THESIS

BIOINFORMATIC IDENTIFICATION AND FUNCTIONAL CHARACTERIZATION OF CYTOKINETIC REGULATORS IN *MTB*

Submitted by

Rebecca M. Crew

Department of Microbiology, Immunology and Pathology

In partial fulfillment of the requirements

For the Degree of Master of Science

Colorado State University

Fort Collins, Colorado

Summer 2013

Master's Committee:

Advisor: Richard Slayden

Mary Jackson
William Hanneman

ABSTRACT

BIOINFORMATIC IDENTIFICATION AND FUNCTIONAL CHARACTERIZATION OF CYTOKINETIC REGULATORS IN *MTB*

A fundamental lack of understanding of *Mtb* regulation during latent tuberculosis infections (LTBI), which comprises the vast majority of tuberculosis cases, has hindered global eradication efforts. To elucidate mechanisms associated with transition to the non-replicating persistent (NRP) state associated with LTBI, we set out to identify regulators involved in cell division control in *Mtb*. Bioinformatic analysis identified *rv1708* as encoding a MinD-like protein putatively involved in septum placement, and *rv2216* as encoding a potential SOS-associated cell division inhibitor, SulA. Bioinformatic-based assessments of orthology revealed a differential lineage than anticipated for the proteins encoded by both open reading frames (ORFs). We describe these two novel regulators in *Mtb* here for the first time. It was found that Rv1708 lacks regions vital for MinD function and shows greater similarity with the Soj protein from *Bacillus* sp. involved in the regulation of sporulation and timing of division. Due to these similarities we have re-named Rv1708 as Soj_{Mtb}. Significantly, Soj_{Mtb} shows potential as a cytokinetic and dormancy regulator both by homology, morphology, and growth kinetic analysis. Overexpression of soj_{Mtb} attenuates growth and elicits filamentation characteristic of a disruption in early division, similar to Soj activity in other organisms. Given the role of Soj in the control of dormancy phenotypes in *Bacillus* sp. we believe Soj_{Mtb} serves as an important regulator during dormancy transitions in mycobacteria, as associated with the development of LTBI.

Although Rv2216 was initially identified by homology to SulA proteins, analysis of orthology indicates greater similarity with a separate group of widely conserved yet poorly defined cell division regulatory proteins. Thusly, we have re-name *rv2216* as *cdr* for cell division regulator. Cdr proteins share

limited similarity to SulA: enough to be mis-identified in organisms lacking a true SulA but insufficient to infer similar functionalities. Cdr proteins are present in hundreds of organisms through different walks of life, yet this work presents the first characterization of their effects on cellular activity. Induction of the SOS response by Mitomycin C treatment did not induce *cdr* expression, supporting our classification Cdr proteins separate from SulA. Overexpression of *cdr* resulted in a bimodal increase in cell length without an apparent effect on growth kinetics, suggesting Cdr stimulation of cellular elongation relative to division. Profiling of cell cycle discriminant genes in response to *cdr* overexpression corroborates this hypothesis, showing an induction of late division events associated with the production of new plasma membrane and cell wall components. Sub-cellular localization studies using an inducible Cdr-GFP fusion protein revealed cell cycle-dependent localization to the inner membrane at sites involved in cell wall and plasma membrane growth and remodeling. Furthermore, global transcriptional analysis revealed a unique profile of adaptive programs associated with hypoxia-associated NRP, *de novo* lipid synthesis and phospholipid/triacylglycerol turnover. These processes are required for normal growth and promote homeostasis during times of stress by preventing and repairing oxidative damage to membrane constituents in diverse organisms. Importantly, Cdr represents a novel regulatory class of proteins with broad representation in all classifications of life, potentially involved in division and stress responses associated with dormancy, and is described here for the first time in *Mtb.* The foundation provided here, both for Soj_{Mtb} and Cdr, provides insight into the regulatory mechanisms employed during NRP transitions associated with LTBI, and will aid in the development and implementation of more targeted studied in the future.

TABLE OF CONTENTS

CHAPTER 4: DISCUSSION

CHAPTER 5: MATERIALS AND METHODS

APPENDIX

LIST OF TABLES

APPENDIX I

APPENDIX II

LIST OF FIGURES

CHAPTER 1

CHAPTER 2

CHAPTER 3

CHAPTER I: INTRODUCTION

1.1 TUBERCULOSIS DISEASE

It has been estimated that one-third of the world's population is infected with *Mycobacterium tuberculosis* with documented cases in nearly every country [1-2]. The World Health Organization (WHO) calculated a prevalence of nearly 12 million tuberculosis (TB) cases in 2011 along with more than 1.4 million deaths, making TB the second leading cause of infectious disease-associated deaths [1]. The primary causative agent of TB in humans is *Mycobacterium tuberculosis* (*Mtb*), a non-motile, obligatory aerobic, weakly gram-positive, acid-fast bacillus belonging to the order Actinomycetales. TB is a chronic inflammatory disease that primarily affects the lungs, also known as pulmonary TB, but is also capable of disseminating and infecting multiple sites in the body including the bones, joints, lymph nodes, kidneys, spleen, genitourinary system, and central nervous system [3]. The bacteria are transmitted in respiratory droplets, and subsequently inhaled into the alveoli of the lungs where they attempt to establish an infection [2, 4]. In a majority of exposure cases the bacteria are cleared by the immune system, however 5-10% of exposures will result in the establishment of a primary infection (Figure 1) [2, 4]. In approximately 90-95% of primary infections the immune response is able to contain the infection leading to the development of what is referred to as a latent TB infection (LTBI)[2, 4]. Latent infections are non-transmissible and present no clinical symptoms of disease. However, upon sufficient immune suppression and other contributing factors the latent infection may reactivate with an average lifetime chance of 5-10% in an otherwise healthy individual [2, 4]. This reactivated or secondary progressive TB regains the capacity for transmission, and presents characteristic disease symptoms including a persistent cough, fever, fatigue, and weight loss eventually resulting in death in nearly half of untreated cases [1].

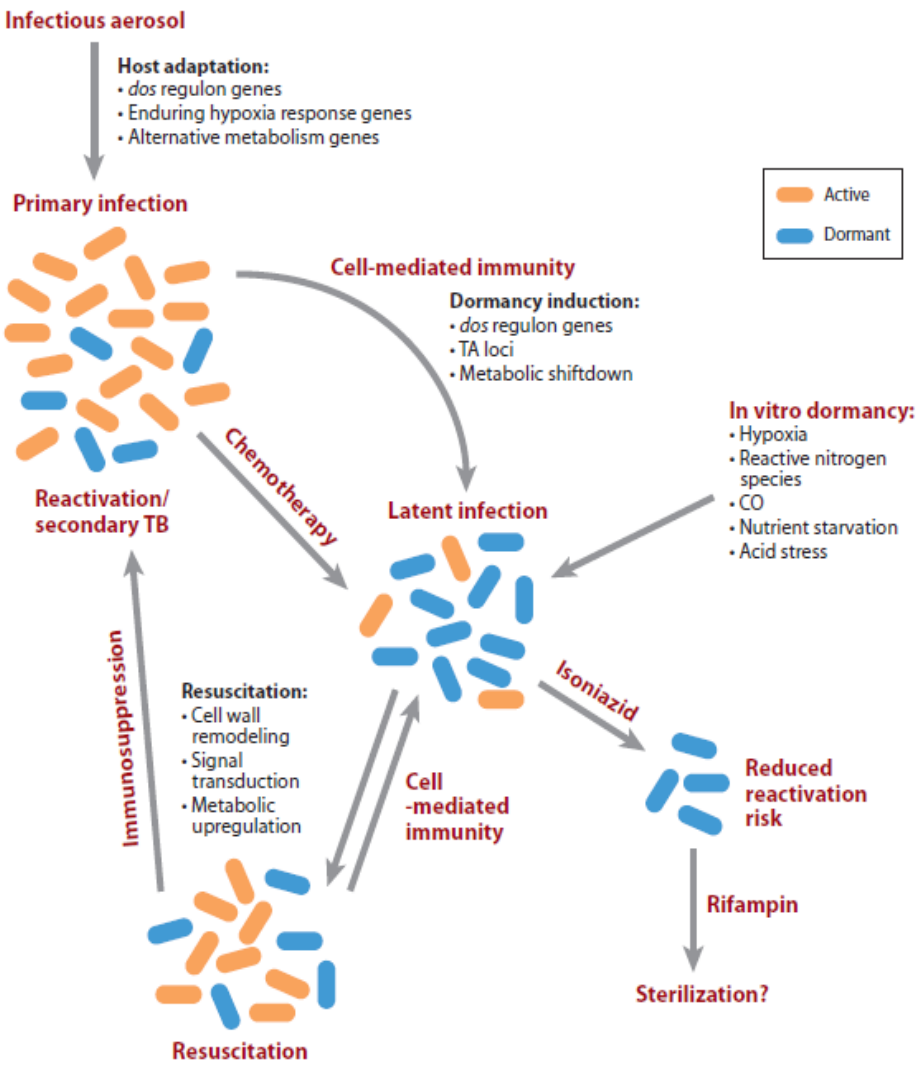

Figure 1.1: *Mtb* infections, TB disease, host-associated stresses and molecular programs important for pathology during infection.

Upon exposure the bacteria employs a variety of regulatory strategies to promote survival within the host. In 5-10% of cases exposure will result in the establishment of primary infection, however a large majority of the time the combination of stresses experienced within the host and cell mediated immunity or chemotherapy will promote bacterial entry into a dormant state characteristic of latent infections. In response to Immunosuppression and other factors the bacteria may exit dormancy accompanied by the induction of specific molecular programs, a process known as resuscitation, to re-establish an active disease state known as reactivated or secondary disease. The entry into and out of a latent infection may repeat for the lifetime of the individual unless all of the infecting bacteria are sterilized. Treatment strategies for latent infections are highlighted, yet importantly, these agents are ineffective against MDR-TB and XDR-TB strains as LTBI are treated with isoniazid therapy, against which these strains are resistant. (Reproduced with permissions from: Michael C. Chao and Eric J. Rubin, Let sleeping dos lie: does dormancy play a role in tuberculosis?," Annual Review of Microbiology, Volume 64, October 2010, Pages 293-311.)

Although evidence of *Mtb* causing disease in the human population dates back as far as 9250-8160 years BC, there were no effective therapeutics available for the treatment of TB until the discovery of Streptomycin in 1946 [5-6]. Yet after more than 9000 years of TB disease and 60 years of antibacterial chemotherapeutics, TB remains a global epidemic. The first observations of drug resistant strains of *Mtb* were made in the early 1940's, which increased in prevalence over the next few decades. The emergence of HIV in the 1980s contributed to a resurgence of TB disease and set back to global eradication efforts on a monumental scale leading the WHO to declare a global health emergency in 1993 [7]. Resistance in *Mtb* was first recognized for the front line therapeutics, termed such as they are the primary choices for treatment of TB infections, including isoniazid (INH), rifampin (RIF), ethambutol (EMB) and pyrazinamide (PZA). This contributed to the development of the multi-drug resistant TB strains (MDR-TB) which display resistance to both INH and RIF (Figure 1.1) [8]. Treatment of MDR-TB requires the use of second-line drugs, restricted in their usage due to limited efficacy, availability or increased toxicity, however resistance to these agents was not far behind as extremely drug resistant (XDR-TB) strains were soon identified displaying the properties of MDR-TB with additional resistance to second-line therapies including a fluoroquinolone and at least one injectable agent [8]. The development and accumulation of antibacterial resistance is due to a number of factors including monotherapy and poor patient compliance to multi-drug treatment regimens [1, 8]. In addition to antibiotic resistance, several other important factors have contributed to the prevalence of this disease. High population density and low socio-economic status increase the probability of contracting the disease and decrease the likelihood of successful diagnosis, prevention and treatment, leading to troublesome infection statistics in third world countries [9-10] . Exacerbating this issue, co-infection with HIV/AIDS makes patients more susceptible to *Mtb* infection, increasing infectivity and lethality [11]. Similarly, diabetes, smoking, age, gender, and ethnicity correlate with susceptibility to *Mtb* infection (Figure 1.1)[3, 12-15]. Importantly, unique aspects of the physiology of the bacterium, such as those of

dormancy and persistence, decrease the efficacy of diagnosis and treatment strategies for patients of all backgrounds.

1.2 TB LATENCY AND *MTB* DORMANCY

The physiological state of the bacterium during a latent infection is poorly understood despite the important role LTBI play in the propagation and maintenance of TB disease in the human population, which has hindered disease management strategies effective against these types of infections. Difficulties in studying the bacteria within the human host have complicated phenotypic characterization during infection, as the removal of the bacteria from the host alters the physiological state, thereby confounding any results obtained by this strategy. However, important conclusions have been drawn from studies utilizing *in vitro* stress models and *in vivo* animal infection models. Previous work has shown the bacteria enter into a dormant or non-replicating persistent (NRP) state that is characterized by reduced metabolic activity and the reduction of bacterial cell division [16-17]. Although the bacteria are not dividing some metabolic activity is essential to retain membrane potential, the loss of which is absolutely lethal to the cell, and to combat various other harsh conditions encountered within the host [17]. Importantly, these bacteria must remain viable through prolonged periods of immune suppression by the host, yet remain capable of resuming growth. This requires the coordination of complex regulatory events including transcriptional activation of dormancy regulons coupled with clearance of the translational machinery to facilitate the specific expression of metabolic and physiologic programs designed to promote survival and retain viability during extended dormancy (Figure 1.1). In addition these processes must be linked with proteomic turnover to allow the reduction and alteration of metabolism and division by the removal of proteins involved in the activation and perpetuation of these processes.

Although the average bacterial genome encodes roughly four thousand open reading frames (ORFs) only 50-75% of these are predicted to be actively expressed at any one time [18]. In response to stress or other stimuli this expression profile of the bacterium is altered, resulting in a unique and specific combination of expressed ORFs for the adaptation to specific environmental conditions [17, 19-28]. A typical global adaptive response, such as those associated with a dormant phenotype, results in an average change in expression of 7-10% of the encoded ORFs in the genome, as observed through global transcriptional profiling by the Slayden lab and others [22-24, 27, 29]. This ability of an organism to maintain a stable genotype while tailoring a specific phenotype in response to stress plays a particularly significant role in bacterial pathogenesis where it allows the bacterium to adapt to harsh or inconsistent environments thereby facilitating evasion of host immune defenses, dissemination, entry and exit from dormancy, and antibacterial tolerance (Figure 1.1)[30]. Such transcriptional changes in a genetically homogenous population have even been shown to contribute to drug resistance observed in MDR-TB strains as well as the tolerant phenotype of antibacterial persisters [31-32].

1.3 CONDITIONS ASSOCIATED WITH DORMANCY

Understanding the environment that bacteria encounter during an infection is key to elucidating the biological process associated with latency and to aid in the development of novel, effective drug treatment strategies. *Mtb* is inhaled in infected respiratory droplets and deposited within pulmonary alveoli where they are taken up by resident macrophages (Figure 1.2). Once engulfed, these bacilli persist and replicate within the phagosome by preventing phagosome-lysosome fusion [33-35]. Upon exposure, most of the bacteria will be effectively destroyed or neutralized by the innate immune response. However, few may persist and thrive at sufficient levels to establish an infection. During early disease stages, also known as the acute phase, the bacteria continue to divide rapidly [36]. Macrophages

and lymphocytes aggregate in an attempt to contain the infection, triggering an inflammatory response [2, 4]. This pathological feature is known as a granuloma and may be visualized in the lung tissue of individuals infected with *Mtb* (Figure 1.2)[2, 4]. The onset of granuloma formation signals the chronic stage of infection [2, 4]. At this time bacterial replication slows considerably, if not all together, as the bacteria enter into NRP, a dormant state capable of enduring prolonged periods of immune-mediated attack [36-37]. Early studies using mouse infection models suggested that the phagosome of alveolar

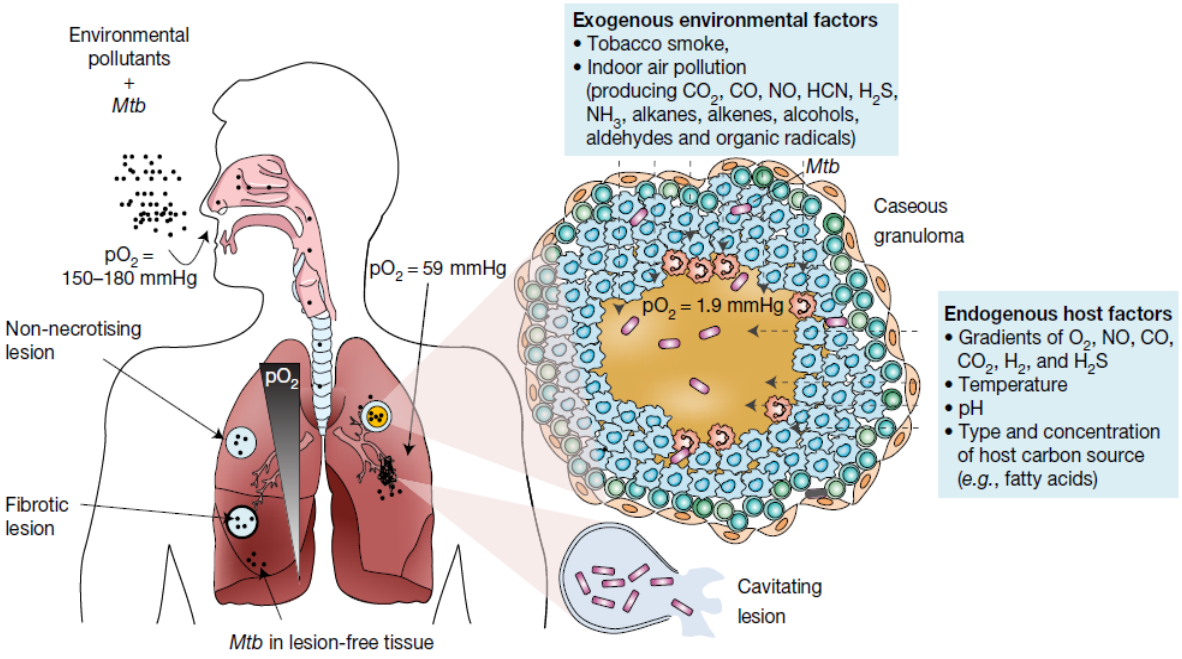

Figure 1.2: *Mtb* infection, oxygen tension and granuloma structure.

Aerosolized *Mtb* has access to abundant oxygen upon infection, however, availability becomes progressively limited as the bacteria penetrate deeper into the lungs and as immune cell infiltration establishes granuloma formation. A nearly 100-fold reduction in oxygen tension is experienced by the bacterium within advanced granulomas. Within these structures a few bacteria may be found within macrophages, but most will be located extracellularly within the necrotic core. (Reproduced with permissions from: Ashwani Kumar, Aisha Farhana, Loni Guidry, Vikram Saini, Mary Hondalus and Adrie J.C. Steyn, "Redoxhomeostasis in mycobacteria: the key to tuberculosis control?," Expert Reviews in Molecular Medicine, Volume 13, e39, December 2011, doi:10.1017/S1462399411002079)

macrophages is the primary niche for the bacterium during disease. However, a recent study using a guinea pig model, which shows more extensive similarity to human TB pathology, found that the majority of the infecting bacteria exist extracallularly within granulomas (Figure 1.2)[38]. This is likely the result of bacterial release upon cellular lysis of damaged macrophages within highly hypoxic granulomatous structures [38].

Although environmental differences exist between various pathological sites of infection, some host-associated stresses are universal. Restriction of purines, various amino acids, vitamins and metals are commonly experience by intracellular pathogens during macrophage infection [39]. Transcriptomic and proteomic profiling of mycobacteria isolated from infected tissues and sputum suggest that the availability of iron, oxygen and sugars are limited for the bacteria during infection (Figure 1.2) as reflected by the increased abundance of enzymatic pathways involved in iron acquisition, alternative respiration, hypoxic tolerance, and fatty acid metabolism associated with glucose starvation [24, 26]. TB granulomas in animal models showing similar pathology to human disease have been identified as hypoxic (Figure 1.2), supporting the role of oxygen limitation in *Mtb* infections [40]. Regulatory programs induced during oxygen limitation, such as the Dormancy (*dos*) regulon (Figure 1.1) expressed during acute hypoxic periods, show an association of molecular programs involved in alternative respiration and energy generation, nitrate reduction, iron acquisition, and chaperone functions as well as lipid and fatty acid utilization [41]. Similarly, the enduring hypoxic response (EHR) (Figure 1.1), required for survival during sustained hypoxia, utilizes programs for alternative respiration and energy generation, iron acquisition, chaperone functions and lipid utilization programs in addition to sulfur metabolism, virulence factors, and detoxification programs [23]. Furthermore, comparisons of transcriptional profiles obtained from infected tissues with *in vitro* stress models limited in oxygen, carbohydrates or essential nutrients reveal significant overlap in adaptive strategies utilized by the bacterium [21-23, 25, 27-28]. These studies illuminate the metabolic programs important for survival

under conditions associated with dormancy and demonstrate an association between anaerobiosis (respiration upon severe oxygen limitation), nutrient limitation, alternative metabolism, redox homeostasis, and antibacterial tolerance (Figures 1.1 and 1.2)[19].

The study of live bacilli within infected tissue has been a troublesome endeavor, but was alleviated by the development of the first defined *in vitro* model of *Mtb* dormancy by Wayne and colleagues [42-43]. The Wayne model employs a gradual depletion of dissolved oxygen by reduction of headspace in a fixed volume of slowly stirred liquid culture. Under such conditions early growth occurs at the standard logarithmic rate, however, a nearly 100-fold reduction in available oxygen coincides with greatly reduced growth rate and morphological alterations [16]. Further reduction of oxygen levels below 0.06% saturation induces transition into a second and more prolonged stage of dormancy analogous to the maintained stage associated with chronic, latent infections [16]. These studies provided valuable insight on the transition of a metabolically active, dividing bacterial population into a dormant phenotype. Importantly, it appears that time is required to prime this bacterial response as an abrupt reduction of oxygen to 0.06% yields reduced survival compared to a gradual reduction [16]. Similarly, the resumption of growth from a dormant state is a progressive process further supporting the notion that the bacteria require time for molecular remodeling during transitions into and out of dormancy (Figure 1.1)[42]. Observations of *Mtb* growth within macrophages indicates that the bacteria are capable of replicating normally during the development of foamy macrophages, a lipid-loaded cell type characteristic of hypoxic granulomas, but cease replication once this maturation is complete, providing a pathological link to the physiological conditions associated with dormancy [40, 44]. Furthermore, cessation of bacterial growth by nutrient or oxygen limitation, entry into stationary phase, or macrophage stimulation induces central metabolic reprogram favoring the utilization of fatty acids over sugar carbon sources, implicating bacterial growth rate as important dormancy factor and highlighting the dynamic interplay between these processes during a latent infection [45].

1.4 BACTERIAL REPLICATION DURING DORMANCY

In bacterial cytokinesis, early division involves replication and segregation of the chromosomes, along with the localization of early-recruitment factors to mid-cell [46-47]. These early division proteins recruit other necessary proteins and provide a scaffold for the formation of the divisome, a multi-protein complex capable of performing the functions necessary for replication [46-47]. FtsZ, a prokaryotic homolog of the eukaryotic protein tubulin, is the most widely conserved division gene across bacterial taxa; present in all but a few select species [48-49]. Under permitting conditions, FtsZ proteins polymerize into protofilaments energized by the hydrolysis of GTP to GDP, which then laterally associate to form protofilament bundles that assemble into a structure called the Z-ring at mid-cell [48, 50-51]. The Z-ring acts as a scaffold for peptidoglycan-synthesizing enzymes that will form the septum dividing the two daughter cells and may also provide the mechanical force required for constriction of the cell wall at the septum [46-47]. This activity is essential for division in diverse bacteria including the mycobacteria [52-55]. Formation of the Z-ring is energetically expensive and has been recognized as the first committed step of the division process; marking the separation of early division events from late division events [46-47, 50]. Once FtsZ polymerization is permitted, the divisome can act in late-division processes and proceed quickly through formation of the Z-ring, septal wall synthesis and daughter cell resolution [46-47].

Replication can be slowed by *Mtb* in response to host-associated stresses, as demonstrated by several *in vitro* models that induce dormancy via exposure to oxidizing agents, oxygen deprivation or nutrient deprivation, supporting a role for slow growth as an adaptive response during infection [56]. Importantly, reduction or cessation of replication is a common adaptive strategy for diverse organisms to protect the cell and retain viability under stressful conditions. For example, slow growing bacteria and non-replicating bacteria in particular, show an increased capacity to withstand sub-lethal concentrations

of antimicrobial agents; a phenomenon termed antibacterial tolerance or persistence [20]. Antibiotic tolerance resulting from a subpopulation displaying a tolerant phenotype despite a susceptible genotype may result from simple metabolic or transcriptional shifts. However, this phenomenon of antibiotic tolerance is likely a secondary effect of the dormant phenotype induced to survive host-associated stresses during infection. Indeed, *Mtb* grown under various models of dormancy displays increased antibacterial tolerance [19-20, 57]. Furthermore, observations of *Mtb* under the Wayne-model of dormancy show completely halted division where the bacilli become elongated, displaying duplicated, segregated chromosomes but no obvious septa, similar to phenotypic observation of bacilli grown within macrophages, indicating that a major regulatory event occur at the level of septation in the transition to dormancy [16, 58].

1.5 FTSZ REGULATION

Filamentation, defined by bacterial elongation in the absence of division, is an adaptive survival strategy for many bacterial pathogens and provides a number of advantages including the prevention of phagocytosis, evasion of innate immune responses, as well as increased resistance to heat, oxidative damage, and antibacterial activity [59]. Such filamentation may result from alterations in cell division protein stoichiometery, or by purposeful septum inhibition either through direct regulation of FtsZ activity or reduction of FtsZ levels in the cell [60-63]. Recent studies have shown that FtsZ levels are reduced during hypoxia and oxidative stress indicating transcriptional control of FtsZ gene expression may be a contributing factor during growth changes, however, it is unlikely to be the only mechanism regulating the activity of FtsZ in the cell [62-63]. It is apparent that under normal resting conditions FtsZ protein levels far exceed those required for spontaneous FtsZ polymerization in the cytoplasm, and that only a fraction of this population are actively involved in formation of the Z-ring [50, 64-65]. Measurable

levels of FtsZ protein are present in bacilli within lung granulomas in the guinea pig model of *Mtb* infection, and comparable levels of FtsZ protein are observed both in bacilli grown in standard media or within macrophages [58, 66]. Importantly, as previously described, *Mtb* observed during dormancy and within macrophages are elongated and devoid of Z-rings, indicating that the transition into non-replicating persistence occurs prior to FtsZ polymerization. Given the importance of FtsZ activity in the cytokinetic process and the steep energetic price of polymerization, it is not surprising that this protein is subject to extensive regulation (Figure 1.3). However, despite the numerous molecular programs identified participating in FtsZ polymerization control in model organisms, few have been found within the mycobacterial genomes [29, 55].

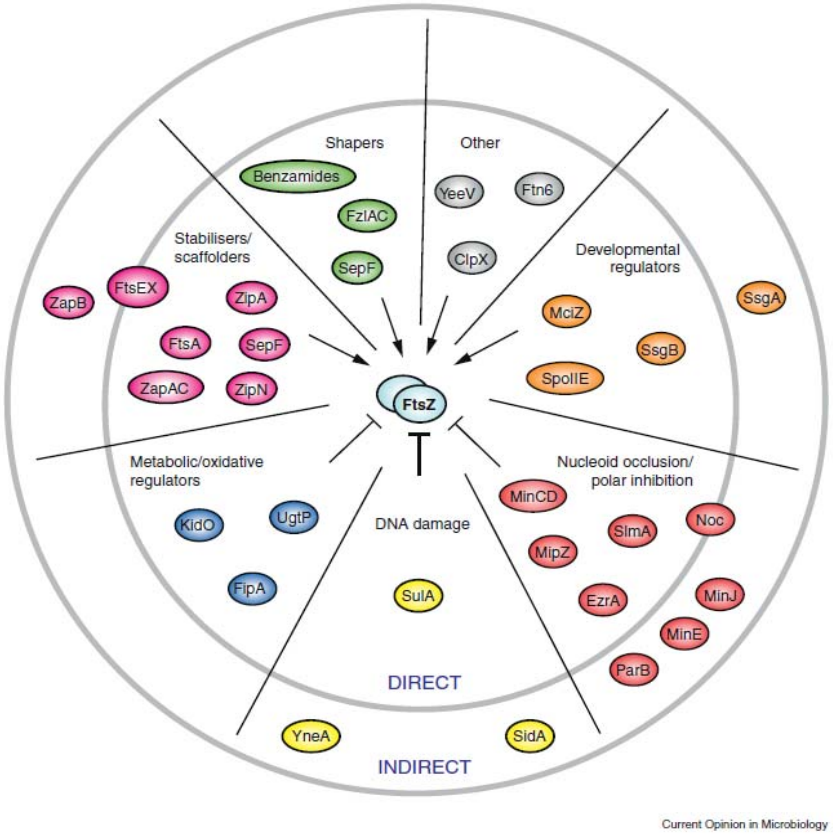

Figure 1.3: FtsZ regulating proteins identified in bacteria.

A list of post-translational FtsZ regulating proteins identified in bacteria as of 2011 sorted by regulation type, interaction, and functional class. (Reprinted from: Clare L Kirkpatrick, Patrick H Viollier, "New(s) to the (Z-)ring", Current Opinion in Microbiology, Volume 14, Issue 6, December 2011, Pages 691-697, ISSN 1369-5274, http://dx.doi.org/10.1016/j.mib.2011.09.011 with permissions from Elsevier)

Lessons on bacterial division in the model organisms *Escherichia coli*, *Bacillus subtilis*, and more recently *Caulobacter crescentus*, have shown that FtsZ polymerization is actively regulated during different steps of division and differentiation as well as in response to metabolic and oxidative stresses (Figure 1.3). A number of FtsZ regulators have been identified and show significant conservation across bacterial taxa, such as FtsA and the FtsEX complex [46, 50]. However the presence or absence of these regulators varies widely across species suggesting a robust diversity of regulatory strategies. Few of the currently known FtsZ regulators have been identified in *Mtb* by bioinformatic-based homology searches. Among those proteins surprisingly absent in mycobacteria is FtsA, the second most well-conserved division protein [46, 50]. FtsA plays a vital role in the recruitment of early divisome components and promoting FtsZ polymerization [46, 50]. The only known polymerization-promoting factor encoded by mycobacterial genomes is SepF, recently identified and characterized in *B. subtilis* as a conserved regulator of early division in gram-positive bacterium [67]. Similarities between the activities of FtsA and SepF proteins provide a possible explanation for how replication in *Mtb* proceeds in the absence of this central regulator, yet it remains unclear if SepF can act alone to coordinate early division-associated recruitment or if it acts in concert with an unidentified molecular program [68]. A novel regulatory strategy was recently discovered in *Mtb*, in which ClpX protease binding to FtsZ proteins appears to negatively affect FtsZ polymerization dynamics by a sequestration or inactivation mechanism rather than proteolytic activity [69].

