ML1419c (60.9 kDa), ML1750c (67.4 kDa), respectively (indicated by black arrows).

4.3.4 Conditional expression of *ml1750c*, *ml1419c*, and *ml0397c* in *P. aeruginosa* PAO1

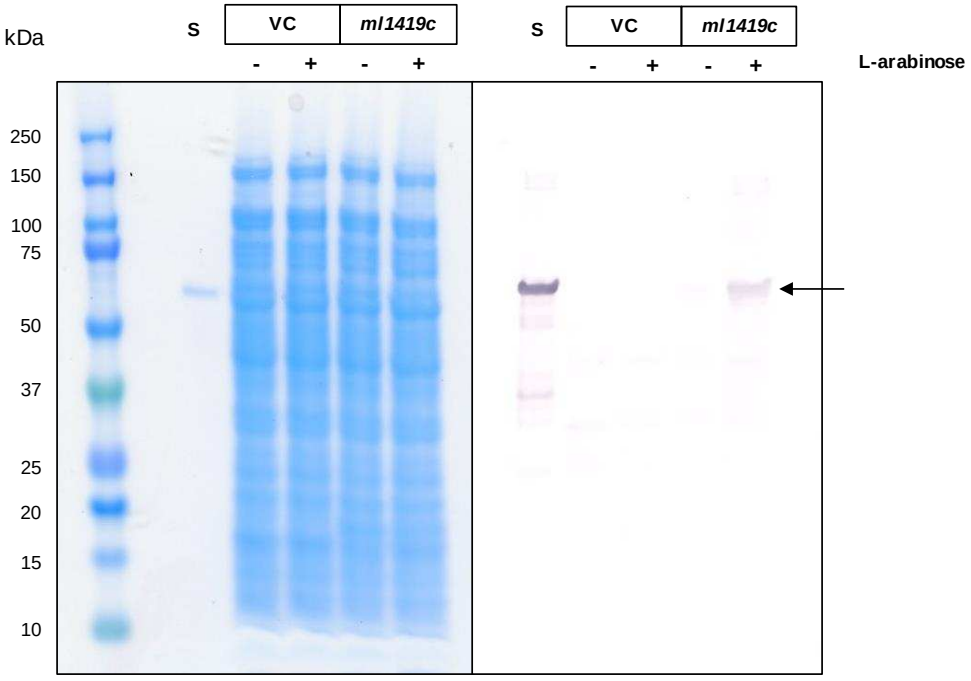
M. leprae cannot be cultured *in vitro*, thus protein functions for this bacterium are typically studied in model organisms (41). The production and function of c-di-GMP has been extensively studied in *P. aeruginosa*, where phenotypes and mutants associated with this second messenger molecule are well described. In this study, *ml1750c*, *ml1419c*, full length *ml0397c* and PAS-GGDEF *ml0397c* were conditionally expressed in *P. aeruginosa* PAO1 under the control of the L-arabinose responsive P_{BAD} promoter (23). Recombinant protein production was assessed by SDS-PAGE and Western blot of whole cell lysates of *P. aeruginosa* strains (Figure 4-7). Proteins of the expected size; ML1750c (67.4 kDa), ML1419c (60.9 kDa), and PAS-GGDEF ML0397c (34 kDa)

that reacted with specific polyclonal serum were observed in the whole cell lysates of *P. aeruginosa* expressing *ml1750c*, *ml1419c*, and PAS-GGDEF *ml0397c*, respectively when grown in the presence of arabinose. Anti-ML0397c (PAS-GGDEF) polyclonal antibody could not detect protein with 63 kDa corresponding to full length ML0397c. Only proteins with a lower molecular weight than expected size were observed (Figure 4-8). It should also be noted that the *P. aeruginosa* expressing full length *ml0397c* had reduced growth compared to PAS-GGDEF *ml0397c* (data not shown). These results indicate that *M. leprae* ML1750c, ML1419c, and PAS-GGDEF ML0397c can be produced in the *P. aeruginosa* heterologous host.

(a)


kDa
S
VC
ml1750c
S
VC
ml1750c
- + - +
- + - +
L-arabinose
250
150
100
75
50
37
25
20
15
10

(b)


kDa
S
VC
ml1419c
S
VC
ml1419c
- + - +
- + - +
L-arabinose
250
150
100
75
50
37
25
20
15
10

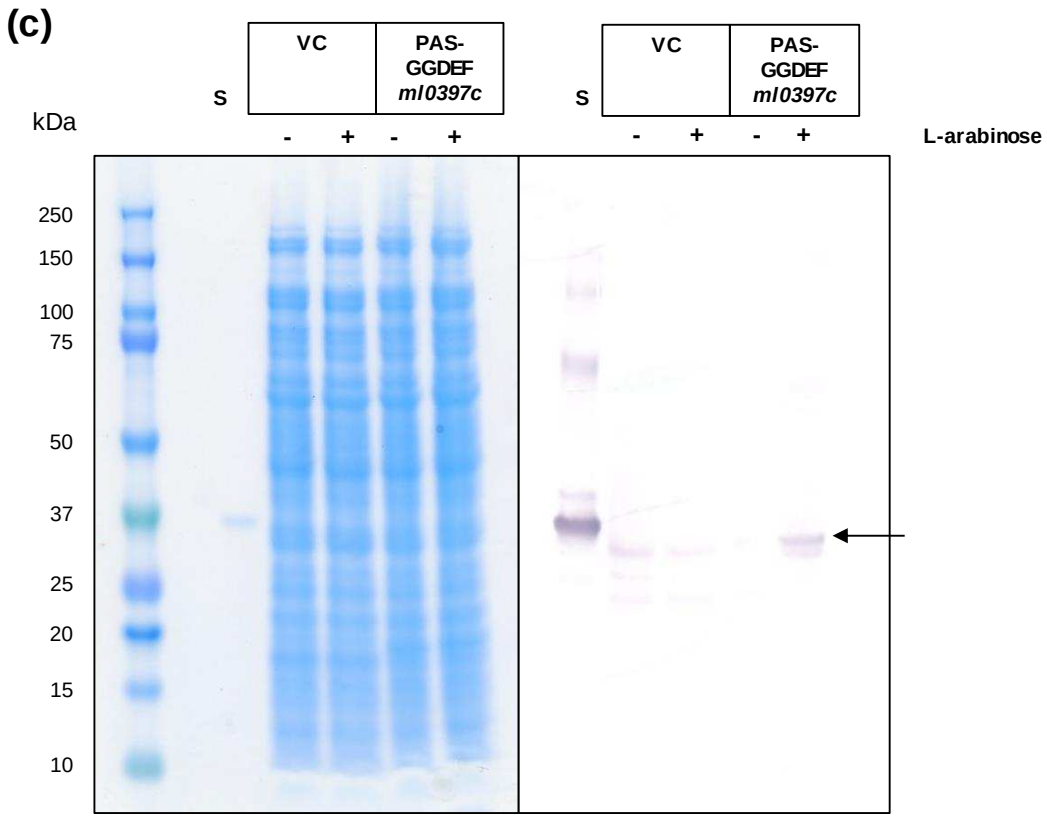


Figure 4-7. Expression of *ml1750c*, *ml1419c*, and PAS-GGDEF *ml0397c* from *M. leprae* in *P. aeruginosa* PAO1

Whole cell lysates (30 µg) of the *P. aeruginosa* PAO1 expressing *ml1750c* (a), *ml1419c* (b), PAS-GGDEF *ml0397c* (c) and pJN105 vector control (VC) grown under the presence (+) and the absence (-) of arabinose were analyzed by SDS-PAGE with Coomassie blue staining (left panel) and Western blot (right panel). Purified recombinant proteins produced in *E. coli* were used as a positive control (lane S).

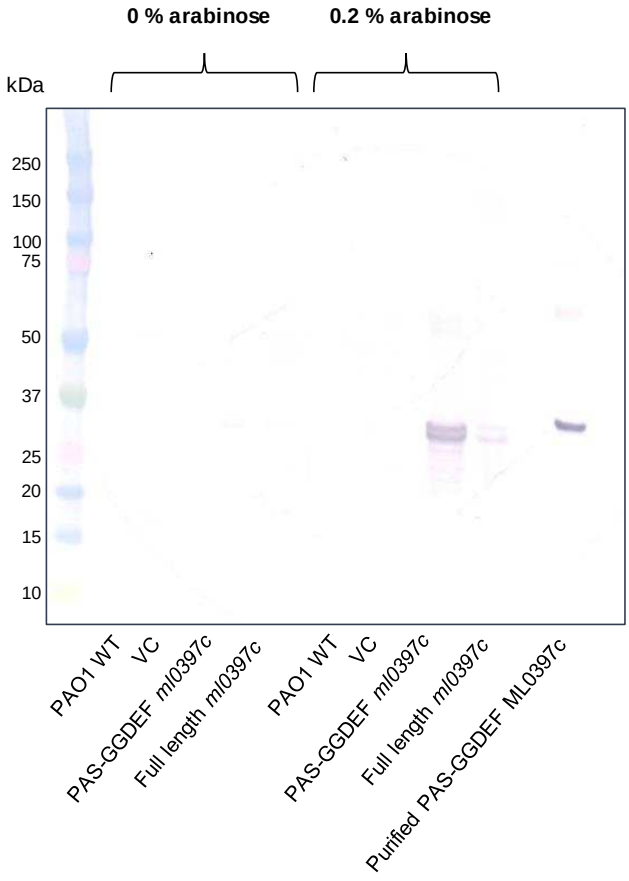


Figure 4-8. Expression of full length *ml0397c* from *M. leprae* in *P. aeruginosa* PAO1
Western blot analysis of whole cell lysate from *P. aeruginosa* PAO1 wild type and *P. aeruginosa* expressing PAS-GGDEF *ml0397c*, full length *ml0397c*, and vector control (VC) grown under the presence and the absence of arabinose. The full length ML0397c (63 kDa) was not detected by anti-ml0397c (PAS-GGDEF) polyclonal serum.

Furthermore, the *ml1419c* and a mutated construct of this gene ($ml1419c_{\Delta GGDEF}$) was also conditionally expressed in *P. aeruginosa* PAO1 under the control of the L-arabinose responsive P_{BAD} promoter (23). Protein of the expected size (approximately 60.9 kDa) that reacted with anti-ML1419c polyclonal serum was observed in the whole cell lysates of *P. aeruginosa* expressing *ml1419c*, or $ml1419c_{\Delta GGDEF}$ when grown in the presence of arabinose (Figure 4-9 (b)). In the absence of arabinose, no or low levels of recombinant protein production was observed. No

products reactive to the anti-ML1419c polyclonal serum were observed for *P. aeruginosa* containing the pJN105 vector control regardless of the presence or absence of arabinose.

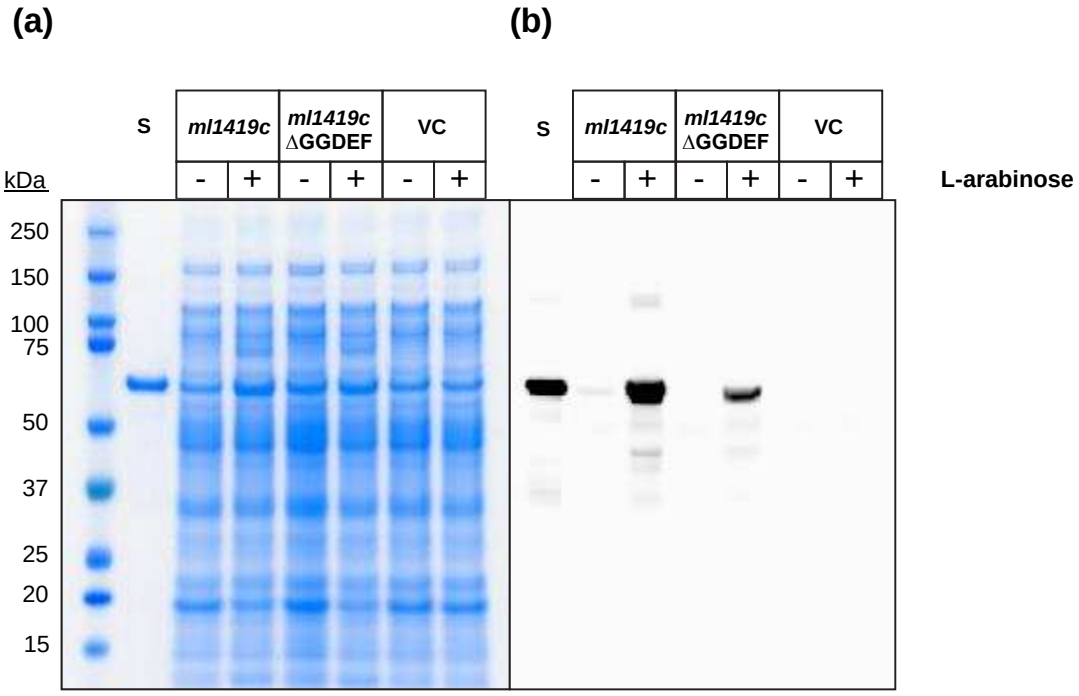


Figure 4-9. Expression of *ml1419c* from *M. leprae* in *P. aeruginosa* PAO1
Recombinant *P. aeruginosa* PAO1 containing *ml1419c*, *ml1419c*$_{\Delta GGDEF}$, or the pJN105 vector (VC) were grown in the presence (+) or absence (-) of 0.2% L-arabinose. Whole cell lysates (5 µg) of the recombinant strains were analyzed by SDS-PAGE with Coomassie blue staining (a) and Western blot (b). Purified recombinant ML1419c produced in *E. coli* was used as a positive control (lane S) for Coomassie blue staining (500 ng) and Western blot (25 ng).

4.3.5 Phenotypic studies of *P. aeruginosa* PAO1 and *E. coli* BL21(DE3) pLysS expressing *ml1750c*, *ml1419c*, and *ml0397c*

The phenotypic studies of *P. aeruginosa* PAO1 and *E. coli* BL21(DE3) pLysS expressing *ml1750c*, *ml1419c*, and PAS-GGDEF *ml0397c* were performed after the confirmation of protein production in these two bacteria. Results of phenotypic studies are shown in Table 4-4. For *ml1750c*, *P. aeruginosa* and *E. coli* expressing *ml1750c* did not show an alteration of phenotypes

corresponding to increased c-di-GMP level. However, *P. aeruginosa* expressing *ml1750c* had decreased biofilm formation compared to vector control (Figure 4-10). For *ml0397c*, *P. aeruginosa* strains expressing full length *ml0397c* or PAS-GGDEF *ml0397c* did not have alteration in swimming motility and biofilm formation, but they produced an intermediate phenotype of colony morphology (Figure 4-11). *E. coli* expressing full length *ml0397c* did not have an alteration in phenotypes. Importantly, *P. aeruginosa* expressing *ml1419c* and *E. coli* expressing PAS-GGDEF *ml1419c* and *ml0397c* has tremendously changed in phenotypes in a manner consistent to increase of c-di-GMP level. Therefore, this study was performed to investigate functions of ML1419c and ML0397c as diguanylate cyclases (DGCs) to produce c-di-GMP in *P. aeruginosa* and *E. coli*, respectively.

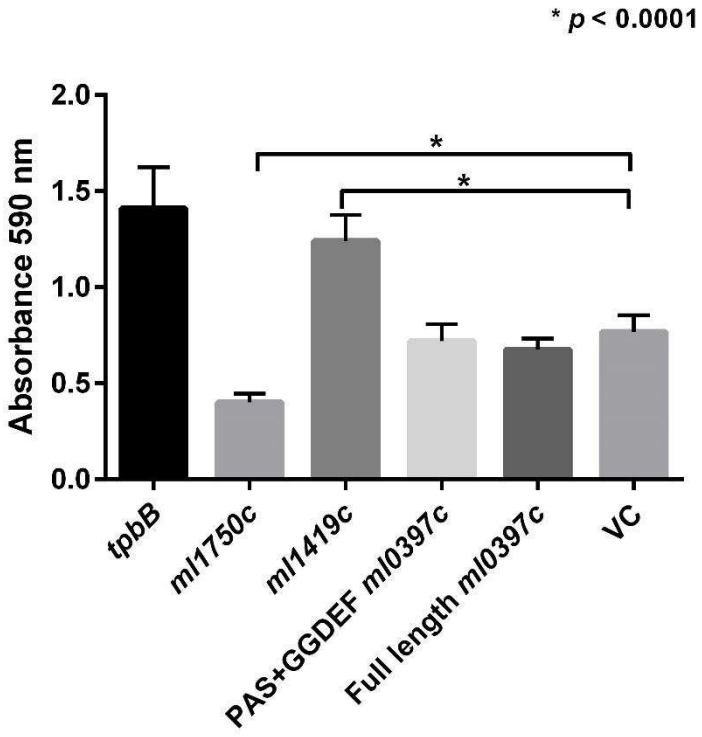


Figure 4-10. Recombinant expression of *ml1419c* and *ml1750c* in *P. aeruginosa* PAO1 alters biofilm formation

Biofilm formation as measured by crystal violet binding was increased in *P. aeruginosa* expressing *tpbB* and *ml1419c* and decreased in *P. aeruginosa* expressing *ml1750c*. The data were averaged from six replicates. Vector control (VC), * $p < 0.0001$.

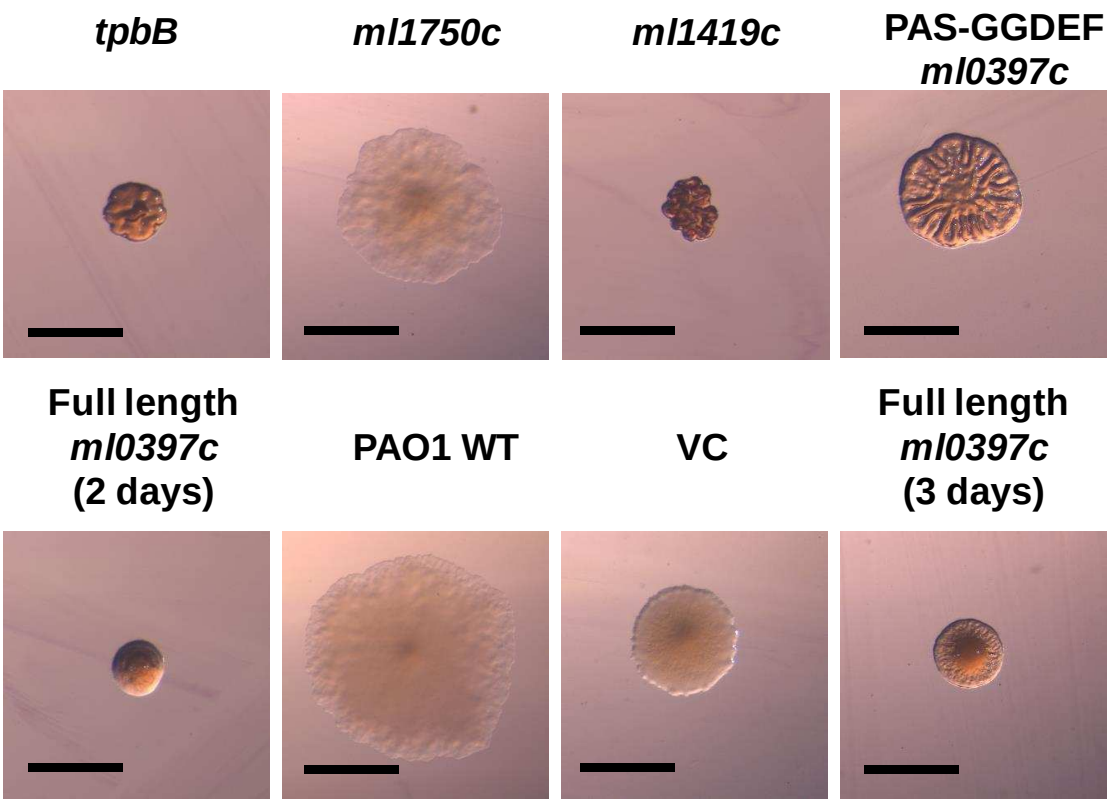


Figure 4-11. Colony morphology of *P. aeruginosa* PAO1 and recombinant strains expressing *M. leprae* DGC genes

Strains were grown on VBMM agar containing Congo red, Brilliant Blue, and 1% L-arabinose (2 days). Rugose colonies were observed in *P. aeruginosa* expressing *tpbB* and *ml1419c*. *P. aeruginosa* expressing PAS-GGDEF *ml0397c* and full length *ml0397c* (3 days) have intermediate colony morphology. *P. aeruginosa* expressing *ml1750c*, PAO1 wild type and vector control (VC) form round colonies with smooth surfaces. Scale bar corresponds to 2 mm.

Table 4-4. Phenotypic studies in *P. aeruginosa* PAO1 and *E. coli* BL21(DE3) pLysS

	***P. aeruginosa* PAO1**					***E. coli* BL21 (DE3) pLysS**	
	Colony morphology	**Congo red binding**	**Swimming motility**	**Twitching motility**	**Biofilm formation**	**Curli formation**	**Cellulose formation**
High c-di-GMP level in bacteria	Rugose	Increased	Inhibited	Inhibited	Increased	Bound Congo red	Fluorescence +
ml1750c	WT[a]	WT	WT	WT	Decreased	WT	WT
ml1419c	Rugose	Increased	Inhibited	Inhibited	Increased	Bound Congo red	Fluorescence +
Full length *ml0397c*	Intermediate	Increased	WT	WT	WT	WT	WT
PAS-GGDEF *ml0397c*	Intermediate	Increased	WT	WT	WT	Bound Congo red	Fluorescence +

[a] Phenotype is similar to vector control or wild type strains; *P. aeruginosa* PAO1 or *E. coli* BL21(DE3) pLysS.

4.3.6 *ml1419c* alters *P. aeruginosa* colony morphology

To provide an initial assessment of whether expression of *ml1419c* resulted in c-di-GMP production in *P. aeruginosa* PAO1, the colony morphology of recombinant *P. aeruginosa* strains was investigated on VBMM Congo red and Brilliant Blue agar plates (Figure 4-12). *P. aeruginosa* PAO1 typically forms round colonies with smooth surfaces and regular borders; however, increased intracellular c-di-GMP levels induced formation of small colonies with wrinkly or rugose colony morphology, and increased Congo red and Brilliant Blue binding that is correlated to the increase of exopolysaccharide (EPS) production (5, 42). The recombinant *P. aeruginosa* PAO1 conditionally expressing *ml1419c* in the presence of arabinose resulted in small and wrinkled colonies exhibiting rugose morphology. These colonies were similar in appearance to the positive control of *P. aeruginosa* expressing *tpbB*, a well-studied DGC. The first two glycine residues of the GGDEF motif of confirmed DGCs participate in binding of the GTP substrate and

the glutamic acid binds Mg^{2+} that is necessary for DGC activity (40, 43). Thus, it was expected that alteration of *P. aeruginosa* colony morphology by *ml1419c* expression would be abrogated by an in-frame deletion of $_{472}$GGDEF$_{476}$ (*ml1419c*$_{\Delta GGDEF}$). *P. aeruginosa* expressing *ml1419c*$_{\Delta GGDEF}$ resulted in larger colonies that resembled the smooth colony morphology of wild-type *P. aeruginosa* PAO1 or vector control (Figure 4-12). These data indicated that recombinant ML1419c functions as a DGC and the GGDEF domain of this protein is essential for this activity.

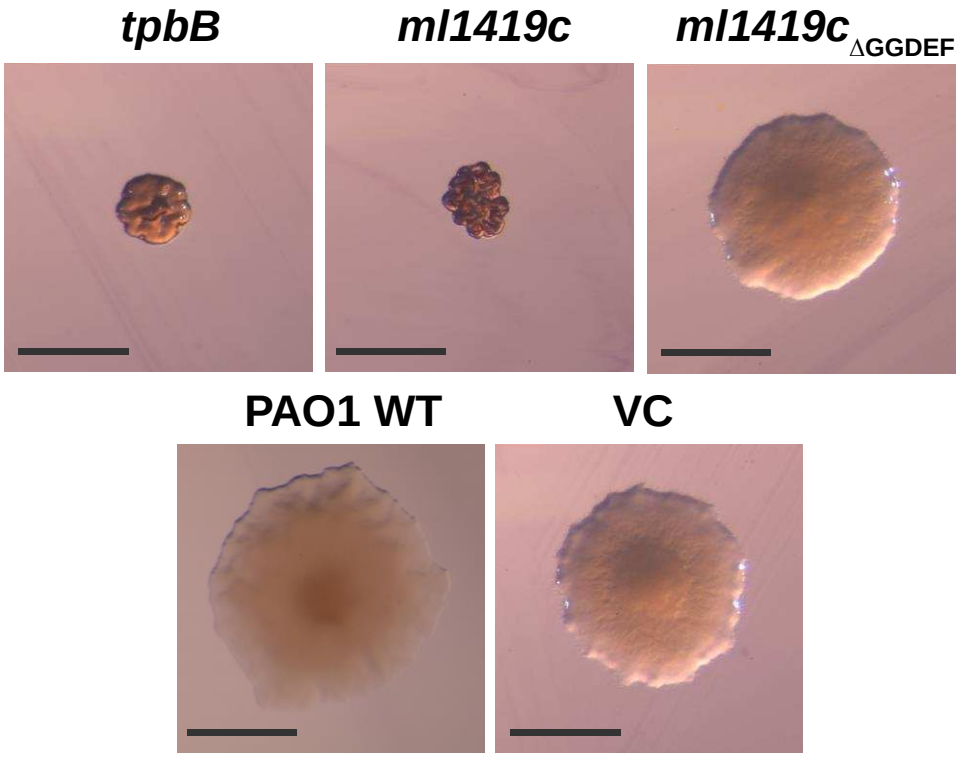


Figure 4-12. Colony morphology of *P. aeruginosa* PAO1 and recombinant strains expressing *ml1419c* constructs
Strains were grown on VBMM agar containing Congo red, Brilliant Blue, and 1% L-arabinose. Rugose colonies were observed in *P. aeruginosa* expressing *tpbB* and *ml1419c*. PAO1 wild type, *P. aeruginosa* expressing *ml1419c*$_{\Delta GGDEF}$, and the vector control (VC) form round colonies with smooth surfaces. Scale bar corresponds to 1 mm.

4.3.7 *ml1419c* expression provides for quantifiable phenotypic differences associated with DGC activity

To provide better assessment of potential c-di-GMP production by ML1419c and the impact of this production on *P. aeruginosa*, several quantifiable phenotypes (twitching motility, swimming motility, and biofilm formation) were measured. Swimming was assessed by quantitatively measuring the swim zone diameter formed by bacteria from the point of inoculation in low-viscosity agar (44). In contrast, twitching motility was quantified by measuring the migration of bacteria that were inoculated at a single point between the agar and the polystyrene petri dish (28, 29). Similar to *P. aeruginosa* expressing *tpbB*, the expression of *ml1419c* suppressed swimming and twitching motility (Figure 4-13 (a), (c)). This is consistent with increased production of c-di-GMP and corresponding inhibition of bacterial flagella and type IV pili function (2, 11). In concordance with decreased DGC activity, alteration of the GGDEF domain in ML1419c (ML1419c$_{\Delta GGDEF}$) resulted in swimming and twitching motility that was similar to wild-type PAO1 and the vector control (Figure 4-13 (a), (c)). The suppression of the swimming and twitching phenotypes of *P. aeruginosa* expressing *ml1419c* was significant when compared to wild-type PAO1 and strains expressing the ΔGGDEF mutation in *ml1419c* (Figure 4-13 (b), (d)). It was also noted that suppression of motility was only observed when *P. aeruginosa* expressing *ml1419c* was grown in the presence of arabinose (data not shown).

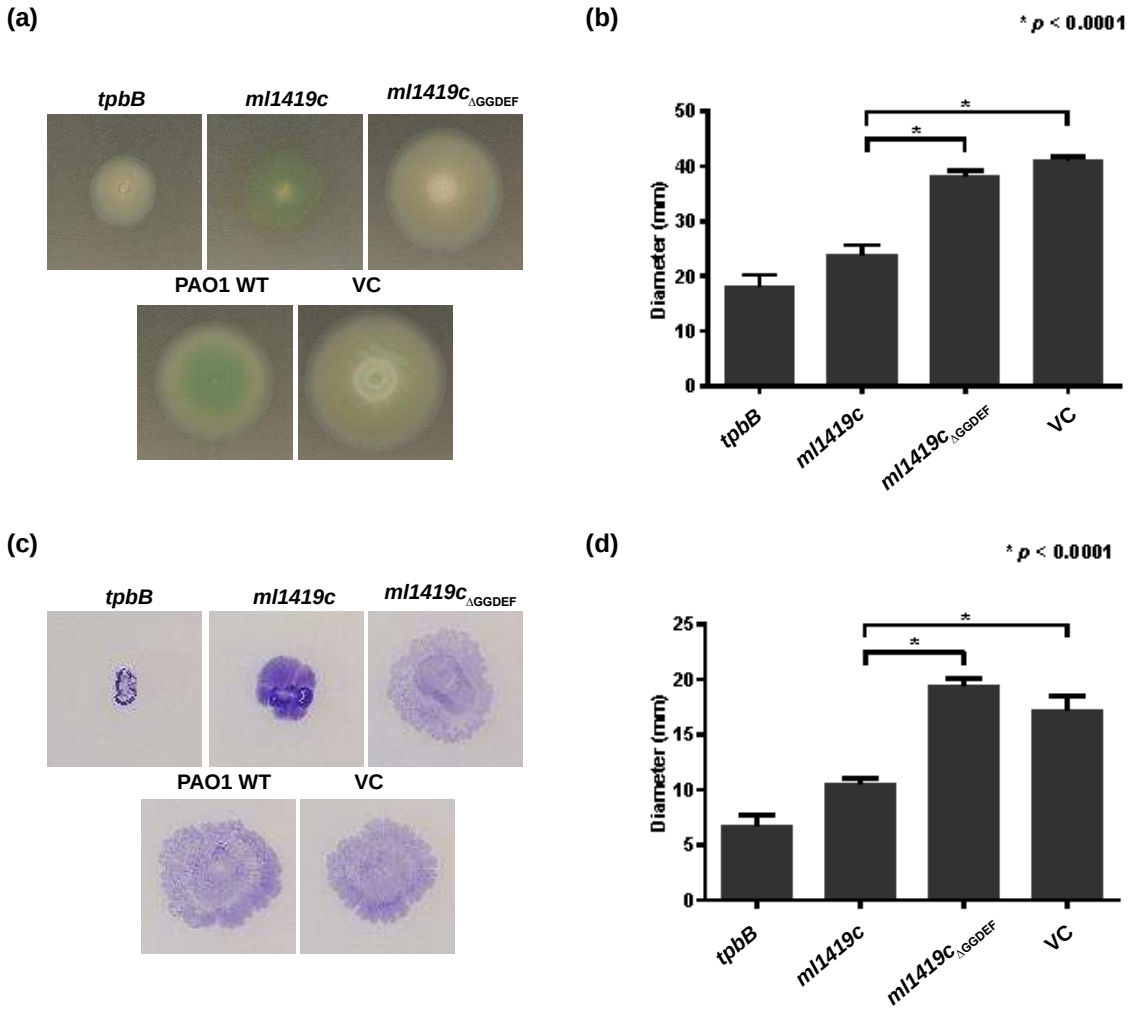


Figure 4-13. *P. aeruginosa* expressing *ml1419c* suppresses motility
(a) Swimming and (c) twitching motility were suppressed in *P. aeruginosa* expressing *tpbB* and *ml1419c* as compared to the wild-type PAO1 strain and vector control (VC) in the presence of 0.2% L-arabinose. Twitching and swimming motility were restored in *P. aeruginosa* expressing mutated *ml1419c* (*ml1419c*$_{\Delta GGDEF}$). The diameter (mm) of (b) swimming and (d) twitching motility zones was measured for four replicates of each strain and averaged, * $p < 0.0001$.

Another well documented activity of c-di-GMP in *P. aeruginosa* is the induction of biofilm formation (5, 19). Elevated cellular levels of c-di-GMP increase the production of biofilm matrix components in *P. aeruginosa* including the Pel and Psl polysaccharides and a biofilm-associated adhesin (4, 5, 12). Under conditions of arabinose induction, *P. aeruginosa* expressing *ml1419c* or

tpbB produced significantly more biofilm as compared to wild-type PAO1 and the vector control (Figure 4-14). As observed with the other phenotypic assays, *P. aeruginosa* biofilm formation was significantly reduced when the GGDEF motif was deleted from ML1419c (ML1419c$_{\Delta GGDEF}$) (Figure 4-14). The effect of *ml1419c* on multiple phenotypes of *P. aeruginosa* PAO1 associated with increased levels of c-di-GMP production provides strong evidence that the *M. leprae* ML1419c functions as a DGC. Additionally, the deletion of GGDEF motif of ML1419c reduced or eliminated *ml1419c* induction of phenotypes associated with elevated cellular levels of c-di-GMP.

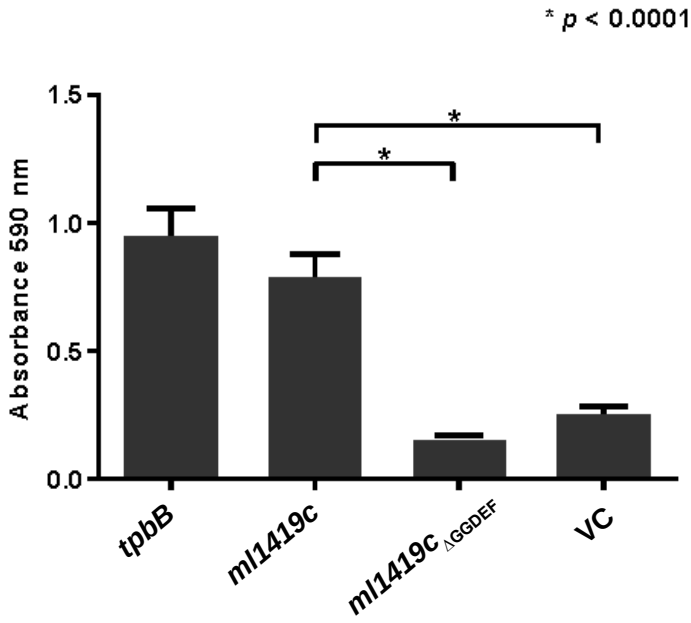


Figure 4-14. Recombinant expression of *ml1419c* in *P. aeruginosa* PAO1 induces biofilm formation

Biofilm formation as measured by crystal violet binding was increased in *P. aeruginosa* expressing *tpbB* and *ml1419c*. Biofilm formation was abrogated in *P. aeruginosa* expressing $ml1419c_{\Delta GGDEF}$. The data were averaged from six replicates. Vector control (VC), * $p < 0.0001$.

4.3.8 ML1419c and PAS-GGDEF ML0397c increased curli and cellulose formation in *E. coli*

The production of proteinaceous extracellular fiber, curli, and cellulose that are the extracellular matrix components (45) was used to provide an initial assessment of whether expression of *ml0397c* resulted in c-di-GMP production in *E. coli*. Typically, c-di-GMP mediates the production of curli and cellulose (20, 46) in *E. coli* resulting in bacteria bound to Congo red and Calcofluor white stain, respectively (21, 47). The *E. coli* expressing *ml1419c* and PAS-GGDEF *ml0397c* exhibited red or light red and white florescent colony in the presence of IPTG. *E. coli* expressing *ml1419c* and PAS-GGDEF *ml0397c* increased the binding of Congo red dye and Calcofluor white stain compared to the vector control under the presence of IPTG.

