MYCOBACTERIOPHAGE LYSINS: BIOINFORMATIC CHARACTERIZATION OF LYSIN A AND IDENTIFICATION OF THE FUNCTION OF LYSIN B IN INFECTION

by

Kimberly Marie Payne

B. S. in Biochemistry & Molecular Biology, The Pennsylvania State University, 2006

Submitted to the Graduate Faculty of

Arts and Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2010

UNIVERSITY OF PITTSBURGH

ARTS AND SCIENCES

This dissertation was presented

by

Kimberly M. Payne

It was defended on

September 30, 2010

and approved by

Jeffrey L. Brodsky, Ph.D., Biological Sciences, University of Pittsburgh

Roger W. Hendrix, Ph.D., Biological Sciences, University of Pittsburgh

Paul R. Kinchington, Ph.D., Biological Sciences, University of Pittsburgh

Jeffrey G. Lawrence, Ph.D., Biological Sciences, University of Pittsburgh

Dissertation Advisor: Graham F. Hatfull, Ph.D., Biological Sciences, University of Pittsburgh

MYCOBACTERIOPHAGE LYSINS: BIOINFORMATIC CHARACTERIZATION OF

LYSIN A AND IDENTIFICATION OF THE FUNCTION OF LYSIN B IN INFECTION

Kimberly Marie Payne, PhD

University of Pittsburgh, 2010

Tuberculosis kills nearly 2 million people each year, and more than one-third of the world's population is infected with the causative agent, *Mycobacterium tuberculosis*. Mycobacteriophages, or bacteriophages that infect *Mycobacterium* species including *M. tuberculosis*, are already being used as tools to study mycobacteria and diagnose tuberculosis. More than 60 mycobacteriophage genomes have been sequenced, revealing a vast genetic reservoir containing elements useful to the study and manipulation of mycobacteria. Mycobacteriophages also encode proteins capable of fast and efficient killing of the host cell. In most bacteriophages, lysis of the host cell to release progeny phage requires at minimum two proteins: a holin that mediates the timing of lysis and permeabilizes the cell membrane, and an endolysin (lysin) that degrades peptidoglycan. Accessory lysis proteins have also been discovered, often with functions specific to that phage's host.

Many lysins of phages infecting Gram-positive bacteria are proving to be potent antibacterials. Further, lysis proteins can provide insight into the properties and composition of the host cell wall. Given the complexity of the mycobacterial cell wall and its medical relevance in tuberculosis as an immunogenic barrier that complicates treatment, as well as the urgent need for new therapeutic options, the mycobacteriophage lysins clearly warrant further scientific investigation.

This work focuses on the mycobacteriophage lysin LysA and the accessory lysis protein LysB. Bioinformatic characterizations show that LysA proteins posess a variety of domains arranged in modular organizations, reflecting extensive recombination within the mycobacteriophage population. In addition to known peptidoglycan-hydrolytic activities, novel cell wall-binding domains are identified, as well as several domains of unknown function found only in mycobacteriophages. LysB proteins are unique to mycobacteriophages and perform a singular role as mycolylarabinogalactan esterases that sever the connection between the mycobacterial outer membrane and the peptidoglycan cell wall complex to ensure efficient lysis and progeny phage release. There is also preliminary evidence of peptidoglycan hydrolytic ability, inducible cell lysis, and growth inhibition of *Mycobacterium smegmatis* by LysA and LysB proteins. These studies suggest that mycobacteriophage lysis proteins can be exploited as useful tools, both in the laboratory and clinical setting.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

xiv

PREFACE

To start, I wish to thank my advisor, Dr. Graham Hatfull, for the invaluable guidance he has provided me as I pursued my doctorate. Graham provided a wealth of knowledge, patience, and insight to my aid, but it was his positive attitude and enthusiasm that truly inspired me, and I hope to emulate his passion for science and sharing it through education and outreach.

I also must thank my thesis committee, to whom I attribute much of my success. My bioinformatic analyses would not have been possible without the help of Dr. Jeffery Lawrence, who patiently taught me not only the methods but the underlying theory as well. Always questioning, Dr. Jeff Brodsky pushed me to think critically and consistently offered insightful advice. I greatly admire his ability to ask these penetrating questions, and I hope to someday be just as dangerously inquisitive. I am inspired by Dr. Roger Hendrix's deep passion for science and the myriad ways he expresses it. I will miss his outrageous phage phacts, as well his clarinet leading the BSO down the hallway. Finally, I'm grateful to Dr. "Kip" Kinchington, both for his advice over the years and later accepting me into his lab as a postdoc. Early on I was impressed by his focus on critical thinking, and I look forward to his mentorship as my career continues.

I was lucky to be in a lab of smart, kind, supportive, and fun people who made my time in grad school so enjoyable. Several people have had a great impact on my life these last four years. Dr. Laura Marinelli has mentored me in every aspect of science, and I could not have accomplished as much as I have without her. Dr. Greg Broussard was a constant source of

support and advice, on science, careers, and life, and it is because of him that I feel prepared to move forward. And it was Amrita Balachandran who kept me sane and was always there for me, even when she "had no words". I greatly admire her kindness, patience, and passion for teaching. I also won't forget the plethora of undergrads and Phagehunters, especially Andrew Savinov, Drew Hryckowian, Tim Sampson, and Charlie Bowman.

I am hugely grateful to the entire Department of Biological Sciences; I can't imagine a more friendly, close-knit, and fun department and my interactions with everyone greatly enriched my graduate experience. I want to acknowledge Crystal Petrone, Cathy Barr, and Pat Dean, who worked behind the scenes to keep my academic life on track, and Dr. Tom Harper for sharing Photoshop tricks and a lot of good conversation. I also wish to thank those who mentored me in teaching, including Drs. Valerie Oke, Karen Arndt, and Melanie Popa. Melanie, I cannot express how thankful I am to have you as a teaching mentor and friend. You are one of the kindest, most generous people I know, and your love of teaching has been an inspiration to me.

Thank you to all those in the Program in Integrative Molecular Biology, especially my own class: Drs. Eric de Groh, Arpana Dutta, Xi Liu, and Naveena Yanamala. As a member of the inaugural class, I've had the privilege of watching our fledgling program grow strong on the success of our exceptional students. We are more than a program: we are our own community.

Lastly, my heartfelt appreciation goes to my family. I love my wonderful, brilliant, amazing husband Jonathan, who gave me unwavering support, constant love, and a really cool last name. And I credit my parents, Paul and Nonette Clemens, with my success, for it was their belief in me that got me this far. I especially wish to thank my father, for all of the walks in the woods, trips to the library, and impulsive experiments that usually involved or resulted in a fire. You taught me to question the world around me and to pursue my passion. Thank you.

1.0   INTRODUCTION

Tuberculosis (TB) is a worldwide health concern; nearly two million people die from TB annually, making it the number one cause of death due to an infectious agent (WHO, 2009). More than one-third of the global population is infected with *Mycobacterium tuberculosis*, and treatment has become increasingly difficult. The complex, waxy cell wall and slow growth of the mycobacterium complicate treatment, and antibiotic resistance is increasing (CDC, 2009; Karakousis et al., 2004). There is an urgent need for new therapeutic options for TB, as well as for continued study to better understand this pathogen.

One source of tools for the study and manipulation of mycobacteria are the mycobacteriophages, phages that infect *Mycobacterium* species. Tailed bacteriophages are the most abundant biological entities in the biosphere and are predicted to be the largest reservoir of genetic information (Hendrix, 2003; Suttle, 2005; Wommack and Colwell, 2000). More than 60 mycobacteriophages have been fully sequenced and 6,858 putative open reading frames identified, the majority of which are of unknown function (Hatfull et al., 2010). Whole phages have been used both as genetic tools and for diagnostic and strain-typing purposes (Lee et al., 2004; Piuri et al., 2009; Snider et al., 1984) but specific phage proteins are beginning to be exploited for the manipulation of mycobacteria (van Kessel and Hatfull, 2007). However, very few studies have attempted to identify cytotoxic proteins in mycobacteriophages (Rybniker et al., 2008).

Bacteriophages were first characterized in the early 1900s, but research into therapeutic application of phages declined when antibiotics were commercialized in the 1940s (Borysowski et al., 2006; Hermoso et al., 2007). With the advent of growing antibiotic resistance, investigation into phage therapies is increasing (Fischetti, 2006). Several studies have focused on the potential of the well-known cytotoxic proteins involved in host lysis (Borysowski et al., 2006; Hermoso et al., 2007; Loessner, 2005). Bacteriophages lyse their hosts at the end of lytic infection by compromising the integrity of the cell wall to release progeny phages. The mechanism of cell lysis by dsDNA-tailed phages involves at minimum a holin and an endolysin (lysin) (Young, 1992). Lysins have a modular design, composed of combinations of peptidoglycan lytic domains and host-specific cell wall-binding domains that have evolved through recombination between the host and other phages (Lopez et al., 1997). Recently, the lysins of phages infecting Gram-positive bacteria have shown promise as highly-specific antibacterial agents in animal disease models and food biocontrol applications (Fischetti, 2006). These lytic proteins are active on antibiotic-resistant strains and do not appear to generate resistance themselves (Loeffler et al., 2001), making them very appealing in the treatment of drug-resistant pathogens.

In addition to the therapeutic-focused research on bacteriophage lysis, studies are being conducted to better understand the underlying mechanism of phage lysis. These studies have identified accessory lysis proteins – in addition to the holin and endolysin – and are discovering their roles in phage lysis. Unlike the holin and lysin, these accessory proteins are restricted to certain phages and appear to have host-specific functions; for example, the Rz/Rz1 proteins, found only in phages infecting Gram-negative bacteria, are responsible for fusing the inner and outer membranes – the latter structure being unique to Gram-negative species (Berry et al.,

2008).   Because of this specificity, characterization of accessory lysis proteins can provide insight into both the mechanism of lysis by a specific phage and the nature of the host's cell wall.

The complex mycobacterial cell wall is of great interest to the medical community and is the focus of much research (Daffe and Draper, 1998; Karakousis et al., 2004; Kremer and Besra, 2005; Mahapatra et al., 2005a). Like all phages, mycobacteriophages must lyse their hosts to release progeny phages, and so they have evolved specific proteins suited to this purpose. Therefore, exploration of the mycobacteriophage lysis proteins serves two purposes: determining the mechanisms by which mycobacteriophages overcome the unique barriers present in the mycobacterial cell wall, as well as identifying cytotoxic proteins that may have potential as therapeutic agents.

My investigations have focused on two mycobacteriophage lysis proteins, the endolysin LysA and the accessory lysis protein LysB. A bioinformatic analysis revealed a great diversity of LysA proteins, with many different predicted lytic activities and several intriguing unknown domains.   In addition, I have determined that LysB functions to remove the barrier of the mycobacterial outer membrane as the final step of lysis.

1.1   BACTERIOPHAGE LYTIC INFECTION

Tailed bacteriophages are believed to be the most abundant biological entities on the planet, predicted to number approximately 10^{31} (Hendrix, 2003; Suttle, 2005; Wommack and Colwell, 2000). This vast population is responsible for an estimated 10^{24} new infections every second (Wilhelm et al., 2002). Bacteriophage infection begins with the adsorption of the phage to the host, usually through specific interactions between tail fiber proteins and a receptor on the

surface of the host cell. Once bound, the phage penetrates the cell wall to inject its genome into the cytoplasm (Figure 1 [1]). At this point, for a temperate phage such as λ, a decision must be made between the lysogenic or lytic pathway (Oppenheim et al., 2005). In lysogeny, the phage genome is integrated into the host chromosome, transforming the bacterium into a lysogen (Figure 1 [2]). The phage genome is then maintained indefinitely as a prophage until it is induced and enters the lytic pathway (Figure 1 [3]).

During lytic development, the virus uses the host machinery, energy, and macromolecular precursors to create new phage particles. The lytic pathway culminates in the production of large numbers of progeny phages and cell lysis to release them from the host. Viral genes are temporally expressed in early and late stages – these can be subdivided for more precise descriptions of specific phage life cycles (Little, 2006). Upon initiation of lytic growth, genes that encode proteins necessary for replication of the viral genome are sequentially transcribed from an early promoter (Figure 1 [4]). Other proteins are also produced that possess a variety of functions in different phages, from the regulation of late gene expression to the degradation of the host chromosome by T-even phages (Warren and Bose, 1968). Once replication is underway, structural and lysis genes are transcribed from a late promoter (Figure 1 [5]). The genomic DNA is packaged into proheads (Figure 1 [6]) that then expand to form rigid icosohedral capsids (Jardine and Anderson, 2006). These capsids are then assembled with tail and tail fiber proteins to form mature virions (Figure 1 [7]).

The final step of lytic infection is the release of the progeny phages from the host cell (Figure 1 [8]). The obstacle to phage escape is the bacterial cell wall, a complex structure involving one or two hydrophobic lipid membranes, a rigid network of peptidoglycan (PG), and various cell wall-associated polymers. Two basic components – the cell membrane and the PG –

are common to all bacteria, and dsDNA bacteriophages have evolved a universal system for tackling these obstacles: the holin and endolysin (Young, 1992). In addition to the holin and endolysin, there are specialized accessory lysis proteins that address unique aspects of the host's cell wall. The major types of bacterial cell walls are discussed below, followed by a description of the mechanisms phages have developed to subvert these barriers to lyse the cells and free the progeny phages to infect new hosts.

Lytic Pathway
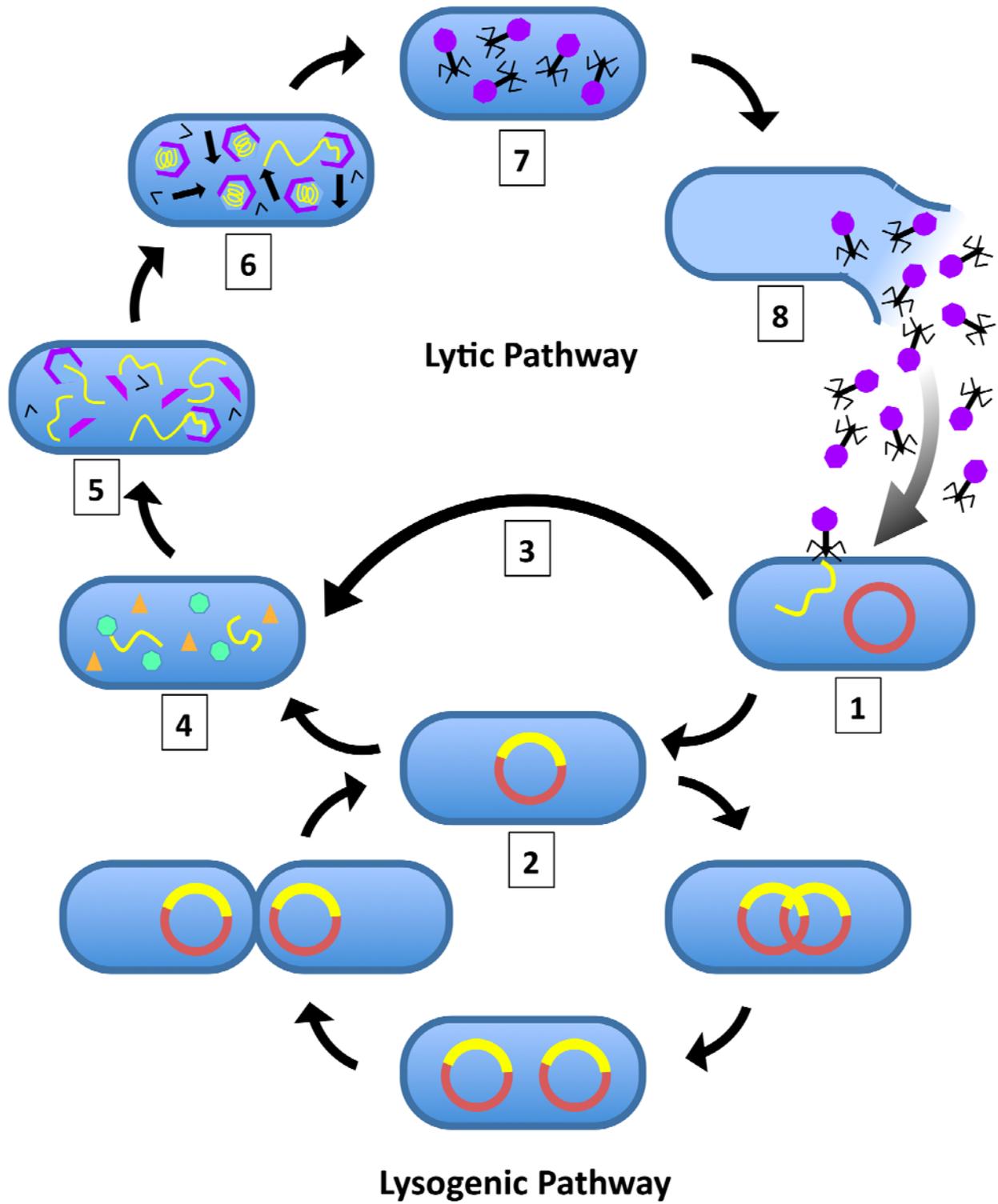
Lysogenic Pathway

Figure 1. Bacteriophage lytic and lysogenic pathways

Figure 1: Temperate bacteriophages can undergo lysogenic or lytic infections. (1) Phage adsorption and injection of genome; (2) Integration of phage genome into host chromosome, beginning the lysogenic pathway; (3) Direct progression to lytic pathway; (4) Early expression of genes to replicate genome; (5) Late expression of genes to create structural and lysis proteins; (6) Formation of proheads and packaging of DNA; (7) Assembly and maturation of new virions; (8) Cell lysis to release progeny phage to initiate new infections.

1.2   THE CELL WALL

The cell wall of bacteria is an extracellular structure composed of complex polymers and macromolecules that protects the cell in two ways. First, it preserves cell integrity via a mesh-like network of peptidoglycan (PG) that surrounds the cell in a structure called the sacculus (Schleifer and Kandler, 1972). This PG scaffold maintains the cell's shape and enables it to withstand internal osmotic pressures up to 25 atmospheres (Vollmer et al., 2008a), while remaining flexible enough to allow for the morphogenic changes required for cell growth and protein secretion (Dijkstra and Keck, 1996). Second, in conjunction other cell wall-associated components, the cell wall provides a permeability barrier that prevents the free diffusion of molecules into the cell. This is accomplished by one or two lipid membranes, both non-PG cell wall polymers and cell wall-associated molecules, and to some extent the PG network (Vollmer et al., 2008a). These and other components of the cell wall that vary between bacteria will be discussed in the sections specific to Gram-negative and Gram-positive species. Meanwhile, the PG is the only cell wall component common to both Gram-negative and Gram-positive bacteria.

1.2.1   Peptidoglycan

PG is a complex polymer of sugars and amino acids found in nearly all prokaryotes (Figure 2). The only exceptions are the *Mycoplasma* and *Planctomyces* species (Moulder, 1993; Tamura et al., 1995), which lack PG, and the presence of PG in the photosynthetic organelles of some eukaryotic algae (Aitken and Stanier, 1979). While the main purpose of the PG network is to

prevent the rupture of the cell due to internal osmotic pressure, the intact structure is surprisingly elastic—capable of reversibly shrinking or expanding up to 30% in *Escherichia coli* (Koch and Woeste, 1992). Gram-positive bacteria have a thick (15-30 nm) PG complex while Gram-negative have a monolayer (~2.5 nm) or at most three layers (Labischinski et al., 1991; Matias et al., 2003). The PG structure is the product of more than 20 enzymes and is largely invariant between bacteria, especially within the Gram-negative group (Schleifer and Kandler, 1972). Most variation is seen in the peptide portions in Gram-positive bacteria (Cummins and Harris, 1956; Schleifer and Kandler, 1972). The pervasiveness of the PG throughout the bacterial kingdom and its absence in eukaryotes make it an attractive target for anti-bacterial therapies. PG is essential to the cell; any inhibition of PG biosynthesis or degradation during cell growth will result in lysis of the cell (Vollmer et al., 2008a). For many decades, antibiotics such as penicillin that target the enzymes involved in PG synthesis have been used to great effect; however, with increasing antibiotic-resistance, new therapies are targeting the unvarying PG structure instead of the biosynthetic enzymes (Fischetti, 2008).

PG is composed of glycan strands that are connected via a short peptide bridge (Figure 2). The glycan strands contain alternating *N*-acetylglucosamine (GlcNAc) and *N*-acetylmuramic acid (MurNAc) residues linked by $\beta(1\rightarrow4)$ bonds (Ghuysen, 1968). Substituting most of the D-lactoyl moieties on the C3 of each MurNAc is a peptide stem of three to five amino acids. The disaccharide of GlcNAc-β-(1$\rightarrow$4)-MurNAc with a peptide moiety is considered a muropeptide (Vollmer et al., 2008a). The peptide stems of one PG strand can cross-link with neighboring peptides to create a network of PG. The type and amount of cross-linking varies for different bacteria, as well as under different growth states and environmental conditions (Fordham and Gilvarg, 1974; Pisabarro et al., 1985). In addition, the C6s of many MurNAcs are substituted

through a phosphodiester linkage with species-specific cell polymers such as the teichoic acids

found in the cell wall of Gram-positive bacteria (Shockman and Barrett, 1983).
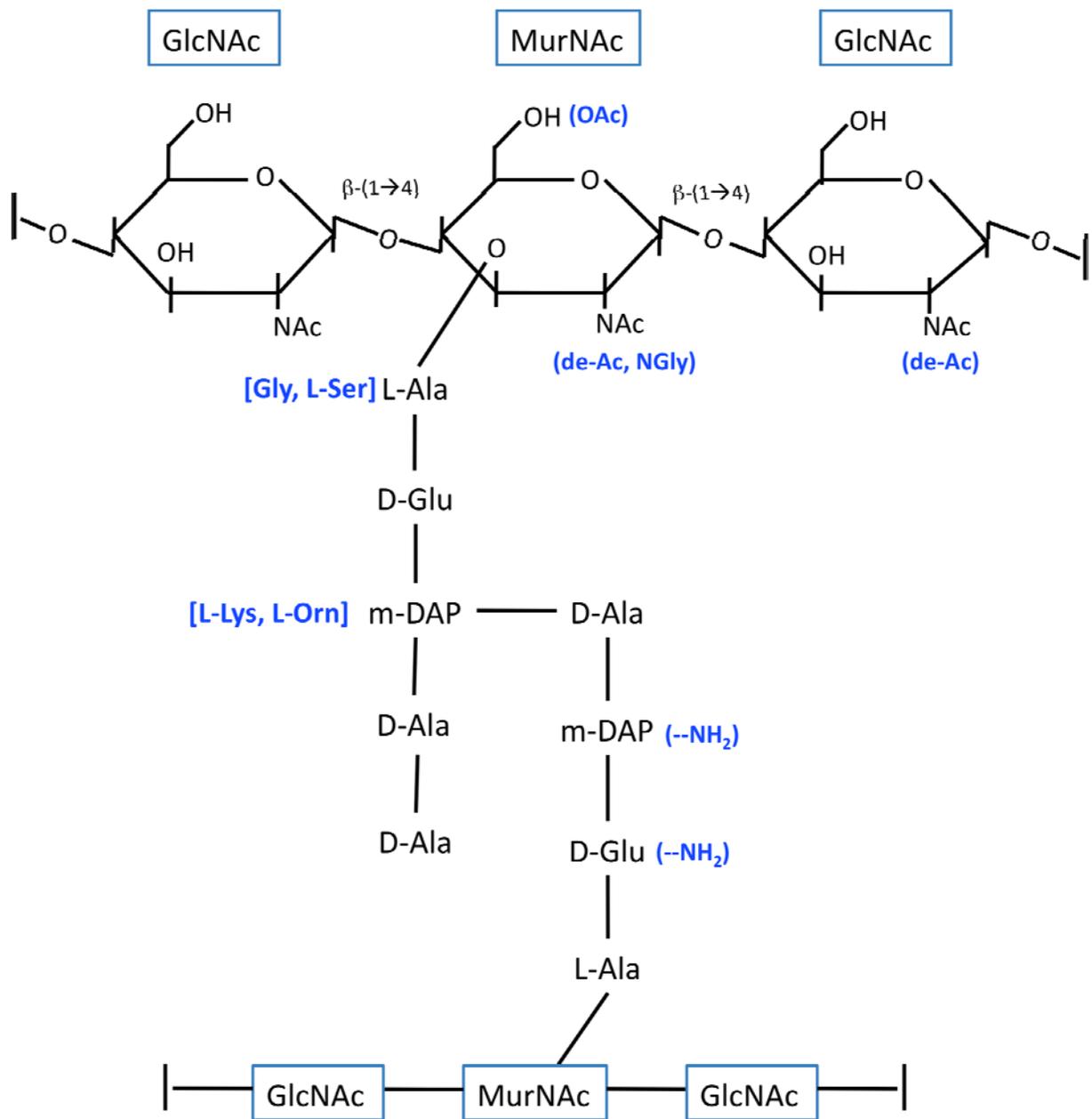
Figure 2. Peptidoglycan Structure

Figure 2: This diagram of peptidoglycan is based on type A1γ (Schleifer and Kandler, 1972). Glycan strands are identified as GlcNAc (*N*-acetylglucosamine) and MurNAc (*N*-acetyl muramic acid) linked with a β-(1→4) bond. In the peptide cross-link, m-DAP represents *meso*-diaminopimelic acid, and all other abbreviations are those standard for isomers of amino acids. Modifications or alternate residues are shown in blue and include de-acetylation of MurNAc and GlcNAc, *O*-acetylation and *N*-glycolylation of MurNAc, amidation of the D-Glu or m-DAP, and common substitutions in the peptide cross-links: L-Lys or L-Orn (ornithine) for m-DAP, and Gly or L-Ser for L-Ala.

1.2.1.1 Glycan strands

The glycan strands comprise the rigid portion of the PG structure and are formed through the oligomerization of monomeric disaccharide peptide units by transglycosylation reactions (Vollmer et al., 2008a). The length of the strands can vary greatly depending on the bacterial species and growth conditions, although this is not correlated with the thickness of the PG (Vollmer et al., 2008a). Glycan strands average between 20 – 40 disaccharide units (Vollmer et al., 2008a); *E. coli* and *Staphylococcus aureus* are within this range with an average of 21 and 18 units, respectively (Quintela et al., 1995; Ward, 1973). At the far ends of the spectrum, *Helicobacter pylori* glycan strands average <10 units (Costa et al., 1999), while those of *Bacillus subtilis* range from 50 – 250 units (Ward, 1973). The nature of the chain ends also varies; Gram-negative and some Gram-positive bacteria glycan chains terminate in 1,6-anhydro-MurNAc residues, which are formed through cleavage by a lytic transglycosylase (Vollmer and Holtje, 2001). The glycan chains of other bacteria are hydrolyzed by a glucosaminidase or lysozyme, resulting in MurNAc or GlcNAc residues containing a free reducing end (-OH) at the C1 or C4 (Vollmer et al., 2008a).

Other modifications are seen on the glycan strands that can affect the stability of the cell wall, increase resistance to degradative enzyme, and confer pathogenic properties (Vollmer, 2008). Many *Bacillus* species have *N*-deacetylases that remove the acetyl group from a portion of their GlcNAc and MurNAc residues (Figure 2), and some Gram-positive bacteria, including some *Micrococcus* and *Streptococcus, O*-acetylate some of the MurNAc residues by adding an extra acetyl group on the C6 (Figure 2) to create 2,6-*N,O*-diacetyl muramic acid (Vollmer, 2008). *N*-glycolylation changes the acetyl group (-COCH$_3$) on the C2 of MurNAc to a glycolyl group (-COCH$_2$OH) (Figure 2). This modification was first described in *M. smegmatis* (Adam et al.,

1969) and is a hallmark of many Actinomycetales genera (Raymond et al., 2005; Vollmer, 2008). The prevalence of these alterations within a cell's PG varies with growth phase and other environmental conditions. All of these slight modifications of the GlcNAc and MurNAc chains subtly change the properties of the cell wall; many appear to increase the stability of the cell wall and all seem to increase the bacterium's resistance to lysozyme (Vollmer, 2008).

1.2.1.2 Peptide cross-links

The peptide cross-links are formed through transpeptidation reactions between two neighboring peptide stems and are responsible for the flexibility of the PG sacculus (Vollmer et al., 2008a). All peptide subunits are formed of alternating L- and D-isomers of amino acids (Figure 2); the specific amino acids are largely uniform in Gram-negative bacteria but can vary among Gram-positive species (Schleifer and Kandler, 1972). Regardless of composition, the muropeptides begin with pentapeptide moieties that are cross-linked by an L,D-transpeptidase to another pentapeptide in a way that causes one pentapeptide to lose its terminal amino acid, becoming a tetrapeptide. Further reduction of the peptide stem by D-Ala-carboxypeptidases can also create tripeptides that are unable to form cross-links (Navarre et al., 1999).

The most common arrangement of amino acids in the peptide stem is (MurNAc)—L-Ala—D-Glu—m-DAP—D-Ala—D-Ala, where m-DAP is *meso*-diaminopimelic acid (Figure 2). This configuration is found in nearly all Gram-negative bacteria, as well as in *Bacillus* and *Mycobacterium* species (Schleifer and Kandler, 1972). Each of these amino acids occupies a particular ordinal position with respect to the MurNAc residue: L-Ala is position 1, D-Glu is position 2, m-DAP is position 3, and the two D-Ala are in positions 4 and 5. The transpeptidation reaction creates a peptide bond between the free carboxyl of the donor in position 4 and the free amino group of the acceptor in position 2 or 3 (Vollmer and Holtje, 2001).

13

The energy for the reaction comes from the release of the D-Ala in position 5 of the donor's pentapeptide. However, there are various possible replacements and alterations for the amino acids in the peptide stem, mostly occurring in Gram-positive species. The L-Ala in position 1 and the two D-Alas in positions 4 and 5 are nearly universal, but other positions 2 and 3 are highly variable. Interestingly, most other amino acids are excluded from the PG, likely based on their structure; *e.g.* branching amino acids (Val, Leu, Thr); aromatic (Phe, Tyr, Trp); sulfur-containing (Met, Cys); His, and Arg (Schleifer and Kandler, 1972).

Schleifer and Kandler (1972) designed a classification system for the peptide cross-links in bacteria based on three characteristics. First are the two groups A and B, defined by the anchoring point of the cross-linkage. Group A cross-linking occurs between the amino acids in position 3 of one peptide stem and position 4 of the other. In the example pentapeptide above, this would involve a link between the m-DAP (position 3) and the first D-Ala (position 4) of another peptide. Group A cross-linking is the most common throughout all bacteria. Group B involves a cross-link between position 2 (D-Glu above) and position 4 (D-Ala); this group is rare, primarily found in coryneform bacteria, and will not be further discussed.

The second level of classification concerns the nature of the cross-link (Schleifer and Kandler, 1972). Four subgroups are identified (A1-4) based on the presence or absence of an interpeptide bridge. Subgroup 1 forms direct cross-links, such as from the γ-carboxyl of the ω-amino group of the diamino acid in position three to the carboxyl group of the D-Ala in position 4. Subgroup 2 forms cross-links from position 3 of one peptide stem through a polymerized peptide subunit to the D-Ala in position 4 of the other peptide stem. Lastly, subgroups 3 and 4 utilize interpeptide bridges such as penta-Gly in *S. aureus* or a second diamino acid.

14

The third level of classification looks at the diamino acid in position 3, which is the most variable component in the peptide stem (Schleifer and Kandler, 1972). Continuing with the most common type of PG, there are three variations of A1 peptide cross-links; A1α has a L-Lys in position 3, A1β has L-Orn, and A1γ has m-DAP. The final distinctions are based on slight modifications to the peptide stem. The m-DAP variation (A1γ), can be divided into types based on the possible amidation of free carboxyls. *E. coli* and *Bacillus megaterium* have no amidation, while some of the peptide subunits of *Bacillus licheniformis* are amidated at the α-carboxyl of D-Glu and of *B. subtilis* are amidated on the free carboxyl group of m-DAP. Lastly, the cell walls of *Lactobacillus*, *Corynebacterium diphtheriae*, and *Mycobacterium* can be amidated at both locations (Mahapatra et al., 2005c).

1.2.1.3 Alterations and Cell Growth

The PG is constantly being turned-over as the cell grows. Elongation of the sacculus is accomplished by incorporating new material at multiple sites along the cell wall (Burman et al., 1983). In Gram-positive bacteria, new PG is inserted along the inner face of the cytoplasmic membrane, displacing older PG outwards (Koch and Doyle, 1985); this PG is recycled into the growth media (Vollmer et al., 2008b). In one generation of *E. coli*, an estimated 40-45% of the PG is released by lytic transglycosylases, amidases, and endopeptidases to be recycled (Goodell, 1985; Vollmer and Holtje, 2001; Vollmer et al., 2008b). Cell wall growth can also be regulated through the action of D,D-carboxypeptidases, which slow growth by removing the D-Ala from peptide pentapeptides, creating a peptide stem that can act as an acceptor for a cross-link but not a donor (Vollmer et al., 2008b). Further removal of the D-Ala in position 4 will leave a tripeptide unable to participate in any cross-linking.

PG-degrading enzymes, principally lytic transglycosylases and amidases, are also responsible for the creation of the large holes in the PG required for various structures and events (Scheurwater et al., 2008; Vollmer et al., 2008b). Degradation of the cell wall PG is absolutely required for cell division, and this is achieved at the septum by amidases and lytic transglycosylases (Vollmer et al., 2008b). Targeted PG degradation is also necessary for the export of large proteins and protein complexes, the assembly of flagella and conjugation pili, and the competency of the cell for transforming extracellular DNA (Dijkstra and Keck, 1996). Finally, as cells enter stationary phase, cross-linking increases (Fordham and Gilvarg, 1974; Pisabarro et al., 1985), creating a more rigid and protective barrier that has increased resistance to lysozymes (Vollmer et al., 2008a).

1.2.2   Gram-negative Cell Wall

The cell wall of Gram-negative bacteria contains little PG (<10% of cell wall weight) (Schleifer and Kandler, 1972), but is rich in lipids. Outside of the cytoplasmic membrane are the periplasmic space, the PG complex, and the asymmetrical outer membrane (Figure 3 A) (Matias et al., 2003). The periplasm is immediately outside of the cytoplasmic (or inner) membrane and extends to the PG layer, and lipoproteins anchor the PG along the outer membrane. Lipoproteins are covalently attached to the m-DAP of the PG peptides through their C-terminal Lys or Arg residues, while the fatty acid portion is inserted into the inner leaflet of the outer membrane (Holtje, 1998). Besides the lipoproteins, the inner leaflet is composed of the fluid phospholipids. The outer leaflet contains membrane proteins like porins, as well as the less flexible lipopolysaccharides, or LPS (Gronow and Brade, 2001).