Importantly, bioinformatic analysis has failed to identify effector proteins involved in two vital regulatory events: correct placement of the septum at mid-cell, and DNA damage-induced replicative cessation, also known as the SOS response (Figure 1.3). Correct localization of the septum to mid-cell is vital to ensure the equal distribution of cytoplasmic contents to the daughter cells and to maintain the integrity of the chromosome by preventing cleavage of the nucleoids. Nuceloid occlusion proteins in the model organisms *E. coli* and *B. subtilis* localize to the chromosome by binding double-stranded DNA

where they act to inhibit FtsZ polymerization in close proximity to the chromosome, thereby averting cleavage due to aberrant cinching. Neither the gram-negative effector protein Noc, nor the gram-positive effector SlmA has been identified in any actinomycete. Another regulatory strategy to ensure proper Z-ring localization is the Min system, a molecular program that promotes Z-ring formation at mid-cell by preventing FtsZ polymerization elsewhere [46, 50]. The inhibition of FtsZ polymerization is carried out by MinC which is activated by interactions with its partner protein MinD [70]. MinD also interacts with the localization factors MinE in gram-negative bacteria, or DivIV in gram-positive bacteria, to establish a concentration gradient of MinC inhibitor that is greatest at the poles and least at mid-cell [70-71]. Alternatively, MipZ plays a comparable role to MinCD in *C. cresenctus*, where this complex is localized by ParB, a nucleoid partitioning protein [72]. Mycobacteria encode for both localization factors DivIV, which has been shown to play a role in peptidoglycan synthesis, and ParB, which acts in chromosome partitioning [73-74]. However, no protein directly involved in regulating FtsZ activity has yet been identified for these important molecular programs. This raises the question: how do mycobacteria regulate FtsZ protein dynamics to maintain these vital cellular processes, and how does such regulation contribute to the development of a non-replicating persistent state?

REFERENCES

1. Organization, W.H., *Global Tuberculosis Report 2012. WHO, Geneva, Switzerland*. 2012, WHO/HTM/TB.
2. Barry, C.E., 3rd, et al., *The spectrum of latent tuberculosis: rethinking the biology and intervention strategies.* Nat Rev Microbiol, 2009. **7**(12): p. 845-55.
3. Renner, L.D. and D.B. Weibel, *Cardiolipin microdomains localize to negatively curved regions of Escherichia coli membranes.* Proc Natl Acad Sci U S A, 2011. **108**(15): p. 6264-9.
4. Russell, D.G., C.E. Barry, 3rd, and J.L. Flynn, *Tuberculosis: what we don't know can, and does, hurt us.* Science, 2010. **328**(5980): p. 852-6.
5. Hershkovitz, I., et al., *Detection and molecular characterization of 9,000-year-old Mycobacterium tuberculosis from a Neolithic settlement in the Eastern Mediterranean.* PLoS One, 2008. **3**(10): p. e3426.
6. Singh, B. and D.A. Mitchison, *Bactericidal activity of streptomycin and isoniazid against tubercle bacilli.* Br Med J, 1954. **1**(4854): p. 130-2.
7. *Tuberculosis Fact Sheet*, NIH, Editor. 2010. p. 2.
8. Chan, E.D. and M.D. Iseman, *Multidrug-resistant and extensively drug-resistant tuberculosis: a review.* Curr Opin Infect Dis, 2008. **21**(6): p. 587-95.
9. Zhang, T., et al., *Persistent problems of access to appropriate, affordable TB services in rural China: experiences of different socio-economic groups.* BMC Public Health, 2007. **7**: p. 19.
10. Mishra, P., et al., *Socio-economic status and adherence to tuberculosis treatment: a case-control study in a district of Nepal.* Int J Tuberc Lung Dis, 2005. **9**(10): p. 1134-9.
11. Corbett, E.L., et al., *The growing burden of tuberculosis: global trends and interactions with the HIV epidemic.* Arch Intern Med, 2003. **163**(9): p. 1009-21.
12. Alisjahbana, B., et al., *The effect of type 2 diabetes mellitus on the presentation and treatment response of pulmonary tuberculosis.* Clin Infect Dis, 2007. **45**(4): p. 428-35.
13. Alcaide, J., et al., *Cigarette smoking as a risk factor for tuberculosis in young adults: a case-control study.* Tuber Lung Dis, 1996. **77**(2): p. 112-6.
14. Connolly, M. and P. Nunn, *Women and tuberculosis.* World Health Stat Q, 1996. **49**(2): p. 115-9.
15. Pareek, M., et al., *Ethnicity and mycobacterial lineage as determinants of tuberculosis disease phenotype.* Thorax, 2012.
16. Wayne, L.G. and C.D. Sohaskey, *Nonreplicating persistence of mycobacterium tuberculosis.* Annu Rev Microbiol, 2001. **55**: p. 139-63.
17. Talaat, A.M., et al., *Mycobacterial bacilli are metabolically active during chronic tuberculosis in murine lungs: insights from genome-wide transcriptional profiling.* J Bacteriol, 2007. **189**(11): p. 4265-74.
18. Passalacqua, K.D., et al., *Structure and complexity of a bacterial transcriptome.* J Bacteriol, 2009. **191**(10): p. 3203-11.
19. Bartek, I.L., et al., *The DosR regulon of M. tuberculosis and antibacterial tolerance.* Tuberculosis (Edinb), 2009. **89**(4): p. 310-6.
20. Deb, C., et al., *A novel in vitro multiple-stress dormancy model for Mycobacterium tuberculosis generates a lipid-loaded, drug-tolerant, dormant pathogen.* PLoS One, 2009. **4**(6): p. e6077.
21. Hampshire, T., et al., *Stationary phase gene expression of Mycobacterium tuberculosis following a progressive nutrient depletion: a model for persistent organisms?* Tuberculosis (Edinb), 2004. **84**(3-4): p. 228-38.
22. Muttucumaru, D.G., et al., *Gene expression profile of Mycobacterium tuberculosis in a non-replicating state.* Tuberculosis (Edinb), 2004. **84**(3-4): p. 239-46.

23. Rustad, T.R., et al., *The enduring hypoxic response of Mycobacterium tuberculosis.* PLoS ONE, 2008. **3**(1): p. e1502.
24. Schnappinger, D., et al., *Transcriptional Adaptation of Mycobacterium tuberculosis within Macrophages: Insights into the Phagosomal Environment.* J Exp Med, 2003. **198**(5): p. 693-704.
25. Taneja, N.K., et al., *Mycobacterium tuberculosis transcriptional adaptation, growth arrest and dormancy phenotype development is triggered by vitamin C.* PLoS One, 2010. **5**(5): p. e10860.
26. Timm, J., et al., *Differential expression of iron-, carbon-, and oxygen-responsive mycobacterial genes in the lungs of chronically infected mice and tuberculosis patients.* Proc Natl Acad Sci U S A, 2003. **100**(24): p. 14321-6.
27. Voskuil, M.I., K.C. Visconti, and G.K. Schoolnik, *Mycobacterium tuberculosis gene expression during adaptation to stationary phase and low-oxygen dormancy.* Tuberculosis (Edinb), 2004. **84**(3-4): p. 218-27.
28. Ward, S.K., et al., *Transcriptional profiling of mycobacterium tuberculosis during infection: lessons learned.* Front Microbiol, 2010. **1**: p. 121.
29. England, K., R. Crew, and R.A. Slayden, *Mycobacterium tuberculosis septum site determining protein, Ssd encoded by rv3660c, promotes filamentation and elicits an alternative metabolic and dormancy stress response.* BMC Microbiol, 2011. **11**: p. 79.
30. Rainey, P.B., et al., *The evolutionary emergence of stochastic phenotype switching in bacteria.* Microb Cell Fact, 2011. **10 Suppl 1**: p. S14.
31. Chatterjee, A., et al., *Global Transcriptional Profiling of Longitudinal Clinical Isolates of Mycobacterium tuberculosis Exhibiting Rapid Accumulation of Drug Resistance.* PLoS One, 2013. **8**(1): p. e54717.
32. Keren, I., et al., *Characterization and transcriptome analysis of Mycobacterium tuberculosis persisters.* MBio, 2011. **2**(3): p. e00100-11.
33. Goren, M.B., et al., *Prevention of phagosome-lysosome fusion in cultured macrophages by sulfatides of Mycobacterium tuberculosis.* Proc Natl Acad Sci U S A, 1976. **73**(7): p. 2510-4.
34. Malik, Z.A., S.S. Iyer, and D.J. Kusner, *Mycobacterium tuberculosis phagosomes exhibit altered calmodulin-dependent signal transduction: contribution to inhibition of phagosome-lysosome fusion and intracellular survival in human macrophages.* J Immunol, 2001. **166**(5): p. 3392-401.
35. Meena, L.S. and Rajni, *Survival mechanisms of pathogenic Mycobacterium tuberculosis H37Rv.* FEBS J, 2010. **277**(11): p. 2416-27.
36. McKinney, J.D., et al., *Persistence of Mycobacterium tuberculosis in macrophages and mice requires the glyoxylate shunt enzyme isocitrate lyase.* Nature, 2000. **406**(6797): p. 735-8.
37. Saunders, B.M., A.A. Frank, and I.M. Orme, *Granuloma formation is required to contain bacillus growth and delay mortality in mice chronically infected with Mycobacterium tuberculosis.* Immunology, 1999. **98**(3): p. 324-8.
38. Hoff, D.R., et al., *Location of intra- and extracellular M. tuberculosis populations in lungs of mice and guinea pigs during disease progression and after drug treatment.* PLoS One, 2011. **6**(3): p. e17550.
39. Appelberg, R., *Macrophage nutriprive antimicrobial mechanisms.* J Leukoc Biol, 2006. **79**(6): p. 1117-28.
40. Via, L.E., et al., *Tuberculous granulomas are hypoxic in guinea pigs, rabbits, and nonhuman primates.* Infect Immun, 2008. **76**(6): p. 2333-40.
41. Park, H.D., et al., *Rv3133c/dosR is a transcription factor that mediates the hypoxic response of Mycobacterium tuberculosis.* Mol Microbiol, 2003. **48**(3): p. 833-43.
42. Wayne, L.G., *Dormancy of Mycobacterium tuberculosis and latency of disease.* Eur J Clin Microbiol Infect Dis, 1994. **13**(11): p. 908-14.

43. Wayne, L.G. and L.G. Hayes, *An in vitro model for sequential study of shiftdown of Mycobacterium tuberculosis through two stages of nonreplicating persistence.* Infect Immun, 1996. **64**(6): p. 2062-9.
44. Peyron, P., et al., *Foamy macrophages from tuberculous patients' granulomas constitute a nutrient-rich reservoir for M. tuberculosis persistence.* PLoS Pathog, 2008. **4**(11): p. e1000204.
45. Shi, L., et al., *Carbon flux rerouting during Mycobacterium tuberculosis growth arrest.* Mol Microbiol, 2010. **78**(5): p. 1199-215.
46. Harry, E., L. Monahan, and L. Thompson, *Bacterial cell division: the mechanism and its precison.* Int Rev Cytol, 2006. **253**: p. 27-94.
47. Hett, E.C. and E.J. Rubin, *Bacterial growth and cell division: a mycobacterial perspective.* Microbiol Mol Biol Rev, 2008. **72**(1): p. 126-56, table of contents.
48. Erickson, H.P., et al., *Bacterial cell division protein FtsZ assembles into protofilament sheets and minirings, structural homologs of tubulin polymers.* Proc Natl Acad Sci U S A, 1996. **93**(1): p. 519-23.
49. Vaughan, S., et al., *Molecular evolution of FtsZ protein sequences encoded within the genomes of archaea, bacteria, and eukaryota.* J Mol Evol, 2004. **58**(1): p. 19-29.
50. Adams, D.W. and J. Errington, *Bacterial cell division: assembly, maintenance and disassembly of the Z ring.* Nat Rev Microbiol, 2009. **7**(9): p. 642-53.
51. Dajkovic, A., et al., *Cross-linking FtsZ polymers into coherent Z rings.* Mol Microbiol, 2010. **78**(3): p. 651-68.
52. Dai, K. and J. Lutkenhaus, *ftsZ is an essential cell division gene in Escherichia coli.* J Bacteriol, 1991. **173**(11): p. 3500-6.
53. Beall, B. and J. Lutkenhaus, *FtsZ in Bacillus subtilis is required for vegetative septation and for asymmetric septation during sporulation.* Genes Dev, 1991. **5**(3): p. 447-55.
54. Dziadek, J., et al., *Conditional expression of Mycobacterium smegmatis ftsZ, an essential cell division gene.* Microbiology, 2003. **149**(Pt 6): p. 1593-603.
55. Slayden, R.A., D.L. Knudson, and J.T. Belisle, *Identification of cell cycle regulators in Mycobacterium tuberculosis by inhibition of septum formation and global transcriptional analysis.* Microbiology, 2006. **152**(Pt 6): p. 1789-97.
56. Beste, D.J., et al., *Transcriptomic analysis identifies growth rate modulation as a component of the adaptation of mycobacteria to survival inside the macrophage.* J Bacteriol, 2007. **189**(11): p. 3969-76.
57. Hu, Y., A.R. Coates, and D.A. Mitchison, *Sterilizing activities of fluoroquinolones against rifampin-tolerant populations of Mycobacterium tuberculosis.* Antimicrob Agents Chemother, 2003. **47**(2): p. 653-7.
58. Chauhan, A., et al., *Mycobacterium tuberculosis cells growing in macrophages are filamentous and deficient in FtsZ rings.* J Bacteriol, 2006. **188**(5): p. 1856-65.
59. Justice, S.S., et al., *Morphological plasticity as a bacterial survival strategy.* Nat Rev Microbiol, 2008. **6**(2): p. 162-8.
60. Sureka, K., et al., *Novel role of phosphorylation-dependent interaction between FtsZ and FipA in mycobacterial cell division.* PLoS One, 2010. **5**(1): p. e8590.
61. Dziadek, J., et al., *Physiological consequences associated with overproduction of Mycobacterium tuberculosis FtsZ in mycobacterial hosts.* Microbiology, 2002. **148**(Pt 4): p. 961-71.
62. Roy, S. and P. Ajitkumar, *Transcriptional analysis of the principal cell division gene, ftsZ, of Mycobacterium tuberculosis.* J Bacteriol, 2005. **187**(7): p. 2540-50.
63. Roy, S., et al., *The ftsZ Gene of Mycobacterium smegmatis is expressed Through Multiple Transcripts.* Open Microbiol J, 2011. **5**: p. 43-53.

64. Kiran, M., et al., *Mycobacterium tuberculosis ftsZ expression and minimal promoter activity.* Tuberculosis (Edinb), 2009. **89 Suppl 1**: p. S60-4.
65. Chen, Y. and H.P. Erickson, *Rapid in vitro assembly dynamics and subunit turnover of FtsZ demonstrated by fluorescence resonance energy transfer.* J Biol Chem, 2005. **280**(23): p. 22549-54.
66. Sharma, D., et al., *Expression of mycobacterial cell division protein, FtsZ, and dormancy proteins, DevR and Acr, within lung granulomas throughout guinea pig infection.* FEMS Immunol Med Microbiol, 2006. **48**(3): p. 329-36.
67. Hamoen, L.W., et al., *SepF, a novel FtsZ-interacting protein required for a late step in cell division.* Mol Microbiol, 2006. **59**(3): p. 989-99.
68. Singh, J.K., et al., *SepF increases the assembly and bundling of FtsZ polymers and stabilizes FtsZ protofilaments by binding along its length.* J Biol Chem, 2008. **283**(45): p. 31116-24.
69. Dziedzic, R., et al., *Mycobacterium tuberculosis ClpX interacts with FtsZ and interferes with FtsZ assembly.* PLoS One, 2010. **5**(7): p. e11058.
70. Gregory, J.A., E.C. Becker, and K. Pogliano, *Bacillus subtilis MinC destabilizes FtsZ-rings at new cell poles and contributes to the timing of cell division.* Genes Dev, 2008. **22**(24): p. 3475-88.
71. Park, K.T., et al., *The Min oscillator uses MinD-dependent conformational changes in MinE to spatially regulate cytokinesis.* Cell, 2011. **146**(3): p. 396-407.
72. Thanbichler, M. and L. Shapiro, *MipZ, a spatial regulator coordinating chromosome segregation with cell division in Caulobacter.* Cell, 2006. **126**(1): p. 147-62.
73. Donovan, C., et al., *A synthetic Escherichia coli system identifies a conserved origin tethering factor in Actinobacteria.* Mol Microbiol, 2012. **84**(1): p. 105-16.
74. Chaudhuri, B.N., et al., *A combined global and local approach to elucidate spatial organization of the Mycobacterial ParB-parS partition assembly.* Biochemistry, 2011. **50**(11): p. 1799-807.

CHAPTER 2: IDENTIFICTION OF DIVISION REGULATORS IN *MTB*

The bioinformatic identification and protein classification presented in this work was performed by Rebecca Crew as presented in the manuscript "*Mycobacterium tuberculosis* septum site determining protein, Ssd encoded by *rv3660c*, promotes filamentation and elicits an alternative metabolic and dormancy stress response" published in BMC Microbiology, 2011 (11). Morpholigical and transcriptional analysis was performed by K. England.

2.1 SEARCHING FOR A MinD HOMOLOG

Searches for putative FtsZ regulators identified *rv3660c* and *rv1708* as encoding putative MinD homologs with similarity to our MinD consensus model. In model bacteria, such as *B. subtilis* and *E. coli*, MinD binds and activates MinC, a FtsZ polymerization inhibiting protein, and aids in the correct localization of this protein through interactions with membrane-associated localization proteins [1]. In *B. subtilis*, MinD localization is dependent upon the membrane-bound DivIV protein, which localizes statically to the poles, while in *E. coli* MinD is localized by the soluble, membrane-associated oscillating protein, MinE [1-3]. Regardless of the terminal localization factor, these systems utilize the same basic mechanism: the establishment of a concentration gradient of FtsZ inhibitor that is greatest at the poles and weakest at mid-cell [1-3]. This helps ensure the correct placement of the septum at mid-cell and prevents the production of Z-rings at aberrant locations which may result in cleavage of the nucleoids and subsequent loss of viability. On the assumption that mycobacteria must possess a molecular mechanism to ensure correct septum placement, we set out to identify homologs of regulating proteins in *Mtb*.

2.2 PRELIMINARY WORK WITH *rv3660c*

Bioinformatic analysis was used to elucidate the functional role of one such putative FtsZ regulator, encoded by *rv3660c*. Preliminary screens using a MinD consensus model generated by a global multiple alignment of annotated MinD proteins pooled from the NCBI Reference Sequence (RefSeq) database identified *rv3660c* as encoding a protein with similarity to MinD in *Mtb* [4]. However, protein functional domain mapping revealed a divergent organization from MinD proteins suggesting a dissimilar functionality. This deduction was further supported by the orthological classification of Rv3660c separate from annotated MinD proteins within the Orthologous Matrix (OMA) Browser database [4]. Instead, Rv3660c is grouped with a set of conserved yet uncharacterized proteins of diverse bacterial origin, defined by OMA group 73337, putatively annotated as septum-site determining proteins [4]. While this annotation seems to indicate that OMA 73337 proteins act as MinD proteins in the role of septal placement, experimental evidence is lacking. Comparisons of the MinD consensus and OMA 73337 consensus sequences with Rv3660c show more extensive similarity in amino acid sequence, size and domain organization within its own OMA group (Figure 2.1) [4]. Furthermore, hierarchical clustering analysis of this dataset grouped Rv3660c with OMA 73337 septum site determining proteins, separate from annotated MinD proteins, indicating that while Rv3660c shows some similarity to MinD, it likely descended from a different ancestral protein and therefore may be involved in different processes or act through an alternative mechanism within the cell [4].

Previous work from the Slayden laboratory has shown a potential involvement with the process of septum formation and placement in both *M. smegmatis* and *Mtb* in that overexpression of *rv3660c* resulted in filamentation (*Mtb* wild type 1.73 ± 0.43 µm compared to 2.53 ± 0.76 µm in the merodiploid) with a decreased frequency of Z-rings, while knockout of *rv3660c* caused mini-cell formation (1.35 ± 0.51 µm) [4]. Furthermore, transcriptional analysis of *Mtb* overexpressing *rv3660c* showed significant

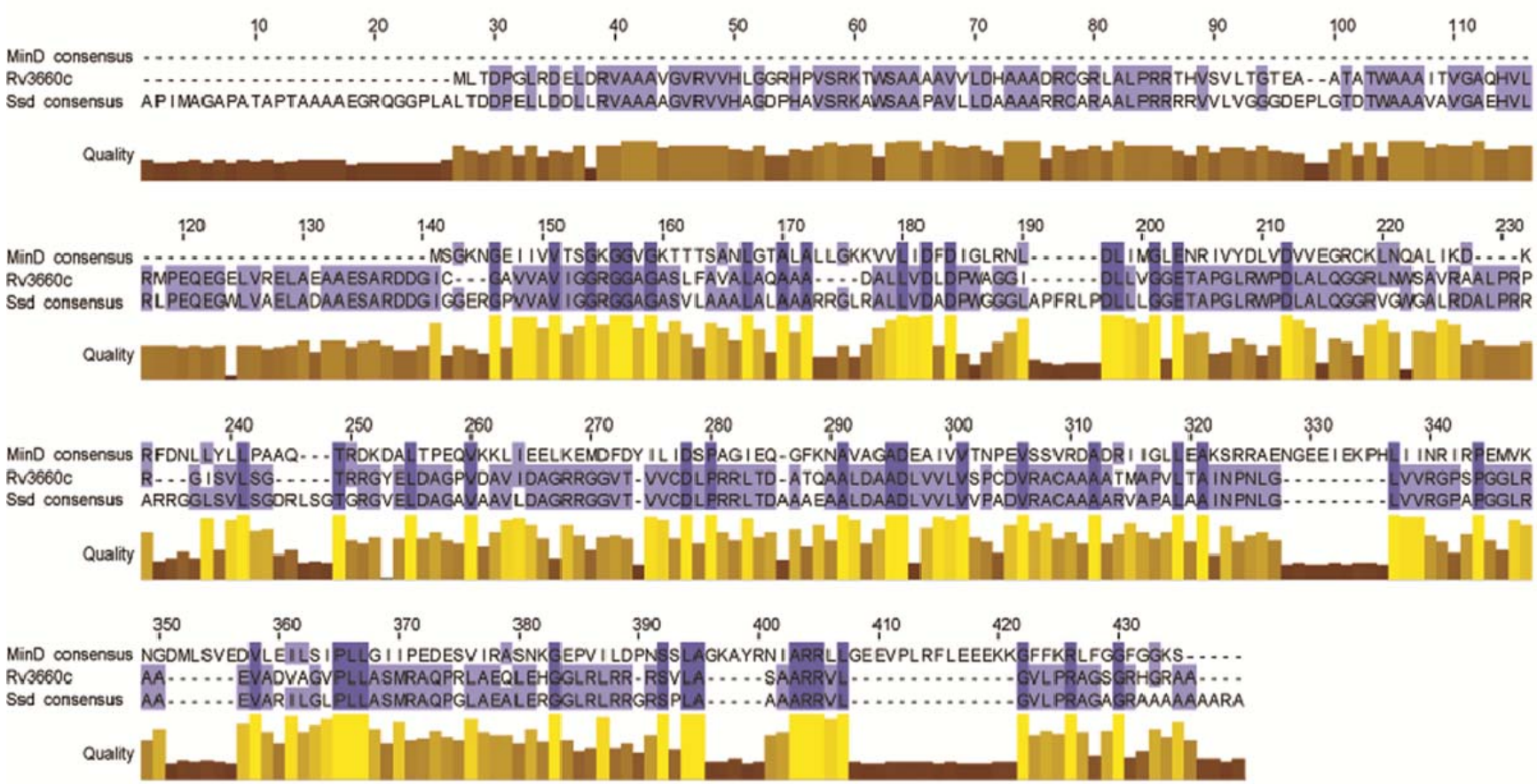

Figure 2.1: *Rv3660c* encodes an Ssd protein in *Mtb*.

Multiple sequence alignment of Rv3660c protein sequence with Ssd (OMA group 73337) and MinD consensus models reveals more extensive similarity within its own protein family. Blue boxes highlight conserved residues with darker shades indicating greater conservation.

induction of dormancy and stress programs including *dosR*-dependent genes and alternative sigma factors, supporting a potential role for Rv3660c in a dormancy-related growth state [4]. This work provides evidence of a regulatory program linking cell division with adaptive programs in *Mtb* that may contribute to the development of a non-replicating persistent phenotype during the transition into dormancy.

2.3 IDENTIFICATION OF *rv1708*

Homology searches identified a second putative MinD homolog encoded by *rv1708* showing 20% similarity, 80% coverage with our MinD consensus model. Although similar searches were unable to identify MinC or MinE homologs encoded in *Mtb*, a DivIV homolog is known to exist in mycobacteria [5]. Assumptions that a septum placement system must exist in *Mtb* led us to reduce the stringency cut-off for protein homology to 20%, which still falls within the 20-30% similarity range of orthologous proteins [6-7]. This also increases the likelihood of false-positive results, therefore further investigations into the nature of Rv1708 were undertaken. Functional domain mapping of Rv1708 revealed a single-domain organization consisting of a Walker-type ATPase domain accounting for approximately 68% of the coding sequence, and an unmapped N-terminal extension. Further characterization using Phyre2 revealed a small domain near the N-terminus that displays weak similarity to sulfurylase proteins, which reduce inorganic sulfate into a form suitable for biosynthesis of the sulfur-containing amino acids cysteine and methionine. However, the role of this domain in Rv1708 remains undefined [8]. Orthology analysis did not differentiate Rv1708 from other ATPase family proteins; grouping Rv1708 along with over 100 other experimentally un-characterized proteins, belonging to OMA group 95163, annotated as ParA chromosomal partitioning proteins, Soj ParA-like initiation inhibitor proteins, MinD septum placement proteins or CbiA cobyrinic acid ac-diamide synthase proteins. This is likely due to the fact that

these different proteins share the same Walker-type ATPase domain that comprises a majority of their protein structure [9-10]. This domain generates energy via ATP hydrolysis which is utilized to elicit a conformational change that alters the proteins interactions with other interaction partners or specific substrates [10]. In the case of MinD, ATP hydrolysis alters its binding affinity for MinC, ParB for ParA, Spo0J for Soj, and the cobyrinic acid substrate for CbiA [11-12]. Unfortunately, simple similarity comparisons are not sufficient to differentiate these proteins, therefore requiring more specific investigations.

2.4 *Rv1708* ENCODES A SOJ-LIKE PROTEIN IN *Mtb*

The MinD protein of *E. coli* has been shown to contain two regions essential activity, termed switch I and switch II, that act together to specifically activate MinC inhibitory activity [13]. Comparisons of these regions with the Rv1708 protein reveal an absence of these regions indicating that Rv1708 is unable to perform the functions characteristic of a MinD protein. Furthermore, MinD proteins are known to possess an aliphatic helix at the C-terminus required for association with membrane phospholipids, and thus correct sub-cellular localization [14]. Comparisons reveal that Rv1708 does not possess such a domain at the C-terminus, further supporting the conclusion that Rv1708 does not function as a MinD (Figure 2.2).

CbiAs, or cobyrinic acid a,c-diamide synthases, are involved in the amidation of cobyrinic acid during the synthesis of vitamin B12 in bacteria [12, 15]. These proteins are larger than most of the other Walker-type ATPases to enhance capacity for substrate binding [12]. Characterization of the CbiA of *Salmonella typhimurium* LT2 identified conserved residues, specifically D97, involved in substrate recognition and binding that are specific to CbiA enzymes compared to other similar type ATPases [12].

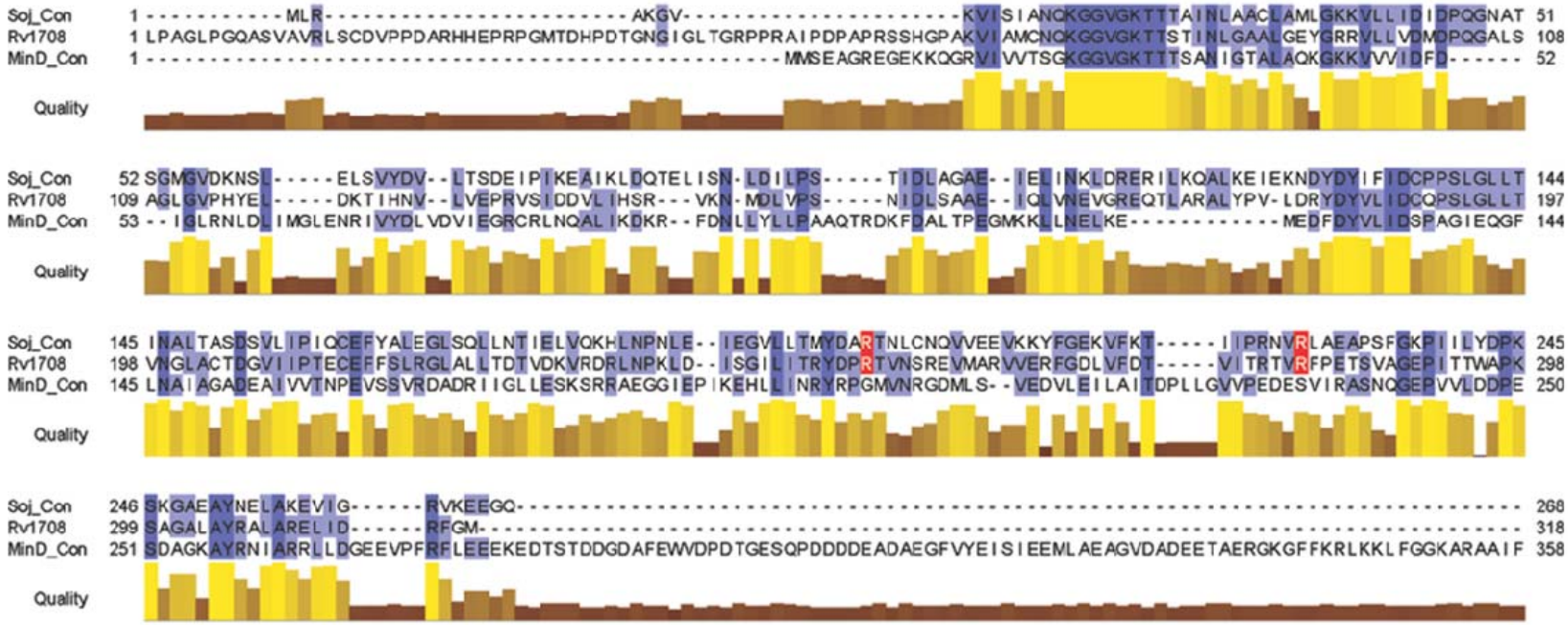

Figure 2.2: R*v1708* encodes a Soj-like protein in *Mtb*.

Local multialignment of Rv1708 with Soj and MinD consensus models. Blue Boxes indicate conserved residues, with darker colors denoting greater conservation. Red boxes indicate Soj-specific residues involved in non-specific DNA binding.

Comparisons of Rv1708 with LT2 CbiA indicate that Rv1708 is unlikely to interact with cobyrinic acid as it lacks this vital residue, indicating that Rv1708 is not a CbiA protein in *Mtb*.

Both ParA and Soj lack the aliphatic helix of MinD proteins and instead display somewhat truncated C-termini with negatively charged amino acids involved in non-specific DNA binding (Figure 2.2). Furthermore, Rv1708 shares more than two-times the percent identity with the ParA/Soj (OMA group 95838) consensus model, 46% similarity with 78% coverage, than with MinD (OMA group 78690) consensus, 20% similarity and 80% coverage. This supports the assignment of Rv1708 as encoding a ParA/Soj-like protein in *Mtb*.