Furthermore, the full length *ml0397c* can develop light orange color compared to the vector control on the Congo red agar plates with low concentrations of IPTG, 0.1mM (Figure 4-15 and 4-16).

It was expected that alteration of curli and cellulose formation by PAS-GGDEF *ml0397c* expression would be abrogated by an in-frame deletion of $_{219}$GGDEF$_{223}$ (PAS-GGDEF *ml0397c*$_{\Delta GGDEF}$) which has been shown to be involved in active site of DGC (40, 43). However, the *E. coli* expressing PAS-GGDEF *ml0397c*$_{\Delta GGDEF}$ was able to develop orange to red colored and white fluorescent colony but less than *E. coli* expressing PAS-GGDEF *ml0397c* (Figure 4-15 and 4-16). It is likely that the high concentration of IPTG heterologous induction induction is toxic to recombinant *E. coli* strains based on the irregular edge of colonies produced in IPTG concentrations $\geq$ 0.5 mM. This was especially true for the *E. coli* expressing full length *ml0397c*. These data indicated that ML0397c functions as a DGC and the GGDEF domain of this protein is likely essential for this activity.

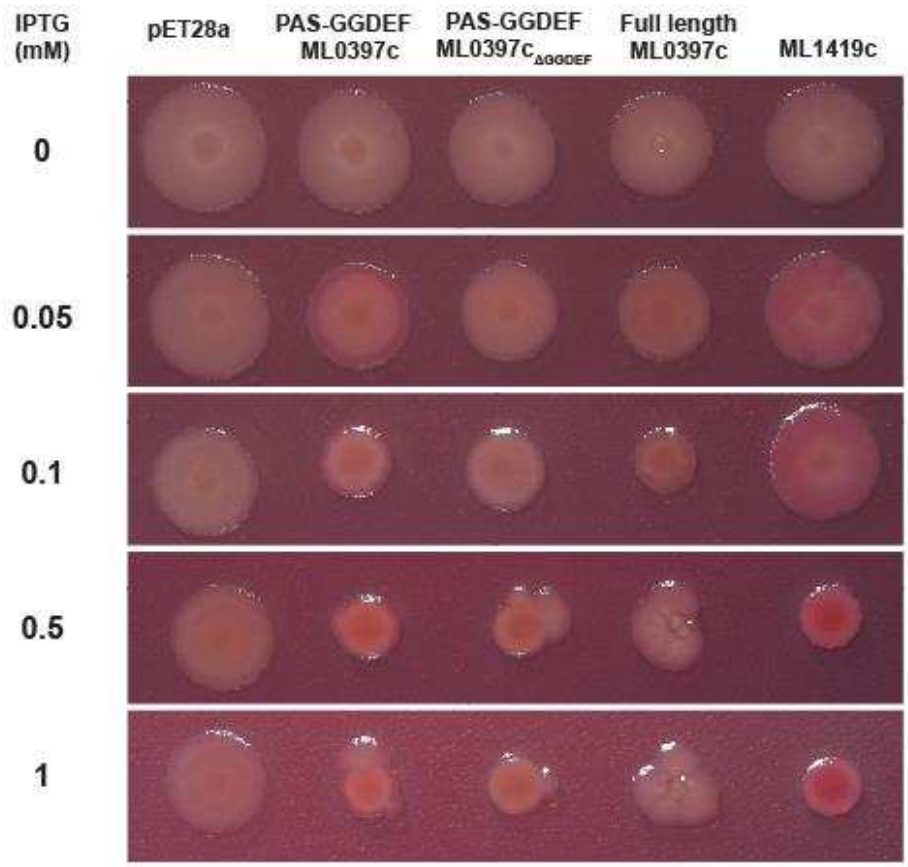


Figure 4-15. Curli formation of recombinant *E. coli* BL21(DE3) pLysS strains
Strains were grown on YESCA agar containing Congo red, kanamycin, chloramphenicol and IPTG. Red color colony was observed in *E. coli* expressing *ml1419c*, PAS-GGDEF *ml0397c* and PAS-GGDEF *ml0397c*$_{\Delta GGDEF}$ but with less color. *E. coli* expressing pET28a vector control and full length *ml0397c* had colorless colonies.

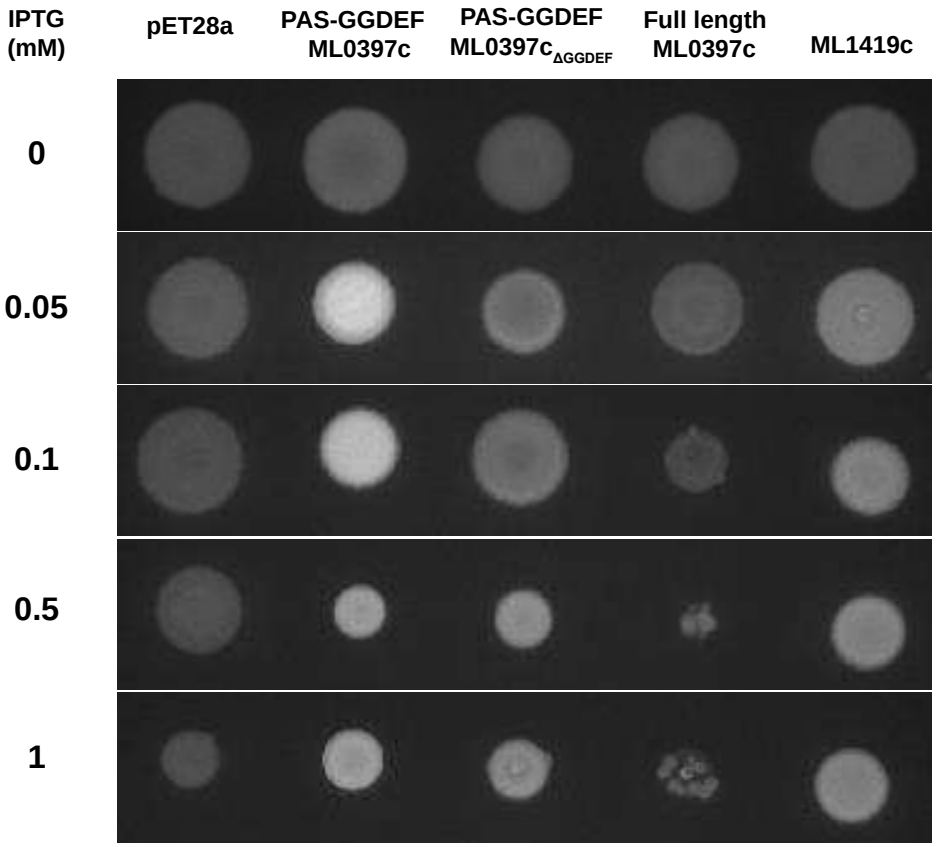


Figure 4-16. Cellulose formation of recombinant *E. coli* BL21(DE3) pLysS strains
Strains were grown on low salt LB agar containing 0.02% Calcofluor white stain. The *E. coli* expressing PAS-GGDEF *ml0397c* exhibited positive fluorescence compared to pET28a vector control. The fluorescence was observed in the *E. coli* expressing *ml1419c* and PAS-GGDEF *ml0397c*$_{\Delta GGDEF}$ with less expression compared to PAS-GGDEF *ml0397c*.

4.3.9 Detection of c-di-GMP *in vivo* by LC-MS

To directly assess the DGC function of ML1419c and ML039c7 in the *P. aeruginosa* or *E. coli* genetic backgrounds, LC-MS was performed to detect and measure the relative abundance of c-di-GMP in *P. aeruginosa* and *E. coli* extracts. Initial analyses of c-di-GMP standard and [^{13}C] adenosine (applied as an internal standard) demonstrated that these two products eluted with retention times of 7.061 and 7.702 min, and yielded *m/z* values of 691.1021 and 269.1065, respectively. Tandem mass spectrometry (MS/MS) fragmentation of c-di-GMP resulted in

transition ions of *m/z* 152.0577, 248.0786, and 540.0566. These transition ions were used to confirm that the parent ion *m/z* 691.1021 represented c-di-GMP (Figure 4-17).

LC-MS analyses of wild-type PAO1 and recombinant *P. aeruginosa* strains revealed the presence of c-di-GMP (Figure 4-18 (a), (b), (c), (d)). The quantitative analyses of the relative c-di-GMP levels were based on the normalized peak area of c-di-GMP to [^{13}C]adenosine. *P. aeruginosa* expressing *ml1419c* significantly increased the abundance (approximately nine fold) of c-di-GMP detected as compared to the vector control (Figure 4-18 (e)). However, the c-di-GMP abundance of *P. aeruginosa* expressing *ml1419c* was lower than *P. aeruginosa* expressing *tpbB* (Figure 4-19). In comparison, there were no significant differences in c-di-GMP abundance between the vector control and *P. aeruginosa* expressing *ml1419c*$_{\Delta GGDEF}$ or PAS-GGDEF *ml0397c* (Figure 4-18 (e)). These data confirm that ML1419c of *M. leprae* functions as a DGC to produce c-di-GMP and that the GGDEF motif is part of the active site domain as shown in other DGCs (39).

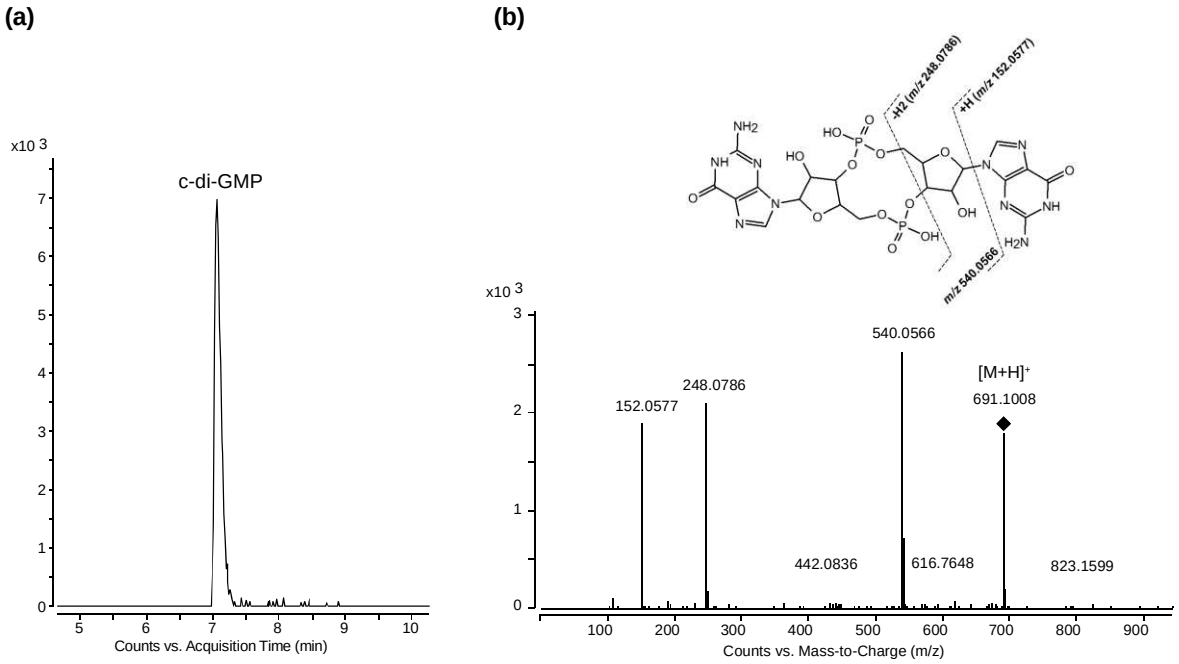
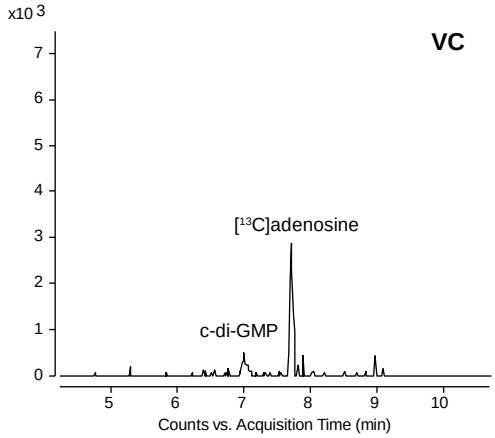


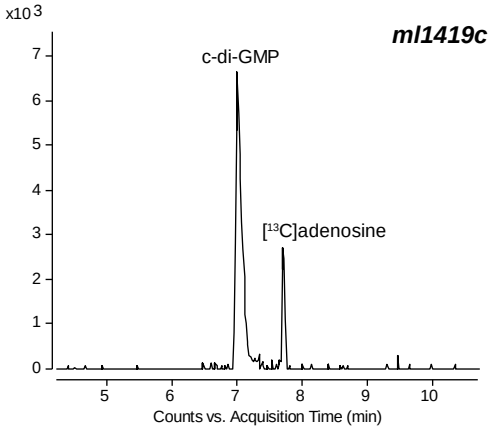
Figure 4-17. Detection of c-di-GMP standard by LC-MS

C-di-GMP standard was separated and detected by LC-MS. (a) The extracted ion chromatogram (EIC) of c-di-GMP (*m/z* 691.102) revealed a retention time of 7.061 min. (b) The c-di-GMP peak was confirmed by tandem mass spectrometry (LC-MS/MS) with the transition ions of *m/z* 152.0577, 248.0786 and 540.0566. The *m/z* 248.0786 fragment ion is diagnostic for c-di-GMP.

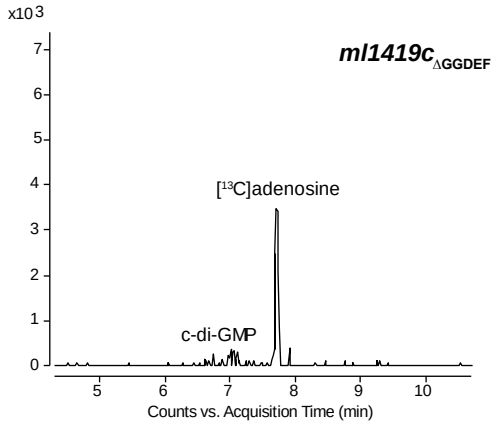
(a)


x10 3
VC
[13C]adenosine
c-di-GMP
Counts vs. Acquisition Time (min)

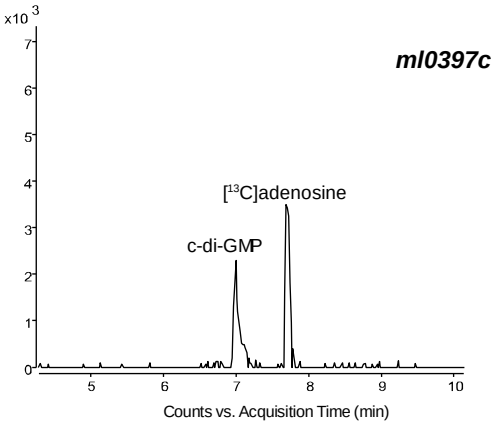
(b)


x10 3
ml1419c
c-di-GMP
[13C]adenosine
Counts vs. Acquisition Time (min)

(c)


x10 3
ml1419cΔGGDEF
[13C]adenosine
c-di-GMP
Counts vs. Acquisition Time (min)

(d)


x10 3
ml0397c
[13C]adenosine
c-di-GMP
Counts vs. Acquisition Time (min)

(e)

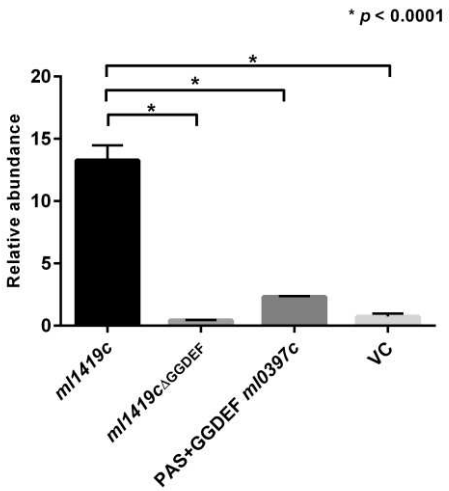

* p < 0.0001
Relative abundance
ml1419c
ml1419cΔGGDEF
PAS+GGDEF ml0397c
VC

Figure 4-18. c-di-GMP detection and relative quantification of c-di-GMP in recombinant *P. aeruginosa* strains
The LC-MS EIC of c-di-GMP (*m/z* 691.102) in extracts of recombinant *P. aeruginosa* strains; (a) vector control (VC), (b) *ml1419c*, (c) $ml1419c_{\Delta GGDEF}$, and (d) PAS-GGDEF *ml0397c*. [^{13}C]adenosine was applied as an internal standard (retention time 7.702 min, *m/z* 269.1065). (e) The relative quantification demonstrated a significant increase in the abundance of c-di-GMP produced in *P. aeruginosa* expressing *ml1419c* as compared to VC and *P. aeruginosa* expressing $ml1419c_{\Delta GGDEF}$, * $p < 0.0001$. Experiments were performed with three biological and three technical replicates.

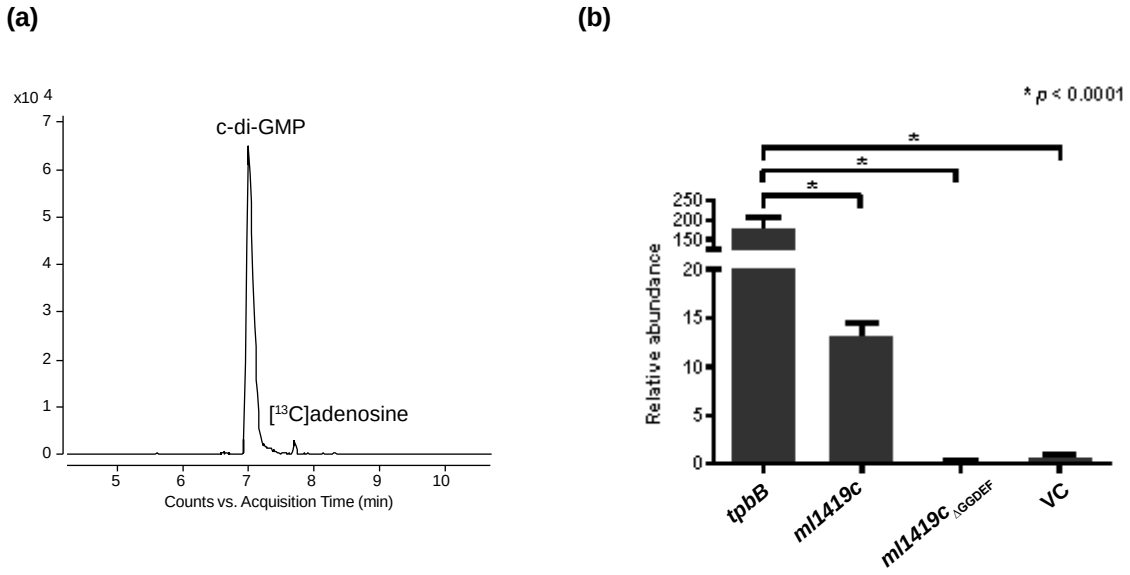


Figure 4-19. c-di-GMP detection by LC-MS and relative quantification of c-di-GMP in *P. aeruginosa* recombinant strains
(a) The EIC of c-di-GMP detected in the extracts from *P. aeruginosa* expressing *tpbB*. (b) The relative quantification of c-di-GMP in *P. aeruginosa* expressing *tpbB* and other recombinant genes.

For assessment c-di-GMP production by *ML0397c* in *E. coli*, the c-di-GMP standard and [^{13}C]adenosine were eluted with retention times of 7.735 and 8.315 min (Figure 4-20), and yielded *m/z* values and the transition ion of *m/z* by LC-MS/MS as shown by ML1419c study (Figure 4-17). The retention times of both c-di-GMP and [^{13}C]adenosine internal standard were shifted from *P. aeruginosa* extract analysis about 0.4 min that may be caused from variation in mobile phase compositions or flow rate and the fact that the LC-MS analyses for the DGC activity of these two

genes products were performed at different times. However, the LC-MS analysis was interfered by an unknown compound with the same mass as [^{13}C]adenosine in LB medium. Consequently, the unknown compound caused ion suppression leading to the absence of [^{13}C]adenosine detection. Therefore, the cell pellets were washed twice with water to get rid of the interfering compound from LB medium.

LC-MS analyses of recombinant *E. coli* strain expressing PAS-GGDEF *ml0397c* revealed the significant increase of detected c-di-GMP abundance (Figure 4-20). In addition, no c-di-GMP was detected in the extracts of *E. coli* expressing *ml1419c* or PAS-GGDEF *ml0397c*$_{\Delta GGDEF}$, or the vector control (Figure 4-20). Furthermore, the c-di-GMP was absent in the extracts of *E. coli* expressing PAS-GGDEF *ml0397c* cultures grown without IPTG induction (data not shown). These data confirm that ML0397c of *M. leprae* functions as a DGC to produce c-di-GMP and that the GGDEF motif is part of the active site domain as shown in other DGCs.

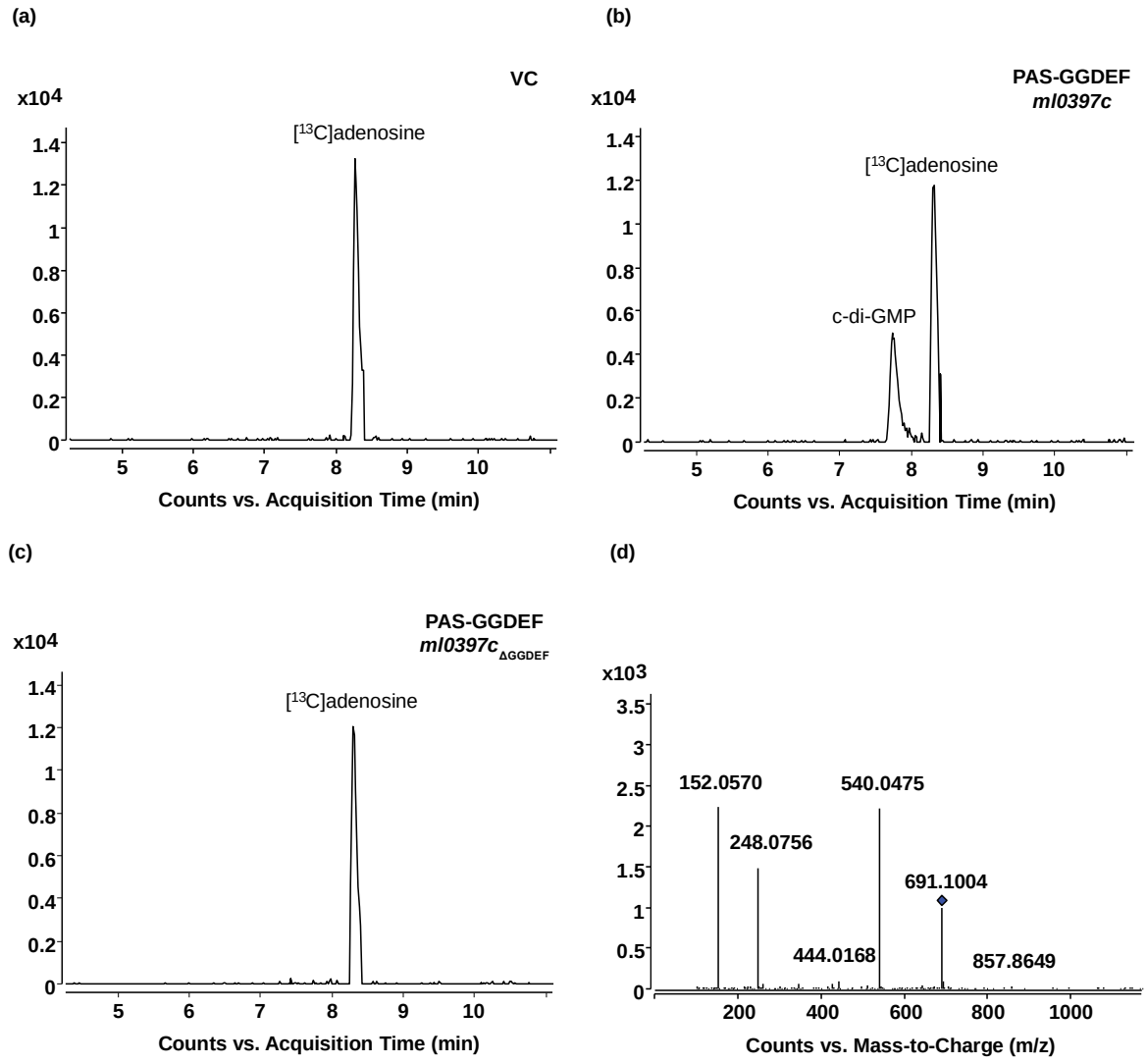


Figure 4-20. c-di-GMP detection by LC-MS of c-di-GMP in *E. coli* recombinant strains
The LC-MS EIC of c-di-GMP (*m/z* 691.102) in extracts of recombinant *E. coli* BL21(DE3) pLysS strains; (a) pET28a vector control (VC), (b) PAS-GGDEF *ml0397c*, and (c) PAS-GGDEF *ml0397c*$_{\Delta GGDEF}$. [^{13}C]adenosine was applied as an internal standard (retention time 8.315 min, *m/z* 269.1065). (d) The c-di-GMP peak of PAS-GGDEF *ml0397c* was confirmed by LC-MS/MS with the transition ions of *m/z* 152.0570, 248.0756 and 540.0475. The *m/z* 248.0756 fragment ion is diagnostic for c-di-GMP. Experiments were performed with three biological and three technical replicates.

4.4 Discussion

This current study demonstrated that ML1419c and ML0397c encoded on the genome of *M. leprae* possess functional DGC activity resulting in the synthesis of c-di-GMP that can be measured by phenotypic or analytical assays. Thus, the *M. leprae* ML1419c protein was renamed as DgcA and the *M. leprae* ML0397c protein was renamed as DgcB. Given the inability to grow or manipulate *M. leprae in vitro,* the use of *P. aeruginosa* and *E. coli* as heterologous expression hosts was critical for demonstrating the enzymatic activity of DgcA and DgcB. DgcA and DgcB from *M. leprae* possess conserved DGC A-site (RFGGDEF) and I-site (RSRD) motifs. The consensus sequence of the A-site motif, RxGG(D/E)EF, has been well studied and is known to participate in protein dimerization, as well as substrate binding and catalytic activity (40, 43). Thus, we hypothesized that deletion of the GGDEF domain would prevent enhanced production of c-di-GMP by recombinant ML1419c and ML0397c and the subsequent loss of *P. aeruginosa* and *E. coli* phenotypes associated with increased DGC activity.

In fact, the recombinant expression of the *M. leprae dgcA* was able to mimic the same phenotypes in *P. aeruginosa* that were induced by over-expression of a well characterized *P. aeruginosa* encoded DGC, *tpbB* (8, 25). Modification and expression of recombinant ML1419c, $ml1419c_{\Delta GGDEF}$, resulted in decreased c-di-GMP levels as compared to that of *P. aeruginosa* expressing full-length *ml1419c*. Likewise, swimming and biofilm phenotypes associated with increased c-di-GMP production were altered. For *dgcB*, *P. aeruginosa* expressing PAS-GGDEF *dgcB* had no changes in phenotypes.

M. leprae DgcA and DgcB were able to alter phenotypes of *E. coli* in a manner consistent of increased c-di-GMP level. However, for DgcB, only the truncated gene containing only PAS and GGDEF domains could be shown to produce DGC activity. The expression of full length *dgcB*

that includes sequences encoding the transmembrane domains was not accomplished. The *E. coli* BL21(DE3) pLysS including other *E. coli* strains adapted for membrane proteins such as Lemo21(DE3) and CD43(DE3) also could not produce full length DgcB in various conditions such as low temperature, and low or high level of IPTG (data not shown). Specifically, the induction of the full length *dgcB* with high concentrations of IPTG (> 0.5 mM) prevented growth in liquid medium (data not shown). This indicates the toxicity of the full protein to *E. coli*. Typically, the expression of transmembrane proteins in heterologous hosts can be toxic for the host, as the transmembrane proteins can result in mis- or unfolded protein (48). Further optimization of the growth conditions, and the selection of a suitable bacterial host and expression vector is required to achieve production of full length recombinant DgcB.

The use of *E. coli* as a model organism for c-di-GMP study is important. Different species of *E. coli* can display various phenotypic characteristics. The *E. coli* BL21(DE3) pLysS contains the T7 RNA polymerase gene that is expressed in under the IPTG-inducible lac UV5 promoter. This subsequently induces a high-level gene expression from T7 promoter driven expression vectors such as pET vectors. In addition, the T7 lysozyme gene in the pLysS plasmid represses the activity of T7 RNA polymerase. This consequently reduces the basal level of protein production from the target gene leading to increase the tolerance to toxic protein (49). For c-di-GMP studies, the *E. coli* BL21(DE3) pLysS strain displayed more tightly control of phenotypic expression compared to *E. coli* BL21(DE3) strain that increased binding to Congo red dye in the absence of IPTG induction (data not shown). The *E. coli* BL21(DE3) pLysS is originated from *E. coli* B strain that lacks the genes cluster for flagellar biosynthesis (50). Therefore, the swimming motility assay that requires flagellar (51) cannot be used to access c-di-GMP function within the *E. coli* BL21(DE3) pLysS strain.

Modification and expression of recombinant PAS-GGDEF *dgcB* (PAS-GGDEF $dgcB_{\Delta GGDEF}$) resulted in absence of c-di-GMP as compared to that of *E. coli* expressing PAS-GGDEF *dgcB*. The inability to detect c-di-GMP in *E. coli* expressing *dgcA*, PAS-GGDEF $dgcB_{\Delta GGDEF}$ and vector control may be related to the c-di-GMP levels were lowere than limit of detection of LC-MS method. The *E. coli* expressing *dgcA* had an alteration in curli formation and cellulose formation, but no c-di-GMP detected by LC-MS. This low level of c-di-GMP in *E. coli* expressing *dgcA* could be caused from the variation in *dgcA* expression at different growth conditions such as medium, duration of growth, inducible condition or the two times washes of bacterial cells with water before c-di-GMP extraction.

Likewise, curli and cellulose formation associated with increased c-di-GMP production of *E. coli* expressing PAS-GGDEF $dgcB_{\Delta GGDEF}$ were reduced, but not absolutely absent. One possible explanation for the intermediate phenotype observed in this study, is that PAS-GGDEF $dgcB_{\Delta GGDEF}$ still functions to interact with target proteins involved in curli and cellulose synthesis in *E. coli.* Data and a supporting model for these types of interactions has been recently proposed to explain how the physical interaction between a DGC and a c-di-GMP binding effector contributes to c-di-GMP signaling (52). The intermediate phenotypes observed for this degenerate GGDEF may allow us to further characterize the functional role of degenerate DGCs in the regulation of c-di-GMP regulated behaviors and better understand cross talk between DGCs and other proteins in future studies.

The expression of *ml1750c* in *P. aeruginosa* PAO1 reduced biofilm formation, but did not alter other phenotypes. This result revealed that the EAL domain of ML1750c is active under this particular growth condition with unknown signaling mechanism. The further study could be performed by expressing *ml1750c* in the overexpressed DGC bacterial strain to elucidate the PDE

activity of ML1750c. Previous studies have shown that the full length protein was important for protein acivity (37). Thus, the inactivation of individual DGC or PDE activity in GGDEF or EAL domain, respectively, with conserved full length protein which would facilitate the study od DGC or PDE protein activity in ML1750c.

Two of the DGC genes from *M. leprae*, (*dgcA* and *dgcB*), were previously found to be expressed during infection in animal models and leprosy patients (15, 16). These data indicate that at various stages during infection, the *M. leprae* DGCs may play important roles in leprosy pathogenesis. The production of DgcA during the early stages of leprosy as described in previous studies (15, 16) and the ability of DgcA to induce a robust immune response (53) indicate an important role for this protein in the pathogens ability to sense and respond to environmental changes during the initial stages of infection. The putative sensing domain of DgcA from *M. leprae* possesses three PAS domains, and two of these PAS domains have conserved heme-binding sites while DgcB harbors one PAS domain with heme-binding pocket. Thus, we hypothesize that DgcA and DgcB likely respond to oxygen tension, nitric oxide and/or carbon monoxide (54). Importantly, the environmental cues that are perceived by the PAS sensor domains of DgcA and DgcB would be expected to alter downstream gene expression and protein function via the activity of DgcA and DgcB derived c-di-GMP. Previous research groups have overexpressed GGDEF proteins similar to DgcA in order to study protein activity in the absence of activating signal (5, 8), and it is likely that elevated protein levels facilitate the dimerization and activation of these DGCs (55). At present, it is unknown whether the DGC activity demonstrated by the recombinant expression of *M. leprae dgcA* in *P. aeruginosa* and *dgcB* in *E. coli* was a result of protein abundance or interaction of the DgcA PAS domains with an environmental signal. Future

structural studies of DgcA and targeted binding assays are required to define the ligands that bind to this protein, and whether they induce or repress DGC activity.

REFERENCES

1. Romling U, Galperin MY, Gomelsky M. 2013. Cyclic di-GMP: the first 25 years of a universal bacterial second messenger. *Microbiol Mol Biol Rev* 77:1-52.

2. Merritt JH, Brothers KM, Kuchma SL, O'Toole GA. 2007. SadC reciprocally influences biofilm formation and swarming motility via modulation of exopolysaccharide production and flagellar function. *J Bacteriol* 189:8154-8164.

3. De N, Navarro MV, Raghavan RV, Sondermann H. 2009. Determinants for the activation and autoinhibition of the diguanylate cyclase response regulator WspR. *J Mol Biol* 393:619-633.

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

5. Hickman JW, Tifrea DF, Harwood CS. 2005. A chemosensory system that regulates biofilm formation through modulation of cyclic diguanylate levels. *Proceedings of the National Academy of Sciences of the United States of America* 102:14422-14427.

6. Gomez MI, Prince A. 2007. Opportunistic infections in lung disease: *Pseudomonas* infections in cystic fibrosis. *Curr Opin Pharmacol* 7:244-251.

7. Lyczak JB, Cannon CL, Pier GB. 2000. Establishment of *Pseudomonas aeruginosa* infection: lessons from a versatile opportunist. *Microbes Infect* 2:1051-1060.

8. Kulasekara H, Lee V, Brencic A, Liberati N, Urbach J, Miyata S, Lee DG, Neely AN, Hyodo M, Hayakawa Y, Ausubel FM, Lory S. 2006. Analysis of *Pseudomonas aeruginosa* diguanylate cyclases and phosphodiesterases reveals a role for bis-(3'-5')-cyclic-GMP in virulence. *Proc Natl Acad Sci U S A* 103:2839-2844.