The negatively-charged LPS cover approximately 75% of the cell's surface, acting as a barrier to hydrophobic substances like detergents and some antibiotics, while the rest of the outer membrane prevents the diffusion of hydrophilic compounds (Gronow and Brade, 2001). The LPS is composed of three parts: lipid A, the core, and the O-antigen (Gronow and Brade, 2001; Raetz and Whitfield, 2002). The O-antigen is a long polysaccharide chain that varies greatly among Gram-negative bacteria, even within a single species. The core links the O-antigen to lipid A and is composed of a combination of several sugars and non-carbohydrate modifications such as phosphate and amino acid substituents. The most conserved region is the lipid anchor (lipid A), a glucosamine-based phospholipid with multiple fatty acid chains that extend into the outer leaflet of the outer membrane. In particular, lipid A plays a role in pathogenesis as an endotoxin (Raetz and Whitfield, 2002). In humans, lipid A and other bacterial cell wall components are recognized by the toll-like receptors (TLRs) of the innate immune system. Upon the binding of these endotoxins, the TLRs initiate the inflammatory response and recruit phagocytic cells (Aderem and Ulevitch, 2000).
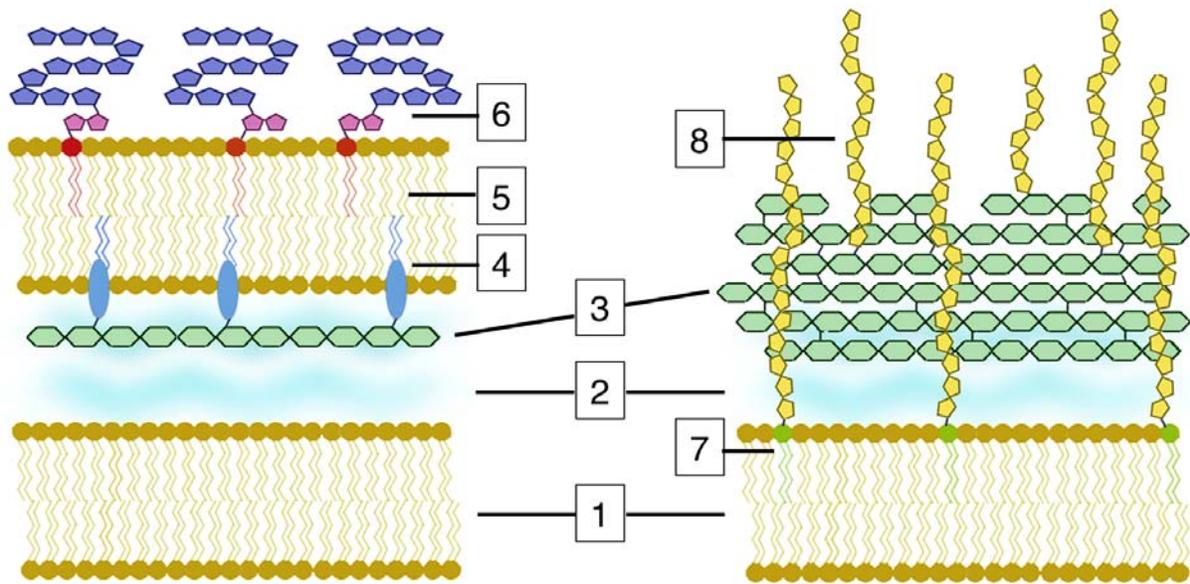
Figure 3. Gram-negative and Gram-positive bacterial cell walls

Figure 3: **A.** Gram-negative cell wall composed of a cell membrane, periplasmic space, thin layer of PG anchored by lipoproteins, and outer membrane with LPS. **B.** Gram-positive cell wall comprising a cell membrane, small periplasmic space, thick layer of PG, and polysaccharides (*e.g.* teichoic acid) extending from the membrane or PG. (1) cell membrane; (2) periplasmic space; (3) peptidoglycan (PG); (4) lipoproteins; (5) outer membrane; (6) lipopolysaccharides (LPS); (7) lipoteichoic acid, with lipid anchor in green; (8) teichoic acid.

1.2.3 Gram-positive Cell Wall

The Gram-positive cell wall is defined by a thick layer of PG that constitutes 30-70% of the total cell wall and surrounds a single cytoplasmic membrane (Figure 3 B) (Shockman and Barrett, 1983). In bacteria lacking a polysaccharide capsule or proteinaceous S-layer, the PG and associated cell wall polymers are the first contact with the external environment. While not a true permeability barrier like the outer or cytoplasmic membranes, this structure does serve to limit the access of larger molecules (Shockman and Barrett, 1983). A bipartite organization of the Gram-positive cell wall has been observed, using cryo-transmission electron microscopy (Matias and Beveridge, 2005, 2006; Zuber et al., 2006), which includes an inner wall zone of low density and an outer wall zone of high density (Figure 3 B). The inner wall zone averages between 16 – 22 nm in *S. aureus* and *B. subtilis* (Matias and Beveridge, 2005, 2006) and appears analogous to the periplasmic space in Gram-negative bacteria (Figure 3 A). It is expected to contain many secreted proteins and new wall polymers in the process of being added to the current PG and polysaccharide network (Vollmer et al., 2008a). Interestingly, several proteins that are found in a soluble form in the periplasm of Gram-negative bacteria have lipid-modified homologs in Gram-positive bacteria (Navarre et al., 1999). The lipid moiety likely serves to retain the proteins near the cytoplasmic membrane, whereas the homologous proteins in Gram-negative bacteria are bound by the inner and outer membranes. Beyond this periplasmic space, the dense outer wall zone likely comprises the PG-polysaccharide polymer complex and ranges between 15 – 30 nm based on the Gram-positive species, growth phase, and environmental conditions (Vollmer et al., 2008a).

Non-PG cell wall polymers account for 10-60% of the cell wall mass and primarily include polysaccharides such as teichoic acid (Schaffer and Messner, 2005). These anionic cell wall polymers are attached either via a phosphodiester bond to the C6 of some MurNAc residues or to a lipid moiety that tethers them to the cytoplasmic membrane (Navarre et al., 1999). The best-studied polysaccharides are the teichoic acids and teichuronic acids found in *Bacillus, Staphylococcus,* and *Micrococcus* (Shockman and Barrett, 1983); the choline-containing polysaccharides in *Streptococcus* (Hermoso et al., 2003); and the arabinogalactan and lipoarabinomannan in *Mycobacterium* and related genera (Brennan and Besra, 1997; Crick et al., 2001; Yagi et al., 2003). These polymers are indispensible for growth, and many proteins that localize to the PG have evolved to bind them, including host PG-remodeling enzymes, like the *Streptococcus pneumoniae* Atl autolysin (Giudicelli and Tomasz, 1984), and several streptococcal phage lysins (Lopez et al., 1997). In addition, many of these polymers are linked to pathogenesis and can illicit an immune response, similar to the LPS of *E. coli* and related Gram-negative bacteria (Aderem and Ulevitch, 2000).

1.3   MYCOBACTERIAL CELL WALL

1.3.1   Overall Organization

Mycobacteria have been traditionally categorized as Gram-positive bacteria due to their extensive network of PG (Brennan and Nikaido, 1995). Mycobacteria have a chemotype IV cell wall defined by the presence of arabinogalactan substituted with mycolic acids (Embley et al., 1986); altogether this is called the mycolic acid-arabinogalactan-peptidoglycan (mAGP) complex (Figure 4 A). However, the mycobacterial cell wall has been likened to that of Gram-negative bacteria, since the covalently-attached mycolic acids intercalate with various lipids to form an asymmetric lipid bilayer analogous to the outer membrane of Gram-negative bacteria (Minnikin et al., 1982). This layer with its associated lipids that create the cell envelope can account for up to 60% of the cell's dry weight (Kremer and Besra, 2005). There is one significant difference between the Gram-negative and mycobacterial cell wall structures: the mycobacterial outer membrane is completely tethered to the PG (Figure 4). As a result, the mycolic acids are essential for outer membrane integrity, as shown with the loss of the outer membrane in *Corynebacterium glutamicum* Δ*pks13*, a mutant that is defective in mycolic acid biosynthesis (Portevin et al., 2004). Mycolic acids are indispensible for the growth of mycobacterial species (Portevin et al., 2004; Portevin et al., 2005), and so experiments disrupting the arabinogalactan-mycolic acid structure are conducted with related *Corynebacterium* species.

Recently, this mycobacterial outer membrane has been visually verified with cryo-electron microscopy (Hoffmann et al., 2008; Zuber et al., 2008). Interestingly, these reports

found that the membrane was not much larger than the plasma membrane; the mycobacterial outer membrane was approximately 15% thicker than the cytoplasmic membrane (8 nm and 7 nm, respectively) (Hoffmann et al., 2008; Zuber et al., 2008). Considering the length of mycolic acids, this observation has led to the proposal of several new models for the membrane (Figure 4). In keeping with the asymmetrical bilayer model, Zuber et al. (2008) has proposed that the longer chains of the mycolic acids fold to stay within the inner leaflet (Figure 4 B). Hoffmann et al. (2008) instead proposed two possible models, one in which the longer meromycolate chain of the mycolic acids extends into the outer leaflet (Figure 4 C), and a second in which the base of the mycolic acids remains in the periplasm with only the ends extending into the inner leaflet (Figure 4 D).
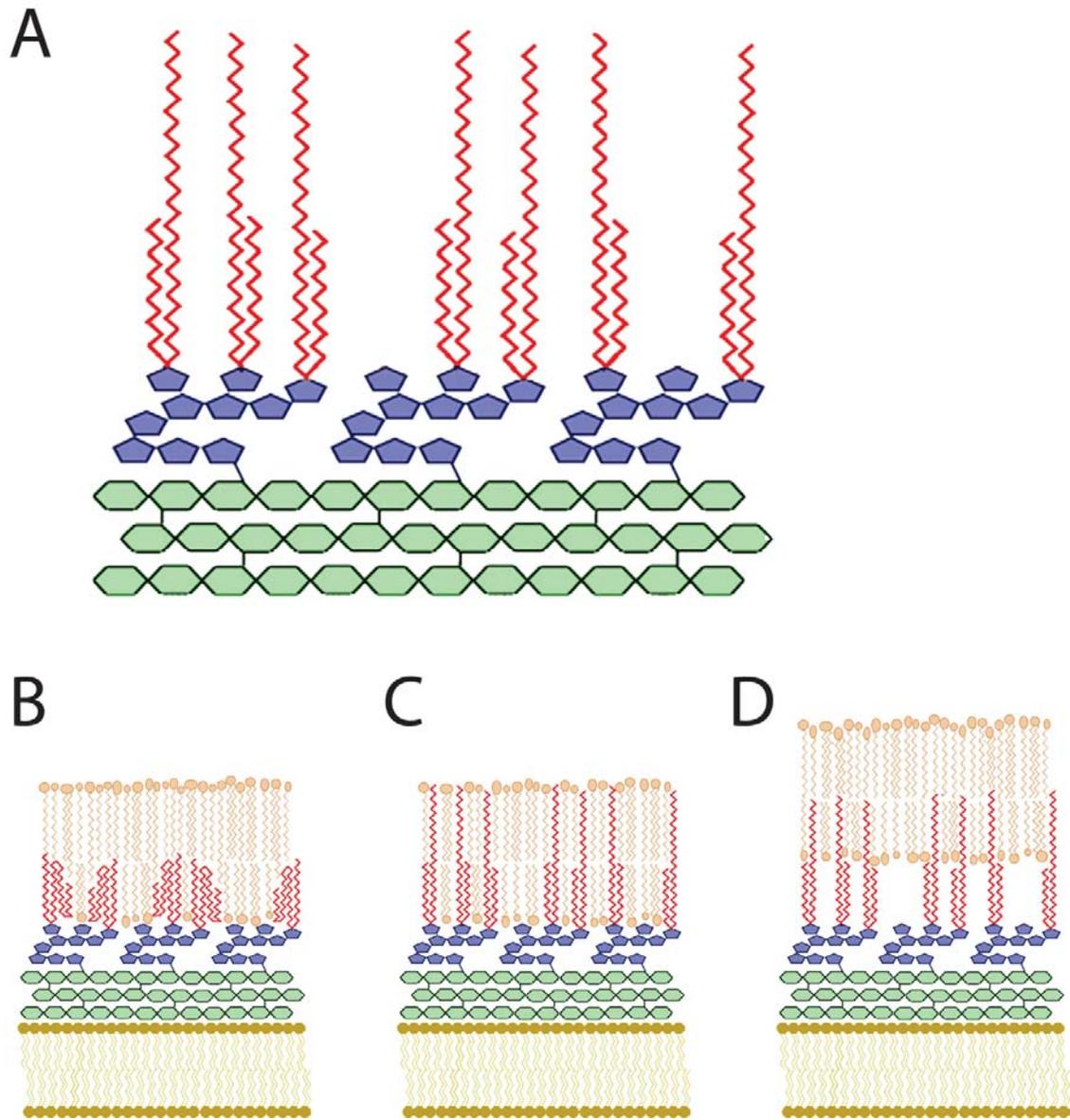
Figure 4. Models of structure of the mycobacterial cell wall

Figure 4: **A.** The mycobacterial cell wall core contains peptidoglycan (green) with branching chains of arabinogalactan (purple) that terminate in mycolic acids (red). **B-D.** The mAGP is outside of the plasma membrane (yellow), and the mycolic acids intercalate with extractable lipids (orange) to form the mycobacterial outer membrane. The specific organization of the outer membrane is still unknown, and based on the size of the membrane as seen with cryo-electron microscopy, three arrangements have been proposed: **B.** the long meromycolate chain of the mycolic acids folds to remain within the inner leaflet (Zuber et al., 2008); **C.** the meromycolate chain extends fully into the outer membrane; **D.** only the region of the meromycolate not interacting with the α-branch interacts with the lipids in the inner leaflet (Hoffmann et al., 2008).

1.3.2 Distinguishing Characteristics

1.3.2.1 Peptidoglycan Modifications

Mycobacterium and other members of the Actinomycetales suborder of Corynebacteriaceae have several significant differences in their PG structure. Approximately 75% of the PG is cross-linked in *Mycobacterium* species, compared to 20-30% in *E. coli* (Mahapatra et al., 2005a). These organisms have type A1γ PG, which is characterized by the inclusion of *meso*-diaminopimelic acid (m-DAP) (Schleifer and Kandler, 1972). The m-DAP is cross-linked to the D-Ala of another tetra-peptide moiety, but about one-third of the cross-linking in *Mycobacterium* species occurs between two m-DAP moieties (Figure 5); this is believed to provide additional rigidity to the PG (Brennan and Nikaido, 1995; Mahapatra et al., 2005a). The cross-linking peptides are also more often amidated on the Glu and m-DAP (Figure 5) (Petit and Lederer, 2000). Additionally, *Mycobacterium* and *Nocardia* have N-glycolylated muramic acids (MurNGly) instead of N-acetylation (Figure 5), a modification that introduces an additional alcohol group and may provide more opportunities for hydrogen bonding, further adding to the stability of the PG (Brennan and Nikaido, 1995; Raymond et al., 2005).
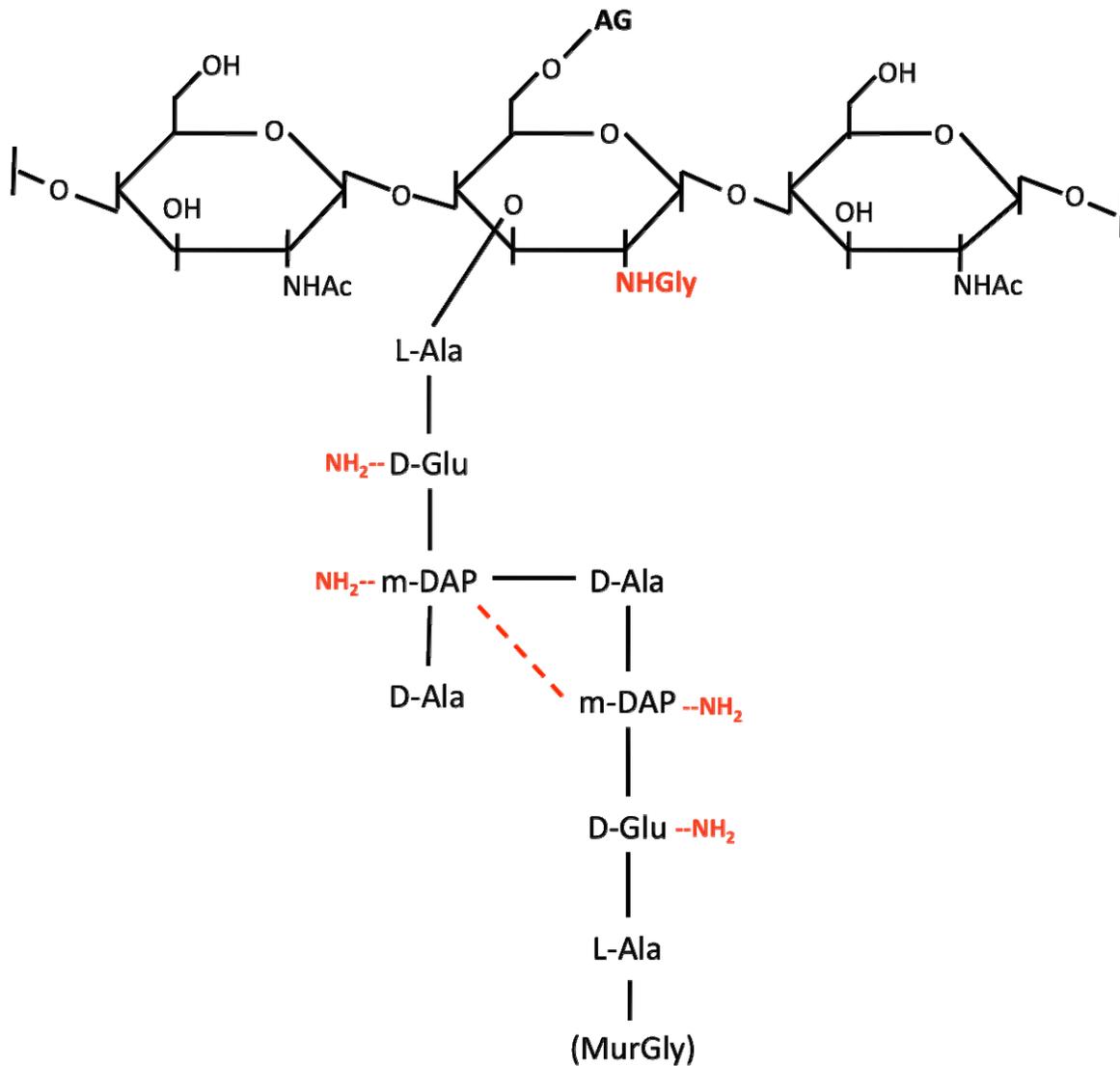
Figure 5. Peptidoglycan modifications in mycobacteria

Figure 5: The peptidoglycan structure of alternating sugars and the A1γ-type tetrapeptide, found in most Gram-negative and some Gram-positive species, are shown with modifications highlighted in red. Instead of N-acetylation, the muramic acids are N-glycolated, and the D-Glu and m-DAP of the peptide cross-link may be amidated. Also, in addition to the standard D-Ala → m-DAP linkage, there are some direct cross-links between two m-DAP (red dotted line).

1.3.2.2 Arabinogalactan

Branching from 10-12% of MurNGly residues (Brennan and Nikaido, 1995) are chains of mycolate-arabinogalactan, or AG (Figure 6 A). The AG is entirely composed of arabinose and galactose in the furanose form (Ara*f* and Gal*f* respectively) (Brennan and Nikaido, 1995). The galactan core is covalently attached to the C6 of MurNGly through a diglycosylphosphoryl bridge (McNeil et al 1990), and extends as a chain of approximately 30 alternating 5- or 6-linked β-D-Gal*f* residues (Daffe et al., 1990). Arabinin chains branch from the C5 of some 6-linked Gal*f* residues and forms 5-linked α-D-Ara*f* extensions. A 3,5-linked α-D-Ara*f* divides the chain into two branches followed by three more 5-linked Ara*f* residues and ending with the non-reducing terminal pentaarabinofuranosyl structure (Figure 6 B). This structure includes a 3,5-linked α-D-Ara*f* branching into two 2-linked β-D-Ara*f* units. Approximately two-thirds of the pentaarabinofuranosyl units are esterified by mycolic acids (McNeil et al., 1991).

A

galactan

pentaarabinofuranosyl
structure

arabinose

diglycosylphosphoryl
bridge

B

Ester linkages

β-D- Gal*f*

β-D- Gal*f*

β-D- Gal*f*

O-(β-D- Gal*f*)₅

β-D- Gal*f*

β-D- Gal*f*

[α-D- Rham*p*]

[α-D- GlcNAc*p*] Linker

MurGly

Mycolic acids

Pentaarabinofuranosyl
structure
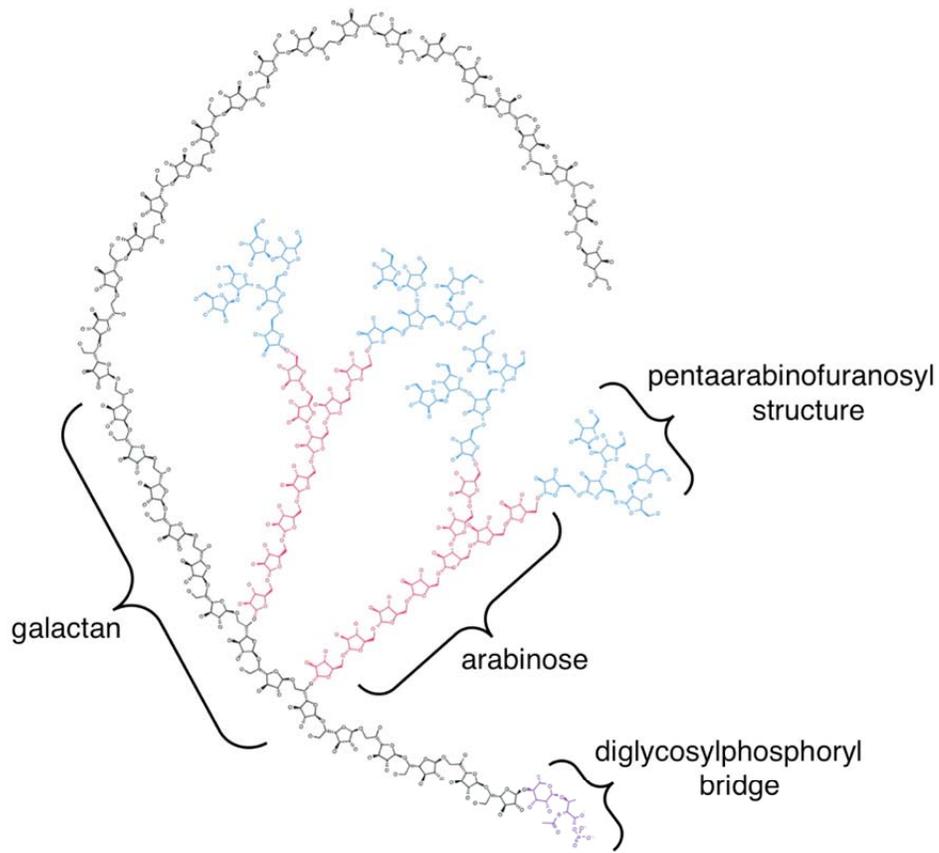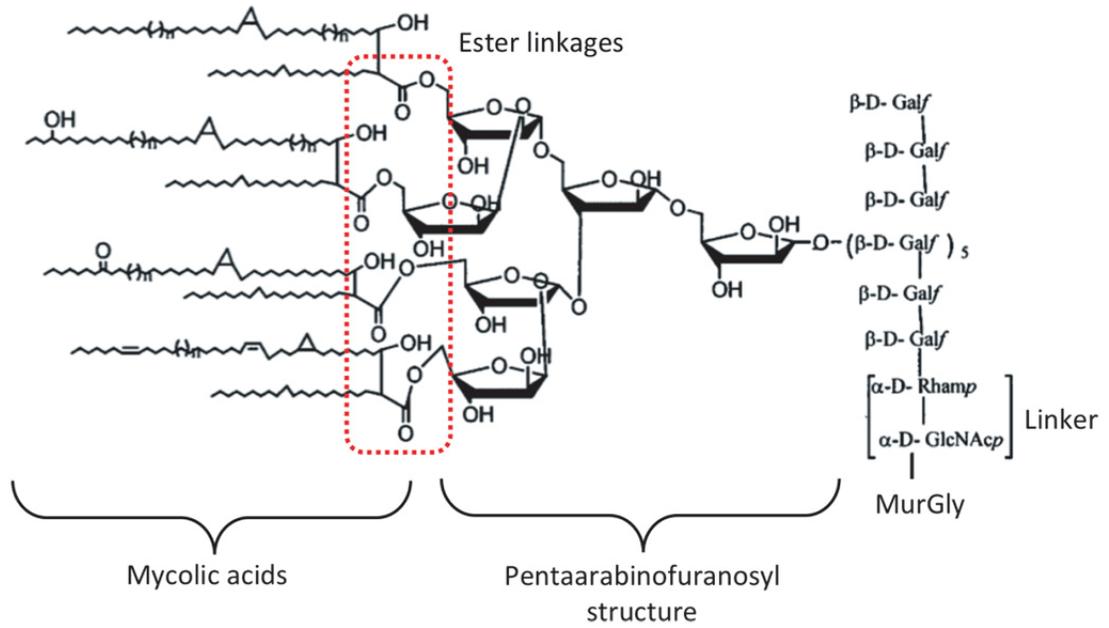
Figure 6. Arabinogalactan chain with branching mycolic acids

Figure 6: **A.** The long chain of approximately 30 galactan sugars (black) is connected to the C6 of some muramic acids through a linker, the diglycosylphosphoryl bridge (purple). From the C5 of some Gal*f* residues are branching chains of Ara*f* (red/blue) (adapted from Crick et al., 2001). **B.** The Ara*f* chains terminate in pentaarabinofuranosyl structures that are esterified (red dotted-line box) with mycolic acids (adapted from Tripathi et al., 2005).

1.3.2.3 Mycolic Acids

The mycolic acids that are characteristic of the mycobacteria are α-alkyl, β-hydroxy C_{60}-C_{90} fatty acids (Figure 7 A), larger than those in the related *Corynebacterium* and *Nocardia* genera (Jacobs, 2000). The saturated α-branch averages C_{20}-C_{25}, while the main chain meromycolic acid moiety (the β-hydroxy branch) averages C_{60} and can contain double bonds, cyclopropane rings, and oxygen functions (Figure 7 B) (Kremer and Besra, 2005). These meromycolate modifications define the types of mycolic acids and can be used to identify mycobacterial species through isolation and analysis by mass spectroscopy (Watanabe et al., 2001) and thin-layer chromatography (Hamid et al., 1993). α-mycolates (those lacking any oxygenation of the meromycolate chain) are found in all mycobacteria, while α'-mycolates are thus far only seen in fast-growing strains (Jacobs, 2000). Oxygenated mycolate types include methoxy, epoxy, keto, and wax ester; keto-mycolates are characteristic of slow-growing mycobacteria species (Jacobs, 2000). In addition to the α and α' mycolates, *M. smegmatis* also has epoxy mycolates (Figure 7 B) (Parish and Stoker, 1998). *M. tuberculosis* contains α, methoxy, and keto mycolates (Figure 7 B) (Parish and Stoker, 1998; Watanabe et al., 2001).
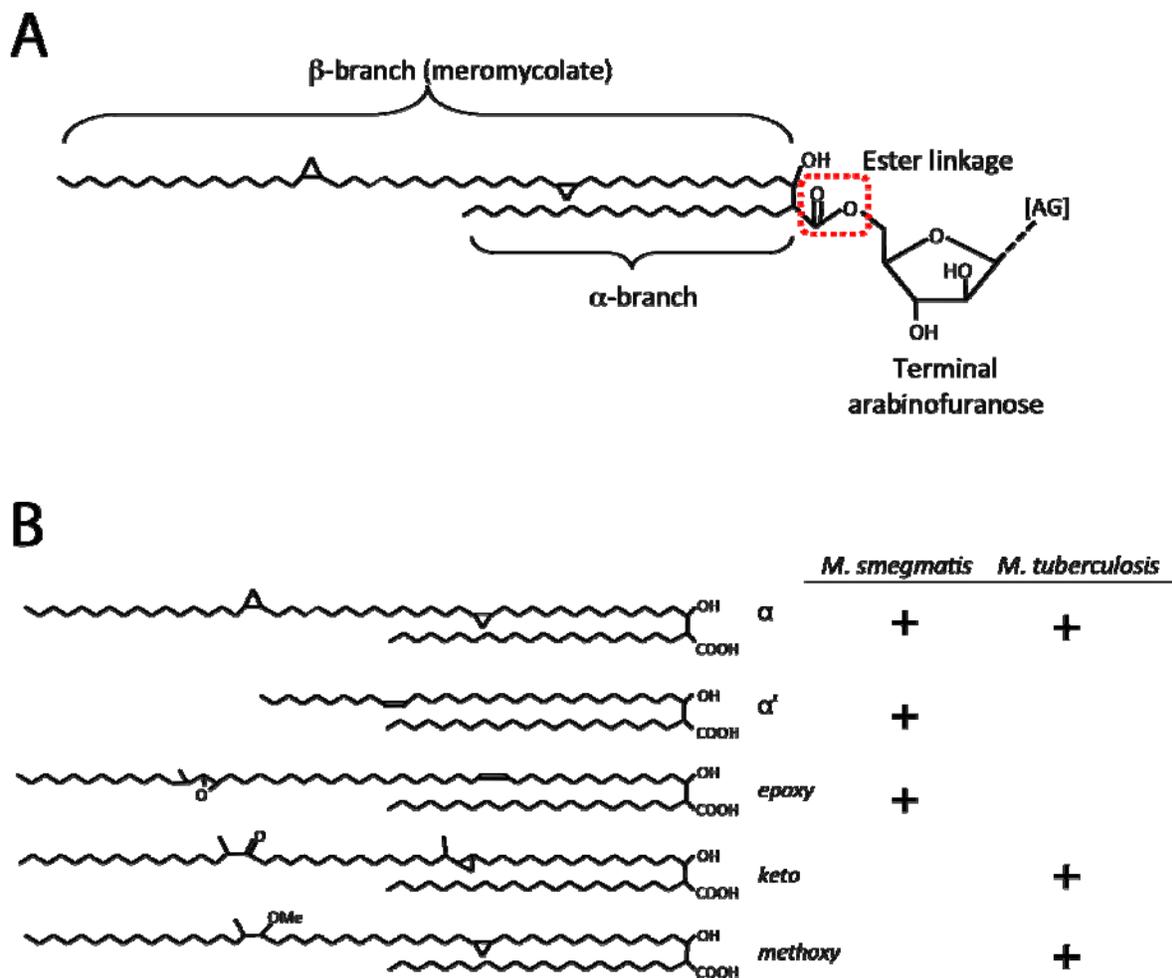
Figure 7. Mycolic acid structures

Figure 7: **A.** A mycolic acid esterified (red dotted-line box) to a terminal arabinofuranose residue. The α-branch and β-branch or meromycolate are indicated. **B.** Several different mycolic acid types that contain cyclopropane rings, double bonds, and oxygen groups. A (+) indicates which mycolic acids are present in *M. smegmatis* or *M. tuberculosis*.

1.3.2.4 Other Components

Lipoarabinomannan

Anchored in the cell membrane is the lipoarabinomannan (LAM), which is analogous to the lipoteichoic acids found in Gram-positive bacteria. LAM consists of a mannose chain extending from phosphatidyl-*myo*-inositol, a phospholipid inserted into the plasma membrane (Briken et al., 2004). The α-1,6-linked mannose backbone is substituted at C2 by a single mannose unit. The arabinin chain extends from the end of the mannose backbone, and consists of α-(1→5)-linked Ara*f* with a tetra-arabinofuranoside and hexa-arabinofuranoside branch (Briken et al., 2004; Chatterjee et al., 1991). The capping of the terminal Ara*f* varies between mycobacterial species; fast-growing species like *M. smegmatis* terminate the arabinin with inositol phosphate caps (PILAM), the pathogenic mycobacteria including *M. tuberculosis* have one to three mannose as caps (ManLAM), and *Mycobacterium chelonae* has been found to lack any capping (AraLAM) (Briken et al., 2004). The differences in capping can influence pathogenesis; ManLAM is shown to inhibit dendritic cell function, better enabling *M. tuberculosis* to evade the immune system (Karakousis et al., 2004).

Extractable Lipids and Permeability

In addition to the covalently attached components, a number of non-covalently attached or extractable lipids have been found associated with the cell walls of mycobacteria. The most common types include lipooligosaccharides, phenolicglycolipids, glycopeptidolipids, sulfolipids, glycerophospholipids, and acetylated trehaloses, such as trehalose 6,6'-dimycolate (TDM), commonly known as cord factor (Brennan and Nikaido, 1995). These lipids associate with the mycolic acids of the mAG and largely make up the outer leaflet of the bilayer, although several

are now believed to associate with the mycolic acids in the inner leaflet as well (Ortalo-Magne et al., 1996). The less-packed lipids of this outer leaflet are more disordered than the closely packed inner mycolic acids, creating a gradient of decreasing fluidity (and permeability) from the outside to the inside of the cell wall. The high lipid content of the cell wall makes it virtually impermeable to hydrophilic compounds, although porins such as MspA allow the passage of small molecules such as glucose (Stahl et al., 2001). Surprisingly, the cell wall is significantly less permeable to hydrophobic compounds than would be predicted, due to the decreasing permeability towards the inner cell wall; however, hydrophobic compounds do traverse the cell wall considerably more easily than hydrophilic ones (Brennan and Nikaido, 1995). The impermeability of the cell wall renders mycobacteria resistant to many antibiotics and chemotherapeutics, contributing to the difficulty of treating TB. Additionally, mycobacteria are resistant to drying, alkaline conditions, and to many disinfectants, making it difficult to prevent transmission in densely populated areas (Brennan and Nikaido, 1995).

1.4   PEPTIDOGLYCAN HYDROLASES

The PG structure of bacterial cells is constantly being remodeled, and this requires a variety of enzymes that break down and recycle pre-existing PG as well as synthesize new PG. PG hydrolases are therefore vital to the regulation of growth of the PG sacculus, which occurs through the turnover and recycling of PG subunits and the elongation and separation of cells during division (Vollmer et al., 2008b). PG hydrolases also participate in the production of signaling molecules; remodeling needed to accommodate large molecular structures such as flagella and secretion systems; stimulation of growth as resuscitation promoting factors; and the changes associated with sporulation and germination in bacteria such as *B. subtilis* (Vollmer et al., 2008b).