At this juncture it would be pertinent to note that there are two types of chromosomally encoded ParA genes. Type Ia possess DNA binding domains that recognize specific DNA sequences which are used for transcriptional control of plasmid-encoded *parAB* operons but may also be found within chromosomes. Alternatively, Type Ib ParA genes display non-specific DNA-binding via negatively charged residue interaction with the positively charged DNA phosphate backbone [16-17]. Functionally, ParA exhibits non-specific DNA binding which localizes it to the chromosome until it encounters its interaction partner ParB, after which the complex acts in concert to ensure proper chromosomal segregation into the forming daughter cells during division [18]. The activities of Soj are more complex, in that Soj has been shown to regulate DnaA activity, thus controlling the timing of cell division initiation. This activity is regulated both by interaction with Spo0J (equivalent to ParB) and ATP [19-20]. Interestingly, Soj has also been implicated in the transcriptional control of sporulation-specific genes in *B. subtilis* through activational control of the checkpoint protein Sda, but has been shown to be dispensable for correct chromosomal segregation during division [20-21]. Comparisons of Rv1708 with these proteins revealed the potential for non-specific DNA binding by Rv1708 through charge-charge interactions, similar to Type Ia ParA and Soj proteins (Figure 2.2)[20]. Unfortunately the biochemical

properties that differentiate Soj activity from ParA activity have yet to be elucidated making classification by bioinformatic analysis troublesome. Despite this, the fact remains that BLAST searches against the NCBI RefSeq database return a list of top hits abundant in Soj rather than ParA proteins, suggesting that whatever the differences are, Rv1708 is more similar to Soj proteins than to ParA proteins. For this reason, along with the following data, we have chosen to refer to Rv1708 as Soj_{Mtb}, as we believe it represents a true Soj ortholog in mycobacteria. This is further supported by the existence of an already functionally characterized *parAB* operon, encoded by ORFs *rv3918c-3917c*, in mycobacteria.

2.5 THE *rv1708* OPERON ARRANGEMENT SUGGESTS CHROMOSOMAL INTERACTIONS

Operon analysis of soj_{Mtb} suggests that it lies within a polycistronic operon with several other genes spanning *rv1708-rv1711*. It is interesting that soj_{Mtb} is predicted to be the first member of the operon as these genes tend to show the highest degree of expression due to their close proximity to the operator/promoter and RNAP loss or stall is more likely to occur with increased distance from this region [22]. This suggests a significant quantity of Soj_{Mtb} protein is required under cellular conditions eliciting expression. Other members of this operon include *scpA* encoded by *rv1709*, *scpB* encoded by *rv1710*, and a conserved hypothetical protein encoded by *rv1711*. ScpA and ScpB form a multi-protein complex with the structural maintenance of chromosomes protein SMC (Rv2922c), which is believed to play an important role in chromosomal organization, compaction and partitioning in the organisms that possess this program [23-24]. Although this complex is well represented in bacteria as well as eukarya its mechanism and role remain poorly understood [24]. In *B. subtilis*, which encodes a Soj/Spo0J system rather than a ParA/ParB system, it has been observed that the SMC/ScpA/ScpB program is indispensable for correct partitioning along with Spo0J, while Soj does not appear to contribute significantly [21]. In

this system it appears that Spo0J recruits SMC to the nucleoid, followed by SMC/ScpA/ScpB complex formation, and together these proteins participate in the processes of chromosomal localization, condensation and segregation, while Soj has developed alternative functionalities [25-28].

As stated previously, proteins encoded within the same operon have a high chance of functionally interacting. This is particularly interesting in lieu of the fact that ScpA has been recognized as the primary mediator of protein-protein interactions for the SMC/ScpA/ScpB complex [29]. ScpA promotes the formation of the SMC complex through interactions with both SMC and ScpB, and also has been shown to interact with several other proteins in *B. subtilis* including several histidine kinases, transcriptional regulators, and various other proteins involved in diverse cellular processes [29]. This indicates a promiscuous coordination of chromosomal organization with a wide range of cellular programs, further supporting the potential for Soj_{Mtb}-ScpA interactions *in vivo*, warranting further investigation in the future. This operon organization suggests a role for Soj_{Mtb} in chromosomal compaction through potential interactions with ScpA, a phenomenon that has been shown to directly contribute to the expressible ORFs by limiting access of RNAP to compacted regions [30]. This is an interesting mechanism of regulation as it allows global and dynamic control of the accessible promoters on the chromosome, potentially contributing to transcriptional control despite a lack of an Sda homolog in *Mtb*.

2.6 *Rv1708* OVEREXPRESSION STALLS GROWTH

Standard growth kinetic analysis was performed as described in section 5.2 using *M. smegmatis* and *Mtb* soj_{Mtb}::pVV16 and vector controls. Continuously stirred broth cultures of both showed a marked reduction in growth rate and saturation density in the merodiploid, indicating that overexpression of soj_{Mtb} disrupts growth and division *in vitro* (Figure 2.3). This result is consistent with a Soj-like activity as

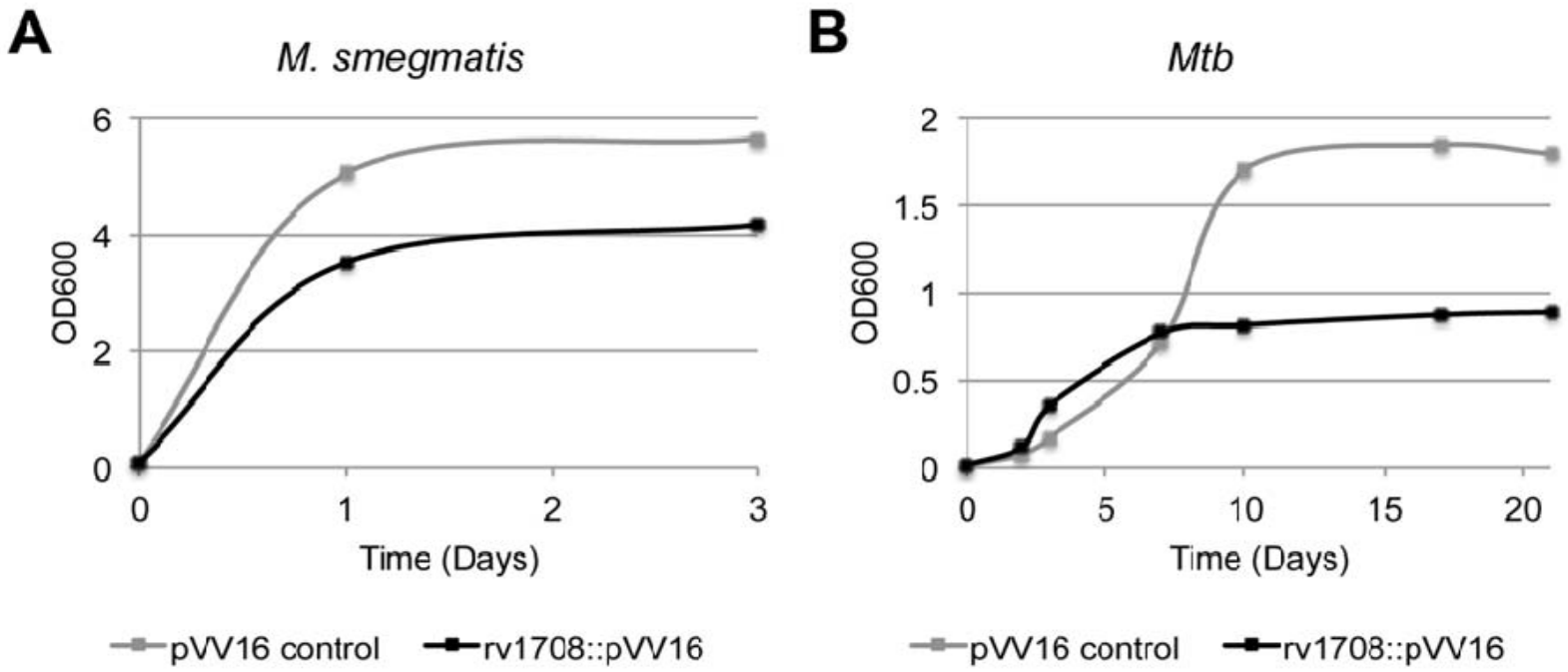

Figure 2.3: Growth of mycobacteria upon *soj*$_{Mtb}$ overexpression.

Growth of *M. smegmatis* (A) and *Mtb* (B) *soj*$_{Mtb}$::pVV16 and vector controls in continuously stirred broth cultures *in vitro* as determined by optical density 600nm readings. Values represent averages from triplicate experiments.

Soj primarily acts as a repressor of DnaA activity when not bound to DNA. A Soj regulatory scheme has been determined and proceeds as follows: monomeric Soj binds ATP, stimulating homodimerization; these homodimers may bind DNA in a cooperative fashion until it encounters Spo0J bound to specific *parS* sites on the chromosome, generally found in the vicinity of the *oriC*, stimulating ATP hydrolysis by Soj which reverts to its monomeric form, and the process repeats [19-20]. Soj has DnaA-activating ability when in the ATP-dependent dimerized form that is bound to DNA, while all other forms are inhibitive of DnaA activity [19-20]. It stands to reason that an overexpression of soj_{Mtb} would result in a small proportion of the DnaA-activating dimerized form, limited by the availability of ATP, and an abundance of DnaA-inhibiting monomers, resulting in a net decrease in replication rate. This would explain the loss of growth kinetics in mycobacteria upon constitutive soj_{Mtb} overexpression.

2.7 CONCLUSIONS

Searches for Min-system homologs failed to identify any components of this classical septum placement system in mycobacteria other than MinD. DivIVA, which has been shown to act as a polar tethering factor for the Min system in *Bacillus* sp. and other gram positive bacteria, appears to have developed an alternative role in *Mtb* where it is required for the polar localization of machinery involved in cell wall synthesis and growth [31]. Our own homology searches identified Rv1708 as encoding a putative MinD-like protein in *Mtb*, however further analysis indicate this protein is more akin to the Soj initiation and sporulation regulator of *Bacillus* sp., leading us to name this protein Soj_{Mtb}. Several lines of evidence suggest a potential for this protein in the regulation of division and chromosomal dynamics. Sporulation shares many similarities with the development of dormancy in mycobacteria, being triggered by the induction of specific molecular programs that result in the formation of a morphologically distinct, metabolically and divisionally quiescent phenotype that can endure prolonged

periods of stress [32]. Soj has been shown to participate in the development of this phenotype through replicative stall and transcriptional control of sporulation-specific genes [20]. The positioning of soj_{Mtb} adjacent to *scpA* provides contextual evidence of a role for this protein in global transcriptional regulation related to the development of NRP phenotypes. Transcriptional analysis of soj_{Mtb} merodiploids performed by the Slayden lab further supports this notion.

As Soj plays a central role in the development of NRP phenotypes in *Bacillus* sp. it stands to reason that Soj_{Mtb} plays an equally important role in mycobacterial dormancy. Soj_{Mtb} displays activity in regulating cell division and has demonstrated the potential for global transcriptional control. It is our hope that the identification and characterization of this novel regulator in mycobacteria will elucidate the processes involved in dormancy transitions. A deeper understanding of the molecular mechanisms contributing to the development of dormant phenotypes associated with latent disease will facilitate the development of novel therapeutic strategies that are more efficient and specific at targeting this bacterial population in the host. Until effective means of treating LTBI have been developed, tuberculosis will remain a prominent disease.

2.8 SUMMARY OF THIS WORK

Based on the assumption that FtsZ regulation is an essential process that must be carried out by the mycobacterial cell to ensure survival and viability of successive generations, homology-based bioinformatic searches were employed to identify ORFs in the *Mtb* H37Rv genome encoding proteins homologous to FtsZ regulators involved in septal placement and the SOS-associated replicative stall. Determination of gene orthology has been a useful tool for the examination of phylogenetic relationships as well as functional predictions of novel proteins. Therefore, protein orthology was examined using a variety of bioinformatic strategies to assess the relationship between experimentally

characterized FtsZ regulators and the putative mycobacterial regulators identified by preliminary consensus-based homology searches. Important factors contributing to protein orthology determination beyond sequence-based similarity include the genomic context and phylogenetic conservation of ORFs as well as the functional domain organization and three-dimensional structure of the proteins they encode. The information obtained from these studies was used to elucidate experimentally obtained data and guide future experimental designs.

These studies identified two proteins, each potentially participating in one of the two essential cellular processes of interest: a MinD-like protein encoded by *rv1708* putatively involved in septum placement, and SulA-like protein encoded by *rv2216* putatively involved in the DNA damage-associated SOS response, which will be discussed in later chapters. In-depth orthological analysis of these proteins indicated an alternative lineage than anticipated. It was determined that *rv1708* encodes a mycobaterial ortholog of the *B. subtilis* Soj protein. Soj in *Bacillus* sp. acts as a regulator of sporulation-specific gene expression and controls the timing of cell division through interactions with DnaA, the initiator of chromosomal replication [19-20]. The development of a smooth filamentous phenotype concurrent with growth attenuation corroborates an activity for Soj_{Mtb} that is similar to Soj in *Bacillus* sp, which inhibits growth prior to Z-ring formation. As Soj is involved in the development of dormant states in *Bacillus* sp., Soj_{Mtb} may play an equivalent role in mycobacteria, lending to the transition into dormancy during a latent infection. Soj_{Mtb} now stands as one of the few sporulation-associated genes identified in the mycobacteria to date where it acts as an important regulator of dormancy phenotypes.

REFERENCES

1. Harry, E., L. Monahan, and L. Thompson, *Bacterial cell division: the mechanism and its precison.* Int Rev Cytol, 2006. **253**: p. 27-94.
2. Gregory, J.A., E.C. Becker, and K. Pogliano, *Bacillus subtilis MinC destabilizes FtsZ-rings at new cell poles and contributes to the timing of cell division.* Genes Dev, 2008. **22**(24): p. 3475-88.
3. Park, K.T., et al., *The Min oscillator uses MinD-dependent conformational changes in MinE to spatially regulate cytokinesis.* Cell, 2011. **146**(3): p. 396-407.
4. England, K., R. Crew, and R.A. Slayden, *Mycobacterium tuberculosis septum site determining protein, Ssd encoded by rv3660c, promotes filamentation and elicits an alternative metabolic and dormancy stress response.* BMC Microbiol, 2011. **11**: p. 79.
5. Donovan, C., et al., *A synthetic Escherichia coli system identifies a conserved origin tethering factor in Actinobacteria.* Mol Microbiol, 2012. **84**(1): p. 105-16.
6. Rost, B., *Twilight zone of protein sequence alignments.* Protein Eng, 1999. **12**(2): p. 85-94.
7. Zvelebil, M.J. and J.O. Baum, *Understanding bioinformatics*. 2008: Garland Science/Taylor & Francis Group.
8. Jaramillo, M.L., et al., *Cloning, expression and bioinformatics analysis of ATP sulfurylase from Acidithiobacillus ferrooxidans ATCC 23270 in Escherichia coli.* Bioinformation, 2012. **8**(15): p. 695-704.
9. Leipe, D.D., et al., *Classification and evolution of P-loop GTPases and related ATPases.* J Mol Biol, 2002. **317**(1): p. 41-72.
10. Pathak, E., N. Atri, and R. Mishra, *Role of highly central residues of P-loop and it's flanking region in preserving the archetypal conformation of Walker A motif of diverse P-loop NTPases.* Bioinformation, 2013. **9**(1): p. 23-8.
11. Leonard, T.A., P.J. Butler, and J. Lowe, *Bacterial chromosome segregation: structure and DNA binding of the Soj dimer--a conserved biological switch.* EMBO J, 2005. **24**(2): p. 270-82.
12. Fresquet, V., L. Williams, and F.M. Raushel, *Mechanism of cobyrinic acid a,c-diamide synthetase from Salmonella typhimurium LT2.* Biochemistry, 2004. **43**(33): p. 10619-27.
13. Zhou, H. and J. Lutkenhaus, *The switch I and II regions of MinD are required for binding and activating MinC.* J Bacteriol, 2004. **186**(5): p. 1546-55.
14. Hu, Z. and J. Lutkenhaus, *A conserved sequence at the C-terminus of MinD is required for binding to the membrane and targeting MinC to the septum.* Mol Microbiol, 2003. **47**(2): p. 345-55.
15. Raux, E., et al., *Cobalamin (vitamin B12) biosynthesis: functional characterization of the Bacillus megaterium cbi genes required to convert uroporphyrinogen III into cobyrinic acid a,c-diamide.* Biochem J, 1998. **335 (Pt 1)**: p. 167-73.
16. Davis, M.A., K.A. Martin, and S.J. Austin, *Biochemical activities of the parA partition protein of the P1 plasmid.* Mol Microbiol, 1992. **6**(9): p. 1141-7.
17. Gerdes, K., J. Moller-Jensen, and R. Bugge Jensen, *Plasmid and chromosome partitioning: surprises from phylogeny.* Mol Microbiol, 2000. **37**(3): p. 455-66.
18. Mierzejewska, J. and G. Jagura-Burdzy, *Prokaryotic ParA-ParB-parS system links bacterial chromosome segregation with the cell cycle.* Plasmid, 2012. **67**(1): p. 1-14.
19. Scholefield, G., J. Errington, and H. Murray, *Soj/ParA stalls DNA replication by inhibiting helix formation of the initiator protein DnaA.* EMBO J, 2012. **31**(6): p. 1542-55.
20. Murray, H. and J. Errington, *Dynamic control of the DNA replication initiation protein DnaA by Soj/ParA.* Cell, 2008. **135**(1): p. 74-84.

21. Ireton, K., N.W.t. Gunther, and A.D. Grossman, *spo0J is required for normal chromosome segregation as well as the initiation of sporulation in Bacillus subtilis.* J Bacteriol, 1994. **176**(17): p. 5320-9.
22. Snyder, L. and W. Champness, *Molecular genetics of bacteria*. 2007: Amer Society for Microbiology.
23. Jensen, R.B. and L. Shapiro, *Cell-cycle-regulated expression and subcellular localization of the Caulobacter crescentus SMC chromosome structural protein.* J Bacteriol, 2003. **185**(10): p. 3068-75.
24. Graumann, P.L. and T. Knust, *Dynamics of the bacterial SMC complex and SMC-like proteins involved in DNA repair.* Chromosome Res, 2009. **17**(2): p. 265-75.
25. Gruber, S. and J. Errington, *Recruitment of condensin to replication origin regions by ParB/SpoOJ promotes chromosome segregation in B. subtilis.* Cell, 2009. **137**(4): p. 685-96.
26. Sullivan, N.L., K.A. Marquis, and D.Z. Rudner, *Recruitment of SMC by ParB-parS organizes the origin region and promotes efficient chromosome segregation.* Cell, 2009. **137**(4): p. 697-707.
27. Autret, S., R. Nair, and J. Errington, *Genetic analysis of the chromosome segregation protein Spo0J of Bacillus subtilis: evidence for separate domains involved in DNA binding and interactions with Soj protein.* Mol Microbiol, 2001. **41**(3): p. 743-55.
28. Dedrick, R.M., H. Wildschutte, and J.R. McCormick, *Genetic interactions of smc, ftsK, and parB genes in Streptomyces coelicolor and their developmental genome segregation phenotypes.* J Bacteriol, 2009. **191**(1): p. 320-32.
29. Dervyn, E., et al., *The bacterial condensin/cohesin-like protein complex acts in DNA repair and regulation of gene expression.* Mol Microbiol, 2004. **51**(6): p. 1629-40.
30. Luijsterburg, M.S., et al., *The architectural role of nucleoid-associated proteins in the organization of bacterial chromatin: a molecular perspective.* J Struct Biol, 2006. **156**(2): p. 262-72.
31. Kang, C.M., et al., *Wag31, a homologue of the cell division protein DivIVA, regulates growth, morphology and polar cell wall synthesis in mycobacteria.* Microbiology, 2008. **154**(Pt 3): p. 725-35.
32. Errington, J., *From spores to antibiotics via the cell cycle.* Microbiology, 2010. **156**(Pt 1): p. 1-13.

CHAPTER 3: THE PUTATIVE SULA/CDR REGULATOR ENCODED BY RV2216

The bioinformatic identification, orthological analysis, protein classification, and sub-cellular localization studies presented in this work were performed by Rebecca Crew as presented in the manuscript "MazF6 toxin of *Mycobacterium tuberculosis* demonstrates antitoxin specificity and is coupled to regulation of cell growth by a Soj-like protein" to be published in BMC Microbiology, 2013. Morpholigical and transcriptional analysis was performed by K. England.

3.1 *Rv2216* ENCODES A PUTATIVE DIVISION REGULATOR *Mtb*

Consensus-based protein BLAST searches found no ORFs encoding proteins that share significant similarity with MinC, MinE, MipZ or YneA family regulators. However searches using SulA and MinD consensus models identified *rv2216*, *rv3660c* and *rv1708* as encoding proteins similar to FtsZ regulators from other bacterial species. For a discussion of the MinD putative homologs encoded by *rv3660c* and *rv1708* refer to Chapter 2. The 301 amino acid protein encoded by *rv2216*, annotated in the current databases as a conserved hypothetical protein of unknown function, was identified as sharing significant homology (33% sequence identity) with SulA proteins however this similarity was limited over the length of the protein (only 9% coverage), warranting further investigation. This difference is not due to size limitations as Rv2216 is contains nearly twice the number of residues as SulA, therefore this low homology is likely due to a different ancestral lineage.

3.2 Rv2216 IS ORTHOLOGOUS TO Cdr PROTEINS

Functional domain mapping of Rv2216 revealed a two domain organization consisting of an NADH/NADPH utilizing domain (PF01370) adopting a Rossman-like fold that comprises a majority of the protein sequence, and a small C-terminal domain of undetermined function and conformation (DUF1731, PF08338) (Figure 3.1). TransMembrane prediction using Hidden Markov Models (TMHMM) analysis through TubercuList indicates the existence of a short, 26 amino acid N-terminal signal sequence. Analysis through the Protein Homology/analogy Recognition Engine V2 (Phyre2) confirmed this domain architecture and suggested a general functional classification as a nucleoside-diphosphate sugar epimerase, related to SulA proteins.

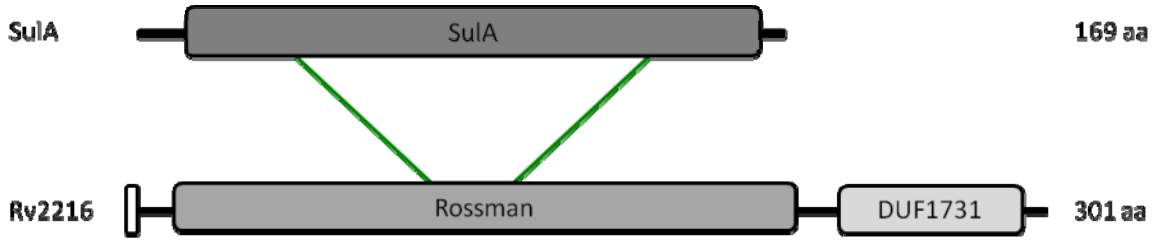

Figure 3.1: Comparative functional domain organizations of SulA and Rv2216.

Rv2216 shares significant yet limited sequence similarity with SulA proteins. This similarity primarily maps to a central catalytic region of the SulA (PF03846) and Rossman (PF01370) domains, indicated by green lines. DUF1731 (PF08338) is unique to Rv2216, as is the N-terminal 26 amino acid signal sequence, indicated by a white box.

The OMA Browser separates Rv2216 orthologous groups (OMA 66279-CLIRTGI and related groups) from SulA proteins (OMA 440861 and others). The groups of orthologous Rv2216-like proteins (Figure 3.2) are composed of conserved hypotheticals of unknown function loosely annotated as

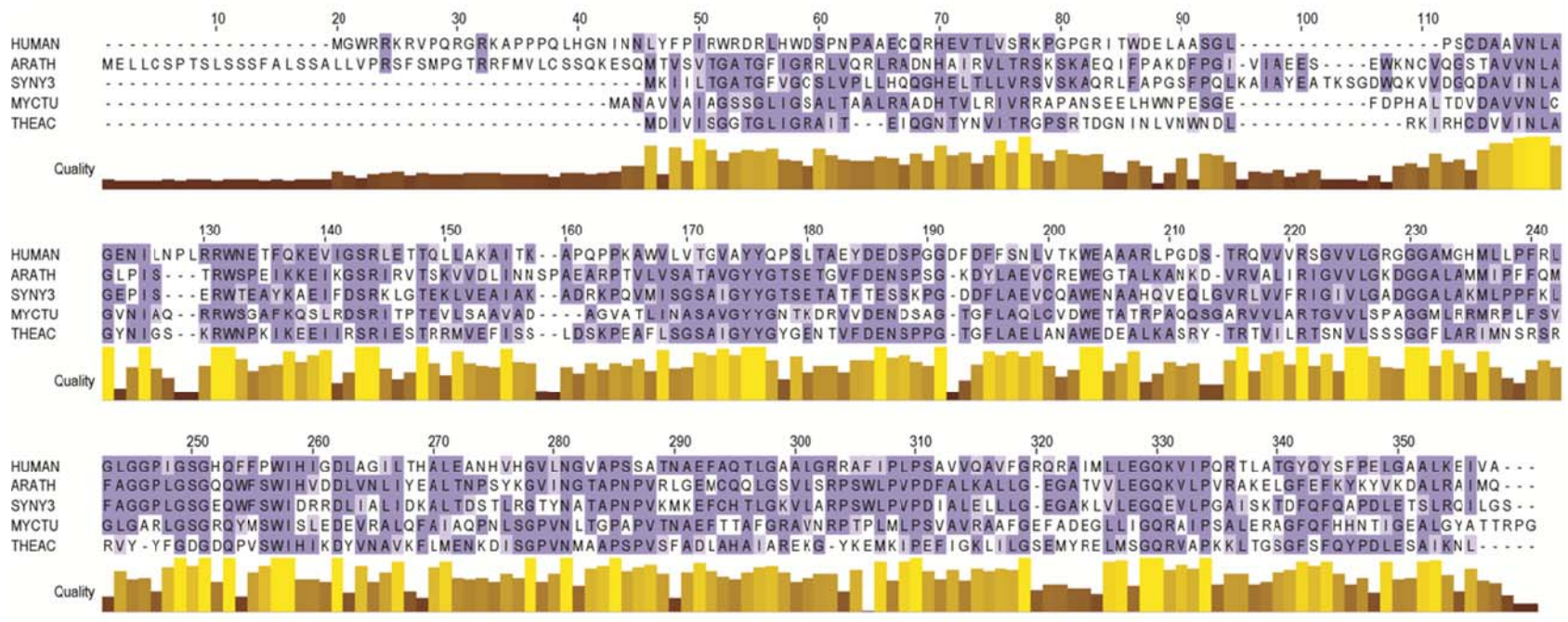

Figure 3.2: Local multialignment of Cdr proteins from diverse organisms.

Select Cdr proteins were chosen from among the predicted 561 orthologs to visualize sequence conservation in diverse species. HUMAN - *H. sapien* (uniprot Q9NRG7); ARATH - *A. thaliana* (uniprot Q9SJU9); SYNCY - *Synechocystis* sp. (uniprot P73467); MYCTU - *M. tuberculosis* (uniprot P67232); and THEAC - *T. acidophilum* (uniprot Q9HIQ8). Shared amino acids are indicated by dark blue and non-identical residues sharing similar biochemical properties are indicated by light blue.

putative cell division inhibitors. Due to a lack of experimental evidence corroborating this functional description we have termed these Rv2216-like proteins as Cdr proteins for cell division regulatory proteins. Interestingly, Cdr proteins show greater conservation than SulA proteins, with 561 predicted orthologs compared to 95 for SulA, and are characterized by the presence or absence of DUF1731 at the C-terminus (Figure 3.2).

3.3 THE *cdr* OPERON CONTAINS ESSENTIAL ORFs IN *Mtb*

The Operon Database (OperonDB) predicts *rv2216* as the second member of a polycistronic operon consisting of *rv2215-rv2219*, composed of the annotated predicted essential ORFs *dlaT*, *lipB*, and *lipA* as well as an unannotated conserved transmembrane protein encoded by *rv2219* (Figure 3.3). This operon prediction is supported by strong positive pearson correlation coefficients ($\geq$ 0.4) from the Tuberculosis Database (TBDB) operon browser for the genes of this operon, suggesting co-regulation of gene expression. Experimental evidence corroborating the existence of this operon is provided by the observation of reduced *rv2216* expression levels in *dlaT* knockout mutants [1].

Figure 3.3: Organization of the *cdr* operon in *Mtb* H37Rv.

DlaT (dihydrolipoamide acytransferase) is a component of the pyruvate dehydrogenase (PDH) and peroxynitrite reductase/peroxidase (PNR/P) protein complexes. The activity of DlaT is activated by the covalent attachment of a lipoic acid moiety, a post-translational modification termed protein lipoylation, performed by LipA and LipB. These proteins are essential for intermediary metabolism and the neutralization of RNI during host infection. *rv2219* encodes a conserved hypothetical protein of unknown function. *dlaT – rv2215*; *cdr – rv2216*; *lipB – rv2217*; *lipA – rv2218*.

3.4 *cdr* OVEREXPRESSION ELICITS FILAMENTATION WITHOUT DIVISION ATTENUATION

To investigate the connection between cell cycle progression and Cdr activity, phenotypic observation as well as growth kinetic analysis was employed upon *cdr* overexpression. *Mtb* and *M. smegmatis cdr* merodiploid strains showed elongated phenotypes when examined by scanning electron microscopy (SEM) (Figure 3. 4, A-D). Measurements of the *Mtb cdr* merodiploid cells showed a moderate and bimodal increase in average cell length from 1-2 µm in vector control strains to 1.5-2.5 µm and 3.5-5.5 µm in the experimental strain (Figure 3. 4, E). The frequency of mid-cell bands representative of Z-rings (Figure 3.4 black arrow) was equivalent to wild type upon *cdr* overexpression. These ultrastructural features are indicative of a disruption in late division events (after Z-ring formation) in *Mtb* [2-4]. This suggests an activity different from SulA proteins which inhibit the formation of Z-ring structures [5]. Surprisingly, no statistically significant difference in growth kinetics could be observed in the *M. smegmatis cdr* merodiploid by optical density 600nm ($O.D._{600nm}$) readings in broth suspension (Figure 3.5).

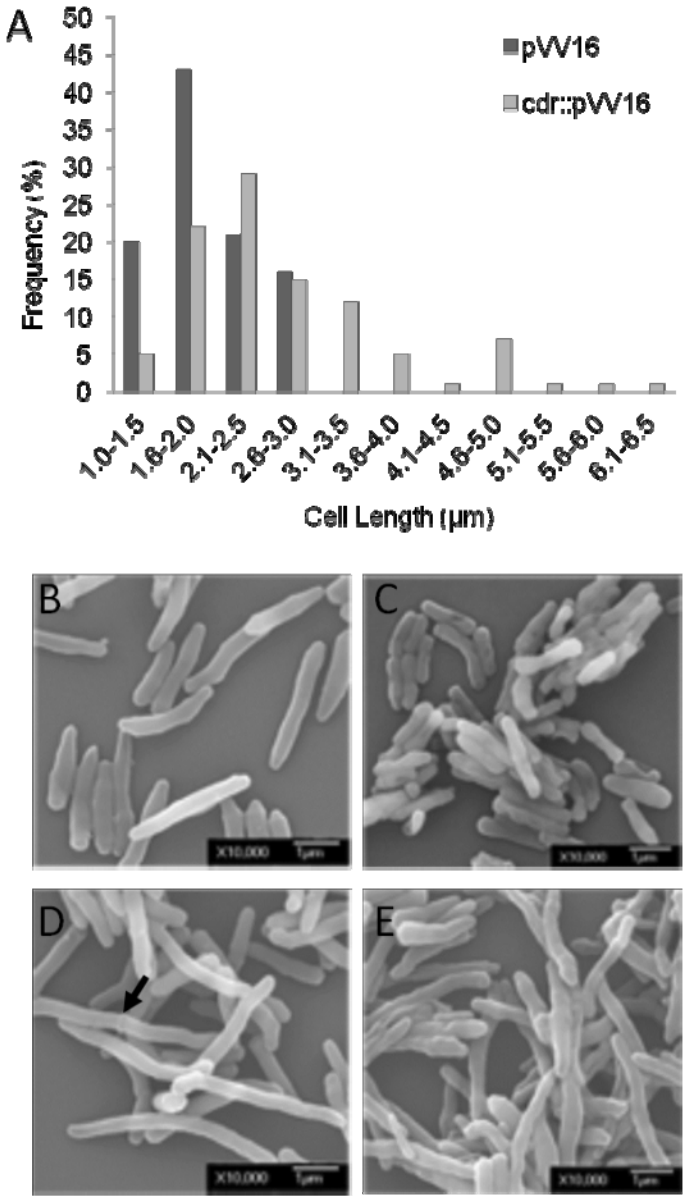

Figure 3.4: Morphological analysis of mycobacterium in response to *cdr* overexpression.