9. Ryan RP, Lucey J, O'Donovan K, McCarthy Y, Yang L, Tolker-Nielsen T, Dow JM. 2009. HD-GYP domain proteins regulate biofilm formation and virulence in *Pseudomonas aeruginosa*. *Environ Microbiol* 11:1126-1136.

10. Mills E, Pultz IS, Kulasekara HD, Miller SI. 2011. The bacterial second messenger c-di-GMP: mechanisms of signalling. *Cellular Microbiology* 13:1122-1129.

11. Simm R, Morr M, Kader A, Nimtz M, Romling U. 2004. GGDEF and EAL domains inversely regulate cyclic di-GMP levels and transition from sessility to motility. *Mol Microbiol* 53:1123-1134.

12. Borlee BR, Goldman AD, Murakami K, Samudrala R, Wozniak DJ, Parsek MR. 2010. *Pseudomonas aeruginosa* uses a cyclic-di-GMP-regulated adhesin to reinforce the biofilm extracellular matrix. *Mol Microbiol* 75:827-842.

13. Rosano GL, Ceccarelli EA. 2014. Recombinant protein expression in *Escherichia coli*: advances and challenges. *Frontiers in Microbiology* 5:172.

14. Hengge R. 2010. Role of cyclic di-GMP in the regulatory networks of *Escherichia coli*, The Second Messenger Cyclic Di-GMP doi:doi:10.1128/9781555816667.ch16. American Society of Microbiology.

15. Williams DL, Slayden RA, Amin A, Martinez AN, Pittman TL, Mira A, Mitra A, Nagaraja V, Morrison NE, Moraes M, Gillis TP. 2009. Implications of high level pseudogene transcription in *Mycobacterium leprae*. *BMC Genomics* 10:397.

16. Spencer JS, Dockrell HM, Kim HJ, Marques MA, Williams DL, Martins MV, Martins ML, Lima MC, Sarno EN, Pereira GM, Matos H, Fonseca LS, Sampaio EP, Ottenhoff TH, Geluk A, Cho SN, Stoker NG, Cole ST, Brennan PJ, Pessolani MC. 2005. Identification of specific proteins and peptides in *Mycobacterium leprae* suitable for the selective diagnosis of leprosy. *J Immunol* 175:7930-7938.

17. Hickman JW, Harwood CS. 2008. Identification of FleQ from *Pseudomonas aeruginosa* as a c-di-GMP-responsive transcription factor. *Mol Microbiol* 69:376-389.

18. Irie Y, Borlee BR, O'Connor JR, Hill PJ, Harwood CS, Wozniak DJ, Parsek MR. 2012. Self-produced exopolysaccharide is a signal that stimulates biofilm formation in *Pseudomonas aeruginosa*. *Proc Natl Acad Sci U S A* 109:20632-20636.

19. Lee VT, Matewish JM, Kessler JL, Hyodo M, Hayakawa Y, Lory S. 2007. A cyclic-di-GMP receptor required for bacterial exopolysaccharide production. *Mol Microbiol* 65:1474-1484.

20. Weber H, Pesavento C, Possling A, Tischendorf G, Hengge R. 2006. Cyclic-di-GMP-mediated signalling within the sigma network of *Escherichia coli*. *Mol Microbiol* 62:1014-1034.

21. Spurbeck RR, Tarrien RJ, Mobley HL. 2012. Enzymatically active and inactive phosphodiesterases and diguanylate cyclases are involved in regulation of motility or sessility in *Escherichia coli* CFT073. *MBio* 3.

22. Claret L, Miquel S, Vieille N, Ryjenkov DA, Gomelsky M, Darfeuille-Michaud A. 2007. The flagellar sigma factor FliA regulates adhesion and invasion of Crohn disease-associated *Escherichia coli* via a cyclic dimeric GMP-dependent pathway. *J Biol Chem* 282:33275-33283.

23. Newman JR, Fuqua C. 1999. Broad-host-range expression vectors that carry the L-arabinose-inducible *Escherichia coli araBAD* promoter and the *araC* regulator. *Gene* 227:197-203.

24. Wolfgang MC, Lee VT, Gilmore ME, Lory S. 2003. Coordinate regulation of bacterial virulence genes by a novel adenylate cyclase-dependent signaling pathway. *Dev Cell* 4:253-263.

25. Ueda A, Wood TK. 2009. Connecting quorum sensing, c-di-GMP, pel polysaccharide, and biofilm formation in *Pseudomonas aeruginosa* through tyrosine phosphatase TpbA (PA3885). *PLoS Pathog* 5:(6): e1000483. doi:1000410.1001371/journal.ppat.1000483.

26. Smith PK, Krohn RI, Hermanson GT, Mallia AK, Gartner FH, Provenzano MD, Fujimoto EK, Goeke NM, Olson BJ, Klenk DC. 1985. Measurement of protein using bicinchoninic acid. *Analytical Biochemistry* 150:76-85.

27. Choi KH, Kumar A, Schweizer HP. 2006. A 10-min method for preparation of highly electrocompetent *Pseudomonas aeruginosa* cells: application for DNA fragment transfer between chromosomes and plasmid transformation. *J Microbiol Methods* 64:391-397.

28. Deziel E, Comeau Y, Villemur R. 2001. Initiation of biofilm formation by *Pseudomonas aeruginosa* 57RP correlates with emergence of hyperpiliated and highly adherent phenotypic variants deficient in swimming, swarming, and twitching motilities. *J Bacteriol* 183:1195-1204.

29. Darzins A. 1993. The Pilg gene product, required for *Pseudomonas aeruginosa* pilus production and twitching motility, is homologous to the enteric, single-domain response regulator CheY. *Journal of Bacteriology* 175:5934-5944.

30. O'Toole GA. 2011. Microtiter dish biofilm formation assay. *J Vis Exp* doi:10.3791/2437:e2437.

31. Irie Y, Parsek MR. 2014. LC/MS/MS-based quantitative assay for the secondary messenger molecule, c-di-GMP. *Methods Mol Biol* 1149:271-279.

32. Marchler-Bauer A, Derbyshire MK, Gonzales NR, Lu S, Chitsaz F, Geer LY, Geer RC, He J, Gwadz M, Hurwitz DI, Lanczycki CJ, Lu F, Marchler GH, Song JS, Thanki N, Wang Z, Yamashita RA, Zhang D, Zheng C, Bryant SH. 2015. CDD: NCBI's conserved domain database. *Nucleic Acids Res* 43:D222-226.

33. Letunic I, Doerks T, Bork P. 2015. SMART: recent updates, new developments and status in 2015. *Nucleic Acids Res* 43:D257-260.

34. Di Tommaso P, Moretti S, Xenarios I, Orobitg M, Montanyola A, Chang JM, Taly JF, Notredame C. 2011. T-Coffee: a web server for the multiple sequence alignment of protein and RNA sequences using structural information and homology extension. *Nucleic Acids Res* 39:W13-17.

35. Krogh A, Larsson B, von Heijne G, Sonnhammer EL. 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol* 305:567-580.

36. Gupta K, Kumar P, Chatterji D. 2010. Identification, activity and disulfide connectivity of c-di-GMP regulating proteins in *Mycobacterium tuberculosis*. *PLoS One* 5:(11): e15072. doi:15010.11371/journal.pone.0015072.

37. Bharati BK, Sharma IM, Kasetty S, Kumar M, Mukherjee R, Chatterji D. 2012. A full-length bifunctional protein involved in c-di-GMP turnover is required for long-term survival under nutrient starvation in *Mycobacterium smegmatis*. *Microbiology* 158:1415-1427.

38. Hong Y, Zhou X, Fang H, Yu D, Li C, Sun B. 2013. Cyclic di-GMP mediates *Mycobacterium tuberculosis* dormancy and pathogenecity. *Tuberculosis (Edinb)* 93:625-634.

39. Ryjenkov DA, Tarutina M, Moskvin OV, Gomelsky M. 2005. Cyclic diguanylate is a ubiquitous signaling molecule in bacteria: insights into biochemistry of the GGDEF protein domain. *J Bacteriol* 187:1792-1798.

40. Wassmann P, Chan C, Paul R, Beck A, Heerklotz H, Jenal U, Schirmer T. 2007. Structure of $BeF3^-$-modified response regulator PleD: implications for diguanylate cyclase activation, catalysis, and feedback inhibition. *Structure* 15:915-927.

41. Monot M, Honore N, Garnier T, Zidane N, Sherafi D, Paniz-Mondolfi A, Matsuoka M, Taylor GM, Donoghue HD, Bouwman A, Mays S, Watson C, Lockwood D, Khamispour A, Dowlati Y, Jianping S, Rea TH, Vera-Cabrera L, Stefani MM, Banu S, Macdonald M, Sapkota BR, Spencer JS, Thomas J, Harshman K, Singh P, Busso P, Gattiker A, Rougemont J, Brennan PJ, Cole ST. 2009. Comparative genomic and phylogeographic analysis of *Mycobacterium leprae*. *Nat Genet* 41:1282-1289.

42. Starkey M, Hickman JH, Ma L, Zhang N, De Long S, Hinz A, Palacios S, Manoil C, Kirisits MJ, Starner TD, Wozniak DJ, Harwood CS, Parsek MR. 2009. *Pseudomonas aeruginosa* rugose small-colony variants have adaptations that likely promote persistence in the cystic fibrosis lung. *J Bacteriol* 191:3492-3503.

43. Chan C, Paul R, Samoray D, Amiot NC, Giese B, Jenal U, Schirmer T. 2004. Structural basis of activity and allosteric control of diguanylate cyclase. *Proc Natl Acad Sci U S A* 101:17084-17089.

44. O'Toole GA, Kolter R. 1998. Flagellar and twitching motility are necessary for *Pseudomonas aeruginosa* biofilm development. *Mol Microbiol* 30:295-304.

45. Barnhart MM, Chapman MR. 2006. Curli biogenesis and function. *Annu Rev Microbiol* 60:131-147.

46. Brombacher E, Baratto A, Dorel C, Landini P. 2006. Gene expression regulation by the curli activator CsgD protein: modulation of cellulose biosynthesis and control of negative determinants for microbial adhesion. *J Bacteriol* 188:2027-2037.

47. Collinson SK, Doig PC, Doran JL, Clouthier S, Trust TJ, Kay WW. 1993. Thin, aggregative fimbriae mediate binding of *Salmonella enteritidis* to fibronectin. *J Bacteriol* 175:12-18.

48. Miroux B, Walker JE. 1996. Over-production of proteins in *Escherichia coli*: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. *J Mol Biol* 260:289-298.

49. Studier FW, Rosenberg AH, Dunn JJ, Dubendorff JW. 1990. Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol* 185:60-89.

50. Yoon SH, Han M-J, Jeong H, Lee CH, Xia X-X, Lee D-H, Shim JH, Lee SY, Oh TK, Kim JF. 2012. Comparative multi-omics systems analysis of *Escherichia coli* strains B and K-12. *Genome Biology* 13:1-13.

51. Wolfe AJ, Visick KL. 2008. Get the message out: Cyclic-di-GMP regulates multiple levels of flagellum-based motility. *Journal of Bacteriology* 190:463-475.

52. Dahlstrom KM, Giglio KM, Collins AJ, Sondermann H, O'Toole GA. 2015. Contribution of physical interactions to signaling specificity between a diguanylate cyclase and its effector. *MBio* 6:01978-01915.

53. Geluk A, van den Eeden SJ, Dijkman K, Wilson L, Kim HJ, Franken KL, Spencer JS, Pessolani MC, Pereira GM, Ottenhoff TH. 2011. ML1419c peptide immunization induces *Mycobacterium leprae*-specific HLA-A*0201-restricted CTL *in vivo* with potential to kill live mycobacteria. *J Immunol* 187:1393-1402.

54. Henry JT, Crosson S. 2011. Ligand-binding PAS domains in a genomic, cellular, and structural context. *Annu Rev Microbiol* 65:261-286.

55. Hallberg ZF, Wang XC, Wright TA, Nan B, Ad O, Yeo J, Hammond MC. 2016. Hybrid promiscuous (Hypr) GGDEF enzymes produce cyclic AMP-GMP (3', 3'-cGAMP). *Proc Natl Acad Sci U S A* 113:1790-1795.

CHAPTER 5 DIGUANYLATE CYCLASES OF *MYCOBACTERIUM LEPRAE* AND EXPRESSION *IN VIVO*

5.1 Introduction

The c-di-GMP-producing DGC activity of *M. leprae* DgcA and DgcB was demonstrated using recombinant gene expression in heterologous hosts, *P. aeruginosa* and *E. coli*. The production of c-di-GMP by these two proteins could alter phenotypes of model organisms in a manner consistent with increased c-di-GMP. However, the physiological function of these proteins in *M. leprae* is unknown. Elucidation of the functional role of c-di-GMP produced by DgcA and DgcB has great impediments, since *M. leprae* cannot be grown *in vitro* and genetic manipulation cannot be performed on this bacterium. The presence of *M. leprae* gene expression or protein production during infection in leprosy patients or infected animal model tissues may imply cues of biological or physiological roles of those *M. leprae* genes and proteins for the pathogenesis of leprosy.

From genome analyses, *M. leprae* harbors 1,604 predicted open reading frames (ORFs) and 1,116 pseudogenes (1, 2). However, all of these predicted ORFs may not result in functional proteins and transcripts (3-5). We hypothesized that the transcriptional analysis of *M. leprae* genes during infection may provide corresponding information to support further protein analyses. These transcriptional analyses may also help develop a better understanding of those genes required during infection, especially at different clinical spectra of leprosy. Previous techniques used for identification of *M. leprae* gene transcripts in infected tissues from athymic nude mice include reverse transcriptase-polymerase chain reaction (RT-PCR) and DNA microarray (3, 4, 6). By these approaches, 211 *M. leprae* gene transcripts expressed during infection in athymic nude mice

were identified (4). However, the gene encoding proteins related to c-di-GMP production and degradation were not noted in this study. From *M. leprae* whole genome DNA microarrays, Williams *et al*. (3) showed that the *M. leprae ml1752c* encoding the EAL protein in addition to *dgcB* and *dgcA* which encode GGDEF proteins were expressed during infection. These results indicate the presence of potentially functional DGC and PDE in *M. leprae*. In our current study, droplet digital polymerase chain reaction (ddPCR) was applied for the investigation of *M. leprae* DGC associated gene expression in *M. leprae* infected foot pads of nude mice which provides high number of viable bacteria. ddPCR has been introduced as a new technique for quantifying gene expression with high precision and can be used for absolute quantification without the need of generating a calibration curve (7, 8).

Although the study of gene transcription can provide the information of expressed genes under defined time or conditions, study of protein production helps to address the potential function of expressed protein and biological characteristics, such as relative abundance of protein, protein localization and interactions, and also mycobacterium-host relationships (9). The identification of *M. leprae* protein production has been widely studied in infected-armadillo tissue, which provides adequate quantities of bacteria for this purpose (10). A total of 37 *M. leprae* proteins have been identified in *M. leprae* subcellular fractions prior to proteomics approaches (11, 12). Those proteins are involved in several *M. leprae* cellular mechanisms including virulence, lipid metabolism, and cell wall and cellular processes (11). The major purpose of those studies was to identify new antigens for *M. leprae* diagnosis. Since the introduction of proteomics, more than a thousand of proteins from *M. leprae* have been identified using bacilli isolated from infected armadillo tissue (11, 13-16). Recently, de Souza *et al.* identified the largest number of *M. leprae* proteins (1,046 protein products) including ML1752c (EAL containing protein) and ML1750c

(GGDEF-EAL containing protein) from *M. leprae* whole cell lysates by LC-MS/MS (16). Despite the high sensitivity of mass spectrometry was used in previous study (17), DgcA and DgcB have not been identified yet in *M. leprae*. In our current study, using antibodies produced against the DGCs of *M. leprae* (Chapter 4) and proteomics approaches, the presence of these proteins was evaluated using armadillo derived bacilli. Peptide sequences were confirmed by mass spectrometry. Data obtained from this study will provide a better understanding of c-di-GMP with primary cues of this second messenger during infection.

5.2 Materials and methods

5.2.1 ddPCR primer design and optimization

Primers were designed by PrimerQuest Tool from Primer3 program or IDT (Integrated DNA technologies, Coralville, Iowa) PrimerQuest tool (https://www.idtdna.com/Primerquest/Home/Index). Specificity of primers was verified by BLAST with NCBI non-redundant database. Primers were targeted to the center of the desired gene sequences, generating product sizes of approximately 80-100 base pairs (bp). Primer sequences used for ddPCR reaction targeting *dgcA*, *ml1750c*, *dgcB*, and *sigA* are listed in Table 5-1. Primers were tested by convectional PCR reaction with Q5 high fidelity DNA polymerase and *M. leprae* gDNA. The product size was confirmed by the presence of correct products with electrophoresis on 1% agarose gel.

Table 5-1. ddPCR primers used in this study

Genes	Primers		PCR product sizes
ml1750c	Forward	5'- CAT ACG CTG AGA ACT GAG GTT G -3'	91 bp
	Reverse	5'- TAC TCA TGT CTA CCT CAG GAA GG -3'	
dgcA	Forward	5'- TTG GAG GAA GGA GTC ATC GTC -3'	88 bp
	Reverse	5'- GGT CAC TGA CGC CAA GAA TG -3'	
dgcB	Forward	5'- CAT CCT TGA GGT GAG ACC TG -3'	88 bp
	Reverse	5'- CTT GTT CTT GGC TGC ATA CAT C -3'	
sigA	Forward	5'- ACA TCG AAC CAG GTG AAA CC -3'	97 bp
	Reverse	5'- GTC GGC TCA ACG ATC TCT TC -3'	

5.2.2 ddPCR reaction optimization

The QX200 ddPCR EvaGreen reaction mixture (Bio-Rad) was used for DNA quantification according to the manufacturer's manual. The 20 µl reaction included 1X QX200 ddPCR EvaGreen Supermix (Bio-Rad), forward primer 100 nM, reverse primer 100 nM and template. Each ddPCR reaction was loaded into a droplet generation cartridge (DG8 Catridge for QX200) (Bio-Rad) which is specific to the QX200 Droplet Generator (Bio-Rad). QX200 Droplet generation oil (70 µl) for EvaGreen Supermix was loaded into oil wells for each channel. Then, the droplet generation cartridge was sealed and placed into QX200 Droplet Generator (Bio-Rad). Each PCR was partitioned into uniform nanoliter-sized, aqueous droplets in oil (~20,000 droplets per 20 µl reaction). Each droplet can contain one or multiple template copies or no template (7, 8).

The generated droplets were transferred to a 96-well PCR plate (Thermo Scientific) that was placed into conventional thermal cycler (C1000 Touch Thermal Cycler, Bio-Rad). Thermal cycling conditions used for ddPCR EvaGreen mixture reaction are shown in Table 5-2. After PCR amplification, each droplet froma a 96-well plate was examined by the QX200 Droplet Reader (Bio-Rad), and data were analyzed by Bio-Rad Quantasoft analysis software. Droplets containing one or more template copies produced positive end-points of fluorescent signal whereas those

without template remained negative after amplification of PCR reaction. The concentration of template was analyzed by Poisson correction for multiple target molecules per droplet (7, 8).

For optimization of new primer sets, an annealing temperature gradient was used for PCR. The temperature gradient ranged from 50ºC to 65ºC. Templates for the PCR reaction were the synthetically double-stranded gBlock gene fragments ordered from IDT shown in Appendix B. An aliquot (10^{-6} ng) of the gBlock fragment was included into ddPCR reaction. The primer concentration optimization was performed after the optimized annealing temperature was selected by using the final concentration of reverse primer (100 nM) as recommended by the manufacturer's protocol for dPCR EvaGreen Supermix and the forward primer at 100, 150, 200, and 250 nM.

Table 5-2. Thermocycling conditions for ddPCR reaction

Cycling step	**Temperature (ºC)**	**Time**	**Ramp Rate**	**Number of cycles**
Enzyme activation	95	5 min		1
Denaturation	95	30 s		40
Annealing	50-65 (Temperature optimization)	1 min	2ºC/s	40
Signal stabilization	4	5 min		1
	90	5 min		1
Hold	4	infinite		1

5.2.3 *M. leprae* RNA purification

M. lepare RNA was extracted from two granulomatous footpad samples of athymic nude mice (designated as FoxN1nu/FoxN1nu), eight months-post infection with 10^7 *M. leprae* (provided by Dr. Maria Angela De Melo Marques, Colorado State University). Mouse footpads were snap-frozen in liquid nitrogen and homogenized by pulverization (18). *M. leprae* RNA was purified with Trizol reagent (Ambion by life technologies, Grand Island, NY) as described in the manufacturer's protocol. Briefly, the phase separation in homogenized tissue was performed by using 200 µl chloroform per 1 ml of Trizol reagent. RNA was separated from aqueous phase of the samples by 500 µl of 100% isopropanol. The RNA was washed with 75% ethanol and resuspended in DNase, RNase free water. Contaminated DNA was removed from RNA by using TURBO DNA-free Kit (Ambion by life technologies) and RNA purified by phenol:chloroform:isoamyl alcohol (125:24:1, v/v) (Fisher Scientific) extraction. RNA quality and quantity was assessed by Agilent 2100 Bioanalyzer (Agilent Technologies) or Nanodrop 2000 spectrophotometer (Thermo Scientific) at an absorbance 260 nM. DNA-free RNA aliquots were stored at -80ºC.

5.2.4 cDNA synthesis

DNA-free RNA (1µg) was converted to cDNA by using specific reverse primer and Transcriptor First Strand cDNA Synthesis Kit (Roche) per the manufacturer's protocol. Briefly, the cDNA synthesis reaction (1 µg of RNA, 2.5 µM reverse specific primer, DNase and RNase free water) was incubated at 65ºC for 10 min. The following reagents were added to the reaction; 1X reverse transcriptor reverse transcriptase reaction buffer, 20 U protector RNase inhibitor, 1 mM dNTPs, and 10 U reverse transcriptase. The mixture was incubated at 50ºC for 30 min, and

reverse transcriptase was inactivated at 85ºC for 5 min. The cDNA was stored at -20ºC until use. Transcription reactions were performed in replicates of three for each RNA extract from the mouse footpads. The reactions also included negative controls. Non-template control (NTC) was included to determine reagent contamination while non-reverse transcriptase control (NRT) was used to determine DNA contamination.

5.2.5 ddPCR reaction

The ddPCR reactions were performed as described in the primer optimization protocol. cDNA template (50 and 100 ng) synthesized from *M. leprae* RNA was used in the first reaction to optimize an appropriate amount of cDNA. For final assays, 50 ng of cDNA was used for the ddPCR reactions. ddPCR reactions were performed in replicates of three for each cDNA sample. The negative controls from cDNA synthesis reactions were also analyzed by ddPCR. Concentrations of forward primer and reverse primer were 250 nM and 100 nM, respectively. Non-template reaction was included as a negative control to determine contamination in ddPCR reagents. Additionally, a positive control with 10^{-6} ng gBlock DNA was included. The selected annealing temperature was 55ºC obtained from the annealing temperature gradient experiment.

5.2.6 Detection of *M. leprae* proteins encoding DGCs in infected animal tissues

5.2.6.1 SDS-PAGE and Western blot analyses

M. leprae subcellular fractions [whole cell sonicate (JS Batch#97), cytosol fraction (MLSA) (JS Batch#95), cell membrane fraction (MLMA) (JS Batch#95), and cell wall fraction (MLCwA) (JS Batch#94)] were provided by Dr. John Spencer, Colorado State University or BEI resources [MLMA (NR-19331, BEI Batch#61181929), and MLCwA (NR-19333, BEI Batch#61391641)]. Protein concentrations of *M. leprae* subcellular fractions were determined

using the BCA assay. Purified recombinant DgcA (25 ng), PAS-GGDEF DgcB (25 ng) and ML1750c (50 ng) and *M. leprae* subcellular fractions (50 µg) were resolved under denaturing conditions on NuPAGE 4-12% bis-tris polyacrylamide gels (Life technologies). Proteins were transferred to PVDF membranes and probed with rabbit anti-DgcA, anti-ML1750c, or anti-DgcB polyclonal sera (1:100,000) (Term-bleed Day 57). A negative control was probed with pre-bleed serum. The secondary antibody was anti-rabbit IgG conjugated to HRP (1:20,000) (Promega). Chemiluminescence was used to visualize reactive proteins by incubating the membrane with luminol-based enhanced chemiluminescence HRP substrate (SuperSignal West Dura Extended Duration Substrate) (Thermo Scientific).

5.2.6.2 In-gel trypsin digestion of *M. leprae* peptides in whole cell sonicate

M. leprae whole cell lysates (50 µg) were resolved under denaturing conditions on NuPAGE 4-12% bis-tris polyacrylamide gels. Proteins were visualized by staining with Coomassie G-250 stain per the manufacturer's protocol (SimplyBlue SafeStain) (Novex by life technologies). The gel bands ranging from 50-75 kDa were cut into three slices (WCS1, WCS2 and WCS3). Gel pieces were destained in 200 µl of 1:1 acetonitrile (ACN)/50 mM ammonium bicarbonate (Ambic) (Sigma Aldrich) at 60°C for 10 min. The washing step was repeated until destining was complete. Gel pieces were dehydrated in 200 µl of 100% ACN for 5 min followed by reduction in 200 µl of 25 nM dithiothreitol (DTT) (GE Healthcare, LifeScience, Pittsburgh, PA) in 50 mM AmBic at 60°C for 20 min. The gel pieces were incubated in 200 µl of 55 mM iodoacetamide (Sigma Aldrich) in 50 mM AmBic in the dark for 20 min at room temperature, washed in ACN:Ambic (1:1) and dehydrated in ACN before trypsin was applied to gel pieces for in-gel digestion. Gel pieces were rehydrated in 20 µl of trypsin solution (12.5 ng/µl trypsin

(Promega) in 50 mM AmBic) at 37°C for overnight. The digestion reactions were collected and the peptides extracted from the gel pieces with 20 µl of 2.5% trifluoroacetic acid (TFA) (Sigma Aldrich) and vortexing for 15 min. The peptide samples were dried by vacuum, and desalted with Pierce C18 Spin Column (Thermo Scientific) per the manufacturer's protocol before submitting for LC-MS/MS at Proteomics and Metabolomics Facility (PMF) at Colorado State University.

5.2.6.3 Mass spectrometry analysis of *M. leprae* peptides from trypsin-digested whole cell sonicate

The trypsin digested samples were submitted to the PMF at Colorado State University. The following protocols were obtained from PMF.

Mass spectrometry analysis

Peptides were purified and concentrated using an on-line enrichment column (Thermo Scientific 5µm, 100 µm ID x 2cm C18 column). Subsequent chromatographic separation was performed on a reverse phase nanospray column (Thermo Scientific EASYnano-LC, 3µm, 75 µm ID x 100mm C18 column) using a 90 min linear gradient from 10%-30% buffer B (100% ACN, 0.1% formic acid) at a flow rate of 400 nanoliters/min. Peptides were eluted directly into the mass spectrometer (Thermo Scientific, Orbitrap Velos) and spectra were collected over a *m/z* range of 400-2000 Da using a dynamic exclusion limit of 2 MS/MS spectra of a given peptide mass for 30 s (exclusion duration of 90 s). The instrument was operated in Orbitrap-LTQ mode where precursor measurements were acquired in the Orbitrap (60,000 resolution) and MS/MS spectra (top 20) were acquired in the LTQ ion trap with a normalized collision energy of 35kV. Compound

lists of the resulting spectra were generated using Xcalibur 2.2 software (Thermo Scientific) with a S/N threshold of 1.5 and 1 scan/group.

Data analysis

Tandem mass spectra were extracted, charge state deconvoluted and deisotoped by ProteoWizard MsConvert (version 3.0). All MS/MS samples were analyzed using Mascot (Matrix Science, London, UK; version 2.3.01). Mascot was set to search the Uniprot_Mycoleprae_022916rev database (unknown version, 3206 entries) assuming a digestion enzyme trypsin. Mascot was searched with a fragment ion mass tolerance of 0.80 Da and a parent ion tolerance of 20 PPM. Oxidation of methionine, carbamidomethyl of cysteine and carboxymethyl of cysteine were specified in Mascot as variable modifications.

Search results from all samples were imported and combined using the probabilistic protein identification algorithms implemented in the Scaffold software (version Scaffold_4.4.1.1, Proteome Software Inc., Portland, OR). Peptide identifications were accepted if they could be established at greater than 95.0% probability by the Peptide Prophet algorithm (19) with Scaffold delta-mass correction. Protein identifications were accepted if they could be established at greater than 95.0% probability and contained at least 2 identified peptides. Protein probabilities were assigned by the Protein Prophet algorithm (20). Proteins that contained similar peptides that could not be differentiated based on MS/MS analysis alone were grouped to satisfy the principles of parsimony.

5.2.7 Statistical analyses

P values were calculated by one-way analysis of variance (ANOVA) followed by Tukey comparison using GraphPad Prism version 6.0 (GraphPad Software). Data were expressed as mean values ± SD. The p value < 0.05 was considered as statistically significant.

5.3 Results

5.3.1 Quantification of expression of genes encoding DGC proteins by ddPCR

5.3.1.1 ddPCR reaction optimization

The single color ddPCR assay was optimized to quantify the expression level of *M. leprae* genes encoding DGC proteins that produce c-di-GMP. The positive (with template) and negative (without template) ddPCR droplets were discriminated by applying the fluorescence amplitude threshold (8). In addition, only the reactions that had more than 10,000 droplets were analyzed by Poisson statistic analysis by QuantaSoft software. The reference gene (*M. leprae sigA*) was used as a positive control for the presence of *M. leprae* RNA. All primers were specific to *ml1750c*, *dgcA*, *dgcB*, or *sigA*. PCR with gel electrophoresis was used to determine the PCR product size and all primer sets provide the correct product size on 1% agarose gel (data not shown).

A temperature gradient was used to examine the annealing temperature that could discriminate the negative droplets (no template) from positive droplets (with template) with a high copy number of PCR products. From the temperature gradient (50ºC - 60ºC), the temperatures that were lower than 60°C provided obvious separation between groups of positive and negative droplets for all genes as shown in Figure 5-1. However, the concentrations (copy number per µl of PCR reaction) of amplified cDNA from temperatures lower than 60°C did not differ significantly. The temperature chosen for ddPCR reactions for all genes was 55°C.

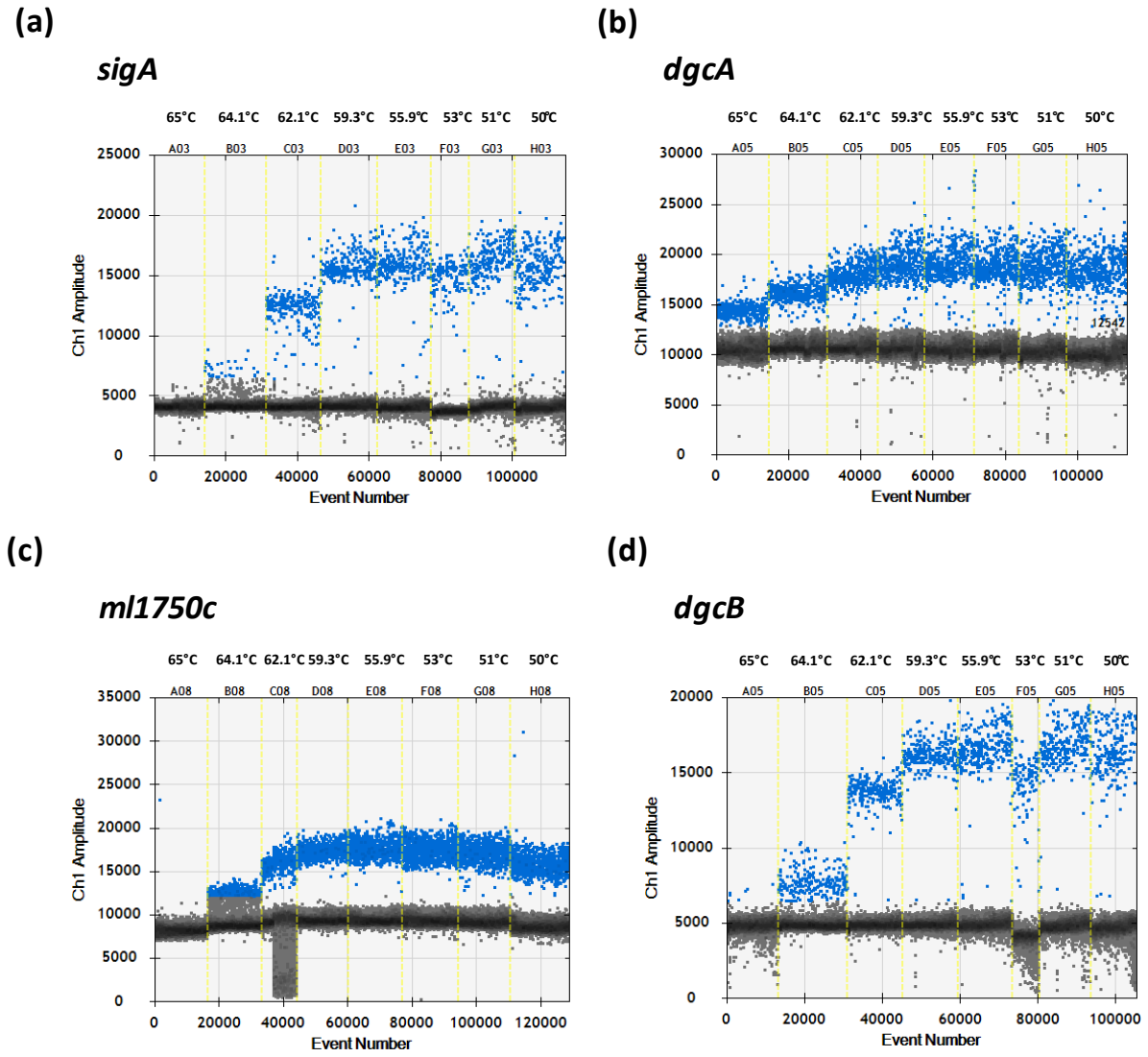


Figure 5-1. Annealing temperature optimization for *sigA*, *dgcA*, *ml1750c* and *dgcB* primer sets
The ddPCR reactions with the annealing temperature gradient for *sigA* (a), *dgcA* (b), *ml1750c* (c), and *dgcB* (d) were analyzed by QuantaSoft software. The positive droplets are depicted in blue color and the negative droplets are in grey color. The X axis represents the fluorescent amplitude of droplets and the Y axis shows event number or droplet number in 20 µl ddPCR reaction.