The PG hydrolases are a large group of enzymes that cleave a subset of the bonds in PG. There are three types of PG hydrolases that target the PG (Figure 8): glycosidases cleave the bonds between the sugar residues, GlcNAc and MurNAc; N-acetyl-L-alanoyl-amidases (amidases) hydrolyze the bond between the glycans and the peptide bridges; and peptidases of varying specificities target the bonds in the peptide region.
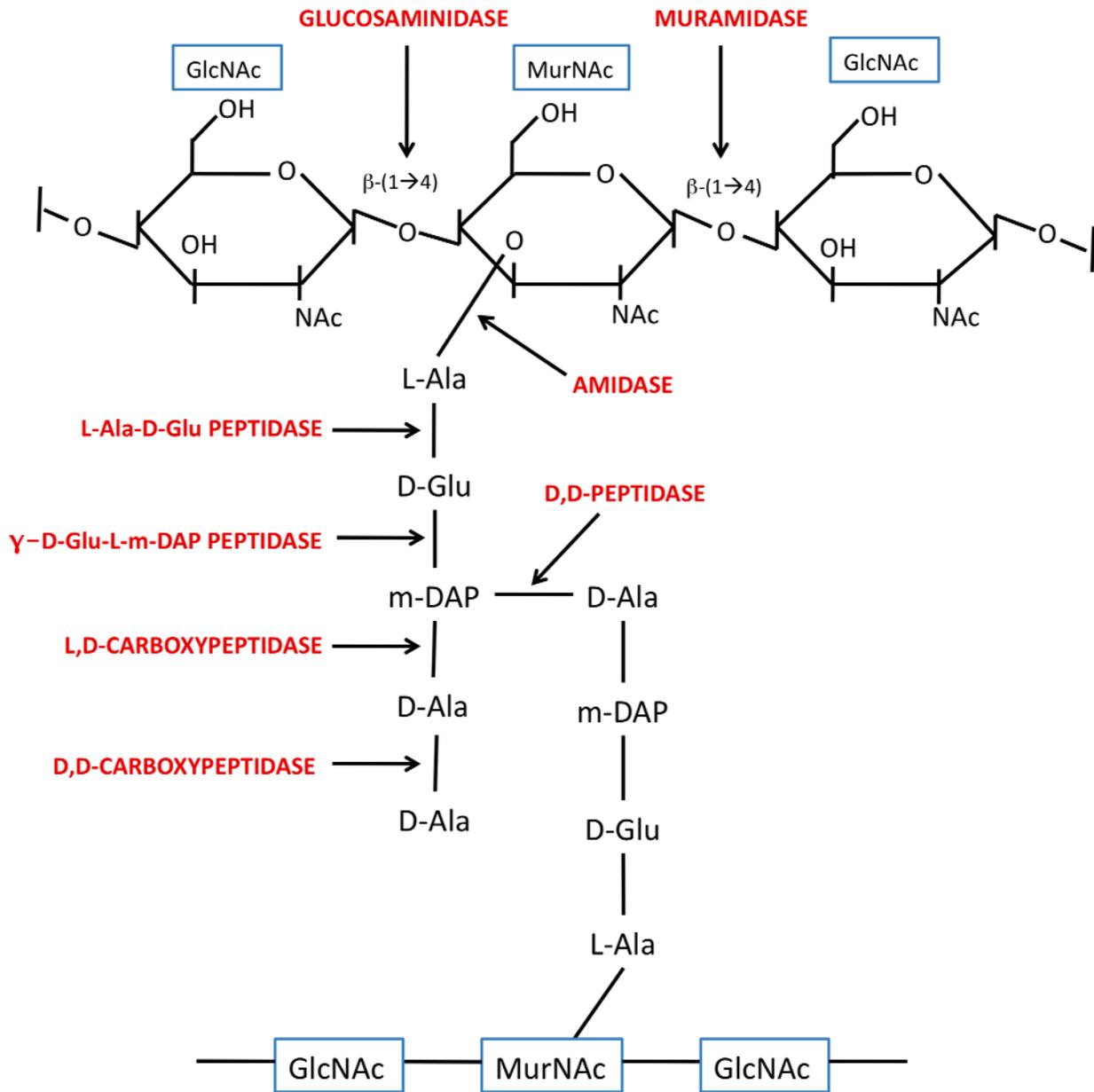
Figure 8. Target bonds of peptidoglycan hydrolases

Figure 8: The three types of PG hydrolase – glycosidase (glucosaminidase and muramidase), amidase, and peptidase – are listed in red with arrows indicating their target bonds in the PG. All hydrolases labeled as peptidases, excluding carboxypeptidases, are in fact endopeptidases.

1.4.1 Glycosidases

The glycan strand across all bacteria contains alternating β-(1→4)-linked *N*-acetylglucosamine (GlcNAc) and *N*-acetylmuramic acid (MurNAc) residues.   PG glycosidases target this β-(1→4) bond; β-*N*-acetylglucosaminidases (glucosaminidases) hydrolyze the GlcNAc-β-(1→4)-MurNAc bond, while β-*N*-acetylmuramidases (muramidases) cleave the MurNAc-β-(1→4)-GlcNAc bond (Figure 8). Glucosaminidases are widely seen in bacteria (Vollmer et al., 2008b), but are not found in many phages, while muramidases have been characterized in many phage lysins (Loessner, 2005). There are two types of muramidases defined by the enzymatic mechanisms that result in different end products (Vollmer et al., 2008b).

1.4.1.1 Lysozymes

The lysozymes are muramidases that hydrolyze the MurNAc-GlcNAc bond, resulting in a MurNAc residue with a reducing end (Figure 9). These enzymes are seen in phages, bacteria, fungi, invertebrates, and vertebrates and have been extensively studied (Jolles and Jolles, 1984). There are four characterized classes of lysozymes, three of which (phage T4 lysozyme or T4L, Hen egg-white lysozyme or HEWL, and goose egg-white lysozyme or GEWL) have a common structural fold containing the catalytic and substrate-binding sites and so appear to have diverged from a common ancestor (Weaver et al., 1984). The fourth class is based on the *Chalaropsis* lysozyme and has an entirely different tertiary structure (Rau et al., 2001). However, all four groups utilize the same acid/base catalytic mechanism, involving two Glu residues for the three similar classes and a Glu and Asp for the *Chalaropsis*-type lysozymes (Hermoso et al., 2003).

1.4.1.2 Lytic Transglycosylases

Lytic transglycosylases also target the MurNAc-GlcNAc bond, but instead of hydrolysis these enzymes cleave the bond and concomitantly form a 1,6-anhydro ring on the MurNAc residue (Figure 9) (Holtje et al., 1975). (While lytic transglycosylases do not operate with a hydrolysis reaction mechanism, I will be referring to the PG-modifying enzymes as a whole as PG hydrolases throughout this writing.) These enzymes possess a similar fold to the goose-type lysozymes (Thunnissen et al., 1995), but posses a single catalytic Glu in the active site, which results in the intramolecular ring formation instead of hydrolysis (Thunnissen et al., 1994). There are four different families of lytic transglycosylases, with the fourth considered to represent transglycosylases from bacteriophage, with the R lysin of phage λ regarded as the prototype (Blackburn and Clarke, 2001).
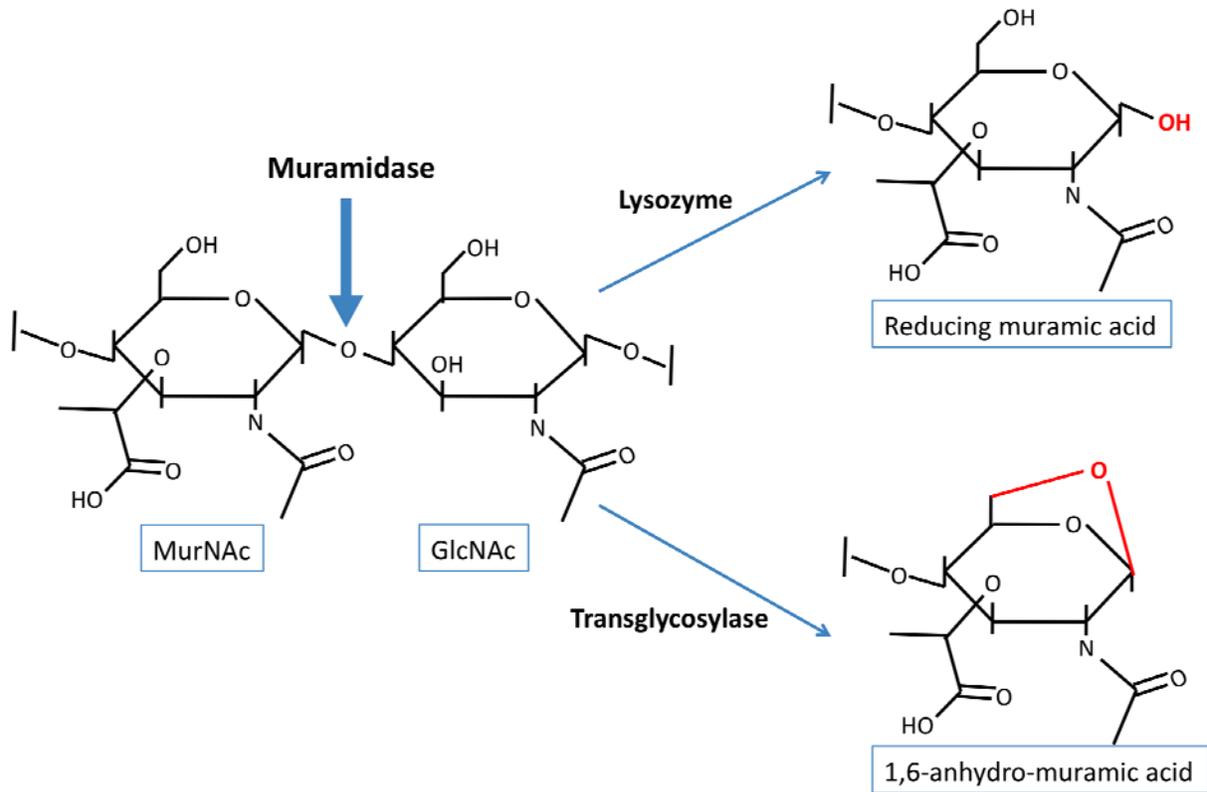
Figure 9. Lysozyme versus transglycosylase activity

Figure 9: Muramidases include two different enzyme types, both of which cleave the MurNAc-β-(1→4)-GlcNAc bond. Lysozymes (top) hydrolyze the bond, resulting in a MurNAc with a reducing end (red). Transglycosylases (bottom) cleave the bond and concomitantly form a 1,6-anhydro ring on the MurNAc (red).

1.4.2 Amidases

N-acetyl-muramoyl-L-Alanine amidases, or simply amidases, hydrolyze the D-lactoyl-L-Ala amide bond (Figure 8) (Vollmer et al., 2008b).   This severs the link between the glycan strands and the peptide bridges in the PG and is one of the most effective ways of degrading PG, as evidenced by its use in bacteria for PG remodeling during cell division (Heidrich et al., 2001) or cell wall recycling (Jacobs et al., 1994) and its widespread presence in the lysins of many phages (Vollmer et al., 2008b).   Many amidases are metallo-enzymes with Zn^{2+} as part of the catalytic mechanism, where it activates a water molecule that then acts as a nucleophile to attack the amide bond (Kerff et al., 2010).   The structures of the lysin of bacteriophage T7 (Cheng et al., 1994) and AmiD of *E. coli* (Kerff et al., 2010), among others, have revealed two His residues to be involved in the Zn^{2+} binding.   Structural similarities have also been observed between these amidases and Zn^{2+} peptidases such as thermolysin (Cheng et al., 1994; Kerff et al., 2010).

1.4.3 Peptidases

There are two primary classes of peptidases that hydrolyze PG: endopeptidases that cleave amide bonds within the peptide bridge, and carboxypeptidases that remove C-terminal amino acids (Holtje, 1995).   Carboxypeptidases have not been seen in phage lysins; the removal of a terminal amino acid would not cause the degradation of the PG polymer (Vollmer et al., 2008b). However, the diversity in amino acid composition of peptide cross-links provides many targets for a variety of endopeptidases (Figure 8).   Some endopeptidases are named for the amino acids between which they cleave, but given the variability in the peptide cross-link most

endopeptidases are instead identified simply by the isomers between which they cleave (Holtje, 1995; Smith et al., 2000). For example, L,D-endopeptidases hydrolyze the bond between the L-amino acid in position one and the D-amino acid in position 2, and D,D-endopeptidases hydrolyze the cross-links made between the D-Ala and m-DAP in most PG (Vollmer et al., 2008b); these are most commonly L-Ala-D-Glu-endopeptidases because L-Ala and D-Glu are the amino acids most commonly found in positions 1 and 2, respectively (Figure 2) (Vollmer et al., 2008a).

Peptidases from all organisms are being indexed in the MEROPS peptidase database, (http://merops.sanger.ac.uk), where they are grouped into families based on sequence similarity, and then into clans of families defined by the catalytic mechanism (Rawlings and Barrett, 1993; Rawlings et al., 2010). The clans are usually identified by the catalytic amino acid residue (*e.g.* Cysteine peptidases) (Rawlings and Barrett, 1993). There are also metallopeptidases, which use a metal ion such as Zn^{2+} as part of the catalytic mechanism. These have two absolutely conserved His residues, similar to Zn^{2+} amidases such as the T7 phage lysin (Cheng et al., 1994).

1.5   BACTERIOPHAGE LYSIS

At the end of an infection or upon the induction of a prophage that results in new virions, phage progeny must escape the host cell. While filamentous phage can extrude through the cell envelope with minimal harm to the cell (Russel, 1995), most phage release uses a lysis mechanism (Young et al., 2000). Lysis of a bacterium requires disrupting the cell wall in order to release the cytoplasmic contents—including progeny phage—into the environment to find new hosts. There are two known strategies for phage lysis (Young et al., 2000). For the simple ssDNA and ssRNA phages a single lysis gene is used (see "Other lysis systems"). The more-complex dsDNA phages employ a multigenic system that at minimum involves a holin and an endolysin.

The holin and lysin work together in two different capacities. Both are expressed in the late stages of infection and accumulate, either in the membrane (holin) or the cytoplasm (lysin). The holin forms lesions in the cell membrane that result in depolarization, thereby halting cellular respiration. Simultaneously, these lesions allow the lysin access to its PG substrate, which it degrades through one or two of several enzymatic activities. The cell wall is no longer able to withstand the internal osmotic pressure and the cell lyses. The triggering time by the holin is very precise and reliable; a synchronized infection can lyse a culture in seconds (Young et al., 2000). Prior to triggering, the lysin accumulates harmlessly in the cytoplasm.

The need for rapid lysis at the optimal time in phage production places great evolutionary pressure on holins to maintain a tight schedule specified by the rate of new phage assembly and maturation, the physiological state of the host, and the average time required for phage to find

41

and infect a new host (Wang, 2006). In lysing the cells too early the maximum production of phage progeny is not reached. However, while delaying lysis allows an increase in the number of new virions it must be balanced against the opportunity to infect new host cells (Young et al., 2000). In an environment of low host cell concentration, it is better to postpone lysis for a linear increase in virions as long as the cell is capable of producing them. In conditions of high host cell concentration there is a strong advantage in decreasing the latent period for faster lysis and new infections, which lead to an exponential increase in phage titer.

1.5.1   The Lysis Cassette

The region of the genome encoding the holin, lysin, and any associated lysis genes is referred to as the lysis cassette. The lysis genes can be identified through mutagenesis by specifically looking for a lysis-defective phenotype without a reduction in the accumulation of intracellular virions (Young, 1992). The majority of studies on lysis cassettes and holins have focused on lambdoid phages such as λ, P22, P1, and T4 (Young, 1992), while those of Gram-positive phages have received less attention; conversely, most research on lysins has focused on those of Gram-positive phages, which show therapeutic potential (Borysowski et al., 2006; Fischetti, 2008; Loessner, 2005; O'Flaherty et al., 2009). As illustrated with the λ lysis cassette (Figure 10), the holin (*S*) and lysin (*R*) genes are generally adjacent with the holin upstream and often overlapping the lysin, even if only by 1 bp (Young, 1992). There are some exceptions to this setup, notably the phage P1 *lydA* and *lydB* holin/anti-holin genes that are separated from the lysin-encoding gene *17* (Schmidt et al., 1996), and the T4 phage holin *t*, which is significantly distanced from the lysin encoded by *e* (Wang et al., 2000). *Rz/Rz1* (discussed below) are a pair of accessory lysis genes with orthologs in other Gram-negative phages (Summer et al., 2007) and

are located downstream of the holin and lysin (Young, 1992). This cluster of genes was designated as a cassette since it can be cloned and expressed to induce lysis apart from any other phage genes (Garrett et al., 1981). While both the holin and lysin are needed for lysis, they operate independently, as can be demonstrated by exchanging various holins and lysins of different phages (Garcia et al., 1997; Young et al., 2000). As a result, the lysis cassette is highly mosaic both between the holin and lysin genes and within the lysins themselves (Garcia et al., 1988; Lopez et al., 1997; Young et al., 2000).
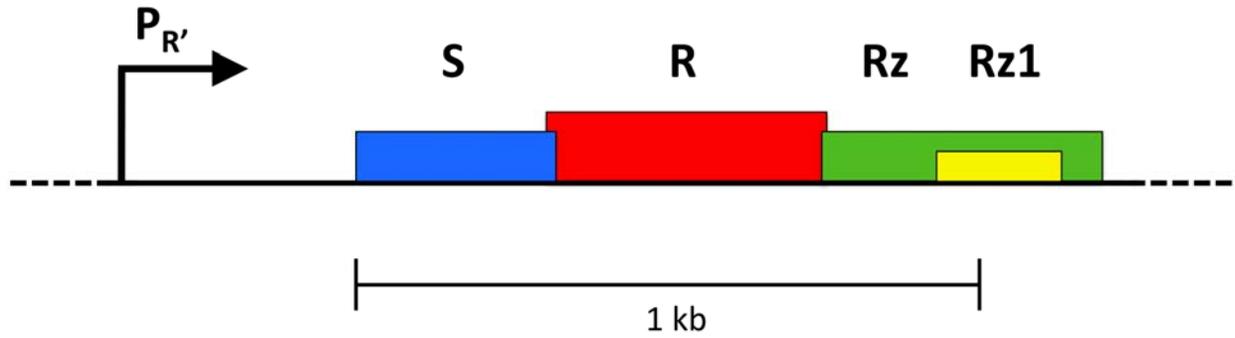
Figure 10. Lambdoid phage lysis cassette

Figure 10: The lysis cassettes of phages λ, P22, and 21 are the best studied and representative of many phages that infect Gram-negative bacteria. Four lysis ORFs are shown: S encodes the holin (blue), R encodes the endolysin (red), and Rz/Rz1 (green/yellow) encode accessory lysis proteins found in phages infecting Gram-negative bacteria (Summers *et al.* 2007). Lysis genes overlap slightly, except for *Rz1*, which is completely embedded in a different reading frame within *Rz*. Transcription begins at the late promoter $P_{R'}$.

1.5.2 Holins

Holins have been traditionally considered the "protein clock" of bacteriophage infections by regulating when to effect host lysis (Wang et al., 2000; Young et al., 2000). These proteins perform two functions: creating a lesion in the inner membrane and timing lysis for the optimal release of phage progeny.

Holins accumulate in the inner membrane throughout late gene expression. At a genetically-specified time the holins disrupt integrity of the membrane to quickly dissipate the membrane potential and activate lysin-mediated PG hydrolysis for swift lysis of the host (Figure 11). This is accomplished by the oligomerization of membrane-embedded holin proteins as shown by electron microscopy and other structural studies (Dewey et al., 2010; Pang et al., 2009; Savva et al., 2008). The resulting lesion depolarizes the membrane, which eliminates the proton motive force required for respiration and the overall integrity of the cell membrane. Consistent with the precise timing achieved by holins in order to maximize phage production, the membrane potential remains constant until seconds before the holins trigger lesion formation (Grundling et al., 2001). In this way holins accumulate in the membrane but do not compromise the biosynthetic capacity of the host to continue producing virions. The precise trigger for lysis is unknown, but may be tied to a critical concentration of holin in the membrane, precipitating a change to form homo-oligomeric lesions. This process that may be opposed by the energy state of the membrane (Grundling et al., 2001). Once one hole forms, the global membrane potential collapses, which would result in the formation of more lesions leading to rapid lysis. It has long been observed that energy poisons (chemicals that disrupt the membrane potential, such as

cyanide and dinitrophenol) can trigger lysis instantaneously (Young, 1992), theoretically by precipitating the formation of holin lesions.

Due to their small size of 60-185 aa (Loessner et al., 1999) and the dynamic evolutionary pressures, holins are a highly diverse family of functional homologs with low sequence similarity (Young et al., 2000). A single mutation can significantly affect lysis timing, giving the phage considerable temporal plasticity to adapt to changing environments and host conditions (Raab et al., 1986; Raab et al., 1988; Zheng et al., 2008b). The majority of current knowledge of holins comes from the study of coliphages such as λ and P21, which contain an *S* gene encoding the holin. Two broad classes of holins have been designated based primarily on predicted transmembrane domains and other structural analyses examining polar, charged, and hydrophobic regions (Wang et al., 2000). Class I holins as represented by S^{λ} are 95-183 aa and possess three transmembrane domains, while class II holins such as S^{21} average 65-95 aa and have two transmembrane domains (Young et al., 2000). A third unique class exists for T, the T4 holin, which is unusually large (218 aa) and highly hydrophilic with a single N-terminal transmembrane domain and the majority of the protein in the periplasm (Ramanculov and Young, 2001).

The canonical holin forms a channel in the cell membrane to grant the lysin access to its substrate (Figure 11 A). As seen by cryo-electron microscopy (Dewey et al., 2010), the S holin of λ forms exceptionally large (80 nm – 1.5 μm) channels that provide lysins access to the PG but that can allow the passage of proteins more than 450 kDa in size (Wang et al., 2003). In stark contrast, the holins of P1 and P21 form smaller "pinholins" (Figure 11 B) (Park et al., 2006). The hole formed by oligomers of S^{21} is only 1.6 nm, not wide enough to allow the passage of any proteins, much less lysins (Pang et al., 2009; Park et al., 2006). These pinholins

are paired with the SAR endolysins (signal-anchor-release, discussed below), which travel across the membrane prior to the triggering of lysis via the *E. coli sec* transport system (Xu et al., 2004). Since the lysin does not need the holin for access to its substrate, these holins are only responsible for achieving precise timing of lysis. There is a gradual spontaneous activation of SAR lysins, but the rapid depolarization of the membrane potential by the pinholin activates the accumulated SAR lysins simultaneously for rapid lysis (Park et al., 2006). As the pinholins are unable to provide a standard lysin access to the PG it is not possible to compliment a S^λ- mutant with S^{21} since there is no passage through the cell membrane for the R lysin. However, a S^{21}- mutant can be complemented with S^λ since a canonical holin can still activate the P21 SAR lysin by disrupting the membrane potential, even if it is not required for substrate access, thus demonstrating an advantage of canonical holins in their ability to function with SAR or cytoplasmic lysins (Park et al., 2006).
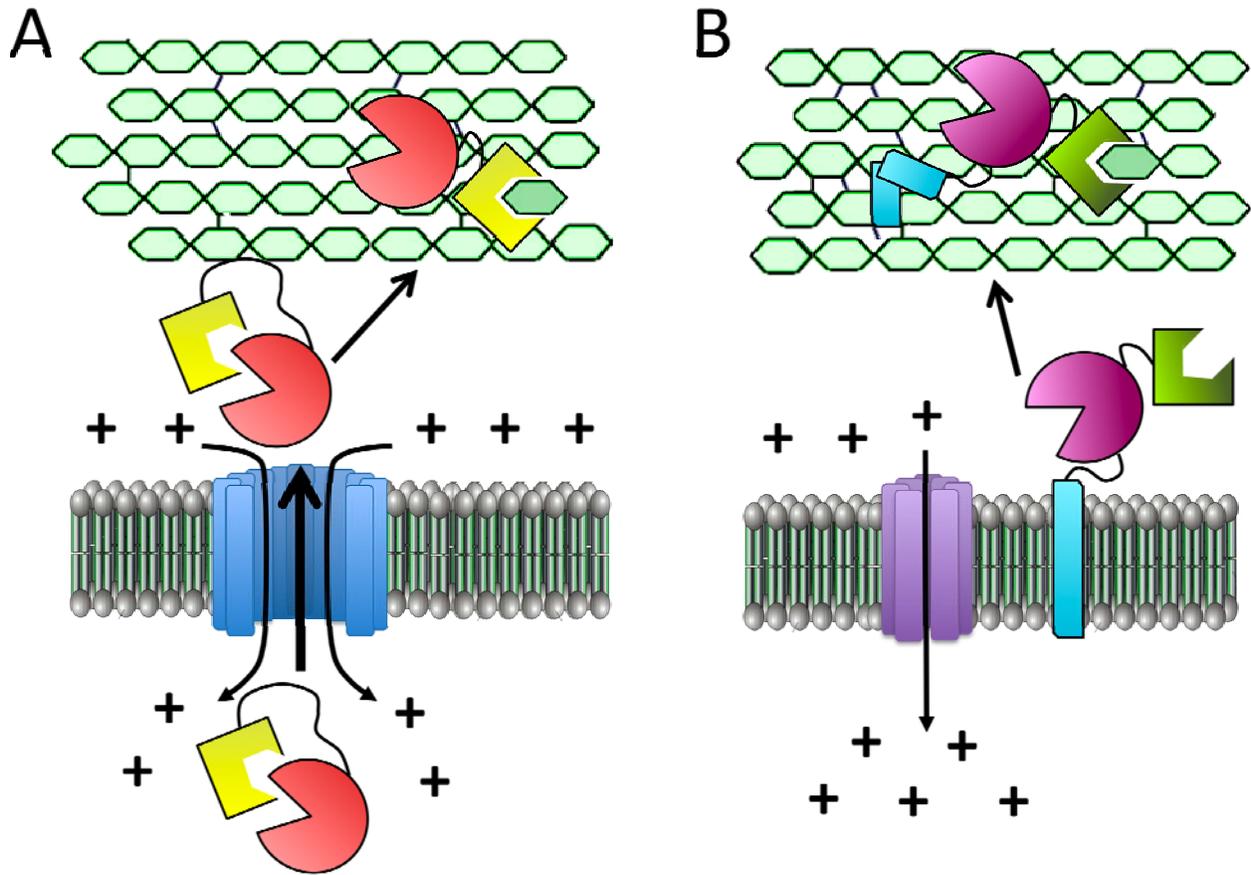
Figure 11. Holin and lysin mechanisms

Figure 11: The holins (blue, purple) form lesions that depolarize the membrane (represented by the movement of [+] across the membrane) and allow the lysins (red/yellow and pink/green) access to the substrate. **A.** Canonical holins form a large lesion (blue) in the cell membrane that depolarizes the proton motive force and allows the lysin (red/yellow) to access the PG. The lytic activity (red domain) may be inhibited until the cell wall-binding domain (yellow) has bound its substrate. **B.** Pinholins form tiny pores (purple) that can only depolarize the membrane. These are paired with signal-anchor-release or SAR-endolysins (pink/green), which are secreted into the periplasm with a transmembrane anchor (cyan) in the cell membrane. Once the pinholin depolarizes the membrane, the SAR endolysin is released from the membrane and can access the PG.

1.5.3 Endolysins

The endolysin, or lysin, works in conjunction with the holin in dsDNA bacteriophage to effect phage release. Lysins degrade the PG cell wall, and once this structure has been compromised the contents of the cell are rapidly expelled due to the high internal osmotic pressure. Importantly, lysins exhibit a narrow spectrum of lytic activity that is determined by the recognition by the lytic domain of unique PG linkages and by the cell wall binding domain (CBD) of exclusive components of the host cell wall (Borysowski et al., 2006).

1.5.3.1 Modular Structure

Lysins are modular proteins typically consisting of two domains, an N-terminal domain with lytic (PG-hydrolyzing) activity and a C-terminal cell wall binding domain (CBD) (Lopez et al., 1997). Lysins have domain-specific sequence similarities to other PG hydrolases (Loessner et al., 1997; Sheehan et al., 1997) and this was proposed to be the result of extensive recombination between phages and their hosts through processes like non-homologous recombination and horizontal gene transfer (Lopez et al., 1997). The Garcia lab has conducted a comprehensive investigation of lysin modularity. Their study of the cell wall hydrolases of *S. pneumoniae* and its phages demonstrates recombination on multiple levels (Figure 12). Several lytic and cell wall binding domains have been identified from the host autolysin and six different phage lysins (Lopez and Garcia, 2004). These include lysozyme, glucosaminidase, and amidase activities paired with either of two CBDs, creating four different two-domain organizations. Recombination events between lysins can be identified in the lysins of phages Cp-1 and Cp-7, which have highly similar lysozyme domains but completely unrelated CBDs (Figure 12). There is also recombination between phage and host; Cpl-1 and the host autolysin, LytA, have

strong C-terminal similarity but completely different N-terminal domains with predicted activities, lysozyme and amidase, respectively (Figure 12). Finally, the N-terminal domain of Pal is similar to the amidases of phages SM1 and BK5-T, which infect *Streptococcus mitis* and *Lactococcus lactis*, respectively. This illustrates that the Pal lysin is a chimeric enzyme of intergeneric origin (Sheehan et al., 1997).
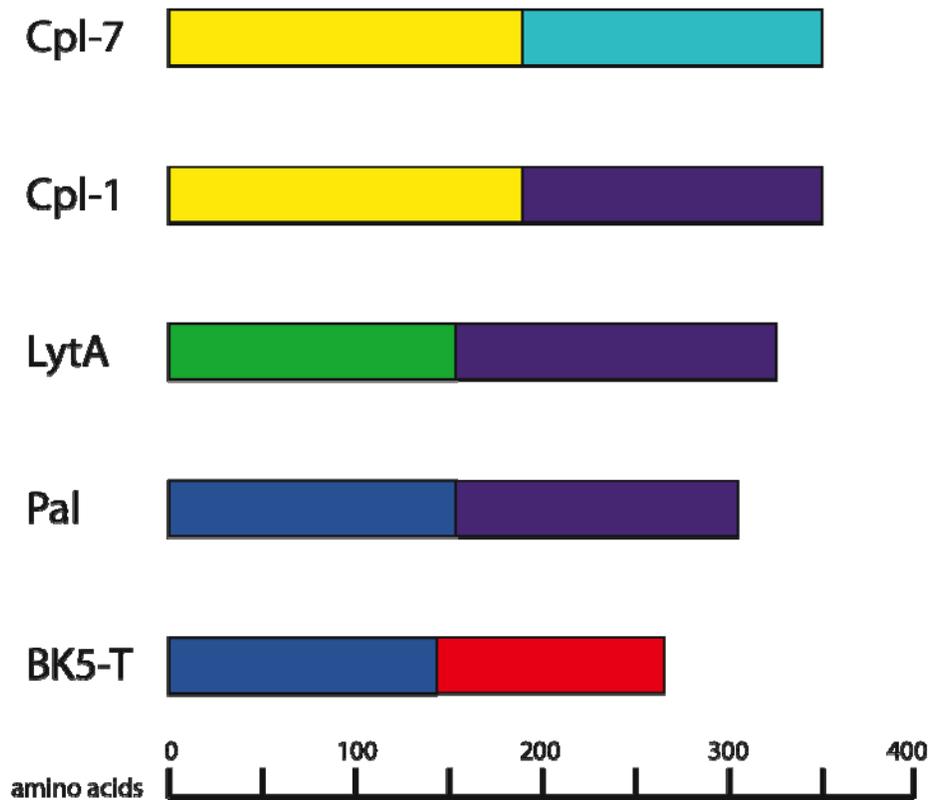
Figure 12. Evidence of recombination in *Streptococcus* phage lysins

Figure 12: Lopez and Garcia (2004) demonstrated the modular nature of phage lysins with examples of recombination between phages, between phage and host, and between phages infecting different genera of bacteria. Matching colors indicate domains with significant similarity and similar function. Cpl-7, Cpl-1, and Pal are lysins of phages that infect *Streptococcus*; LytA is an autolysin of *S. pneumoniae*; and the final lysin is from phage BK5-T, which infects *L. lactis*. The domains include muramidase (pfam01183, yellow); amidase_2 (pfam01510, green); amidase_5 (pfam05382, blue); Cpl-7 binding domain (cyan); choline-binding domain (purple); and BK5-T lysin binding domain.

The modular structure of lysins has also been experimentally determined by X-ray crystallography (Hermoso et al., 2003; Korndorfer et al., 2006), which consistently reveals a two-domain enzyme with the N- and C-terminal domains connected by a short linker. Further, deletion mutants have been used to show that many domains can function independently (Cheng and Fischetti, 2007; Loessner et al., 2002). Chimeric lysins have been experimentally generated that possess the properties of the individual domains used to construct them (Croux et al., 1993a, b; Diaz et al., 1991; Sanz et al., 1996). In one instance a chimeric lysin with two lytic domains of differing activities paired with a single CBD was created and the specific functions of each domain were retained (Sanz et al., 1996). This bi-functional organization was later observed in several natural phage lysins (below), although it does not appear to be common. It should also be noted that the majority of lysins studied are from Gram-positive phages; lysins from phages infecting Gram-negative species, such as the λ R transglycosylase and the T7 lysozyme, tend not to be modular but have a more globular structure (Bienkowska-Szewczyk et al., 1981; Cheng et al., 1994). However, two lysins have recently been identified in phages infecting *Pseudomonas aeruginosa* that display a modular structure with an N-terminal binding domain and C-terminal transglycosylase activity (Briers et al., 2007).

1.5.3.2 Lytic Activity

Phage lysins are dependent on the permeabilization of the cell membrane by a holin to degrade the cell wall. If lysin is expressed in the absence of any holin, the addition of CH_3Cl (which artificially permeabilizes the membrane) will instantly induce lysis (Young, 1992). The most basic arrangement includes a canonical holin like S^λ (discussed above) forming a large lesion in the cell membrane that allows passage of lysins and access to their PG substrate while simultaneously dissipating the membrane potential of the cell (Figure 11 A). However, a second

method has recently been identified in which the holin is not responsible for the lysin to access the PG but is still required for efficient lysis. An N-terminal hydrophobic domain that could serve as an export signal sequence was first seen in the fOg44 phage infecting the Gram-positive *Oenococcus oeni* (Sao-Jose et al., 2000) and then analyzed in *E. coli* phage P1 (Xu et al., 2005; Xu et al., 2004). With this sequence the P1 Lyz lysin could be transported into the periplasm using the host *sec* system, where it then remained tethered via a transmembrane domain to the cell membrane. Once the membrane was depolarized—by a holin, CH_3Cl, or an energy poison—the Lyz was released to degrade the PG. This N-terminal region was termed a signal-anchor-release (SAR) domain and has since been identified in many lysins, including coliphages P21, Mu, and T1 (Xu et al., 2004). SAR lysins are often paired with pinholins (Figure 11 B), which are only capable of depolarizing the membrane (Pang et al., 2009), but they can be used with canonical holins as well. However there are several lysins for which the associated holin is not identified.

All four PG-hydrolyzing activities—amidase, muramidase, glucosaminidase, and peptidase—have been identified in lysins, although amidase and lysozyme activities are the most common (Loessner, 2005). The amide bond connecting the MurNAc to the L-Ala of the peptide is the most consistently present link in bacterial PG, which can differ in the nature of the peptide bridges or modifications of the glycan chains (Schleifer and Kandler, 1972). Certain peptide bridge-specific peptidase activities are specifically adapted to the PG of a host, introducing one method of ensuring specificity in addition to the CBD (discussed below). For example, L-alanoyl-D-glutamate peptidase activity has thus far only been found in *Listeria* phage (Loessner, 2005). Lastly, while lysozyme-type muramidases that hydrolyze the MurNAc—GlcNAc bond are common, transglycosylases appear to be very rare. Transglycosylase activity is the only lytic

activity that is not technically hydrolysis, since the reaction forms a 1,6-anhydrobond in MurNAc. The transglycosylase motifs are actually very common in phage, but they are found in the tail proteins and believed to assist in entry during phage infection (L. Marinelli and G. Hatfull, manuscript in preparation, Moak and Molineux, 2000, 2004). The R lysin of phage λ and KZ144 and EL188 of the *P. aeruginosa* phages, φKZ and EL are transglycosylases (Bienkowska-Szewczyk et al., 1981; Briers et al., 2007), but these are from phages infecting Gram-negative bacteria. To my knowledge, no Gram-positive phage transglycosylase lysins have been characterized.