(A) Measurements of 100 *M. smegmatis* cells over 3 different fields of view each were averaged and plotted according to length. Vector controls (pVV16) averaged 1-2 µm in mid-log grown cultures compared to mean lengths of 1.5-2.5 µm and 4.5-5.5 µm in *cdr*::pVV16. (B-D) SEM of *Mtb and M. smegmatis* pVV16 (A and B) and *cdr*::pVV16 (C and D).

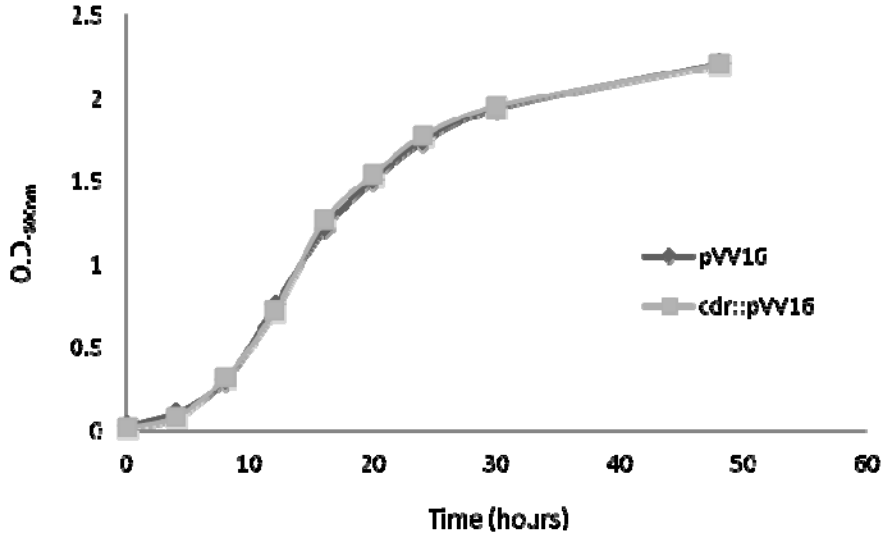

Figure 3.5: Standard growth kinetics of *M. smegmatis* upon *cdr* overexpression.

Average values were determined from three independent replicates and plotted to visualize growth kinetics of pVV16 and *cdr*::pVV16 in *M. smegmatis* by $O.D._{600nm}$ readings of broth suspensions.

3.5 Cdr DOES NOT PARTICIPATE IN THE SOS RESPONSE

Previous studies have found that induction of the SOS response inhibits Z-ring formation through the activity of SulA protein in the cell [5]. To assess the connection between Cdr and SulA protein activity, we induced the SOS response using sub-inhibitory concentrations of the DNA cross-linking agent mitomycin C (MMC) [6]. Phenotypic observation by SEM of *Mtb* following MMC treatment revealed an extremely filamentous phenotype, devoid of apparent Z-ring structures, indicative of disruption in early division events (Figure 3.6, A-B). qRT-PCR analysis of SOS and cell cycle discriminant genes revealed the successful induction of DNA repair genes characteristic of the SOS response by this treatment as well as a down regulation of cell division genes, indicating that a disruption in cell division is experienced as a result of SOS program induction under these conditions (Figure 3.6, C-D).

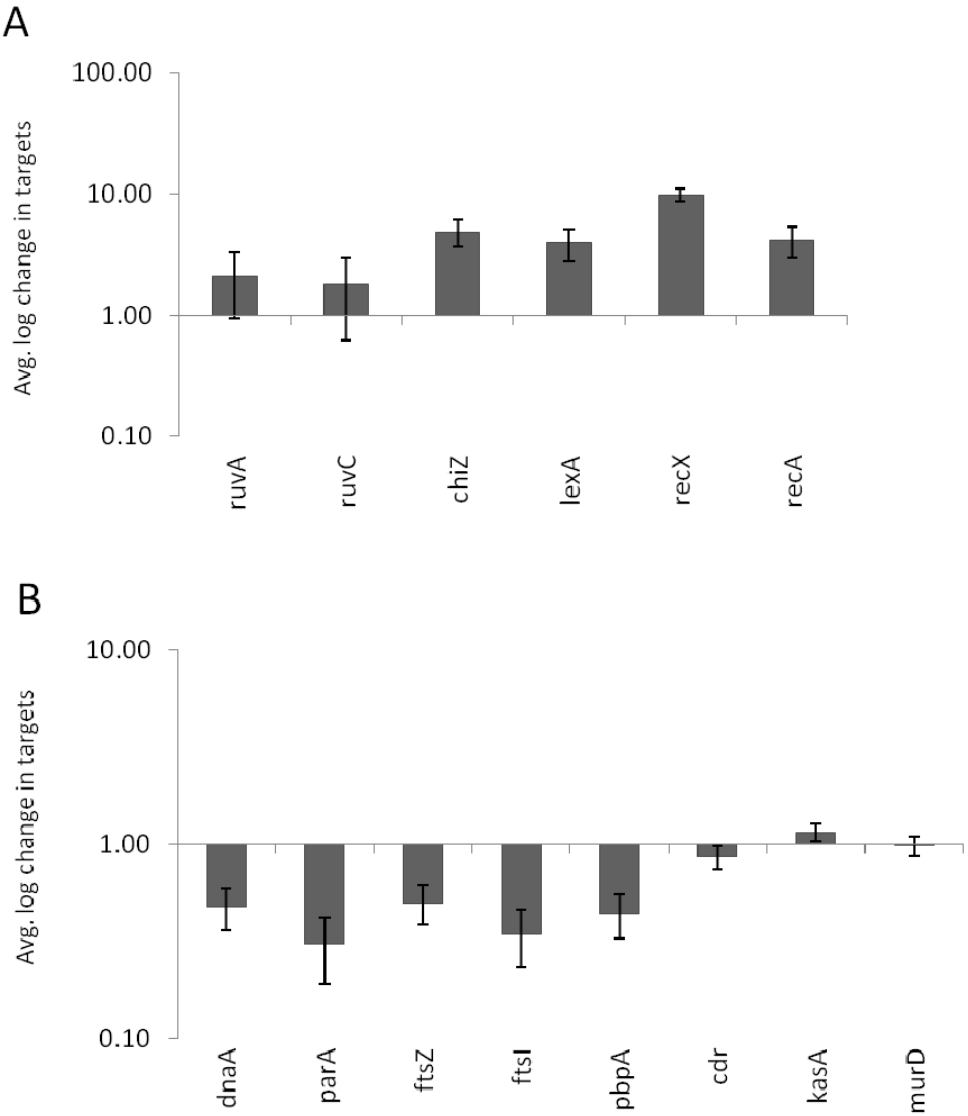

Figure 3.6: qRT-PCR analysis of SOS and cell division discriminant genes in response to MMC on *Mtb*.

Transcriptional analysis of SOS response (A) and cell cycle (B) discriminant genes in response to treatment with sub-inhibitory concentrations of MMC on *Mtb*. Values are averaged from two independent replicates, normalized to *sigA*, and represented by the mean log change in expression between treated and un-treated H37Rv.

Importantly, the lack of *cdr* induction following MMC treatments indicates that *cdr*, unlike *sulA*, is not a component of the DNA damage repair response, also known as the SOS response (Figure 3.6, D). This observation further corroborates a divergent activity and level of regulation for Cdr and SulA proteins.

3.6 *Cdr* OVEREXPRESSION STIMULATES LATE DIVISION PROCESSES

As Cdr involvement in cell cycle progression was unclear from standard growth kinetic analysis, transcriptional profiling of cell cycle discriminant genes was performed in *cdr* merodiploids to further elucidate potential cell cycle disruption in response to *cdr* overexpression. A general trend could be observed showing a suppression of early division gene expression and an upregulation of late division genes (Figure 3.7). Profiling performed by the Slayden lab in response to FtsI inhibition displayed a similar cellular response, as the disruption of this late division protein results in a suppression of fellow late division proteins. Conversely, similar studies on FtZ-inhibited bacteria displays an inverse pattern of expression in that inhibition of the early division protein FtsZ causes a down-regulation of fellow early division proteins. These cellular responses are likely the result of an attempt to maintain appropriate stiochiometry of proteins involved in the same cellular processes (late-late or early-early), while attempting to overcome the stall in division by up-regulating the opposite programs, ie. inhibition of late division induces early division programs and vice versa. Together these studies indicate that increased quantity of Cdr in the bacteria results in a disturbance of late division, post Z-ring, events [3-4]. Alternatively, as late cell division processes are involved the production and remodeling of plasma membrane and cell wall material, the elongated morphology of the *cdr* merodiploid may be due to enhanced membrane and cell wall synthesis.

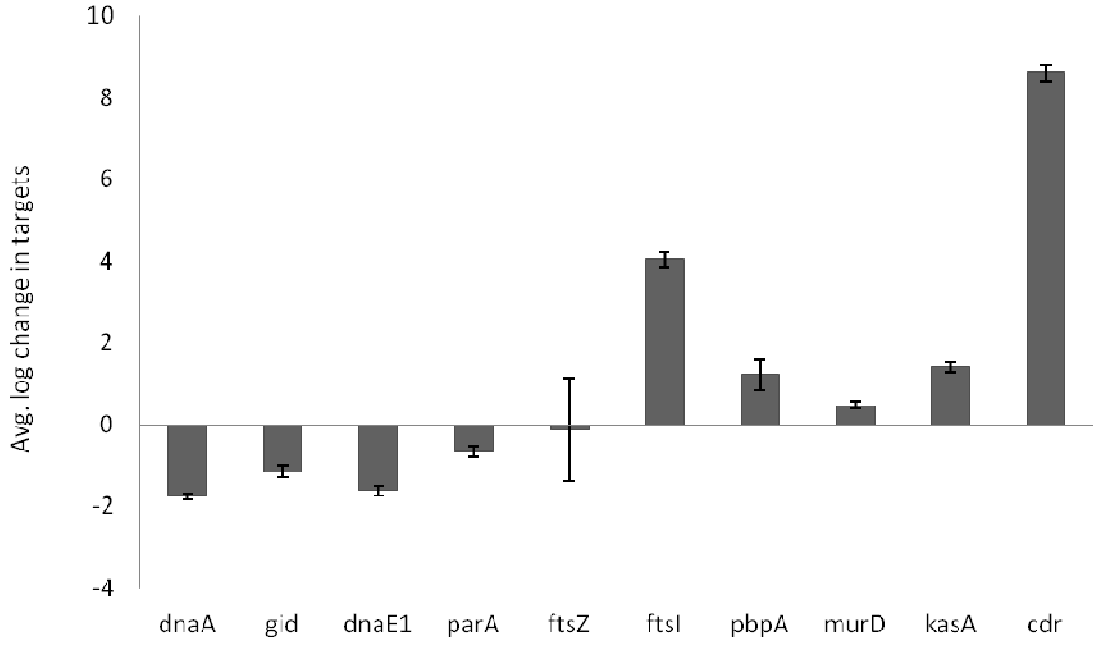

Figure 3.7 qRT-PCR analysis of cell cycle discriminant genes in *cdr*::pVV16 *Mtb*.

Transcriptional analysis of cell cycle genes involved in chromosomal replication (*dnaA*, *gid*, *dnaE1*), segregation (*parA*), Z-ring formation (*ftsZ*), and septum cell wall synthesis (*ftsI*, *pbpA*, *murD*, *kasA*). Values are averaged from two independent replicates, normalized to *sigA*, and represented by the mean log change in expression between *cdr*::pVV16 and wild type strains.

3.7 *Cdr* OVEREXPRESSION ELICITS METABOLIC REPROGRAMMING

Whole genome transcriptional profiling was employed to elucidate the molecular programs associated with Cdr activity. Upon *cdr* overexpression 1,048 ORFs were transcriptionally active (SNR>2, p value $\leq$0.05) with a total of 277 displaying a $\geq$ 2-fold increase in expression (Appendix A). In total, this expression profile shows a global change of ~7% of the encoded ORFs in the *Mtb* genome, which is consistent with cellular responses to disruption of division and induction of adaptive programs [2]. Genes associated with the functional classes of hypotheticals and conserved hypotheticals as well as

intermediary metabolism and respiration were the most differentially regulated, followed by cell wall processes and lipid metabolism (Figure 3.8).

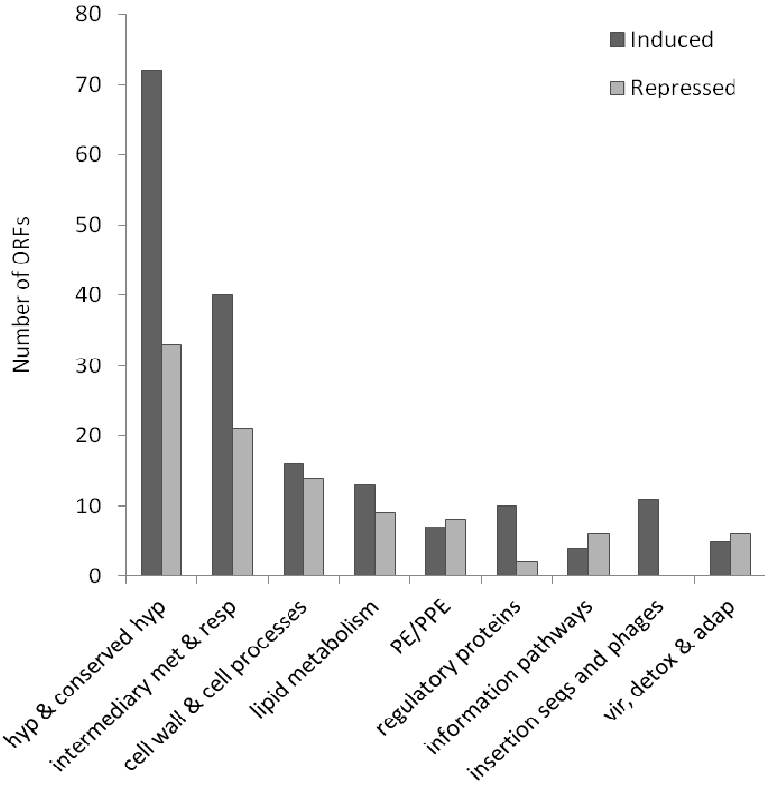

Figure 3.8 Differentially expressed ORFs in *cdr*::pVV16 in *Mtb* sorted by functional class.

The 277 genes showing statistically significant changes in expression (SNR>2, p value $\leq$0.05, log ratio $\geq$2.0) in *cdr*::pVV16 *Mtb* sorted by functional class display an induction of poorly defined cellular programs and intermediary metabolism and respiration, followed by differential regulation of genes involved in cell wall processes and lipid metabolism.

PAthosysTems Resource Integration Center (PATRIC) pathway comparisons of the significantly differentially expressed ORFs revealed a unique induction of specific molecular programs associated with oxidative phosphorylation (Figure 3.9), *de novo* fatty acid synthesis (Figure 3.10), and glycerophospholipid synthesis (Figure 3.11). Figures display Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways, with Enzyme Commission (EC) numbers to represent enzymatic reactions. The major

protein components of the oxidative phosphorylation systems were down-regulated in the merodiploid with the exception of NADH Dehydrogenase [1.6.99.3] which was surprisingly induced (Figure 3.9). In general this indicates a reduced capacity for aerobic respiration upon *cdr* overexpression in *Mtb*. The FasI system of *de novo* fatty acid synthesis was significantly up-regulated upon *cdr* overexpression, along with specific programs favoring the production of hexadecanoic and octadenacoic acids (palmitic and stearic acid, respectively) used as precursors for long-chain fatty acid synthesis, the primary acyl components of membrane phospholipids (Figure 3.10) [7]. The profile of glycerophospholipid metabolism showed wide-spread changes favoring the utilization of lysophospholipid moieties (LPLs), which serve as important intermediates during *de novo* phospholipid synthesis or are produced by membrane damage. Furthermore, processes creating and utilizing lysophospholipid moieties (LPLs), phospholipids containing only 1 acyl chain which represent important intermediates during phospholipid synthesis and remodeling, were up-regulated in the *cdr* merodiploid (Figure 3.11). Specifically, the ECs [2.3.1.15], [2.3.1.51], and [2.3.1.-] were up while [2.7.8.5] and [2.7.8.-] were down (Figure 3.11). Programs allowing the inter-conversion of different phospholipid species were generally down-regulated (Figure 3.11) [8].

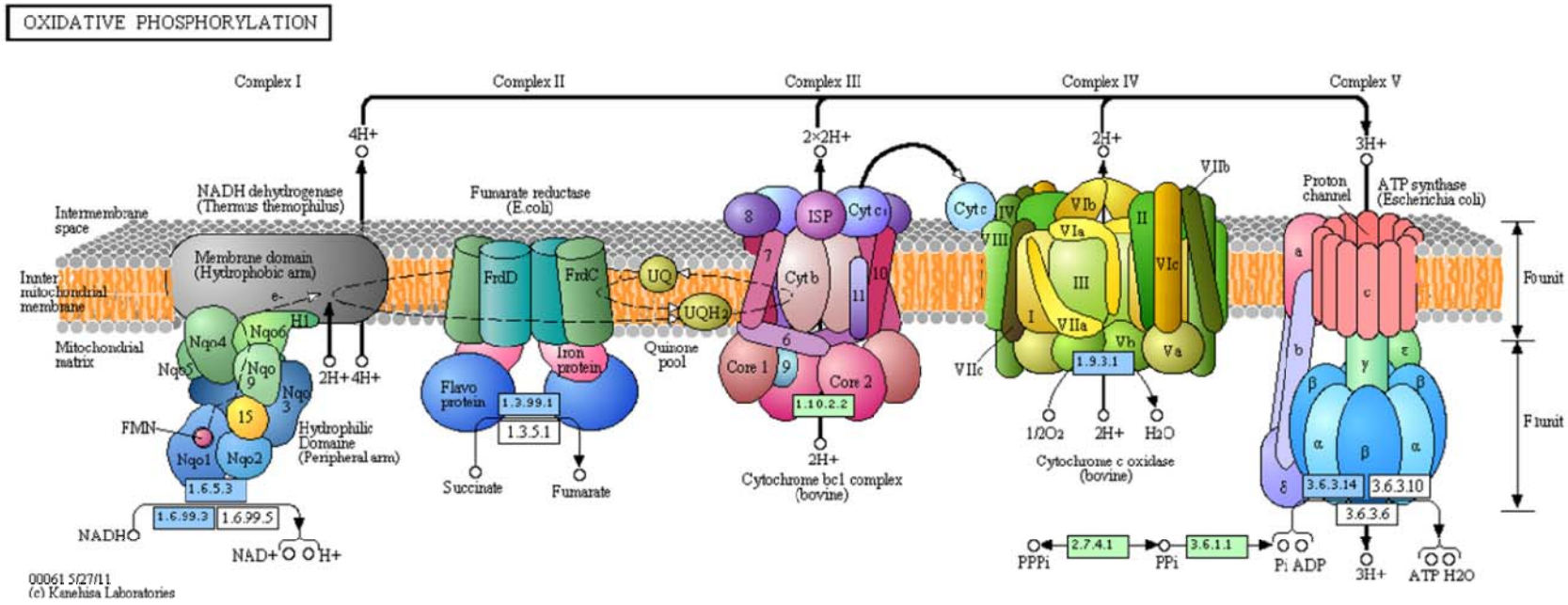

Figure 3.9: Oxidative phosphorylation KEGG pathways differentially regulated in *cdr*::pVV16 *Mtb*.

The PATRIC comparative pathways tool was used to visualize KEGG pathways differentially regulated in the *cdr*:pvv16 *Mtb* strain, as determined by microarray analysis. The 277 genes displaying statistically significantly altered expression levels (SNR>2, p value $\leq$0.05, log ratio $\geq$2.0) were used for this analysis. Oxidative phosphorylation genes that were differentially expressed are represented by blue EC numbers, representative of the enzymatic reactions their encoded proteins participate in. Green shaded EC boxes indicate processes present in *Mtb* but not differentially regulated in the *cdr*:pvv16 strain. White boxes indicate ECs for which no proteins have yet been annotated in *Mtb*. All blue boxes were negatively regulated by *cdr* overexpression with the exception of EC [1.6.99.3], which was up-regulated.

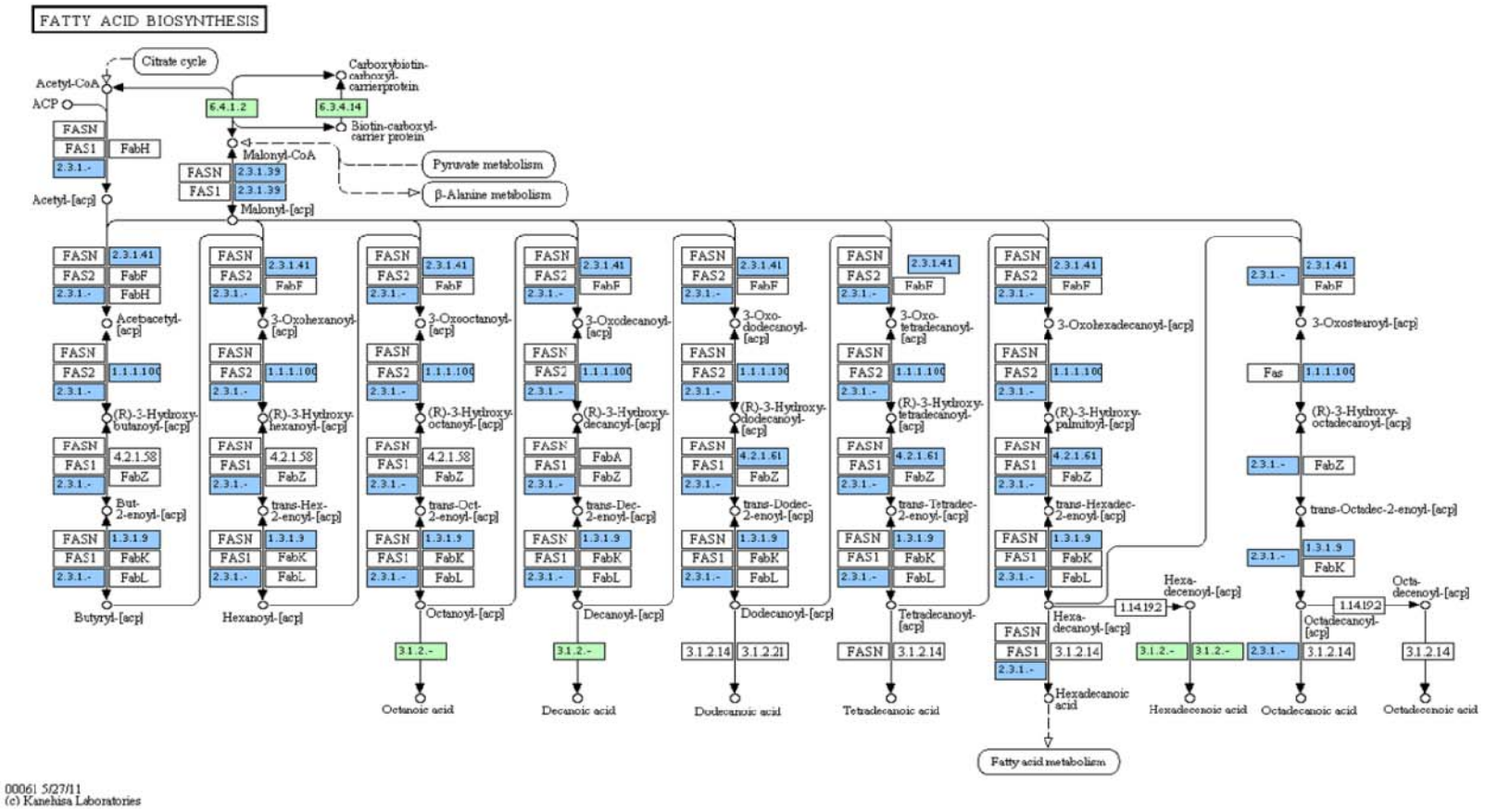

Figure 3.10: Fatty acid synthesis KEGG pathways differentially regulated in *cdr*::pVV16 *Mtb*.

De novo fatty acid synthesis genes that were induced are represented by blue EC numbers, representative of the enzymatic reactions their encoded proteins participate in. Green shaded EC boxes indicate processes present in *Mtb* but not differentially regulated in the *cdr*:pVV16 strain. White boxes indicate ECs for which no proteins have yet been annotated in *Mtb*.

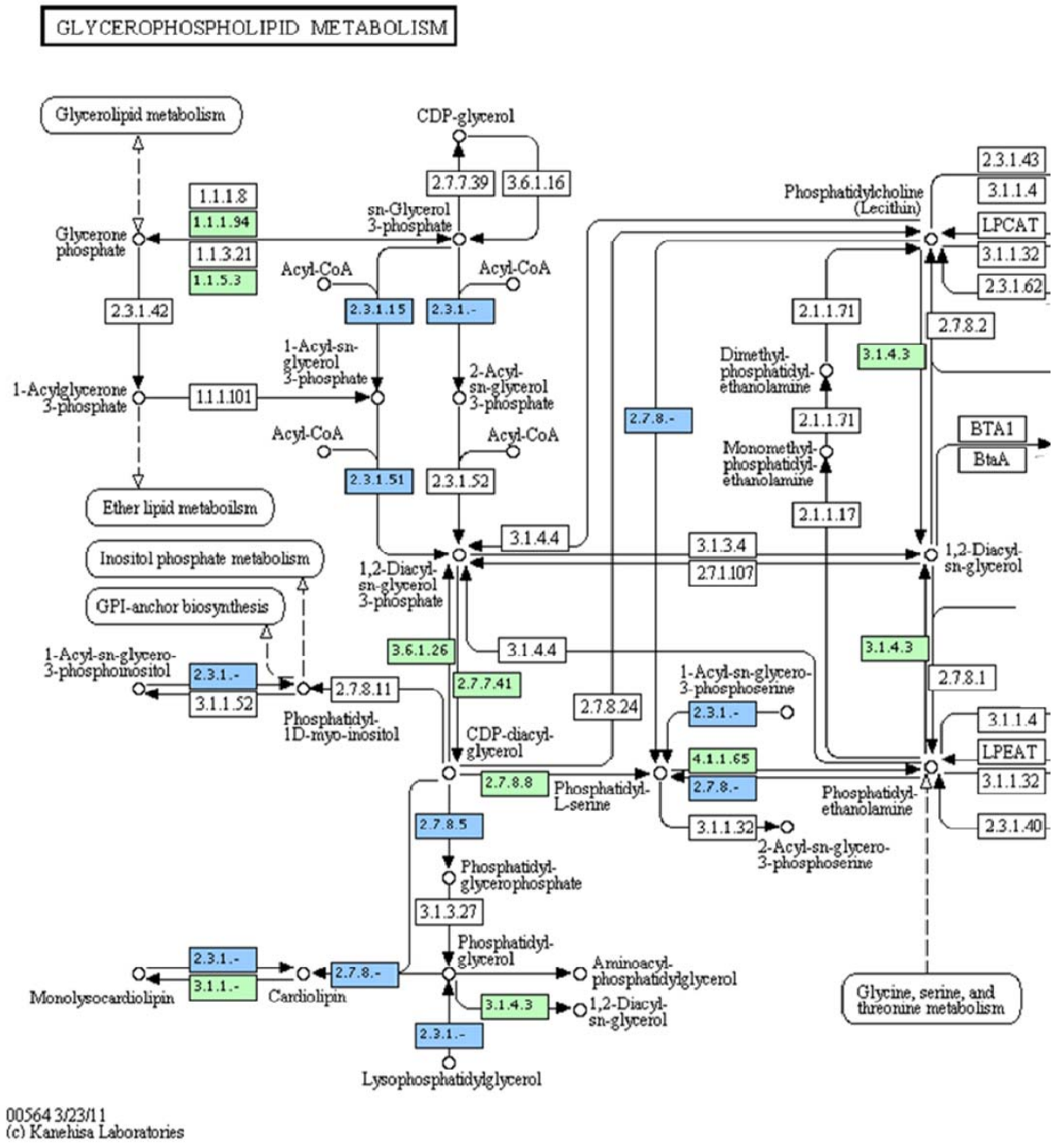

Figure 3.11 Glycerophospholipid metabolism KEGG pathway differential regulation in *cdr*::pVV16 *Mtb*.

Up-regulated pathways involved in LPL utilization include ECs [2.3.1.15], [2.3.1.51], and [2.3.1.-] and down-regulated phospholipid interconversion reactions [2.7.8.5] and [2.7.8.-]. Blue ECs indicate processes differentially regulated in *cdr*::pVV16, green indicates non-differentially regulated processes, and white show enzymatic reactions catalyzed by no known annotated proteins in *Mtb*.

3.8 Cdr-GFP LOCALIZES TO EARLY DIVISION SITES AND REGIONS OF ACTIVE GROWTH

To elucidate sub-cellular locations of Cdr protein and examine any potential roles for Cdr at the division site, a recombinant fluorescently tagged Cdr protein was expressed in *M. smegmatis* and visualized using fluorescent microscopy. Cdr-GFP localized dynamically and consistently in a cell-cycle dependent manner. Traditionally the first step of the division cycle is marked by the replication of the bacterial chromosome (Figure 3.12, B step 1), this is followed by segregation of the chromosomes to the poles as the cell continues to elongate (Figure 3.12, B step 2). Once the chromosomes have been cleared from the mid-cell site the Z-ring may begin to form, after which synthesis of new cell wall material on the Z-ring scaffold proceeds quite quickly (Figure 3.12, B step 3). Finally, the daughter cells separate (Figure 3.12, B step 4) and the process repeats from the beginning.

A Cdr-GFP localization scheme becomes apparent after organizing the sub-cellular localization patterns according to their progression through the cell cycle. Relative to the above steps, Cdr-GFP begins near a polar position with minor foci apparent at the mid-cell site as the DNA content of the cell increases (Figure 3.12, B step 1). As the chromosomes begin to segregate, the mid-cell presence of Cdr-GFP wanes and the remaining foci re-locate to an extreme polar position (Figure 3.12, B step 2). Differential daughter cell morphologies have commonly been observed during mycobacterial replication, where resultant cells from a single round of division may appear in either a linear or V-shaped morphology, however an explanation for this phenomenon has yet to be described [9]. Interestingly, in this phase, which appears to mark the initiation of the V-shaped phenotype, a differential behavior of Cdr-GFP foci may be observed in the two daughter cells. While the polar foci of the future linear cell remains at a polar position, Cdr-GFP splits along the outer elbow of the future snapping cell (arrow-heads) indicating a potential role for this protein in the morphological variability observed during mycobacterial division (Figure 3.12, B step 2). After this point the two future cells

display a similar Cdr-GFP localization patterns. Once the Z-ring has been established and the septum has formed, Cdr-GFP foci re-localize to a quarter-cell position (Figure 3.12, B step 3), a location representative of the future division site of the almost fully formed daughter cells, as the final steps of division will complete quickly, and a new round of division will soon commence. Finally, as the daughter cells begin to separate, the quarter-cell foci persist and novel foci develop at the newly formed cell poles (Figure 3.12, B step 4). Presumably, as the daughter cells complete their separation, marking the completion of a single round of division, the new polar foci develop further while the quarter-cell foci relocate to the old cell pole, and the cycle continues anew from step 1.