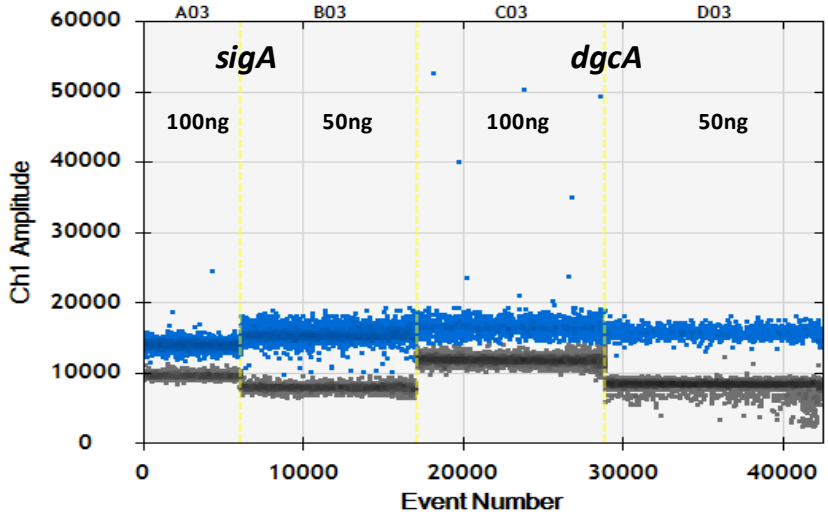
The effect of primer concentration on the ddPCR reactions was determined. Various concentrations of forward primers were used in ddPCR with gBlock as a template. The concentration of reverse primer was fixed at 100 nM. The forward primer concentrations (100, 150, 200, 250 nM) did not have a significant effect to the discrimination between positive and

negative droplets as well as the determination of DNA concentration in ddPCR reaction (data not shown). As a result, the low concentration of forward primer, 100 nM, was selected for ddPCR reaction.

5.3.1.2 ddPCR optimization for *M. leprae* cDNA

The appropriate amount of cDNA for ddPCR was determined using two different cDNA amounts, 100 ng and 50 ng. The separation between positive and negative droplets was reduced in 100 ng cDNA in ddPCR reaction compared to 50 ng cDNA (Figure 5-2 (a)). The poor separation was due to the high fluorescent background of negative droplets. Therefore, 50 ng of cDNA was chosen for the final assays. The concentration of cDNA (copies per µl of ddPCR reaction) is shown in Figure 5-2 (b). 50 ng cDNA is sufficient for quantitation and statistical analyses.

(a)



(b)

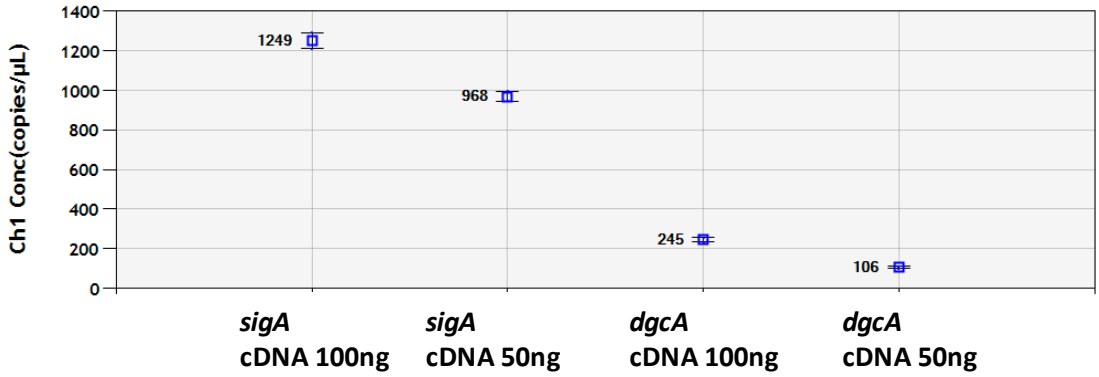


Figure 5-2. Optimization of *M. leprae* cDNA amount for ddPCR reaction
The ddPCR amplitude result (a) shows the decreased separation between negative and positive droplets in reaction with 100 ng cDNA for both *sigA* and *dgcA* genes. (b) the concentrations of DNA are shown in copies per µl of ddPCR reaction. The positive droplets are depicted in blue color and the negative droplets are in grey color. The X axis represents the fluorescent amplitude of droplets and the Y axis shows event number or droplet number in 20 µl ddPCR reaction.

5.3.2 Expression of *M. leprae ml1750c*, *dgcA*, and *dgcB* during infection

Quantitative levels of transcripts from *ml1750c*, *dgcA*, *dgcB*, and *sigA* were determined from RNA extracted from athymic nude mice at eight months-post infection. From ddPCR reactions, all three *M. leprae* genes predicted to encode DGCs were expressed during infection. However, the transcript level of each gene was different. The results of ddPCR reactions from one cDNA samples are depicted in Figure 5-3 and the copy number of nucleic acid detected from 50 ng cDNA (three cDNAs from one mouse footpad) is shown in Table 5-3 and Figure 5-4. The *sigA* reference gene had the highest number of transcripts compared to *dgcA*, *ml1750c*, and *dgcB*, $p < 0.0001$. Additionally, the *ml1750c* transcripts were observed to be greater quantity as compared to *dgcA* and *dgcB*, $p < 0.05$. In conclusion, the *ml1750c*, *dgcA* and *dgcB* were expressed during infection with different level of expression. However, they had lower level of expression compared to the *sigA* reference gene (Figure 5-3).

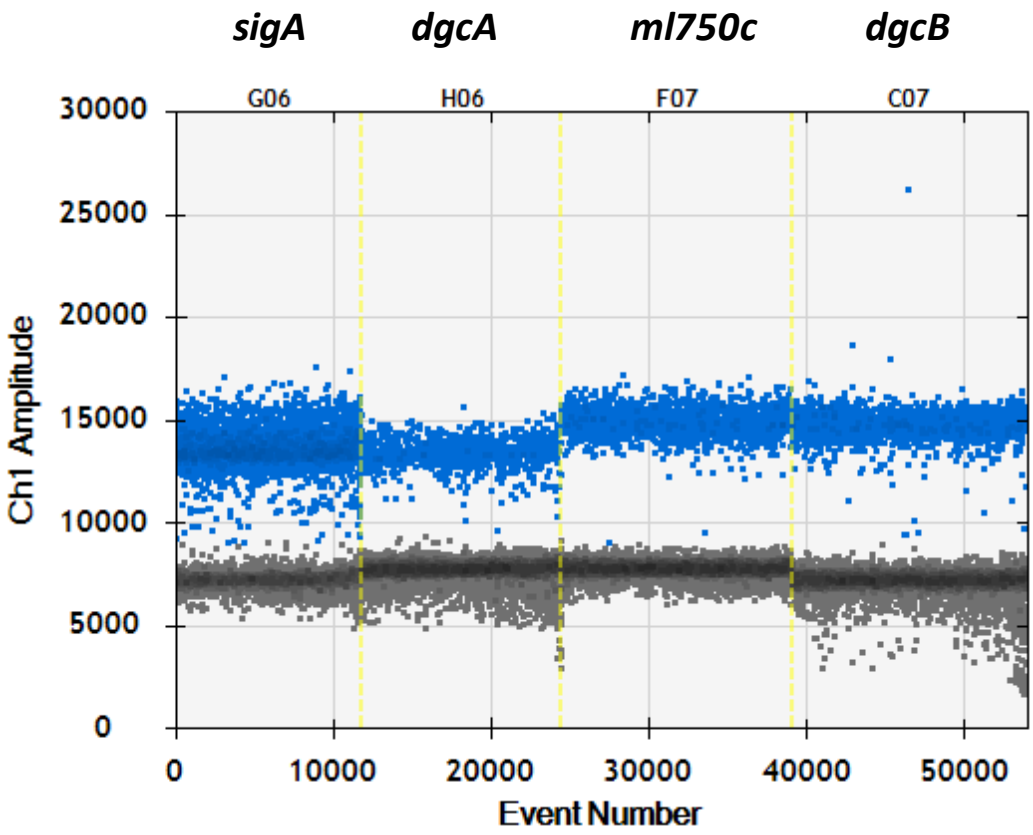


Figure 5-3. ddPCR results of *sigA*, *dgcA*, *ml1750c*, and *dgcB* gene expression
Gene expression determined by ddPCR revealed that all *M. leprae* genes encoding GGDEF proteins were expressed during infection. *sigA* was used as a reference gene. The fluorescent positive droplets of PCR that contained *M. leprae* cDNA is demonstrated in blue. Negative droplets without template are showed in grey. The X axis represents the fluorescent amplitude of droplets and the Y axis shows event number or droplet number in 20 µl ddPCR reaction.

Table 5-3. Absolute quantification (copies/µl) of *sigA*, *dgcA*, *ml1750c*, and *dgcB* from two mouse footpads

Genes	Mouse footpad No. 1			Mouse footpad No. 2		
	cDNA-1	cDNA-2	cDNA-3	cDNA-1	cDNA-2	cDNA-3
sigA	1014.667	1241.5	773	773.33333	843	652.6667
dgcA	71.9	124.75	117.3333	88	106.33333	107.3333
ml1750c	241.6667	301	332.6667	303.33333	372.33333	254
dgcB	121.3333	104.75	164.3333	89	151.33333	86

*Each number is an average from three ddPCR reactions.

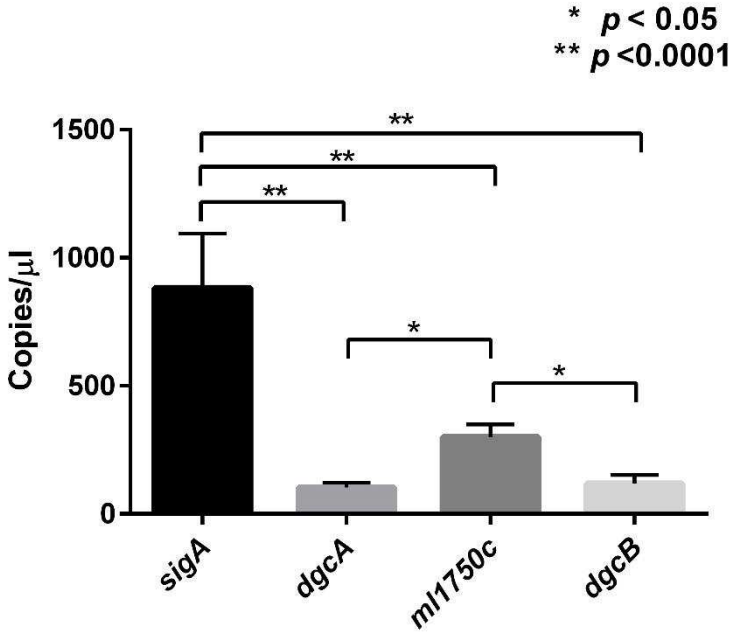


Figure 5-4. Absolute quantification of *sigA*, *dgcA*, *ml1750c*, and *dgcB*
Bar graphs depicting transcript concentrations of *sigA*, *dgcA*, *ml1750c*, and *dgcB* transcripts detected in infected mouse foot pads (copies/µl). Each bar graph represents average concentrations from six cDNA synthesized from two RNA preparations for each gene. * $p < 0.05$, ** $p < 0.0001$.

5.3.3 Detection of *M. leprae* ML1750c, DgcA, and DgcB in *M. leprae* subcellular fractions

DGC protein production in *M. leprae* was determined using specific antibodies. The anti-ML1750c, anti-DgcA, and anti-DgcB polyclonal sera were used to monitor ML1750c, DgcA, and DgcB in *M. leprae* subcellular fractions derived from infected armadillo. Purified proteins from *E. coli* were used as positive controls. From the Western blot analysis, the anti-ML1750c and anti-DgcA polyclonal antibody were able to detect proteins in *M. leprae* whole cell lysate, cell wall and cell membrane fractions at the proteins sizes corresponding to ML1750c (67.4 kDa) and DgcA (60.9 kDa), respectively (Figure 5-5 (A)). However, the positive results for cell wall and cell membrane were not consistant for all sample batches. Cell wall and cell membrane fractions obtained from different batches did not show specific bands for all DGC proteins (Figure 5-5 (B)). DgcA and ML1750c were not detected by pre-term bleed D0 (data not shown). The anti-DgcB polyclonal serum could not detect proteins that have size ~ 63.9 kDa or 34 kDa corresponding to full length DgcB or truncated protein without transmembrane in all subcellular fractions. These experiments demonstrated that ML1750c and DgcA are produced during infection in armidillos.

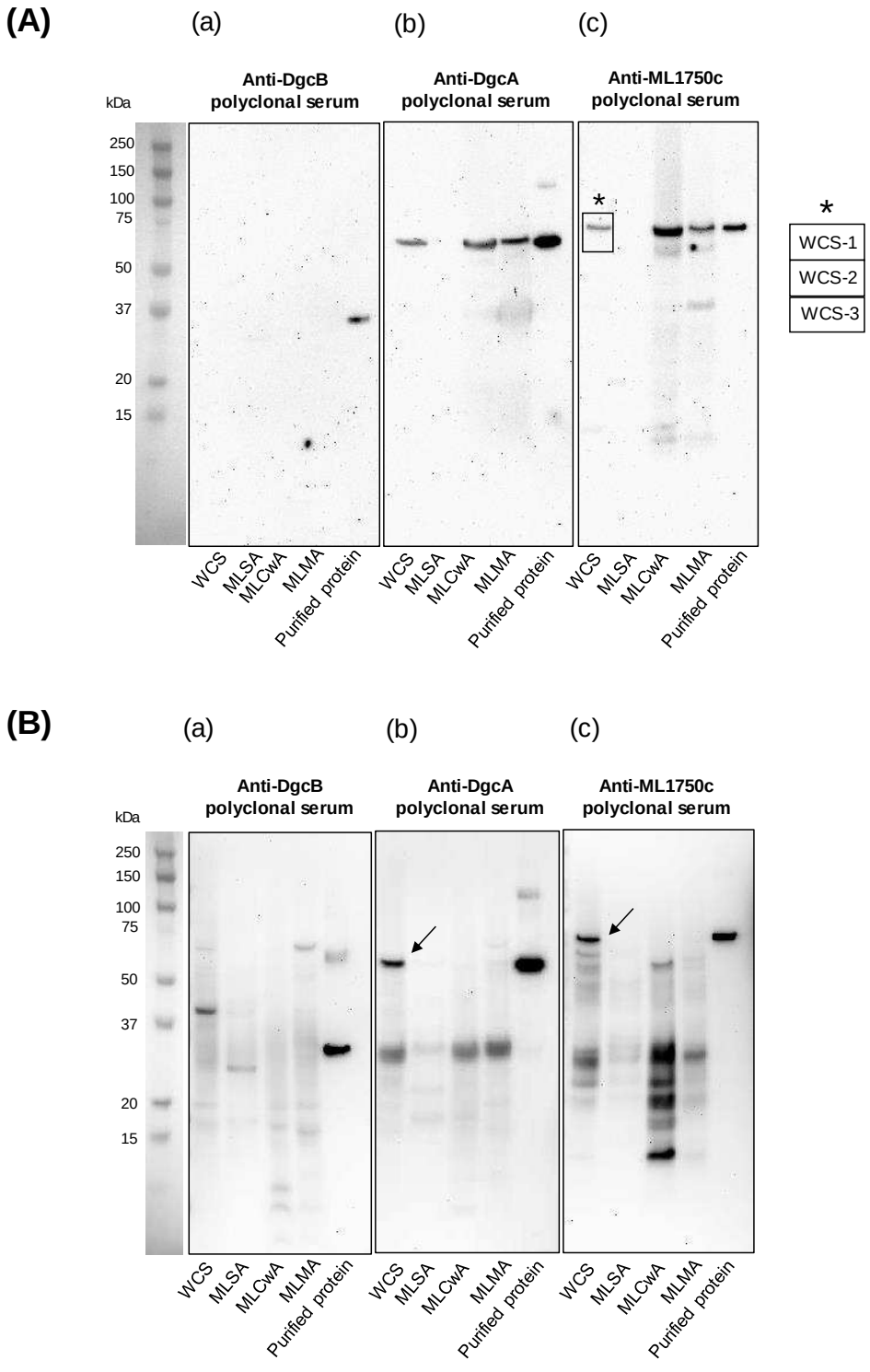


Figure 5-5. Detection of DgcB, DgcA, and ML1750c in *M. leprae* subcellular fractions
Western blot analyses of whole cell sonicate (WCS), cytosolic fractions (MLSA), cell wall (MLCwA) and cell membrane (MLMA) detected with anti-DgcB (a), anti-DgcA (b) and anti-ML1750c (c) polyclonal sera. Purified proteins represent positive controls for the DgcB (34 kDa

for PAS-GGDEF DgcB), DgcA (60.9 kDa) and ML1750c (67.4 kDa). The positive results for MLCwA and MLMA were not consistant for all sample batches. (A) the batches of samples are WCS JS Batch#97, MLSA JS Batch#95, MLCwA BEI Batch#61391641, and MLMA BEI Batch#61181929. (B) the batches of samples are WCS JS Batch#97, MLSA JS Batch#95, MLCwA JS Batch#94, and MLMA JS Batch#95. Boxed area represents fraction that were subsequently purified and further analyzed by LC-MS/MS. Non-specific binding of polyclonal sera to proteins in subcellular fractions were observed.

5.3.4 Detection of *M. leprae* ML1750c, DgcA and DgcB by mass spectrometry

Mass spectrometry was performed to identify peptides related to DGC DgcA, and DgcB and predicted DGC ML1750c from *M. leprae* whole cell sonicate derived from infected armadillo. The protein bands on Coomassie stained gel ranged from 50-75 kDa were selected for trypsin digestion (Figure 5-5 (A)). A total of 240 proteins were identified with greater than 95.0% probability and contained at least 2 identified peptides (Appendix C) as assigned by the Protein Prophet algorithm (20). There are seven proteins identified in *M. leprae* whole cell sonicate that were not identified in previous studies including DgcA (11, 13-16) (Table 5-4).

Table 5-4. New *M. leprae* proteins identified in this study by LC-MS/MS

Identified proteins	Leproma	Molecular Weight (kDa)	Unique peptide count		
			WCS1	WCS2	WCS3
DNA primase	ML0833	71	8		
tRNA-2-methylthio-N(6)-dimethylallyladenosine synthase	ML0989	56			4
Probable GTP pyrophosphokinase	ML0491	87	2		2
Probable ferredoxin/ferredoxin-NADP reductase	ML2134	60			4
Putative phospho-sugar mutase	ML0706	57			3
DNA ligase	ML1705	76			2
Possible regulatory protein	ML1419 (DgcA)	61			3

M. leprae DGC (DgcA) and a hybrid DEGC-PDE (ML1750c) were identified by mass spectrometry from *M. leprae* whole cell sonicate derived from infected armadillo. There are 11, 13, 6 peptides of ML1750c with 100% protein identification probability identified in samples WCS1, WCS2, and WCS3, respectively. For DgcA, there were two peptides with 100% and 99% protein identification probability identified in WCS3. These data were consistant with predicted size of ML1750c (67.4 kDa) and DgcA (60.9 kDa). Percent coverages of identified peptides in ML1750c are 21%, 27% and 9.6% in samples WSC1, WSC2, and WSC3, respectively, and in ML1419c is 5.3% (Table 5-5). All identified peptides of ML1750c and DgcA are listed in Appendix D.

Table 5-5. ML1750c and DgcA protein identification by LC-MS/MS

Samples	Identified proteins	Protein identification probability (%)	Percent coverage (%)	Unique peptide count
WCS-1	ML1750c	100	21	11
WCS-2	ML1750c	100	27	13
WCS-3	ML1750c	100	9.6	6
WCS-3	DgcA	100	5.3	2

This study demonstrated that the *M. leprae* ML1750c and DgcA are produced during infection. Specifically, this study is likely the first to report that DgcA is produced during infection. For the DgcB, no evidence has been shown in protein production from both techniques, which corresponded to previous studies (11, 13-16).

5.4 Discussion

The knowledge about c-di-GMP in mycobacteria especially *M. leprae* is still limited. In previous chapters, the DGC activity of DgcA and DgcB has been shown to be functional and active in surrogate hosts. However, the biological and physiological functions of these proteins in *M. leprae* had not been investigated. Studies of the *M. leprae* genome and proteome have been accomplished and proved important information about potential genes and proteins involved in several cellular mechanisms of *M. leprae* including virulence, protein secretion, lipid metabolism, cell and cell wall process and respiration (3, 4, 11, 13-16). All of the information gained from those studies could provide cues for the proteins required for intracellular survival and development of disease. Nevertheless, the knowledge about c-di-GMP signaling and other aspects of this molecule in *M. leprae* is extremely limited.

A previous study by William et al. (3) has revealed the expression of *dgcA*, *dgcB* genes encoding DGC and the *ml1752c* gene encoding PDE in infected mice. This study exhibited the potential for the existence of c-di-GMP turnover in *M. leprae* cell. Therefore, the transcriptional study of genes encoding DGC was performed for *dgcA*, *dgcB* and especially *ml1750c* which has not previously been shown to be transcribed during infection. The reverse transcriptase reaction and quantitative analysis of transcripts by ddPCR were used to monitor the gene expression. These experiments indicated that the transcripts of *ml1750c*, *dgcA*, and *dgcB* in the infected mouse footpads and the *ml1750c* had the highest level of transcript compared to *dgcA* and *dgcB*. To gain a better understanding about these proteins, their production during infection was determined. Firstly, the detection of these proteins by specific polyclonal antibody was performed. The anti-ML1750c and anti-DgcA polyclonal serum detected proteins with estimated molecular masses that correspond to the theoretical masses of ML1750c and DgcA in *M. leprae* whole cell sonicate. Furthermore, these two proteins were also observed in *M. leprae* cell wall and cell membrane fractions. Interestingly, ML1750c and DgcA did not have domains associated to membrane or cell wall structure, but are present in these two cellular fractions. However, the prediction by PSORTb Subcellular Localization Prediction Tool (http://www.psort.org/psortb/) predicted these two proteins are associated to cytoplasmic membrane. From previous studies, the Rv1354c and MSMEG_2196, homologs of ML1750c are present in membrane fraction of *M. tuberculosis* and *M. smegmatis*, respectively (21, 22). Importantly, from protein-protein interaction study, Cui et al. has shown that the Rv1354c has an interaction with ATB-binding cassette (ABC) transporter in the membrane. It was proposed to be potentially involved in membrane-associated signaling pathway that responds to environmental stresses in *M. tuberculosis* (23). Therefore, the ML1750c and DgcA could be associated to membrane or cell wall by an interaction with other proteins, and

thus explaining their detection in membrane or cell wall fractions. The inconsistency of protein detection in different samples may be caused from the variation in sample preparation and the quantity of proteins in each sample that may be lower than the limit of detection for the immunoblotting method.

The confirmation of the presence of ML1750c and DgcA was performed by mass spectrometry (LC-MS/MS). The polyacrylamide gel was used to separate proteins by molecular mass. This use of gel electrophoresis prior to mass spectrometry analysis removes low molecular weight impurities such as detergents and buffer components that can interfere mass spectrometry analysis as well as enriches for target protein (24). Additionally, the polyacrylamide matrix can store protein in femtomole level. Thus, it decreases the loss of protein during separation (25). From this experiment, the ML1750c and DgcA peptides were identified and the proteins were identifed with high probability. However, only two peptides were identified for DgcA. This could result from the low amount of protein in prepared samples. Moreover, the loss of peptides during sample preparation and incomplete proteolytic digestion by trypsin could lead to the failure of peptide identification. Lastly, the ionization of peptides from DgcA may not have been successful due to peptide properties that require different instrument setting (17, 26). All of these reasons may explain the absence of DgcB in the samples. Nevertheless, DgcB may be produced under different conditions during infection. Therefore, further investigation is needed. This protein identification study was correlated to the datat from gene expression study which *ml1750c* has the greatest level of expression compared to *ml1419c* and *ml0397c*. The proteomic study and expression of gene encoding DGC of *M. leprae* during different stages of infection may provide valuable perspective about proteins related to individual clinical spectrum of leprosy.

The expression of *M. leprae* gene encoding DGC as well as the production of DGC proteins during infection provides new perspectives about c-di-GMP molecules in *M. leprae*. From these data, we concluded that the *M. leprae* has high potential to produce c-di-GMP during infection. Therefore, the c-di-GMP may play important roles for pathogenesis or disease development.

REFERENCES

1. Cole ST, Eiglmeier K, Parkhill J, James KD, Thomson NR, Wheeler PR, Honore N, Garnier T, Churcher C, Harris D, Mungall K, Basham D, Brown D, Chillingworth T, Connor R, Davies RM, Devlin K, Duthoy S, Feltwell T, Fraser A, Hamlin N, Holroyd S, Hornsby T, Jagels K, Lacroix C, Maclean J, Moule S, Murphy L, Oliver K, Quail MA, Rajandream MA, Rutherford KM, Rutter S, Seeger K, Simon S, Simmonds M, Skelton J, Squares R, Squares S, Stevens K, Taylor K, Whitehead S, Woodward JR, Barrell BG. 2001. Massive gene decay in the leprosy bacillus. *Nature* 409:1007-1011.

2. Monot M, Honore N, Garnier T, Zidane N, Sherafi D, Paniz-Mondolfi A, Matsuoka M, Taylor GM, Donoghue HD, Bouwman A, Mays S, Watson C, Lockwood D, Khamesipour A, Dowlati Y, Jianping S, Rea TH, Vera-Cabrera L, Stefani MM, Banu S, Macdonald M, Sapkota BR, Spencer JS, Thomas J, Harshman K, Singh P, Busso P, Gattiker A, Rougemont J, Brennan PJ, Cole ST. 2009. Comparative genomic and phylogeographic analysis of *Mycobacterium leprae*. *Nat Genet* 41:1282-1289.

3. Williams DL, Slayden RA, Amin A, Martinez AN, Pittman TL, Mira A, Mitra A, Nagaraja V, Morrison NE, Moraes M, Gillis TP. 2009. Implications of high level pseudogene transcription in *Mycobacterium leprae*. *BMC Genomics* 10:397.

4. Williams DL, Torrero M, Wheeler PR, Truman RW, Yoder M, Morrison N, Bishai WR, Gillis TP. 2004. Biological implications of *Mycobacterium leprae* gene expression during infection. *J Mol Microbiol Biotechnol* 8:58-72.

5. Suzuki K, Nakata N, Bang PD, Ishii N, Makino M. 2006. High-level expression of pseudogenes in *Mycobacterium leprae*. *FEMS Microbiol Lett* 259:208-214.

6. Spencer JS, Dockrell HM, Kim HJ, Marques MA, Williams DL, Martins MV, Martins ML, Lima MC, Sarno EN, Pereira GM, Matos H, Fonseca LS, Sampaio EP, Ottenhoff TH, Geluk A, Cho SN, Stoker NG, Cole ST, Brennan PJ, Pessolani MC. 2005. Identification of specific proteins and peptides in *Mycobacterium leprae* suitable for the selective diagnosis of leprosy. *J Immunol* 175:7930-7938.

7. Hindson CM, Chevillet JR, Briggs HA, Gallichotte EN, Ruf IK, Hindson BJ, Vessella RL, Tewari M. 2013. Absolute quantification by droplet digital PCR versus analog real-time PCR. *Nat Meth* 10:1003-1005.

8. Hindson BJ, Ness KD, Masquelier DA, Belgrader P, Heredia NJ, Makarewicz AJ, Bright IJ, Lucero MY, Hiddessen AL, Legler TC, Kitano TK, Hodel MR, Petersen JF, Wyatt PW, Steenblock ER, Shah PH, Bousse LJ, Troup CB, Mellen JC, Wittmann DK, Erndt NG, Cauley TH, Koehler RT, So AP, Dube S, Rose KA, Montesclaros L, Wang S, Stumbo DP, Hodges SP, Romine S, Milanovich FP, White HE, Regan JF, Karlin-Neumann GA, Hindson CM, Saxonov S, Colston BW. 2011. High-throughput droplet

digital PCR system for absolute quantitation of DNA copy number. *Anal Chem* 83:8604-8610.

9. James P. 1997. Protein identification in the post-genome era: the rapid rise of proteomics. *Q Rev Biophys* 30:279-331.

10. Balamayooran G, Pena M, Sharma R, Truman RW. 2015. The armadillo as an animal model and reservoir host for *Mycobacterium leprae*. *Clin Dermatol* 33:108-115.

11. Marques MA, Espinosa BJ, Xavier da Silveira EK, Pessolani MC, Chapeaurouge A, Perales J, Dobos KM, Belisle JT, Spencer JS, Brennan PJ. 2004. Continued proteomic analysis of *Mycobacterium leprae* subcellular fractions. *Proteomics* 4:2942-2953.

12. Parkash O, Singh BP. 2012. Advances in Proteomics of *Mycobacterium leprae*. *Scand J Immunol* 75:369-378.

13. Pessolani MC, Brennan PJ. 1996. Molecular definition and identification of new proteins of *Mycobacterium leprae*. *Infect Immun* 64:5425-5427.

14. Marques MA, Chitale S, Brennan PJ, Pessolani MC. 1998. Mapping and identification of the major cell wall-associated components of *Mycobacterium leprae*. *Infect Immun* 66:2625-2631.

15. Marques MA, Neves-Ferreira AG, da Silveira EK, Valente RH, Chapeaurouge A, Perales J, da Silva Bernardes R, Dobos KM, Spencer JS, Brennan PJ, Pessolani MC. 2008. Deciphering the proteomic profile of *Mycobacterium leprae* cell envelope. *Proteomics* 8:2477-2491.

16. de Souza GA, Softeland T, Koehler CJ, Thiede B, Wiker HG. 2009. Validating divergent ORF annotation of the *Mycobacterium leprae* genome through a full translation data set and peptide identification by tandem mass spectrometry. *Proteomics* 9:3233-3243.

17. Aebersold R, Mann M. 2003. Mass spectrometry-based proteomics. *Nature* 422:198-207.

18. Su AI, Wiltshire T, Batalov S, Lapp H, Ching KA, Block D, Zhang J, Soden R, Hayakawa M, Kreiman G, Cooke MP, Walker JR, Hogenesch JB. 2004. A gene atlas of the mouse and human protein-encoding transcriptomes. *Proc Natl Acad Sci U S A* 101:6062-6067.

19. Keller A, Nesvizhskii AI, Kolker E, Aebersold R. 2002. Empirical atatistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Analytical Chemistry* 74:5383-5392.

20. Nesvizhskii AI, Keller A, Kolker E, Aebersold R. 2003. A statistical model for identifying proteins by tandem mass spectrometry. *Anal Chem* 75:4646-4658.

21. Mawuenyega KG, Forst CV, Dobos KM, Belisle JT, Chen J, Bradbury EM, Bradbury AR, Chen X. 2005. *Mycobacterium tuberculosis* functional network analysis by global subcellular protein profiling. *Mol Biol Cell* 16:396-404.

22. Sharma IM, Prakash S, Dhanaraman T, Chatterji D. 2014. Characterization of a dual-active enzyme, DcpA, involved in cyclic diguanosine monophosphate turnover in *Mycobacterium smegmatis*. *Microbiology* 160:2304-2318.

23. Cui T, Zhang L, Wang X, He ZG. 2009. Uncovering new signaling proteins and potential drug targets through the interactome analysis of *Mycobacterium tuberculosis*. *BMC Genomics* 10:118.

24. Shevchenko A, Tomas H, Havlis J, Olsen JV, Mann M. 2006. In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat Protoc* 1:2856-2860.

25. Shevchenko A, Loboda A, Ens W, Schraven B, Standing KG, Shevchenko A. 2001. Archived polyacrylamide gels as a resource for proteome characterization by mass spectrometry. *Electrophoresis* 22:1194-1203.

26. Link AJ, Eng J, Schieltz DM, Carmack E, Mize GJ, Morris DR, Garvik BM, Yates JR. 1999. Direct analysis of protein complexes using mass spectrometry. *Nat Biotech* 17:676-682.

CHAPTER 6 EVALUATION OF *M. SMEGMATIS* AS A MODEL FOR DIGUANYLATE CYCLASE FUNCTION

6.1 Introduction

The inability of *M. leprae* to grow *in vitro* demands alternative model organisms or surrogate hosts for studying protein function and genetic manipulation. The utilization of other *Mycobacterium* species as a model organism for *M. leprae* study is theoretically ideal because of their closely related genomes. *M. smegmatis* is a rapid growing mycobacterium, which requires less than seven days to grow in optimal conditions (1-3). In addition, it is a non-pathogenic mycobacterium and is effective for genetic manipulation. Specifically, the *M. smegmatis* strain mc^2155 can be efficiently transformed with plasmid vectors by electroporation (4). Therefore, *M. smegmatis* has been used as a model organism for research on *M. tuberculosis* and other mycobacteria, including *M. leprae*.

For *M. smegmatis*, this organism harbors only one hybrid GGDEF-EAL containing protein, which is MSMEG_2196. The MSMEG_2196 protein has been revealed to be a bi-functional protein that can produce and degrade c-di-GMP. *M. smegmatis* c-di-GMP is proposed to be involved in long-term survival under starvation (5), and can alter cell length as well as colony morphology (6). Recently, c-di-GMP has been shown to be required for biofilm formation and sliding motility related to glycopeptidolipids (GPLs) and polar lipid production by *M. smegmatis* (7, 8). The LtmA transcription factor has been identified as a c-di-GMP receptor in *M. smegmatis* that regulates the expression of genes related to lipid transport (9). *M. smegmatis* may serve as an alternative model organism for studies on *M. leprae* c-di-GMP because it contains one GGDEF protein that when eliminates would result in the construction of a strain without DGC activity.

Overexpression of *M. leprae* genes encoding DGC in *M. smegmatis* would provide valuable information about protein functions and c-di-GMP roles in the environmental niche closely related to *M. leprae* cells. This study will focus on the ability of *M. smegmatis* to produce c-di-GMP and the biological roles of *M. leprae* proteins ML1750c, DgcA, and DgcB in *M. smegmatis*.

6.2 Materials and methods

6.2.1 Bacterial strains, genomic DNA, and growth conditions

M. smegmatis mc^2155 and recombinant strains were grown at 37°C in 7H9 medium containing 0.2% glycerol and 0.05% Tween 80. Kanamycin (50 µg/ml) (Sigma Aldrich) and hygromycin (50 µg/ml) (EMD Millipore) were used for selection of Δ*msmeg_2196 M. smegmatis* and recombinant *M. smegmatis* strains, respectively. The following reagents were obtained through BEI Resources, NIAID, NIH: Genomic DNA from *M. leprae*, Strain NHDP, NR-19350 and Strain Thai-53, NR-19352.

6.2.2 Construction of *dgcA* and *dgcB* plasmids and recombinant *M. smegmatis* strains

Heterologous gene expression in *M. smegmatis* was accomplished by using the constitutively expressed vector pVV16, containing the *hsp60* promoter. *M. leprae dgcA* and *dgcB* were amplified by PCR from NHDP63 genomic DNA of *M. leprae* with Q5 high fidelity DNA polymerase (New England BioLabs), the *dgcA* forward and *dgcA* reverse primers and *dgcB* forward and *dgcB* reverse primers, respectively (Table 6-2). Underlined sequences represent *Nde*I and *Hin*dIII sites in the forward and reverse primers, respectively. The 1692-bp *dgcA* and 927-bp *dgcB* fragments were cloned into pVV16 using the *Nde*I and *Hin*dIII sites to generate pMRLB118

and pMRLB119. The pMRLB118 and pMRLB119 plasmids were then electroporated in to competent *M. smegmatis* mc^2155 and Δ*msmeg_2196*.