Several bifunctional lysins have been identified that have two lytic domains, similar to those previously created experimentally. Intriguingly, the N-terminal domain for these lysins is an endo-peptidase, usually a D-alanyl-glycyl endo-peptidase (CHAP) (Bateman and Rawlings, 2003), and all are from phages infecting *Staphylococcus aureus* or *Streptococcus agalactiae*. Of those infecting *S. aureus*, LysK of phage K, LysH4 of φH5, and the lysin of φ11 have CHAP and amidase activities as well as an SH3b CBD (Becker et al., 2009; Horgan et al., 2009; Navarre et al., 1999; O'Flaherty et al., 2005; Obeso et al., 2008). PlyGBS of the *S. agalactiae*-infecting phage B30 possesses CHAP peptidase and lysozyme domains with an SH3b CBD (Cheng and Fischetti, 2007; Donovan et al., 2006a; Pritchard et al., 2004). The strangest of the set is the lysin of LambdaSa2, which infects *S. agalactiae*. This lysin has an N-terminal D-γ-Gln-L-Lys endo-peptidase and a C-terminal glucosaminidase domain but in between these domains are two Cpl-7 binding domains, motifs that bind choline first identified in the *S. pneumococcal* phage Cp-7 lysin (Donovan and Foster-Frey, 2008; Pritchard et al., 2007). The separate domains of φ11 lysin, LysK, and PlyGBS have activity *in vitro* as determined by reaction products and site-directed mutagenesis studies (Donovan et al., 2006a). However, when truncation and deletion

mutants were made and activity was tested on live cells, for every single bifunctional lysin the peptidase domain was responsible for all, or nearly all, of the activity, and the isolated central domain held little or no activity (Becker et al., 2009; Donovan and Foster-Frey, 2008; Donovan et al., 2006a; Horgan et al., 2009). The presence or absence of the CBD varied in its effect on the activity (discussed below).

1.5.3.3 Cell Wall Binding Domains

The lytic domain of a lysin is likely to be active on a variety of bacteria due to the common structure of PG; however, lysins are highly specific, occasionally even to the level of certain serovars of a host species (Loessner et al., 2002). The CBD binds cell wall components that are often unique to the host, such as the teichoic acids found in some *Streptococcus* species. The CBD may inhibit lysin activity until it binds this substrate, and in many instances the CBD enhances the lytic activity. Often the lytic domain simply lacks sufficient substrate binding affinity even in proximity to its substrate (Loessner, 2005). In contrast, the affinity of the binding domains of Ply118 and Ply500 were comparable to affinity-matured antibacterial antibodies (Loessner et al., 2002). In addition to swapping CBDs to create chimeric lysins with changing host specificities, CBD deletion studies have also been completed. Oddly, the effects on the activity of a lysin with or without the CBD vary widely. The loss of the CBD decreases activity for *Streptococcus* phage lysins Cpl-1, Cpl-7, and LambdaSa2, *Listeria* phage lysins Ply118 and Ply500, and *Bacillus* phage lysin PlyG (Donovan and Foster-Frey, 2008; Garcia et al., 1990; Kikkawa et al., 2008; Loessner et al., 2002). However, the lytic activity was unchanged or even increased when the CBD was deleted for *Listeria* phage lysin Ply511*, Bacillus* phage lysins PlyL and Ply21, *Streptococcus* phage lysin PlyGBS and B30 lysin, and *Staphylococcus* phage lysin LysK (Becker et al., 2009; Cheng and Fischetti, 2007; Donovan et

al., 2006a; Gaeng et al., 2000; Horgan et al., 2009; Low et al., 2005). It is interesting to note that several of the lysins that lose activity contain CBD that are highly specific to their host bacteria, namely the choline-binding motifs for Cpl-1, teichoic acid-binding motifs for Cpl-7 and LambdaSa2 (Lopez and Garcia, 2004), the *Listeria*-specific carbohydrate for Ply118 and Ply500 (Loessner et al., 2002), and the unknown *Bacillus cereus*-specific substrate for PlyG (Kikkawa et al., 2008). Further, PlyGBS, B30 lysin, and LysK all share an SH3b CBD, a binding domain in bacteria that is homologous to eukaryotic SH3 domains (Whisstock and Lesk, 1999).

In addition to enhancing lysin activity, the CBD restricts the range of susceptible bacteria by inhibiting activity in the absence of a cell wall binding substrate that is specific to the host (Hermoso et al., 2003; Korndorfer et al., 2006). The inhibitory effect is largely mediated by the CBD allosterically binding and suppressing the lytic activity, as has been demonstrated in the *Bacillus anthracis* prophage Ply21 (Low et al., 2005). These cell wall binding substrates are often molecules unique to the host. Pneumococcal lysins such as Cpl-1 have evolved specific motifs to bind choline, a component of the teichoic acids of *S. pneumoniae* (Hermoso et al., 2003; Lopez et al., 1997). These lysins are dependent on choline-binding for activity, while other lysins with choline-independent CBDs such as CPL-7, can lyse *S. pneumoniae* with ethanolamine in place of choline. Studies with pneumococcal lysins have also found host-specificity to be a transferable property. Fusing the catalytic domain of the LYC autolysin from *Clostridium acetobutylicum* to the choline-specific CBD of CPL1 created a protein capable of lysing *S. pneumoniae* but not *C. acetobutylicum* (Croux et al., 1993a). These and other studies associate host-specificity of lytic activity primarily—but not solely—with the CBD.

1.5.3.4 Lysins with additional activities

While not common, several lysins are known or suspected of possessing additional activities beyond the degradation of PG. Most notorious and thus far unique is the bifunctional T7 lysozyme (technically an amidase) that also plays a role in transcriptional regulation by binding to the T7 RNA polymerase (Moffatt and Studier, 1987); oddly, the T7 lysin is not modular but appears to be a globular protein that performs both of these functions (Cheng et al., 1994). Two other lysins, namely the T4 lysozyme (During et al., 1999) and the lysin of a *Bacillus amyloliquefaciens* phage (Morita et al., 2001b), appear to have membrane-permeabilization abilities. Both have C-terminal sequences that are similar to cationic antimicrobial peptides known to disrupt the cell membrane (Epand and Vogel, 1999). Interestingly, the *B. amyloliquefaciens* phage lysin was able to lyse cells without a holin and showed some microbicidal activity against *E. coli* when applied externally (Morita et al., 2001a; Morita et al., 2001b; Orito et al., 2004), lending hope to the idea of phage lysin therapy against Gram-negative bacteria.

1.5.4   Other Lysis Proteins

While the minimum lysis cassette of dsDNA phages requires a holin and a lysin (Young et al., 2000), additional lysis proteins have been identified in many lysis cassettes (Catalao et al., 2010; Gil et al., 2008; Payne et al., 2009; Summer et al., 2007). For example, nearly all Gram-negative phages appear to contain one or two additional lysis proteins that address the barrier of the outer membrane (Summer et al., 2007); these are never found in a Gram-positive phage. Deletion of these and other accessory lysis proteins results in a mild lysis phenotype that can be observed as differences in plaque morphology or in the timing of lysis, but this does not completely prevent

the phages from forming plaques (Catalao et al., 2010; Payne et al., 2009; Young, 2002; Zhang and Young, 1999). In addition, research on lysis has primarily focused on dsDNA phages, but the mechanism of lysis by a few single-stranded (ss) nucleotide phages have been characterized (Bernhardt et al., 2000; Bernhardt et al., 2001b). These phages employ single-gene lysis systems which target PG synthesis enzymes instead of the PG structure itself.

1.5.4.1 Anti-holins

Many holin systems also include an inhibitor, or anti-holin, that aids in the regulation of the timing of lysis. A common feature in S^λ, S^{21}, and other holins of both classes is a dual-start motif that creates two nearly-identical proteins with opposing functions from a single gene by utilizing alternate start codons (Barenboim et al., 1999; Blasi and Young, 1996). In S^λ, translation of the S gene from the first or third start codons produces $S^\lambda 107$ (Met-1—Lys-2—Met-3…) or $S^\lambda 105$ (Met-3…), respectively. $S^\lambda 105$ is the effector form of the holin that creates the actual lesions, while $S^\lambda 107$ is an anti-holin. The production of $S^\lambda 105$ or $S^\lambda 107$ is determined by RNA secondary structure and the relative ratios influence the timing of lysis, with a greater proportion of $S^\lambda 105$ accelerating lysis onset (Wang et al., 2000). However, not all anti-holins are encoded within the holin gene. Many are encoded by separate genes, as is the case with the phage P1 *lydA* and *lydB* holin/anti-holin (Schmidt et al., 1996).   Also unique is the method of inhibition in T4 phage. A situation known as lysis inhibition (LIN) occurs when there is a secondary infection of the cell by another T4 phage (Doermann, 1948). This event overrides the lysis schedule programmed into T and virions continue to be made. Lysis can be delayed for hours if there continues to be secondary infections. LIN is accomplished by the periplasmic RI protein, which acts as an anti-holin to directly inhibit T (Ramanculov and Young, 2001). The LIN phenomenon demonstrates the evolved ability of T4 to respond to environmental signals and

adjust its infection accordingly; secondary infections would indicate a lack of host cells, and so the phage chooses to postpone lysis.

1.5.4.2 Rz/Rz1 and Spanins

In addition to the S holin and R lysin, two additional proteins, Rz/Rz1, are present in the λ lysis cassette whose functions were only recently elucidated (Berry et al., 2008). *Rz* is directly after the *R* lysin gene and codes for a 153 aa protein. Fully embedded in the *Rz* gene in the +1 reading frame is *Rz1* which encodes a smaller 60 aa protein (Figure 10) (Hanych et al., 1993). These proteins appear to be nonessential for lysis as nonsense mutations in the gene confer a conditional lysis phenotype only in the presence of >10 mM Mg^{2+} (Young et al., 1979). Rz is a predicted transmembrane protein (Summer et al., 2007) and Rz1 is a lipoprotein that localizes to the outer membrane (Kedzierska et al., 1996) and has been demonstrated to fuse liposome membranes (Bryl et al., 2000). Further, there is evidence for an interaction between Rz and Rz1 proteins; studies with homologs encoded by *18.5* and *18.7* in T7 indicated a protein interaction (Bartel et al., 1996), and complementations of Rz/Rz1 with λ, P22, and PRD1 require that the cognate pair be used (Krupovic et al., 2008; Markov et al., 2004).

    λ Rz/Rz1 and equivalents (hereafter known collectively as Rz or Rz1) have been predicted in over 120 phages infecting Gram-negative bacteria (Summer et al., 2007), but none are seen in Gram-positive bacteria. The homologs were identified based on similarity by sequence or secondary structure to known Rz and Rz1 proteins such as those in λ, P2, and T7. Specifically, the Rz protein needed to include a single N-terminal transmembrane domain and the Rz1 protein was predicted to be an outer membrane lipoprotein. From the search, three architectural genetic classes were determined: embedded, where the *Rz1* is entirely within *Rz* (λ, T7); overlapped, where *Rz1* extends past the end of *Rz* (P2); and separated *Rz* and *Rz1* genes (T4)

59

(Summer et al., 2007).   Among some genomes lacking Rz/Rz1 proteins, a single protein was predicted following the holin and lysin that resembled a fusion of the Rz and Rz1 structures and is predicted to span the entire periplasmic space, earning it the name "spanin" (Summer et al., 2007).   The spanin of T1 has been shown to complement a λ Rz⁻Rz1⁻ mutant, confirming it as a functional analog of Rz/Rz1 (Summer et al., 2007).

The current model for the involvement of Rz/Rz1 places it at the final step in the phage infection cycle (Berry et al., 2008). Lysis begins with the permeabilization of the cell membrane by the holin, followed by lysin-dependent degradation of the PG layer.   For Gram-negative bacteria, this still leaves the outer membrane as a barrier to phage release.   The Rz/Rz1 (or spanin) proteins fuse the inner and outer membranes to effectively remove all barriers to phage release (Berry et al., 2008).   While the Rz/Rz1 proteins function as a cognate pair, they are mechanistically independent of the holin and lysin (Berry et al., 2008), which allows for recombination that leads to the mosaicism observed in phage lysis cassettes.

1.5.4.3 Non-peptidoglycan Hydrolytic Lysis Methods

All dsDNA bacteriophage employ the holin-endolysin system to lyse host cells at the end of infection.   However, phages with smaller single-stranded (ss) nucleic acid genomes including *Microviridae*, *Leviviridae*, and *Alloleviridae* only require a single protein to effect lysis (Bernhardt et al., 2002).   The timing of lysis in these single gene systems is dependent upon the level of gene expression, the activity of the protein, and the rate of cell division (Bull et al., 2004).

The *Microvirus* φX174 has a ssDNA genome of only 5386 bp encoding 11 genes.   The lysin gene *E*, is fully embedded in the +1 reading frame of gene *D* and creates a 91 residue protein with an N-terminal hydrophobic domain and positively charged C-terminus (Young et

al., 2000). The mechanism of action of E was recently determined through genetic, biochemical, and structural studies (Bernhardt et al., 2000; Bernhardt et al., 2001a; Zheng et al., 2008a; Zheng et al., 2009). E blocks cell wall synthesis by inhibiting MraY (Bernhardt et al., 2001a), a translocase that catalyzes the formation of the first lipid-linked precursor in cell wall synthesis (Ikeda et al., 1991). By targeting a host protein, resistant strains with mutations in MraY can be isolated (Bernhardt et al., 2001b), an event that is not observed in holin-lysin systems, which target the PG structure itself (Loeffler et al., 2001; Schuch et al., 2002). Similar to cell wall synthesis inhibitors like the antibiotics penicillin, E is dependent on cell growth and division, and E-mediated lysis is most commonly observed along the septum between newly dividing cells where new cell wall is being incorporated (Witte et al., 1998). Consistent with the dependence of the phage on active cell growth, the addition of a protein synthesis inhibitor or energy poison stops lysis; in contrast, in holin-based systems the depolarization of the membrane by an energy poison will trigger lysis (Young, 1992).

The E protein is capable of replacing the lysis cassette of λ with minor plaque phenotypes (Zheng et al., 2008b). However, the latent period is determined by the quantity of E protein, whereas the latent period for a holin is almost completely dependent on the specific allele. This may indicate that E has lesser control over the lysis of its host in comparison to holins, which have increased plasticity with the ability to adjust the timing of lysis through small mutations and thus adapt to changing environments (Zheng et al., 2008b).

Two ssRNA phages also have alternative mechanisms of lysis (Bernhardt et al., 2002). The ssRNA *Allolevivirus* Qβ contains only 4 genes in a 4217 nucleotide genome, none of which are a lysis protein (Bernhardt et al., 2001b). Instead, the capsid maturation protein A_2 also fulfills the role of lysis, in addition to its functions in adsorption to the host sex pilus and the

protection of the virion genome from ribonucleases. Resistant host mutants could also be recovered from Qβ infection and were found to have mutations in the *murA* gene, which catalyzes the first committed step in PG biosynthesis (Bernhardt et al., 2001b). The virion-associated A_2 binds directly to MurA effectively titrating it from the cell wall synthesis pathway leading to an induction of lysis. However, another study found a frame-shifted protein near the replicase in Qβ that appears to cause rapid lysis in high sodium chloride environments (Nishihara et al., 2004), so the circumstances of lysis in Qβ remain unresolved. MS2, a 3569 nucleotide *Levivirus* phage, possesses a lysis gene, *L*, embedded out of frame in the coat and replicase genes (Bernhardt et al., 2002). L is a transmembrane protein with a hydrophilic N-terminus and hydrophobic C-terminus, and requires a high level of synthesis to effect a gradual lysis (Bernhardt et al., 2002). Currently the mechanism of lysis by L is unknown.

1.5.5   Mycobacteriophage Lysins

Isolation and characterization of mycobacteriophages began in the 1940s (Froman et al., 1954; Gardner and Weiser, 1947; Whittaker, 1950) with a focus on typing clinical isolates of mycobacterial pathogens (Jones, 1975; Snider et al., 1984) and, later, as tools for genetic manipulation of mycobacteria (Jacobs, 2000; Jacobs et al., 1989; Snapper et al., 1990; van Kessel and Hatfull, 2007). As of this writing, 60 mycobacteriophage genomes have been characterized (Hatfull et al., 2010); the majority of these phage sequences are the product of the Phagehunter educational research program conducted in the Hatfull lab (Hatfull et al., 2010; Pedulla et al., 2003). To the extent that studies have identified putative lysis genes, they have done so based on homology to known PG hydrolases.

Mycobacteriophage proteins are grouped into "phamilies" corresponding to specific levels of sequence similarity (Hatfull et al., 2008), and there are currently more than 1500 such phamilies (Hatfull et al., 2010). Pham7, the phamily comprising LysAs, is one of only three phamilies with a representative from every phage genome. There is, however, no single sequence element common to all Pham7 members; indeed, there is extensive mosaicism throughout Pham7 (Hatfull et al., 2010; Hatfull et al., 2006; Pedulla et al., 2003).

One of the first observations made about the mycobacteriophage endolysins regarded the modularity of the proteins. Mediavilla *et al.* (2000) noted that the central regions of Bxb1 gp8 and D29 gp10 were 59% identical, while sequences within the same ORF to either side were unrelated. This central region shows weak homology to chitinase family proteins, a hydrolytic activity never before seen in phage lysins. Further, the C-termini of Bxb1 gp8 and D29 gp10 did not show significant similarity, but instead a 100 aa region in the C-terminus of Bxb1 gp8 had 39% identity with the C-terminus of TM4 gp29 (Mediavilla et al., 2000). The interchanging of regions to generate variation in the phage genome – even within a single gene – is consistent with the mosaic nature of phage genomes and their evolution by illegitimate recombination (Lawrence et al., 2002).

Focused study of mycobacteriophage lysins began with the unsequenced mycobacteriophage Ms6 (Garcia et al., 2002). Garcia *et al*. (2002) identified a lysis cassette that contained a set of 5 co-transcribed ORFs in different overlapping reading frames (*e.g.* the TGA stop codon of ORF2 overlapped the ATG start codon of ORF3). ORF2 was identified as the LysA based on sequence similarity to known lysins with amidase activity, and on the lysis of *E. coli* expressing the Ms6 *lysA* upon the addition of chloroform to permeabilize the membrane (Garcia et al., 2002). ORF4 is a holin; it shares similarity with a holin of the *L. lactis* phage r1t,

was predicted to have two transmembrane domains similar to class II holins, and was able to complement a λ *S* mutant (Garcia et al., 2002). The functions of ORF1, ORF3, and ORF5 were less clear; all three had homologs in other mycobacteriophages but showed little similarity outside of these phages, leading to the theory that they had mycobacteria-specific functions in infection (Garcia et al., 2002). In addition to the lytic activity shown by Ms6 ORF2 (now Ms6 gp2), the Giles LysA was determined to be an essential gene for plaque formation (Marinelli et al., 2008), which is consistent with its vital role in lysis. Very recently, the TM4 gp29 LysA has been cloned and shown to have PG hydrolytic activity on zymograms and by chloroform-induced lysis of *E. coli* expressing TM4 gp29 (Henry et al., 2010). However, these analyses have been restricted to only a few select LysA proteins, and no characterization of the diversity preliminarily observed in these proteins has been conducted.

Beyond attaining its label based on its location downstream of the Ms6 *lysA* gene, little study has been conducted on LysB until very recently. The occurrence of lysis proteins in addition to endolysins and holins is not unusual, but the LysB proteins did not resemble antiholins, Rz/Rz1 or spanins, or any other proteins previously associated with lysis. Pham9 comprises the LysB proteins (Hatfull et al., 2006), which are found in 56 of the 60 published phage genomes (Hatfull et al., 2010). The *lysB* gene is always close to *lysA*, despite the difference in the location of the lysis cassette in the broader mycobacteriophage genome organization. The lysis genes generally follow those involved in virion structure and assembly, but in Cluster A mycobacteriophages the lysis cassette preceeds the structural genes (Payne et al., 2009); however, regardless of location in the genome, *lysB* remains linked with *lysA*. The LysB homolog in Ms6, gp3 (previously Ms6 ORF3), has been identified as a lipolytic enzyme based on its conserved signature motif Gly-X-Ser-X-Gly and the hydrolysis of lipase and

esterase substrates, including Tween-20, Tween-80, tributyrin and various *p*-nitrophenol fatty acid esters (Gil et al., 2008). However, the natural substrate and the purpose of the LysB proteins in mycobacteriophage infection were not determined.

Very recently, the role of Ms6 ORF1, now gp1, has been identified to be that of a chaperone for the LysA, gp2 (Catalao et al., 2010). Portrayed as playing a role similar to the SAR domain of other endolysins, gp1 is required for sec-mediated export of the gp3 LysA into the periplasm (Catalao et al., 2010). Gp1 shares structural characteristics with type III secretion system chaperones and formed a dimer to bind gp3 and translocate it across the membrane. The N-terminal 60 amino acids of the gp3 LysA were necessary and sufficient for binding and export. Similar to the pinholins and SAR-endolysins (Pang et al., 2009), a holin was not necessary for access to the PG substrate, and lysis could be induced simply by disrupting the membrane potential (Catalao et al., 2010). A Ms6 *gp1* deletion mutant was able to form plaques with a lesser efficiency and appeared to have a reduced burst size, although the cause of the latter was not determined (Catalao et al., 2010). Homologs of gp1 are found in 18 out of 60 phage genomes, primarily in Sub-cluster F1 phages (Hatfull et al., 2010). The identification of gp1 as an ancillary lysis protein playing a novel role as a lysin chaperone in mycobacteriophage further expands the known mechanisms of phage lysis.

1.5.6   Lysin Applications

Phage therapy has been pursued since the late 1800's, although Felix d'Herelle first described the bacteriophages and identified their viral nature in 1917 (D'Herelle, 2007), and subsequently attempted to use phage in the treatment of dysentery. A review of the history of phage therapy can be found in the review by Summers, W.C. (2001), and O'Flaherty *et al.* (2009) have written

an excellent overview of more recent advances.   Lysins have only recently been examined as potential antimicrobials.   While the lytic ability of isolated lysins was recognized in 1957 (Krause, 1957), it was not until 2001 that the therapeutic potential of lysins was first demonstrated (Nelson et al., 2001).

1.5.6.1 Therapeutic Potential

Lysins are being tested in two therapeutic capacities: the prevention of colonization by pathogenic bacteria and the elimination of infection.   Both of these functions were shown in the first therapeutic application of a lysin from the *Streptococcus* phage C1, which infects groups A, C, and E *Streptococcus* species (Nelson et al., 2001).   The purified lysin was able to rapidly sterilize bacteria in culture, protect the oral cavity of mice from colonization with a single dose, and eliminate all detectable bacteria in previously colonized mice.   This amazing demonstration combined with more than a decade of lysin study led many groups to examine the activity of other phage lysins on other pathogenic bacteria.   To date, the majority of these studies have targeted *S. pneumoniae* (Entenza et al., 2005; Grandgirard et al., 2008; Jado et al., 2003; Loeffler et al., 2003; Loeffler and Fischetti, 2003; Loeffler et al., 2001; McCullers et al., 2007), Group B streptococcus (Cheng et al., 2005), *S. aureus* (Rashel et al., 2007), *Enterococcus faecalis and faecium* (Yoong et al., 2004), and *B. anthracis* (Schuch et al., 2002).   These studies are reviewed in depth by Borysowski *et al.* (2005).

The prevention of colonization is important because it is the starting point for an infection and contributes to the spread of the disease in the community (Bogaert et al., 2004; von Eiff et al., 2001)).   The first lysin study aimed to prevent colonization of Group A streptococci in the upper respiratory system of mice (Nelson et al., 2001).   Nelson *et al.* used the C1 lysin as a prophylactic agent that was applied to a mucosal surface, namely the oral cavity, to great effect.

Another example is the lysin PlyGBS that is active against Group B streptococci (Cheng et al., 2005), a pathogen that can colonize human genitalia. One goal is to develop PlyGBS as a potential prophylactic against infection by Group B streptococci in newborns, whose mothers may be vaginally colonized during pregnancy. In addition to its lytic ability, PlyGBS is highly specific and does not lyse commensal bacteria (Cheng et al., 2005). The narrow spectrum lytic activity is one of several advantages that lysins have over broad-spectrum antibiotics.

Treating bacterial infections with lysins poses more difficulty because of the many host interactions possible (see below). Still, lysins have proven effective in treating infections of pneumococcal bacteria and *B. anthracis*, even when delivered intravenously and intraperitoneally (Jado et al., 2003; Loeffler et al., 2003). However, systemically treating bacteremia or other internal infections with lysins presents complications with the immune system that are avoided or diminished by topical application, although thus far there have been no problematic effects observed (Borysowski et al., 2006).

In more difficult cases of bacterial infection it can be beneficial to use multiple treatments. Co-administration of Pal and Cpl-1, two lysins specific to *S. pneumoniae* but targeting different bonds in the cell wall, showed a synergistic lethal effect (Loeffler et al., 2003). In addition, several studies have found that lysins can work synergistically with antibiotics to clear infections, provided that the bacterial strain is sensitive to the antibiotic (Djurkovic et al., 2005; Loeffler and Fischetti, 2003). Adding a second bactericidal agent reduces the opportunity for resistance to arise, and in attacking the cell wall, the lysin may make access easier for the antibiotic. This cooperative treatment could help keep antibiotics as a viable treatment option and reduce the spread of resistance.

1.5.6.2 Biocontrol of Food

The specificity of lysins provides a unique opportunity for biological control of unwanted bacteria. Applications of lysins in the food industry are already a reality and the efforts have been pioneered by the dairy industry. *Listeria monocytogenes,* the causative agent of listeriosis, is a serious contamination concern in dairy cultures (McLauchlin et al., 2004). Loessner's group has spent considerable time investigating the properties of *Listeria* phage (Loessner et al., 1999; Loessner et al., 2002; Loessner et al., 1995) and has designed a elegant system to prevent contamination of *Lactococcus* cultures with *L. monocytogenes*. Rather than simply add purified lysin, they engineered *L. lactis* cells to secrete Ply118 and Ply511 into the medium, leading to lysis of any *Listeria monocytogenes* cells while leaving *L. lactis* and other bacteria intact (Gaeng et al., 2000). This approach is also being considered for biocontrol of *C. perfringens*, a pathogen that is problematic in the poultry industry (Zimmer et al., 2002).

In fact, genetic engineering of lysins has progressed beyond the level of bacteria. Transgenic plants that express lysin genes are being created. Notably is the T4 lysozyme potato, which produced the T4 lysin to protect against damage caused by the phytopathogen *Erwinia carotovora* (de Vries et al., 1999). When damaged, the plant cells released the lysin and hydrolyzed the bacterial cells. An even more ambitious undertaking was the engineering of cows that secreted lysostaphin (a bacterial autolysin similar to phage lysins) into the milk (Wall et al., 2005). The milk was able to kill *S. aureus* and the cows were protected from *S. aureus* mastitis. As concerns about the amount of antibiotics being given to livestock continue and pesticide-resistant crop pathogens are becoming more frequent, biocontrol of our food sources with phage lysins is an appealing option.

68

1.5.6.3 Other Applications

There are a variety of applications being pursued for lysins beyond medical treatment and biocontrol of food. Control of *B. anthracis*, the causative agent of anthrax, was of great interest after concerns of bioterrorism in the U.S. Postal Service in late 2001 to early 2002. The PlyG lysin was isolated from the γ phage infecting *B. anthracis* and found to have exogenous lytic activity both on vegetative cells and the more dangerous spores (Schuch et al., 2002). Further, due to the high specificity of the lysin, it was able to detect *B. anthracis* in samples, first by adding PlyG and then measuring any ATP release from lysing cells using a luciferase assay (Schuch et al., 2002) and later by using the high-affinity CBD alone for rapid and precise detection (Fujinami et al., 2007). High-affinity CBDs are also being used in research to label cells and immobilize them on a solid surface (Loessner et al., 2002).

Lysins can also be used to lyse their own cells in a controlled manner. One interesting application of this is in cheese ripening. The lysis of the *L. lactis* cells in the starter culture of cheese releases intracellular enzymes that are involved in flavor formation. By inducing autolysis through the expression of *lytA* and *lytH*, the lysin and holin of the lactococcal phage ϕUS3 the cheese could theoretically ripen more quickly (de Ruyter et al., 1997). Another novel application of lysin-mediated autolysis is the release of a cell's cytoplasmic contents. This technique has been employed in attempts to generate vaccines and more directly induce the host immune response. The single lysis protein E of ϕX174 has been expressed within *E. coli* and *H. pylori* to create "ghosts," empty cell envelopes that are vaccine candidates (Haidinger et al., 2003; Jalava et al., 2003; Panthel et al., 2003). Another novel approach termed "bactofection" uses the intracellular pathogenic *L. monocytogenes* as a DNA delivery system by lysin-mediated spontaneous autolysis once within the mammalian cells (Pilgrim et al., 2003). A similar

approach was taken with *Vibrio cholerae* and *Salmonella enterica* expressing the λ S and R lysis genes that act as mucosal DNA delivery vectors (Jain and Mekalanos, 2000).

It is worth noting that these prior examples have all either utilized both a holin and lysin or the single lysis gene *E* from φX174 (Bernhardt et al., 2001a). Currently most applications of lysins involve exogenous exposure, which is effective for the accessible PG of Gram-positive bacteria. However, there are advances being made in adapting lysins for use with Gram-negative bacteria. The T4 lysozyme (During et al., 1999) and the *B. amyloliquefaciens* phage lysin (Morita et al., 2001b) have C-terminal sequences that are similar to cationic antimicrobial peptides, and it appears they may be capable of permeabilizing the outer membrane to induce lysis of Gram-negative bacteria exogenously (Morita et al., 2001a; Morita et al., 2001b; Orito et al., 2004). It is possible to engineer lysins with C-terminal peptides that enhance lysis, as has been shown with lysozyme and *E. coli* (Ibrahim et al., 1994), and so it may be possible to take advantage of these known anti-microbial regions to confer the ability to lyse Gram-negative bacteria on other lysins.

1.5.6.4 Considerations

The use of lysins in the above applications raises several questions and concerns. Most important is the safety of their use in medical therapy. Lysins target PG, a structure exclusively found in bacteria, so there is minimal damage to be expected from non-specific action on non-bacterial host substrates, especially when the lytic activity is combined with the high specificity conferred by the CBD. That being said, at least one endopeptidase, the *S. aureus* autolysin lysostaphin, has shown the ability to degrade elastin, a glycine-rich polymer in mammalian tissues (Park et al., 1995). The larger variety of substrate bonds targeted by peptidases may make them less ideal for such treatments as opposed to muramidases and amidases that target

bonds unique to PG (Borysowski et al., 2006). However, no topical or systemic applications of lysins, even lysostaphin, have shown any toxicity or irritation to date, even after multiple exposures (Climo et al., 1998; Dajcs et al., 2000; Loeffler et al., 2003; Loeffler et al., 2001; Nelson et al., 2001).

Interactions with the immune system are a concern for lysin therapies in two ways. Proteins are naturally immunogenic (able to elicit an immune response); they are recognized as foreign and stimulate the humoral response to produce antibodies designed to inactive them (Fischetti, 2005). In addition, lysins will have to survive the body's basic defenses, such as mucosal barriers, secreted proteases, and inhibitory pH environments such as the stomach. As studied by repeated injections in treating mice and rabbits, introducing lysins systemically into the body did generate antibodies but did not significantly reduce the therapeutic activity (Jado et al., 2003; Loeffler et al., 2003). The antibacterial activity against some *Streptococcus* and *Bacillus* species by lysins was not neutralized by the antibodies, potentially due to the very high affinity interaction of the lysin with the cell wall (Fischetti, 2005). Regardless, the potential inactivation of lysins that provoke an immune response is still a concern. One approach to reduce the immunogenicity of lysins is conjugation to polyethylene glycol (PEG). The addition of PEG to a protein both reduces antibody binding, inhibits proteolysis and clearance, and impairs uptake by dendritic cells for antigen processing (Delgado et al., 1992). PEGylation of lysostaphin has already been proven to be effective at reducing antibody binding and increasing the half-life in the serum (Walsh et al., 2003).

The second immune-centered concern for lysin treatment is the release of pro-inflammatory cell wall components as a result of bacterial lysis. The immune system is hypersensitive to many bacterial molecules including endotoxins like lipopolysaccharide,

71

teichoic and lipoteichoic acids, and PG. A sudden massive release of such material can result in septic shock that cascades into a large-scale inflammatory response that can lead to multiple organ failure and death (Nau and Eiffert, 2002). Again, no evidence of any bacteriolytic side effects has been observed in animal model studies to date (Jado et al., 2003; Loeffler et al., 2003).

A major problem in current bacterial treatments is resistance to antibiotics (Di Perri and Bonora, 2004; Lowy, 2003; Ong et al., 2004; Razavi et al., 2007; Saenz et al., 2004). It is in their inability to generate resistance that lysins trump antibiotics (Loeffler et al., 2001; Schuch et al., 2002). Since lysins are essential for the lytic phage life cycle they have evolved to target molecules essential to their phages' hosts (Borysowski et al., 2006). Instead of targeting a single protein in a synthesis pathway, endolysins act upon the structure that is the result of dozens of enzymes and is largely conserved across the entire bacterial kingdom (Schleifer and Kandler, 1972). Further, the CBD of lysins binds to highly host-specific components of the cell wall. For example, most *Streptococcus* phage lysins bind teichoic acids containing choline, which is necessary for viability (Garcia et al., 1988; Hermoso et al., 2003). Another cell wall component bound by lysins is polyrhamnose, which is important for the growth of Group A *Streptococcus* species (Fischetti, 2003). However, while there are effective antibiotics against both Gram-postive and Gram-negatice bacteria, lysins remain ineffective as a treatment of Gram-negatice infections due to their inability to reach the PG.

One factor that can affect a bacterium's susceptibility to lysins is the bacterial growth phase. Lytic activity was decreased against stationary phase bacteria (Loeffler et al., 2001; Pritchard et al., 2004). This is likely because of the changes in the cell wall associated with entering a non-replicating state: increased cross-linking of the PG, deacetylation of the amino

sugars, and increases in cell wall-associated proteins and polysaccharides (Borysowski et al., 2006). However, this is not unique to lysins, since these cells are also less susceptible to many antibiotics (Tuomanen et al., 1986) and lytic phage (Levin and Bull, 2004).

While lysins possess many qualities that would make them a good potential antimicrobial, there are still potential problems are disadvantages when compared to phage therapy and antibiotics. A major advantage of treatment through phage therapy is the self-replication of the therapeutic agent, which is not possible with purified antibiotics or lysins (O'Flaherty et al., 2009). While antibiotics may have other side effects, they are largely resistant to the body's defenses; hostile environments, proteases, and the antibodies produced by the humoral immune system may inactivate phages and lysins (Borysowski et al., 2006). Phages and antibiotics have also been identified that are active against Gram-negative bacteria, while little research has produced lysins that can effectively lyse these bacteria upon exogenous application (Briers et al., 2007).

In summary, no single antimicrobial agent is perfect, but there is much potential in phage lysins in the treatment, control, and study of bacteria for numerous reasons: (i) lysins achieve rapid lysis both *in vitro* and *in vivo*; (ii) unlike antibiotics, lysins do not appear to generate resistance; (iii) lysins target a structure that is only found in bacteria; (iv) due to the evolved specificity of the CBD, lysins have a very narrow spectrum of activity, preventing complications from the destruction of indigenous microflora; (v) to date, lysins appear to be safe for topical and systemic administration and are not significantly impaired by the immune system; (vi) the ease of genetic engineering produces opportunities to create lysins with optimal lytic ability and potentially exogenous activity against Gram-negative bacteria; (vii) the vast number of

bacteriophage ($\sim 10^{31}$) (Hendrix, 2003) provides an immense reservoir for new lysins, as well as other antimicrobials.