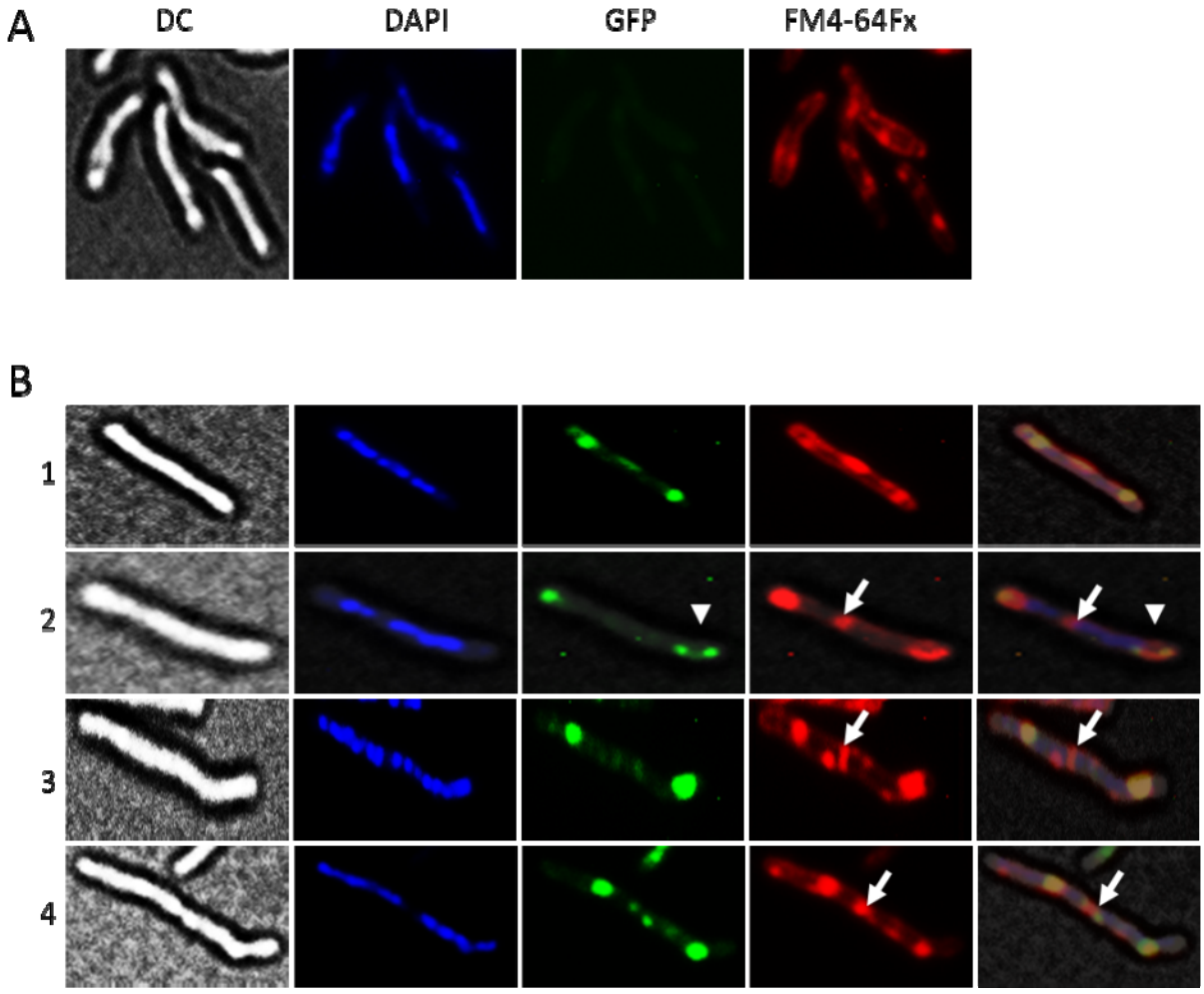

Figure 3.12: Cdr-GFP localization in *M. smegmatis*.

Fluorescent microscopy of *M. smegmatis* stained with DAPI nucleic acid stain to visualize chromosomes and lipophilic FM 4-64 Fx dye to visualize membranes and septal structures in wild type (A) and *cdr*::pMCSU7 (B) strains. *Cdr*-GFP expression was induced for 6 hours prior to fixation and imaging. Localization patterns of Cdr-GFP are arranged by progression through the cell cycle (1-4) as described in the text. White arrows indicate division sites (2-4) and white arrow-heads indicate splitting foci at the outer edge of early snapping daughter cell (2). As the best results were visualized at 6-hours post-induction no significant morphological elongation was observed in Cdr-GFP expressing bacterium.

REFERENCES

1. Shi, S. and S. Ehrt, *Dihydrolipoamide acyltransferase is critical for Mycobacterium tuberculosis pathogenesis.* Infect Immun, 2006. **74**(1): p. 56-63.
2. Slayden, R.A., D.L. Knudson, and J.T. Belisle, *Identification of cell cycle regulators in Mycobacterium tuberculosis by inhibition of septum formation and global transcriptional analysis.* Microbiology, 2006. **152**(Pt 6): p. 1789-97.
3. Slayden, R.A. and J.T. Belisle, *Morphological features and signature gene response elicited by inactivation of FtsI in Mycobacterium tuberculosis.* J Antimicrob Chemother, 2009. **63**(3): p. 451-7.
4. Respicio, L., et al., *Characterizing septum inhibition in Mycobacterium tuberculosis for novel drug discovery.* Tuberculosis (Edinb), 2008. **88**(5): p. 420-9.
5. Mukherjee, A., C. Cao, and J. Lutkenhaus, *Inhibition of FtsZ polymerization by SulA, an inhibitor of septation in Escherichia coli.* Proc Natl Acad Sci U S A, 1998. **95**(6): p. 2885-90.
6. Rand, L., et al., *The majority of inducible DNA repair genes in Mycobacterium tuberculosis are induced independently of RecA.* Mol Microbiol, 2003. **50**(3): p. 1031-42.
7. Moo-Young, M., *Comprehensive Biotechnology (Online)*. 2011: Elsevier Science.
8. Nohturfft, A. and S.C. Zhang, *Coordination of lipid metabolism in membrane biogenesis.* Annu Rev Cell Dev Biol, 2009. **25**: p. 539-66.
9. Thanky, N.R., D.B. Young, and B.D. Robertson, *Unusual features of the cell cycle in mycobacteria: polar-restricted growth and the snapping-model of cell division.* Tuberculosis (Edinb), 2007. **87**(3): p. 231-6.

CHAPTER 4: DISCUSSION

Since the early 1900's concerted scientific efforts have aimed to eradicate prevalent infectious diseases from the human population [1]. In those early days *Mtb* was considered an ideal candidate for targeted eradication because it is a human-specific bacterial pathogen, diagnosable by the PPD (purified protein derivative) test, and treatable with antibiotics [2]. Since then, it has become apparent that eradication of tuberculosis from the human population will not be as simple as anticipated due in part to the phenomena of bacterial persistence associated with chronic *Mtb* infection [3]. Contributing to the global disease burden is this vast number of chronic infections, a form that is particularly difficult to diagnose due to a lack of disease symptoms and difficult to treat as the bacteria are less responsive to chemotherapeutics during this time. The transition from active TB to a chronic infection, or LTBI, requires a stall in replication and down-shift of metabolic processes for the bacteria to enter into a non-replicating persistent (NRP) state. It is believed that the transition into NRP is a controlled process by the bacteria, whereby purposeful inhibition of cell division and metabolic shift-down is utilized to tolerate stressful conditions experienced within the host [4-5]. As replicative stall is most commonly the result of protein-targeted inhibition of FtsZ polymerization in diverse bacteria in relation to pathological and stress-associated adaptations, we set out to identify the protein inhibitors of cell division in *Mtb* [6-8].

4.1 Rv2216 IS AN ORTHOLOG OF A HIGHLY CONSERVED FAMILY OF Cdr PROTEINS

Initial bioinformatic screens utilizing consensus-based modeling and reciprocal best hits strategies identified *rv2216* as encoding a protein sharing significant similarity with SulA, the most widespread and well-described cell division inhibiting protein. However, questions were raised in regards to the strength of this match as the coverage was poor with only 9% of the Rv2216 protein

mapping to SulA, however, this is not due to length limitations as Rv2216 is nearly twice as long as SulA.. Further investigations into the nature of Rv2216 revealed a unique and divergent functional domain organization from SulA (Figure 3.1) that is characteristic of a highly conserved yet poorly defined set of proteins vaguely annotated as putative cell division inhibitors. Experimental evidence corroborating this classification is lacking and seems to be due primarily to the weak similarity these proteins share with SulA division inhibitors. Despite ambiguity regarding exactly how Rv2216-like proteins affect division, there is evidence to support an involvement in the division processes of the organisms in which they are encoded; therefore we have chosen to define this protein family as Cdr proteins, referring to their involvement in cell division regulation [9-10].

While *sulA* is primarily encoded by gram-negative alphaproteobacteria, *cdr* is more ubiquitous and may be found in extremely diverse organisms ranging from archaea and bacteria to higher mammals (Figure 3.2). Further evidence of the differential lineage and activity of Cdr and SulA is revealed by the existence of true Cdr homologs in SulA-containing organisms such as *E. coli.* The *E. coli* SulA (uniprot P0AFZ5) shares 30% identity and 9% coverage with Rv2216, similar to comparisons between Cdr and SulA consensus models. However, *E. coli* also encodes a true Cdr homolog, currently annotated *yfcH* (uniprot P177775), which shares more extensive similarity (39%) and coverage (95%) with Rv2216 than SulA, further supporting a different lineage for these proteins. In higher organisms Cdrs are plastid or mitochondrial targeted proteins indicating early development and dissemination of these proteins as these structures are the remnants of ancient symbiotic bacterium [9-12]. The surprising level of conservation at both the organismal and sequence level through millions of years of evolution suggests a vital role for Cdr in the survival and/or propagation of the organisms in which they are encoded.

4.2 Cdr DOES NOT PARTICIPATE IN THE SOS RESPONSE

To explore potential functional similarities between Cdr and SulA proteins, Cdr expression was assessed by qRT-PCR analysis following induction of the SOS response in *Mtb*. The SOS response is a well characterized molecular program designed to coordinate an inhibition of cell division, elicited by SulA activity, in response to DNA damage and to allow DNA repair to occur prior to the resurgence of division thus ensuring chromosomal viability and genetic fitness for future generations [13]. Traditionally, the SOS response is induced through UV irradiation or oxidative damage, however, we used the DNA cross-linking agent mitomycin C (MMC) to elicit a similar response in *Mtb* [13-14]. The classic components of a DNA damage-associated response were successfully up-regulated subsequent to MMC treatment, concurrent with a repression of cell division genes indicating a disruption in division is associated with this process in *Mtb,* similar to SOS program induction in other bacteria. However, *cdr* gene expression was not induced. This suggests a differential level of regulation for *cdr* than *sulA*, presumably accompanied by a differential activity of the two proteins encoded by these genes, and indicates that Cdr likely does not play a role in the DNA damage response of mycobacteria.

4.3 Cdr STIMULATES CELL GROWTH RELATIVE TO DIVISION

Few studies examining Cdr function have been performed, however among these select experiments a common theme is evident: alterations of Cdr protein level coincides with division defects [9-10]. To substantiate this idea in mycobacteria a constitutive *cdr*-overexpression system was engineered into *M. smegmatis* and *Mtb,* followed by phenotypic observation using SEM in conjunction with standard growth curve analysis by $O.D._{600nm}$ readings. As cell division proteins are known to exist in a very specific stoichiometry, alterations of cell division protein levels disturbs the relative balance of the

division machinery resulting in cell division defects, apparent through direct observation by the formation of long filamentous cell shapes [15]. Historically, this method was used to identify proteins involved in cell division in bacteria, which subsequently bear the title Fts standing for filament forming temperature-sensitive proteins [15]. Furthermore, the point at which replication is perturbed (early versus late) may be implied by the relative frequency of ultrastructural features indicative of Z-rings on these bacterial filaments. In general, cell division processes including chromosomal replication, segregation, and FtsZ polymerization to form the Z-ring constitute early division events, while processes subsequent to Z-ring formation, including the production of septal membrane and cell wall material as well as resolution of the daughter cells, constitute late division events. In *Mtb*, disruption of early division by the inhibition of FtsZ protein polymerization by small molecule inhibitors leads to the formation of a smooth filamentous phenotype, accompanied by a 10-fold or greater reduction in the frequency of Z-ring structures [16]. Conversely, similar inhibition of the late division protein FtsI results in filamentous cells displaying concentric Z-rings at normal levels [17].

With these criteria in mind, observations of *M. smegmatis* and *Mtb cdr* overexpressing strains suggested an involvement in late cell division events. Surprisingly, standard growth kinetic analysis of these strains did not support the conclusion that abundant Cdr disrupts division, as experimental strains grew at a rate comparable to controls. While filamentation may contribute to higher $O.D._{600nm}$ readings despite a lower cell density, the filamentation observed as a result of *cdr* overexpression was moderate and the growth rate maintained wild type levels even when enumerated by colony forming units (cfu). However, it is important to note that the conditions of growth for our *cdr* merodiploid were laboratory adapted and not representative of the conditions the bacteria would experience within the host. Future studies utilizing *in vitro* models of dormancy may be more appropriate to address the specific questions

of this study: whether Cdr is involved in the regulation of cell division during chronic infection and how this process might be involved in the transition into and out of a dormant state.

A possible explanation for filamentation while maintaining normal growth parameters may be an increase in cell growth rate relative to division cycle progression. While most rod shaped bacteria seem to coordinate their cell cycle relative to cell length, mycobacteria appear to coordinate on a time-based scale as synchronous replication may be maintained despite differential daughter cell lengths [4]. This hypothesis is consistent with the morphological characterization that identified elongated cells displaying a frequency of Z-rings that was comparable to wild type.

In an attempt to further delineate the level of involvement of Cdr in division, transcriptional profiling of cell cycle discriminant genes by qRT-PCR was undertaken in the *cdr* merodiploid *Mtb* strain. Previous studies by the Slayden lab performing similar profiling on *Mtb* following treatment with FtsZ and FtsI inhibitors have found a trend in the expression profile of cell division genes in response to early or late division disruption [16-17]. Following FtsZ inhibition, early division genes show increased expression while late division genes display repressed expression. In contrast, treatments with FtsI inhibitors cause an induction of late division genes and a repression of early division genes. Likely these responses are compensatory mechanisms employed by the cell in an attempt to overcome the inhibition and allow rapid continuation of growth once the inhibition in lifted. The Cdr-specific profile appeared strikingly similar to a FtsI-inhibited profile which might suggest a disruption of late division events, however this conclusion is inconsistent with our previous findings that division is not significantly inhibited by *cdr* overexpression. Alternatively, as late division proteins are involved in the processes of membrane and cell wall biogenesis and turnover, the Cdr-induced profile may reflect a stimulation of cellular elongation rather than an inhibition in late division. Further studies are required to delineate this difference.

4.4 Cdr-GFP LOCALIZES TO REGIONS OF GROWTH AND DIVISION

To further elucidate the role of Cdr in the cell, sub-cellular localization studies were undertaken in the fast-growing species *M. smegmatis* using an inducible fluorescent recombinant protein expression system in conjunction with fluorescent microscopy. Studies have shown that *M. smegmatis* shares the same division machinery and profile, including its own *cdr* ortholog, despite growing at a rate 8 times faster than *Mtb*, and thus *M. smegmatis* has commonly been used to study replication for the range of *Mycobacterium* sp. [4-5, 18-19]. Variable localization patterns were observed in logarithmically grown cultures, however universal observations include the formation of distinct Cdr-GFP foci, numbering between 2-4 foci per cell, found near polar or quarter-cell positions. These locations include regions of active growth and early division sites as quarter-cell positions are representative of future division sites during the subsequent round of division in the resultant daughter cells [18-19]. This is consistent with the conclusions of Joyce et al. who determined that the division site is accurately placed at mid-cell, the location of which is determined very early in the division process [19]. Despite somewhat contradictory conclusions obtained using nascent cell wall staining rather than recombinant PBP1 localization, it stands to reason that priming of the future mid-cell site may occur during late stages of the previous round of replication [5]. In addition, it was observed that Cdr-GFP foci co-localize in most cases with FM 4-64 Fx styryl dye hot-spots. FM 4-64 Fx is a fixable analog of the FM 4-64 lipophilic dye, which is marketed for the non-specific staining of membranes. However, due to a net positive charge of the dye molecules it displays preferential staining for anionic (acidic) lipids, which in bacterial membranes primarily includes phosphatidylglycerol (PG^{-1}) and cardiolipin (CL^{-2}) [20-21]. The co-occurrence of Cdr-GFP foci and FM 4-64 Fx hotspots as well as the increased presence of these hot-spots in the *cdr*::pMCSU7 strain indicate a connection between increased Cdr protein levels and anionic phospholipid domains.

As cell division is a dynamic, continuously active process, the localization of several cell division proteins has likewise been shown to be dynamic, displaying different localizations at different stages throughout the cell cycle [22-26]. With this in mind, recombinant *M. smegmatis* were sorted and analyzed according to progressive division stages, revealing a distinct and consistent localization pattern. Based on this assessment it appears that Cdr-GFP localizes near the poles early in the division process: prior to chromosomal segregation but subsequent to DNA replication (Figure 4.1, B 1). As segregation commences, the Cdr-GFP foci move to more extreme polar positions, displaying differential behavior in daughter cells of differential morphology (Figure 4.1, B 2). The occurrence of V-shaped cells is a unique feature of mycobacterial division where daughter cells from a single replication cycle may result in either a linear or V-shaped morphology, lending to morphological heterogeneity of the bacterial population [18]. This divisional heterogeneity has also been shown to be linked to differential growth rates and antibiotic susceptibility of the resultant daughter cells, and is believed to be a contributing factor to bacterial survival during dormancy and the associated difficulties in treating latent infections [4]. Early works suggested that creation of this V-shaped morphology, referred to as snapping, occurred during the final stages of cell division as a result of un-equal cell wall rupture during separation of the daughter cells, however these conclusions were primarily presumptions based upon electron microscopy of asynchronously dividing populations [18, 27-28]. More recent studies using time-lapse microscopy of single cells reveal that the formation of these V-shaped cells occurs from the poles during early elongation stages of division, analogous to Figure 3.12, B step 2 presented here [4, 19]. Furthermore, specific staining of novel cell wall and plasma membrane production indicates that these processes occur exclusively at the cell poles and outer edge of the V-elbow, indicating that this morphology is the result of increased synthesis on one side of the cell relative to the other [18]. Therefore, the splitting of Cdr-GFP foci observed along the outer edge of the elbow in (Figure 4.1, B 1) indicates a role for Cdr in the morphological determination of the daughter cells, as this phenomenon

was not observed in the linear daughter cell, which may contribute to antibacterial tolerance. After initiation of the snapping morphology Cdr-GFP appears to act similarly at either end relative to mid-cell, migrating from the poles to a quarter-cell position as the septum forms (Figure 4.1, B 3). As stated previously, once the Z-ring has been established, the cell division process is committed and the septum forms quite rapidly. Therefore, the quarter-cell positions of the daughter cells at this point (Figure 4.1, B 3) represent an early mid-cell position for the next round of replication that will soon commence. Lastly, as the daughter cells begin to separate the Cdr-GFP foci migrate back toward a polar position as novel foci develop at the newly formed cell poles (Figure 4.1, B 4). A simplified localization scheme has been provided to aid in visualization of the likely patterns that would be observed in a model bacillus such as *E. coli* compared with the unusual features identified in *Mycobacterium* sp. (Figure 4.2). It is important to note the differences in *E. coli* growth occurs from the middle of the cell while in Mycobacterium sp. growth occurs from the poles, therefore actual Cdr-GFP localizations may appear quite different in *E. coli* than suggested in Figure 4.1 A. The symmetrical division depiction is included to allow easier visualization of the dynamic division-associated localization patterns in a dividing bacillus while reducing the complexity from differential daughter cell morphologies.

4.5 Cdr INDUCES METABOLIC PROGRAMS INVOLVED IN LIPID SYNTHESIS

To explore the molecular mechanisms associated with Cdr activity, global transcriptional analysis was performed on the *cdr* merodiploid *Mtb* strain. After stringent significance cut-offs were applied a total of 277 ORFs showed statistically significant differential regulation in response to *cdr* overexpression (Appendix A). The activity of the proteins encoded by these genes relate to Cdr protein activity either directly, through protein interactions, or indirectly, downstream effects, and thus are

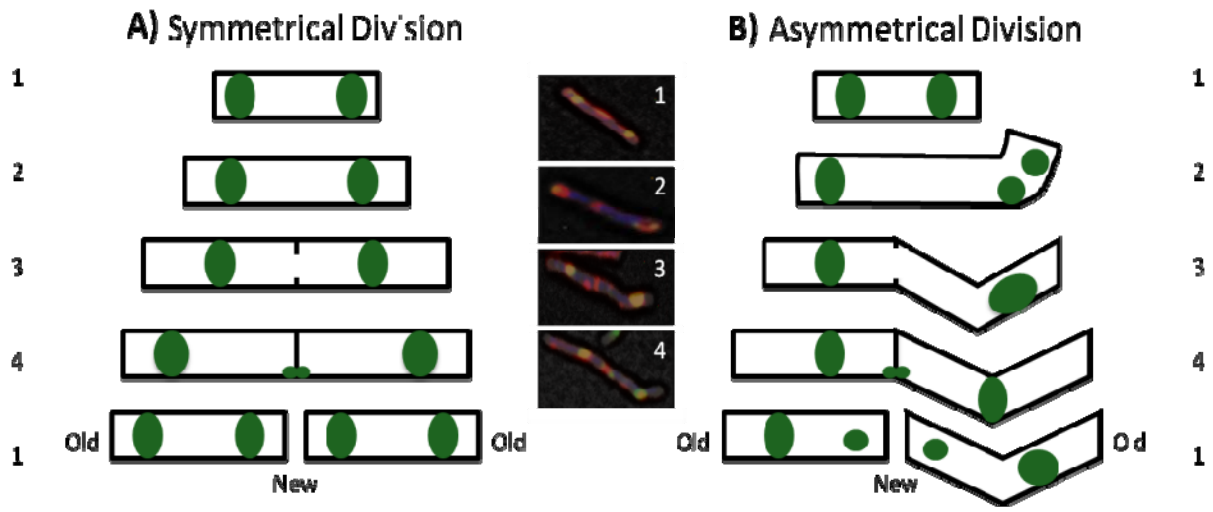

Figure 4.1: Dynamic localization scheme of Cdr-GFP.

A predicted localization pattern that may be observed in other bacilli that do not exhibit unique mycobacterial features of division (A) provided along-side a simplified view of the localizations of Cdr-GFP and morphologies observed in mycobacterial division (B). For more in-depth discussion please refer to text section 4.4.

indicative of the Cdr's effect on the cell. Similarly, the most differentially regulated functional classes included intermediary metabolism and respiration followed by cell wall processes, hypothetical and conserved hypothetical classes. This is similar to transcriptional profiles of *Mtb* during a chronic guinea pig infection model which show the most differentially regulated protein classes belonging to intermediary metabolism and respiration and cell wall processes [29]. The most drastically different profile was observed for genes encoding hypothetical and conserved hypothetical proteins indicating that *cdr* is likely a component of previously undefined cellular processes. As described in section 5.1, difficulties in accurate annotation often result in mis-annotation or non-annotation of ORFs. As of 2012 the H37Rv genome over 27% of the encoded ORFs were annotated as hypotheticals, therefore it is not surprising that a significant proportion of any cellular response is largely comprised of such proteins

[30]. Enzymes associated with intermediary metabolism and respiration were the next most differentially regulated, indicating that *cdr* activity elicits specific metabolic reprogramming. Pathway mapping revealed a near universal down-regulation of oxidative phosphorylation systems, possibly suggesting a negative association between the role of Cdr and aerobic respiration. In support of this, several hypoxia-specific genes were up-regulated including *dosR*, the acute hypoxia program regulator, and *nrdZ*, a hypoxia-induced type II ribonucleotide reductase [31-32].

Interestingly, alterations in lipid metabolism showed very specific patterns of induction. Specifically, *de novo* fatty acid synthesis showed widespread induction including the *fasI* fatty acid synthase, homologous to eukaryotic *fas* systems [33]. In mycobacteria, which encode FasI single-enzyme and FasII multi-enzyme systems, FasI is utilized primarily for *de novo* lipid synthesis of long-chain fatty acids (C16-C18) used for phospholipid and triacylglyceride (TAG) synthesis, and very long-chain fatty acids (C24-C26) which may be further elongated by FasII for the production of mycolic acid outer cell wall components [34]. The primary product of FasI are the saturated palmitic (16:0) and stearic (18:0) acids however, *desA3* desaturase was also up-regulated converting stearic acid to oleic acid (18:1) [35]. This provides the basic building blocks for *de novo* phospholipid synthesis as phospholipids are generally composed of a saturated and unsaturated acyl chain, of which palmitic and oleic acids are the most common at the *sn*-1 and *sn*-2 positions respectively [36-37].

Other enzymes associated with acyl-ACP production were significantly up-regulated, including *accA1*, *kasA*, and *plsB1*. Similar to the regulation of *fasI*, the activity of these enzymes is inhibited by an abundance of acyl-ACP, therefore a consistent induction of these programs, which should produce acyl-ACP, indicates the fatty acids produced by FasI are quickly being utilized by the cell [38]. The fate of these *de novo* synthesized fatty acids may be elucidated by the concurrent up-regulation of *plsB1*, a putative glycerol-3-phosphate acyltransferase (GPAT), the first step in glycerolipid synthesis.

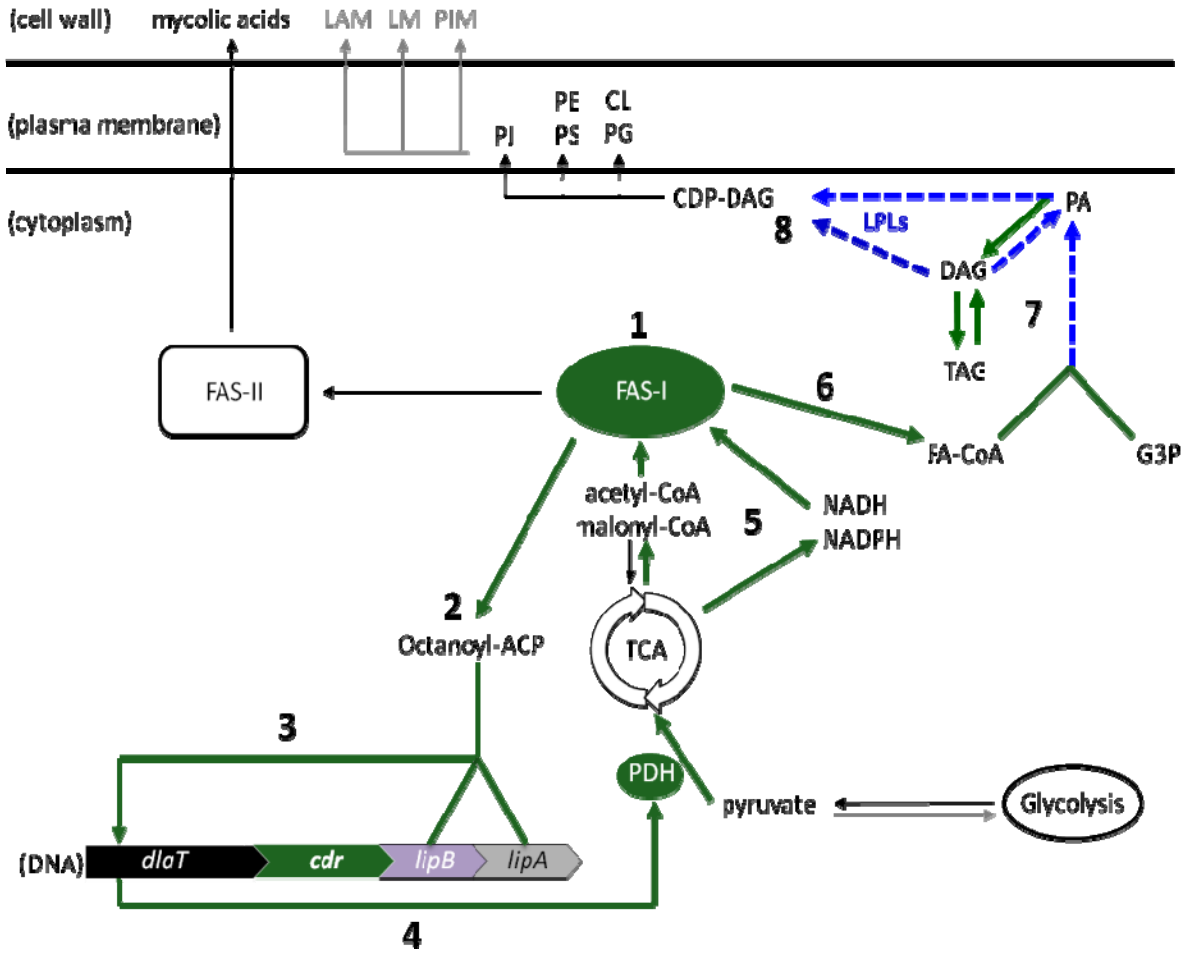

Figure 4.2: Regulatory scheme associated with Cdr activity in *Mtb*.

A visualization of the molecular programs differentially regulated in *cdr*::pVV16 *Mtb* associated with lipid synthesis as determined by global transcriptional profiling and pathways analysis. Green arrows indicate induced processes, grey indicates repressed processes and black indicates a lack of statistically significant differential regulation. Blue segmented arrows indicate sources of LPL production. Cdr overexpression up-regulated *fasI* mediated *de novo* fatty acid synthesis (1) to provide the medium chain fatty acids (2) utilized by operon partners LipBA to activate DlaT (3) component of the pryruvate dehydrogenase (PDH) complex (4). PDH and TCA cycle activity provide the metabolic substrates for continued FasI fatty acid (FA) production (5). FasI also produces long-chain monounsaturated fatty acids used during TAG and phospholipid synthesis (6). TAG turnover provides an additional pool of fatty acids and LPLs (7) for phospholipid acyl remodeling and synthesis (8).

This system is analogous to eukaryotic GPATs which have been shown to utilize glycerol-3-phosphate

(G3P) and long-chain acyl-CoA to produce lysophosphatidic acid (LPA, referred to in Figure 3.11 as 1(or

2)-Acy-*sn*-glycerol-3-phosphate), a precursor to phospholipid and TAG biosynthesis [39]. As seen in Figure 3.11, enzymatic programs creating phospholipids from lysophospholipids (LPLs) were specifically up-regulated, suggesting phospholipid remodeling or damage is occurring at this time, as LPLs are important metabolic intermediates during phospholipid metabolism and synthesis [40]. A down-regulation of *pgsA* phosphatidylinositol synthase, an enzyme required for the production of phosphatidylinositol (PI) which may be further processed into more complex lipids, may indicate a more specific focus on the production of other phospholipid moieties such as the anionic lipid classes PG and CL, consistent with our observation of FM 4-64 Fx staining patterns [41].