6.2.3 qRT-PCR primer design and optimization for detection of *msmeg_2196* and *sigA*

Primers were designed by the IDT PrimerQuest tool (https://www.idtdna.com/Primerquest/Home/Index). The specificity of primers was verified by BLAST with NCBI's non-redundant database. The primers amplify the center of the target gene sequence, with product sizes of approximately 90-100 bp. Primer sequences used for qRT-PCR reaction are *msmeg_2196* (qRT-PCR) forward and *msmeg_2196* (qRT-PCR) reverse primers targeting *msmeg_2196* gene and *sigA* (qRT-PCR) forward and *sigA* (qRT-PCR) reverse primers targeting *siA* gene (Table 6-2).

The efficiency of qRT-PCR was tested by the LightCycler 480 instrument (Roche) using PCR containing SYBR Green I Master (Roche), forward and reverse primers (400 nM), and synthetic DNA (from IDT) at final concentration of 10^{-1}-10^{-7} ng/reaction. The PCR products were confirmed by the presence of bands at the appropriate size after electrophoresis on a 1% agarose gel.

6.2.4 *M. smegmatis* RNA purification

M. smegmatis mc^2155 was grown in 7H9 broth containing 0.2% glycerol and 0.05% Tween 80. The stationary phase culture was diluted 1:100 (OD_{600} ~0.05) in fresh 7H9 broth containing 0.2% glycerol and 0.05% Tween 80 and grown at 37ºC with shaking at 200 rpm. The cell pellets were collected from independent *M. smegmatis* cultures at different time points: 24 h, 48 h, 72 h, and 96 h, in replicates of four. The bacterial pellets were resuspened in Trizol reagent (Invitrogen,

Grand Island, NY) (1 ml per pellet from 10 ml culture). Cell lysis was performed by bead beating with 10 cycles of 45 sec on, and 3 min off, on ice. To separate RNA, the homogenized samples were incubated at room temperature for 5 min followed by the addition of 200 µl of chloroform per 1 ml of Trizol reagent (Ambion by life technologies). The samples were mixed by shaking vigorously for 15 seconds, followed by incubation for 2–3 min at room temperature. The RNA phase in the top aqueous layer was separated by centrifugation at 12,000xg for 15 min at 4°C. The RNA was then isolated by transferring the aqueous phase to new tubes containing 500 µl of 100% isopropanol and 100 µl of 3 M sodium acetate. The RNA was precipitated at 4ºC overnight and subsequently centrifuged at 12,000xg for 15 min at 4°C. The RNA pellets were washed with 75% cold ethanol. To remove gDNA, the RNA was treated with TURBO DNA-free DNase (Ambion) twice, as the manufacturer's manual, and purified by phenol:chloroform:isoamyl alcohol (125:24:1, v/v) (Fisher Scientific). Lastly, RNA quality and quantity was checked by NanoDrop 2000 spectrophotometer (Thermo Scientific) and analyzed on a 1.2% agarose gel stained with ethidium bromide. The resulting DNA-free RNA aliquots were stored at -80ºC.

6.2.5 cDNA synthesis

DNA-free RNA (1 µg) was converted to cDNA by using specific reverse primers to each gene (Table 6-2). The cDNA synthesis reactions containing 1 µg RNA, 2 µM reverse primer, 500 nM dNTP mix (Invitrogen), and DNase-RNase free H_2O, were incubated for five min at 65ºC, followed by the addition of 1X First-strand buffer (Invitrogen), and 10 mM dithiothreitol (DTT) (Invitrogen). The reaction incubated at 37ºC for two min, followed by the addition of 1 µl of M-MLV reverse transcriptase (200 U) (Invitrogen), and an additional incubation at 37ºC for five min. The reaction was inactivated by heating at 70ºC for 15 min. cDNA was stored at -20ºC until use.

The reactions included negative controls including a non-template control (NTC), and a non-reverse transcriptase control (NRT). The NTC was included to identify contamination in reverse transcription reactions, and the NRT excluded reverse transcriptase to determine if there was any DNA contamination.

6.2.6 qRT-PCR of *msmeg_2196* and *sigA*

The real-time PCR reactions were composed of LightCycler 480 SYBR Green I Master, 100 ng of cDNA, 400 µM of forward and reverse primers (Table 6-2), and DNase-RNase free water were analyzed by the LightCycler 480 instrument. The reactions were done in triplicate. The temperature program used for real-time PCR was shown in Table 6-1. The NTC and NRT were included in the experiments.

Table 6-1. Temperature program used for real-time PCR by LightCycler 480 instrument

Program name	**Pre-incubation**	**Analysis Mode**	**None**
Target (°C)	Acquisition Mode	Hold (hh:mm:ss)	Ramp Rate (°C/s)
95	None	0:05:00	4.4

Program name	**Amplification**		
Cycles	45	Analysis Mode	Quantification
Target (°C)	Acquisition Mode	Hold (hh:mm:ss)	Ramp Rate (°C/s)
95	None	0:00:10	4.4
60	None	0:00:10	2.2
72	Single	0:00:10	4.4

Program name	**Melting Curve**			
Cycles	1	Analysis Mode	Melting Curves	
Target (°C)	Acquisition Mode	Hold (hh:mm:ss)	Ramp Rate (°C/s)	Acquisition (per °C)
95	None	0:00:10	4.4	
65	None	0:00:10	2.2	
97	Continuous		0.11	5

Program name	**Melting Curve**		
Cycles	1	Analysis Mode	Melting Curves
Target (°C)	Acquisition Mode	Hold (hh:mm:ss)	Ramp Rate (°C/s)
40	None	0:00:30	2.2

6.2.7 *msmeg_2196* gene replacement by homologous recombination

M. smegmatis with a *msmeg_2196* marked mutant (Δ*msmeg_2196*) was constructed by homologous recombination using the kanamycin resistance gene replacement strategy (10). The allelic exchange selection was performed using a thermosensitive counter selectable (*ts/sacB*) plasmid, pPR27 (11). The upstream 506-bp gene fragment (F_u) and downstream 468-bp gene fragment (F_d) to *M. smegmatis msmeg_2196* were generated by Q5 high fidelity DNA polymerase. The primers for F_u construction are *msmeg_2196* F_u forward and *msmeg_2196* F_u reverse primers that have *Not*I and *BamH*I sites (underlined sequences), respectively. For the F_d construction, the *msmeg_2196* F_d forward and *msmeg_2196* F_d reverse primers that have *Not*I and *Spe*I sites underlined were utilized (Table 6-2).

The two resulting gene fragments were ligated to a 816-bp kanamycin (Kan^R) resistance cassette obtained from pUC4K (provided by Dr. Mary Jackson, Colorado State University) at *BamH*I and restriction sites. Consequently, the kanamycin resistance gene was flanked by two gene fragments upstream and downstream of *msmeg_2196* (F_u-Kan^R-F_d) (Figure 6-1). The constructed F_u-Kan^R-F_d gene fragment was cloned into pPR27 vector (11) at *Not*I and *Spe*I restriction sites generating pMRLB120, which was transferred into electrocompetent *M. smegmatis* mc^2155 for an allelic exchange. The pPR27 contains a mycobacterial thermosensitive origin of replication (*ts*) that is able to replicate at low temperatures (30 ºC - 32ºC), but not at high temperatures (39 ºC or 42 ºC) in fast growing mycobacteria. In addition, pPR27 vector has *sacB* counter-selectable marker and *xylE* gene, which are used to achieve allelic replacement by selection with sucrose and as a colored marker for vector delivery into the host (11). The homologous recombination was initially selected by catechol colony staining. The selected colonies were grown in LB broth containing 50 µg/ml kanamycin and 0.05% Tween 80

at 30ºC. Double crossover clones were selected on LB agar containing 50 µg/ml kanamycin and 10% sucrose, grown at 42ºC and confirmed by white colony color when exposed to catechol. An allelic exchange at the *msmeg_2196* locus (Δ*msmeg_2196* mutant strain) was confirmed by PCR and sequencing.

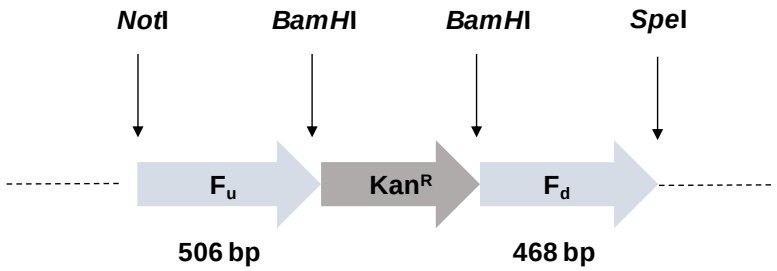


Figure 6-1. Construction of disrupted *msmeg_2196* by kanamycin resistance gene replacement

6.2.8 *M. smegmatis* genomic DNA extraction

Genomic DNA was extracted from the Δ*msmeg_2196* mutant strain as described previously with minor modifications (11). Mycobacterial cells were grown in 7H9 broth containing 0.2% glycerol and 0.05% Tween 80 and harvested by centrifugation at 3,000 rpm, 4°C for 10 min from 10 ml cultures. The mycobacterial cells were incubated at 37°C, rocking overnight in Solution I containing 25% sucrose, 50 mM Tris pH 8, 50 mM ethylenediaminetetraacetic acid (EDTA), and 0.5 µg/µl of lysozyme in DNase-free water. Then, Solution II is added, containing 100 mM Tris pH 8, 0.4 µg/µl proteinase K (Ambion by life technologies), and 1% sodium dodecyl sulfate (SDS) in DNase-free water. The samples were incubated at 55°C for 4 h. DNA extraction was accomplished by using 500 µl of phenol:chloroform:isoamyl (25:24:1) (Amresco, Cleveland, OH), and the aqueous phase containing nucleic acid was transferred to new tubes. The DNA was precipitated with 3 M sodium acetate (Sigma Aldrich) and 100% ethanol. The DNA pellets were

collected by centrifugation and resuspended in DNase-free water. The DNA was stored at 4°C until analysis.

6.2.9 Confirmation of the Δ*msmeg_2196* mutant strain by PCR reaction and sequencing

The DNA sequence of Δ*msmeg_2196 M. smegmatis* was confirmed PCR using Q5 high fidelity DNA polymerase (New England BioLabs) Δ*msmeg_2196 M. smegmatis* genomic DNA, Δ*msmeg_2196* forward and Δ*msmeg_2196* reverse primers that bind to upstream and downstream of *msmeg_2196* (approximately 700-bp upstream and downstream to *msmeg_2196*) (Table 6-2). The PCR product was confirmed by gel electrophoresis on a 1% agarose gel. In addition, the PCR product was cloned into pGEM-T-easy vector (Promega) and sent for DNA sequencing at Proteomics and Metabolomics facility (PMF), Colorado State University.

Table 6-2. Primers used in this study

Primers	**Sequences**
dgcA forward	5'- AAGCATATGGTGTTGGAGACG GTGCGTAG -3'
dgcA reverse	5'- GAGAAGCTTGCTAGGTTGTTGGTTGAACG -3'
dgcB forward	5'- AAGCATATGGCCGACATTACATCAGAGG -3'
dgcB reverse	5'- GAGAAGCTTCGTGACGGCACATTGTTTC -3'
msmeg_2196 (qRT-PCR) forward	5'-CGGTATCGCACAAGGCATTC-3'
msmeg_2196 (qRT-PCR) reverse	5'-AACCCTTGGCCGTCAATA-3'
sigA (qRT-PCR) forward	5'-GCCGAGAAGGGCGAGAAG-3'
sigA (qRT-PCR) reverse	5'-GGTTCGCCTCCAGCAGATG-3'
msmeg_2196 Fu forward	5'- ATTTGCGGCCGCCCGAGGTGATGGACAAG -3'
msmeg_2196 Fu reverse	5'- CGGGATCCGACTGCCATGAGCTGAGTG -3'
msmeg_2196 Fd forward	5'- CGGGATCCGTGAGATGCAGACCCTTCTTG -3'
msmeg_2196 Fd reverse	5'- GGACTAGTCGAGATCATCGAG GCGAACG -3'
Δ*msmeg_2196* forward	5' ACGGTACTGAGCGATCGTTCCGAAATG -3'
Δ*msmeg_2196* reverse	5'- ACC AGGGAGTATCTCGACGAACTCGTG -3'

6.2.10 Phenotypic assays of *M. smegmatis* strains

6.2.10.1 Growth rate study

The growth rate of *M. smegmatis* mc^2155 and Δ*msmeg_2196* was assessed by diluting stationary phase culture (1:100) into 10 ml of Middlebrook 7H9 medium (Becton Dickinson) containing 0.2% glycerol or 2% glucose, 0.05% Tween 80, and the appropriate antibiotic. The cultures were grown at 37°C with shaking at 200 rpm. Bacterial growth was determined by measuring absorbance of turbidity at OD_{600} at different time points. The experiment was performed in triplicate.

6.2.10.2 Colony morphology

The colony morphology of Δ*msmeg_2196 M. smegmatis* was assessed on Middlebrook 7H10 agar (Difco, Becton Dickinson) containing 0.5% glycerol and the appropriate antibiotic by spreading stationary-phase cultures on the medium. The colony morphology was observed after incubation at 37°C for four days. Additionally, the determination of colony morphology after long-term incubation was performed on 1.5% Middlebrook 7H9 agar (Difco) containing 2% glucose (6). The stationary phase cultures were diluted to OD_{600} ~0.5, and 5 µl of diluted culture was inoculated on the agar and grown at 37°C for two weeks and four weeks to determine the colony morphology of Δ*msmeg_2196* mutants after long-term incubation.

6.2.10.3 Biofilm formation

Biofilm formation was assessed as previously described (7, 12, 13) with modifications. Stationary phase *M. smegmatis* cultures were washed with modified Sauton's medium (0.05% potassium phosphate, 0.05% magnesium sulfate, 0.4% L-Asparagine, 0.005% ferric ammonium

citrate, 0.2% citric acid, 4.76% glycerol, pH 7) and diluted to an OD_{600} of ~0.1 in modified Sauton's medium containing the appropriate antibiotic. Aliquots (200 µl) of diluted cultures were added to 96-well polystyrene plates (Microtest flat bottom (#351172), Becton Dickinson) in replicates of six. Plates were sealed with Parafilm M and incubated at 37°C for three days. The 96-well plates were washed twice with water, stained with 0.1% crystal violet for 45 min, and washed twice with water. Bound crystal violet was solubilized with 30% acetic acid and the absorbance was measured at 590 nm.

6.2.10.4 Pellicle formation

Pellicle formation was assessed by inoculating 20 µl of diluted cultures as described in biofilm formation study into 5 ml modified Sauton's medium without Tween 80 in 12 ml culture tubes (KIMAX, Kimble Chase, Rockwood, TN). The pellicle tubes were incubated at 37°C for three days without disturbance.

6.2.10.5 Sliding motility

Sliding motility was assessed as previously described (14) with minor modifications. Sliding motility was assayed by inoculating 3 µl of a *M. smegmatis* stationary phase culture onto low-viscosity 7H9 agar (0.3% ultrapure agarose agar (Invitrogen)) containing 2% glucose or without glucose. The sliding motility was observed after growth at 37°C for 72 h. The pictures of sliding motility were taken with white epi illumination under the Gel Doc XR imager (Bio-Rad).

6.2.11 c-di-GMP detection in *M. smegmatis* by LC-MS analysis

M. smegmatis mc^{2}155 and Δ*msmeg_2196* mutant were grown in 1 L 7H9 medium containing 0.2% glycerol at 37°C with shaking at 200 rpm. Cells were harvested at different time points; 24 h, 48 h, 72 h, and 96 h. *M. smegmatis* mc^{2}155 was also cultivated in low carbon source (0.02% glycerol) 7H9 medium to determine c-di-GMP production under starvation as described from a previous study (5). The c-di-GMP extraction and detection was performed following the protocol in Chapter 3 (3.2.2 and 3.2.5).

6.2.12 Statistical analyses

P values were calculated by unpaired t test statistic or one-way of variance (ANOVA) using GraphPad Prism version 6.0 (GraphPad Software). Data were expressed as mean values ± SD. The p value < 0.05 was considered statistically significant.

6.3 Results

6.3.1 Construction of Δ*msmeg_2196* mutant

The level of c-di-GMP in *M. smegmatis* has been shown to be influenced by DGC and PDE activities from the bi-functional MSMEG_2196 (5). The deletion of MSMEG_2196 would eliminate physiological and biological roles regarding this protein in *M. smegmatis*. *M. smegmatis* Δ*msmeg_2196* would provide advantages as a surrogate host to study *M. leprae* DGC encoding gene under a background without DGC activity. The construction of *M. smegmatis* Δ*msmeg_2196* was accomplished by introduction of the delivery vector, pPR27, into *M. smegmatis* mc^{2}155. The disrupted *msmeg_2196* (with the kanamycin resistance gene) replaced the wild type *msmeg_2196* using homologous recombination. The confirmation of the Δ*msmeg_2196* mutant was performed

by PCR with primers that bind to sequences ~700- bp upstream and downstream to *msmeg_2196*. The PCR product sizes acquired from wild type *msmeg_2196* and Δ*msmeg_2196* mutant were 3,134-bp and 2,681-bp, respectively (Figure 6-2). From DNA sequence analysis, the Δ*msmeg_2196* mutant had the kanamycin resistance gene sequence replacing *msmeg_2196* (data not shown). These experiments indicate the construction of *msmeg_2196* disruption by the kanamycin resistance gene. The *M. smegmatis* Δ*msmeg_2196* strain was further investigated for phenotypic characterization.

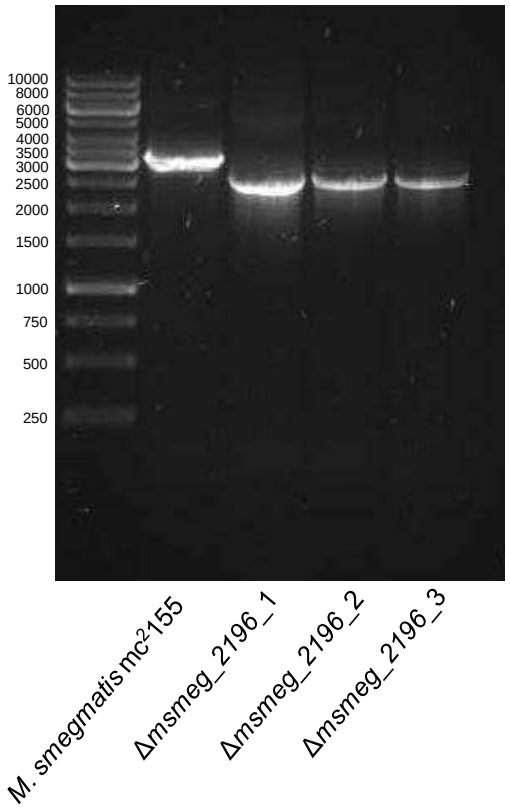


Figure 6-2. Electrophoresis gel of PCR products amplified from *M. smegmatis* mc^{2}155 and Δ*msmeg_2196*
Gel electrophoresis showing PCR products of the *msmeg_2196* gene from *M. smegmatis* mc^{2}155 and three clones of Δ*msmeg_2196* mutants using primers binding to upstream and downstream sequences of the *msmeg_2196* gene.

6.3.2 Phenotypic characterization of the *M. smegmatis* Δ*msmeg_2196*

6.3.2.1 Growth rate

Growth rates of *M. smegmatis* mc^{2}155 and the *msmeg_2196* mutant were determined under two nutrient rich conditions (either with 0.2% glycerol or 2% glucose). The Δ*msmeg_2196* mutant had a similar growth pattern compared to *M. smegmatis* mc^{2}155. The disruption of *msmeg_2196* did not show significant alteration in growth profiles under rich nutrient conditions (Figure 6-3).

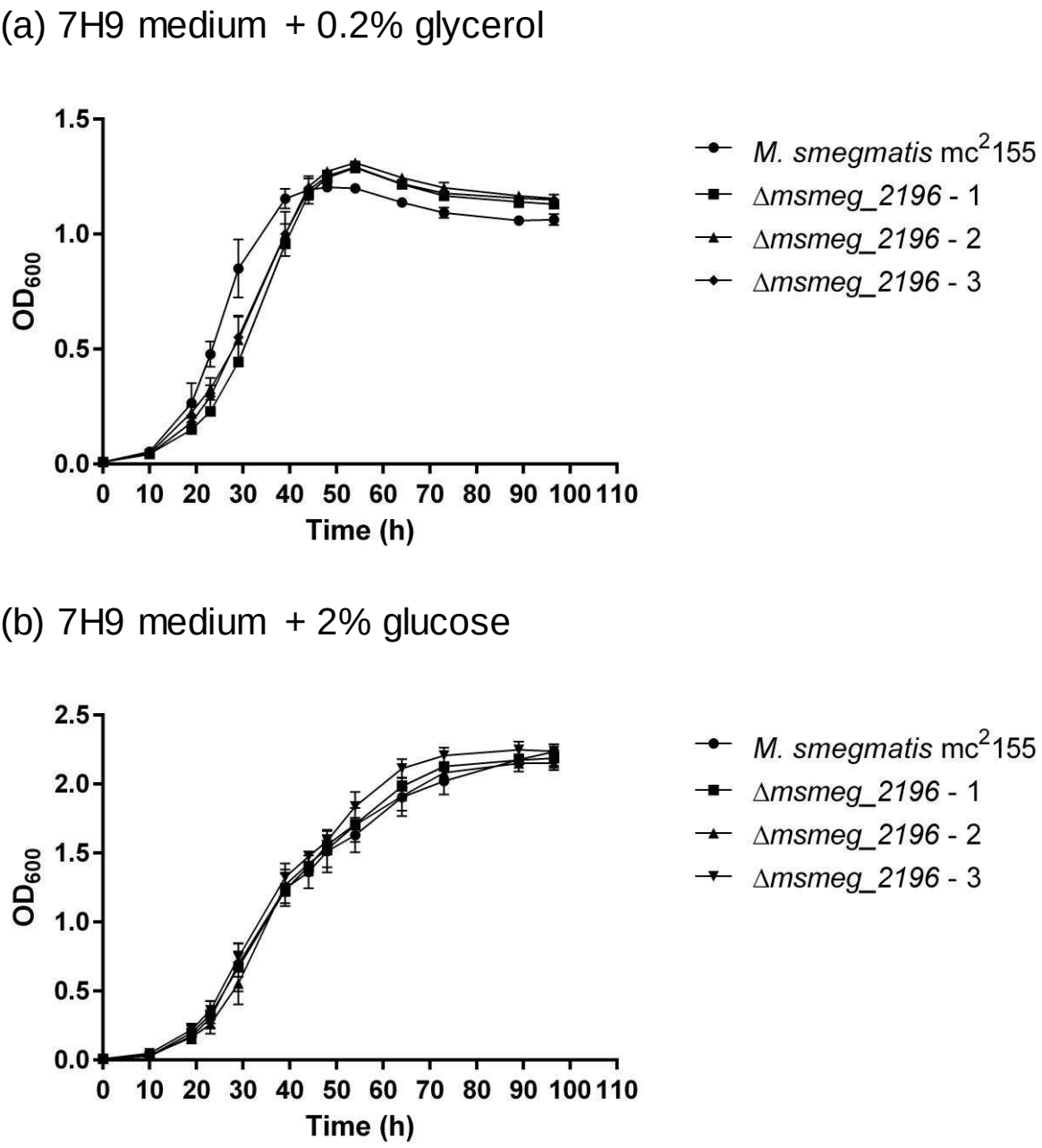


Figure 6-3. Growth curves of *M. smegmatis* mc^{2}155 and Δ*msmeg_2196*
Grow curves of *M. smegmatis* mc^{2}155 and Δ*msmeg_2196*: (a) 7H9 medium containing 0.2% glycerol or (b) 7H9 medium containing 2% glucose. X-axis represents time (hours). Y-axis represents absorbance values at 600 nm. Shown data are mean values with SD. In each condition, three clones of the Δ*msmeg_2196* mutants were tested.

6.3.2.2 Colony morphology

The alteration in colony morphology of the Δ*msmeg_2196* mutant was assessed on 7H10 containing 0.5% glycerol grown at 37°C for four days. Colony morphology of the Δ*msmeg_2196* mutant was not altered from that of *M. smegmatis* mc^2155. Both strains formed rough colonies after a 4-day incubation (Figure 6-4). Long-term incubation (2 weeks and 4 weeks) did not impact on colony morphology of the Δ*msmeg_2196* mutant when compared to the *M. smegmatis* mc^2155 strain (Figure 6-5). Both bacterial strains formed irregular-edge colonies with wrinkled surface.

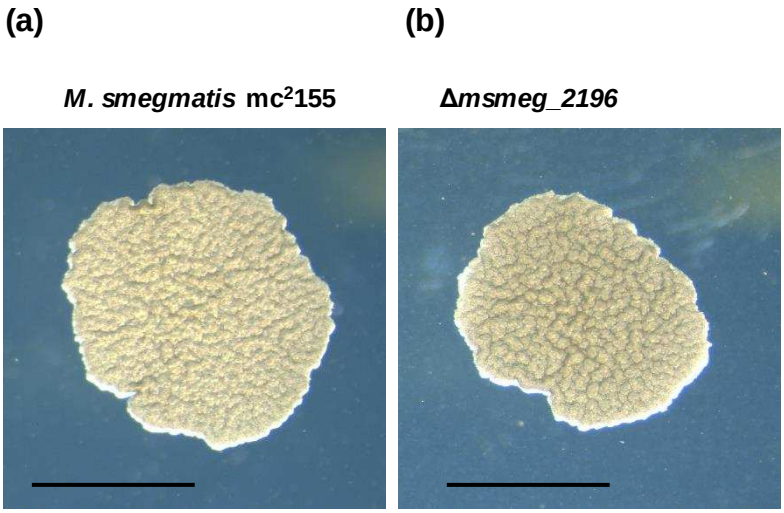


Figure 6-4. Colony morphology of *M. smegmatis* mc^2155 and Δ*msmeg_2196*
Colony morphology of (a) *M. smegmatis* mc^2155 and (b) Δ*msmeg_2196* mutant. Scale bar corresponds to 2 mm.

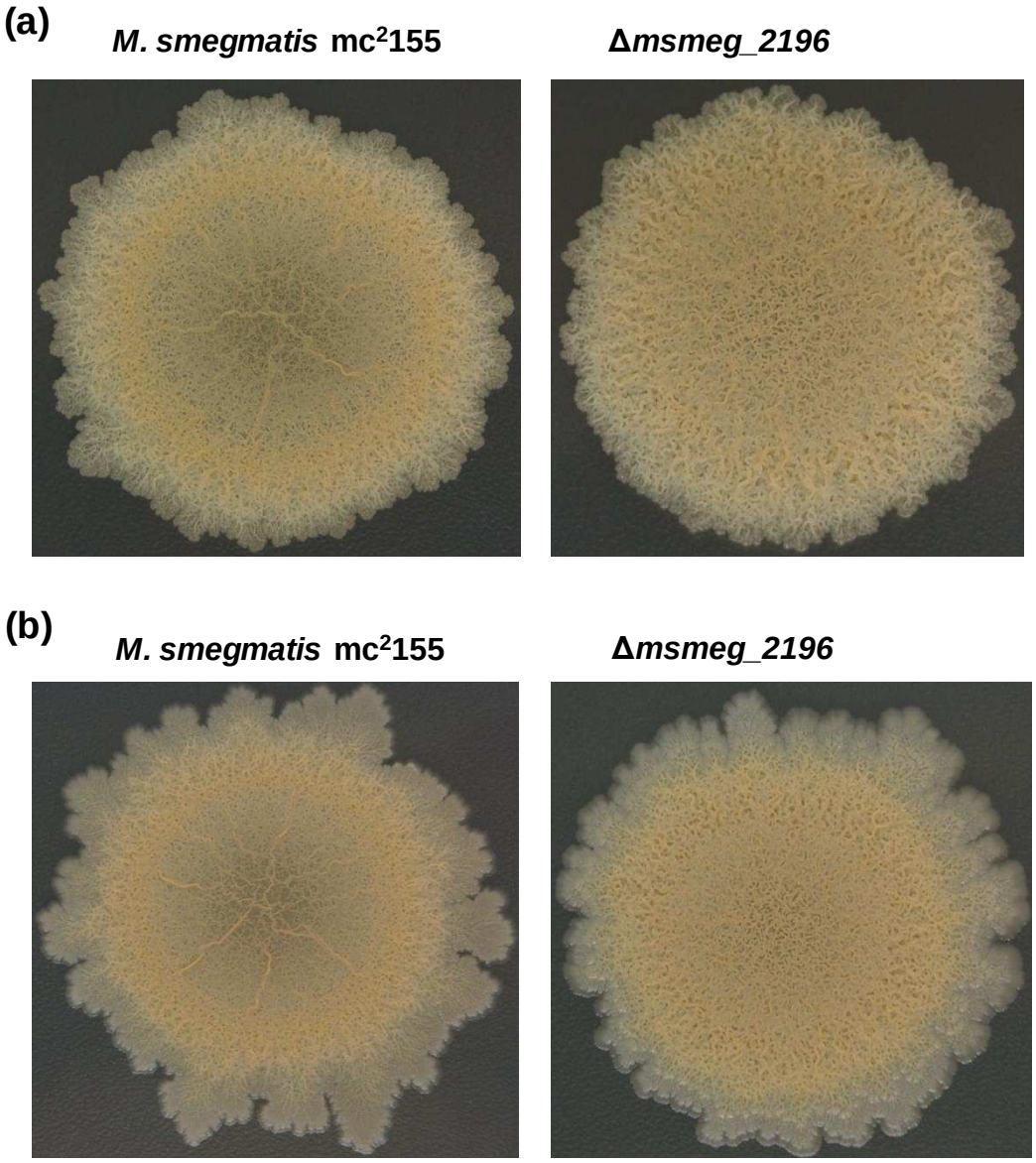


Figure 6-5. Colony morphology of *M. smegmatis* mc^{2}155 and Δ*msmeg_2196* after two and four weeks of incubation

6.3.2.3 Pellicle and biofilm formation

The pellicle and biofilm formation of *M. smegmatis* mc^{2}155 and the Δ*msmeg_2196* mutant was assessed as growth of bacterial cells on the surface of liquid culture and attachment of bacterial cells on the surface of the medium. The Δ*msmeg_2196* mutant had decreased pellicle formation compared to the *M. smegmatis* mc^{2}155 strain (Figure 6-6) grown in Sauton's medium. In addition, quantitative measurement of biofilm formation was performed by using crystal violet staining to attached cells on the surface of 96-well plates. Results show that the disruption of *msmeg_2196* had also significantly decreased biofilm formation in Sauton's medium (Figure 6-7), $p < 0.0001$. These results suggest that the *msmeg_2196* gene is related to pellicle and biofilm formation.

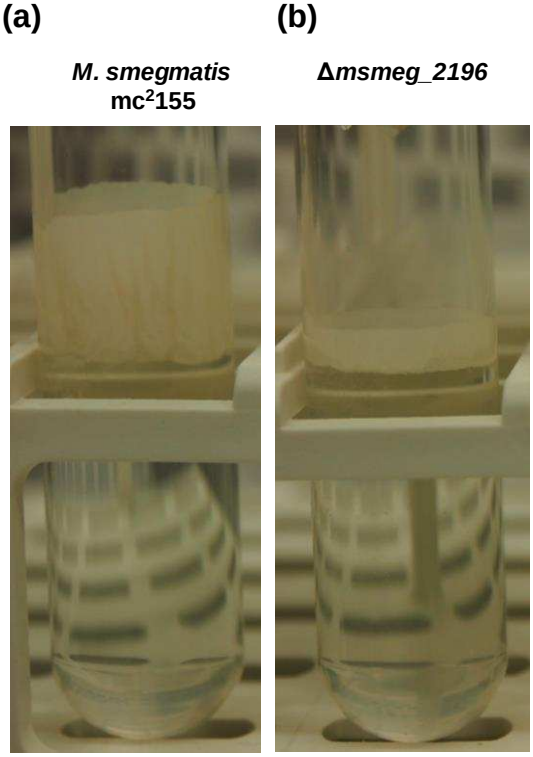


Figure 6-6. Pellicle formation of *M. smegmatis* mc^{2}155 and Δ*msmeg_2196*

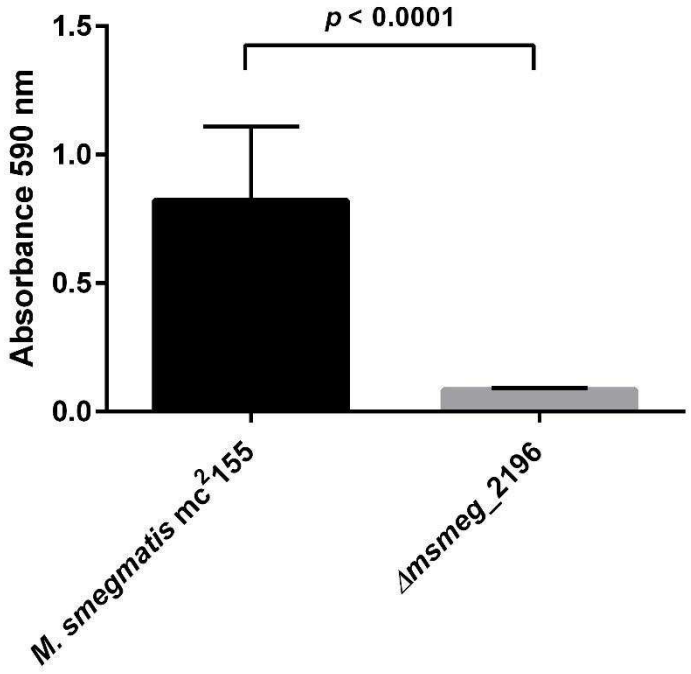


Figure 6-7. Biofilm formation of *M. smegmatis* mc^{2}155 and Δ*msmeg_2196*

6.3.2.4 Sliding motility

Sliding motility is a flagellum-independent spreading mechanism on the surface utilized by *Mycobacterium* spp. (14). This sliding motility was determined on the low viscosity 7H9 agar with (2% glucose) and without a carbon source (7). Deletion of *msmeg_2196* in *M. smegmatis* did not have an impact to sliding motility. The Δ*msmeg_2196* mutant was able to spread outwards from the inoculation site on the agar surface in a similar manner to *M. smegmatis* mc^{2}155 under both conditions, with and without carbon source (Figure 6-8). On non-carbon source plates, the spreading of colonies seemed to be a thinlayer while the spreading on rich medium had cell increased cell density (14).