1.6   SUMMARY

Past studies of dsDNA bacteriophage lysis have centered on the phages of Gram-negative bacteria, such as λ and T4 (Calendar, 2006). However, over the past two decades this research has accelerated, resulting in new models of phage lysis that reflect its regulation by holins and execution by endolysin proteins (Young, 1992). Concurrently, research on the endolysins of Gram-positive bacteria has also expanded, revealing their modular organization, the variety of their activities, and their potential as highly specific and effective anti-bacterial agents (Fischetti, 2008). In more recent years, the role of several accessory lysin proteins has been determined (Berry et al., 2008); indeed new ones are still being identified (Catalao et al., 2010), which illustrates just how complex the phage lysis system is, and how much remains to be understood. Little research has been conducted on mycobacteriophage lysis proteins, and in addition to the benefits of increased knowledge of lysis strategies, the threat posed by drug-resistant forms of TB warrants the investigation of these proteins as potential therapeutics or research tools. My thesis research has focused on two lysis proteins that are found in mycobacteriophages: LysA and LysB.

1.6.1 Specific Aim 1: Bioinformatic characterization of the mycobacteriophage Lysin A proteins.

Mycobacteriophage LysA proteins are expected to serve the role of the endolysin, which hydrolyzes the PG to induce host lysis. More than 60 mycobacteriophage genomes have been sequenced, and each has a LysA homolog. This wealth of sequence data provides an opportunity to bioinformatically examine these proteins, which appear to have complex architectures generated by recombination and composed of a variety of domains. In addition to the expected PG hydrolytic domains, several domains of unknown function that are unique to mycobacteriophages are also present. The results of this bioinformatic characterization are described in Chapter 2. This initial survey of the LysA proteins lays the foundation for future research, including an examination of phage evolutionary mechanisms that generate these mosaic proteins, and investigation of the domains unique to mycobacteriophage endolysins and implications for lysis of mycobacteria. Mycobacteriophage lysins also have potential as therapeutics, and preliminary data shows some anti-mycobacterial activity by some LysAs; these bioinformatic studies will enable directed genetically engineering of lysins for use in research and antimicrobial therapy.

1.6.2. Specific Aim 2: Determination of the function of mycobacteriophage Lysin B proteins.

Prior studies have indicated that LysB possesses lipolytic activity characteristic of a serine esterase (Gil et al., 2008). However, these studies utilized artificial substrates, and so the natural

target in the context of mycobacterial infection is unknown. We hypothesized that the ester bond covalently linking the mycolic acids to the rest of the cell wall is the target of LysB. We tested this by cloning the mycobacteriophage D29 LysB protein and assaying its activity on purified mAGP. The results of these studies comprise the first half of Chapter 3. This study has revealed the first known protein with mycolylarabinogalactan esterase activity. LysB may be useful in research methods including isolation of mycobacterial cell wall fractions, and preliminary studies indicate some anti-mycobacterial activity, which should be further characterized.

1.6.3. Specific Aim 3: Identification of the role of Lysin B in mycobacteriophage infections.

While a LysA protein is present in all mycobacteriophage genomes, only 56 out of the 60 surveyed mycobacteriophage genomes lack a LysB homolog, suggesting that the protein is not essential for lytic infection. To examine the role of LysB, we created a *lysB* deletion mutant in mycobacteriophage Giles. We compared the mutant phage to the wildtype Giles phage in its ability to form plaques, efficiently lyse cells, and release progeny phages. The second part of Chapter 3 documents these results, which provide further insight into the complex system required for the lysis of the host mycobacterial and further emphasize the importance of its cell wall as a barrier.

2.0   CHARACTERIZATION OF THE MODULAR LYSA PROTEINS

2.1   INTRODUCTION

Endolysin proteins, which function to degrade the host cell wall at the completion of lytic growth, are found in all tailed dsDNA bacteriophages. Recent work has focused on characterizing these proteins from the phages of a number of Gram-positive bacteria, including *S. aureus, S. pneumoniae,* and *L. monocytogenes* (Borysowski et al., 2006; Loessner, 2005; Lopez and Garcia, 2004). However**,** little study has been done on mycobacteriophage endolysins, the LysA proteins, which have been identified largely based on homology to known peptidoglycan (PG) hydrolases (Garcia et al., 2002; Hatfull et al., 2006). The first test of lytic activity was done by Garcia *et al*. (2002), who examined the lysis cassette of Ms6 and identified gp2 (ORF2). This was characterized as having lytic activity based on the observation that *E. coli* cells overexpressing ORF2 lysed upon the addition of chloroform, which permeabilizes the membrane and presumably allows the lysin access to its substrate. Additional zymography studies of several LysAs (Bxz1 gp236, Che8 gp32, Corndog gp69, D29 gp10) have also shown PG hydrolytic activity on *M. luteus* PG (Appendix A.1) (Payne et al., 2009). Most recently, TM4 gp29 has been shown by Henry *et al*. (2010) to have PG hydrolytic activity by promoting cell lysis in the chloroform assay and clearing on a zymogram of *M. luteus* PG. While these studies demonstrate empirically that at least several of the LysAs of mycobacteriophage have PG

hydrolytic activity, neither the specific enzymatic mechanisms nor cell wall binding domains have been characterized. There is, however, a wealth of raw sequence data available, which indicates a great diversity of activities within the LysA proteins themselves. A bioinformatic analysis of the more than 60 LysA proteins published (as of July 2010) will provide insight into the structure and diversity of these proteins, allowing a more targeted approach to examinations of their activity and potential application in the future.

2.2   LYSAS SHOW EXTENSIVE MODULARITY

Phage lysins are established to be modular with two domains: an N-terminal PG hydrolytic domain and a C-terminal cell wall binding domain (Lopez et al., 1997). Very few lysins have been identified that have three domains, and in every instance these lysins possess two lytic domains of differing specificities (Cheng and Fischetti, 2007; Navarre et al., 1999; Pritchard et al., 2004). In light of these precedents, the mycobacteriophage LysA proteins are exceptional in the predominance of the three-domain architecture, and in the prevalence of domains of unknown function and LysAs with no identifiable PG hydrolytic activity.

2.2.1   Organization of LysAs

In order to further characterize the mycobacteriophage LysAs, BLASTp searches were performed to find conserved domains with either predicted PG hydrolase activity or sequence similarity to regions in other known lysins. While many LysAs did have hits to PG hydrolytic domains, most also had domain-sized regions with no identifiable sequence similarity outside of other mycobacteriophage LysAs. In many cases, these were conveniently on the N-terminal or C-terminal side of a predicted PG hydrolytic domain, effectively separating the protein into three modules. In order to characterize these distinct regions, a domain was defined as a region of significantly shared amino acid sequence identity (e-value < $1e^{-05}$, ClustalW similarity >20%) between two or more mycobacteriophage LysAs. Similar to the mycobacteriophage Pham

concept (Pedulla et al., 2003), claim to a domain required having similarity to at least one other defined region in another mycobacteriophage LysA. As a result, two LysAs may be entirely unrelated by the above criteria but are still categorized as LysAs based on domain-specific similarities to a third LysA. Up to three iterations of PSI-BLAST analysis that only included other mycobacteriophage LysA hits in subsequent iterations were used to ensure that all domains were distinct from each other. Care was taken with all domains to ensure that a function was not attributed based solely on similarity to other proteins; several N- and C-terminal domains shared similarity with proteins annotated as PG hydrolases, but the shared region did not include the PG hydrolytic activity. This was necessary to avoid a "guilt by association" error that is an important consideration when bioinformatically analyzing mosaic proteins. Applying this definition to more than 60 LysAs, we found 13 shared domains, as well as three domain-sized regions that as yet have no match in other mycobacteriophage LysAs. The identified domains consist of four unknown N-terminal domains (labeled N1-N4), six domains with BLAST-predicted PG hydrolytic activity (Ami1, Ami2, GH19, GH25, M23, and TG), and three separate C-terminal domains (labeled C1-C3); all of which will be discussed in more detail below. These domains are distributed across 20 mycobacteriophage clusters and sub-clusters (groups of phage with a certain level of genome sequence similarity) and four singleton phages (Table 1).

Table 1. LysA domain structures organized by phage cluster

Cluster	Phage	gp #	N-terminal						Catalytic					C-terminal					
			N1	N2	N3	N4	Ami I	Ami II	GH19	GH25	M23	TG	NONE	C1	C2	C3	LGFP	Unmatched	2-Dom.
A1	Bethlehem	8	+	-	-	-	-	-	+	-	-	-	-	+	-	-	-	-	-
	Bxb1	8	+	-	-	-	-	-	+	-	-	-	-	+	-	-	-	-	-
	DD5	10	+	-	-	-	-	-	+	-	-	-	-	+	-	-	-	-	-
	Jasper	10	+	-	-	-	-	-	+	-	-	-	-	+	-	-	-	-	-
	KBG	10	+	-	-	-	-	-	+	-	-	-	-	+	-	-	-	-	-
	Lockley	9	+	-	-	-	-	-	+	-	-	-	-	+	-	-	-	-	-
	Solon	9	+	-	-	-	-	-	+	-	-	-	-	+	-	-	-	-	-
	U2	7	+	-	-	-	-	-	+	-	-	-	-	+	-	-	-	-	-
A2	Che12	11	-	-	-	+	-	-	+	-	-	-	-	+	-	-	-	-	-
	D29	10	-	-	-	+	-	-	+	-	-	-	-	+	-	-	-	-	-
	Hammer*	13	+	-	-	-	-	-	-	-	-	+	-	+	-	-	-	Catalytic	-
	L5	10	-	-	-	+	-	-	-	-	-	-	+	+	-	-	-	-	Y
	Pukovnik	11	+	-	-	-	-	-	+	-	-	-	-	+	-	-	-	-	-
A3	Bxz2	11	-	-	-	-	-	-	+	-	-	-	-	+	-	-	-	N-term	-
B1	Chah	50	-	-	-	-	+	-	-	-	+	-	-	-	+	+	-	-	-
	Orion	49	-	-	-	-	+	-	-	-	+	-	-	-	+	+	-	-	-
	PG1	49	-	-	-	-	+	-	-	-	+	-	-	-	+	+	-	-	-
B2	Qryzula	44	+	-	-	-	-	-	+	-	-	-	-	+	-	-	-	-	-
	Rosebush	46	+	-	-	-	-	-	+	-	-	-	-	+	-	-	-	-	-
B3	Phaedrus	44	-	-	+	-	+	-	-	-	-	-	-	-	+	-	-	-	-
	Pipefish	48	-	-	+	-	+	-	-	-	-	-	-	-	+	-	-	-	-
B4	Cooper	44	-	+	-	-	-	+	-	-	-	-	-	+	-	-	-	-	-
	Nigel	42	-	+	-	-	-	+	-	-	-	-	-	+	-	-	-	-	-
C1	Bxz1	236	-	+	-	-	-	+	-	-	-	-	-	+	-	-	-	-	-
	Cali	240	-	+	-	-	-	+	-	-	-	-	-	+	-	-	-	-	-
	Catera	239	-	+	-	-	-	+	-	-	-	-	-	+	-	-	-	-	-
	Rizal	239	-	+	-	-	-	+	-	-	-	-	-	+	-	-	-	-	-
	ScottMcG	242	-	+	-	-	-	+	-	-	-	-	-	+	-	-	-	-	-
	Spud	242	-	+	-	-	-	+	-	-	-	-	-	+	-	-	-	-	-
C2	Myrna	243	+	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	Y
D	Adjutor	36	+	-	-	-	-	-	-	-	-	-	+	-	+	+	-	-	Y
	Butterscotch	35	+	-	-	-	-	-	-	-	-	-	+	-	+	+	-	-	Y
	Gumball	34	+	-	-	-	-	-	-	-	-	-	+	-	+	+	-	-	Y
	PBI1	35	+	-	-	-	-	-	-	-	-	-	+	-	+	+	-	-	Y
	Plot	36	+	-	-	-	-	-	-	-	-	-	+	-	+	+	-	-	Y
	Troll4	35	+	-	-	-	-	-	-	-	-	-	+	-	+	+	-	-	Y
E	244	34	-	-	-	+	-	-	-	+	-	-	-	+	-	-	-	-	-
	CJW1	32	-	-	-	+	-	-	-	+	-	-	-	+	-	-	-	-	-
	Kostya	33	-	-	-	+	-	-	-	+	-	-	-	+	-	-	-	-	-
	Porky	31	-	-	-	+	-	-	-	+	-	-	-	+	-	-	-	-	-
F1	Boomer	32	-	+	-	-	+	-	-	-	-	-	-	-	+	-	-	-	-
	Che8	32	-	+	-	-	+	-	-	-	-	-	-	-	+	-	-	-	-
	Fruitloop	29	-	-	+	-	+	-	-	-	-	-	-	-	+	-	-	-	-
	Llij	30	-	-	+	-	+	-	-	-	-	-	-	-	+	-	-	-	-
	Pacc40	30	-	+	-	-	-	+	-	-	-	-	-	-	+	-	-	-	-
	PMC	30	-	-	+	-	+	-	-	-	-	-	-	-	+	-	-	-	-
	Ramsey	32	-	+	-	-	+	-	-	-	-	-	-	-	+	-	-	-	-
	Tweety	30	-	-	+	-	+	-	-	-	-	-	-	-	+	-	-	-	-
F2	Che9d	35	-	-	-	-	-	+	+	-	-	-	-	-	+	-	-	-	-
G	BPs	27	-	-	-	-	+	-	-	-	+	-	-	-	+	+	-	-	-
	Halo	27	-	-	-	-	+	-	-	-	+	-	-	-	+	+	-	-	-
H1	Konstantine	39	+	-	-	-	-	-	-	-	-	-	+	-	+	+	-	-	Y
	Predator	30	+	-	-	-	-	-	-	-	-	-	+	-	+	+	-	-	Y
H2	Barnyard	39	-	-	-	-	-	+	-	-	-	-	-	-	-	-	+	C-term	Y
I	Brujita	29	-	-	-	+	-	-	-	-	-	-	+	-	-	+	-	C-term	Y
	Che9c	25	+	-	-	-	-	-	-	-	-	-	-	-	+	+	-	-	-
J	Omega	50	-	-	-	+	-	-	-	+	-	-	-	+	-	-	-	-	-
K	TM4	29	-	+	-	-	+	-	-	-	-	-	-	+	-	-	-	-	-
None	Corndog	69	-	-	+	-	+	-	-	-	-	-	-	-	+	-	-	-	-
	Giles	31	-	-	-	-	+	-	-	-	+	-	-	-	+	+	-	-	-
	Wildcat	49	-	-	-	-	+	-	+	-	-	-	-	+	-	-	-	-	-

*unpublished
Unmatched – domain-sized regions with no similarity to other LysAs
2-Dom. – LysA with only two domains

2.2.2 Identified Domains and Combinations

These domains (and three unknown regions) combine to make 19 unique organizations (Figure 13). The most common pattern, observed 9/19 times, consists of an N-terminal domain (N1 – N4), a central lytic domain with predicted PG hydrolytic activity, and a C-terminal domain (Figure 13 C, D, F, H, I, L, M, P, R); in addition, Bxz2 gp11, though possessing an unmatched N-terminal region, also appears to have a three-domain organization (Figure 13 E). Four other organizations contain three domains; each has two predicted lytic domains in combination with a C1-C3 C-terminal domain. Interestingly, the central catalytic activity in these two-lytic domain lysins is always an amidase, while the N-terminal lytic domain is either a GH19 muramidase in Che9d gp35 (Figure 13 G) and Wildcat gp49 (Figure 13 S), or an M23 peptidase in Cluster G and Sub-cluster B1 (Figure 13 K) and in Giles gp31 (Figure 13 J).

Lastly, there are five organizations that appear to contain only two domains. Three consist of one of the unknown N-terminal domains (N1 or N4) matched with a C-terminal domain or region [N1-C3 for Cluster D, Sub-cluster H1, and Che9c gp25 (Figure 13 Q); N4-C1 for L5 gp10 (Figure 13 N); and N4 with an unmatched C-terminal region containing a PG-binding domain for Brujita gp29 (Figure 13 B)]. Myrna gp243 (Figure 13 O) has an N-terminal domain followed by a lytic domain, but it lacks a recognizable C-terminal domain or region; however, it has a central region of considerable size (~100 aa) that has no homology to any known protein. Barnyard gp39 (Figure 13 A) is the only LysA that resembles canonical bacteriophage lysins: it contains a predicted lytic domain at its N-terminus and predicted binding region at the C-terminus with identifiable binding motifs, but this C-terminal region is unmatched to other LysAs.
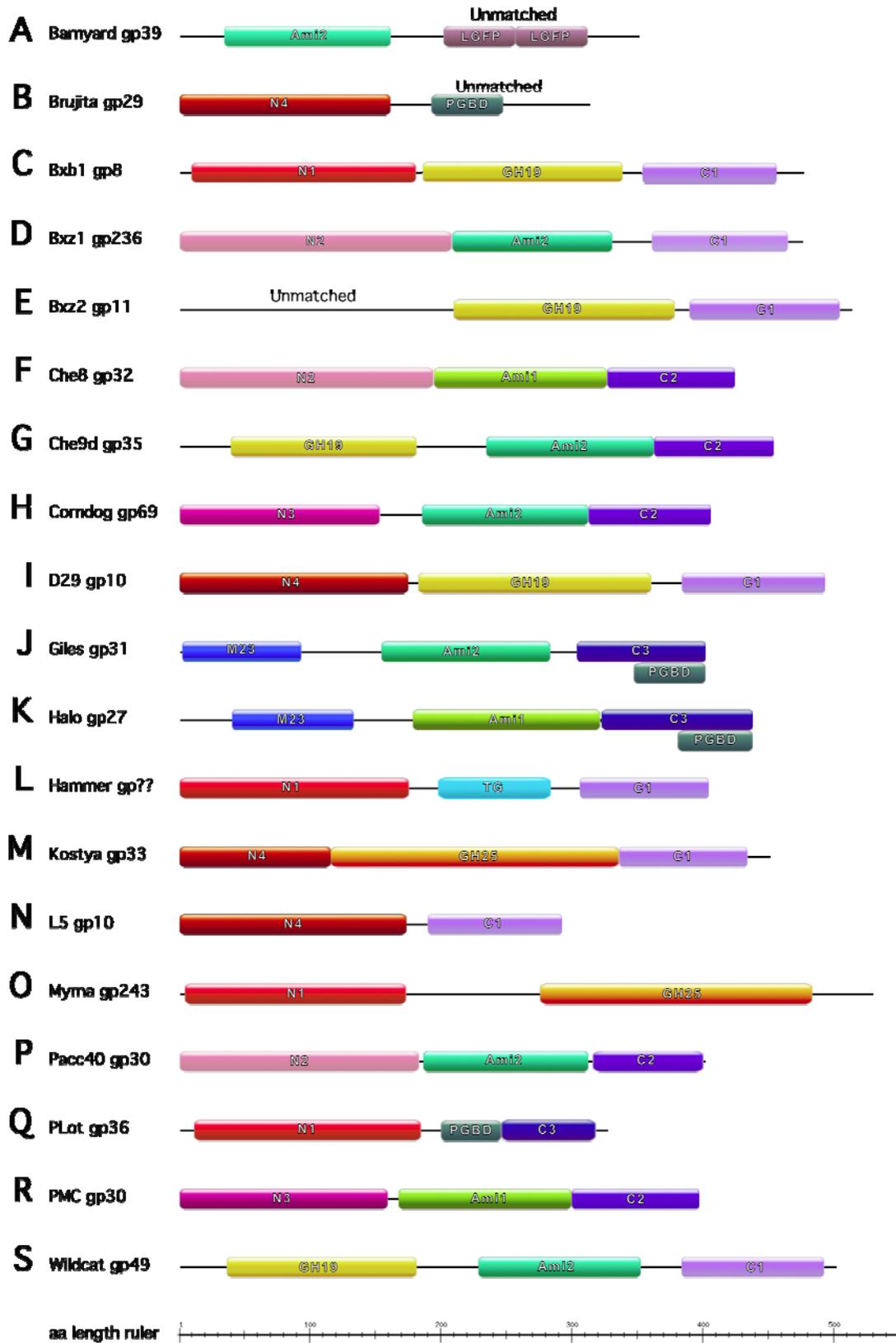
Figure 13. LysA modular organizations

Figure 13: Through combinations of 13 domains and three unmatched regions, 19 different organizations are observed in the 70 mycobacteriophage LysAs. Representative phage LysAs are shown for each organization. N1-N4 (red and pink shades) and C1-C3 (purple shades) are N-terminal and C-terminal regions, respectively, with no predicted function. Predicted catalytic functions include Ami1 and Ami2 (amidase 1, green; and amidase 2, aquamarine), GH19 (glycoside hydrolase family 19, yellow), GH25 (glycoside hydrolase family 25, orange), M23 (M23 family peptidase, blue), and TG (lytic transglycosylase, cyan). Also identified are predicted binding motifs, PGBD (PG binding domain, gray) and LGFP repeats (rose).

2.2.3   Evidence of Recombination in LysAs

The mosaicism of mycobacteriophage genomes (Hatfull, 2010) is echoed on a smaller scale in the LysA proteins. The distribution of different domains throughout the LysA proteins speaks to the variety in the mycobacteriophage metaproteome, but it is interesting to note that members of specific phage clusters do not necessarily possess the same LysA organizations; evidence of recombination can be found within and between clusters and sub-clusters of phage (Table 1). This is especially noticeable in Sub-clusters A2 and F1. Among the six members of Sub-cluster A2, there are four different organizations (Figure 13 C, I, L, N) made by combining five different domains: N1, N4, GH19, TG, and C1 (specific domain functions are discussed below). The differences arise from the swapping between the N1 or N4 domains, as well as an exchange of the more common GH19 domain for a transglycosylase (TG) in Hammer gp13 (Figure 13 L, Figure 14). Further, in comparing the lysis cassette of L5 gp10 (Figure 13 N) with fellow members of Sub-cluster A2, it appears that the central lytic domain has simply been lost (Figure 14), similar to the absence of LysB in the lysis cassette of Che12 (Chapter 3).
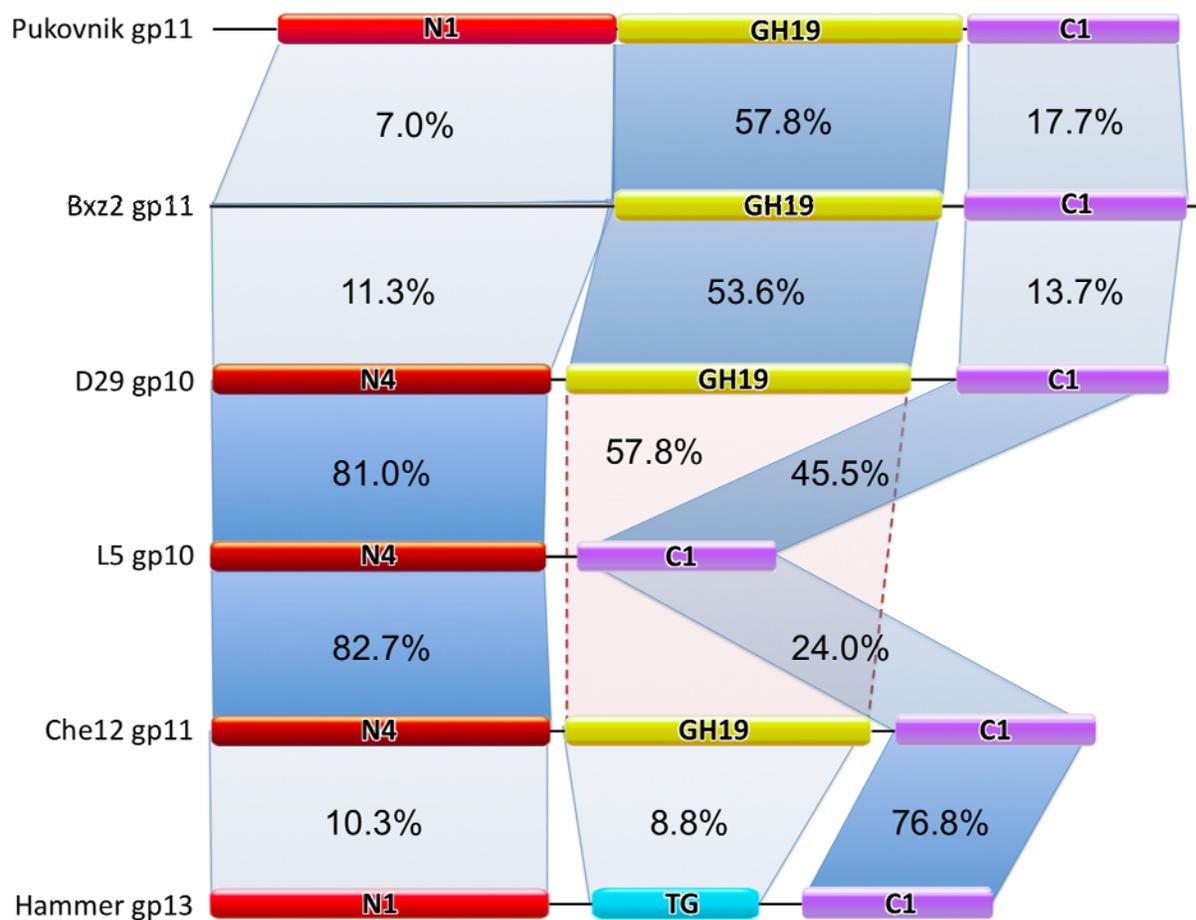
Figure 14. Domain organization in Sub-cluster A2 LysAs

Figure 14: A comparison of the domains present in the LysAs of Sub-cluster A2 phages with the degree of sequence similarity based on a ClustalW alignment indicated in boxes between the regions. The similarity between the GH19 domain of D29 gp10 and Che12 gp11 is shown in the red box with the dotted lines.

In the F1 sub-cluster (Table 1) there are three different organizations (Figure 13 F, P, R), resulting from combinations of five domains, including N2, N3, the two predicted amidases Ami1 and Ami2, and C2. Most LysAs in this phage sub-cluster appear to only interchange the N-terminal domain, with the exception being Pacc40 gp30 (Figure 13 P), in which the Ami1 lytic domain has been replaced by Ami2. The C-terminal domain is constant in all Sub-cluster F1 LysAs, including Pacc40 gp30. Portions of the F1 LysAs are similar to other LysAs, such as Bxz1 LysA and other LysAs of Sub-cluster C1 phages (Figure 13 D), and the singletons TM4 (Figure 13 D) and Corndog gp69 (Figure 13 H).

The relationships of four of these proteins both as a whole and by individual domains were analyzed for evidence of recombination: Che8 gp32 (N2-A1-C2); Corndog gp69 (N3-A2-C2); TM4 gp29 (N2-A2-C1); and Tweety gp30 (N3-A1-C2) (Figure 15). Maximum likelihood phylogenies were determined for the N-terminal and amidase domains using PHYML (Guindon and Gascuel, 2003) (Figure 15). Bootstrap values were calculated to test support for the predicted phylogeny (Felsenstein, 1985), and an SH test was performed to test the likelihood of alternate topologies explaining the sequence data (Shimodaira and Hasegawa, 1999). The grouping of LysAs is different for each domain (Figure 15 B, C). For example, the tree of the N-terminal domains groups those of Che8 gp32 and TM4 gp29 together, both of which are identified as having N2 domains, while the LysAs with N3 domains, Tweety gp29 and Corndog gp69, are also distinct from the N2 LysAs and are clustered together. However, an evaluation of the Ami1 and Ami2 domains shows a completely different phylogeny, grouping Che8 gp32 with Tweety gp30 and Corndog gp69 with TM4 gp29. The most parsimonious phylogenies (Figure 15 B, C) are strongly supported by bootstrap values of 100, and the SH test rejects alternate topologies (SH test *P*-values < 0.001). These results support recombination events as an

explanation for the observed modularity of the LysA proteins. Overall, given the variety of domain organizations and the evidence of recombination, LysAs appear to be the most varied and complex phage lysins yet studied.
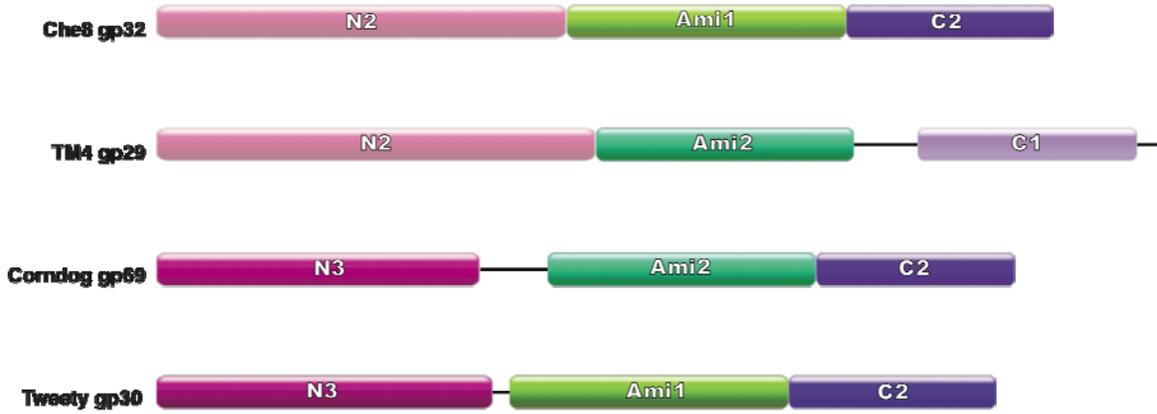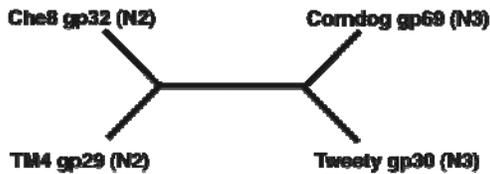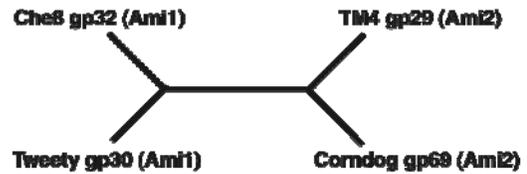
Figure 15. Recombination of domains between LysAs

Figure 15: ClustalW alignments of four LysA proteins were used to construct a phylogenetic tree for the N-terminal and amidase domains. A. Proteins are identified by phage name and gp#. B. and C. Tree illustrating the most parsimonious phylogeny for the (B) N-terminal and (C) amidase domains. Bootstrap values are 100 for the division of recombination between the domains, and an SH test rejected alternate topologies (*P*-value <0.001). Branch lengths do not represent evolutionary distance.

2.3 PEPTIDOGLYCAN HYDROLYTIC ACTIVITIES PREDICTED IN LYSAS

There are multiple predicted activities associated with mycobacteriophage LysAs. The majority of LysAs contain one putative catalytic domain; however, LysAs containing two or, surprisingly, zero have also been observed. The predicted PG activities were initially determined based on amino acid sequence similarity to known conserved domain families with a specific activity, in addition to sequence identity with proteins in GenBank that are annotated as having or predicted to have a specific activity. The domains were more closely examined by conducting multiple alignments and motif searches, and by identifying conservation residues important to activity.

2.3.1 Amidases

One of the most common types of PG hydrolase activity seen throughout phage lysins is the N-acetylmuramoyl-L-alanine amidase (Loessner, 2005). Amidases hydrolyze the bond between N-acetylmuramoyl and L-amino acids to separate the glycan strands from the cross-linking peptides (Figure 8). Similar to other known phage lysins, the abundance of this activity throughout LysAs may speak to its effectiveness (Figure 13 A, D, F, G, H, J, K, P, R, S), and these domains are found in combination with every other domain except N1, N4, GH25, and TG. In mycobacteriophage LysAs, these common regions are recognized – with e-values ranging from $1e^{-10}$ to $1e^{-05}$ – as having conserved domains, including the Amidase_2 (pfam01510) and the Peptidoglycan Recognition Protein (PGRP) superfamily (cd06583), which itself includes Zn-dependent amidases (EC: 3.5.1.28), such as λ T7 lysozyme (Cheng et al., 1994), AmiD of *E. coli* (Kerff et al., 2010), and PlyL of *Bacillus anthracis* phage Ba02 (Low et al., 2005). The LysA amidase domains average 140 aa in length and are always in the center of the LysA, with the

exception of Barnyard gp39, which has only two domains total, with an N-terminal amidase activity (Figure 13 A).

Comparison of the amidase domains within the LysAs reveals what appear to be two distinct groups, here labeled as Ami1 (Figure 13 F, K, R) and Ami2 (Figure 13 A, D, G, H, J, P, S). There are no common hits between Ami1 and Ami2 domains, even after three iterations of PSI-BLAST and no significant similarity seen in ClustalW alignments (Figure 16), making it appear that the introduction of these amidase domains took place in two independent events and came from different sources. Indeed, members of the Ami1 group show little similarity outside of other mycobacteriophage LysAs; the few similar proteins revealed by PSI-BLAST are predicted lysins of *Corynebacterium* and *Propionibacterium* phage. Ami2 appears less distanced from non-mycobacteriophage proteins and shows similarity to hypothetical *Mycobacterium* species proteins, including *Mycobacterium tuberculosis* H37Rv hypothetical protein Rv3594, as well as a variety of amidases and PG-binding domain-containing proteins in other bacteria.
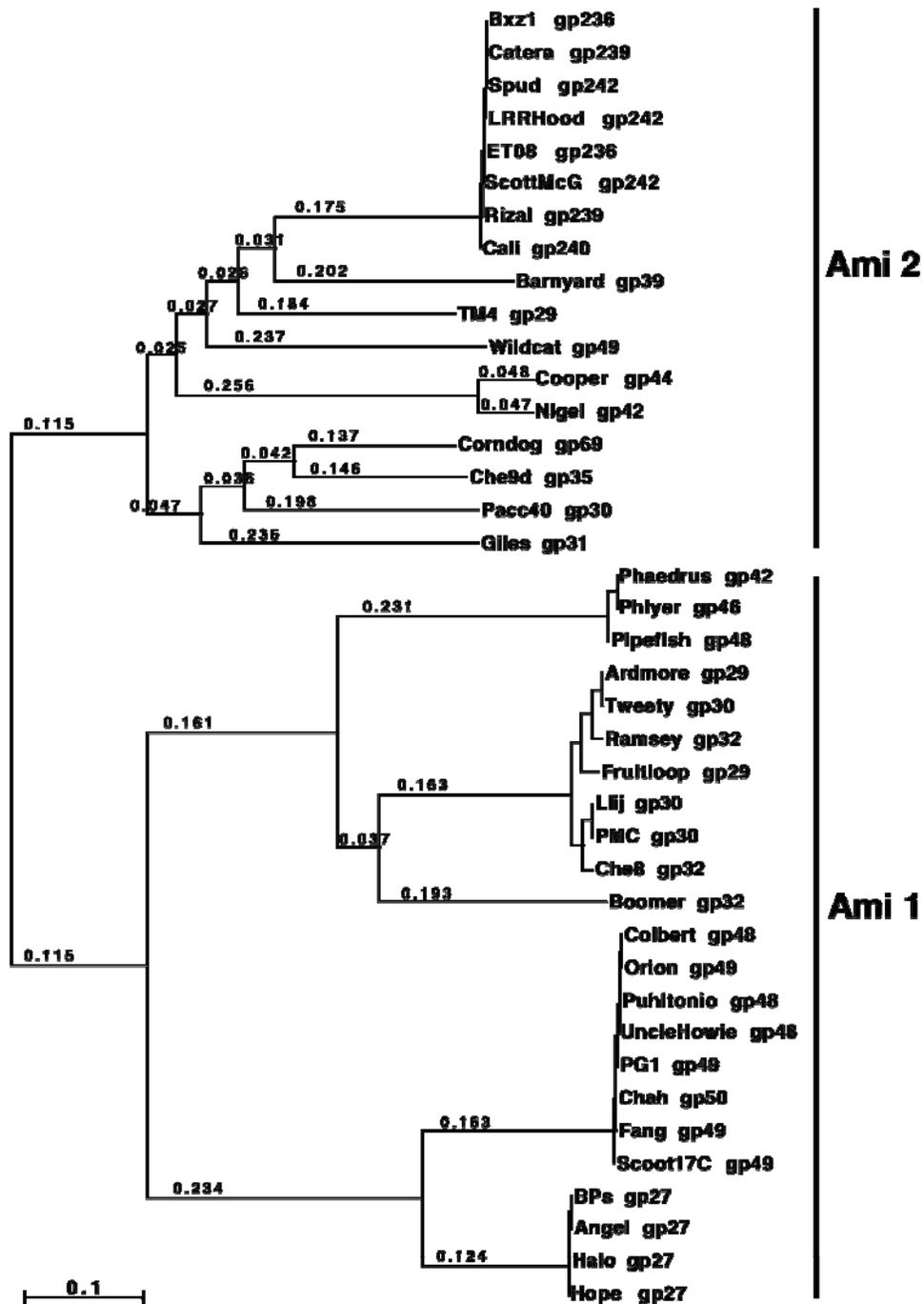
Figure 16. Phylogenetic Tree of LysA amidase domains

Figure 16: A neighbor-joining tree based on a ClustalW alignment of all amidase domains found in LysAs. Numbers indicate distance between nodes.