4.6 Cdr CONTEXT WITHIN THE *cdr* OPERON

Further delineation of the roles of Cdr may be provided by considering Cdr protein in the context of its operon partners. The polycistronic operons of prokaryotes allows coupling of intimately associated molecular programs by providing a single operator/promoter for the expression of adjacent genes [42]. Polycistronic operons are transcribed as a single mRNA, which is processed to produce separate proteins [42]. While in many cases operon partners may functionally interact, this is not a universal guarantee. However, their functionalities are temporally linked with each other providing clues as to their conditions of expression and associated molecular programs [42]. The compositions of these operons are even more important in intracellular pathogenic bacteria infecting a narrow host range, such as *Mtb*, as it has been observed that these bacteria show greatly reduced genome sizes, likely due to the predictability of the stresses experienced during their life cycle and the ability to obtain simple nutrients directly from host tissues [43]. Furthermore, the ancient and persistent presence of *Mtb* in the

human population has provided ample time for adaptation to these conditions resulting in transcriptional coupling of related genes [44].

The genomic context of the *cdr* operon was investigated and found to be highly conserved across mycobacteria. Characterized members of this operon include the upstream ORF *rv2215* encoding the dihydrolipoamide acytransferase *dlaT* and downstream ORFs *rv2217-rv2218* encoding the lipoate biosynthesis proteins *lipB* and *lipA* respectively. DlaT comprises the E2 component of two separate multi-protein complexes: 1) the pyruvate dehydrogenase complex (PDH), and essential component of intermediary metabolism which oxidizes pyruvate in acetyl CoA, linking glycolysis with the TCA cycle, and 2) the peroxynitrite reductase/peroxidase complex (PNR/P) required for the neutralization of highly toxic peroxynitrite species during infection [45-46]. *DlaT* knockout strains show reduced pathology in the mouse model of *Mtb* infection, with particular difficulties in establishing a persistent infection comparable to latent disease [47]. For these reasons DlaT has been considered as an attractive target for the treatment of LTBI, with DlaT-specific inhibitors showing activity against dormant bacilli [48]. This suggests a role for Cdr protein during standard growth conditions as well as during a prolonged and latent infection. The enzymes LipA and LipB act in concert to carry out the process of protein lipoylation, a post-translational modification that attaches a short (C8) acidic lipid called lipoic acid to a target protein, resulting in functional activation [49]. Specifically, LipB octanoyl transferase binds octanoyl-ACP, a medium chain lipid generated during FasI *de novo* fatty acid synthesis, and covalently attaches it to apoDlaT, after which LipA lipoate synthase converts octanoate to lipoate through the addition of two sulfur atoms [49]. DlaT is the only lipoylated protein identified in *Mtb* to date which raises the central question of Cdr's contribution or connection to the intimately associated molecular program involving DlaT and LipBA [47].

FasI induction by Cdr provides a direct benefit to its operon partners by stimulating the production of the octanoic acid precursors required for lipoylation and activation of DlaT. DlaT activity is essential for active metabolism, such as during periods of growth which require the production and insertion of new plasma membrane and cell wall material, of which the first step of production requires FasI-generated fatty acids [50]. Additionally, DlaT activity is required for complete pathogenesis during infection, a time that is associated with reduced bacillary growth and division [47]. The role of DlaT at this time is primarily related to the neutralization of peroxynitirite species, potent oxidizing agents generated from the cross-reaction of reactive oxygen intermediates (ROI) and reactive nitrogen intermediates (RNI) by immune effector cells [46]. The oxidative stresses experienced within the host can be extremely damaging to the cell wall and membrane, therefore modification and reparation of these structures would remain an essential process despite a lack of growth and division. The biochemical properties of acidic phospholipids, particularly CL which has a net -2 charge and 2 monounsaturated acyl chains, make them more susceptible to peroxidation, relative to our observations of Cdr-GFP at regions enriched in CL content [51-52]. As re-acylation enzymes catalyzing the regeneration of LPLs into phospholipids were induced by Cdr activity, it can be postulated that Cdr plays an important role in membrane regeneration and repair during infection.

In addition to the contribution by Cdr to membrane homeostasis, this activity may also play a role in preventing oxidative damage caused by excess FFA that are not linked to biosynthetic carrier proteins in the cell. Such FFA may be generated during acyl editing by lipases, such as PLA and Tgs, which cleave single acyl chains from phospholipids or TAG to create LPLs or DAG and a FFA moiety. It has been shown that these FFAs cause oxidative damage to the cell which may be further exacerbated by the presence of RNI [53-54]. This is supported by the observation that RNI and FFA have been shown to contribute to the immune-mediated killing and clearance of *Mtb* [54]. Thus the transcriptional coupling

of FFA utilization and neutralization of RNI-associated oxidizing agents, such as within the *dlaT-cdr* operon, would promote oxidative homeostasis for *Mtb* within the host. In addition to the oxidative damage experienced within the bacteria, PLA production of FFAs plays a central role in the respiratory burst defense strategy of macrophages, contributing to the clearance of several important pathogens including *Rhodococcus equi*, *P. aeruginosa*, and *Mtb* [55-56]. It has been shown that FFAs are abundant within the phagosomes of macrophages, a location in which the *Mtb* bacteria reside [57-58]. Potentially, the assimilation of excess FFAs into phospholipids and TAG would reduce their abundance, thereby limiting their deleterious effects on oxidative damage, providing an integral link between lipid metabolism and oxidative damage and further substantiating the connection of these processes in the *cdr* operon and the Cdr-induced transcriptional profile.

4.7 PHOSPHOLIPID COMPOSITION IS AN IMPORTANT CELLULAR REGULATORY STRATEGY

The phospholipid composition of the plasma membrane has far-reaching effects on cellular processes, serving as an important modulator for the spatial and temporal organization of the division and respiratory machinery [59]. The specific phospholipid composition has been shown to change in a cell-cycle dependent manner, and may influence cell division progression [21, 60]. The localization and activity of DnaA, the initiator of DNA replication which acts as the division initiation timer, is dependent upon acidic phospholipid domains [61-63]. Such domains have a similar localization and activation effect on MinD, spatial regulator of the MinC-based septum placement system, (however, no MinC homologs have been identified in mycobacteria) and in FtsY targeting of proteins to the membrane [59, 64-66]. Likewise, acidic phospholipids domains are required for the correct insertion and activation of respiratory complexes including the nitrate reductase complex, NarGHI required for alternative

respiration during hypoxia, and the cytochrome complexes III, IV and V energy production systems [67-69].

Studies primarily performed in *Saccharomyces cerevisiae* have shown an intimate connection between TAG and phospholipid metabolism [70-72]. This connection is not surprising since both products are produced from diacylglycerol (DAG) precursors and share a common acyl pool for synthesis and editing of these species [70-72]. It was found that reduction of TAG hydrolysis results in reduced phospholipid synthesis, and alterations in their acyl chain compositions [70]. In *Mtb*, TAG accumulation is specifically linked with pathological conditions associated with host infection and dormancy phenotypes [73]. Bacterial TAG and macrophage-generated lipid bodies (LBs), small pseudo-organelles consisting of phospholipids surrounding a layer of steryl esters (SE) and an inner layer of TAG along with some LB-specific proteins, serve as an important source of nutrition and lipids for Mtb during infection [74]. The production of these LBs is specifically induced by *Mtb* infection, leading to the development of foamy macrophages, termed such due to a high content of LBs in the cytoplasm lending them a bubbly appearance, and may share these lipid contents with the phagosomal compartment in-which the bacteria reside [74]. Furthermore, it has been shown that *Mtb* can access these lipids, break them down into smaller constituents, and re-form them into TAG within the bacterium [73]. This implicates TAG synthesis and degradation along with acyl editing as central processes during infection. In this study we found that overexpression of *cdr* up-regulates *tgs1* and *lipY*, encoded by *rv3130c* and *rv3097c*, the primary genes involved in TAG biosynthesis and degradation, respectively [75]. The concurrent synthesis and hydrolysis of TAG in the bacterium suggests TAG is being utilized as a fatty acid sink for alterations in the available acyl chains for phospholipid editing rather than a storage system or energy source, suggesting a potential link between the Cdr-induced cellular response and dormancy responses in *Mtb*.

Interestingly, *Bacteroides* sp. encode a modified Cdr protein that exists with a C-terminal fusion to a Lipocalin2 domain (PF08212) (uniprot Q89Z84). The *E. coli* Blc protein, is a single-domain protein encoding a Lipocalin2 domain, and is currently the only experimentally characterized bacterial protein of this type [76]. Blc, which stands for bacterial lipocalin, belongs to a family of proteins that generally display lipid-binding properties [77] . Investigations identified a specific preference by this protein to bind LPLs over phospholipids and FAs, with a slight preference for shorter alkyl chain lengths (palmitic C16 > oleic C18) [76]. In addition, expression of this protein has been primarily observed during stationary growth and osmotic stress, times that are associated with cell envelope stress [78]. The un-cleaved N-terminal signal sequence of *Mtb* Cdr is believed to be involved in the specific targeting of the protein to the membrane, and indeed it has been found to be enriched in the membrane fraction [79-80]. These observations implicate a functional role for Cdr at the membrane where it may act to monitor glycerophospholipid metabolism.

4.8 Cdr LITERATURE AND CONCLUSIONS

Through phospholipid turnover and acyl editing it is possible for a bacterium to alter its plasma membrane composition and fluidity, enhancing survival under temperature and osmotic stresses. Phospholipid editing may be achieved through the *de novo* synthesis of phosphate head groups and acyl chains for the production of novel phospholipids, as well as by acyl editing and interconversion of existing phospholipid species. Although interconversion programs were generally down-regulated upon *cdr* overexpression, *de novo* phospholipid synthesis was up-regulated along with TAG hydrolysis, indicating supplementation of the free acyl pool with TAG-derived lipids. This process is concurrent with the continual production of FFA and LPL moieties, both of which have been found to be inhibitive to

growth and division in large quantities [50, 57, 60]. The maintenance of division kinetics observed upon *cdr* overexpression may be explained by the rapid assimilation of these species during phospholipid and TAG turnover.

The studies presented here are the first to examine the Cdr protein gene function in significant detail. Despite the prevalence of these proteins in biological systems, very little is known about the nature of these unusual enzymes. The limited yet significant similarity of these proteins with SulA has led to frequent mis-annotations and incorrect assumptions as to the role these proteins play in the cell. Currently only two independent experimental studies have been performed on Cdr-type proteins, both of which focus on the homolog in *Arabidopsis thaliana,* intending to describe a SulA protein in this species (which we will refer to as AtCdr) [9-10]. Although much of the evidence obtained from these two studies is contradictory, some common conclusions can be made. Firstly, AtCdr is a nuclear-encoded plastid-targeted protein that is exclusively expressed in tissues containing actively dividing plastids, indicating an involvement in the plastid division process [9-10]. Second, AtCdr does not behave in a manner akin to SulA division inhibitors, substantiated by a lack of demonstrable interaction with either AtFtsZ-1 or AtFtsZ-2 [9-10]. Finally, AtCdr protein show differential accumulation between plastids, with more extreme changes in protein level correlating with more severe division defects [9-10]. Maple et al. published first in May 2004 that AtCdr must act as a positive division regulator as overexpression failed to induce division defects, while knockdown of protein levels by 90% or more caused heterologous division defects resulting in the reduction of chloroplasts per cell from near 80 to 1-2 in a majority of the experimental plant lines [10]. Interestingly, this knockdown was the result of an attempted overexpression system that resulted in co-suppression of both the exogenous and endogenous Atcdr genes, while their attempts to knockdown expression using antisense Atcdr was insufficient to elicit phenotypic changes despite an almost 70% reduction in expression levels [10]. Later, Raynaud et al.

published a similar paper on this same protein using an *A. thaliana* background for their overexpression system and a *Synechocystsis* sp. PCC6803 background for knockout studies [9]. Although unable to form a homogenous knockout mutant, heterodiploid knockout mutants were generated and displayed an inability to completing late-stage cell division, indicating this protein may be involved in closure at the division site and separation of the daughter plastids [9]. Using an overexpression system similar to Maple et al. they observed plastid division defects that reduced the numbers of plastids per cell in a heterogonous manner [9]. Importantly, similar levels of Atcdr mRNA were detected in experimental and control strains however AtCdr protein was undetectable in the overexpression strain [9]. They concluded that this was likely due to translational or post-translational degradation of the protein, which indicates that this strain may have been experiencing a Cdr deficiency rather than an abundance [9]. It seems that Cdr protein levels are monitored and controlled by the cell via extensive post-transcriptional methods. The mammalian Cdr proteins display several different methods of post-translational modification including phosphorylation, uniquitylation and acetylation, the last two of which generally act as signals for proteolytic degradation [81-83]. This further supports the important role for Cdr proteins in the cell, as the extensive level of regulation indicates tight control of protein levels are required for optimal functioning. Significantly, the summation of the above conclusions are consistent with the activity of Cdr protein as a regulator of phospholipid remodeling, as the ratio of acidic phospholipids in the membrane may directly affect the success of division events [60].

Our investigations into the nature of the Cdr protein encoded by *rv2216*, will provide a foundation for a deeper understanding of this important yet poorly described family of proteins. The importance of the Cdr protein function is demonstrated by its substantial presence in diverse organisms and high degree of conservation at the sequence level, as well as by the extensive methods of regulation observed in mammalia. The role of Cdr in *Mtb* demonstrate some unique characteristics as our data

suggests it plays a role in the mycobacteria-specific process of morphological heterogeneity, and currently stands as the first protein identified as participating directly in this unique phenomenon. Furthermore, the Cdr-induced transcriptional profile suggests its role in *Mtb* is linked to conditions associated with dormant phenotypes, such as those displayed during latent infections. The proposed role for Cdr in influencing phospholipid and neutral lipid metabolism remains relevant in diverse life forms, from prokaryotes through higher eukarya. Although much work remains to be done to elucidate the mechanism of Cdr activity, the work presented here is the first to explore a functional role for Cdr by examining its effect on cellular transcription. It is our belief that this work provides valuable insight into elucidating the complex regulatory strategies utilized in diverse life forms by characterizing this novel family of regulators. A deeper understanding of the functions and roles of Cdr proteins would not only aid in the characterization of dormant phenotypes in mycobacteria that contribute to persistent infections and treatment failure, but also to host defense mechanisms associated with lipid metabolite inflammatory mediators as the human cdr is associated with these processes [11]. Although further studies are needed, the body of work presented here provides the foundation for targeted studies aimed at elucidating the precise roles of this family of important proteins.

REFERENCES

1. Aylward, B., et al., *When is a disease eradicable? 100 years of lessons learned.* Am J Public Health, 2000. **90**(10): p. 1515-20.
2. *Prospect of tuberculosis eradication.* Public Health Rep, 1963. **78**(6): p. 507-509.
3. Sacchettini, J.C., E.J. Rubin, and J.S. Freundlich, *Drugs versus bugs: in pursuit of the persistent predator Mycobacterium tuberculosis.* Nat Rev Microbiol, 2008. **6**(1): p. 41-52.
4. Aldridge, B.B., et al., *Asymmetry and aging of mycobacterial cells lead to variable growth and antibiotic susceptibility.* Science, 2012. **335**(6064): p. 100-4.
5. Singh, B., et al., *Asymmetric growth and division in Mycobacterium spp.: compensatory mechanisms for non-medial septa.* Mol Microbiol, 2013. **88**(1): p. 64-76.
6. Adams, D.W. and J. Errington, *Bacterial cell division: assembly, maintenance and disassembly of the Z ring.* Nat Rev Microbiol, 2009. **7**(9): p. 642-53.
7. Kirkpatrick, C.L. and P.H. Viollier, *New(s) to the (Z-)ring.* Curr Opin Microbiol, 2011. **14**(6): p. 691-7.
8. Justice, S.S., et al., *Morphological plasticity as a bacterial survival strategy.* Nat Rev Microbiol, 2008. **6**(2): p. 162-8.
9. Raynaud, C., et al., *An Arabidopsis homolog of the bacterial cell division inhibitor SulA is involved in plastid division.* Plant Cell, 2004. **16**(7): p. 1801-11.
10. Maple, J., et al., *GIANT CHLOROPLAST 1 is essential for correct plastid division in Arabidopsis.* Curr Biol, 2004. **14**(9): p. 776-81.
11. Gallwitz, M., J.M. Reimer, and L. Hellman, *Expansion of the mast cell chymase locus over the past 200 million years of mammalian evolution.* Immunogenetics, 2006. **58**(8): p. 655-69.
12. Miyagishima, S.Y., *Mechanism of plastid division: from a bacterium to an organelle.* Plant Physiol, 2011. **155**(4): p. 1533-44.
13. Shinagawa, H., *SOS response as an adaptive response to DNA damage in prokaryotes.* EXS, 1996. **77**: p. 221-35.
14. Rand, L., et al., *The majority of inducible DNA repair genes in Mycobacterium tuberculosis are induced independently of RecA.* Mol Microbiol, 2003. **50**(3): p. 1031-42.
15. Harry, E., L. Monahan, and L. Thompson, *Bacterial cell division: the mechanism and its precison.* Int Rev Cytol, 2006. **253**: p. 27-94.
16. Respicio, L., et al., *Characterizing septum inhibition in Mycobacterium tuberculosis for novel drug discovery.* Tuberculosis (Edinb), 2008. **88**(5): p. 420-9.
17. Slayden, R.A. and J.T. Belisle, *Morphological features and signature gene response elicited by inactivation of FtsI in Mycobacterium tuberculosis.* J Antimicrob Chemother, 2009. **63**(3): p. 451-7.
18. Thanky, N.R., D.B. Young, and B.D. Robertson, *Unusual features of the cell cycle in mycobacteria: polar-restricted growth and the snapping-model of cell division.* Tuberculosis (Edinb), 2007. **87**(3): p. 231-6.
19. Joyce, G., et al., *Cell division site placement and asymmetric growth in mycobacteria.* PLoS One, 2012. **7**(9): p. e44582.
20. Muchova, K., A.J. Wilkinson, and I. Barak, *Changes of lipid domains in Bacillus subtilis cells with disrupted cell wall peptidoglycan.* FEMS Microbiol Lett, 2011. **325**(1): p. 92-8.
21. Barak, I. and K. Muchova, *The role of lipid domains in bacterial cell processes.* Int J Mol Sci, 2013. **14**(2): p. 4050-65.

22. Weiss, D.S., et al., *Localization of FtsI (PBP3) to the septal ring requires its membrane anchor, the Z ring, FtsA, FtsQ, and FtsL.* J Bacteriol, 1999. **181**(2): p. 508-20.
23. Plocinska, R., et al., *Septal localization of the Mycobacterium tuberculosis MtrB sensor kinase promotes MtrA regulon expression.* J Biol Chem, 2012. **287**(28): p. 23887-99.
24. Dasgupta, A., et al., *The serine/threonine kinase PknB of Mycobacterium tuberculosis phosphorylates PBPA, a penicillin-binding protein required for cell division.* Microbiology, 2006. **152**(Pt 2): p. 493-504.
25. Dziedzic, R., et al., *Mycobacterium tuberculosis ClpX interacts with FtsZ and interferes with FtsZ assembly.* PLoS One, 2010. **5**(7): p. e11058.
26. Ginda, K., et al., *ParA of Mycobacterium smegmatis co-ordinates chromosome segregation with the cell cycle and interacts with the polar growth determinant DivIVA.* Mol Microbiol, 2013. **87**(5): p. 998-1012.
27. Dahl, J.L., *Electron microscopy analysis of Mycobacterium tuberculosis cell division.* FEMS Microbiol Lett, 2004. **240**(1): p. 15-20.
28. Hett, E.C. and E.J. Rubin, *Bacterial growth and cell division: a mycobacterial perspective.* Microbiol Mol Biol Rev, 2008. **72**(1): p. 126-56, table of contents.
29. Kruh, N.A., et al., *Portrait of a pathogen: the Mycobacterium tuberculosis proteome in vivo.* PLoS One, 2010. **5**(11): p. e13938.
30. Doerks, T., et al., *Annotation of the M. tuberculosis hypothetical orfeome: adding functional information to more than half of the uncharacterized proteins.* PLoS One, 2012. **7**(4): p. e34302.
31. Trauner, A., et al., *The dormancy regulator DosR controls ribosome stability in hypoxic mycobacteria.* J Biol Chem, 2012. **287**(28): p. 24053-63.
32. Dawes, S.S., et al., *Ribonucleotide reduction in Mycobacterium tuberculosis: function and expression of genes encoding class Ib and class II ribonucleotide reductases.* Infect Immun, 2003. **71**(11): p. 6124-31.
33. Fernandes, N.D. and P.E. Kolattukudy, *Cloning, sequencing and characterization of a fatty acid synthase-encoding gene from Mycobacterium tuberculosis var. bovis BCG.* Gene, 1996. **170**(1): p. 95-9.
34. Gago, G., et al., *Fatty acid biosynthesis in actinomycetes.* FEMS Microbiol Rev, 2011. **35**(3): p. 475-97.
35. Zheng, C.J., et al., *Fatty acid synthesis is a target for antibacterial activity of unsaturated fatty acids.* FEBS Lett, 2005. **579**(23): p. 5157-62.
36. Chang, Y. and B.G. Fox, *Identification of Rv3230c as the NADPH oxidoreductase of a two-protein DesA3 acyl-CoA desaturase in Mycobacterium tuberculosis H37Rv.* Biochemistry, 2006. **45**(45): p. 13476-86.
37. Mazzella, N., et al., *Bacterial phospholipid molecular species analysis by ion-pair reversed-phase HPLC/ESI/MS.* J Lipid Res, 2004. **45**(7): p. 1355-63.
38. Rock, C.O. and S. Jackowski, *Forty years of bacterial fatty acid synthesis.* Biochem Biophys Res Commun, 2002. **292**(5): p. 1155-66.
39. Wendel, A.A., T.M. Lewin, and R.A. Coleman, *Glycerol-3-phosphate acyltransferases: rate limiting enzymes of triacylglycerol biosynthesis.* Biochim Biophys Acta, 2009. **1791**(6): p. 501-6.
40. Bergethon, P.R., *The Physical Basis of Biochemistry: The Foundations of Molecular Biophysics*. 2010: Springer New York.
41. Jackson, M., D.C. Crick, and P.J. Brennan, *Phosphatidylinositol is an essential phospholipid of mycobacteria.* J Biol Chem, 2000. **275**(39): p. 30092-9.
42. Snyder, L. and W. Champness, *Molecular genetics of bacteria*. 2007: Amer Society for Microbiology.

43. Moran, N.A., *Microbial minimalism: genome reduction in bacterial pathogens.* Cell, 2002. **108**(5): p. 583-6.
44. Hershkovitz, I., et al., *Detection and molecular characterization of 9,000-year-old Mycobacterium tuberculosis from a Neolithic settlement in the Eastern Mediterranean.* PLoS One, 2008. **3**(10): p. e3426.
45. Tian, J., et al., *Mycobacterium tuberculosis appears to lack alpha-ketoglutarate dehydrogenase and encodes pyruvate dehydrogenase in widely separated genes.* Mol Microbiol, 2005. **57**(3): p. 859-68.
46. Bryk, R., et al., *Metabolic enzymes of mycobacteria linked to antioxidant defense by a thioredoxin-like protein.* Science, 2002. **295**(5557): p. 1073-7.
47. Shi, S. and S. Ehrt, *Dihydrolipoamide acyltransferase is critical for Mycobacterium tuberculosis pathogenesis.* Infect Immun, 2006. **74**(1): p. 56-63.
48. Bryk, R., et al., *Selective killing of nonreplicating mycobacteria.* Cell Host Microbe, 2008. **3**(3): p. 137-45.
49. Spalding, M.D. and S.T. Prigge, *Lipoic acid metabolism in microbial pathogens.* Microbiol Mol Biol Rev, 2010. **74**(2): p. 200-28.
50. Nohturfft, A. and S.C. Zhang, *Coordination of lipid metabolism in membrane biogenesis.* Annu Rev Cell Dev Biol, 2009. **25**: p. 539-66.
51. Dacaranhe, C.D. and J. Terao, *A unique antioxidant activity of phosphatidylserine on iron-induced lipid peroxidation of phospholipid bilayers.* Lipids, 2001. **36**(10): p. 1105-10.
52. Schlame, M., *Cardiolipin synthesis for the assembly of bacterial and mitochondrial membranes.* J Lipid Res, 2008. **49**(8): p. 1607-20.
53. Akaki, T., et al., *Effector molecules in expression of the antimicrobial activity of macrophages against Mycobacterium avium complex: roles of reactive nitrogen intermediates, reactive oxygen intermediates, and free fatty acids.* J Leukoc Biol, 1997. **62**(6): p. 795-804.
54. Akaki, T., et al., *Comparative roles of free fatty acids with reactive nitrogen intermediates and reactive oxygen intermediates in expression of the anti-microbial activity of macrophages against Mycobacterium tuberculosis.* Clin Exp Immunol, 2000. **121**(2): p. 302-10.
55. Adolph, S., et al., *Unsaturated fatty acids as modulators of macrophage respiratory burst in the immune response against Rhodococcus equi and Pseudomonas aeruginosa.* Free Radic Biol Med, 2012. **52**(11-12): p. 2246-53.
56. Shimizu, T. and H. Tomioka, *[Novel type of antimicrobial mechanism in host macrophages against mycobacterial infections].* Nihon Hansenbyo Gakkai Zasshi, 2009. **78**(3): p. 283-91.
57. Freeman, B.A. and W.S. Lynn, *Fatty acid secretion and metabolism in 'activated' rabbit alveolar macrophages.* Biochim Biophys Acta, 1980. **620**(3): p. 528-37.
58. Kanetsuna, F., *Bactericidal effect of fatty acids on mycobacteria, with particular reference to the suggested mechanism of intracellular killing.* Microbiol Immunol, 1985. **29**(2): p. 127-41.
59. Lam, V.Q., et al., *Lipid activation of the signal recognition particle receptor provides spatial coordination of protein targeting.* J Cell Biol, 2010. **190**(4): p. 623-35.
60. Michel, G.P., et al., *Is there a correlation between membrane phospholipid metabolism and cell division?* Ann Inst Pasteur Microbiol, 1985. **136A**(1): p. 111-8.
61. Yamamoto, K., M. Rajagopalan, and M. Madiraju, *Phospholipids promote dissociation of ADP from the Mycobacterium avium DnaA protein.* J Biochem, 2002. **131**(2): p. 219-24.
62. Yamamoto, K., et al., *Modulation of Mycobacterium tuberculosis DnaA protein-adenine-nucleotide interactions by acidic phospholipids.* Biochem J, 2002. **363**(Pt 2): p. 305-11.
63. Maloney, E., et al., *Localization of acidic phospholipid cardiolipin and DnaA in mycobacteria.* Tuberculosis (Edinb), 2011. **91 Suppl 1**: p. S150-5.

64. Mazor, S., et al., *Mutual effects of MinD-membrane interaction: II. Domain structure of the membrane enhances MinD binding.* Biochim Biophys Acta, 2008. **1778**(11): p. 2505-11.
65. de Leeuw, E., et al., *Anionic phospholipids are involved in membrane association of FtsY and stimulate its GTPase activity.* EMBO J, 2000. **19**(4): p. 531-41.
66. Erez, E., et al., *Genetic evidence for functional interaction of the Escherichia coli signal recognition particle receptor with acidic lipids in vivo.* J Biol Chem, 2010. **285**(52): p. 40508-14.
67. Arias-Cartin, R., et al., *Cardiolipin-based respiratory complex activation in bacteria.* Proc Natl Acad Sci U S A, 2011. **108**(19): p. 7781-6.
68. Pinheiro, T.J. and A. Watts, *Lipid specificity in the interaction of cytochrome c with anionic phospholipid bilayers revealed by solid-state 31P NMR.* Biochemistry, 1994. **33**(9): p. 2451-8.
69. Subramanian, M., A. Jutila, and P.K. Kinnunen, *Binding and dissociation of cytochrome c to and from membranes containing acidic phospholipids.* Biochemistry, 1998. **37**(5): p. 1394-402.
70. Rajakumari, S., R. Rajasekharan, and G. Daum, *Triacylglycerol lipolysis is linked to sphingolipid and phospholipid metabolism of the yeast Saccharomyces cerevisiae.* Biochim Biophys Acta, 2010. **1801**(12): p. 1314-22.
71. Gaspar, M.L., et al., *Coordination of storage lipid synthesis and membrane biogenesis: evidence for cross-talk between triacylglycerol metabolism and phosphatidylinositol synthesis.* J Biol Chem, 2011. **286**(3): p. 1696-708.
72. Mora, G., M. Scharnewski, and M. Fulda, *Neutral lipid metabolism influences phospholipid synthesis and deacylation in Saccharomyces cerevisiae.* PLoS One, 2012. **7**(11): p. e49269.
73. Daniel, J., et al., *Mycobacterium tuberculosis uses host triacylglycerol to accumulate lipid droplets and acquires a dormancy-like phenotype in lipid-loaded macrophages.* PLoS Pathog, 2011. **7**(6): p. e1002093.
74. Melo, R.C. and A.M. Dvorak, *Lipid body-phagosome interaction in macrophages during infectious diseases: host defense or pathogen survival strategy?* PLoS Pathog, 2012. **8**(7): p. e1002729.
75. Deb, C., et al., *A novel lipase belonging to the hormone-sensitive lipase family induced under starvation to utilize stored triacylglycerol in Mycobacterium tuberculosis.* J Biol Chem, 2006. **281**(7): p. 3866-75.
76. Campanacci, V., et al., *The membrane bound bacterial lipocalin Blc is a functional dimer with binding preference for lysophospholipids.* FEBS Lett, 2006. **580**(20): p. 4877-83.
77. Bishop, R.E., *The bacterial lipocalins.* Biochim Biophys Acta, 2000. **1482**(1-2): p. 73-83.
78. Bishop, R.E., et al., *Stationary phase expression of a novel Escherichia coli outer membrane lipoprotein and its relationship with mammalian apolipoprotein D. Implications for the origin of lipocalins.* J Biol Chem, 1995. **270**(39): p. 23097-103.
79. Malen, H., et al., *Definition of novel cell envelope associated proteins in Triton X-114 extracts of Mycobacterium tuberculosis H37Rv.* BMC Microbiol, 2010. **10**: p. 132.
80. de Souza, G.A., et al., *Bacterial proteins with cleaved or uncleaved signal peptides of the general secretory pathway.* J Proteomics, 2011. **75**(2): p. 502-10.
81. Hornbeck, P.V., et al., *PhosphoSitePlus: a comprehensive resource for investigating the structure and function of experimentally determined post-translational modifications in man and mouse.* Nucleic Acids Res, 2012. **40**(Database issue): p. D261-70.
82. Wagner, S.A., et al., *A proteome-wide, quantitative survey of in vivo ubiquitylation sites reveals widespread regulatory roles.* Mol Cell Proteomics, 2011. **10**(10): p. M111 013284.
83. Danielsen, J.M., et al., *Mass spectrometric analysis of lysine ubiquitylation reveals promiscuity at site level.* Mol Cell Proteomics, 2011. **10**(3): p. M110 003590.

CHAPTER 5: MATERIALS & METHODS

5.1 HISTORY OF PREDICTING UNANNOTATED OPEN READING FRAMES

The advent of rapid and accurate whole-genome sequencing techniques have facilitated a great number of advances in the field of cellular biology. Bacterial genomes have been the most widely studied due to their impact on public health and applications for biotechnology, as well as their moderate genome size. Unfortunately, the usefulness of such data is limited by the functional annotation of the open reading frames (ORFs) comprising these genomes. Considering that a single bacterial genome is on average composed of around 4,000 ORFs, such studies have led to the development of massive genomic databases containing exponentially increasing quantities of putatively identified ORFs which comprise the large majority of all currently sequenced genomes.