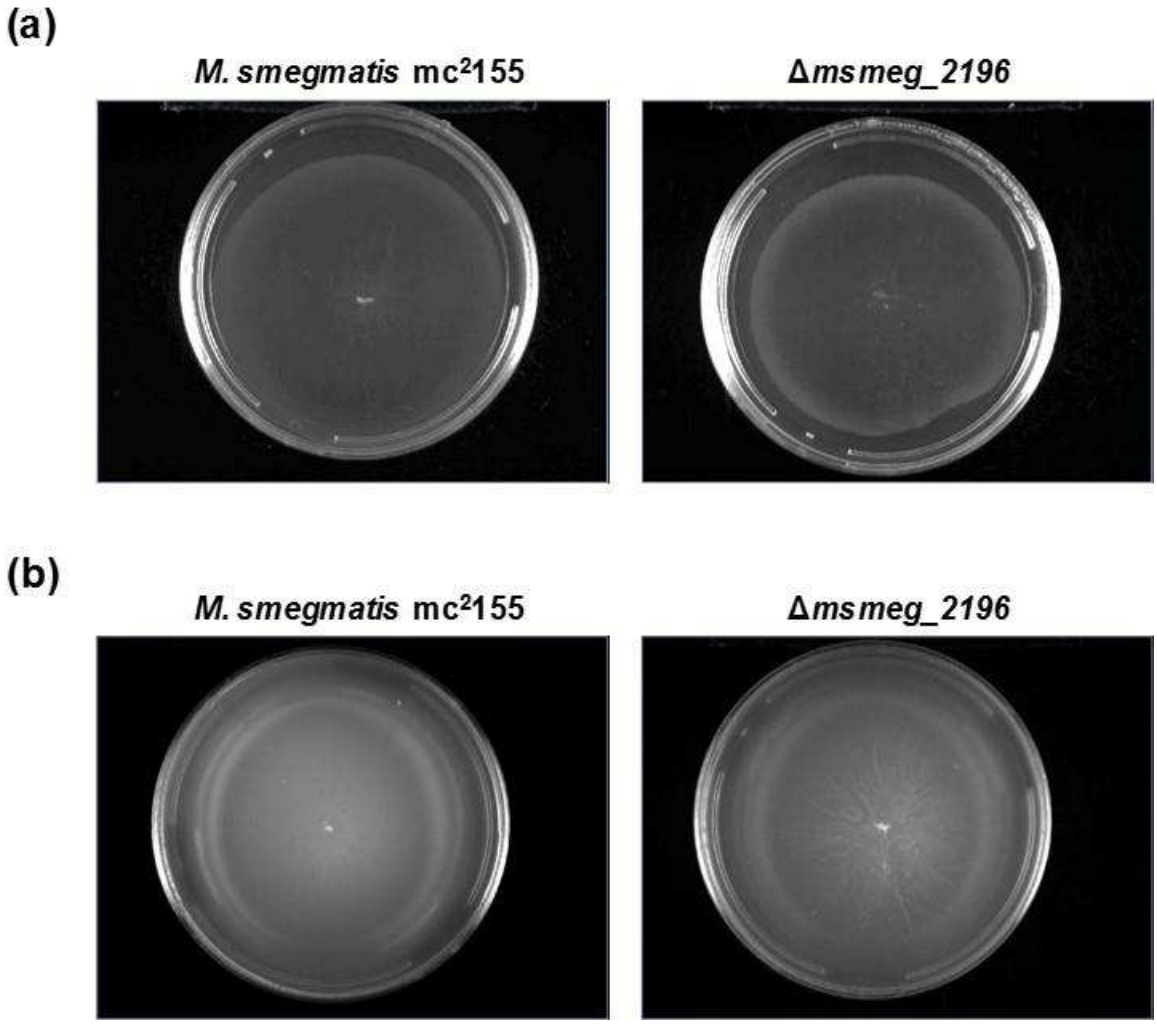


Figure 6-8. Sliding motility of *M. smegmatis* mc^{2}155 and Δ*msmeg_2196*
Sliding motility plates of *M. smegmatis* mc^{2}155 and Δ*msmeg_2196* on 7H9 agar without (a) and with carbon source (2% glucose) (b).

6.3.2.5 c-di-GMP production by *msmeg_2196* in *M. smegmatis*

The production of c-di-GMP in *M. smegmatis* mc^{2}155 was determined by LC-MS. The c-di-GMP was extracted from 1-3 mg of *M. smegmatis* cells grown from log-phase to stationary-phase by perchloric acid, as described in Chapter 3. Since MSMEG_2196 is the only DGC found in *M. smegmatis*, the introduction of the Δ*msmeg_2196* mutation was expected to abrogate all c-di-GMP production. c-di-GMP was not detected in both extracts from *M. smegmatis* mc^{2}155 and

the Δ*msmeg_2196* mutant at different growth phases. In addition, the c-di-GMP extraction was performed in *M. smegmatis* mc^2155 cells grown in 7H9 medium with a low level of carbon source, 0.02% glycerol, as described in a previous study (5) and expected to have increased c-di-GMP levels compared to *M. smegmatis* mc^2155 cells grown in 7H9 medium with 0.2% glycerol (5). However, there was no c-di-GMP detected in those samples (data not shown).

6.3.3 *msmeg_2196* expression *in vitro*

The expression of *msmeg_2196* in *M. smegmatis* grown *in vitro* was determined by real-time PCR. The expression level was determined based on the point (Cp) where the fluorescence of each transcript rises above the background fluorescence (15). Relative quantification of *msmeg_2196* gene expression at different time points was determined by a comparison to transcript levels at 24 h, and normalized to reference *sigA* gene transcripts. The quantitative data is reliable only when the efficiency of real-time PCR reactions for both the target gene and reference gene is optimal or identical. The real-time PCR efficiency from primers targeting *msmeg_2196* and *sigA* was calculated from a slope of linear regression obtained from the LightCycler software according to the equation $E = 10^{(-1/slope)}$ (16). The efficiency of the real-time PCR reaction from *msmeg_2196* and *sigA* is 90% and 85%, respectively (Figure 6-9). No primer-dimers were observed from melting curve analysis. The relative quantification of *msmeg_2196* expression at different time points was calculated using the $2^{-\Delta\Delta Cp}$ method (16, 17) by comparison to *msmeg_2196* transcript at 24 h and normalized by *sigA* transcript (Figure 6-10). From relative quantification, *msmeg_2196* has highest expression level at 24 h or early log-phase of growth profile compared to other time points. Expression was decreased at 48 h, 72 h, and 96 h.

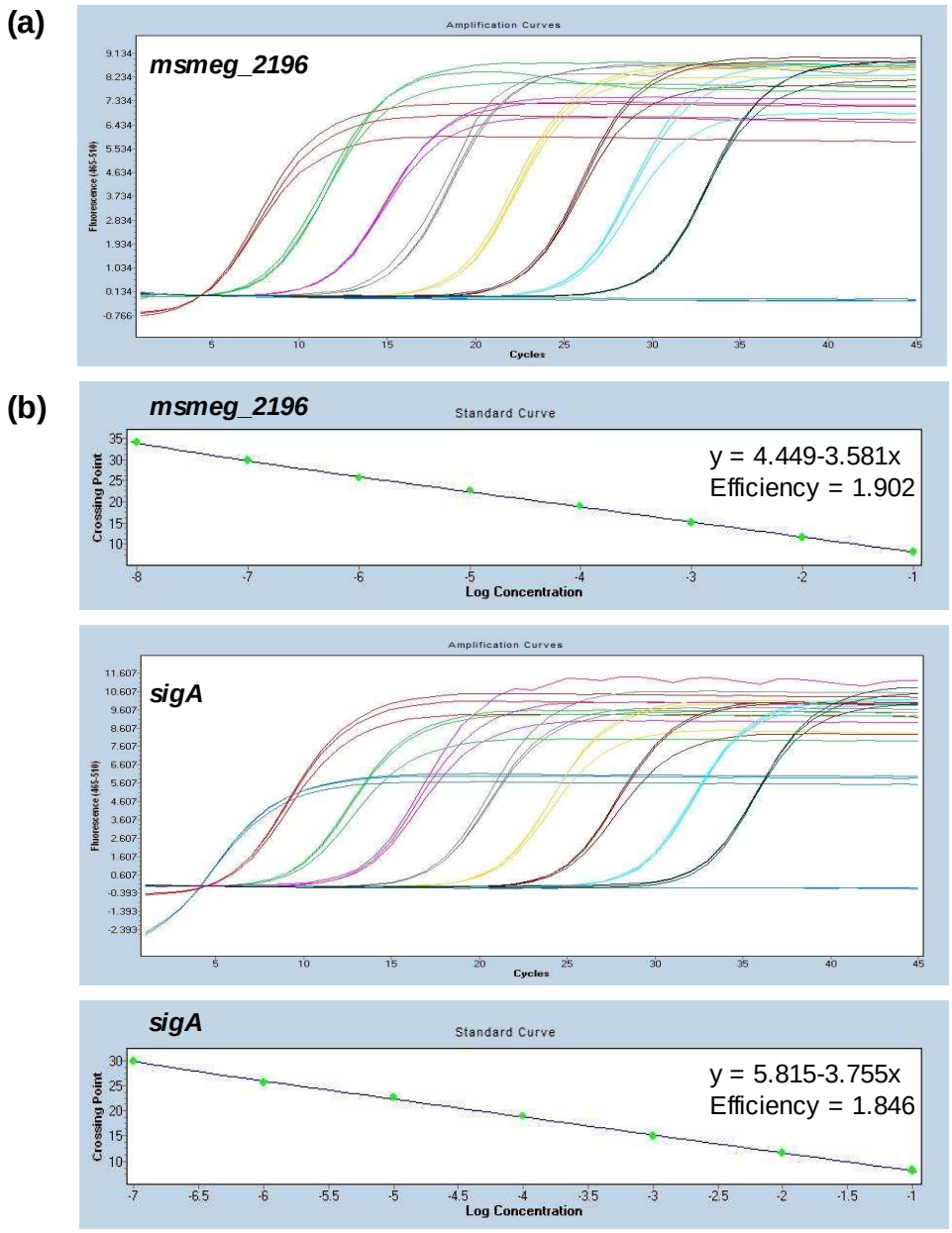


Figure 6-9. Real-time PCR efficiency
The amplification curves of *msmeg_2196* (a) and *sigA* (c) exhibited cycles where fluorescence rises above the background (Cp). The calibration curve of *msmeg_2196* (b) and *sigA* (d) showed the linear equation and efficiency of real-time PCR reaction.

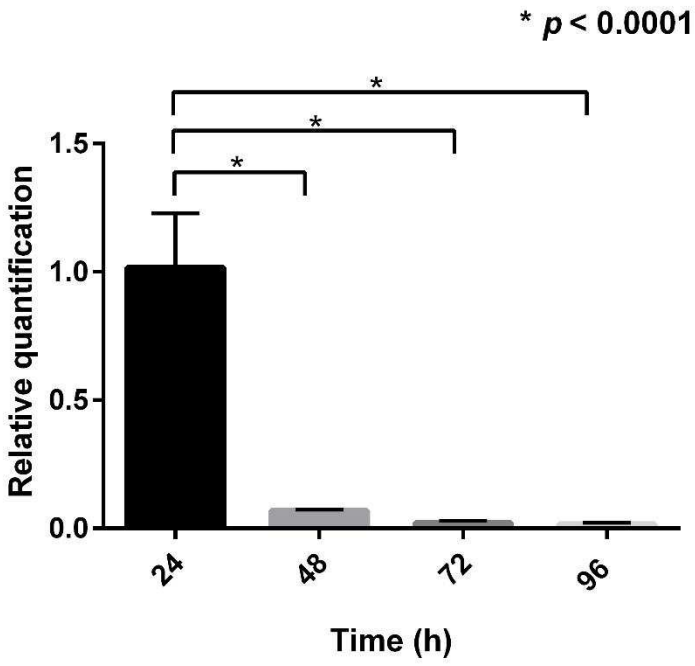


Figure 6-10. Relative quantification of *msmeg_2196* expression
Bar graphs showing relative quantification of *msmeg_2196* expression at 24 h, 48 h, 72 h, and 96 h. * $p < 0.0001$.

6.3.4 Recombinant overexpression of *dgcA* and *dgcB* in *M. smegmatis*

The production of recombinant protein was observed in the expression of *dgcA* and *dgcB* in *M. smegmatis* using a *hsp60* promoter. Phenotypic characterizations of recombinant overexpression of *dgcA* and *dgcB* in *M. smegmatis* were preliminarily determined by colony morphology and sliding motility. The colony morphology was observed on 7H10 containing 0.5% glycerol. *M. smegmatis* mc^2155 and the Δ*msmeg_2196* mutant expressing *dgcA* (mc^2155::*dgcA* and Δ*msmeg_2196*::*dgcA*) and PAS-GGDEF *dgcB* (mc^2155::*dgcB* and Δ*msmeg_2196*::*dgcB*) exhibited altered colony morphology. The colony morphology of overexpressed *dgcA* and PAS-GGDEF *dgcB* strains had smaller, more prominent wrinkles on the surface, and an irregular shape compared to the vector control (pVV16) (Figure 6-11). It was also noted that the introduction of pVV16 into *M. smegmatis* and growth under selection with hygromycin decreased the colony size

compared to *M. smegmatis* mc^{2}155 and the Δ*msmeg_2196* mutant without the vector (Figure 6-4 and Figure 6-11).

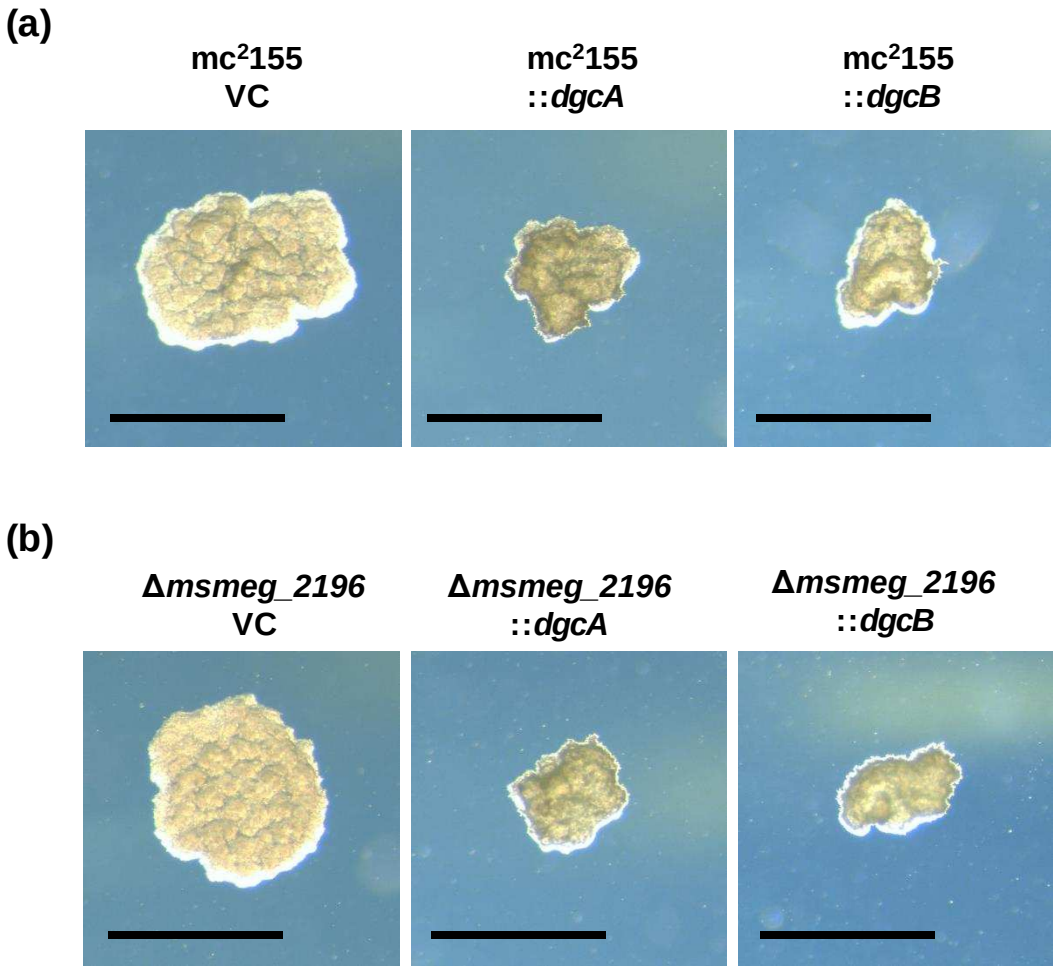


Figure 6-11. Colony morphology of *M. smegmatis* strains expressing *dgcA* and *dgcB*
M. smegmatis mc^{2}155 (a) and Δ*msmeg_2196* mutant (b) expressing *dgcA* and *dgcB* had an alteration of colony morphology compared to vector control (VC). Scale bar corresponds to 2 mm.

The sliding motility was assessed to determine the impact of *dgcA* and PAS-GGDEF *dgcB* expression in *M. smegmatis* mc^{2}155 and the Δ*msmeg_2196* mutant. Both *M. smegmatis* mc^{2}155 and the Δ*msmeg_2196* mutant expressing *dgcA* had a decreased sliding motility on the surface of sliding agar without a carbon source compared to vector control (Figure 6-12). On the sliding agar with 2% glucose, *M. smegmatis* strains expressing *dgcA* had decreased sliding motility with more

densely packed areas in the center surrounded with spreading cells (Figure 6-13). *M. smegmatis* strains expressing PAS-GGDEF *dgcB* had a similar pattern of sliding motility as the vector control on sliding agar with and without a carbon source (Figure 6-12 and 6-13).

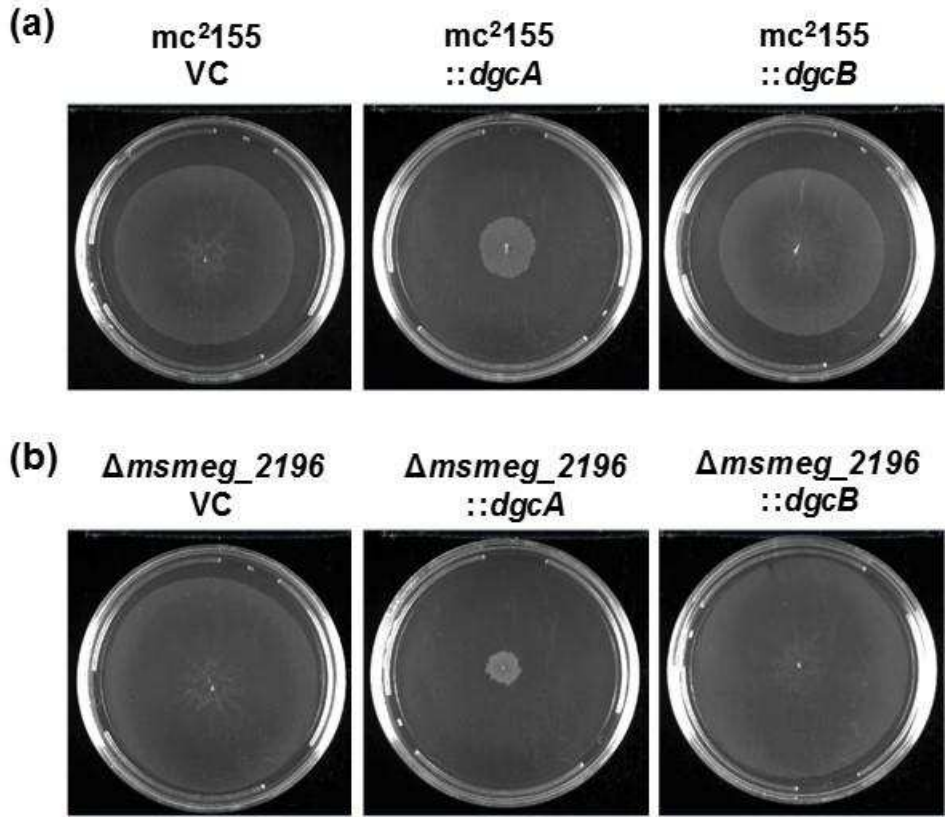


Figure 6-12. Sliding motility of *M. smegmatis* strains expressing *dgcA* and PAS-GGDEF *dgcB* on 7H9 agar without a carbon source
M. smegmatis mc^{2}155 (a) and Δ*msmeg_2196* mutant (b) expressing *dgcA*, PAS-GGDEF *dgcB*, and vector control (VC).

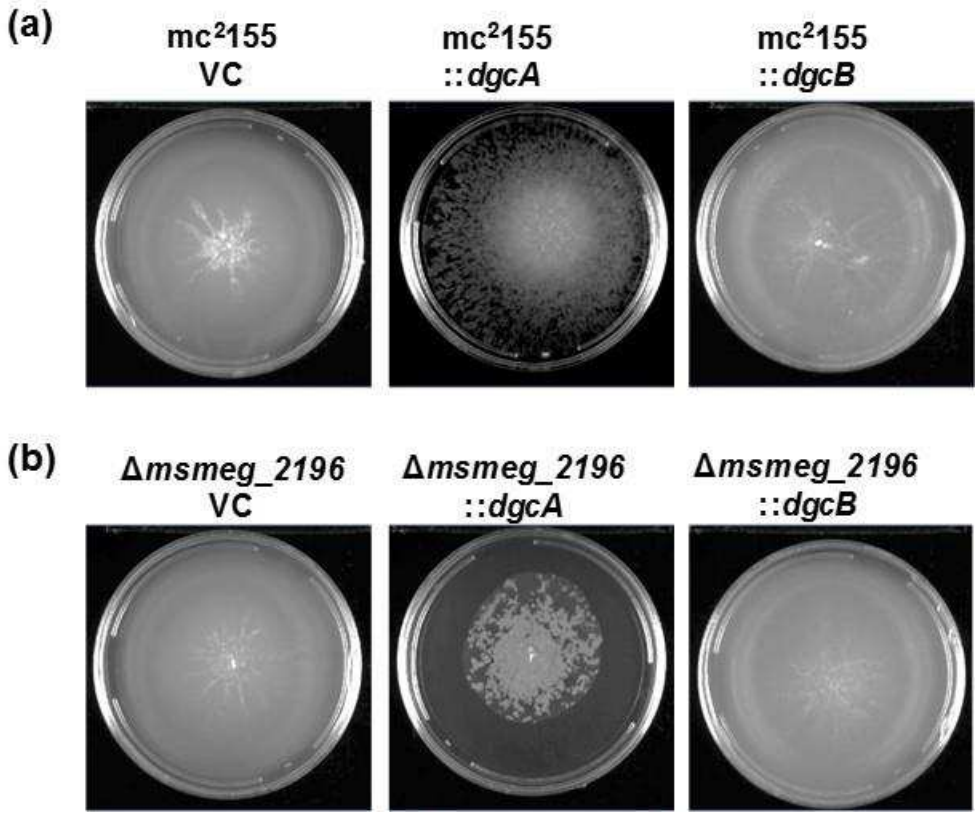


Figure 6-13. Sliding motility of *M. smegmatis* strains expressing *dgcA* and *dgcB* on 7H9 agar with 2% glucose
M. smegmatis mc^{2}155 (a) and Δ*msmeg_2196* mutant (b) expressing *dgcA*, PAS-GGDEF *dgcB*, and vector control (VC).

6.4 Discussion

The investigation of *M. smegmatis* as a model organism to study *M. leprae* was performed by phenotypic characterization and c-di-GMP detection by LC-MS. The disruption of *msmeg_2196*, the only GGDEF protein in *M. smegmatis*, was performed by gene replacement with a kanamycin resistance gene. The expression of *msmeg_2196* was determined by qRT-PCR and results show that *msmeg_2196* is expressed during growth *in vitro*. The highest expression level was observed at early log-phase of the growth curve and decreased after entering stationary phase.

Deletion of the *msmeg_2196* gene did not show significant effect on colony morphology, sliding motility, and growth profile of Δ*msmeg_2196* compared to *M. smegmatis* mc^{2}155. The

finding of unchanged colony morphology corresponded to previous studies that showed alterations in colony morphology only in overexpressed *msmeg_2196* in *M. smegmatis* mc^2155 after long periods of incubation (5, 7). However, the Δ*msmeg_2196* mutant did show significantly decreased pellicle and biofilm formation in modified Sauton's medium, indicating an alteration of lipid composition in cell walls (7, 18, 19). The complementation of Δ*msmeg_2196* mutant with *msmeg_2196* gene is required to determine the restoration of biofilm and pellicle formation in complemented strain.

Biofilm formation of mycobacteria is not related to the production of exopolysaccharides (20). In *M. smegmatis*, the initial attachment during biofilm formation and sliding motility involves glycopeptidolipids (GPLs) (14, 18), present in the outermost layer of cell envelopes (21). In addition, mycobacterial biofilm maturation is associated with mycolyl diacylglycerol production regulated by Lsr2 (22) and free mycolic acids (19, 23), which is regulated by GroEL1 chaperone protein (19) and MmpL11 mycolic acid-containing lipid transporter (24). Previous studies have found that deletion of *msmeg_2196* reduced GPLs and polar lipid production (7).

Previous studies have shown an alteration of sliding motility in Δ*msmeg_2196* mutant on agar without a carbon source (5, 7). However, we did not observe an alteration of sliding motility in Δ*msmeg_2196* mutant. There are many factors that could impact the sliding motility, which are concentration and wetness of agar surface, nutrient components of medium, and humidity during culture (14). Further investigation is required in order to elucidate the absence of change in sliding motility of the Δ*msmeg_2196* mutant.

The LC-MS method could not detect c-di-GMP in extracts from up to three grams of *M. smegmatis* mc^2155 cells grown in different medium conditions and time points. Since the bacterium has only one GGDEF protein, *msmeg_2196*, it consequently had low level of c-di-GMP

in cells as observed in previous studies (5, 25) and this c-di-GMP level could be lower than limit of detection of the LC-MS method used. Therefore, a technique with higher sensitivity is required.

The overexpression of both *dgcA* and *dgcB* have shown alteration in phenotypes. The change in colony morphology of *M. smegmatis* expressing *dgcA* and *dgcB* and sliding motility of *M. smegmatis* expressing *dgcA* could be related to the impacts of protein interactions or c-di-GMP to lipid production in *M. smegmatis*. Further investigation in biofilm and pellicle formation would provide more clues about the impact of *dgcA* and *dgcB* on *M. smegmatis* lipid production. In addition, inactivation of *dgcA* and *dgcB* activities by GGDEF motif deletion would help to reveal that those phenotypic changes truly resulted from c-di-GMP production if colony morphology and sliding motility of recombinant strains are similar to *M. smegmatis* mc^2155. Finally, increased c-di-GMP level in *M. smegmatis* expressing *dgcA* and *dgcB* would significantly confirm that these phenotypic alterations in cells are truly caused by c-di-GMP.

In conclusion, *M. smegmatis* has potential to be a model organism for *M. leprae* c-di-GMP study. However, it needs further optimization for phenotypic studies which are growth conditions such as medium and temperature. c-di-GMP extraction and detection methods also need to be carefully optimized in order to have a better understanding of c-di-GMP roles in this organism, as well as roles of *M. leprae* genes encoding DGC.

REFERENCES

1. Stahl DA, Urbance JW. 1990. The Division between fast-growing and slow-growing species corresponds to natural relationships among the mycobacteria. *Journal of Bacteriology* 172:116-124.

2. Rogall T, Wolters J, Flohr T, Bottger EC. 1990. Towards a phylogeny and definition of species at the molecular level within the genus *Mycobacterium*. *Int J Syst Bacteriol* 40:323-330.

3. Reyrat JM, Kahn D. 2001. *Mycobacterium smegmatis*: an absurd model for tuberculosis? *Trends Microbiol* 9:472-474.

4. Snapper SB, Melton RE, Mustafa S, Kieser T, Jacobs WR, Jr. 1990. Isolation and characterization of efficient plasmid transformation mutants of *Mycobacterium smegmatis*. *Mol Microbiol* 4:1911-1919.

5. Bharati BK, Sharma IM, Kasetty S, Kumar M, Mukherjee R, Chatterji D. 2012. A full-length bifunctional protein involved in c-di-GMP turnover is required for long-term survival under nutrient starvation in *Mycobacterium smegmatis*. *Microbiology* 158:1415-1427.

6. Sharma IM, Prakash S, Dhanaraman T, Chatterji D. 2014. Characterization of a dual-active enzyme, DcpA, involved in cyclic diguanosine monophosphate turnover in *Mycobacterium smegmatis*. *Microbiology* 160:2304-2318.

7. Gupta KR, Kasetty S, Chatterji D. 2015. Novel functions of (p)ppGpp and cyclic di-GMP in mycobacterial physiology revealed by phenotype microarray analysis of wild-type and isogenic strains of *Mycobacterium smegmatis*. *Appl Environ Microbiol* 81:2571-2578.

8. Gupta KR, Baloni P, Indi SS, Chatterji D. 2016. Regulation of growth, cell shape, cell division, and gene expression by second messengers (p)ppGpp and cyclic di-GMP in *Mycobacterium smegmatis*. *J Bacteriol* 198:1414-1422.

9. Li W, He ZG. 2012. LtmA, a novel cyclic di-GMP-responsive activator, broadly regulates the expression of lipid transport and metabolism genes in *Mycobacterium smegmatis*. *Nucleic Acids Res* 40:11292-11307.

10. Kendall SL, Frita R. 2008. Construction of targeted mycobacterial mutants by homologous recombination p297-310, T Mycobacteria Protocols, vol 465.

11. Pelicic V, Jackson M, Reyrat JM, Jacobs WR, Jr., Gicquel B, Guilhot C. 1997. Efficient allelic exchange and transposon mutagenesis in *Mycobacterium tuberculosis*. *Proc Natl Acad Sci U S A* 94:10955-10960.

12. O'Toole GA. 2011. Microtiter dish biofilm formation assay. *J Vis Exp* doi:10.3791/2437:e2437.

13. Flores-Valdez MA, de Jesus Aceves-Sanchez M, Pedroza-Roldan C, Vega-Dominguez PJ, Prado-Montes de Oca E, Bravo-Madrigal J, Laval F, Daffe M, Koestler B, Waters CM. 2015. The cyclic di-GMP phosphodiesterase gene Rv1357c/BCG1419c affects BCG pellicle production and *in vivo* maintenance. *IUBMB Life* 67:129-138.

14. Martinez A, Torello S, Kolter R. 1999. Sliding motility in mycobacteria. *J Bacteriol* 181:7331-7338.

15. Wong ML, Medrano JF. 2005. Real-time PCR for mRNA quantitation. *Biotechniques* 39:75-85.

16. Pfaffl MW. 2001. A new mathematical model for relative quantification in real-time RT–PCR. *Nucleic Acids Research* 29:e45.

17. Livak KJ, Schmittgen TD. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. *Methods* 25:402-408.

18. Recht J, Martinez A, Torello S, Kolter R. 2000. Genetic analysis of sliding motility in *Mycobacterium smegmatis*. *J Bacteriol* 182:4348-4351.

19. Ojha A, Anand M, Bhatt A, Kremer L, Jacobs WR, Jr., Hatfull GF. 2005. GroEL1: a dedicated chaperone involved in mycolic acid biosynthesis during biofilm formation in mycobacteria. *Cell* 123:861-873.

20. Zambrano MM, Kolter R. 2005. Mycobacterial biofilms: A greasy way to hold it together. *Cell* 123:762-764.

21. Ortalo-Magne A, Lemassu A, Laneelle MA, Bardou F, Silve G, Gounon P, Marchal G, Daffe M. 1996. Identification of the surface-exposed lipids on the cell envelopes of *Mycobacterium tuberculosis* and other mycobacterial species. *J Bacteriol* 178:456-461.

22. Chen JM, German GJ, Alexander DC, Ren H, Tan T, Liu J. 2006. Roles of Lsr2 in colony morphology and biofilm formation of *Mycobacterium smegmatis*. *J Bacteriol* 188:633-641.

23. Ojha AK, Trivelli X, Guerardel Y, Kremer L, Hatfull GF. 2010. Enzymatic hydrolysis of trehalose dimycolate releases free mycolic acids during mycobacterial growth in biofilms. *J Biol Chem* 285:17380-17389.

24. Pacheco SA, Hsu FF, Powers KM, Purdy GE. 2013. MmpL11 protein transports mycolic acid-containing lipids to the mycobacterial cell wall and contributes to biofilm formation in *Mycobacterium smegmatis*. *J Biol Chem* 288:24213-24222.

25. Bharati BK, Swetha RK, Chatterji D. 2013. Identification and characterization of starvation induced msdgc-1 promoter involved in the c-di-GMP turnover. *Gene* 528:99-108.

CHAPTER 7 DISCUSSION AND CONCLUSIONS

M. leprae is an obligate intracellular pathogen and is uncultivable *in vitro*. Its unique genome with massive gene decay leads to the defects in several metabolic pathways compared to other mycobacteria. Consequently, *M. leprae* requires diverse resources from the host, and utilizes various mechanisms for survival in host cells. The study of *M. leprae* has been tremendously impaired since genetic manipulation cannot be performed with this bacterium. There are several aspects of *M. leprae* that require further investigation such as *M. leprae*–host interactions, disease pathogenesis related nerve damage, and transmission routes. The current treatments for leprosy have been used for decades and the development of drug resistance leads to difficulty in disease management and control. Better understanding of the host-bacterium interaction and also the discovery of new, specific *M. leprae* antigens or compounds would be valuable for the development of new drugs, diagnostic tools, and vaccines. After the whole genome of *M. leprae* was sequenced (1), 1,604 protein encoding genes were revealed. The production of some proteins have been verified by proteomic approaches (2-6). It is a challenging situation for researchers to investigate the molecular and cellular functions of these encoded proteins, including the protein interactions in the complicated physiology processes of *M. leprae*.

The study of cyclic nucleotides and dinucleotides in bacteria, including mycobacteria, is attractive and has been widely studied in various physiological aspects such as bacterial pathogenesis. Specifically, c-di-GMP was recognized in 1987 and several aspects such as signaling mechanisms, receptors, and biological roles for this molecule have been revealed in *P. aeruginosa* and other bacteria (7-13). The formation of c-di-GMP is catalyzed via DGC, an enzyme that utilizes GTP as the substrate and possesses a conserved GG(D/E)EF motif which is part of the

active site residue in GGDEF domain (14, 15). The degradation of c-di-GMP is mediated by PDE proteins that possess EAL domains which are responsible for the depletion of c-di-GMP and conversion to a linear nucleotide 5′phosphoguanylyl 3′5′ guanosine (pGpG) (16), and proteins containing HD-GYP domains are also able to hydrolyze c-di-GMP to two molecules of guanosine monophosphate (GMP) (17). Nevertheless, the study of c-di-GMP in mycobacteria is still limited, especially for *M. leprae* for which no c-di-GMP studies have been reported. This study is significant since it is the first report that *M. leprae* has an ability to produce c-di-GMP.

M. leprae has the potential to produce and degrade c-di-GMP because it harbors two GGDEF containing proteins (DgcA and DgcB), one hybrid GGDEF-EAL containing protein (ML1750c), and one single EAL containing protein (ML1752c). The ML1750c is a hybrid DGC-PDE protein possessing an N-terminal GAF sensor domain as well as GGDEF and EAL domains. DgcA and DgcB are not encoded by *M. smegmatis* or *M. tuberculosis*. DgcA possesses three sequential PAS signaling domains N-terminal to the GGDEF domain. Two DgcA heme-binding sites in PAS domains were identified. DgcB harbors a single N-terminal PAS sensor domain with a heme-binding site linked to a GGDEF domain and ten C-terminal transmembrane domains. Comparison of the genomes of the two primary mycobacterial pathogens, *M. leprae* and *M. tuberculosis*, reveals that *M. leprae* has a significantly smaller genome, a relatively large number of pseudogenes, fewer functional proteins, and fewer transcription factors (1). Consequently, *M. leprae* is refractory to *in vitro* growth and has evolved into an obligate intracellular pathogen. Given the narrow biological niche of *M. leprae*, it is intriguing that this pathogen harbors three coding sequences for known or predicted DGCs (*dgcA*, *dgcB* and *ml1750*c), while *M. tuberculosis* has only one protein, Rv1354c a homologue of ML1750c. A coding sequence for a ML1750c homolouge is also the only DGC found in the genome of *M. smegmatis* (*msmeg_2196*) (18), a

nonpathogenic saprophyte commonly used as a model to define gene function of mycobacterial pathogens (19, 20). Rv1354c and MSMEG_2196 have been confirmed to have DGC and PDE activity, and associated phenotypes (18, 21-24).