Central to amidase activity is the Zn^{2+} binding site. The structures of T7 lysozyme, PlyL, and AmiD have three Zn^{2+}-coordinating residues: His17-His122-Cys130 for T7 lysozyme, His29-His129-Cys137 for PlyL, and His35-His151-Asp161 for AmiD (Cheng et al., 1994; Kerff et al., 2010; Low et al., 2005). An alignment of the Ami1 and Ami2 domains (Figure 17) shows a conservation of these residues. Somewhat surprisingly, however, the third Zn^{2+}-binding residue for LysA amidases is an Asp similar to the bacterial AmiD as opposed to the Cys found in the two phage lysins. Additional residues involved in catalysis include a Lys (Lys128 in T7 lysozyme, Lys135 in PlyL, Lys159 in AmiD) and a Tyr or Glu (Tyr46 in T7 lysozyme, Glu90 in PlyL, and Glu104 in AmiD). The Lys residue is conserved in Ami2, exactly 2 residues before the catalytic Asp (Figure 17 B), but there is a His in this position in Ami1 (Figure 17 A). A conserved Tyr residue is seen approximately 30 aa from the first His in Ami1 (Figure 17 A), and there is both a conserved Glu and partially conserved Tyr about 70 and 60 aa away from the catalytic His (Figure 17 B). In AmiD, Thr37 (near the catalytic His35) is involved in substrate binding (Kerff et al., 2010). A conserved Thr is observed in both Ami1 and Ami2, appearing immediately after the predicted catalytic His in Ami1 and after the two conserved His residues in Ami2 (Figure 17). Based on these alignments, it is likely that both Ami1 and Ami2 are catalytically active.

A
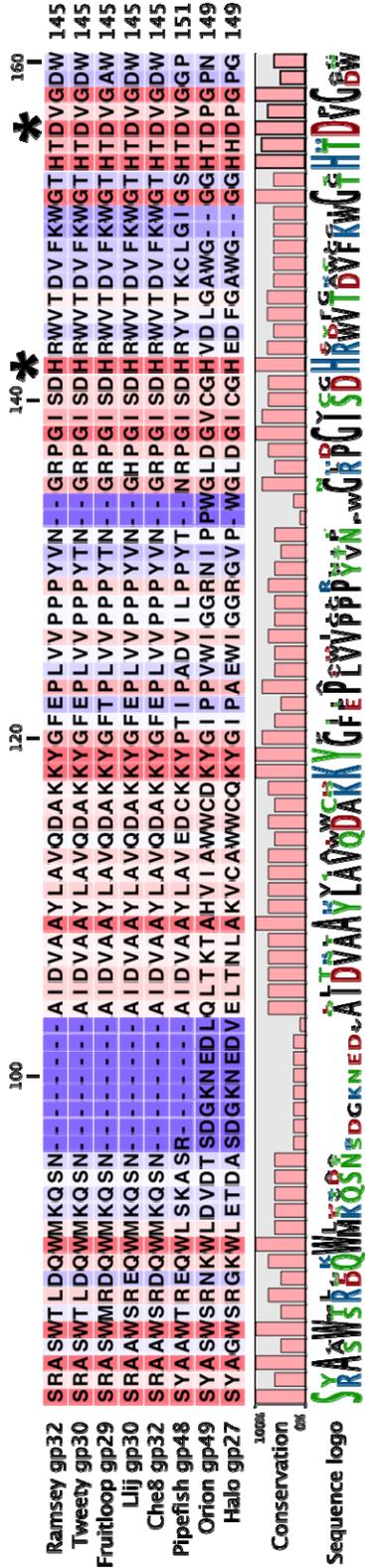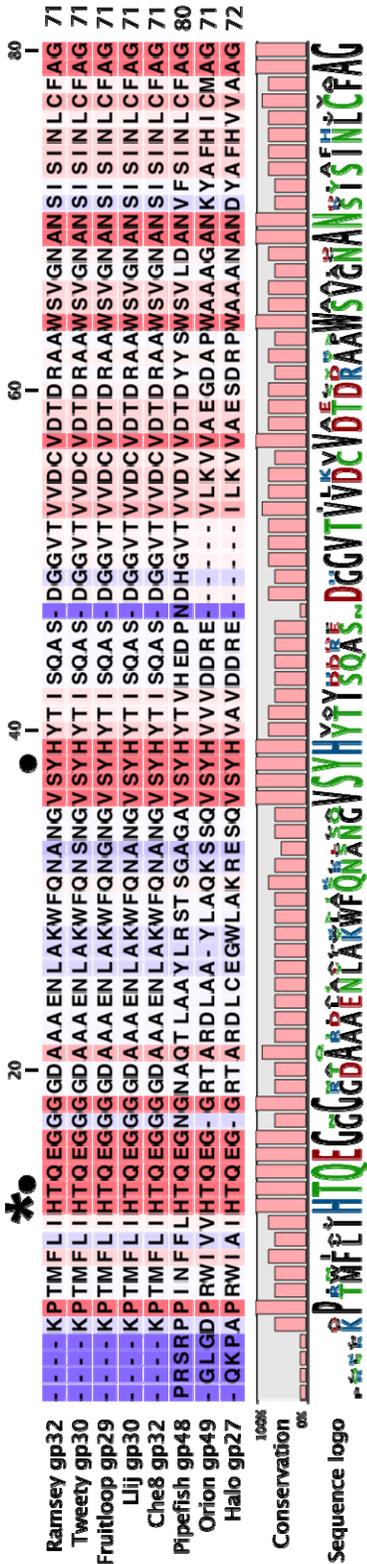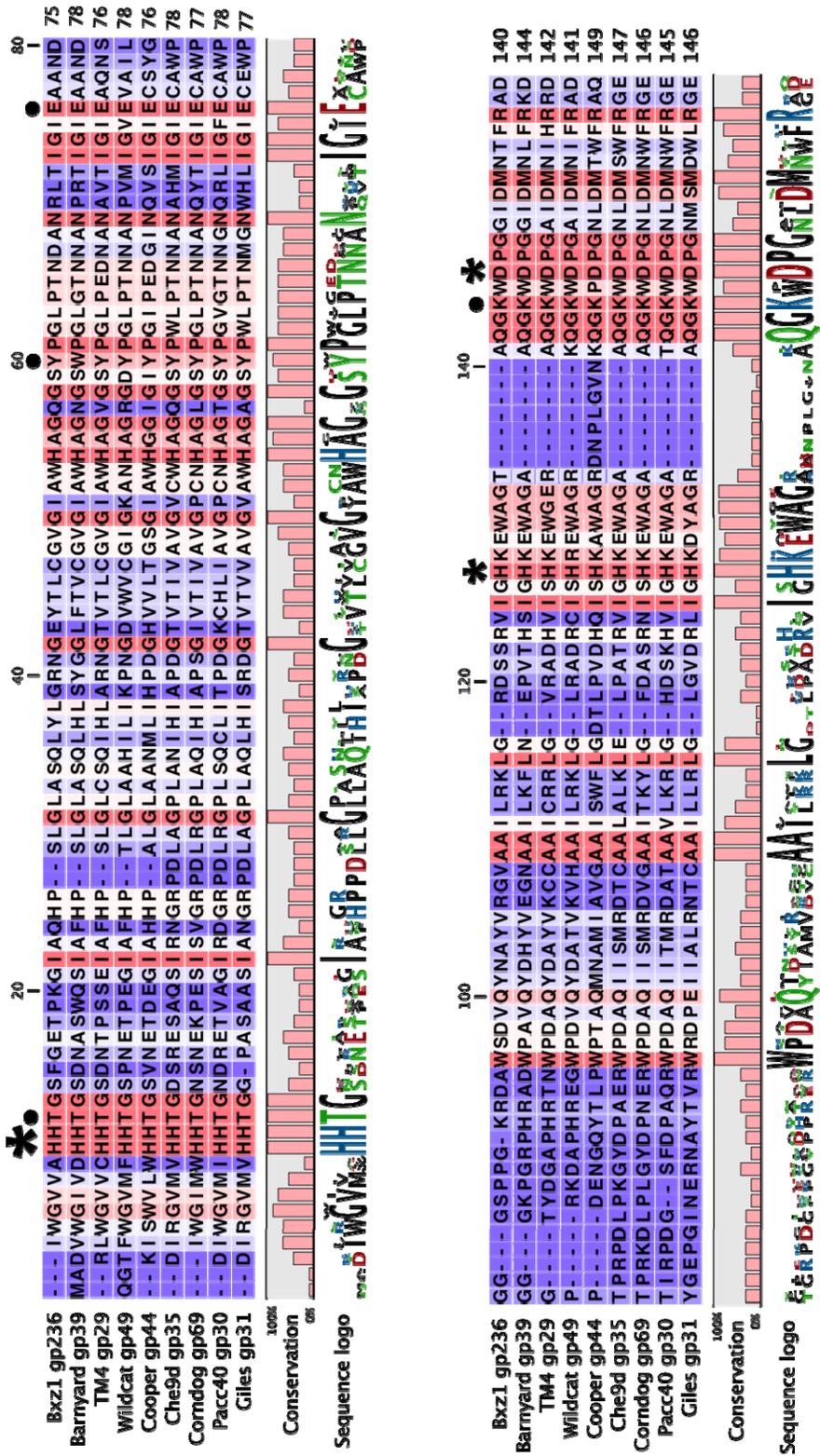
Figure 17. Ami1 and Ami2 domains

Figure 17: ClustalW alignments of the Ami1 and Ami2 domains. **A.** Alignment of unique Ami1 domains, with asterisks highlighting the Zn^{2+}-binding residues (His; His; Asp) and periods identifying other conserved residues seen in amidase family members implicated in catalysis (Tyr) and substrate-binding (Thr). **B.** Alignment of unique Ami2 domains, with asterisks highlighting the Zn^{2+}-binding residues (His; His; Asp); the first asterisk marks the two His residues, one of which is the first predicted Zn^{2+}-binding His. Periods identify other residues that may contribute to catalysis (Tyr or Glu, Lys) or substrate-binding (Thr).

2.3.2 Glycosidases

Glycosidase activity is the next most common hydrolytic activity identified in bacteriophage lysins, predominantly those infecting streptococcal species (Lopez et al., 1997; Pritchard et al., 2004). This is somewhat curious considering the wide array of glycosidases seen in host bacteria (Vollmer et al., 2008b) from which phage lysins are presumed to have been acquired via recombination. While there are two bonds known to be targeted by glycosidases, those found in phage lysins are primarily endo-N-acetylmuramidases (Loessner, 2005), which hydrolyze the ß-(1→4) linkages between MurNAc and GlcNAc (Figure 8). There are two types of muramidases: lysozymes, which release MurNAc fragments with a reducing end, and transglycosylases, which cleave the MurNAc-GlcNAc bond and concomitantly form a 1,6-anhydro ring on the MurNAc, leaving no reducing end (Figure 9). Several LysAs are predicted to possess lysozyme activity (EC 3.2.1.17), specifically those containing glycoside hydrolase family 25 (GH25) and glycoside hydrolase family 19 (GH19) domains, and recently a single LysA was identified as having transglycosylase activity.

2.3.2.1 Glycoside Hydrolase Family 19 Muramidases

Several LysAs possess domains that are identified as part of the lysozyme_like superfamily (cl00222), which broadly includes lytic transglycosylases, some lysozymes, some chitinases, and several glycosidases. In the LysAs these domains average 150 aa in size, and they identify as belonging to COG3179 and Glycoside Hydrolases Family 19 chitinases (cd00325, pfam00182), with e-values ranging from $1e^{-26}$ to $1e^{-10}$. By definition, chitinases (EC 3.2.1.14) digest polymers of chitin made up of repeating units of N-acetylglucosamine, which are not found in bacterial

PG.   However, GH19 chitinases have recently been identified in *Streptomyces* species and now in other actinobacteria (Kawase et al., 2004; Watanabe et al., 1999).   Currently, 43 phages are listed in the CaZy database as having a GH19 domain, the majority of which are found in mycobacteriophage or *Pseudomonas* phage.   However, only one phage chitinase – ORF187 of φRSL1, which infects the phytopathogenic bacterium *Ralstonia solanacearum* – has been appears to have hydrolytic activity though it has not been verified as a lysozyme (Yamada et al., 2010) and shows little sequence similarity to any LysAs. It was also previously noted that Bxb1 gp8 and L5 gp10 show similarity to chitinases (Kawase et al., 2004).

Two of the best-studied GH19 chitinases are the plant chitinase from *Carica papaya* (Huet et al., 2008) and *Streptomyces griseus* chitinase C (Watanabe et al., 1999). These proteins were aligned with the GH19 domains of mycobacteriophage LysAs (Figure 18), revealing that both of the catalytic Glu residues are conserved.   Several conserved motifs were previously identified as present across all GH19 chitinase family proteins (Huet et al., 2008; Udaya Prakash et al., 2010), and some of these overlap with motifs found in actinobacterial GH19 chitinases (Kawase et al., 2004). The first motif, containing one of the catalytic Glu residues, is partially conserved when compared with signature sequences, referred to as C1 (Kawase et al., 2004) and M2 (Udaya Prakash et al., 2010). The second conserved region, labeled as C2 or M1, begins and ends with two Tyr residues and is more strongly conserved.   While the full motifs are not observed, the most conserved residues for the motifs C3, C4/M4, and M5 are visible (Kawase et al., 2004; Udaya Prakash et al., 2010); the motif M3 is not seen (Udaya Prakash et al., 2010).

The GH19 domain is found in the center of LysAs in Cluster A and Sub-cluster B2 (Figure 13 C, E, I), and is also in the N-terminus of Che9d gp35 (Figure 13 G) and Wildcat gp49 (Figure 13 S).   Interestingly, only four GH19 chitinase domains have been predicted in

Mycobacterium species as identified in CaZy; one is a bacteriophage protein in *M. abscessus*, and the others are predicted chitinases in *Mycobacterium* sp. JLS, KMS, and MCS. When the GenBank database was examined with BLASTp, the sequence of GH19 in LysAs shows the greatest similarity to the chitinases in various species of *Pseudomonas*, a proteobacterium, rather than to those from *Mycobacterium* or to the numerous chitinases from *Streptomyces* (another member of the order Actinomycetales) that have been more extensively studied (Kawase et al., 2004; Watanabe et al., 1999). This may indicate intergeneric recombination, or genetic exchange between organisms belonging to different genuses or other higher taxonomic categories, a phenomenon that has been observed in other phage lysins (Sheehan et al., 1997). Still, the pervasiveness of the domain throughout LysAs suggests that the activity is highly effective for mycobacterial lysis.

Figure 18. Glycoside hydrolase family 19 domains

Figure 18: A ClustalW alignment of the LysA GH19 domains and regions of *Carica papaya* chitinase and *Streptomyces griseus* chitinase C proteins. The asterisks identify the two conserved catalytic Glu residues. Regions of sequence with significant similarity to signature chitinase motifs identified by Kawase et al. (C1-4) and Udaya Prakash et al. (M1-5) are identified. Periods indicate conserved residues that are found in the motifs listed above them; the overall sequence of these motifs is not conserved. More conserved residues are shaded red while less conserved residues are shaded blue.

2.3.2.2 Glycoside Hydrolase Family 25 Muramidases

GH25 enzymes are only known to possess lysozyme activity (Henrissat et al., 1995). Many glycosidases found in other phages, including those infecting *S. pneumoniae* (Hermoso et al., 2003) and *B. anthracis* (Porter et al., 2007), are classified as GH25 lysozymes. In mycobacteriophage LysAs, the majority of GH25 domains are in the center of the protein, with the exception of Myrna gp243, where the GH25 domain appears towards the C-terminus (Figure 13 O). GH25 domains are the largest domains seen in LysAs, with an average size of 200 aa, and all GH25-containing LysAs fall into the GH25_muramidase superfamily (cl10448), with e-values averaging $1e^{-04}$; an exception is Omega gp50, which shows significantly more similarity to the GH25 superfamily and specific conserved domains, with e-values ranging from $1e^{-22}$ to $1e^{-11}$.

The conserved domain subfamilies differ slightly; GH25 domains in Cluster E LysAs (Figure 13 M) are in the GH25_PlyB-like (cd06523) sub-family, a group whose prototype is the PlyB lysin that is active on *B. anthracis* (Porter et al., 2007). The Myrna (cluster C2) LysA, gp243 (Figure 13 O), contains a GH25_Lyc-like conserved domain (cd06525), while other Cluster C1 phages have Ami2 domains. In addition to those seen for the above GH25 LysAs, the conserved domains associated with Omega gp50 (Figure 13 M) include GH25_YegX-like (cd06524), GH25_LysA-like (cd06417), and smaller regions with similarity to GH25_LytC-like (cd06414), GH25_CH-type (cd06412), GH25_Cpl1-like (cd06415), and GH25_AtlA-like (cd06522). These sub-families encompass many other phage endolysins and host autolysins, including PlyB (Porter et al., 2007), Cpl-1 (Garcia et al., 1987), AtlA (Oshida et al., 1995), and LytC (Smith et al., 2000). A PSI-BLAST of the above LysA GH25 domains reveals homology to many hypothetical *M. tuberculosis* proteins and a few *Rhodococcus and Nocardia* proteins.

The GH25 domain of Omega gp50 shows high homology to diverse bacterial proteins from *Corynebacterium* and other Actinomycetales, as well as *Bacillus*, and *Clostridium* species.

The catalytic residues of GH25 lysozymes consist of an Asp and a Glu, with the Glu being part of a D-X-E motif. GH25 is proposed to function similar to other lysozymes, *via* a general acid/base catalytic mechanism using Glu as an acid proton donor and an Asp as a nucleophile/base (Hermoso et al., 2003). In the crystal structures of the *B.anthracis* phage lysin PlyB and pneumococcoal lysin Cpl-1, these catalytic residues have been identified as Asp6/Glu92 (Porter et al., 2007) and Asp10/Glu94 (Hermoso et al., 2003), respectively. In addition to the catalytic residues, two more Asps proposed to be part of the catalytic machinery are conserved throughout the GH25 family: Asp90 of the D-X-E motif and Asp171 in PlyB (Porter et al., 2007). Additionally, Tyr59 and Tyr125 of the pneumococcal phage lysin Cpl-1 interact with the MurNAc, and are also conserved across GH25 family members (Perez-Dorado et al., 2007). An alignment of representative LysA GH25 domains and the *M. tuberculosis* H37Ra hypothetical protein MtubH3_23920 (which has 46% sequence identity with the GH25 of Kostya gp33) shows a conservation of all of these residues (Figure 19) in the same order and approximate spacing relative to PlyB and Cpl-1.
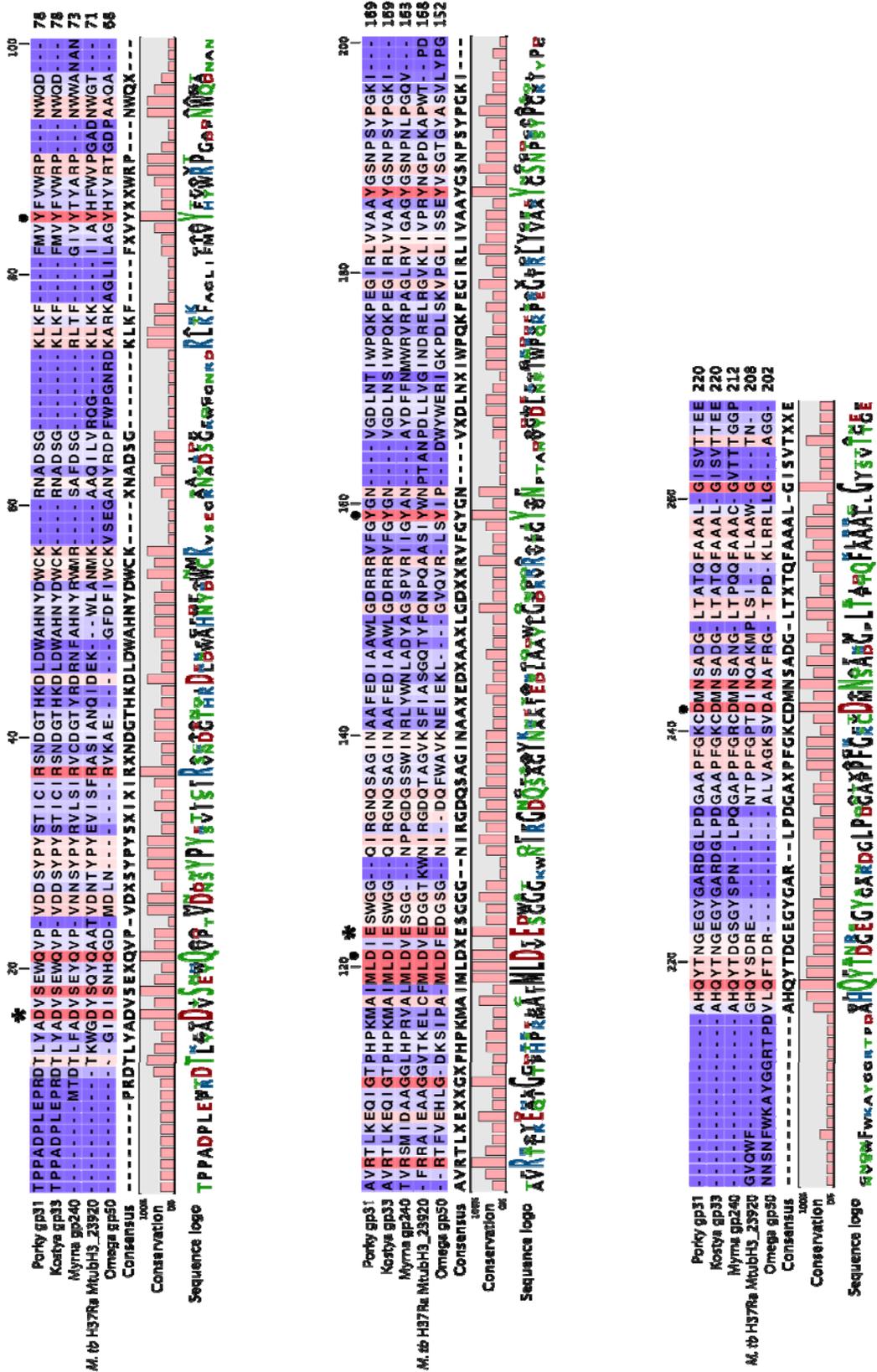
Figure 19. Glycoside hydrolase family 25 domains

105

Figure 19: A Clustal alignment of the LysA GH25 domains and a region of *M. tuberculosis* H37Ra hypothetical protein MtubH3_23920. The asterisks identify the catalytic Asp and Glu residues, the latter of which is in the conserved D-x-E motif. Periods indicate other conserved residues that are found in throughout members of cd00599 GH25_muramidase (Marchler-Bauer et al., 2009). More conserved residues are shaded red while less conserved residues are shaded blue.

2.3.2.3 Transglycosylase

A third glycosidase activity is predicted in the central domain of a single LysA, Hammer gp13 (Figure 13 L), which has conserved domain hits to Transglycosylase (pfam06737) and SLT (pfam01464), both of which are in the lysozyme_like superfamily (cl00222). Lytic transglycosylases (TG) are muramidases that target the same MurNAc-B-(1$\rightarrow$4)-GlcNAc bond as lysozymes, but instead of hydrolyzing the bond, they catalyze an intramolecular transglycosylation of the glycosyl bond onto the C6 hydroxyl group of the muramic residues to form a 1,6-anhydromuramic acid (Scheurwater et al., 2008). While rare in lysins – aside from several Gram-negative lysins, including R of phage λ (Bienkowska-Szewczyk et al., 1981) and KZ144 of *Pseudomonas* phage ϕKZ – transglycosylase motifs are commonly seen in the tip of phage tape measure proteins, including several in mycobacteriophage (Marinelli, 2008). The phages are believed to use the transglycosylase activity to pierce the bacterial cell wall during the injection of phage DNA into the host (Marinelli, 2008; Moak and Molineux, 2000, 2004).

In fact, the TG domain of Hammer gp13 shows 71% sequence identity with C-terminal regions in the tape measure proteins of Cluster D mycobacteriophages, suggesting that recombination between two mycobacteriophage genes of different functions may have created a TG-containing LysA or vice versa. The tape measure protein of Hammer (gp29) is similar to those in other Cluster A phages, and shows no significant similarity to Hammer gp13. The TG domain of Hammer also has similarity to several bacterial proteins, including an unknown secreted protein in *M. smegmatis* mc^{2}155 (37% identity, e-value = 2e^{-04}), the resuscitation-promoting factor RpfA of *Corynebacterium glucuronolyticum* ATCC 51866 (38% identity, e-value = 8e^{-04}), and the N-terminus of *Streptomyces sp.* AA4 transglycosylase domain protein

(44% identity, 4e-value = 4e^{-06}). Rpf proteins actively degrade PG to facilitate growth of dormant bacteria (resuscitation), and have been well-studied in *M. tuberculosis*, which encodes multiple Rpf proteins (Kana and Mizrahi, 2010), and Rpf-motifs have been identified in the tape measure proteins of mycobacteriophage (Marinelli, 2008; Pedulla et al., 2003).

A ClustalW alignment (Figure 20) of the transglycosylase regions of these three proteins and two mycobacteriophage tape measure proteins (PLot gp28 from Sub-cluster D and Pacc40 gp14 from Sub-cluster F1) shows a conserved Glu, which is the single catalytic residue in transglycosylases, as opposed to the Glu-Glu or Glu-Asp catalytic pairs found in most other lysozymes (Thunnissen et al., 1995). There are also several consensus motifs specific to the lytic transglycosylase families (Blackburn and Clarke, 2001). These include the Ser adjacent to the catalytic Glu, which together are identified as motif I and are observed in all lytic transglycosylases, as well as the later Gly and Gln in a G-[G/L]-[F/G]-Q, similar to the G-[L/I]-[M/W]-[Q/M] sequence of motif II. Interestingly, based on these motifs and other conserved residues of the four families of transglycosylases, the LysA and related TG domains from Figure 20 most closely resemble family 1 bacterial transglycosylases (Blackburn and Clarke, 2001) rather than family 4 λ phage transglycosylases. Additional residues in Hammer gp13 match those present in >80% of the sequences of sub-family 1C (Figure 20). In sum, the overall sequence similarity to other known transglycosylase domains and the conservation of catalytic Glu strongly support the classification of the central domain of Hammer gp13 as a transglycosylase, the first identified in any mycobacteriophage LysA.
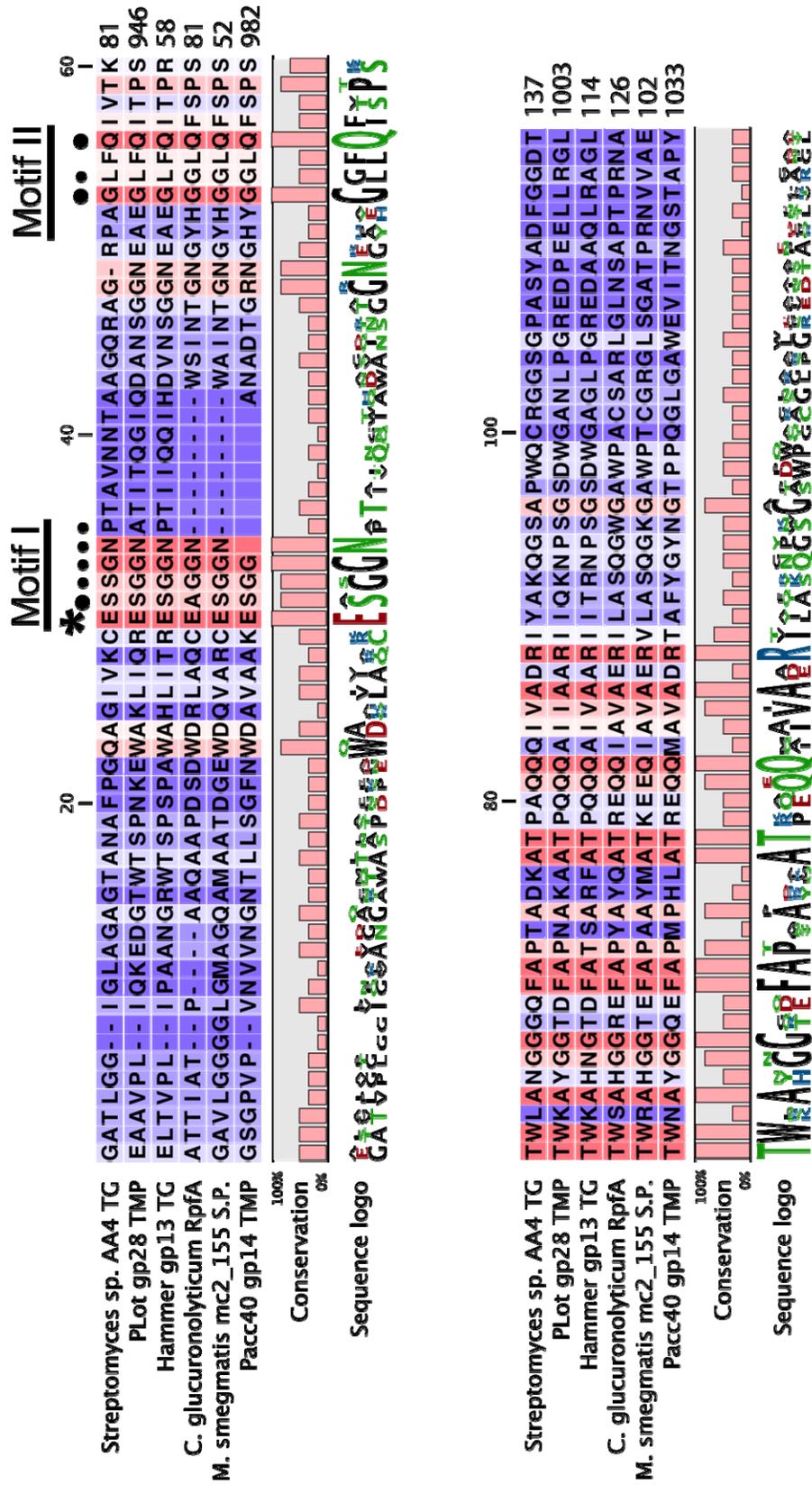
Figure 20. Transglycosylase domains

Figure 20: A Clustal alignment of the Transglycosylase domains from Hammer gp13 LysA, the tape-measure proteins PLot gp29 and Pacc40 gp14, the resuscitation-promoting factor RpfA of *Corynebacterium glucuronolyticum* ATCC 51866, the transglycosylase domain protein of *Streptomyces sp.* AA4, and a secreted protein of *M. smegmatis* mc^{2}155. The asterisk identifies the single catalytic Glu residue. Motifs characteristic of family 1 transglycosylases are labeled (Blackburn and Clarke, 2001). Large periods indicate conserved residues that are found in all members of transglycosylase family 1, specifically the E-S and G-X-X-Q regions in each motif. Smaller periods indicate the residues in Hammer gp13 that are conserved in >80% of the members of family 1C. More conserved residues are shaded red while less conserved residues are shaded blue.

2.3.3 Peptidase

There were relatively few peptidases found in the LysAs, compared to their prevalence in the few three-domain lysins from other phages (Cheng and Fischetti, 2007; Donovan and Foster-Frey, 2008; Navarre et al., 1999). The LysA peptidase domains average 90 aa in length and are all found at the N-terminus preceding an amidase domain – either Ami1 in Cluster G and Sub-cluster B1 LysAs (Figure 13 K) or Ami2 in Giles gp31 (Figure 13 J). These peptidases are predicted to belong to the peptidase family M23 (cl09532, pfam01551), a group of zinc metallopeptidases that include bacterial PG hydrolases such as lysostaphin and ß-lytic metallopeptidases (MEROPS). The structure of LytM, a lysostaphin-like peptidase from *Staphylococcus aureus*, has been solved (Odintsov et al., 2004). The conserved Zn^{2+} binding cleft contains four conserved residues: N117; H210; D214; and H293, which is the second His in the H-X-H motif commonly found in Zn^{2+} binding proteins. The M23 peptidase domains in LysAs do contain the zinc-binding H-X-H motif (Figure 21). Additionally there are conserved Arg, His, and Asp residues seen in LysA M23 domains; however, they neither follow the order seen in LytM nor share flanking sequence similarity, so it is questionable whether these are the conserved residues and whether M23 peptidases are active. The most closely related proteins outside of other LysAs are various M23 family peptidases, beginning with an M23 peptidase of *Corynebacterium jeikeium* ATCC 43734, with an e-value of $7e^{-25}$. The closest *Mycobacterium* hit is a peptidase in *Mycobacterium sp.* KMS with an e-value of $1e^{-11}$. There are also some peptidase motifs found in the tape measure proteins of phages, such as motif 3 mycobacteriophage TM4 (Piuri and Hatfull, 2006), but the M23 domains do not show significant similarity to any non-LysA mycobacteriophage proteins.
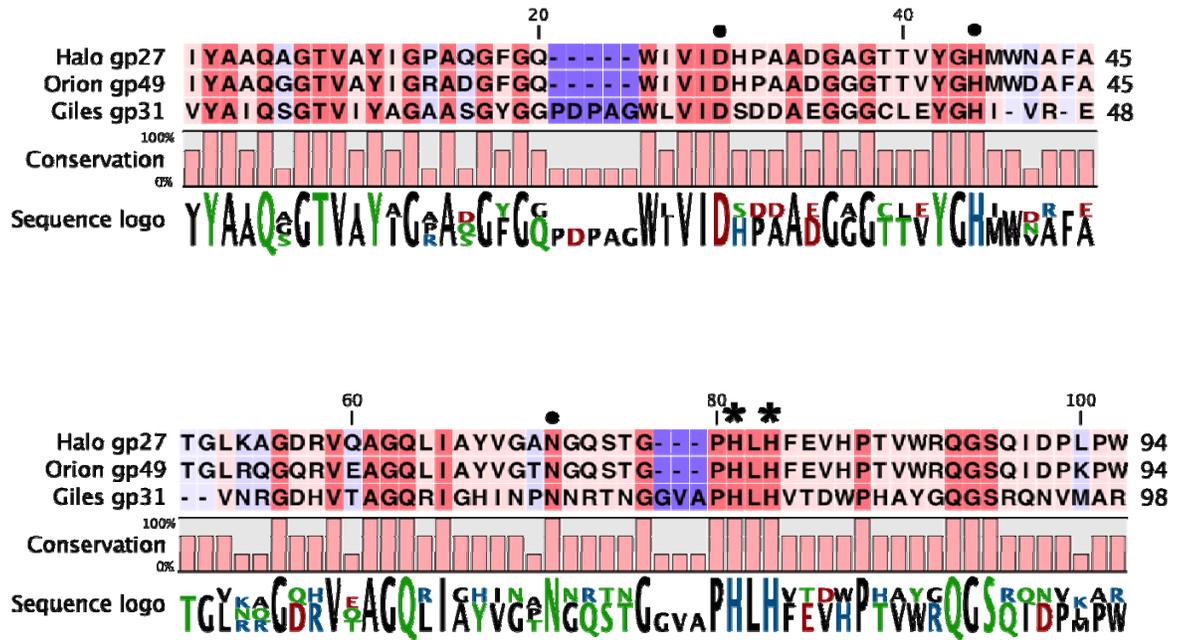
Figure 21. M23 peptidase domains

Figure 21: A ClustalW alignment of three representative M23 family peptidase domains from Giles gp31, Halo gp27, and Orion gp49. More conserved residues are shaded red while less conserved residues are shaded blue. The asterisks highlight the Zn^{2+}-binding H-X-H motif. Periods identify conserved His, Arg, and Asp residues (Odintsov et al., 2004).