The general annotation pipeline from genomic sequence involves an initial detection of potential ORFs followed by a functional annotation based upon nucleotide sequence similarity to characterized genes in other species. Most of this process is automated, where researchers can scan a genome through a program that will identify potential coding sequences as well as perform basic similarity searches to assign putative functions. Putative annotations are based upon the function of characterized proteins sharing sequence similarity with the protein of interest and assume that sequence similarity indicates homology, a qualitative term that refers to a common ancestry between a set of proteins or genes [1-2]. Homologous sequences may be orthologs or paralogs depending upon whether the evolutionary divergence occurred as a result of speciation or gene duplication, respectively [1-2]. While homology does not necessarily infer a common function, similarities in functionalities are often observed for homologous proteins [1-3]. Shared functionality is more common among orthologs than paralogs since duplicate genes within the same genome do not experience the same selective pressure and thus are free to accumulate mutations conferring unique functionalities [1-2]. Early

assessments of homology center upon quantifiable values of sequence similarity above significance cut-offs. It has been determined that 20-30% proteins sequence similarity may be shared between homologous proteins from distantly related organisms [3-4]. Therefore, we chose a 20% shared protein similarity cut-off, equivalent to roughly 90% similarity at the nucleotide level, to prevent the loss of potentially distantly related homologs. The notable difference between similarity cut-offs for protein and gene pair-wise identities is due to the redundancy of the code, where the 25 amino acid code of protein sequences reduces the likelihood of random similarity compared to the four nucleic acid code of gene sequences [4]. Importantly, setting the cut-off value this low also increases the likelihood of false-positive identifications, therefore putatively identified regulators must be verified more in-depth analysis.

Protein function is most accurately predicted from similarity in protein structure. While it is generally true that two proteins sharing extensive sequence similarity will adopt the same three-dimensional structure, this is not always the case since many proteins possess multiple functional domains, the combination and orientation of which may confer unique functionalities [4]. Proteins sharing almost no sequence identity may adopt similar functional folds and perform comparable roles within the cell [4]. Furthermore, there is always the risk of similarity by chance for small proteins during homology searches, such as those performed by annotation software [3-4]. While current annotation methods provide a quick and simple indication of an organism's proteomic potential, incorrect annotations or non-annotations are incredibly common, yet performing more in-depth analyses is too cumbersome to perform on a whole-genome scale. Often, correct annotation and assessment of poorly described ORFs falls onto individual researchers.

5.2 IDENTIFICATION OF PUTATIVE REGULATORS

Annotated cell division regulator protein sequences were obtained from the NCBI Reference Sequence (RefSeq) Database [5]. These sequences were compiled to generate consensus models using the MAFFT algorithm through Jalview V2 software with separate multiple alignments for annotated MinD, MinC, MinE, MipZ, YneA and SulA proteins [6]. These consensus models, which act as representative protein sequences for their protein family, were then used to search the H37Rv genome using a pBLAST algorithm and top hits were searched back against the RefSeq Database using a reciprocal best hits approach. Similarity assessments were performed at the protein level when applicable to increase stringency with significance cut-offs of $\geq$20% identity. Importantly, although the NCBI RefSeq database provides a comprehensive collection of curated non-redundant protein sequences, a general lack of experimental evidence corroborating protein classifications and a high frequency of mis-annotations reduce the strength of generated consensus models and increase the likelihood of false-positive identification of homologs. For this reason the top hits were further examined by orthological analysis, yet considered possible homologs of FtsZ regulators in *Mtb*.

5.3 DETERMINATIONS OF ORTHOLOGY

The Orthology Matrix (OMA) browser repository (Dec 2012) collects non-redundant, highly accurate whole genome sequences, contained over 1,300 fully sequenced genomes including 135 eukaryotes, 107archaea, and 1,078 bacteria [7]. The OMA Browser is a web-based computational matrix that aims to identify and define orthologous and paralogous relationships between proteins within an extensive reference dataset [7]. The OMA dataset is comprised of over 1000 completely sequenced genomes obtained from GenomeReviews, Ensembl, EnsemblGenomes and JGI, and is comprehensive of prokaryotic, eukaryotic and archaeal origins [7]. This is one of the most extensive and complete

orthology datasets available and is modified regularly to include newly sequenced genomes and update current sequences as more accurate sequencing technology becomes available. Additionally, the computational matrix of the OMA Browser accounts for pair-wise sequence similarities as well as evolutionary distances between the organisms of interest and is highly stringent in its assessment [7]. A unique feature of the OMA Browser is the formation of orthologous groups of proteins, where-in each group contains only one protein per species and each pair of proteins within the group is orthologous [7]. Therefore, to create stronger and more dependable reference models consensus models were generated from OMA groups containing functionally characterized MinD, MinC, MinE, MipZ, YneA and SulA proteins respectively. These models were again used to pBLAST the *Mtb* H37Rv proteome to ensure the initial identification of ORFs encoding putative FtsZ regulators. Top hits from the initial screen were examined through OMA Browser and cross-compared against parent OMA groups as well as reference OMA groups to assess the relationship between putatively identified mycobacterial regulators and experimentally characterized regulators from other species.

Confirmed hits were assessed for placement within polycistronic operons to elucidate correlated molecular programs and associated conditions of expression using the Operon DataBase (OperonDB), a computational database evaluating placement, distance and phylogenetic conservation of adjacent ORFs to predict operons, as well as the Operon Browser feature of the TB DataBase (TBDB) which predicts operons based upon expression correlation [8-9]. This was used to provide a context of molecular programs associated with the function of our putative regulators as members of polycictronic operons are known to be co-transcribed and therefore expressed under the same cellular conditions [10].

In addition to pair-wise sequence identity, protein homology may be assessed by similarities in functional domains, and tertiary structures as well as the presence or absence of insertions and deletions in the coding sequence. To identity and map protein functional domains, putative regulatory

sequences were submitted to the Protein families database (Pfam) sequence search [11]. Pfam is free web-based software that utilizes a massive cross-comparison of UniProtKB and NCBI RefSeq protein sequences using the HMMER3 program to identify conserved protein domains [11]. A protein domain is intended to represent a functional unit, involved in a particular aspect of the work of that molecule. Different combinations of domains and their relative orientations, also known as domain architecture, are believed to generate particular and unique functionalities. Pfam allows the identification of such functional domains within target sequences with quantifiable similarity values and provides relevant information regarding phylogenetic conservation, UniProtKB and NCBI RefSeq links to similar proteins and a curated description of protein family features and activities when possible [11].

To further corroborate our identifications and to help elucidate protein homology and functionality, putative mycobacterial regulators were also submitted to the protein fold recognition software Phyre2 (Protein Homology/analogY Recognition Engine V2.0) [12]. Unique to the Phyre2 engine is its comprehensive algorithm scoring protein similarity based upon a combination of pair-wise sequence identity, the absence or presence and frequency of sequence insertions or deletions, functional domains, and tertiary structure of mappable protein folds [12]. This allows for some assessment of orthology, but is most important for functionality assessment and structural prediction of novel proteins.

5.4 GROWTH AND CONSTRUCTION OF RECOMBINANT STRAINS

Mtb H37Rv and *M. smegmatis* mc^2155 were used for strain construction. In all cases, mycobacteria were grown at 37^oC on Middlebrook 7H11 agar media containing 10% OADC and 0.2% glycerol, or with shaking in Middlebrook 7H9 liquid media broth containing 10% ADC, 0.05% Tween80 and 0.2% glycerol.

Strains overexpressing our putative regulators were generated to examine the effects of disrupting these protein levels in the cell as stoichiometric disturbances of cell division proteins are known to elicit specific phenotypic characteristics in bacteria [13-14]. For the creation of these strains putative regulator coding sequences (ORFs: *rv2216*, *rv1708*) were PCR amplified using Accuprime pfx DNA polymerase (Invitrogen™) from purified genomic DNA and cloned into the NdeI and HindIII site of the extrachromosomal mycobacterial constitutive expression vector pVV16 (Appendices C and D). Constructs were transformed into One Shot® TOP10 chemically competent *E. coli* (Invitrogen™) and grown on LB agar containing 50 µg/ml kanamycin for selection. Successful constructs, as confirmed by sequence analysis, were transformed into electro-competent *Mtb* and *M. smegmatis* and grown in the presence of 25 µg/ml kanamycin for selection.

Fluorescently tagged protein expression strains were created for sub-cellular localization studies using the pMCSU7 vector (Appendix D). This is an extrachromosomal GFP fusion Gateway-compatible recombination vector containing the *Streptomyces coelicolor tetO* promoter derived from plasmid tcp3, as created by the Slayden laboratory. Putative regulator CDSs were PCR amplified as described previously, and cloned into pMCSU7 using Gateway technology (BP Clonase Enzyme Mix Invitrogen™ 11789-021, LR Clonase Enzyme Mix Invitrogen™ 11791-043) (Appendix C). Constructs were transformed into *E. coli* and mycobacteria as described previously and selected by kanamycin resistance.

5.5 SUB-CELLULAR LOCALIZATION STUDIES

Fluorescently-fused recombinant proteins were expressed and visualized by fluorescent microscopy to assess the sub-cellular localization and potential regions of activity of putative regulators to further assess involvement in the cell division process. For this, recombinant protein expression was induced from pMCSU7 constructs with 50 ng/µl Anhydrotetracycline (Clontech 631310) in mid-log *M.*

smegmatis strains for six hours, and fixed in 4% paraformaldehyde. Cells were resuspended in 5 μg/ml FM 4-64 Fx (Invitrogen™) in HBSS, incubated on ice for two minutes, and washed three times with HBSS. Dyed cells were air dry fixed to glass slides, coated with Vectashield Hard Set containing DAPI (Vector Laboratories) and covered with glass coverslips. Prepared slides were stored at 4^{o}C and imaged within 24 hours.

Fluorescent imaging was performed using an Olympus Ix7 microscope with a Retiga 2000R camera (Qimaging) and Slidebook software (Intelligent Imaging Innovations Inc.). All images were obtained at 100x magnification using FITC filters for GFP visualization, TRITC filter for FM 4-64 Fx visualization, and DAPI filters for nucleic acid visualization. Images were optimized using Adobe Photoshop CS4 software.

5.6 SOS RESPONSE INDUCTION

The SOS response was induced followed by transcriptional quantification of our putative regulators to explore regulatory similarities between Rv2216 and the SOS-responsive FtsZ inhibitor SulA identified in other bacteria. Mitomycin C (MMC) treatment was used at a final concentration of 0.2 μg/ml to induce the SOS response in *Mtb* [15]. Early-log phase *Mtb* H37Rv, $O.D._{600nm}$ 0.1-0.2, was grown in the presence of MMC over a concentration range of 200 μg/ml to 0.2 μg/ml by 2-fold serial dilutions. IC_{50}, defined as the concentration of drug required to reduce bacterial replication rates by 50% over a 7 day period, was performed in triplicate. Viability testing was performed by direct plating and colony forming unit (cfu) enumeration at 1, 3 and 5 days post-treatment with 0.2 μg/ml and 0.1 μg/ml MMC.

5.7 SCANNING ELECTRON MICROSCOPY

Overexpression strains were observed by scanning electron microscopy (SEM) to reveal morphological alterations in response to increased quantities of our putative regulators as filamentation has been associated with a disruption of cell division [16]. Recombinant *M. smegmatis* and *Mtb* containing pVV16, *rv2216*::pVV16 or *rv1708*::pVV16 and grown to mid-log phase were collected, washed 3 times in PBS at pH 7.4, then fixed in 2.5% gluteraldehyde in Buffer A (0.1 M potassium phosphate at pH 7.4, 1 mM $CaCl_2$ and 1 mM $MgCl_2$) at 4^{o}C for 24 hours. Fixed cells were washed three times in Buffer A, treated with 1% OsO_4 in Buffer A for 30 minutes at 4^{o}C, and washed three final times in Buffer A before dehydration by 20-100% graded ethanol treatments. Images were obtained using a JOEL JEM - 100CX electron microscope.

5.8 TOTAL RNA ISOLATION

Total cellular RNA was isolated from mycobacterial cultures for utilization in transcriptional analysis. Recombinant and wild type *Mtb* was grown to mid-log phase under various experimental conditions and harvested by centrifugation before resuspension in TRIzol® Reagent (Invitrogen®). Cell lysis was performed by bead beating with 0.1 mm Zirconia silicon beads three times in one minute increments, resting on ice between. Chloroform was added to the sample followed by centrifugation to separate total RNA, as outlines in the TRIzol® protocol. For RNA cleanup the aqueous layer was extracted and combined with an equal volume of 70% ethanol before addition to an RNeasy Mini kit (Qiagen) column.

5.9 QUANTITATIVE REAL-TIME PCR ANALYSIS

Quantitative real-time PCR was used in conjunction with reverse transcriptase reactions (qRT-PCR) to quantify the expression of cell division genes upon putative regulator overexpression and to quantify putative regulators, cell division genes, and SOS response genes upon MMC treatment. cDNA was generated from 5 ug of total RNA using the SuperScript™ III First-Strand Synthesis System (Invitrogen™). For qPCR, primers were designed using Primer3 software (Appendix C) and reactions performed using SYBR® Green qPCR SuperMix UDG (Invitrogen™) [17]. Reactions were performed using the iCycler iQ real-time PCR detection system and software (Bio-Rad Laboratories) with the following program: 55°C for 5 minutes, 95°C for 2 minutes, 45 cycles of 95°C for 15 seconds, 60°C for 30 seconds, and 72°C for 45 seconds, followed by melt-curve analysis beginning at 30^{o}C and increasing by 1^{o}C every 10 seconds for 65 cycles. Reactions were performed in triplicate on two independent RNA preparations. Primer efficiencies were determined using 1 ng, 10 ng and 100 ng of genomic DNA as template to generate standard curves. Linear regressions generated from these standard curves were used to determine the number of specific targets per sample using the ΔΔCt method.

5.10 MICROARRAY ANALYSIS

Global transcriptional profiling of *Mtb* strains overexpressing our putative regulators was performed by microarray analysis to assess the molecular programs affected by, and associated with, our putative regulator's activity within the cell. Microarrays were obtained through the Colorado State University TB Research Materials and Vaccine Testing Contract and post-processes using succinic anhydride. Fluorescently labeled cDNA was generated from 5 ug of total RNA in the presence of Cy3- or Cy-5 tagged dUTPs and hybridized to microarray slides at 42^{o}C for 2 hours [18]. Microarray slides were read using an Axon Genepix scanner and raw fluorescent intensities were recorded for data analysis.

Data sets for three biological replicates were combined for data reduction, normalization and analysis. Fluorescent intensities for each Cy-3 and Cy-5 channel were normalized to the mean channel intensity, analyzed by ANOVA single factor analysis, and used to generate natural log-based expression ratios for each gene. Significance cut-offs were set at >1.5-fold change in expression, with an associated p-value of $\leq$0.05.

REFERENCES

1. Koonin, E.V., *An apology for orthologs - or brave new memes.* Genome Biol, 2001. **2**(4): p. COMMENT1005.
2. Jensen, R.A., *Orthologs and paralogs - we need to get it right.* Genome Biol, 2001. **2**(8): p. INTERACTIONS1002.
3. Rost, B., *Twilight zone of protein sequence alignments.* Protein Eng, 1999. **12**(2): p. 85-94.
4. Zvelebil, M.J. and J.O. Baum, *Understanding bioinformatics*. 2008: Garland Science/Taylor & Francis Group.
5. Pruitt, K.D., et al., *NCBI Reference Sequences (RefSeq): current status, new features and genome annotation policy.* Nucleic Acids Res, 2012. **40**(Database issue): p. D130-5.
6. Waterhouse, A.M., et al., *Jalview Version 2--a multiple sequence alignment editor and analysis workbench.* Bioinformatics, 2009. **25**(9): p. 1189-91.
7. Altenhoff, A.M., et al., *OMA 2011: orthology inference among 1000 complete genomes.* Nucleic Acids Res, 2011. **39**(Database issue): p. D289-94.
8. Pertea, M., et al., *OperonDB: a comprehensive database of predicted operons in microbial genomes.* Nucleic Acids Res, 2009. **37**(Database issue): p. D479-82.
9. Reddy, T.B., et al., *TB database: an integrated platform for tuberculosis research.* Nucleic Acids Res, 2009. **37**(Database issue): p. D499-508.
10. Snyder, L. and W. Champness, *Molecular genetics of bacteria*. 2007: Amer Society for Microbiology.
11. Punta, M., et al., *The Pfam protein families database.* Nucleic Acids Res, 2012. **40**(Database issue): p. D290-301.
12. Kelley, L.A. and M.J. Sternberg, *Protein structure prediction on the Web: a case study using the Phyre server.* Nat Protoc, 2009. **4**(3): p. 363-71.
13. Sureka, K., et al., *Novel role of phosphorylation-dependent interaction between FtsZ and FipA in mycobacterial cell division.* PLoS One, 2010. **5**(1): p. e8590.
14. Dziadek, J., et al., *Physiological consequences associated with overproduction of Mycobacterium tuberculosis FtsZ in mycobacterial hosts.* Microbiology, 2002. **148**(Pt 4): p. 961-71.
15. Dullaghan, E.M., P.C. Brooks, and E.O. Davis, *The role of multiple SOS boxes upstream of the Mycobacterium tuberculosis lexA gene--identification of a novel DNA-damage-inducible gene.* Microbiology, 2002. **148**(Pt 11): p. 3609-15.
16. Justice, S.S., et al., *Morphological plasticity as a bacterial survival strategy.* Nat Rev Microbiol, 2008. **6**(2): p. 162-8.
17. http://frodo.wi.mit.edu/cgi-bin/primer3/primer3-www.cgi.
18. Brown, P. *Microarray Experimental Protocols*. 2001; Available from: http://cmgm.stanford.edu/pbrown/protocols/index.html.

APPENDIX I: CDR MICROARRAY DATA SET

Table I.1: Mean Log_2 expression of ORFs in *cdr* ::pVV16 *Mtb* from Microarray analysis.

ORF[b]	Name[b]	Functional Class[b]	Gene product	Avg. Log2 expression	Variance
Rv0024	Rv0024	virulence, detoxification, adaptation	putative p60 homologue	1.87	0.0098
Rv0064	Rv0064	cell wall and cell processes	possible membrane protein	2.12	0.0068
Rv0135c	Rv0135c	regulatory proteins	putative transcriptional regulator	2.46	0.0209
Rv0136	Rv0136	intermediary metabolism and respiration	cytochrome p450	3.20	0.0046
Rv0140	Rv0140	conserved hypotheticals	conserved hypothetical protein	2.65	0.0439
Rv0168	Rv0168	virulence, detoxification, adaptation	part of mce1 operon	-1.02	0.0093
Rv0181c	Rv0181c	conserved hypotheticals	conserved hypothetical protein	1.48	0.0161
Rv0190	Rv0190	conserved hypotheticals	conserved hypothetical protein	1.02	0.0140
Rv0237	lpqI	cell wall and cell processes	beta-hexosaminidase precursor	-1.11	0.0069
Rv0282	Rv0282	conserved hypotheticals	conserved hypothetical protein	-1.85	0.0066
Rv0283	Rv0283	conserved hypotheticals	conserved hypothetical protein	-2.07	0.0052
Rv0284	Rv0284	conserved hypotheticals	conserved hypothetical protein	-1.87	0.0006
Rv0285	PE	PE/PPE	PE-family protein	-2.09	0.0059
Rv0286	PPE	PE/PPE	PPE-family protein	-1.94	0.0063
Rv0287	Rv0287	regulatory proteins	conserved hypothetical protein	-2.45	0.0035
Rv0288	Rv0288	conserved hypotheticals	conserved hypothetical protein	-2.46	0.0032
Rv0290	Rv0290	cell wall and cell processes	unknown hydrophobic protein	-2.04	0.0131
Rv0342	Rv0342	conserved hypotheticals	conserved hypothetical protein	2.61	0.0302
Rv0381c	Rv0381c	unknown	hypothetical protein	2.76	0.0013
Rv0410c	pknG	regulatory proteins	serine-threonine protein kinase	1.67	0.0191
Rv0411c	glnH	cell wall and cell processes	putative glutamine binding protein	1.82	0.0470
Rv0412c	Rv0412c	cell wall and cell processes	unknown probable membrane protein	1.45	0.0477
Rv0464c	Rv0464c	conserved hypotheticals	conserved hypothetical protein	1.00	0.0364
Rv0514	Rv0514	cell wall and cell processes	possible membrane protein	-1.25	0.0128
Rv0563	htpX	virulence, detoxification, adaptation	probable (transmembrane) heat shock protein	2.41	0.0451
Rv0570	nrdZ	intermediary metabolism and respiration	ribonucleotide reductase, class II	2.40	0.0050
Rv0588	Rv0588	virulence, detoxification, adaptation	part of mce2 operon	2.56	0.0095
Rv0589	mce2	virulence, detoxification, adaptation	cell invasion protein	2.16	0.0025
Rv0624	Rv0624	conserved hypotheticals	conserved hypothetical protein	1.08	0.0000
Rv0633c	Rv0633c	unknown	hypothetical protein	0.96	0.0149
Rv0652	rplL	information pathways	50S ribosomal protein L7/L12	-1.64	0.0099
Rv0660c	Rv0660c	conserved hypotheticals	conserved hypothetical protein	1.02	0.0128
Rv0662c	Rv0662c	unknown	hypothetical protein	-1.03	0.0319
Rv0671	lpqP	cell wall and cell processes	probable esterase	1.76	0.0003
Rv0678	Rv0678	unknown	hypothetical protein	1.95	0.0185
Rv0694	lldD1	intermediary metabolism and respiration	L-lactate dehydrogenase (cytochrome)	-1.24	0.0150
Rv0709	rpmC	information pathways	50S ribosomal protein L29	-1.10	0.0262
Rv0724	sppA	intermediary metabolism and respiration	protease IV, signal peptide peptidase	0.98	0.0328
Rv0743c	Rv0743c	unknown	hypothetical protein	1.33	0.0202
Rv0745	Rv0745	conserved hypotheticals	conserved hypothetical protein	-1.18	0.0071
Rv0749	Rv0749	conserved hypotheticals	conserved hypothetical protein	1.07	0.0027
Rv0765c	Rv0765c	intermediary metabolism and respiration	short-chain alcohol dehydrogenase family	3.71	0.0003
Rv0769	Rv0769	intermediary metabolism and respiration	similar to 7-alpha-hydroxysteroid dehydrogenase	1.49	0.0416
Rv0786c	Rv0786c	conserved hypotheticals	conserved hypothetical protein	1.15	0.0283
Rv0791c	Rv0791c	intermediary metabolism and respiration	possible monooxygenasemonoxygenase	3.63	0.0151
Rv0814c	sseC2	intermediary metabolism and respiration	thiosulfate sulfurtransferase	-1.07	0.0521
Rv0839	Rv0839	conserved hypotheticals	conserved hypothetical protein	1.64	0.0076
Rv0851c	Rv0851c	intermediary metabolism and respiration	Short-chain dehydrogenases/reductase	3.41	0.0001
Rv0855	far	lipid metabolism	fatty acyl-CoA racemase	1.24	0.0105
Rv0921	Rv0921	insertion seqs and phages	resolvase	1.06	0.0280
Rv0926c	Rv0926c	conserved hypotheticals	conserved hypothetical protein	1.69	0.0002
Rv0936	pstA2	cell wall and cell processes	PstA component of phosphate uptake	2.51	0.0275
Rv1034c	Rv1034c	insertion seqs and phages	hypothetical protein	1.25	0.0034
Rv1035c	Rv1035c	insertion seqs and phages	hypothetical protein	0.97	0.0216
Rv1037c	Rv1037c	conserved hypotheticals	conserved hypothetical protein	-1.89	0.0417
Rv1040c	PE	PE/PPE	PE-family protein	-1.13	0.0303
Rv1043c	Rv1043c	unknown	hypothetical protein	1.86	0.0219
Rv1063c	Rv1063c	conserved hypotheticals	conserved hypothetical protein	1.88	0.0099
Rv1078	pra	conserved hypotheticals	conserved hypothetical protein	-1.71	0.0402
Rv1082	Rv1082	intermediary metabolism and respiration	similar to S. lincolnensis lmbE	1.07	0.0265
Rv1102c	Rv1102c	conserved hypotheticals	conserved hypothetical protein	1.02	0.0004
Rv1128c	Rv1128c	insertion seqs and phages	REP-family protein	1.20	0.0073
Rv1129c	Rv1129c	regulatory proteins	transcriptional regulator (PbsX/Xre family)	1.55	0.0020
Rv1134	Rv1134	unknown	hypothetical protein	1.54	0.0081
Rv1152	Rv1152	regulatory proteins	transcriptional regulator (GntR family)	1.02	0.0001
Rv1198	Rv1198	conserved hypotheticals	conserved hypothetical protein	-1.45	0.0306
Rv1209	Rv1209	conserved hypotheticals	conserved hypothetical protein	-1.05	0.0386
Rv1224	Rv1224	conserved hypotheticals	conserved hypothetical protein	1.08	0.0097
Rv1235	lpqY	cell wall and cell processes	possible role in sugar transport	1.45	0.0420
Rv1286	cysN	intermediary metabolism and respiration	ATP:sulphurylase subunit 1	1.85	0.0034
Rv1302	rfe	cell wall and cell processes	hypothetical protein	-1.19	0.0049
Rv1308	atpA	intermediary metabolism and respiration	ATP synthase [alpha] chain	-1.08	0.0040

Rv1312	Rv1312	conserved hypotheticals	conserved hypothetical protein	-1.16	0.0161
Rv1329c	dinG	information pathways	probable ATP-dependent helicase	0.98	0.0192
Rv1334	Rv1334	conserved hypotheticals	conserved hypothetical protein	1.10	0.0005
Rv1343c	Rv1343c	conserved hypotheticals	conserved hypothetical protein	-1.11	0.0011
Rv1346	fadE14	lipid metabolism	acyl-CoA dehydrogenase	-1.07	0.0354
Rv1367c	Rv1367c	cell wall and cell processes	probable penicillin binding protein	3.56	0.0173
Rv1386	PE	PE/PPE	PE-family protein	-1.45	0.0528
Rv1396c	PE_PGRS	PE/PPE	PE_PGRS-family protein	1.14	0.0529
Rv1435c	Rv1435c	conserved hypotheticals	conserved hypothetical protein	-1.22	0.0020
Rv1436	gap	intermediary metabolism and respiration	glyceraldehyde 3-phosphate dehydrogenase	-1.25	0.0323
Rv1471	trxB	intermediary metabolism and respiration	thioredoxin reductase	2.07	0.0439
Rv1477	Rv1477	virulence, detoxification, adaptation	putative exported p60 protein homologue	-1.36	0.0471
Rv1536	ileS	information pathways	isoleucyl-tRNA synthase	1.09	0.0115
Rv1551	plsB1	lipid metabolism	glycerol-3-phosphate acyltransferase	3.45	0.0094
Rv1579c	Rv1579c	insertion seqs and phages	phiRV1 phage related protein	1.25	0.0380
Rv1580c	Rv1580c	insertion seqs and phages	phiRV1 phage related protein	1.64	0.0044
Rv1581c	Rv1581c	insertion seqs and phages	phiRV1 phage related protein	2.27	0.0421
Rv1584c	Rv1584c	insertion seqs and phages	phiRV1 phage related protein	2.90	0.0015
Rv1585c	Rv1585c	insertion seqs and phages	phiRV1 phage related protein	2.26	0.0009
Rv1593c	Rv1593c	conserved hypotheticals	conserved hypothetical protein	1.03	0.0226
Rv1599	hisD	intermediary metabolism and respiration	histidinol dehydrogenase	1.63	0.0149
Rv1600	hisC	intermediary metabolism and respiration	histidinol-phosphate aminotransferase	0.96	0.0417
Rv1601	hisB	intermediary metabolism and respiration	imidazole glycerol-phosphate dehydratase	1.37	0.0272
Rv1603	hisA	intermediary metabolism and respiration	phosphoribosylformimino-5-aminoimidazole	1.02	0.0389
Rv1613	trpA	intermediary metabolism and respiration	tryptophan synthase [alpha] chain	-1.48	0.0073
Rv1620c	cydC	intermediary metabolism and respiration	ABC transporter	1.11	0.0000
Rv1621c	cydD	intermediary metabolism and respiration	ABC transporter	1.39	0.0547
Rv1627c	Rv1627c	lipid metabolism	lipid carrier protein	1.14	0.0009
Rv1653	argJ	intermediary metabolism and respiration	glutamate N-acetyltransferase	1.63	0.0036
Rv1704c	cycA	cell wall and cell processes	transport of D-alanine, D-serine and glycine	-1.24	0.0025
Rv1712	cmk	intermediary metabolism and respiration	cytidylate kinase	-1.45	0.0499
Rv1724c	Rv1724c	unknown	hypothetical protein	1.32	0.0017
Rv1733c	Rv1733c	cell wall and cell processes	possible membrane protein	1.03	0.0004
Rv1786	Rv1786	intermediary metabolism and respiration	Probable ferredoxin	1.45	0.0027
Rv1792	Rv1792	conserved hypotheticals	conserved hypothetical protein	-1.36	0.0031
Rv1793	Rv1793	conserved hypotheticals	conserved hypothetical protein	-1.29	0.0087
Rv1809	PPE	PE/PPE	PPE-family protein	1.40	0.0070
Rv1813c	Rv1813c	conserved hypotheticals	conserved hypothetical protein	1.25	0.0002
Rv1854c	ndh	intermediary metabolism and respiration	probable NADH dehydrogenase	1.59	0.0145
Rv1874	Rv1874	unknown	hypothetical protein	1.24	0.0085
Rv1875	Rv1875	conserved hypotheticals	conserved hypothetical protein	1.14	0.0470
Rv1886c	fbpB	lipid metabolism	antigen 85B, mycolyltransferase	-1.08	0.0217
Rv1887	Rv1887	unknown	hypothetical protein	-1.24	0.0298
Rv1896c	Rv1896c	conserved hypotheticals	conserved hypothetical protein	1.82	0.0392
Rv1918c	PPE	PE/PPE	PPE-family protein	1.36	0.0042
Rv1919c	Rv1919c	cell wall and cell processes	weak similarity to pollen antigens	-1.11	0.0347
Rv1932	tpx	virulence, detoxification, adaptation	thiol peroxidase	-1.30	0.0330
Rv1957	Rv1957	unknown	hypothetical protein	1.23	0.0022
Rv1982c	Rv1982c	conserved hypotheticals	conserved hypothetical protein	1.24	0.0211
Rv1997	ctpF	cell wall and cell processes	probable cation transport ATPase	3.03	0.0004
Rv2016	Rv2016	unknown	hypothetical protein	2.76	0.0334
Rv2028c	Rv2028c	conserved hypotheticals	conserved hypothetical protein	1.19	0.0022
Rv2030c	Rv2030c	conserved hypotheticals	conserved hypothetical protein	1.46	0.0271
Rv2031c	hspX	virulence, detoxification, adaptation	14kD antigen, heat shock protein Hsp20 family	1.56	0.0131
Rv2032	Rv2032	conserved hypotheticals	conserved hypothetical protein	3.01	0.0004
Rv2033c	Rv2033c	unknown	hypothetical protein	1.73	0.0217
Rv2052c	Rv2052c	unknown	hypothetical protein	1.46	0.0123
Rv2127	ansP	cell wall and cell processes	L-asparagine permease	-1.29	0.0508
Rv2135c	Rv2135c	conserved hypotheticals	conserved hypothetical protein	1.23	0.0022
Rv2147c	Rv2147c	unknown	hypothetical protein	1.12	0.0312
Rv2153c	murG	cell wall and cell processes	transferase in peptidoglycan synthesis	1.27	0.0001
Rv2188c	Rv2188c	conserved hypotheticals	conserved hypothetical protein	-1.06	0.0301
Rv2190c	Rv2190c	virulence, detoxification, adaptation	putative p60 homologue	-1.83	0.0519
Rv2212	Rv2212	conserved hypotheticals	conserved hypothetical protein	0.97	0.0521
Rv2216	Rv2216	conserved hypotheticals	conserved hypothetical protein	10.59	0.0418
Rv2218	lipA	intermediary metabolism and respiration	lipoate biosynthesis protein A	1.11	0.0016
Rv2230c	Rv2230c	conserved hypotheticals	conserved hypothetical protein	2.65	0.0054
Rv2235	Rv2235	unknown	hypothetical protein	1.24	0.0350
Rv2242	Rv2242	unknown	hypothetical protein	1.59	0.0533
Rv2244	acpM	lipid metabolism	acyl carrier protein (meromycolate extension)	-1.25	0.0269
Rv2246	kasB	lipid metabolism	[beta]-ketoacyl-ACP synthase (meromycolate	-1.01	0.0528
Rv2254c	Rv2254c	unknown	hypothetical protein	1.51	0.0135
Rv2304c	Rv2304c	unknown	hypothetical protein	2.34	0.0005
Rv2346c	Rv2346c	conserved hypotheticals	conserved hypothetical protein	-1.20	0.0066
Rv2364c	bex	intermediary metabolism and respiration	GTP-binding protein of Era/ThdF family	1.15	0.0316
Rv2370c	Rv2370c	conserved hypotheticals	conserved hypothetical protein	4.56	0.0009
Rv2371	PE	PE/PPE	PE-family protein	-1.12	0.0040
Rv2372c	Rv2372c	conserved hypotheticals	conserved hypothetical protein	1.35	0.0379
Rv2386c	trpE2	intermediary metabolism and respiration	anthranilate synthase component I	-2.74	0.0130
Rv2399c	cysT	cell wall and cell processes	sulphate transport system permease protein	0.99	0.0000