Although phenotypic characterization related to c-di-GMP has been conducted in *M. smegmatis* (18, 21, 22, 25), the low level of intracellular c-di-GMP is a major impediment for detection. From our study, the production of recombinant protein was observed with the expression of *dgcA* in *M. smegmatis* using a *hsp60* promoter; howerver the detection of c-di-GMP production was not reproducible with a reasonable number of cells. As described in the published literature, detection and quantification of c-di-GMP in *M. tuberculosis* or *M. smegmatis* requires large amount of cells. Up to 3 g wet weight (18, 23, 24) is required and does not provide a robust system to correlate c-di-GMP levels with observed phenotypes. Based on the limit-of-detection of the mass spectrometry method developed in this study, this would equate to less than 10 molecules of c-di-GMP per cell. This is unusual with what is known about the cellular concentrations required for functional signaling and response for systems that use c-di-GMP. The methods that have been employed for c-di-GMP extraction and quantification have significant deficiencies in the extended length of time required during processing, which results in the inability to quench cellular metabolism and rapid degradation of c-di-GMP.

The overexpression of *dgcA* and *dgcB* in *M. smegmatis* provides evidences that these proteins may be involved in lipid metabolism. The *dgcA* and *dgcB* expression in *M. smegmatis* changed the colony morphology or sliding motility of *M. smegmatis* which is related to lipid components in cell envelope (26, 27). A *msmeg_2196* knockout strain had reduced biofilm and pellicle formation, whereas the sliding motility did not changed. The reduction of sliding motility in *M. smegmatis* strains expressing *dgcA* and also the unchanged sliding motility in *M. smegmatis*

Δ*msmeg_2196* is correlated to physiological role of c-di-GMP in other bacteria. Specifically, c-di-GMP is able to inhibit several motility types (28-31). The determination of the lipids alterated by the overexpression of *dgcA* and *dgcB* in *M. smegmatis* would be valuable to clearly define the function of these proteins. From a previous study in *M. bovis* BCG Pasteur 1173P2, it was demonstrated that c-di-GMP was able to alter lipid production and biofilm formation (32). Thus, the determination of the impact of DgcA and DgcB on biofilm formation in *M. smegmatis* would be advantageous. The knowledge of a connection to biofilm formation in *M. smegmatis* could provide an implication of their involvement in *M. leprae* pathogenesis since biofilm formation has been associated with bacterial adaptation and survival under harsh environments such as starvation and the presence of antibiotics (33-35). However, the study of *M. leprae* biofilm formation has not been reported.

Alternative model organisms, which are *P. aeruginosa* and *E. coli*, have been utilized and provide reliable and reproducible results. *P. aeruginosa* and *E. coli* were used as the heterologous expression host for *dgcA* and *dgcB* in order to accurately detect c-di-GMP production because *P. aeruginosa* and *E. coli* are well-established model organisms for the study of c-di-GMP-mediated phenotypes. The selection of *P. aeruginosa* and *E. coli* was also based on the facile and well-characterized systems to assess c-di-GMP-mediated phenotypes, and for established methods established for c-di-GMP extraction and quantification. The study of another GGDEF containing protein, ML1750c, is more complicated because GGDEF domain is coupled with an EAL domain, which can degrade c-di-GMP. The elucidation of ML1750c functions could be accomplished by applying a molecular genetic approach to construct inactive DGC or PDE activity.

The investigation *M. leprae* c-di-GMP receptors or ligands is important for c-di-GMP signaling in *M. leprae*. c-di-GMP interacts with several classes of receptors. These cellular

effectors can be transcriptional factors (30, 36), proteins in the PilZ domain family (37), degenerate GGDEF or EAL domains (38, 39), riboswitches (40) and the I-site (14, 41). Nevertheless, these receptors could not be easily predicted by bioinformatics analysis (42). For example, the study of c-di-GMP receptor (LtmA) in *M. smegamtis* was performed by screening transcription factor library and investigating individual protein-c-di-GMP interaction by a cross-linking assay, and surface plasmon resonance (SPR) was used to confirm LtmA as a c-di-GMP receptor (23). An alternative approach that could be initially used to identify *M. leprae* c-di-GMP receptors is bioinformatics analysis such as a Blast analysis against well-described c-di-GMP receptors. This method could be performed to identify potential *M. leprae* proteins that may function as c-di-GMP receptors in *M. leprae*. The binding of c-di-GMP to those predicted receptors then could be studied by protein-ligand interation assays such as equilibrium dialysis, SPR, and differential radial capillary action of ligand assay (DRaCALA).

The conformation rearrangements of active site in GGDEF domain is mediated by the response to environamental stimuli perceived by sensory domains (42). DgcA has three PAS domains which may be involved in a complex signaling mechanism and DgcB harbors only one PAS domain. NCBI database shows that DgcA contains two heme-binding sites in PAS domains and DgcB harbors one heme-binding site in its PAS sensor domain. Thus, these two proteins have potential to bind heme and sense oxygen (43). To identify signaling molecules binding to sensory domains in GGDEF-containing proteins; heterologous studies could be performed to determine c-di-GMP production under various conditions such as aerobic, microaerophilic and anaerobic conditions (32). Further studies such as protein crystallography, would provide more information about protein structure, interaction with other molecules, protein functions, prediction of protein localization, and also signaling molecules binding to sensory domains (42, 44, 45).

Gene expression and protein production studies have exhibited the possibility that *M. leprae* proteins synthesizing c-di-GMP could be active and functioning during infection. All *M. leprae* genes encoding GGDEF-containing proteins are expressed during infection and two of the proteins, ML1750c and DgcA were detected in the whole cell lysate of *M. leprae*. These observations provide support for the hypothesis that *M. leprae* c-di-GMP might be involved disease pathogenesis, especially c-di-GMP produced from DgcA, previously described as a T-cell antigen in the early phase of leprosy (46).

To fully define the function and significance of DGC activity in *M. leprae*, an alternative model system that allows for genetic analyses of individual DGCs is required. Toward this goal, all of the known and predicted DGCs of *M. leprae* are encoded in the genomes of *Mycobacterium lepromatosis* (another bacterium restricted to *in vivo* growth) and *Mycobacterium haemophilum* (an opportunistic pathogen). Homologs of DgcA and ML1750c are also encoded in several environmental mycobacteria (*Mycobacterium rhodesiae* NBB3, *Mycobacterium chubuense*, *Mycobacterium chlorophenolicum* and *Mycobacterium rufum*). Evolutionary reduction of the number c-di-GMP signaling pathways is generally believed to be inversely correlated with a bacterium's need and ability to adapt to rapidly changing environmental conditions (42). Therefore, the conservation of DGC genes in *M. leprae* and the presence of these genes in opportunistic pathogens and several environmental *Mycobacterium* spp. is indicative of an essential role for DGCs in the signaling response required for *M. leprae* to survive as an obligate intracellular pathogen. Interestingly like *M. leprae*, *M. haemophilum* is a pathogen of the skin and displays optimal growth at 30°C (47). Recently, it has been reported that *M. haemophilum* can be genetically manipulated to express foreign genes (48), and efforts are now under way to use this *Mycobacterium* sp. as a model to further define the physiological functions and signaling events

associated with DgcA and other DGCs of *M. leprae*. Direct evidence of c-di-GMP production by *M. leprae* would be a significant step for future investigations that target c-di-GMP mechanisms such as its environmental signaling, receptors, and the biological roles. However, for this to occur, new tools that allow for the more sensitive detection of c-di-GMP in tissue samples are required and would need to be integrated with analyses in a model organism such as *M. haemophilum*.

REFERENCES

1. Cole ST, Eiglmeier K, Parkhill J, James KD, Thomson NR, Wheeler PR, Honore N, Garnier T, Churcher C, Harris D, Mungall K, Basham D, Brown D, Chillingworth T, Connor R, Davies RM, Devlin K, Duthoy S, Feltwell T, Fraser A, Hamlin N, Holroyd S, Hornsby T, Jagels K, Lacroix C, Maclean J, Moule S, Murphy L, Oliver K, Quail MA, Rajandream MA, Rutherford KM, Rutter S, Seeger K, Simon S, Simmonds M, Skelton J, Squares R, Squares S, Stevens K, Taylor K, Whitehead S, Woodward JR, Barrell BG. 2001. Massive gene decay in the leprosy bacillus. *Nature* 409:1007-1011.

2. Pessolani MC, Brennan PJ. 1996. Molecular definition and identification of new proteins of *Mycobacterium leprae*. *Infect Immun* 64:5425-5427.

3. Marques MA, Chitale S, Brennan PJ, Pessolani MC. 1998. Mapping and identification of the major cell wall-associated components of *Mycobacterium leprae*. *Infect Immun* 66:2625-2631.

4. Marques MA, Espinosa BJ, Xavier da Silveira EK, Pessolani MC, Chapeaurouge A, Perales J, Dobos KM, Belisle JT, Spencer JS, Brennan PJ. 2004. Continued proteomic analysis of *Mycobacterium leprae* subcellular fractions. *Proteomics* 4:2942-2953.

5. Marques MA, Neves-Ferreira AG, da Silveira EK, Valente RH, Chapeaurouge A, Perales J, da Silva Bernardes R, Dobos KM, Spencer JS, Brennan PJ, Pessolani MC. 2008. Deciphering the proteomic profile of *Mycobacterium leprae* cell envelope. *Proteomics* 8:2477-2491.

6. de Souza GA, Softeland T, Koehler CJ, Thiede B, Wiker HG. 2009. Validating divergent ORF annotation of the *Mycobacterium leprae* genome through a full translation data set and peptide identification by tandem mass spectrometry. *Proteomics* 9:3233-3243.

7. Abel S, Chien P, Wassmann P, Schirmer T, Kaever V, Laub MT, Baker TA, Jenal U. 2011. Regulatory cohesion of cell cycle and cell differentiation through inter linked phosphorylation and second messenger networks. *Molecular Cell* 43:550-560.

8. Irie Y, Borlee BR, O'Connor JR, Hill PJ, Harwood CS, Wozniak DJ, Parsek MR. 2012. Self-produced exopolysaccharide is a signal that stimulates biofilm formation in *Pseudomonas aeruginosa*. *Proc Natl Acad Sci U S A* 109:20632-20636.

9. Kulasekara H, Lee V, Brencic A, Liberati N, Urbach J, Miyata S, Lee DG, Neely AN, Hyodo M, Hayakawa Y, Ausubel FM, Lory S. 2006. Analysis of *Pseudomonas aeruginosa* diguanylate cyclases and phosphodiesterases reveals a role for bis-(3'-5')-cyclic-GMP in virulence. *Proc Natl Acad Sci U S A* 103:2839-2844.

10. Merritt JH, Brothers KM, Kuchma SL, O'Toole GA. 2007. SadC reciprocally influences biofilm formation and swarming motility via modulation of exopolysaccharide production and flagellar function. *J Bacteriol* 189:8154-8164.

11. Simm R, Morr M, Kader A, Nimtz M, Romling U. 2004. GGDEF and EAL domains inversely regulate cyclic di-GMP levels and transition from sessility to motility. *Mol Microbiol* 53:1123-1134.

12. Paul K, Nieto V, Carlquist WC, Blair DF, Harshey RM. 2010. The c-di-GMP binding protein YcgR controls flagellar motor direction and speed to affect chemotaxis by a "backstop brake" mechanism. *Mol Cell* 38:128-139.

13. Tischler AD, Camilli A. 2004. Cyclic diguanylate (c-di-GMP) regulates *Vibrio cholerae* biofilm formation. *Mol Microbiol* 53:857-869.

14. Chan C, Paul R, Samoray D, Amiot NC, Giese B, Jenal U, Schirmer T. 2004. Structural basis of activity and allosteric control of diguanylate cyclase. *Proc Natl Acad Sci U S A* 101:17084-17089.

15. Wassmann P, Chan C, Paul R, Beck A, Heerklotz H, Jenal U, Schirmer T. 2007. Structure of $BeF3^-$ -modified response regulator PleD: implications for diguanylate cyclase activation, catalysis, and feedback inhibition. *Structure* 15:915-927.

16. Schmidt AJ, Ryjenkov DA, Gomelsky M. 2005. The ubiquitous protein domain EAL is a cyclic diguanylate-specific phosphodiesterase: Enzymatically active and inactive EAL domains. *Journal of Bacteriology* 187:4774-4781.

17. Ryan RP, Fouhy Y, Lucey JF, Crossman LC, Spiro S, He YW, Zhang LH, Heeb S, Camara M, Williams P, Dow JM. 2006. Cell-cell signaling in *Xanthomonas campestris* involves an HD-GYP domain protein that functions in cyclic di-GMP turnover. *Proc Natl Acad Sci U S A* 103:6712-6717.

18. Bharati BK, Sharma IM, Kasetty S, Kumar M, Mukherjee R, Chatterji D. 2012. A full-length bifunctional protein involved in c-di-GMP turnover is required for long-term survival under nutrient starvation in *Mycobacterium smegmatis*. *Microbiology* 158:1415-1427.

19. Reyrat JM, Kahn D. 2001. *Mycobacterium smegmatis*: an absurd model for tuberculosis? *Trends Microbiol* 9:472-474.

20. Singh AK, Reyrat JM. 2009. Laboratory maintenance of *Mycobacterium smegmatis*. *Curr Protoc Microbiol* Chapter 10:Unit10C 11.

21. Gupta KR, Baloni P, Indi SS, Chatterji D. 2016. Regulation of growth, cell shape, cell division, and gene expression by second messengers (p)ppGpp and cyclic di-GMP in *Mycobacterium smegmatis*. *J Bacteriol* 198:1414-1422.

22. Gupta KR, Kasetty S, Chatterji D. 2015. Novel functions of (p)ppGpp and cyclic di-GMP in mycobacterial physiology revealed by phenotype microarray analysis of wild-type and isogenic strains of *Mycobacterium smegmatis*. *Appl Environ Microbiol* 81:2571-2578.

23. Li W, He ZG. 2012. LtmA, a novel cyclic di-GMP-responsive activator, broadly regulates the expression of lipid transport and metabolism genes in *Mycobacterium smegmatis*. *Nucleic Acids Res* 40:11292-11307.

24. Hong Y, Zhou X, Fang H, Yu D, Li C, Sun B. 2013. Cyclic di-GMP mediates *Mycobacterium tuberculosis* dormancy and pathogenecity. *Tuberculosis (Edinb)* 93:625-634.

25. Sharma IM, Prakash S, Dhanaraman T, Chatterji D. 2014. Characterization of a dual-active enzyme, DcpA, involved in cyclic diguanosine monophosphate turnover in *Mycobacterium smegmatis*. *Microbiology* 160:2304-2318.

26. Martinez A, Torello S, Kolter R. 1999. Sliding motility in mycobacteria. *J Bacteriol* 181:7331-7338.

27. Recht J, Martinez A, Torello S, Kolter R. 2000. Genetic analysis of sliding motility in *Mycobacterium smegmatis*. *J Bacteriol* 182:4348-4351.

28. Bordeleau E, Purcell EB, Lafontaine DA, Fortier LC, Tamayo R, Burrus V. 2015. Cyclic di-GMP riboswitch-regulated type IV pili contribute to aggregation of *Clostridium difficile*. *J Bacteriol* 197:819-832.

29. Purcell EB, McKee RW, McBride SM, Waters CM, Tamayo R. 2012. Cyclic diguanylate inversely regulates motility and aggregation in *Clostridium difficile*. *J Bacteriol* 194:3307-3316.

30. Krasteva PV, Fong JC, Shikuma NJ, Beyhan S, Navarro MV, Yildiz FH, Sondermann H. 2010. *Vibrio cholerae* VpsT regulates matrix production and motility by directly sensing cyclic di-GMP. *Science* 327:866-868.

31. Spurbeck RR, Tarrien RJ, Mobley HL. 2012. Enzymatically active and inactive phosphodiesterases and diguanylate cyclases are involved in regulation of motility or sessility in *Escherichia coli* CFT073. *MBio* 3.

32. Flores-Valdez MA, de Jesus Aceves-Sanchez M, Pedroza-Roldan C, Vega-Dominguez PJ, Prado-Montes de Oca E, Bravo-Madrigal J, Laval F, Daffe M, Koestler B, Waters CM. 2015. The cyclic di-GMP phosphodiesterase gene Rv1357c/BCG1419c affects BCG pellicle production and *in vivo* maintenance. *IUBMB Life* 67:129-138.

33. Teng R, Dick T. 2003. Isoniazid resistance of exponentially growing *Mycobacterium smegmatis* biofilm culture. *FEMS Microbiology Letters* 227:171-174.

34. Zambrano MM, Kolter R. 2005. Mycobacterial biofilms: A greasy way to hold it together. *Cell* 123:762-764.

35. Ojha AK, Baughn AD, Sambandan D, Hsu T, Trivelli X, Guerardel Y, Alahari A, Kremer L, Jacobs WR, Jr., Hatfull GF. 2008. Growth of *Mycobacterium tuberculosis*

biofilms containing free mycolic acids and harbouring drug-tolerant bacteria. *Mol Microbiol* 69:164-174.

36. Hickman JW, Harwood CS. 2008. Identification of FleQ from *Pseudomonas aeruginosa* as a c-di-GMP-responsive transcription factor. *Mol Microbiol* 69:376-389.

37. Amikam D, Galperin MY. 2006. PilZ domain is part of the bacterial c-di-GMP binding protein. *Bioinformatics* 22:3-6.

38. Jenal U, Malone J. 2006. Mechanisms of cyclic-di-GMP signaling in bacteria. *Annu Rev Genet* 40:385-407.

39. Navarro MV, De N, Bae N, Wang Q, Sondermann H. 2009. Structural analysis of the GGDEF-EAL domain-containing c-di-GMP receptor FimX. *Structure* 17:1104-1116.

40. Kellenberger CA, Wilson SC, Hickey SF, Gonzalez TL, Su Y, Hallberg ZF, Brewer TF, Iavarone AT, Carlson HK, Hsieh YF, Hammond MC. 2015. GEMM-I riboswitches from *Geobacter* sense the bacterial second messenger cyclic AMP-GMP. *Proc Natl Acad Sci U S A* 112:5383-5388.

41. Christen B, Christen M, Paul R, Schmid F, Folcher M, Jenoe P, Meuwly M, Jenal U. 2006. Allosteric control of cyclic di-GMP signaling. *J Biol Chem* 281:32015-32024.

42. Romling U, Galperin MY, Gomelsky M. 2013. Cyclic di-GMP: the first 25 years of a universal bacterial second messenger. *Microbiol Mol Biol Rev* 77:1-52.

43. Henry JT, Crosson S. 2011. Ligand-binding PAS domains in a genomic, cellular, and structural context. *Annu Rev Microbiol* 65:261-286.

44. Gilles-Gonzalez MA, Gonzalez G. 2004. Signal transduction by heme-containing PAS-domain proteins. *J Appl Physiol (1985)* 96:774-783.

45. Sondermann H, Shikuma NJ, Yildiz FH. 2012. You've come a long way: c-di-GMP signaling. *Curr Opin Microbiol* 15:140-146.

46. Spencer JS, Dockrell HM, Kim HJ, Marques MA, Williams DL, Martins MV, Martins ML, Lima MC, Sarno EN, Pereira GM, Matos H, Fonseca LS, Sampaio EP, Ottenhoff TH, Geluk A, Cho SN, Stoker NG, Cole ST, Brennan PJ, Pessolani MC. 2005. Identification of specific proteins and peptides in *Mycobacterium leprae* suitable for the selective diagnosis of leprosy. *J Immunol* 175:7930-7938.

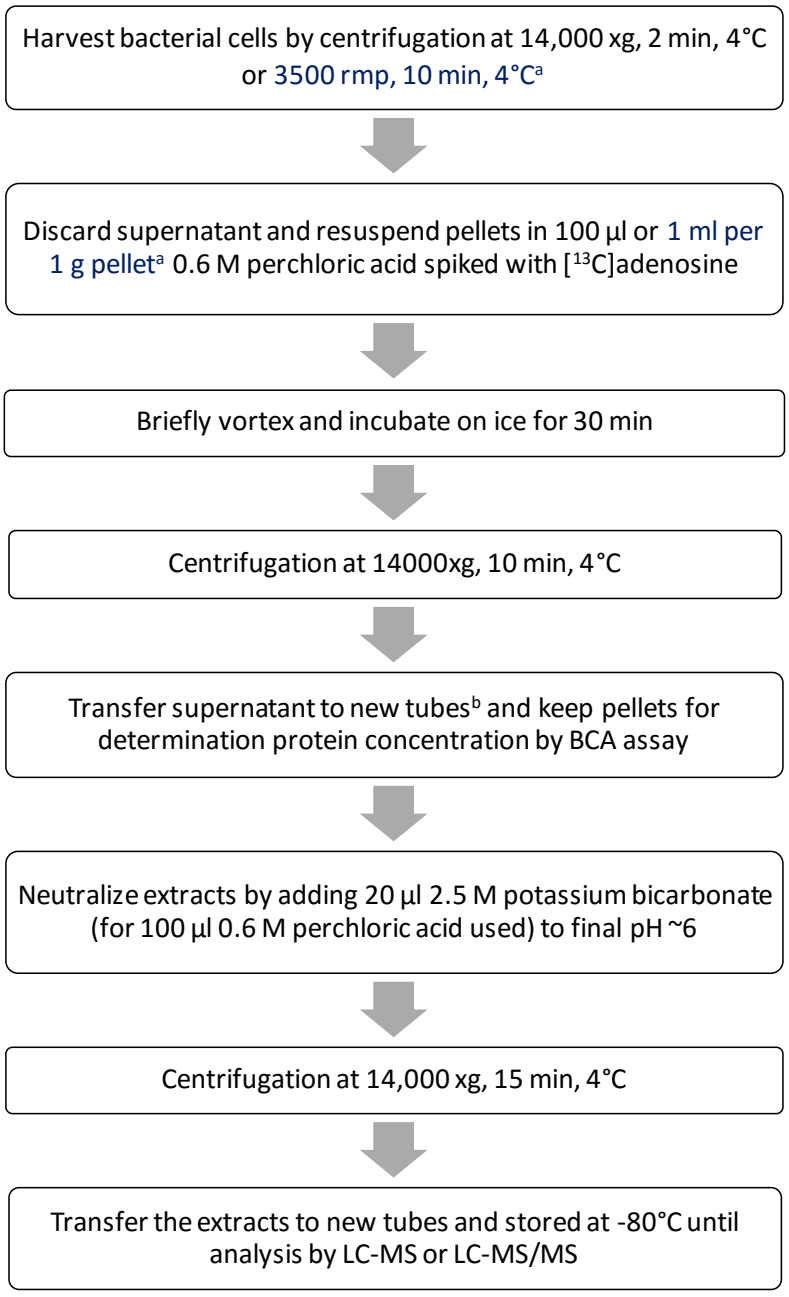
47. Sompolinsky D, Lagziel A, Naveh D, Yankilevitz T. 1978. *Mycobacterium haemophilum* sp. nov., a new pathogen of humans. *International Journal of Systematic and Evolutionary Microbiology* 28:67-75.

48. Tufariello JM, Kerantzas CA, Vilcheze C, Calder RB, Nordberg EK, Fischer JA, Hartman TE, Yang E, Driscoll T, Cole LE, Sebra R, Maqbool SB, Wattam AR, Jacobs

WR, Jr. 2015. The complete genome sequence of the emerging pathogen *Mycobacterium haemophilum* explains its unique culture requirements. *MBio* 6:e01313-01315.

APPENDIX A

Workflow for cyclic nucleotide/dinucleotide extraction



[a]The protocol is modified for *M. tuberculosis* and *M. smegmatis*. [b]The tubes used for neutralization step should be 15 ml or 50 ml falcon tubes for samples less than 5 ml and more than 5 ml, respectively, to prevent loss of samples from the overflow of bubbles generated from neutralization reaction.

APPENDIX B

DNA sequences of *ml1750c*, *dgcA*, *dgcB*, and *sigA* gene fragments used in qRT-PCR and ddPCR reactions

M. smegmatis msmeg_2196

3'-CGGTATCGCACAAGGCATTCCGGGTAAGGATTCAACGTCCGACGTGCTCAACCGG
GCCGACCACGCGGTATTGACGGCCAAGGGTTC-5'

M. smegmatis sigA

3'-GCCGAGAAGGGCGAGAAGCTGCCAGTGCAGCAGCGCCGCGACATGCAGTGGATC
TGCCGCGACGGCGACCGCGCCAAAAACCATCTGCTGGAGGCGAACC-5'

M. leprae ml1750c

3'GCTGCTGAAAGCCTAACGGTTATTGCCCGTTTCGGCGGAGACGAGTTTGTCGTAGT
GCCCGCTGAGTCGGTGTCGGTTGATGTTGCAGAGTCGTTTGCCCATCGCTTGCAGAC
GCGGCTTCAAAAGCAAGTGGTGATCGACGGCGAAATACTCACCCGCACCGTCAGCA
TTGGTGTTGCCACCGGTCTTCCGGGACGGGATACCACGTCGGATCTGCTACGTTGGG
CTGATCACGCGGCGCTGTCGGCGAAAAGCGACGGCAGCAAGGTCGTGGTTCTTGAC
CACGGGATCAGCGCACAGCATACGCTGAGAACTGAGGTTGAACTGCACTTAGCAGG
GATGATCGATACTGACCTGGTGCTGCACTACCTTCCTGAGGTAGACATGAGTACCGG
CAAAGTCTTGGGCACCGAGGCCTTGGTGCGCTGGCAGCACCCGACACGGGGGTTGC
TATTTCCCGATTCATTTATCCCAGTGGCGGAATCGATTAATCTAGCAGG- 5'

***M. leprae ml1419c* (*dgcA*)**

3'-GCATCAGACGTCCTAGGCATGCCGATCGGTGAAGCTGTAGGCGCGCCGCTGGATC
CGGGCGCCATAATCACCGAGGGTGGGGTGCTAAGTACTACCCAGTACGCCTCGGAT
GGCATGGCTCTGGCCGTGCGGGTGTCTGCGGCCGTGATGGAAAACGGTTTCGTGCTG
GTGTGCAGCGATTACACCGCGCTGCTCCGAGTCGAGCGCCGCTTCCAAACTGTCGTC
GCATTATTGGAGGAAGGAGTCATCGTCCTGGACGGCGACGGCTACGTCGAATCAGT
GAACCCTGCCGCCCTGCGCATTCTTGGCGTCAGTGACCGAAGTGCCCTCGATGACCC
CGTCCGGCGGGCAGCAACGCTTCCGATGTACGATGCCCAAGGCCGGCTGCTTGACC
ACGGTCGGGGGGCTGTGCGCGAGTTTTTGCGGACCGGTACCCCAAAAACTGGCTTCA
TCGTCGGTATCGATCGTCCTAGCGACGGGACACGGGTTTGGTTGTCAGT-5'

***M. leprae ml0397c* (*dgcB*)**

3’GCTGGGCGGAGACGAGTTCATCGTACTACTCCGGGGTCCCCTTTCGGACATGAATG
CTAACGACGTCGCAAAAAGACTGCACACAACACTGTCCGAGTCACTCGTGGTCGAC
CAGCTCACGGTGCCTATCGGAGCGAGTGTTGGCATCCTTGAGGTGAGACCTGACGAT
CGACGGCGGGCCGCCGACATTTTGCGTGATGCCGACTCCGCGATGTATGCAGCCAA
GAACAAGAAACAATGTGCCGTCACGCCGCAGCAGCTGGTACCGTTCGTCGCGCTGA
TCGCACTGTTTGTGTTTTTTACAGCTGCGGCCGGAGCTAAGTTCTATGCACCATCAAA
TCTGCTGGTCATTCTCCAACAAACCGTGGTGCTGGCGATCGTCGGATACGGCATGAC
TTTCGTTATCATGGCAGGCTCCGTTGAACTATCGGTGGGCTCGATCGTCGCATTGACT
GGAGTCACCGCGGCACTCGTGGCTGCACAGAATCAATTCGCTGCGA-5’

M. leprae sigA

3’GCCCAGCGAGGACATCGACATCGACATCGAAGCCGCCGACCTCGGTCTCGACGAC
CTCGAAGACGACGACGTAGTCACCGACATCGAACCAGGTGAAACCGAGGATGGCGA
AGCTACTGCAGCCACCAAGACCGACGAAGCTTCCACCGATGATGACGAAGAGATCG
TTGAGCCGACCGAAAAAGACAAGGCCTCCGGCGATTTCGTCTGGGACGAAGATGAA
TCCGAGGCGTTGCGTCAGGCCCGTAAGGATGCCGAGCTCACCGCATCGGCAGACTC
AGTTCGCGCGTACCTCAAACAGATCGGCAAGGTGGCCCTGCTCAACGCCGAGGAAG
AGGTTGAGCTGGCGAAGCGGATCGAGGCTGGCCTATACGCCACGCAGCTAATGACC
GAGCTGTCTGAGCGGGGCACCAAACTACCCACCGCCCAGCGCCGCGACATGATGTG
GATCTGTCGCGACGGCGATCGTGCGAAAAACCATCTGTTGGAAGCCAACCTGC-5’

APPENDIX C

Detected proteins in *M. leprae* whole cell lysate derived from infected armadillo by mass spectrometry

	Identified Proteins	UniProt Accession Number	Leproma	Molecular Weight	Unique peptide count		
					WCS 1	WCS 2	WCS 3
1	60 kDa chaperonin 2	CH602_MYCLE	ML0317	57 kDa	59	81	154
2	DNA-directed RNA polymerase subunit beta	RPOC_MYCLE	ML1890	147 kDa	64	73	88
3	Putative fatty oxidation complex alpha subunit	Q9CBD8_MYCLE	ML2161	76 kDa	41	51	57
4	Chaperone protein DnaK	DNAK_MYCLE	ML2496	67 kDa	56	92	90
5	60 kDa chaperonin 1	CH601_MYCLE	ML0381	56 kDa	42	29	132
6	Isocitrate lyase	ACEA_MYCLE	ML1985	68 kDa	33	67	64
7	Polyketide synthase	Q49934_MYCLE	ML2353	160 kDa	65	55	62
8	Putative mycocerosic synthase	Q9CD78_MYCLE	ML0139	225 kDa	57	49	67
9	Fatty acid synthase	Q7AQ85_MYCLE	ML1191	328 kDa	56	48	44
10	Citrate synthase	Q7APY8_MYCLE	ML2130	48 kDa	16	16	27
11	UPF0182 protein ML0644	Y644_MYCLE	ML0644	108 kDa	24	22	34
12	Acyl-CoA dehydrogenase	Q7APU4_MYCLE	ML2563	67 kDa	19	64	26
13	DNA-directed RNA polymerase subunit beta	RPOB_MYCLE	ML1891	130 kDa	40	31	42
14	Probable cysteine desulfurase 1	CSD1_MYCLE	ML0842	65 kDa	37	7	17
15	Chaperone protein HtpG	HTPG_MYCLE	ML1623	74 kDa	36	22	48
16	Elongation factor Tu	EFTU_MYCLE	ML1877	44 kDa	16	17	17
17	Putative oxidoreductase	Q7AQE6_MYCLE	ML0862	65 kDa	21	28	27
18	Phthiodiolone/phenolphthiodiolone dimycocerosates ketoreductase	PHKR_MYCLE	ML0131	41 kDa	7	6	9
19	Chaperone protein ClpB	CLPB_MYCLE	ML2490	93 kDa	30	22	28
20	Putative long-chain-fatty-acid-CoA ligase	Q9CCE6_MYCLE	ML0887	64 kDa	14	44	21
21	Elongation factor G	EFG_MYCLE	ML1878	77 kDa	17	24	20
22	Malate synthase G	MASZ_MYCLE	ML2069	80 kDa	18	23	32
23	Adenosylhomocysteinase	SAHH_MYCLE	ML0771	54 kDa	10	3	48
24	Polyribonucleotide nucleotidyltransferase	PNP_MYCLE	ML0854	82 kDa	26	12	12

25	Protease IV, signal peptide peptidase	Q7AQ20_MYCLE	ML1839	64 kDa	8	18	49
26	ATP synthase subunit beta	ATPB_MYCLE	ML1145	53 kDa	11	8	35
27	5-methyltetrahydropteroyltriglutamate--homocysteine methyltransferase	METE_MYCLE	ML0961	82 kDa	14	7	13
28	30S ribosomal protein S1	RS1_MYCLE	ML1382	53 kDa	8	10	39
29	ATP synthase subunit alpha	ATPA_MYCLE	ML1143	60 kDa	4	10	31
30	Acyl-CoA synthetase	Q9CD79_MYCLE	ML0138	63 kDa	11	32	30
31	Bifunctional thioredoxin reductase/thioredoxin	TRXB_MYCLE	ML2703	49 kDa	13	11	29
32	Polyketide synthase	Q49933_MYCLE	ML2355	233 kDa	15	18	27
33	MoxR homolog 1	Q9CBL7_MYCLE	ML1810	41 kDa			6
34	Possible oxidoreductase	Q7APU2_MYCLE	ML2565	47 kDa		4	8
35	Uncharacterized protein	Q9CBP5_MYCLE	ML1751	78 kDa	33	11	13
36	Phosphoenolpyruvate carboxykinase [GTP]	PCKG_MYCLE	ML2624	68 kDa	35	23	9
37	Multifunctional 2-oxoglutarate metabolism enzyme	KGD_MYCLE	ML1095	137 kDa	25	8	18
38	Major membrane protein I	MMP1_MYCLE	ML0841	34 kDa	8	5	8
39	NAD(P) transhydrogenase subunit beta	Q9CCZ8_MYCLE	ML2634	49 kDa	13	13	18
40	Aconitate hydratase	Q9CBL3_MYCLE	ML1814	103 kDa	15	13	24
41	Cyclopropane mycolic acid synthase 2	CMAS2_MYCLE	ML2426	35 kDa	11	12	12
42	Putative cytochrome p450	Q9CBE7_MYCLE	ML2088	47 kDa	6	6	9
43	Putative acyl-CoA synthetase	Q9CDB2_MYCLE	ML0100	70 kDa	41	13	15
44	Succinate dehydrogenase flavoprotein subunit	Q9CCM1_MYCLE	ML0697	64 kDa	12	22	19
45	Acetyl-/propionyl-coenzyme A carboxylase alpha chain	BCCA_MYCLE	ML0726	64 kDa	15	17	18
46	Acyl-CoA synthase	Q9CD27_MYCLE	ML2546	59 kDa	9	17	19
47	Glutamine--fructose-6-phosphate aminotransferase [isomerizing]	GLMS_MYCLE	ML0371	68 kDa	17	31	15
48	Probable cation-transporting ATPase I	CTPI_MYCLE	ML2671	167 kDa	20	9	21
49	Uncharacterized protein	Q9CBV5_MYCLE	ML1536	68 kDa	18	24	18
50	DNA-directed RNA polymerase subunit alpha	RPOA_MYCLE	ML1957	38 kDa	9	7	8
51	Putative polyketide synthase	Q9CD81_MYCLE	ML0135	220 kDa	12	9	26