2.4   THE MYSTERIOUS N-TERMINAL DOMAINS

The majority of LysA organizations (13/19; Figure 13 B, C, D, F, H, I, L, M, N, O, P, Q, R) possess one of four unrelated N-terminal domains, labeled N1 – N4. The LysAs that lack one of these domains either have a second predicted catalytic domain at the N-terminus (discussed above), have an N-terminal region that is unknown and does not match another identified LysA in a mycobacteriophage, thus not qualifying as a domain (Bxz2 gp11, Figure 13 E), or do not appear to have a third N-terminal domain (Barnyard gp39, Figure 13 A). There are no predicted signal peptides (predicted with SignalP Server, v. 2.0), signal-anchor-release (SAR) domains or transmembrane domains (predicted with TMHMM Server v. 2.0). A search for conserved motifs was conducted using Multiple Em for Motif Elicitation (MEME 4.4.01, http://meme.nbcr.net) to find the top three significant (lowest e-value) motifs of length no greater than 50 aa across the domains. The sequences included in the search were representative (*e.g.* the N2 domain of Bxz1 gp236 is identical or nearly identical to all other Sub-cluster C1 phage N2 domains, and so Bxz1 gp236 was chosen to be representative, so as not to skew the search results). Any sequences that were identified only in a subset of protein domains that were nearly identical – and thus did not represent a signature sequence found across varied proteins in that domain – were ignored.

N1 domains average 160 aa in size and are only paired with glycosidases or alone with a C-terminal domain in a two-domain LysA (Table 2). In the N1 LysAs possessing three domains, N1 is followed by GH19-C1 (Cluster A and Sub-cluster B2, Figure 13 C) or by TG-C1 (Hammer gp13, Figure 13 L). Otherwise N1 is paired with C3 to create the two domain lysins lacking a predicted catalytic domain in Clusters D and H1, as well as Che9c gp25 (Figure 13 Q). The N1

domains show limited sequence similarity to hypothetical *Streptomyces* and *Rhodococcus* proteins. A MEME search identified three conserved motifs found in all N1 domains in the same order (Figure 22 A, B).

Table 2. N-terminal domain characterization

Domain	Average Size (aa)	Associated Domains	Clusters, Sub-clusters, singleton phage	GenBank Database similarities		Motifs	Motif E-value
N1	160	*Lytic:* GH19, GH25, TG, and alone *C-term:* C1, C3	A1, A2 [Hammer gp13, Pukovnik gp11], B2, D, H1, Myrna gp243, Che9c gp25	*Rhodococcus/Streptomyces hypothetical proteins*	1	[VI][EN][PA][LMR][RGI]D[AT][DQ][SEI][GA]C[WY]T[PD]TNS[VM]A[TN]SNH[LN][SG][GA]TA[MI]D[LY]NW[NS]DHP[FY][QG][KV]	1.30E-162
					2	SP[KR]D[ES]MHF[QE][LM][GH][YF][AN][TE][GY][GH]N	4.10E-67
					3	GW[PR][ME]C[DN][ALT][ADS][ED][CL][DE][YW][TA][AKPT][VI]PGT[SNP][FVL][KTR]L[GP Q][IVFIL][RAQ][KR]G[QISP][LFNS][AVI]IL[KR Q][AS][FL][AIL][AWR][DR][LF]HN]	2.70E-120
N2	215	*Lytic:* Ami1, Ami2 *C-term:* C1, C2	B4, C1, F1 [Boomer gp32, Che8 gp32, Pacc40 gp30, Ramsey gp32], TM4 gp29	(with Bxz1) *Penicillium chrysogenum* Pc12g06060 (5e-22), Aspergillus fumigatus cons. Hypo. Prot. (2e-21), Tuber melanosporum Hypo.prot. 5e-17, other bacterial hypothetical proteins	1	ANR[KS][VD]P[ED]SL[KN]YPHD[AS][VE][GS][SDN]D[GY][DR]SVG[IY]FQQ[QR]	1.60E-74
					2	[WS][KNS][TAR][MQ][ADE][APV][GN][DK][L S][AC]Q[AKR][VI]Q[RQ]S[AG][FYV]PD[RA]YA[KQ][RW][WMV][ADP][EL]A	3.70E-38
					3	[WY][GP][TD][ITVJA][AD][CDR]MD[PL]A[KR][SA]A[AR][QM]FF[ET]R[LM]	8.10E-21
N3	150	*Lytic:* Ami1, Ami2 *C-term:* C2	B1, F1 [Fruitloop gp29, Lilj gp30, PMC gp30, Tweety gp30], Corndog gp69	Nothing significant	1	[GTW][YIK][AD][QN][GQV][RAD][GQR][AIS][YT][LP][TAH][ER][1AF][KD][AGS][RA][1AG][AY][TQ][IRV][KST][EGP][FYC][LF][DA][KI][YML][SWD][PVT]	2.30E-87
					2	[WV][RK][AR][KR][PS][GP][AN][SH][TG]D[W [LG][ND]I[CAF][WML]QQ[AR]P[NG][WE][PE]SA[DA][YA]	7.30E-51
N4	150	*Lytic:* GH19, GH25, and alone *C-term:* C1, PGBD (Brujita gp29)	A2 [Che12 gp11, D29 gp10, L5 gp10], E, Brujita gp29, Omega gp50	Nothing significant	1	[HQ][GH][PG]GGG[AV][NY]SH[MT][AW][CG][TE]L[DM][GT]M[DL][EP]	2.90E-49
					2	R[AT][DM][TW][VI][DGIN]RPS[Y][AG][ST][AT][WE][AS][YF][LR][PL][GD][PH][1K][VI]	1.10E-32
					3	Y[GD]GA[FR][AT]WN[N][DP][PR]L[FS][HT]D[CF][SW][GY]L[DV][AL][QK]	1.80E-33

115

A

Motif 1

Motif 2

Motif 3

B

Name	Combined *p*-value	Motif Location
Bxb1 gp8	2.10e-82	
Che9c gp25	3.19e-75	
Hammer gp13	5.20e-80	
Jasper gp10	1.88e-14	
KBG gp10	3.07e-73	
Konstantine gp39	7.48e-85	
Myrna gp240	8.82e-33	
PLot gp36	6.71e-87	
Predator gp30	1.24e-88	
Pukovnik gp11	7.57e-39	
Rosebush gp46	1.58e-75	
Skipole gp11	1.19e-85	

Motif 1 Motif 2 Motif 3

N1 Domain Motifs

C

Motif 1

Motif 2

Motif 3

D

Name	Combined *p*-value	Motif Location
Boz1 gp236	1.82e-56	
Chc8 g32	2.43e-49	
Cooper gp42	7.30e-39	
TM4 gp29	9.66e-51	
Pacc40 gp30	9.09e-61	
P. chrysogenum Pc12g06060	7.68e-67	
A. fumigatus hypo. prot.	2.21e-70	
T. melanosporum hypo. prot.	1.14e-62	

Motif 1   Motif 2   Motif 3

N2 Domain Motifs

E

Motif 1

Motif 2

F

Name	Combined *p*-value	Motif Location
Pipefish gp48	6.87e-33	
Tweety gp30	3.45e-79	
PMC gp30	1.34e-80	
Corndog gp69	2.48e-47	
Fruitloop gp29	1.34e-80	

Motif 1   Motif 2

N3 Domain Motifs

Figure 22. N-terminal domain motif maps

Figure 22: **A., C., E., G.** Sequence logos of motifs identified by MEME (http://meme.nbcr.net). In each column, all letters with observed frequencies greater than 0.2 are shown; less-frequent letters are not included in the RE (N-terminal Domain Table). **B., D., F., H.** Block diagram of the domain with blocks summarizing the locations of identified motifs and the combined *p*-value. The height of the block is proportional to the *p*-value; only blocks with a *p*-value > 0.0001 are shown, and overlapping blocks will only display the one with the lower *p*-value.

The N2 domains are the largest at around 215 aa and are only associated with amidases (Table 2). N2 precedes Ami1-C2 in Sub-cluster F1 (Figure 13 F), Ami2-C1 in Sub-cluster B4 and TM4 gp29 (Figure 13 D), and Ami2-C2 in Pacc40 gp30 (Figure 13 P). A BLASTp of several N2 domains found sequence similarity to various bacterial proteins and even several fungal proteins. Many were predicted peptidases, often M23/M37 family metallopeptidases similar to those found in the N-terminus of other LysAs; however, the region of similarity was always separate from the peptidase domain or any other identified regions. Three fungal proteins had significant similarity to the N2 domain of Bxz1: Pc12g06060 of *Penicillium chrysogenum* Wisconsin 54-1255 (e-value = $5e^{-22}$); a conserved hypothetical protein of *Aspergillus fumigatus* Af293 (e-value = $2e^{-21}$); and a hypothetical protein of *Tuber melanosporum* Mel28 (e-value = $5e^{-17}$). An analysis of the N2 domains and the fungal sequences by MEME revealed three conserved motifs, all identifiable in both LysA N2 domains and fungal proteins, and always in the same order (Figure 22 C, D).

N3 domains average 150 aa in size and are also only associated with amidases (Table 2). N3 is found with Ami1-C2 in Sub-clusters B1 and F1 (Figure 13 R) and with Ami2-C2 in Corndog gp69 (Figure 13 H). There is no significant similarity to any proteins in GenBank outside of the LysAs. Two significant motifs were identified by MEME (Figure 22 E, F); the first motif is repeated two or three times in these domains, while the second is only present once, but is between two of the repeated sequences. Repeated motifs in a binding domain are not uncommon, as seen with the choline-binding motifs in many autolysins and phage lysins of *Streptococcus* (Lopez et al., 1997), so this may support a binding role for the N3 domain.

N4 is similar to N1 in its association solely with glycosidases or its presence alone with a C-terminal domain (Table 2). N4 domains average 150 aa and are paired with GH19-C1 in Sub-

cluster A2 (Figure 13 I), GH25-C1 in Cluster E and Omega gp50 (Figure 13 M), and alone with

a C1 domain in L5 gp10 (Figure 13 N). Like N3, no significant matches to the N4 domain are

found using BLAST against the GenBank database. There is significant variation between the

different N4 domains in the presence, and location, of three motifs identified by MEME (Figure

22 G, H). The first and second motif are found in all N4 domains, but their order is variable; the

first motif precedes the second in Omega gp50, Kostya gp33, CJW1 gp33, and Brujita gp29, and

the reverse is seen in D29 gp10, Che12 gp11, and L5 gp10. The third motif is absent from

Omega gp50, Kostya gp33, and CJW1 gp32. Similar to the first motif in N3, the third motif in

N4 is also repeated twice, with one copy at the beginning of the domain and a second copy at the

end.

Most intriguing are the two organizations of LysAs that consist of two domains, neither

of which are predicted to be catalytic: either N1 or N4 paired with a C-terminal domain (Figure

13 B, I, N, O, Q; Table 2). In the absence of a predicted PG hydrolase, this could indicate a lytic

function for N1 or N4 that is, as yet, uncharacterized. While there are many LysAs with the N1-

C2 organization, the single two-domain N4 protein, L5 gp10, is unique in its close resemblance

to D29 gp10. The two have 81% sequence identity in their N4 domain and 46% identity in the

C1 domain, and they have extensive synteny along their genomes (Figure 14). Given the

presence of GH19 in the LysAs of similar phage, including Che12 gp11, D29 gp10, Bxz2 gp11,

and Pukovnik gp11, it appears that the GH19 domain was simply lost from L5 gp10. However,

there is experimental evidence that L5 gp10 is catalytically active within mycobacteria, although

the substrate is not known (Appendix A.3).

2.5   C-TERMINAL DOMAINS AND CELL WALL BINDING MOTIFS

All phage lysins characterized to date possess a C-terminal cell wall binding domain that is an important determinant of enzyme specificity (Loessner et al., 2002; Lopez et al., 1997). The three C-terminal domains seen in mycobacteriophage LysAs are expected to serve the same function (Table 3). The C1 domains are moderately divergent but share no similarity outside of LysA proteins. Three conserved motifs were identified by MEME (Figure 23 A, B), always in the same order. However, they are not evenly distributed; several C1 domains (Bxz1 gp236, CJW1 gp32, Cooper gp44, Porky gp31, L5 gp10, D29 gp10, and Rosebush gp46) lack the third motif. Additionally, Wildcat gp49 lacks the first motif, and D29 gp10 and Pukovnik gp11 only contain the second and third motifs, respectively.

The majority of C2 domains only show similarity to some predicted hydrolase and hypothetical *Rhodococcus* proteins, but the C2 of Corndog gp69 is similar, with an e-value = $1e^{-09}$, to the hypothetical protein Rv3594 of *M. tuberculosis* H37Rv and to a metallopeptidase from *Rhodococcus jostii* RHA1, with an e-value of $8e^{-04}$. The *M. tuberculosis* protein Rv3594 resembles a two-domain lysin, with an N-terminal amidase that is similar to LysA Ami2 domains and a C-terminus that resembles C2. The metallopeptidase of *R. jostii* RHA1 contains both an N-terminal M23 peptidase and central GH25 lysozyme domain in addition to a C-terminal region similar to Corndog gp69 C2. These two bacterial protein domains were included in the MEME search for motifs. Of three motifs found, none is present in all sequences (Figure 23 C, D). The first, and largest, motif is present in members of Sub-cluster F1 (Boomer gp32, Che8 gp32, Fruitloop gp29, Ramsey gp32, and Tweety gp30) as well as in Pipefish gp48 and Phaedrus gp42,

although the latter two display weak similarity (2.72 e^{-03} and 1.68e^{-02}, respectively), and this is the only motif they possess.   The second motif is in every domain except for Pipefish gp48 and Phaedrus gp42.   This motif is the only one present with two copies in the *R jostii* metallopeptidase. The third motif was only found in Corndog gp69, Che9d gp35, Pacc40 gp30, and *M. tuberculosis* Rv3594. These four domains all have the same configuration, lacking the first motif but possessing the second and third motifs, with the third motif preceding the second motif.

Table 3. C-terminal domain characterization

Domain	Average Size (aa)	Associated Domains	Clusters, Subclusters, singleton phage	GenBank Database similarities	Motifs	Motif E-value
C1	100	*N-term:* N1, N2, N4 *Lytic:* GH19, GH25, Ami2, TG, and alone	A, B2, B4, C1, E, Omega gp50, TM4 gp29, Wildcat gp49	Nothing significant	1 SR[SR][PIS][LYF][RA][TAH][PEL][GD]EGE[IV]DTxAG[ILF] 2 [NLS][TV]D[GA][NM][VL]H[EIV]Q[YLV]VE[ER][LQ]AR[AM][GY][DHV][DQR]P 3 FLSAL[TS][DP]AEQRELL[DN]L[AL]RQ[LQ][AW]	1.50E-37 2.40E-34 2.20E-28
C2	85	*N-term:* N2, N3 *Lytic:* Ami1,Ami2, GH19	B3, F1, and Corndog gp69	Nothing significant	1 WFPWDYF[TA]ERVNHWA[NA]GGKTEPEPPKVKRFPDDW[TS]DRE[IL][LA]VE[IT]LRQ[LQ]RGY 2 [IT]GWPQLGG[RK]T[LV]VDA[LV][AG][AE][IL]GEK 3 GKWDPGNL[DS]M[NDS]WFR[GA]E[VI][QAR]KD[MLT]	4.90E-153 3.50E-73 6.40E-25
C3	110	*N-term:* N1 *Lytic:* Ami1, Ami2, M23, and alone	B1, D, G, H1, Che9c gp25, Giles gp31	Peptidoglycan binding domain 1 proteins	1 [LK][DV][VA]DG[IS][VFR]G[PK]xT[AEV][AE][AK][LVI] 2 F[VT][YP][ST][HS][ED][DE]M[VL]KQ[LV]WEQ[AL]FGPQAKGW[PE][DAS]L[LF] 3 AEYGIVL[AT]G[EAQ]T[AT]PT[TAS]PP	9.70E-69 1.10E-40 1.90E-10

C1 Domain Motifs

C

Motif 1

Motif 2

Motif 3

D

Name	Combined *p*-value	Motif Location
Boomer gp32	2.04e-67	
Che8 gp32	7.22e-73	
Che9d gp35	4.99e-41	
Corndog gp69	2.62e-41	
Fruitloop gp29	3.01e-72	
Pacc40 gp30	7.45e-40	
Pipefish gp48	2.72e-03	
Phaedrus gp42	1.68e-02	
PMC gp30	1.17e-66	
Ramsey gp32	9.62e-56	
Tweety gp30	4.68e-71	
M. tuberculosis Rv3594	1.56e-33	
R. jostii metallopeptidase	4.80e-17	

Motif 1   Motif 2   Motif 3
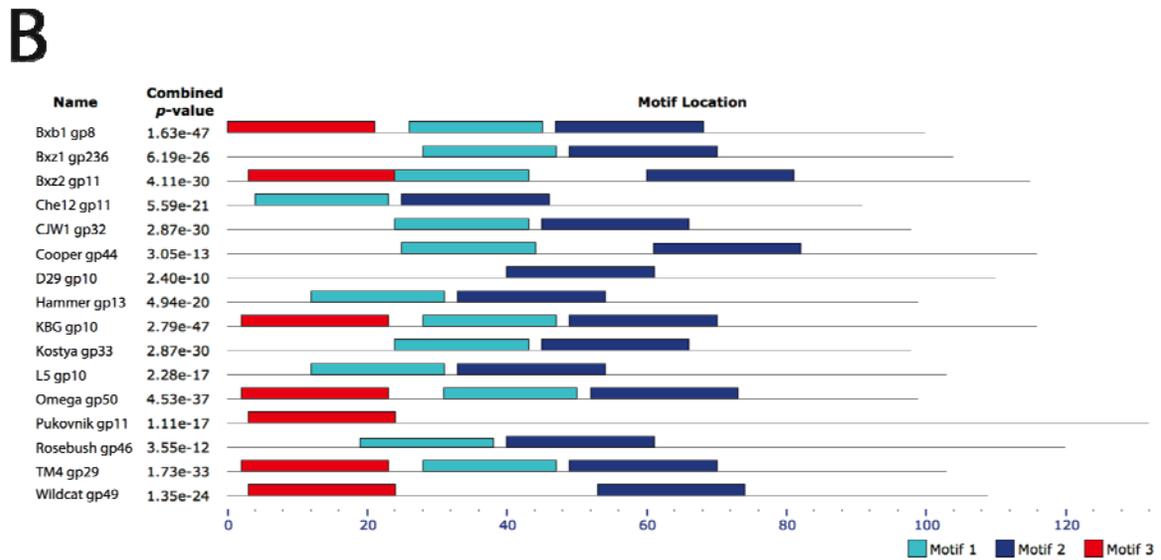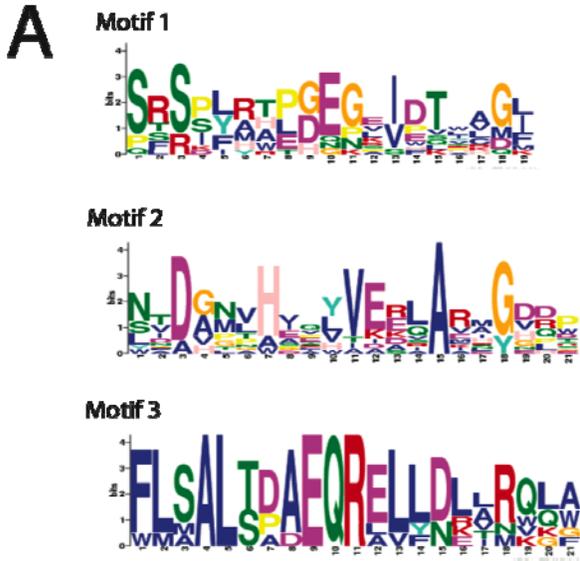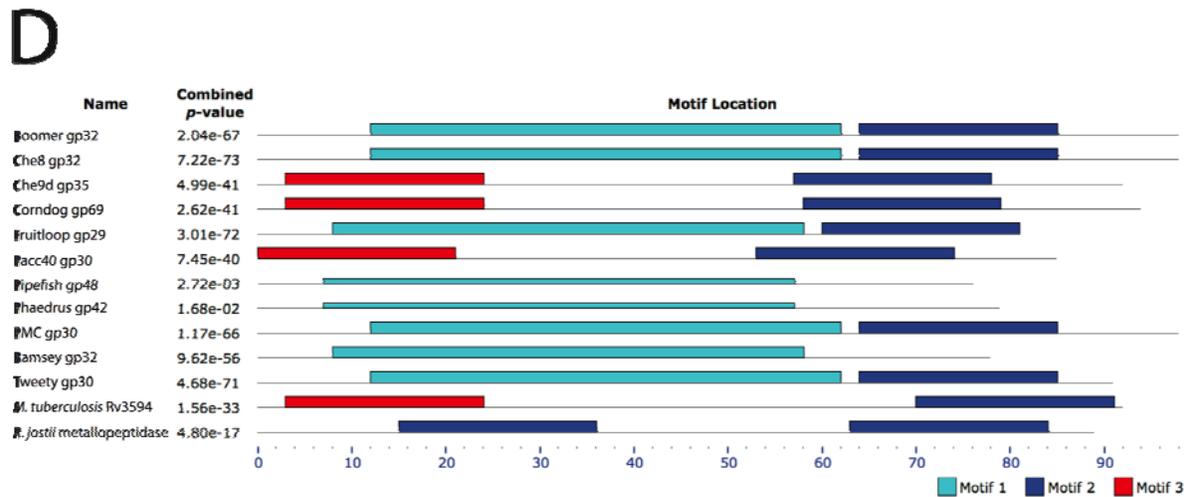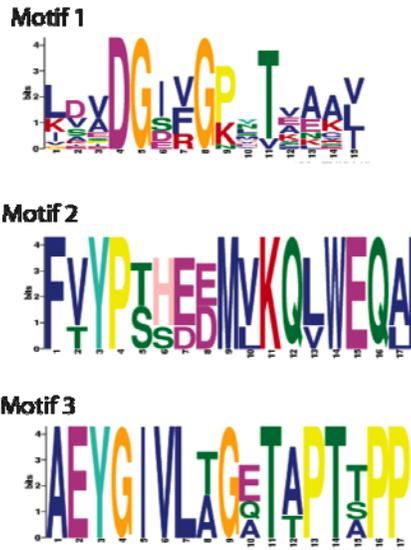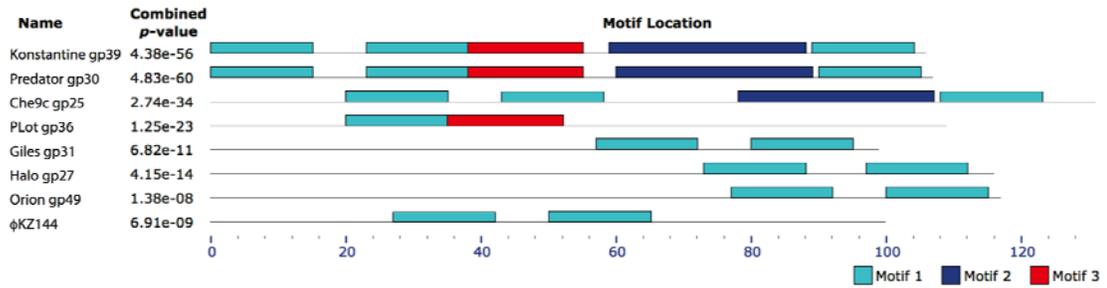
C2 Domain Motifs

Figure 23. C-terminal domain motif maps

Figure 23: **A., C., E.** Sequence logos of motifs identified by MEME (http://meme.nbcr.net). In each column, all letters with observed frequencies greater than 0.2 are shown; less-frequent letters are not included in the RE (C-terminal Domain Table). **B., D., F.** Block diagram of the C-terminal domains, with blocks summarizing the locations of identified motifs and the combined *p*-value. The height of the block is proportional to the *p*-value; only blocks with a *p*-value > 0.0001 are shown, and overlapping blocks will only display the one with the lower *p*-value.

The mycobacteriophage LysA C3 domain contains PG binding motifs (PG_binding_1 superfamily, cl03228; PG_Binding_1, Pfam01471), which supports a predicted role for these C-terminal domains in cell wall binding. The C3 domain is only found in LysAs with peptidase activity (Sub-cluster B1, Cluster G, and Giles gp31; Fig 1J, K) or in two domain LysAs with an N1 domain (Cluster D, Sub-cluster H1, Che9c gp25; Fig 1Q) or N4 domain (Brujita gp29; Fig 1B). Notably, this specific conserved domain has been described in two Gram-negative *Pseudomonas aeruginosa* phages, φKZ and EL, whose lysins KZ144 and EL188 have N-terminal PG_binding_1 domains and C-terminal lytic transglycosylase domains (Briers et al., 2007). A MEME analysis of the C3 domains and the N-terminus of KZ144 (Figure 23 E, F) reveals a single motif (motif 1) that is repeated two to three times in the majority of the domains (though only one copy appears in PLot gp36). This repeated sequence is a good candidate for a PG binding motif, similar to the choline-binding motif repeat identified in Cpl-1 and other streptococcal phage lysins (Hermoso et al., 2003). A second and third motif were also found between, or after, the repeats of the first motif. Both are found in Konstantine gp39 and Predator gp30, while only the second motif is present in Che9c gp25, and only the third motif is present in PLot gp36.

Three other LysAs contain regions in their C-terminus that do not match any other proteins. Similar to the C3 domain, Brujita gp29 (Figure 13 B) also has the same predicted PG binding motif as C3, but otherwise its C-terminal region does not have similarity to any LysAs or other proteins. Myrna gp243 does not have a predicted C-terminal domain (Figure 13 O), and indeed, the catalytic GH25 domain is predicted to extend to just 15 amino acids before the end of the protein, leaving little room for a binding domain. However, there is a central region of about 100 aa that shows no similarity to any other protein, yet is large enough to conceivably contain a

binding domain. One example of a central cell wall binding domain has been found in the lysin of the LambdaSa2 prophage from *Streptococcus agalactiae*, which has two Cpl-7 binding domains between an endo-peptidase and a glucosaminidase (Pritchard et al., 2007), so it is possible that a binding region exists in the center of the Myrna LysA.

Barnyard gp39 has a C-terminus with two predicted LGFP repeats (pfam08310; superfamily cl07065) (Figure 13 A). These 54 aa repeats were first identified in the N-terminus of PS1 (Adindla et al., 2004), a mycolyltransferase in *C. glutamicum* that bears homology to the mycobacterial antigen 85, albeit in a separate region from the LGFP repeats(Puech et al., 2000). LGFP repeats have also been identified in some mycobacterial proteins such as *M. tuberculosis* Rv3811, a protein that also includes an Ami_2 domain. Currently Barnyard gp39 is the only LGFP-containing protein listed under the virus taxonomy in InterPro IPR013207, although this may not consider possible prophages among the 195 bacterial entries.

2.6   CONCLUSIONS

In summary, I have characterized more than 70 of the complex mycobacteriophage LysA proteins, breaking them into 19 organizations comprised of 13 domains and several unmatched regions, as well as identified five predicted PG degrading activities associated with these proteins. While this has increased our understanding of the LysAs, the mysterious N-terminal domains are novel within the greater set of bacteriophage lysins and present an attractive object for further study. Currently, I am examining the activity of various LysAs with special attention to the function of the N-terminal domains. As the LysAs have evolved over great spans of time to efficiently and effectively lyse their target mycobacterial hosts, understanding these proteins on a molecular level may provide insight into the mycobacterial cell wall, the mechanism of mycobacteriophage lysis and evolution, and potential therapies or research tools for diseases like tuberculosis.

2.6.1   The Unique Structure of Mycobacteriophage LysAs

A modular theory for the construction of lysins was proposed some time ago (Diaz et al., 1990; Garcia et al., 1988; Lopez et al., 1997), and as more lysins are sequenced – including the mycobacteriophage LysAs – this theory is further validated. With more than 60 sequenced mycobacteriophage LysAs to compare, the extensive recombination is obvious and impressive, with 13 unique modules and 19 different combinations identified to date.

One thing that makes LysAs unique is their preponderance of three-domain organizations; other phage lysins possess only two domains, with some rare exceptions (Cheng and Fischetti, 2007; Donovan and Foster-Frey, 2008; Navarre et al., 1999; Pritchard et al., 2004). This would indicate some sort of advantage for phage infecting mycobacteria in having three domains, potentially tied to the unknown function of the N-terminal domains. PG hydrolase domains identified by BLASTp most commonly comprise the central domain, except in the rare LysAs with two identifiable lytic domains: Cluster G, Sub-cluster B1, Che9d gp35, Giles gp31, and Wildcat gp49. There are four different N-terminal domains of unknown function that are most frequently paired with a central catalytic activity, and almost every LysA has a C-terminal domain predicted to bind the cell wall. Interestingly, the majority of two-domain LysAs consist of one of the unknown N-terminal domains and a C-terminal domain. Only Barnyard gp39 is directly analogous in structure to other phage lysins with an N-terminal Ami2 activity and C-terminal region with LGFP motif repeats known to specifically bind structures found in mycobacteria or similar bacteria.

Bacteriophage evolution is thought to occur through both homologous and illegitimate or non-homologous recombination (Hendrix et al., 2000). Evidence in broad support of recombination leading to the variety of LysA organizations was provided above. For the ease of characterization, somewhat arbitrary boundaries based on sequence similarity were used, but the exact locations of recombination between LysAs cannot be readily identified. However, the predicted secondary structure could be analyzed in future studies, provided the predictions are accurate enough for a comparison and domains are neither too similar nor too distantly related. If recombination only occurs at domain boundaries, then the secondary structure similarities between two LysAs should change abruptly when there is a change in the domain. Alternately, if

the location of recombination is random with selection later preserving combinations that are functional, then secondary structure may be conserved across some of the above identified domain boundaries. The latter result would support the theory that mosaic phage genomes are in part the result illegitimate or non-homologous reocombination (Hendrix et al., 2000).

The consistent presence of a third N-terminal domain is not entirely unprecedented; several other lysins in streptococcal and staphylococcal phages have three domains in which two are PG hydrolases (Cheng and Fischetti, 2007; Donovan and Foster-Frey, 2008; Navarre et al., 1999; Obeso et al., 2008; Pritchard et al., 2004; Pritchard et al., 2007). As most lysins only contain one lytic domain, it is reasonable to suspect that lysins such as PlyGBS (Cheng and Fischetti, 2007; Pritchard et al., 2004) are exceptions to the rule. These bifunctional lysins are likely the result of the extensive recombination of lytic and binding modules seen throughout phage lysins, resulting in a dual-catalytic product that, at least in the circumstances where it was isolated, did not significantly decrease the fitness of the phage. It is interesting to note that in all of these bifunctional lysins, the N-terminal lytic domain is an endo-peptidase, specifically a D-alanyl-glycyl (CHAP) peptidase (Bateman and Rawlings, 2003) in all except λSa2. However, there are no predicted CHAP domains in any LysAs so far characterized. Additionally, all of the second lytic domains are amidases for lysins of Staphylococcal phages, and glycosidases for *Streptococcus* phages. The majority of the LysAs for which a lytic activity can be identified in the N-terminal domain are also predicted peptidases, although they are Zn^{2+} metallopeptidases (M23 family). Additionally the second domain in most of these LysAs is an Ami1 or Ami2. The exceptions to this are Che9d gp35 and Wildcat gp49, which have N-terminal GH19 and central Ami2 domains.

134

Most studies with purified bifunctional lysins found that only one domain was active; in all cases, the active domain was the N-terminal endo-peptidase domain, which had full or increased activity when expressed alone or with only the cell wall binding domain (Cheng and Fischetti, 2007; Donovan and Foster-Frey, 2008; Donovan et al., 2006a; Navarre et al., 1999). The central amidase or glycosidase domains, even if active in the full protein or *in vitro*, were largely inactive with live bacteria when expressed alone or with the cell wall binding domain. It would be interesting to test the separate activities in the LysAs with two identified catalytic domains; the N-terminal M23 peptidases have the conserved zinc-binding motifs required for activity, and the N-terminal GH19 domains of Wildcat gp49 and Che9d gp35 have the conserved Glu residues as well, suggesting that both domains may be catalytically active. It is less certain if the Ami1 and Ami2 domains are active based solely on sequence alignments; however, several LysAs containing these domains have demonstrated PG hydrolytic activity on zymograms (Appendix A.1).

2.6.2 The Prevalence of Various Peptidoglycan Hydrolytic Domains and Their Predicted Sources

The array of PG hydrolytic activities seen in LysAs is comparable to that seen in most phage lysins, and the bonds targeted are also those most commonly hydrolyzed by other phage lysins (Figure 24). The overall conservation of catalytic and other important residues suggests that these domains are capable of hydrolyzing PG, and several preliminary empirical studies support this (Appendix A.1). However, the origins and evolutions of the different domains are most intriguing. One curiosity is why the LysA proteins (and other non-mycobateriophage lysins) have such a variety of PG hydrolytic activities when all have the same goal with minor

differences in targets. Given the apparent inability for the host to develop resistance to the catalytic activity of a lysin, it is unlikely that the variety is an evolutionary insurance against the generation of a mutant host population that is resistant to a particular lysin. Alternatively, it is possible that the different activities address the minor variations in PG between different hosts or during different stages of host growth. For example, the degree of cross-linking increases as a bacterium enters stationary phase, and thus lysins with peptidase activity may have an advantage in lysins such a host.

Figure 24. Mycobacterial peptidoglycan structure and lysin targets

Figure 24: Diagram of mycobacterial PG and enzyme targets. A diagram of chemotype A1γ PG including modifications seen in mycobacteria (glycolation of muramic acids and amidation of D-Glu and m-DAP). In addition to the cross-linking bond between m-DAP and D-Ala, mycobacteria also have direct m-DAP – m-DAP links (dotted line). Possible hydrolases are listed with an arrow to their target bond. Those with activities identified in LysAs are in red, and the corresponding activity listed in parentheses below.