Rv2412	rpsT	information pathways	30S ribosomal protein S20	-1.03	0.0297
Rv2450c	Rv2450c	conserved hypotheticals	conserved hypothetical protein	-2.07	0.0476
Rv2459	Rv2459	cell wall and cell processes	probable drug efflux protein	-1.27	0.0043
Rv2465c	rpi	intermediary metabolism and respiration	phosphopentose isomerase	1.23	0.0345
Rv2466c	Rv2466c	conserved hypotheticals	conserved hypothetical protein	1.91	0.0108
Rv2483c	Rv2483c	intermediary metabolism and respiration	possible transferase	1.17	0.0091
Rv2495c	pdhC	intermediary metabolism and respiration	dihydrolipoamide acetyltransferase	1.11	0.0068
Rv2497c	pdhA	intermediary metabolism and respiration	pyruvate dehydrogenase E1 [alpha] subunit	1.32	0.0008
Rv2499c	Rv2499c	intermediary metabolism and respiration	putative aldehyde dehydrogenase	1.01	0.0363
Rv2501c	accA1	lipid metabolism	acetyl/propionyl-CoA carboxylase [alpha] subunit	1.26	0.0417
Rv2524c	fas	lipid metabolism	fatty acid synthase	1.18	0.0070
Rv2528c	mrr	information pathways	restriction system protein	-1.07	0.0033
Rv2532c	Rv2532c	unknown	hypothetical protein	1.32	0.0180
Rv2550c	Rv2550c	unknown	hypothetical protein	1.22	0.0031
Rv2612c	pgsA	lipid metabolism	CDP-diacylglycerol-glycerol-3-phosphate	-1.14	0.0060
Rv2624c	Rv2624c	conserved hypotheticals	conserved hypothetical protein	2.85	0.0002
Rv2626c	Rv2626c	conserved hypotheticals	conserved hypothetical protein	2.55	0.0327
Rv2629	Rv2629	unknown	hypothetical protein	2.20	0.0079
Rv2630	Rv2630	unknown	hypothetical protein	2.30	0.0522
Rv2659c	Rv2659c	insertion seqs and phages	phiRV2 integrase	2.75	0.0252
Rv2662	Rv2662	unknown	hypothetical protein	2.86	0.0002
Rv2675c	Rv2675c	intermediary metabolism and respiration	putative methyltransferase	1.74	0.0352
Rv2685	arsB	cell wall and cell processes	probable arsenical pump	-1.42	0.0042
Rv2721c	Rv2721c	conserved hypotheticals	conserved hypothetical protein	-1.06	0.0016
Rv2735c	Rv2735c	unknown	hypothetical protein	1.44	0.0341
Rv2775	Rv2775	unknown	hypothetical protein	1.83	0.0378
Rv2784c	lppU	cell wall and cell processes	lipoprotein	1.28	0.0466
Rv2825c	Rv2825c	conserved hypotheticals	conserved hypothetical protein	1.03	0.0396
Rv2831	echA16	lipid metabolism	enoyl-CoA hydratase/isomerase superfamily	-1.22	0.0406
Rv2860c	glnA4	intermediary metabolism and respiration	proable glutamine synthase	1.49	0.0063
Rv2870c	Rv2870c	conserved hypotheticals	conserved hypothetical protein	1.13	0.0001
Rv2877c	Rv2877c	cell wall and cell processes	possible mercury resistance transport system	1.43	0.0063
Rv2892c	PPE	PE/PPE	PPE-family protein	4.54	0.0511
Rv2901c	Rv2901c	unknown	hypothetical protein	-1.06	0.0017
Rv2913c	Rv2913c	intermediary metabolism and respiration	probable D-amino acid aminohydrolase	1.81	0.0000
Rv2918c	glnD	intermediary metabolism and respiration	uridylyltransferase	1.57	0.0446
Rv2920c	amt	cell wall and cell processes	putative ammonium transporter	1.06	0.0117
Rv2940c	mas	lipid metabolism	mycocerosic acid synthase	-1.59	0.0091
Rv2947c	pks15	lipid metabolism	polyketide synthase	-2.37	0.0082
Rv2948c	fadD22	lipid metabolism	acyl-CoA synthase	-1.36	0.0100
Rv2951c	Rv2951c	intermediary metabolism and respiration	putative oxidoreductase	-1.74	0.0095
Rv2952	Rv2952	intermediary metabolism and respiration	glycosyltransferase	-1.21	0.0013
Rv2953	Rv2953	conserved hypotheticals	conserved hypothetical protein	-1.23	0.0151
Rv2969c	Rv2969c	cell wall and cell processes	possible transmembrane domain	-1.04	0.0102
Rv2973c	recG	information pathways	ATP-dependent DNA helicase	1.59	0.0541
Rv2986c	hupB	information pathways	DNA-binding protein II	-1.52	0.0032
Rv2989	Rv2989	regulatory proteins	transcriptional regulator (IclR family)	1.13	0.0315
Rv3019c	Rv3019c	cell wall and cell processes	similar to Esat6	-2.74	0.0012
Rv3029c	fixA	intermediary metabolism and respiration	electron transfer flavoprotein [beta] subunit	-1.11	0.0092
Rv3043c	ctaD	intermediary metabolism and respiration	cytochrome c oxidase polypeptide I	-1.04	0.0376
Rv3047c	Rv3047c	unknown	hypothetical protein	1.01	0.0006
Rv3060c	Rv3060c	regulatory proteins	transcriptional regulator (GntR family)	2.75	0.0081
Rv3077	atsF	intermediary metabolism and respiration	proable arylsulfatase	-1.56	0.0501
Rv3088	Rv3088	conserved hypotheticals	conserved hypothetical protein	1.12	0.0451
Rv3093c	Rv3093c	unknown	hypothetical protein	-1.95	0.0508
Rv3097c	PE	PE/PPE	PE-family protein	1.92	0.0214
Rv3116	moeB	intermediary metabolism and respiration	molybdopterin biosynthesis	-1.33	0.0118
Rv3127	Rv3127	conserved hypotheticals	conserved hypothetical protein	2.70	0.0009
Rv3130c	Rv3130c	conserved hypotheticals	conserved hypothetical protein	1.83	0.0031
Rv3131	Rv3131	conserved hypotheticals	conserved hypothetical protein	2.88	0.0111
Rv3133c	Rv3133c	regulatory proteins	two-component response regulator	1.78	0.0040
Rv3146	nuoB	intermediary metabolism and respiration	NADH dehydrogenase chain B	-1.32	0.0373
Rv3148	nuoD	intermediary metabolism and respiration	NADH dehydrogenase chain D	-1.07	0.0001
Rv3150	nuoF	intermediary metabolism and respiration	NADH dehydrogenase chain F	-1.30	0.0004
Rv3155	nuoK	intermediary metabolism and respiration	NADH dehydrogenase chain K	-2.09	0.0291
Rv3156	nuoL	intermediary metabolism and respiration	NADH dehydrogenase chain L	-2.06	0.0043
Rv3159c	PPE	PE/PPE	PPE-family protein	1.51	0.0181
Rv3169	Rv3169	conserved hypotheticals	conserved hypothetical protein	1.38	0.0147
Rv3189	Rv3189	conserved hypotheticals	conserved hypothetical protein	3.66	0.0001
Rv3219	whiB1	regulatory proteins	WhiB transcriptional activator homologue	-1.10	0.0117
Rv3229c	desA3	lipid metabolism	acyl-[ACP] desaturase	3.03	0.0069
Rv3234c	Rv3234c	conserved hypotheticals	conserved hypothetical protein	-1.03	0.0148
Rv3244c	lpqB	cell wall and cell processes	lipoprotein	2.09	0.0306
Rv3245c	mtrB	regulatory proteins	sensor histidine kinase	1.10	0.0080
Rv3270	ctpC	cell wall and cell processes	cation transport ATPase	1.37	0.0059
Rv3287c	rsbW	information pathways	anti-sigma B factor	2.62	0.0025
Rv3288c	Rv3288c	conserved hypotheticals	conserved hypothetical protein	2.07	0.0001
Rv3289c	Rv3289c	unknown	hypothetical protein	1.13	0.0071
Rv3290c	lat	intermediary metabolism and respiration	lysine-[epsilon] aminotransferase	1.42	0.0159
Rv3307	deoD	intermediary metabolism and respiration	probable purine nucleoside phosphorylase	-1.93	0.0000

Rv3319	sdhB	intermediary metabolism and respiration	succinate dehydrogenase B	-1.18	0.0098
Rv3327	Rv3327	insertion seqs and phages	hypothetical protein	1.78	0.0411
Rv3331	sugI	cell wall and cell processes	probable sugar transport protein	-1.38	0.0087
Rv3334	Rv3334	regulatory proteins	transcriptional regulator (MerR family)	1.57	0.0125
Rv3346c	Rv3346c	conserved hypotheticals	conserved hypothetical protein	-2.06	0.0015
Rv3347c	PPE	PE/PPE	PPE-family protein	1.29	0.0414
Rv3354	Rv3354	conserved hypotheticals	conserved hypothetical protein	1.85	0.0435
Rv3403c	Rv3403c	unknown	hypothetical protein	-1.30	0.0029
Rv3425	PPE	PE/PPE	PPE-family protein	-1.84	0.0142
Rv3445c	Rv3445c	unknown	hypothetical protein	1.66	0.0092
Rv3463	Rv3463	intermediary metabolism and respiration	probable neuraminidase	1.49	0.0000
Rv3478	PPE	PE/PPE	PPE-family protein	-1.77	0.0029
Rv3483c	Rv3483c	conserved hypotheticals	conserved hypothetical protein	-1.09	0.0041
Rv3532	PPE	PE/PPE	PPE-family protein	1.00	0.0004
Rv3543c	fadE29	lipid metabolism	acyl-CoA dehydrogenase	1.48	0.0123
Rv3544c	fadE28	lipid metabolism	acyl-CoA dehydrogenase	1.27	0.0343
Rv3545c	Rv3545c	intermediary metabolism and respiration	cytochrome p450	1.03	0.0477
Rv3548c	Rv3548c	intermediary metabolism and respiration	short-chain alcohol dehydrogenase family	1.46	0.0065
Rv3550	echA20	lipid metabolism	enoyl-CoA hydratase/isomerase superfamily	1.91	0.0011
Rv3551	Rv3551	intermediary metabolism and respiration	possible glutaconate CoA-transferase	1.54	0.0390
Rv3552	Rv3552	unknown	hypothetical protein	1.38	0.0152
Rv3553	Rv3553	intermediary metabolism and respiration	similar to dioxygenasesdiooxygenases	1.03	0.0444
Rv3560c	fadE30	lipid metabolism	acyl-CoA dehydrogenase	1.01	0.0154
Rv3561	fadD3	lipid metabolism	acyl-CoA synthase	1.50	0.0186
Rv3577	Rv3577	unknown	hypothetical protein	5.13	0.0121
Rv3597c	lsr2	conserved hypotheticals	conserved hypothetical protein	-1.25	0.0007
Rv3619c	Rv3619c	conserved hypotheticals	conserved hypothetical protein	-1.40	0.0021
Rv3622c	PE	PE/PPE	PE-family protein	-1.20	0.0015
Rv3648c	cspA	virulence, detoxification, adaptation	cold shock protein, transcriptional regulator	-1.43	0.0205
Rv3685c	Rv3685c	intermediary metabolism and respiration	Probable cytochrome P-450	1.07	0.0185
Rv3703c	Rv3703c	conserved hypotheticals	conserved hypothetical protein	-1.10	0.0008
Rv3732	Rv3732	conserved hypotheticals	conserved hypothetical protein	2.19	0.0129
Rv3752c	Rv3752c	intermediary metabolism and respiration	probable cytidine/deoxycytidylate deaminase	0.99	0.0471
Rv3763	Rv3764	cell wall and cell processes	19 kDKD	-1.27	0.0002
Rv3778c	Rv3778c	conserved hypotheticals	conserved hypothetical protein	-1.13	0.0001
Rv3808c	Rv3808c	unknown	hypothetical protein	-1.10	0.0187
Rv3841	bfrB	intermediary metabolism and respiration	bacterioferritin	2.83	0.0013
Rv3854c	Rv3854c	intermediary metabolism and respiration	probable monooxygenase	2.11	0.0111
Rv3855	Rv3855	regulatory proteins	putative transcriptional regulator	2.07	0.0242
Rv3864	Rv3864	conserved hypotheticals	conserved hypothetical protein	-1.16	0.0388
Rv3869	Rv3869	conserved hypotheticals	conserved hypothetical protein	1.06	0.0245
Rv3883c	Rv3883c	intermediary metabolism and respiration	probable secreted protease	-1.14	0.0014
Rv3884c	Rv3884c	conserved hypotheticals	conserved hypothetical protein	-1.17	0.0448
Rv3922c	Rv3922c	virulence, detoxification, adaptation	possible hemolysin	-1.37	0.0122
Rv3924c	rpmH	information pathways	50S ribosomal protein L34	-1.11	0.0494

[a] *cdr* recombinant *Mtb* strain compared to control H37Rv strain at 37°C.

[b] ORFs, gene names and functional class are as annotated at http://genolist.pasteur.fr/TubercuList/

APPENDIX II: PRIMERS

Table II.1: Primer sequences used for plasmid construction.

Gene	Construct		Primer Sequence
soj_{Mtb}	pVV16	Forward	5' - CTGGTACATATGTTGCCTGCGGGTCTCCC
		Reverse	5' - AAGCTCAAGCTTCATGCCAAATCGGTCGATC
soj_{Mtb}	pMCSU7	Forward	5' - GGGGCAACTTTGTACAAAAAAGTTGCCCATATGCCTGCGGGTCTCCCGG
		Reverse	5' - GGGGCAACTTTGTACAAGAAAGTTGCAAGCTTCATGCCAAATCGGTCGATC
cdr	pVV16	Forward	5' - CTGGTACATATGATGGCCAACGCCGTTGTC
		Reverse	5' - AAGCTCAAGCTTCTAGCCGGGCCGGGTG
cdr	pMCSU7	Forward	5' - GGGGCAACTTTGTACAAAAAAGTTGCCCATATGGCCAACGCCGTTGTCGC
		Reverse	5' - GGGGCAACTTTGTACAAGAAAGTTGCAAGCTTGCCGGGCCGGGTGGTG

Table II.2: Primer sequences used for qRT-PCR.

Gene	Rv #		Primer Sequence
ruvA	*rv2593c*	Forward	5' - GGTGTGGGCTACCGAGTGAA
		Reverse	5' - AGTCCTCGCGCACAATCATC
ruvC	*rv2594c*	Forward	5' - CGACGTGCATTTCCATACCC
		Reverse	5' - CTTGCAGCGCAAGGATTTTG
chiZ	*rv2719c*	Forward	5' - ACGGGGGAGTCCCTGTATGA
		Reverse	5' - GTGTCTGCAGGCCGTTGAGT
lexA	*rv2720*	Forward	5' - AAGGGCTACCTACGCCGTGA
		Reverse	5' - CAGGGACAAAGGTGGGTTCC
recX	*rv2736c*	Forward	5' - ACTTCTGAGCGCGAAGAGCA
		Reverse	5' - AATACCCGGTTGCCGATGTC
recA	*rv2737c*	Forward	5' - AACTCCGCAAGGGAGACAGG
		Reverse	5' - TCCGATCAGGTAGCCAAGCA
dnaA	*rv0001*	Forward	5' - GCATTCAAACGCAGCTACCG
		Reverse	5' - TGTTGGCATTGTGCAAGGTG
parA	*rv3918c*	Forward	5' - TGATGATCCCGATCCAATGC
		Reverse	5' - CCGGCCGTCATACATGGTAA
ftsZ	*rv2150c*	Forward	5' - CCTCATCGTGATTCCCAACG
		Reverse	5' - TTGATTAGACCCGGGGTGGT
ftsI	*rv2163c*	Forward	5' - GAAACGCGGTCATCTTGGTG
		Reverse	5' - CGTCGGTGACCTTGAGTTGG
pbpA	*rv0016c*	Forward	5' - CTCCGTCCTACGACCCCAAC
		Reverse	5' - GAGATGGCACGGTTGGTCAG
cdr	*rv2216*	Forward	5' - TACCCCGCTGATGTTGCCTA
		Reverse	5' - CTCGCCAATGGTGTTGTGGT
kasA	*rv2245*	Forward	5' - TGCTCATCGAGACGGAGGAG
		Reverse	5' - CTACCGGCACGAACACCATC
murD	*rv2155c*	Forward	5' - CGATGTCCCAGTCGTTCAGG

		Reverse	5’ - AACGGTCTCACCGGCTTTGT
gid	*rv3919c*	Forward	5’ - CCGCGAAGTCGGTAGGCTAT
		Reverse	5’ - GGCTCCGCTACCGATATCCA
dnaE1	*rv1547*	Forward	5’ - AAAGACAAGGCCGCGGTTTA
		Reverse	5’ - CGACCGATGCGAAGTTCAAG
sigA	*Rv2703*	Forward	5’ - TTCGCGCCTACCTCAAACAG
		Reverse	5’ - GCTAGCTCGACCTCTTCCTCG

APPENDIX III: CONSTRUCT MAPS

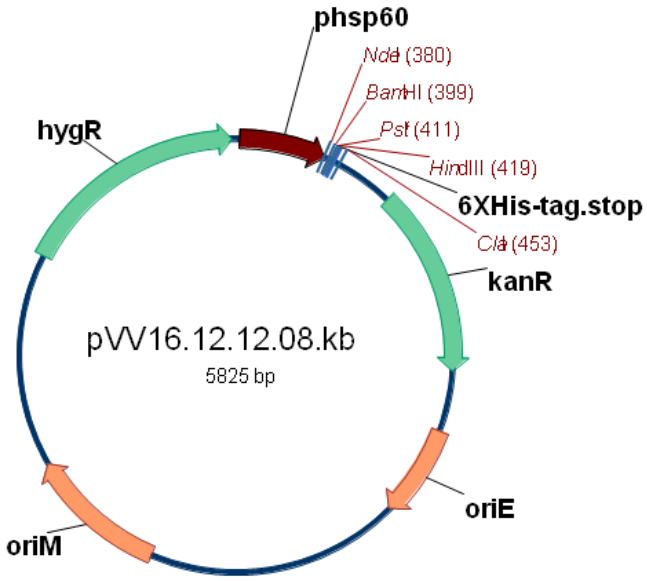

Figure III.1: Construct map of pVV16.

Map of extrachromosomal pVV16 construct used for overexpression studies in mycobacteria. This construct acts as a shuttle vector between *E. coli* and mycobacteria hosts through the dual possession of an *oriE E. coli* replicon and *oriM* mycobacterial replicon. NdeI and HinDIII of the multiple cloning site were used for insertion of ORFs. Phsp60 is the *Mycobacterium bovis* BCG hsp60 promoter which allows constitutive high levels of gene expression within a mycobacterial host. A polyhistidine tag is encoded to allow expression of a C-terminally his-tagged protein for protein purification and immunohistochemistry applications. KanR and HygR allow for selection using kanamycin or hygromycin B, respectively.

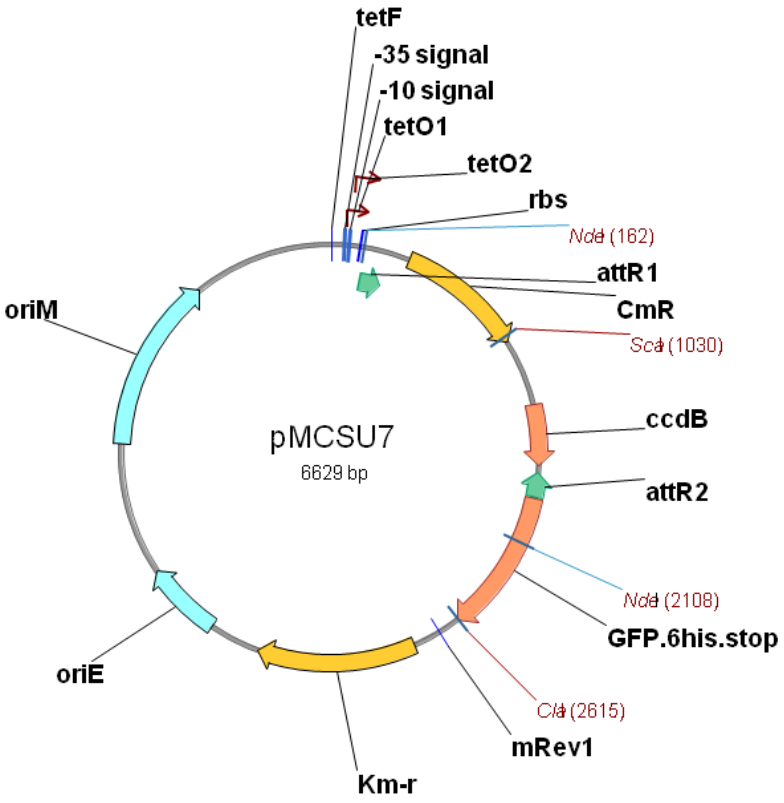

Figure III.2: Construct map of pMCSU7.

Map of extrachromosomal pMCSU7 construct used for sub-cellular localization studies. *oriE* and *oriM* allow utilization as an *E. coli* mycobacterial shuttle vector. This vector propagates in *E. coli ccdB* survival T1-R host under kanamycin+chloramphenicol selection allowing for negative and positive selection after recombination. *attR1* and *attR2* enable Gateway-compatible cloning of ORFs into the vector. KmR allows for kanamycin resistance selection for successful transformations into mycobacteria. *Streptomyces coelicolor tetO* promoter derived from plasmid tcp3 allows for an inducible expression system using anhydrotetracycline. Genes are expressed containing a C-terminal GFP tag followed by a poly histidine tag for protein purification and immunohistochemistry.

APPENDIX IV: CURRICULUM VITAE

Rebecca M. Crew
M.S.

4822 W 6th Street Rd
Greeley, CO 80634
(720)-352-7920
Rebecca.Crew@ColoState.edu

Education:

Masters in Science, Expected August 2013
Dept: Microbiology, Immunology and Pathology
Colorado State University
Fort Collins, Colorado

Bachelors in Science, Received April 2008
Option: Cellular Biology
Minor: Chemistry
Fort Lewis College
Durango, Colorado

Experience:

Graduate Research Assistant, Slayden Lab, Department of Microbiology, Immunology and Pathology at Colorado State University: August 2010 – August 2013
Research: Characterization of cell division regulators in *Mycobacterium tuberculosis* identified by bioinformatic analysis.
Skills: Recombinant DNA techniques, transformation of bacteria, RNA, DNA, and protein isolation, western blots, semi-quantitative and quantitative real-time PCR, site-directed mutagenesis, fluorescence microscopy, recombinant protein expression and purification.

Graduate Teaching Assistant, Department of Microbiology, Immunology and Pathology at Colorado State University: August 2012 – January 2013
Courses: Functional Genomics
Responsibilities: Aided graduate students in the selection, use and application of modern techniques in the field including basic bioinformatics, microarray, NextGen sequencing, mass spectrometry, fluorescent techniques, proteomics, qRT-PCR, and data analysis.

Undergraduate Research Assistant, Jackson Lab, Department of Microbiology, Immunology and Pathology at Colorado State University: August 2009 – August 2010
Research: Inhibition of epoxide hydrolases of *Mycobacterium tuberculosis* by urea derivatives.
Skills: Maintenance of bacterial cultures, aseptic techniques, recombinant DNA techniques, transformation of bacteria, DNA isolation, PCR, *in vitro* drug sensitivity assay.

Awards and Honors:

Graduate Research Assistantship from Department of Microbiology, Immunology and Pathology, 2010-13

Graduated with B.S. at Fort Lewis College Suma Cum Laude, 2008

Tri-Beta Honor Society, 2006-8

Outstanding Junior Biology Student, 2006
Dean's List, 2004-5
First Generation Student Program - Undergraduate Full Tuition Scholarship, 2004-8

John F Reed Honors Program, 2004-8

Publications:

Crew R, Slayden RA. Antibacterial persistence and adaptations to non-replicating persistence in Mycobacterium tuberculosis: a recent review. (*in preparation*)

Crew R, England K, Ramirez M, Slayden RA: A conserved cell division regulator in Mycobacterium tuberculosis, encoded by rv2216 regulates septation and metabolism. (*in preparation*)

Ramirez, MV, **Crew R**, England K, and Slayden RA. MazF6 toxin of Mycobacterium tuberculosis is coupled to regulation of cell growth by a Soj-like protein. (*in preparation*)

Slayden RA, Jackson M, Zucker J, Ramirez MV, Dawson CC, **Crew R**, Sampson NS, Thomas ST, Jamshidi N, Sisk P, Caspi R, Crick DC, McNeil MR, Pavelka MS, Niederweis M, Siroy A, Dona V, McFadden J, Boschoff H, Lew JM: **Updating and curating metabolic pathways of TB**. *Tuberculosis (Edinb)* 2013, **Jan 31** doi: 10.1016/j.tube.2012.11.001 [epub ahead of print].

Grzegorzewicz AE, Kordulakova J, Jones V, Born SE, Belardinelli JM, Vasquie A, Gundi VA, Madacki J, Slama N, Laval F, Vaubourgeix J, **Crew R**, Gicqel B, Daffe M, Morbidoni HR, Brennan PJ, Quemard A, McNeil MR, Jackson M: **A common mechanism of inhibition of the Mycobacterium tuberculosis mycolic acid biosynthetic pathway by isoxyl and thiacetazone**. *J Biol Chem* 2012, **287**(46): 38434-38441.

Brown JR, North EJ, Hurdle JG, Morisseau C, Scarborough JS, Sun D, Kordulakova J, Scherman MS, Jones V, Grzegorzewics A, **Crew R**, Jackson M, McNeil MR, Lee RE: **The structure-activity relationship of urea derivatives as anti-tuberculosis agents**. *Bioorg Med Chem* 2011, **19**(18): 5585-5595.

England K, **Crew R**, Slayden RA: **Mycobacterium tuberculosis septum site determining protein< Ssd encoded by rv3660c, promotes filamentation and elicits an alternative metabolic and dormancy stress response**. *BMC Micro Biol*

2011, **Apr 19** 11:79.

Selected Presentations and Posters:

"Cell division and Non-Replicating Persistence: Identification of Cell Division Regulators in *Mtb*" – Graduate Student Seminar Series, February 28th, 2012

"Characterization Identification of a Novel Conserved Regulator in *Mycobacterium tuberculosis*" – Mycobacterial Research Laboratories Seminar Series, January 22th, 2013

"Cell division control and adaptations to dormancy in *M. tuberculosis*" – Graduate Student Seminar Series, April 23rd, 2013

"Identification of a Novel Conserved Cell Division Regulator, Encoded by *rv2216*, in *Mtb*" – Poster presented at the Center for Environmental Medicine Annual Spring Symposium, May 8th, 2013

Characterization of a Widely Conserved Division Regulator, Encoded by *rv2216*, in *Mtb*" – Poster to be presented at the Front Range Mycobacteria Conference, June 21st, 2013

LIST OF ABBREVIATIONS

ANOVA	Analysis of variance between groups
BLAST	Basic local alignment search tool
cfu	Colony forming unit
CL	Cardiolipin
CoA	Coenzyme A
DAG	Diacylglycerol
DAPI	4',6-diamidino-2-phenylindole
dos	*dosR*-regulon
EC	Enzyme commission number
EHR	Enduring hypoxic response
EMB	Ethambutol
FA	Fatty acid
FAS	Fatty acid synthase
FFA	Free fatty acid
G3P	Glycerol-3-phosphate
GFP	Green fluorescent protein
GPAT	Glycerol-3-phosphate acyltransferase
HBSS	Hank's balanced salt solution
HIV/AIDs	Human immunodeficiency virus/Acquired immunodeficiency syndrome
INH	Isoniazid
KEGG	Kyoto encyclopedia of genes and genomes
LAM	Lipoarabinomannan
LB	Lipid bodies
LM	Lipomannan
LPA	Lysophosphatidic acid
LPL	Lysophospholipid
LTBI	Latent tuberculosis disease
MAFFT	Multiple sequence alignment by the fast Fourier transform
MDR-TB	Multi-drug resistant tuberculosis
MMC	Mitomycin C
Mtb	*Mycobacterium tuberculosis*
NRP	Non-replicating persistence
$O.D._{600nm}$	Optical density at 600nm
OMA	Orthologous matrix
OperonDB	Operon database
ORF	Open reading frame
PATRIC	Pathosystems resource integration center
PDH	Pyruvate dehydrogenase complex
PE	Phosphatidylethanolamine
Pfam	Protein families

Phyre2	Protein homology/analogy recognition engine V2.0
PI	Phosphatidylinositol
PIM	phosphatidylinositol mannoside
PLA	Phospholipase A
PMP	Polymethylated polysaccharides
PNR/P	Peroxynitrite reductase/peroxidase complex
PS	Phosphatidylserine
PZA	Pyrazinamide
qRT-PCR	quantitative real-time PCR in conjunction with reverse-transcriptase PCR
RefSeq	Reference sequence database
RIF	Rifampin
RNAP	RNA-polymerase
RNI	Reactive nitrogen intermediates
ROI	Reactive oxygen intermediates
SE	Steryl esters
SEM	Scanning electron microscopy
TAG	Triacylglycerol
TB	Tuberculosis disease
TBDB	Tuberculosis database
TCA	Tricarboxylic acid cycle
TMHMM	Transmembrane hidden markov model
UV	Ultraviolet radiation
WHO	World health organization
XDR-TB	Extremely drug resistant tuberculosis