52	Uncharacterized protein	Q9CBP6_MYCLE	ML1750	67 kDa	21	24	11
53	Polyketide synthase	Q9CB70_MYCLE	ML2354	193 kDa	11	18	17
54	Alpha-keto-acid decarboxylase	KDC_MYCLE	ML2167	61 kDa		9	30
55	Putative asparagine synthetase	Q9CCF2_MYCLE	ML0874	76 kDa	43	12	6
56	Serine/threonine-protein kinase PknB	PKNB_MYCLE	ML0016	66 kDa	25	3	3
57	Protease II	Q7APX4_MYCLE	ML2226	81 kDa	7		
58	Phosphoenolpyruvate carboxylase	CAPP_MYCLE	ML0578	103 kDa	15	12	10
59	Uncharacterized protein ML0055	Y055_MYCLE	ML0055	62 kDa	4	16	27
60	Dihydrolipoyl dehydrogenase	DLDH_MYCLE	ML2387	50 kDa		4	6
61	ABC-transporter protein, ATP binding component	Q7AQ80_MYCLE	ML1248	62 kDa	4	27	26
62	Alkyl hydroperoxide reductase	Q9CBF5_MYCLE	ML2042	21 kDa	7	7	13
63	Probable propionyl-CoA carboxylase beta chain 5	PCCB_MYCLE	ML0731	59 kDa		6	36
64	DNA polymerase III, subunit [gamma/tau]	Q7APW2_MYCLE	ML2335	66 kDa	21	10	10
65	Glycine--tRNA ligase	Q9CCG4_MYCLE	ML0826	53 kDa	5		18
66	Phthiotriol/phenolphthiotriol dimycocerosates methyltransferase	PHMT_MYCLE	ML0130	31 kDa	12	5	10
67	Probable acyl-CoA dehydrogenase fadE25	ACDP_MYCLE	ML0737	42 kDa	2		4
68	Trans-acting enoyl reductase	TAER_MYCLE	ML0129	45 kDa	7		4
69	Polyketide synthase	Q7APV8_MYCLE	ML2357	200 kDa	17	12	16
70	Probable zinc metalloprotease	Q7APT3_MYCLE	ML2613	75 kDa	7		11
71	Uncharacterized protein	Q9CCI1_MYCLE	ML0798	65 kDa		38	
72	Alanine--tRNA ligase	SYA_MYCLE	ML0512	98 kDa	7	9	14
73	DNA gyrase subunit A	GYRA_MYCLE	ML0006	141 kDa	12	6	18
74	Uncharacterized protein ML2537	Y2537_MYCLE	ML2537	69 kDa	17	8	17
75	S-adenosylmethionine synthase	METK_MYCLE	ML0544	43 kDa	6		3
76	Aspartate--tRNA(Asp/Asn) ligase	SYDND_MYCLE	ML0501	65 kDa	19	16	
77	Isocitrate dehydrogenase [NADP]	Q7APS3_MYCLE	ML2672	83 kDa	5	3	8
78	Putative monooxygenase	Q7AQN8_MYCLE	ML0065	56 kDa			21
79	Probable acyl-CoA synthase	Q7APV7_MYCLE	ML2358	63 kDa	3	19	11

80	Polyketide synthase	Q9CDB1_MYCLE	ML0101	193 kDa	5	7	7
81	Putative acyl-CoA synthetase	Q9CD84_MYCLE	ML0132	74 kDa	16	18	
82	Alanine dehydrogenase	Q9CBV6_MYCLE	ML1532	39 kDa			3
83	Possible secreted hydrolase	Q9CBT1_MYCLE	ML1633	57 kDa			35
84	Proteasome-associated ATPase	ARC_MYCLE	ML1316	67 kDa	11	6	
85	Uncharacterized protein	Q9CBK6_MYCLE	ML1898	50 kDa			6
86	Glutamine synthetase	Q9CCD8_MYCLE	ML0925	54 kDa	2	4	17
87	Transcription termination/antitermination protein NusA	NUSA_MYCLE	ML1558	38 kDa			3
88	Possinble serine protease	Q7AQ93_MYCLE	ML1078	56 kDa			11
89	Putative fatty-acid--CoA ligase fadD21	FAD21_MYCLE	ML1234	63 kDa		20	14
90	Ubiquinol-cytochrome c reductase cytochrome b subunit	QCRB_MYCLE	ML0879	62 kDa	4	4	
91	Protein translocase subunit SecA 1	SECA1_MYCLE	ML0779	106 kDa	2	5	17
92	Possible long-chain acyl-CoA synthase	Q9CC34_MYCLE	ML1346	108 kDa	10	6	12
93	Lipoprotein LpqB	LPQB_MYCLE	ML0775	62 kDa		8	17
94	Trehalose-phosphate synthase	OTSA_MYCLE	ML2254	56 kDa	5	6	14
95	Inosine-5'-monophosphate dehydrogenase	IMDH_MYCLE	ML0387	55 kDa			23
96	Glycerol-3-phosphate acyltransferase	PLSB_MYCLE	ML1246	87 kDa	6	4	16
97	18 kDa antigen	18KD_MYCLE	ML1795	17 kDa	6	7	7
98	Valine--tRNA ligase	SYV_MYCLE	ML1472	99 kDa	6	6	6
99	Dihydroxy-acid dehydratase	ILVD_MYCLE	ML2608	59 kDa			22
100	RNA polymerase sigma factor SigA	Q7AQB1_MYCLE	ML1022	63 kDa	2		25
101	Putative acyl CoA carboxylase [beta] subunit	Q9CDB0_MYCLE	ML0102	56 kDa			34
102	Putative beta-ketoadipyl CoA thiolase	Q9CBD7_MYCLE	ML2162	42 kDa			3
103	Probable iron-sulphur-binding reductase	Q9CB20_MYCLE	ML2501	96 kDa	6	2	9
104	Pyridine transhydrogenase subunit [alpha]1	Q9CCZ6_MYCLE	ML2636	38 kDa			3
105	Uncharacterized protein	Q9CBW5_MYCLE	ML1512	59 kDa			15
106	Glucose-6-phosphate isomerase	G6PI_MYCLE	ML0150	61 kDa			21
107	Mycocerosic acid synthase (Polyketide synthase)	Q49624_MYCLE	ML1229	226 kDa	10	5	10
108	10 kDa chaperonin	CH10_MYCLE	ML0380	11 kDa	6	6	9

109	Homoserine dehydrogenase	DHOM_MYCLE	ML1129	46 kDa	2		3
110	Putative ferredoxin-dependent glutamate synthase	Q9CDD5_MYCLE	ML0061	166 kDa	9	8	8
111	D-3-phosphoglycerate dehydrogenase	SERA_MYCLE	ML1692	54 kDa			22
112	Probable ATP-dependent Clp protease ATP-binding subunit	CLPC_MYCLE	ML0235	94 kDa	9		15
113	Transketolase	TKT_MYCLE	ML0583	76 kDa	7		12
114	Serine hydroxymethyltransferase	GLYA_MYCLE	ML1953	45 kDa		3	
115	Trigger factor	TIG_MYCLE	ML1481	51 kDa			13
116	Glutamine-dependent NAD(+) synthetase	NADE_MYCLE	ML1463	76 kDa	11		5
117	Argininosuccinate synthase	ASSY_MYCLE	ML1412	44 kDa	4		
118	ATP synthase subunit b-delta	ATPFD_MYCLE	ML1142	49 kDa	3		
119	Transcription-repair-coupling factor	Q9CD43_MYCLE	ML0252	131 kDa	3		14
120	Phosphomethylpyrimidine synthase	THIC_MYCLE	ML0294	60 kDa			14
121	Leucine--tRNA ligase	SYL_MYCLE	ML0032	108 kDa	3	5	3
122	Bifunctional purine biosynthesis protein PurH	PUR9_MYCLE	ML0161	55 kDa			11
123	Signal recognition particle protein	SRP54_MYCLE	ML1622	55 kDa			11
124	CTP synthase	PYRG_MYCLE	ML1363	64 kDa	5	3	6
125	Threonine--tRNA ligase	SYT_MYCLE	ML0456	78 kDa	3		7
126	Putative membrane protein	Q7APW0_MYCLE	ML2347	38 kDa		3	3
127	Carbamoyl-phosphate synthase large chain	CARB_MYCLE	ML0536	120 kDa	5	6	6
128	Probable DNA helicase	Q9CBE0_MYCLE	ML2157	61 kDa			21
129	Phenylalanine--tRNA ligase beta subunit	SYFB_MYCLE	ML1402	89 kDa	4	4	8
130	Translation initiation factor IF-2	IF2_MYCLE	ML1556	97 kDa	8		10
131	Putative thiosulfate sulfurtransferase	THTR_MYCLE	ML2198	31 kDa	6		8
132	Methionine--tRNA ligase	SYM_MYCLE	ML0238	60 kDa			17
133	Probable cytochrome c oxidase subunit 1	COX1_MYCLE	ML1728	64 kDa			4
134	Uncharacterized protein	Q49942_MYCLE	ML2346	34 kDa		3	3
135	Putative cholesterol oxidase	Q9CCV1_MYCLE	ML0389	62 kDa			12
136	ABC transporter	Q9CCF9_MYCLE	ML0848	77 kDa	7	7	
137	Adenylosuccinate lyase	Q9CBC1_MYCLE	ML2230	51 kDa			3
138	Bacterioferritin	BFR_MYCLE	ML2038	18 kDa			6

139	Aldehyde dehydrogenase	Q9CCZ5_MYCLE	ML2639	54 kDa			5
140	Uncharacterized protein	Q9CD30_MYCLE	ML2535	146 kDa	2	5	8
141	Chromosomal replication initiator protein DnaA	DNAA_MYCLE	ML0001	56 kDa			16
142	3-hydroxyacyl-CoA dehydrogenase	Q9CB39_MYCLE	ML2461	31 kDa	4	2	3
143	Uncharacterized protein	Q7AQ69_MYCLE	ML1320	56 kDa			7
144	Possible dienelactone hydrolase	Q9CC04_MYCLE	ML1444	25 kDa	3	4	
145	Tryptophan synthase beta chain	TRPB_MYCLE	ML1272	45 kDa	5		
146	Acetolactate synthase	ILVB_MYCLE	ML1696	67 kDa	8	9	
147	DNA gyrase subunit B	GYRB_MYCLE	ML0005	75 kDa	8		4
148	Fumarate hydratase class II	Q7AQ14_MYCLE	ML1947	50 kDa			10
149	Probable membrane protein	Q9CBV2_MYCLE	ML1539	54 kDa			7
150	Putative amino acid decarboxylase	Q9CCR8_MYCLE	ML0524	107 kDa	3	3	3
151	Acetyl/propionyl CoA carboxylase [beta] subunit	Q7AQ32_MYCLE	ML1657	50 kDa			4
152	Possible coenzyme F420-dependent oxidoreductase	Q9CCV8_MYCLE	ML0348	38 kDa	3		
153	Methylmalonyl-CoA mutase, [beta] subunit	Q9CBM6_MYCLE	ML1800	67 kDa	4	7	
154	Polyketide synthase	Q7APV9_MYCLE	ML2356	164 kDa	7	7	7
155	Elongation factor 4	LEPA_MYCLE	ML0611	71 kDa	9	3	3
156	CDP-diacylglycerol-glycerol-3-phosphate	Q7APZ2_MYCLE	ML2081	23 kDa			3
157	Uncharacterized protein	Q7AQ79_MYCLE	ML1249	178 kDa			4
158	Lysine--tRNA ligase	SYK_MYCLE	ML0233	56 kDa			9
159	Possible dehydrogenase	Q7APZ9_MYCLE	ML2061	50 kDa	2		5
160	Pyruvate dehydrogenase E1 component	Q9CBS8_MYCLE	ML1651	105 kDa	3		
161	Phthiocerol/phthiodiolone dimycocerosyl transferase	PAPA5_MYCLE	ML2349	46 kDa			3
162	DNA-binding protein HU homolog	DBH_MYCLE	ML1683	21 kDa			4
163	Glyceraldehyde-3-phosphate dehydrogenase	G3P_MYCLE	ML0570	36 kDa			2
164	Probable cytosol aminopeptidase	AMPA_MYCLE	ML0864	54 kDa			7
165	Sensor histidine kinase MtrB	MTRB_MYCLE	ML0774	62 kDa			10
166	UPF0051 protein ML0593	Y593_MYCLE	ML0593	96 kDa			9
167	DNA primase	DNAG_MYCLE	ML0833	71 kDa	8		

168	Aspartyl/glutamyl-tRNA(Asn/Gln) amidotransferase subunit B	GATB_MYCLE	ML1700	55 kDa			14
169	Protein translocase subunit SecF	SECF_MYCLE	ML0488	50 kDa		2	
170	Glycine dehydrogenase (decarboxylating)	GCSP_MYCLE	ML2072	101 kDa			4
171	Serine/threonine-protein kinase PknG	PKNG_MYCLE	ML0304	83 kDa	8		
172	Putative acyl-CoA synthetase	Q9CD50_MYCLE	ML0243	59 kDa			10
173	Ribonucleoside-diphosphate reductase subunit alpha	RIR1_MYCLE	ML1734	82 kDa	3		3
174	Succinate-semialdehyde dehydrogenase [NADP(+)]	GABD1_MYCLE	ML2573	49 kDa			7
175	Glycerol-3-phosphate dehydrogenase	GLPD_MYCLE	ML0713	63 kDa			10
176	Acetolactate synthase II	Q7APZ1_MYCLE	ML2083	58 kDa			6
177	Arginine--tRNA ligase	SYR_MYCLE	ML1127	60 kDa			9
178	Putative oxidoreductase	Q7AQJ6_MYCLE	ML0315	33 kDa			4
179	Protein translocase subunit SecD	SECD_MYCLE	ML0487	61 kDa	2	2	5
180	Uncharacterized protein	Q7APW4_MYCLE	ML2332	15 kDa	2		
181	Isoleucine--tRNA ligase	SYI_MYCLE	ML1195	120 kDa	5		
182	UPF0051 protein ML0594	Y594_MYCLE	ML0594	42 kDa		3	6
183	Putative dihydrolipoamide acyltransferase	Q7AQE7_MYCLE	ML0861	55 kDa		3	
184	DNA topoisomerase 1	TOP1_MYCLE	ML0200	104 kDa	5		
185	Methylmalonyl-CoA mutase, [alpha] subunit	Q9CBM7_MYCLE	ML1799	82 kDa		6	2
186	Proline--tRNA ligase	SYP_MYCLE	ML1553	54 kDa			6
187	Possible SpoIIIE-family membrane protein	Q7AQ47_MYCLE	ML1543	148 kDa			8
188	Probable extracellular solute-binding dependent transport lipoprotein	Q9CC82_MYCLE	ML1121	60 kDa			8
189	Putative carbohydrate kinase	Q9CCF3_MYCLE	ML0873	34 kDa	2	5	4
190	Diacylglycerol O-acyltransferase	Q7AQ82_MYCLE	ML1244	52 kDa			5
191	Putative acyl-CoA dehydrogenase	Q7AQG8_MYCLE	ML0661	50 kDa			2
192	Possible GTP-binding, protein elongation factor	Q9CBX2_MYCLE	ML1498	68 kDa	6		4
193	Glutamate-ammonia-ligase adenylyltransferase	GLNE_MYCLE	ML1630	110 kDa	4		
194	Probable transmembrane protein	Q9CB64_MYCLE	ML2400	30 kDa			4
195	Uncharacterized protein	Q9CDB7_MYCLE	ML0093	72 kDa	5		

196	Acyl-CoA synthase	Q9CCY8_MYCLE	ML2661	58 kDa			7
197	1-deoxy-D-xylulose-5-phosphate synthase	DXS_MYCLE	ML1038	69 kDa	8		
198	Uncharacterized protein	Q9CCL8_MYCLE	ML0703	46 kDa			6
199	Probable M18 family aminopeptidase 2	APEB_MYCLE	ML2213	46 kDa			2
200	3-deoxy-D-arabino-heptulosonate 7-phosphate synthase (DAHP synthetase)	Q7AQD4_MYCLE	ML0896	50 kDa			7
201	tRNA-2-methylthio-N(6)-dimethylallyladenosine synthase	MIAB_MYCLE	ML0989	56 kDa			4
202	Putative two-component system sensor kinase	Q9CCH8_MYCLE	ML0803	54 kDa			6
203	UvrABC system protein A	UVRA_MYCLE	ML1392	106 kDa	5		
204	Acyl-CoA synthase	Q7AQ08_MYCLE	ML1994	57 kDa			9
205	Putative secreted protein	Q9CD09_MYCLE	ML2594	54 kDa			8
206	Polyphosphate kinase	PPK_MYCLE	ML1681	83 kDa			3
207	Probable GTP pyrophosphokinase	RELA_MYCLE	ML0491	87 kDa	2		2
208	Uncharacterized glycosyl hydrolase ML0392	Y392_MYCLE	ML0392	88 kDa	3		3
209	PcnA	Q9CCY1_MYCLE	ML2697	54 kDa			4
210	Possible uroporphyrin-III C-methyltransferase	Q9CB60_MYCLE	ML2420	59 kDa			4
211	Possible transferase	Q7AQ81_MYCLE	ML1245	62 kDa			5
212	Protein translocase subunit SecA 2	SECA2_MYCLE	ML2082	85 kDa			8
213	Glutamate--tRNA ligase	SYE_MYCLE	ML1688	55 kDa			6
214	Probable ferredoxin/ferredoxin--NADP reductase	FPRB_MYCLE	ML2134	60 kDa			4
215	Probable transport protein MmpL4	MMPL4_MYCLE	ML2378	105 kDa	3		
216	Elongation factor Ts	EFTS_MYCLE	ML1597	29 kDa	3		
217	Possible transcriptional regulator	Q9CBP3_MYCLE	ML1753	119 kDa	2		
218	Ribose-phosphate pyrophosphokinase	KPRS_MYCLE	ML0248	36 kDa			3
219	Uncharacterized protein	Q9CDD7_MYCLE	ML0052	65 kDa		3	
220	Putative secreted protein	Q9CD12_MYCLE	ML2591	54 kDa			4
221	Putative phospho-sugar mutase	Q9CCL7_MYCLE	ML0706	57 kDa			3
222	Conserved membrane protein	Q9CD29_MYCLE	ML2536	58 kDa			7

223	2-succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexene-1-carboxylate synthase	MEND_MYCLE	ML2270	58 kDa	5
224	Possible two-component response regulatory protein	Q9CC47_MYCLE	ML1286	23 kDa	2
225	Putative secreted protein	Q9CD11_MYCLE	ML2592	57 kDa	3
226	DNA translocase FtsK	FTSK_MYCLE	ML0977	90 kDa	4
227	Pup--protein ligase	PAFA_MYCLE	ML1328	51 kDa	4
228	NADPH-ferredoxin reductase FprA	FPRA_MYCLE	ML0666	50 kDa	4
229	3-isopropylmalate dehydratase large subunit	LEUC_MYCLE	ML1685	51 kDa	5
230	DNA ligase	DNLJ_MYCLE	ML1705	76 kDa	2
231	Putative enoyl-CoA hydratase	Q9CD94_MYCLE	ML0120	30 kDa	2
232	Anthranilate synthase component 1	TRPE_MYCLE	ML1269	57 kDa	3
233	ATP synthase gamma chain	ATPG_MYCLE	ML1144	33 kDa	3
234	Uncharacterized protein	Q9CBR4_MYCLE	ML1714	57 kDa	4
235	Uncharacterized lipoprotein ML0489	Y489_MYCLE	ML0489	58 kDa	3
236	Possible regulatory protein	Q7AQ57_MYCLE	ML1419	61 kDa	3
237	Methyl mycolic acid synthase 1	Q9CBK4_MYCLE	ML1900	33 kDa	2
238	Conserved membrane protein	Q7AQ46_MYCLE	ML1544	54 kDa	2
239	30S ribosomal protein S7	RS7_MYCLE	ML1879	18 kDa	2
240	Uncharacterized protein	Q9CCT2_MYCLE	ML0508	35 kDa	2

APPENDIX D

ML1750c and DgcA peptides detected by mass spectrometry

ML1750c													
Samples	**Sequences**	**Mascot Ion score**	**Mascot Delta Ion Score**	**NTT**	**Observed**	**Actual Mass**	**Charge**	**Delta Da**	**Delta PPM**	**Retention Time**	**TIC**	**Start**	**Stop**
WCS1	(K)QVVIDGEILTR(T)	11.08	11.03	2	621.8566	1,241.70	2	0.000534	0.4298	2,700.24	44,388.10	287	297
WCS1	(R)IVAEEQLR(Y)	38.19	26.19	2	479.2725	956.53	2	0.0012	1.253	1,299.14	45,100.80	169	176
WCS1	(R)IVAEEQLR(Y)	28.23	16.64	2	479.2731	956.532	2	0.00242	2.527	1,297.28	35,777.50	169	176
WCS1	(R)LIEVFADR(L)	50.07	50.07	2	481.7692	961.524	2	0.000408	0.4239	2,515.29	47,322.50	234	241
WCS1	(R)LIEVFADR(L)	46.03	46.03	2	481.7687	961.523	2	-0.000568	-0.5901	2,516.16	93,645.50	234	241
WCS1	(K)EPAVLRPEPANADYQR(R)	35.64	35.64	2	609.311	1,824.91	3	-0.000876	-0.4797	1,745.74	47,927.10	94	109
WCS1	(K)EPAVLRPEPANADYQR(R)	29.64	29.01	2	609.3116	1,824.91	3	0.00114	0.6244	1,748.57	97,431.30	94	109
WCS1	(R)YLAEHDDLTGLLNR(R)	40.6	30.81	2	543.946	1,628.82	3	0.000237	0.1455	2,924.87	40,295.20	177	190
WCS1	(R)YLAEHDDLTGLLNR(R)	40.49	30.17	2	543.9453	1,628.81	3	-0.002145	-1.316	2,926.70	52,095.40	177	190
WCS1	(R)GPVTLLFLGLDR(L)	20.87	17.69	2	650.8848	1,299.76	2	-0.000268	-0.206	4,618.14	6,691.07	207	218
WCS1	(R)EAAESLTVIAR(F)	57.55	47.03	2	580.319	1,158.62	2	-0.00096	-0.8278	2,085.45	14,159.80	244	254

WCS1	(R)EAAESLTVIAR(F)	47.41	31.58	2	580.3196	1,158.62	2	0.000138	0.1191	2,084.73	13,775.00	244	254
WCS1	(K)QVVIDGEILTR(T)	41.47	40.1	2	621.8563	1,241.70	2	-7.59E-05	-0.06111	2,702.09	184537	287	297
WCS1	(R)TVSIGVATGLPGR(D)	61.81	55.74	2	614.3572	1,226.70	2	0.001456	1.186	2,518.13	46,054.60	298	310
WCS1	(R)TVSIGVATGLPGR(D)	47.03	38.82	2	614.3572	1,226.70	2	0.001456	1.186	2,515.14	18,334.00	298	310
WCS1	(K)VVVLDHGISAQHTLR(T)	45.75	44.74	2	548.9774	1,643.91	3	-0.000447	-0.2717	1,932.20	35,086.00	334	348
WCS1	(K)VVVLDHGISAQHTLR(T)	41.72	41.72	2	548.9777	1,643.91	3	0.000285	0.1733	1,934.98	40,372.60	334	348
WCS1	(K)VLGTEALVR(W)	66.78	53.99	2	479.2898	956.565	2	-0.000532	-0.5555	2,028.74	128862	378	386
WCS1	(K)VLGTEALVR(W)	48.74	39.25	2	479.2904	956.566	2	0.000628	0.6559	2,026.81	54,429.70	378	386
WCS1	(R)SAGVGLDALLR(I)	48.6	33.35	2	536.3115	1,070.61	2	-0.000198	-0.1847	3,224.18	163425	430	440
WCS1	(R)SAGVGLDALLR(I)	49.88	32.39	2	536.3114	1,070.61	2	-0.00032	-0.2986	3,235.80	78,641.70	430	440
WCS2	(K)EPAVLRPEPANADYQR(R)	31.78	31.78	2	609.3115	1,824.91	3	0.000588	0.3221	1,755.41	56,434.50	94	109
WCS2	(K)EPAVLRPEPANADYQR(R)	31.19	31.19	2	609.3113	1,824.91	3	0.000222	0.1216	1,757.24	99,816.50	94	109
WCS2	(R)IVAEEQLR(Y)	49.82	36.12	2	479.2726	956.531	2	0.001506	1.573	1,303.04	33,273.20	169	176
WCS2	(R)IVAEEQLR(Y)	39.82	25.94	2	479.2726	956.531	2	0.001444	1.508	1,302.10	26,514.80	169	176
WCS2	(R)YLAEHDDLTGLLNR(R)	29.8	21.7	2	543.9467	1,628.82	3	0.002067	1.268	2,932.55	28,119.70	177	190
WCS2	(R)YLAEHDDLTGLLNR(R)	28.9	20.6	2	543.9464	1,628.82	3	0.001152	0.7069	2,930.88	26,177.10	177	190
WCS2	(R)GPVTLLFLGLDR(L)	69.99	69.99	2	650.8853	1,299.76	2	0.000708	0.5443	4,605.64	24,677.40	207	218

WCS2	(R)LIEVFADR (L)	47.28	43.98	2	481.7687	961.523	2	-0.000568	-0.5901	2,518.21	26,966.90	234	241
WCS2	(R)LIEVFADR (L)	46.01	46.01	2	481.7692	961.524	2	0.000408	0.4239	2,519.28	40,658.90	234	241
WCS2	(R)EAAESLTVIAR(F)	66.51	54.75	2	580.3203	1,158.63	2	0.001604	1.383	2,089.38	16,664.70	244	254
WCS2	(R)EAAESLTVIAR(F)	93.45	70.54	2	580.3201	1,158.63	2	0.001236	1.066	2,090.47	15,296.50	244	254
WCS2	(K)QVVIDGEILTR(T)	49.52	48.01	2	621.8561	1,241.70	2	-0.000442	-0.3556	2,695.39	83,441.10	287	297
WCS2	(K)QVVIDGEILTR(T)	40.13	34.77	2	621.8564	1,241.70	2	0.000168	0.1352	2,696.54	57,225.10	287	297
WCS2	(K)QVVIDGEILTR(T)	34.23	25.11	2	622.3632	1,242.71	2	1.014	8.836	2,723.35	20,261.30	287	297
WCS2	(R)TVSIGVATGLPGR(D)	42.71	37.37	2	614.3564	1,226.70	2	-0.000254	-0.2068	2,521.21	15,840.50	298	310
WCS2	(R)TVSIGVATGLPGR(D)	32.91	27.96	2	614.3568	1,226.70	2	0.000602	0.4904	2,519.14	13,611.90	298	310
WCS2	(K)VVVLDHGISAQHTLR(T)	17.3	14.05	2	548.9774	1,643.91	3	-0.00063	-0.383	1,951.53	18,118.80	334	348
WCS2	(K)VLGTEALVR(W)	57.85	42.77	2	479.2897	956.565	2	-0.000776	-0.8103	2,029.17	46,940.60	378	386
WCS2	(K)VLGTEALVR(W)	61.21	40.67	2	479.2901	956.566	2	0.00014	0.1463	2,030.11	83,192.90	378	386
WCS2	(R)SAGVGLDALLR(I)	62.77	36.94	2	536.3153	1,070.62	2	0.007494	6.993	3,228.23	644302	430	440
WCS2	(R)SAGVGLDALLR(I)	55.38	31.98	2	536.3115	1,070.61	2	-7.59E-05	-0.07087	3,225.37	207753	430	440
WCS2	(K)SFVAELGSNASDLAIVR(A)	41.8	40.77	2	874.9626	1,747.91	2	0.00027	0.1544	3,789.04	7,492.96	529	545

WCS2	(R)AQGFLLSRPVDGAAmESLLAK(G)	33.82	30.68	2	730.7227	2,189.15	3	-0.005362	-2.448	3,276.87	65,068.70	580	600
WCS2	(R)AQGFLLSRPVDGAAmESLLAK(G)	31.01	18.24	2	730.723	2,189.15	3	-0.004447	-2.03	3,280.65	39,331.60	580	600
WCS3	(R)IVAEEQLR(Y)	34.37	24.08	2	479.2734	956.532	2	0.003154	3.294	1,295.60	49,685.10	169	176
WCS3	(R)IVAEEQLR(Y)	32.14	22.73	2	479.2701	956.526	2	-0.00356	-3.718	1,294.69	45,744.70	169	176
WCS3	(R)LIEVFADR(L)	50.24	50.24	2	481.7686	961.523	2	-0.00069	-0.7168	2,515.98	46,415.70	234	241
WCS3	(R)LIEVFADR(L)	64.16	64.16	2	481.7694	961.524	2	0.000896	0.9309	2,517.27	38,970.40	234	241
WCS3	(R)EAAESLTVIAR(F)	62.31	40.17	2	580.3203	1,158.63	2	0.001604	1.383	2,085.44	16,697.70	244	254
WCS3	(R)EAAESLTVIAR(F)	45.53	38.89	2	580.3193	1,158.62	2	-0.000472	-0.407	2,084.43	22,758.50	244	254
WCS3	(K)QVVIDGEILTR(T)	49.21	46.92	2	621.8541	1,241.69	2	-0.00447	-3.597	2,700.36	28,183.20	287	297
WCS3	(R)TVSIGVATGLPGR(D)	39.14	31.89	2	614.3562	1,226.70	2	-0.00062	-0.505	2,522.04	35,760.60	298	310
WCS3	(R)TVSIGVATGLPGR(D)	18.04	13.54	2	614.3571	1,226.70	2	0.00109	0.8879	2,524.79	12,890.40	298	310
WCS3	(K)VLGTEALVR(W)	53.05	44.31	2	479.2903	956.566	2	0.000384	0.4011	2,028.13	52,038.00	378	386
WCS3	(K)VLGTEALVR(W)	56.07	41.87	2	479.2902	956.566	2	0.0002	0.2089	2,027.22	60,104.40	378	386

DgcA													
Samples	Sequence	Mascot Ion score	Mascot Delta Ion Score	NTT	Observed	Actual Mass	Charge	Delta Da	Delta PPM	Retention Time	TIC	Start	Stop
WCS3	(R)EVGDASQPSEAVLLR(L)	46.76	34.36	2	785.907	1,569.80	2	-0.000436	-0.2775	2,390.98	18,560.00	99	113
WCS3	(R)EVGDASQPSEAVLLR(L)	39.4	26.54	2	785.9076	1,569.80	2	0.000664	0.4227	2,393.65	18,252.90	99	113
WCS3	(K)TGFIVGIDRPSDGTR(V)	20.59	8.42	2	530.9467	1,589.82	3	0.00185	1.163	2,354.77	25,574.70	338	352

*All peptides have Mascot identify score = 25.

LIST OF ABBREVIATIONS

(p)ppGpp	Guanosine pentaphosphate
AFB	Acid-fast bacilli
AG	Arabinogalactan
AHL	Acyl homoserine lactones
ANOVA	One-way analysis of variance
BB	Mid-borderline leprosy
BCA	Bicinchoninic acid
BL	Borderline lepromatous leprosy
BLAST	Basic Local Alignment Search Tool
BT	Borderline tuberculoid leprosy
cAMP	Cyclic adenosine 3’5’ monophosphate
CAP	Catabolite gene-activator protein
CDD	Conserved Domain Database
c-di-AMP	Cyclic dimeric adenosine monophosphate
c-di-GMP	Cyclic diguanosine monophosphate
CE	Collision energy
CGD	Chronic granulomatous disease
cGMP	Cyclic guanosine 3’5’ monophosphate
CRP	cAMP-receptor protein
CV	Coefficient of variation
DAC	Diadenylate cyclase

DAG	Diacylglycerol
DC	Dendritic cells
DC-SIGN	Dendritic cell-specific intercellular adhesion molecule-grabbing nonintegrin
ddPCR	Droplet digital PCR
DGC	Diguanylate cyclase
DTT	Dithiothreitol
ECF	Extra-cytoplasmic sigma factor
EDTA	Ethylenediaminetetraacetic acid
EIC	Extracted ion chromatogram
EIS	Electrochemical impedance spectroscopy
ENL	Erythema nodosum lepromasum
EPS	Exopolysaccharide
ESI	Electrospray ionization
FAD	Flavin adenosine dinucleotide
FMN	Flavin mononucleotide
FRET	Fluorescence resonance energy transfer
GM-CSF	Granulocyte-macrophage colony-stimulating factor
GMP	Guanosine monophosphate
GPL	Glycopeptidolipid
HPLC	High-performance liquid chromatography
IFN	Interferon
IL	Interleukin
IMAC	Immobilized metal affinity chromatography

IP	Inositol trisphosphate
IPTG	Isopropyl β-D-thiogalactoside
LAM	Lipoarabinomamnan
LC-MS	Liquid chromatography-coupled mass spectrometry
LL	Lepromatous leprosy
LM	Lipomannan
LOD	Limit of detection
LOQ	Limit of quantification
LPS	Lipopolyssacharide
MALDI-TOF	Matrix-assisted laser desorption/ionization-time of flight mass spectrometry
MB	Multibacillary leprosy
MDR	Multidrug efflux pump
MDT	Multidrug therapy
min	Minute
MLCwA	*M. leprae* cell wall fraction
MLMA	*M. leprae* cell membrane fraction
MLSA	*M. leprae* cytosol fraction
NCBI	National Center for Biotechnology information
NRT	Non-reverse transcriptase control
NTC	Non-template control
PAMP	Pathogen-associated molecular pattern
pApA	Phosphoadenylyl adenosine
PB	Paucibacillary leprosy

PCR	Polymerase chain reaction
PDE	Phosphodiesterase
PDIM	Phthiocerol dimycocerosate
PG	Peptidoglycan
PGL	Phenolic glycolipid
pGpG	5′phosphoguanylyl 3′5′ guanosine
PIMS	Phosphatidylinositol mannosides
ppGpp	Guanosine-3,5-bis(pyrophosphate) or guanosine tetraphosphate
PRR	Pattern recognition receptor
qRT-PCR	Quantitative reverse transcription polymerase chain reaction
Q-TOF	Quadrupole time-of-flight
RLEP	*M. leprae*-specific repetitive element
RNI	Reactive nitrogen intermediates
RNS	Reactive nitrogen species
ROI	Reactive oxygen intermediates
ROS	Reactive oxygen and nitrogen species
rpm	revolutions per minute
RR	Reversal reaction
RSH	RelA-SpoT homologue
SAH	Small alarmone hydrolases
SAS	Small alarmone synthase
SDS	Sodium dodecyl sulfate
SNR	Signal to noise ratio

STING	Stimulation of interferon gene
STPK	Serine/threonine protein kinase
TNF	Tumor necrosis factor
TT	Tuberculoid leprosy
VBMM	Vogel-Bonner Minimal Medium
WCS	Whole cell sonicate (*M. leprae*)