The predicted amidase domains (Ami1 and Ami2) appear to be unrelated (Figure 16), showing no similarity even after three iterations of PSI-BLAST, despite both being identified as belonging to the Amidase_2 superfamily. These domains appear to have different origins, suggesting two independent acquisition events in the past; Ami2 shows similarity to mycobacterial proteins, but Ami1 only weakly resembles proteins from *Corynebacterium* and *Propionibacterium*. Alignments of Ami1 and Ami2 domains revealed several conserved residues, including the essential Zn^{2+}-binding residues: His, His, and Asp or Cys (Cheng et al., 1994; Kerff et al., 2010; Low et al., 2005). Additionally, other important residues implicated in catalysis and substrate-binding (Kerff et al., 2010) are nearly all conserved in both Ami1 and Ami2. The retention of these important residues—despite their divergent sequences—suggests that both Ami1 and Ami2 are capable of hydrolyzing the N-acetylmuramoyl-L-alanine amide bond (Figure 24). Further, there is also empirical support for the activity of Ami-containing LysAs, although it has not been determined if the activity is specifically amidase. Purified Ami1-LysA Che8 gp32 and Ami2-LysAs Bxz1 gp236 and Corndog gp69 show PG hydrolytic activity on zymograms (Appendix A.1).

The presence of Glycoside Hydrolase family 19 chitinase domains in LysAs is interesting, considering both the primary activity of GH19 hydrolases and their seeming rarity in the mycobacterial genome and in known phage lysins. Chitinases generally hydrolyze the ß-1, 4-linkages in the polymers of N-acetylglucosamine that make up chitin, a component of fungal cell walls that is not present in bacteria. GH19 chitinases have been found in *Streptomyces* (Watanabe et al., 1999) and subsequently other actinobacteria (Kawase et al., 2004), and due to their similarity to the class IV chitinases found in higher plants, it is proposed that the presence of GH19 is the result of horizontal gene transfer between *Actinobacteria* and higher plants, rather

138

than from another bacterial source (Udaya Prakash et al., 2010). While bacteria can use chitin as an energy and carbon source (Cohen-Kupiec and Chet, 1998), some chitinases have shown activity characteristic of muramidases like GH25 (Flach et al., 1992; Jolles and Jolles, 1984), and they share structural similarity with lysozymes (Holm and Sander, 1994). As phages are not likely to encounter chitin in any situation where their lysin would also be needed, we expect that the GH19-like domains in LysAs have lysozyme-like activity. The conservation of predicted catalytic Glu residues, as well as sequence similarity with several GH19 chitinase-specific motifs (Kawase et al., 2004; Udaya Prakash et al., 2010), strongly suggests that these domains are catalytically active. The GH19-containing LysA D29 gp10 has also shown activity on zymograms (Appendix A.1).

GH25 lysozymes are common both in bacterial cell wall remodeling enzymes (Holtje and Fastrez, 1996) and in many phage lysins (Hermoso et al., 2003; Porter et al., 2007), and so their presence in LysAs is not surprising. Of the four families of lysozymes with known PG hydrolase activity, GH25 belongs to the *Chalaropsis* lysozyme family (Ch-type), which is distinct from the other three families prototyped by hen egg-white (HEWL, c-type), goose egg-white (GEWL, g-type) and T4 lysozyme (T4L, t-type). All of these have a Glu-Asp acid/base catalytic mechanism, but only HEWL, GEWL, and T4L appear to have diverged from a common ancestor (Weaver et al., 1984). While these latter three families share a common core of two α-helices and a three-stranded β-sheet (Monzingo et al., 1996), the Ch-type lysozymes show a completely different tertiary structure with a β/α barrel fold as seen with *Streptomyces coelicolor* cellosyl (Rau et al., 2001). Thus while the target and catalytic mechanism might be the same, the structure distinguishes GH25 lysozymes from the GH19 chitinases. Again, the notable aspect of this domain in LysAs is the seeming evolutionary divergence from a common amidase ancestor.

The GH25 of Omega gp50 appears much more closely related to known GH25-containing proteins, while other LysA GH25 domains in Cluster E and Myrna gp243 domains are significantly different from that in Omega gp50, implying a separation far in the past, if not two separate instances of acquisition.

The transglycosylases share a lysozyme-like fold and target the same bond as muramidases (Figure 9), with the only difference in activity being the enzymatic mechanism – cleavage results in formation of 1,6-anhydromuramyl residues rather than muramic acid residues with reducing ends (Scheurwater et al., 2008). The reason for this difference is attributed to the presence of a single catalytic Glu instead of a Glu-Asp or Glu-Glu pair found in other lysozymes. Transglycosylase domains have been identified in phages for some time, but most are present as part of the tape measure protein and may be involved in phage entry (Lehnherr et al., 1998; Marinelli, 2008). The precise function of these tape measure protein-associated hydrolases has not been definitely shown. However, they may facilitate penetration of the cell wall during phage DNA injection, and possibly stimulate bacterial growth, similar to the Rpf proteins (Marinelli et al., 2008; Moak and Molineux, 2000, 2004; Pedulla et al., 2003; Piuri and Hatfull, 2006). Hammer gp13 is also the first transglycosylase identified in any mycobacteriophage LysA, and the strong similarity between the TG of Hammer gp13 and the portions of tape measure proteins of Cluster D phage argue for a recombination event within mycobacteriophage genomes. It is curious that more transglycosylases are not seen in phage endolysins but rather are restricted to the tape measure proteins. Most transglycosylase lysins are found in phages that infect Gram-negative hosts, including the λ R endolysin (Bienkowska-Szewczyk et al., 1981), and the *Pseudomonas* phage PHIKZ144 lysin (Briers et al., 2007); there are a few putative transglycosylase lysins in some Staphylococcal and *Lactococcal* phage as well (InterPro

140

IPR008258). Based on the general distribution of transglycosylases in phages infecting Gram-positive and Gram-negative bacteria, and the novelty of finding a transglycosylase domain in any mycobacteriophage LysA, transglycosylase activity may be ill-suited – but still viable – for mycobacteriophage lysis.

The presence of an M23 family peptidase in mycobacteriophage lysins is initially confusing. The best-studied M23 peptidases, lysostaphin (Heinrich et al., 1987), ALE-1 (Lu et al., 2006), and LytM (Odintsov et al., 2004), prefer Gly_5 peptide bridges and hydrolyze Gly-Gly bonds, which are not seen in *Mycobacterium* PG. However, neither are Gly_5 peptide links seen in *Bacillus subtilis*, but this bacterium also has lysostaphin homologs (Smith et al., 2000) and type A1γ PG, the same type as mycobacteria (Schleifer and Kandler, 1972). Based on alignments (Spencer et al., 2010), structural similarity to other peptidases (Odintsov et al., 2004), and activity in *B. subtilis* (Horsburgh et al., 2003), it is now thought that M23 family peptidases can act as L-Ala-D-Glu endopeptidases. If the M23 peptidase domains are catalytically active in LysAs, we expect that they have this specific endopeptidase activity (Figure 24). Indeed, many recent structural studies of metallopeptidases have discovered similar folds in peptidases based around a Zn^{2+} binding site that, nonetheless, have distinctly different targets (Bochtler et al., 2004; Hooper, 1994). Further, metallopeptidase motifs have been found in the tape measure proteins of several mycobacteriophages, and PG hydrolytic activity was demonstrated on zymograms (Piuri and Hatfull, 2006); however, there is no significant similarity seen to the M23 domains in LysAs.

2.6.3 Speculations on the Possible Functions of the N-terminal Domains

The vast majority of LysAs, both two- and three-domain, contain N-terminal domains of unknown function. There are several possible functions for these four different domains, none of which need be mutually exclusive. The existence of LysAs with two catalytic domains suggests that these N-terminal domains may have a lytic function, potentially with activity that is specific to mycobacteria, similar to what has been observed for LysB (see Chapter 3). N2 and N3 are always present in front of a PG hydrolase, so a lytic role may be more likely with the N1 and N4 domains, which are seen in LysAs without any other predicted PG hydrolytic activity (Figure 13 B, N, Q). Indeed, lytic activity has been observed with D29 gp10, Kostya gp33, and L5 gp10 upon endogenous expression in *M. smegmatis* (Appendix A.3). All of these proteins share an N4 domain, and L5 gp10 is a two-domain LysA lacking any other predicted activity.

No transmembrane domains have been observed in these domains, making it unlikely that they function as SAR domains or holins. However, in the case of N3 a potential role has already been identified (Catalao et al., 2010). Very recently, Catalao *et al.* identified a role for gp1 of Ms6, a 77 aa protein immediately upstream of the gp2 LysA in an overlapping alternate reading frame. Gp1 acts as a LysA chaperone that uses the sec transport system to translocate Ms6 gp2 across the membrane into the periplasm. Here, Ms6 gp2 appears to be inactive until the membrane is depolarized, a function theoretically carried out by the Ms6 holin, gp4, similar to the manner in which pinholins like S^{21} activate SAR-endolysins (Xu et al., 2004). The first 60 amino acids of Ms6 gp2 are necessary and sufficient for gp1 binding and export; this may represent a role for the N3 domain. According to a BLASTp search, Ms6 gp1 homologs are also found upstream of the LysA proteins of several mycobacteriophages in Sub-cluster F1 (Fruitloop, Tweety, Llij, Che8, Pacc40, Ramsey, and Boomer) with e-values $< 1e^{-15}$. The LysAs

of these phages have N2 and N3 domains (Table 2), and so this may be the role of the N2 and N3 domains; however, this does not account for the full size of the domains, which average 215 and 150 residues, respectively (Table 2). Experimental evidence may indicate a greater similarity to SAR domain lysins than expected; in Corndog gp69 (N3-Ami2-C2), zymogram assays found highly active degradation products that were discovered to have lost most of the N3 domain (Appendix A.1). Similar fragments were found in zymograms with Bxz1 gp236 (N2-Ami2-C1) (Appendix A.1), but the loss of the N2 domain has not been verified. Export by a chaperone is not needed for *in vitro* hydrolytic studies of purified LysA, and so this domain may have been unnecessary or even inhibitory. It is even possible that, in mycobacteriophage infections, activation of the LysA after membrane depolarization results in the cleavage of this region or some other conformational change that allows for increased lytic activity.

There are still a number of other possible functions for the N-terminal domains that should be kept in mind. It has been suggested that, in addition to the host specificity conferred by the cell wall binding domain, endopeptidases may also confer specificity, since they target the peptide cross-links, a structure that is more specific to a bacterial species (Lu et al., 2006). Such peptidases are found in the N-terminus of several bifunctional lysins and a few LysAs. The N-terminal domain may require binding to specific substrates to better position the protein in relation to the substrate and could even sterically inhibit hydrolytic activity until it has bound, similar to the role observed for several other phage lysin cell wall binding domains (Low et al., 2005). The zymogram studies that have so far been performed on several LysAs have used *Micrococcus luteus* PG as a substrate (see Appendix A.1, Payne et al., 2009). In addition to lacking unique molecules like arabinogalactan and mycolic acids, *M. luteus* belongs to PG group A2, which has an interpeptide bridge in its PG cross-links. The A1γ group, containing

Mycobacterium and *Bacillus* species and many Gram-negative bacteria, has directly cross-linked *meso*-diaminopimelyl-containing PG. If the activity is specific to the peptide cross-link, or if hydrolytic activity is inhibited until the N-terminal domain is allowed to bind its mycobacteria-specific target, then the absence of this binding substrate could be limiting the activity seen in the zymograms. In assays with other lysins, a loss of the C-terminal domain could result in an increase in increased or broader hydrolytic activity (Becker et al., 2009; Cheng and Fischetti, 2007; Donovan et al., 2006a; Gaeng et al., 2000; Horgan et al., 2009; Low et al., 2005). This role may be supported for the N2 and N3 domains by zymography studies (Appendix A.1), although there is already evidence that these domains may serve a function in export (Catalao et al., 2010).

Another possibility is that the LysA is proteolytically processed to produce two proteins, one being the N-terminal domain and the other now resembling the canonical two-domain PG hydrolytic lysins. This has been observed for the *Staphylococcus aureus* autolysin Atl, which is processed into two functional proteins, an amidase and a glucosaminidase (Oshida et al., 1995). Alternatively, this second protein could also act as a chaperone for the catalytically active lysin, as was recently described for the Ms6 gp1 protein (Catalao et al., 2010), except that the chaperone is encoded as a part of the LysA. However, while there are homologs of this protein in several LysAs of phages in Sub-cluster F1, there is no similarity between Ms6 gp1 and any N-terminal domains.

Another alternative is that these N-terminal domains act to promote lysis in some way other than direct PG hydrolysis. It is possible that these domains act on another substrate, such as the arabinogalactan, but unless this is combined with another PG hydrolytic activity, it is not likely to result in efficient lysis. Hydrolyzing the arabinogalactan from the MurNAc residues of

the PG should have a similar effect to the release of the mycolic acids from arabinogalactan by LysB (Chapter 3), which, when lacking the PG hydrolase activity associated with the LysA protein, is insufficient to promote lysis (Marinelli et al., 2008; Payne et al., 2009). Since N1 and N4 are found in LysA proteins lacking any other predicted hydrolytic activity, it is less likely that they would have this activity, at least on PG incorporated into the cell wall.

It is still possible that these domains interact with other host proteins to indirectly cause PG hydrolysis. It has recently been observed in *S. pneumoniae* that the host autolysin, LytA, can contribute to, and even replace, the function of the lysin in SV1 phage infection (Frias et al., 2009). It is believed that the depolarization of the cell membrane by the holin protein results in the activation of LytA, which is present, but held inactive within the PG. However, while an SV1 lysin deletion mutant was able to propagate with the aid of LytA, mycobacteriophage Giles*ΔlysA* was unable to form plaques without complementation (Marinelli et al., 2008), suggesting that there may not be a convenient mycobacterial autolysin available for mycobacteriophage.

The N-terminal domains may also target the PG and other cell wall precursor components before they are assembled into the cell wall structure, or they may target the cell wall synthesis proteins themselves, as seen with the lysins of ϕX174 and phage Qβ (Bernhardt et al., 2001a; Bernhardt et al., 2001b).   Interfering in cell wall synthesis at many steps can be lethal, and the target need not be the PG or the enzymes that synthesize it; mycolic acids and arabinogalactan are also essential for mycobacterial survival. For example, the common anti-tuberculosis drug isoniazid binds InhA to halt mycolic acid synthesis, quickly leading to cell lysis (Quemard et al., 1995). However, the host is able to mutate its proteins to prevent such binding, an adaptation that is creating isoniazid-resistant *M. tuberculosis*. Further, being intracellular, this system

would not require a holin, and so theoretically the timing of phage lysis would not be determined by the holin. These are both counter to the design of lysins, which, at the optimal time in phage production, as dictated by the holin, target and degrade an essential and complex structure that is the product of dozens of enzymes and largely invariable throughout prokaryotes. In fact, the inability of bacteria to develop resistance to lysins is one reason that lysins are being investigated as a new class of antimicrobials (Borysowski et al., 2006; Fischetti, 2006).

Lastly, the LysA N-terminal domains may also have functions that more indirectly, or even completely unrelated, to lysis. The classic example of this is the T7 phage lysozyme. In addition to being an amidase that degrades the PG, the T7 lysozyme binds the T7 RNA polymerase to inhibit transcription late in infection (Moffatt and Studier, 1987). However, there are not many lysins known to have extra functions like this, and T7 lysozyme, like many Gram-negative phage lysins, does not have a modular configuration, but rather is more globular in structure (Cheng et al., 1994). Another secondary function characterized in a few phage endolysins is an ability to permeabilize the membrane *via* a short sequence similar to cationic antimicrobial peptides (Epand and Vogel, 1999). Such activities have been identified in the C-terminus of T4 lysozyme (During et al., 1999) and in the lysin of a *Bacillus amyloliquefaciens* phage (Morita et al., 2001a; Morita et al., 2001b). These sequences make it possible to permeabilize a Gram-negative membrane through exogenous application of the lysins (Orito et al., 2004). There is little reason to suspect that a mycobacteriophage would encode a lysin that could kill the host exogenously; however, it is still possible that such a region exists that would contribute in other ways to the disruption of the cell envelope.

146

2.6.4 The Potential for Novel Binding Motifs in LysA C-terminal Domains

The C-terminal domains of the mycobacteriophage LysA proteins may present new binding motifs unique to mycobacteria or closely related genera. Lysin cell wall binding domains often target molecules that are unique to the host bacteria (Hermoso et al., 2003; Loessner et al., 2002; Lu et al., 2006). For example, phage lysins Cpl-1 and Pal bind the teichoic acids of *S. pneumoniae* and possess within their C-terminus, the same repeated choline-binding motifs that target teichoic acids as are seen in the host autolysin, LytA (Lopez et al., 1997). The only LysA with a binding motif homologous to a known binding region in a related bacterium is in Barnyard gp39, which has the LGFP motif. This 54 aa repeat is found in several hypothetical *M. tuberculosis* proteins, as well as in the *Corynebacterium* PS1 cell surface protein, which resembles the mycolyltransferase *M. tuberculosis* antigen 85 complex proteins (Adindla et al., 2004). The long arabinogalactan chain covalently attached to the mycobacterial or corynebacterial PG would seem a reasonable binding substrate for LGFP motifs. While there is no indication as to the binding nature of the C1 domain, the similarity between several C2 domains and *M. tuberculosis* H37Rv R3594 and the metallopeptidase of *R. jostii* RHA1 may represent a novel binding motif for *Mycobacterium* and related bacteria. However, an interesting correlation was observed when the exogenous lytic ability of several purified LysAs was tested on *Propionibacterium acnes*, a Gram-positive bacterium in the order Actinomycetales (Appendix A.2). *P. acnes* does not have A1γ-type PG and lacks mycolylarabinogalactan (Kamisango et al., 1982), and yet it was quickly lysed by Bxz1 gp236 (N2-Ami2-C1) and D29 gp10 (N4-GH19-C1). These two LysAs share no domains except for C1; the other LysAs tested included Che8 gp32 (N2-Ami1-C2) and Corndog gp69 (N3-Ami2-C2), neither of which showed strong lytic

activity. Therefore it is possible that the C1 cell wall binding domain is less specific, allowing Bxz1 gp236 and D29 gp10 a broader range of targets.

The PG binding domain of C3, specifically PG_binding_1 (Pfam01471), is not present in many other phage lysins, but has been characterized in φKZ and EL. These phages infect the Gram-negative *Pseudomonas aeruginosa*, and the lysins KZ144 and EL188 were able to lyse a broad range of outer membrane-permeabilized Gram-negative bacteria, including other *Pseudomonas*, *E. coli*, and *Salmonella typhimurium* (Briers et al., 2007). In addition, large amounts of KZ144 were able to lyse the Gram-positive *B. subtilis*, but not other Gram-positives such as *Micrococcus lysodeikticus* or *Staphylococcus aureus*. The common feature of Gram-negative bacteria and *B. subtilis* is their PG chemotype (Schleifer and Kandler, 1972). All Gram-negative bacteria are chemotype A1 (Schleifer and Kandler, 1972), and KZ144 specifically bound the directly cross-linked N-acetylated chemotype A1γ (Briers et al., 2007), which is present in *P. aeruginosa, E. coli,* and *B. subtilis* (Schleifer and Kandler, 1972). Thus, the PG_binding_1 domain is implicated in the specific binding of chemotype A1γ PG. Notably, *Mycobacterium* species also have chemotype A1γ PG. However there are two primary differences between *Mycobacterium* and *B. subtilis* or Gram-negative A1γ PG: the cross-links of mycobacterial PG are more fully amidated, and the muramic acid residues are N-glycolated instead of N-acetylated (Figure 5). It is also interesting to note that the C3 domain is found in all LysAs that contain M23 peptidase domains; it may be ideally situated to position the peptidase close to its substrate.

2.6.5 Summary

This characterization of the mycobacteriophage LysA proteins has revealed these to be extensively modular proteins constructed from a variety of domains, both with predictable and unknown function, in an architecture rarely seen in the lysins of other bacteriophages. Of those domains with a predicted activity, alignments show that most, or all, of the catalytic residues are conserved, indicating that these domains are likely to be enzymatically active, and there is already some experimental evidence supporting the PG hydrolytic ability of several LysAs (Appendix A.1) (Garcia et al., 2002; Henry et al., 2010; Payne et al., 2009). The N-terminal domains seen in most of these proteins are completely uncharacterized, and several are unique to mycobacteriophage LysAs. The existence of LysAs possessing one of these N-terminal domains in the absence of any predicted catalytic activity is intriguing, and these domains may have new PG hydrolytic activities or other novel functions. The C-terminal domains, which are likely involved in binding, may provide information about important molecules in the mycobacterial cell wall and the protein motifs that are able to specifically bind them. Further, once more fully characterized, these domains may prove useful in research involving labeling, discriminating, and/or binding of mycobacteria, as is currently done with the binding domains of lysins targeting *B. anthracis* and *L. monocytogenes* (Kretzer et al., 2007; Schuch et al., 2002). Ideally this bioinformatic analysis of LysAs will be a starting point for many future studies to harness the unique properties of these proteins for research and therapeutic use, perhaps in conjunction with the other mycobacteriophage lysin, LysB.

3.0 DETERMINATION OF THE ENZYMATIC ACTIVITY AND ROLE OF MYCOBACTERIOPHAGE LYSB PROTEIN

3.1 INTRODUCTION

LysB was originally identified as a potential lysis protein by virtue of its association with the predicted endolysin, LysA (Garcia et al., 2002); however, there has been little study of the LysB protein until recently. The presence of *lysB* within the mycobacteriophage lysis cassettes suggested a role in phage lysis and escape from the cell, as all known proteins involved in lysis to date address the barrier of the cell wall in some form, but its specific function was unknown. The mycobacteriophages each encode a LysA protein, and in some cases a putative holin has also been identified, and these provide a means with which to traverse the peptidoglycan and cell membrane, respectively. However, the mycobacterial outer membrane presents a final barrier to phage release, similar to the outer phospholipid bilayer membrane of Gram-negative bacteria. Just as Gram-negative phages have additional lysis proteins (e.g. the Rz/Rz1 proteins and spanins) to address the outer lipid bilayer (Berry et al., 2008; Summer et al., 2007), it follows that mycobacteriophages may need additional proteins to specifically overcome the mycobacterial outer membrane. However, unlike the unattached outer lipid bilayer of Gram-negative bacteria, the mycobacterial outer membrane is covalently connected to the peptidoglycan. The outermost layer consists of mycolic acids that are integrated into the outer

membrane through weak interactions with the other hydrophobic molecules present, and these mycolic acids are in turn esterified to the arabinogalactan chains that are attached to the peptidoglycan (Brennan and Nikaido, 1995; McNeil et al., 1991).

While not originally identified as LysB, ORF3 (gp3) in mycobacteriophage Ms6 was also recognized to be a part of the Ms6 lysis cassette in addition to a LysA (gp2) and holin (gp4), although its function was unknown (Garcia et al., 2002). A limited sequence alignment with seven other mycobacteriophage LysBs (Gil et al., 2008) found a motif (G-X-S-X-G) that is characteristic of lipolytic enzymes, such as carboxyesterases (EC 3.1.1.1) and lipases (EC 3.1.1.3), and biochemical characterization of the Ms6 LysB showed it to be a lipolytic enzyme (Gil et al., 2008). This class of enzymes hydrolyzes ester bonds to create two products, one with an alcohol group and the other with a carboxylic acid. Lipolytic enzymes have a catalytic triad containing the serine found in the G-X-S-X-G motif in addition to an Asp and His (Carvalho et al., 1999; Gupta et al., 2004; Longhi and Cambillau, 1999). The specific motif seen in LysB proteins is Gly-Tyr-Ser-Gln-Gly, which most closely resembles that of cutinases (EC 3.1.1.74), although the surrounding sequence shows no significant similarity (Garcia, 2002). Cutinases and other serine esterases belong to the superfamily of α/β hydrolases, which contain the catalytic serine in a sharp turn between an α-helix and a β-sheet (Arpigny and Jaeger, 1999; Carvalho et al., 1999; Ollis et al., 1992).

Straddling the categories of serine esterases and lipases, which hydrolyze short (C≤10) and long (C≥10) acylglycerol chains, respectively (Arpigny and Jaeger, 1999), cutinases generally prefer more intermediate chain lengths but have a broader range of activity (Carvalho et al., 1999; Longhi and Cambillau, 1999). Similar to many lipases and esterases, cutinases can hydrolyze a variety of artificial substrates that contain ester bonds (Gilham and Lehner, 2005).

Indeed, studies with Ms6 LysB demonstrated lipolytic activity on several artificial substrates, including Tween-20, Tween-80, tributyrin, triolein, and various *p*-nitrophenyl esters (Gil et al., 2008). Cutinases are normally produced by pathogenic fungi to degrade the protective plant polymer, cutin (Carvalho et al., 1999), which is a complex fatty acid polymer containing many ester bonds. However, cutin is not expected to be encountered by mycobacteriophages.

The only ester bond that has been identified in the covalently attached mycobacterial cell wall is the mycolyl-arabinogalactan bond. Thus, we predict that this is the target of the mycobacteriophage LysB protein (Figure 6). Further, it is likely that hydrolysis of this bond could remove, or at least compromise, the barrier of the mycobacterial outer membrane and thus aid in the release of new phage progeny. In order to test this hypothesis, we conducted a multi-approach characterization of LysB proteins by performing a bioinformatic analysis of 56 LysB proteins, determining the enzymatic activity on both artificial and host substrates, and identifying the role in mycobacteriophage lysis using *lysB*-deletion mutant phage.

3.2   BIOINFORMATIC CHARACTERIZATION OF LYSB PROTEINS

Putative LysB proteins have been detected in 56 out of the 60 fully sequenced mycobacteriophages. Each of these is located within the lysis cassette adjacent to the *lysA* gene with no more than four intervening genes, some of which are putative holins. We performed BLAST analyses, ClustalW alignments, and conserved domain searches for each of the 56 LysB proteins. In addition, we examined the four phage that lack a recognizable LysB protein.

3.2.1 Sequence alignments and conserved domains

The LysB proteins range from 254 (D29 gp12) to 451 (PG1 gp50) residues, and while they do not appear modular like LysAs (Hatfull *et al.*, 2006), they are still highly diverse. A ClustalW sequence alignment (Figure 25) reveals numerous gaps, and by plotting the phylogenetic relationships (Figure 26), one can see that the LysB proteins are highly divergent, with many of these displaying <20% sequence identity to other members of the phamily. Several LysB proteins (D29 gp12, L5 gp12, CJW1 gp35, Porky gp35, 244 gp35, Kostya gp36, Omega gp53, and MAV_0786) have large gaps of 60-120 residues at their N-termini (Figure 25). There is also a Gly-rich region of ~30 residues towards the C-terminus in four LysBs (PG1 gp50, Orion gp50, Chah gp51, and TM4 gp30 [sequence does not fully align in this ClustalX alignment]), with smaller regions of ~10 residues also visible in members of clusters D and G and Giles gp32. The function of this region is unknown. In the search for conserved domains using BLAST against the GenBank database, 14 LysB proteins, including D29 gp12, were predicted to contain a cutinase domain (Pfam01083), and Wildcat gp52 is predicted to have an esterase/lipase domain (Pfam c109107) (Figure 26). These are all predicted lipolytic functions, which is consistent with the activity reported for Ms6 LysB (Gil *et al.*, 2008). In addition, eight LysB proteins have a predicted peptidoglycan-binding domain (pfam01471) at their N-terminus (Figure 26). The LysB proteins listed above, whose N-terminal regions are truncated compared to other LysBs, may be missing this binding region.

G-x-Ser-x-G

Figure 25. Alignment of LysB protein using ClustalW

159

Figure 25: The 56 LysB proteins from 60 sequenced mycobacteriophage genomes, mycobacteriophage Ms6, and a putative prophage of *M. avium* were aligned using ClustalW. The catalytic Ser, Asp, and His are identified with asterisks. Also identified are the G-X-S-X-G motif with the catalytic Ser, the G-X-P motif, and the Gly-rich region. Phage clusters and sub-clusters are listed on the left. More conserved residues are shaded red while less conserved residues are shaded blue.
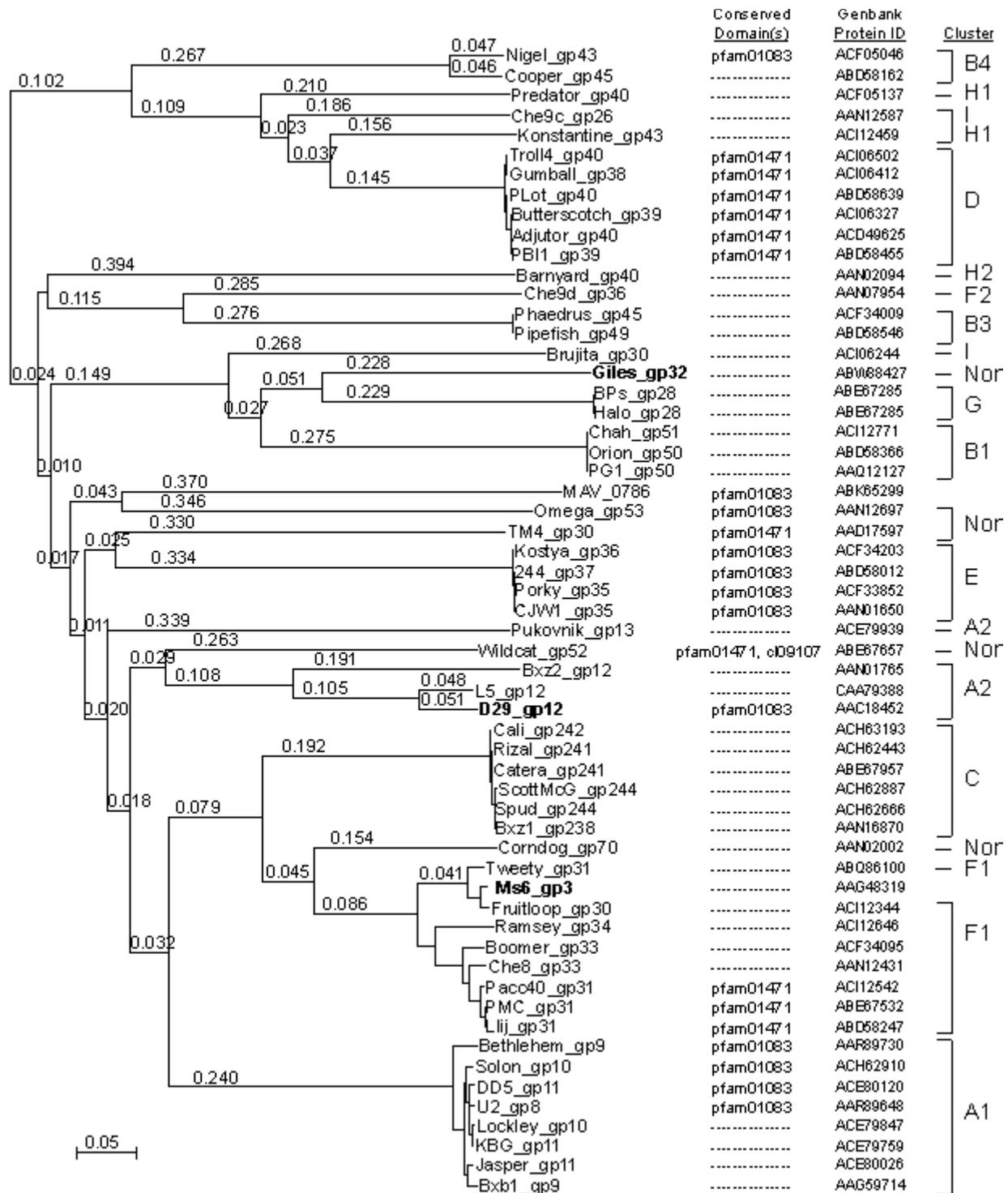
Figure 26. Phylogenetic relationships of LysB proteins

Figure 26: The 56 LysB proteins of sixty completely sequenced mycobacteriophage, mycobacteriophage Ms6, and of a putative M. avium prophage (MAV-0786) were aligned in ClustalW and displayed using Njplot. Conserved domains identifiable in each LysB protein, are shown to the right: pfam01083, Cutinase superfamily; pfam01471, putative peptidoglycan binding domain; cl09107, Esterase_lipase superfamily. The Genbank protein ID numbers are shown on the right, and the three LysB proteins that have been examined experimentally are shown in bold.

3.2.2 Conserved residues

Mycobacteriophage LysB proteins contain three absolutely conserved residues (Figure 26), which include the predicted active site serine (position 82 in D29 LysB) plus a GXP motif approximately 40 residues C-terminal to it (residues 117-119 in D29 LysB). The GXP motif is not absolutely conserved in all serine esterases, and its function is unknown. The lipase- and cutinase-like conserved domains suggest that LysB is a serine esterase, which typically have a G-X-S-X-G motif containing the catalytic Ser. This motif is observed, and almost completely conserved, in the LysB proteins (Figure 26).

By conducting a ClustalW alignment with a greater number of sequences the remaining catalytic residues are more easily identified. Through a limited alignment, three conserved aspartic acids were previously proposed to be involved in catalysis (Gil *et al.*, 2008); however, in this larger alignment of 58 proteins, only the aspartic acid corresponding to position 166 in D29 LysB is highly conserved, with the single departure being the substitution of a glutamic acid in the LysB protein encoded by a putative *Mycobacterium avium* prophage (Figure 25). This difference could indicate an evolutionary divergence that has inactivated the protein within a quiescent prophage, but since Asp → Glu is a largely conserved substitution that retains the important negative charge, the protein may still be active. A conserved His residue is found towards the end of the LysBs, significantly further from the Asp as compared to cutinases due to the presence of a large insertion (Figure 25). However, the crystal structure of the D29 LysB (see Section 3.4), suggests that this His residue (at position 240) is indeed part of the catalytic triad. Interestingly, there is a largely conserved Tyr two or three residues C-terminal to the Asp and His, respectively, corresponding to positions 168 and 243 in D29 LysB, as well as several

163

highly conserved Gly residues. These may be involved in structural stability of the protein or aid in correctly positioning active site residues, as is seen for Ser in the G-Y-S-Q-G motif in most LysBs.

3.2.3 Mycobacteriophage lysis cassettes lacking a LysB protein

Unlike the LysA proteins, there is not an identifiable LysB homolog in every mycobacteriophage genome. Mycobacteriophages Che12, Rosebush, Qyrzula and Myrna lack *lysB* in or near their lysis cassettes (Figure 27), and it is unclear whether these phages encode alternative enzymes performing analogous functions. Aside from Rosebush and Qyrzula, which are highly similar and constitute the only members of the subcluster B2, these phages are not closely related to each other on a genomic level; however, they can be compared to more similar relatives. For example, the genes surrounding the lysis cassette of Che12 are syntenic with other phages in the subcluster A2, including L5, D29, Pukovnik and Bxz2 (Figure 27 A); it appears that Che12 has simply lost its *lysB*. Rosebush and Qyrzula depart from their close relatives of the cluster B phages in the inclusion of a gene immediately downstream of *lysA* that is lacking in other Cluster B phages, all of which do encode LysB (Figure 27 B). This predicted protein is smaller than any LysB proteins (178 aa) and may be a holin; it possesses two transmembrane domains like class II holins (Wang et al., 2000) and shows distant similarity to the D29 putative holin (gp11), while the location of any putative holins in the other Cluster B phages is unknown. In Myrna, neither of the two ORFs (*244*, *245*) immediately downstream of *lysA* is related to any other mycobacteriophage proteins (Figure 27 C). However, Myrna gp244 has similarity to the N-terminal region of the large (823 aa) *Rhodococcus* protein RHA1_ro08121 that contains both M23 family peptidase (Pfam01551) and glycoside hydrolase family 25 muramidase (cd00599)

164

motifs in its C-terminus. These same peptidoglycan-hydrolyzing activities are predicted in several LysAs, suggesting that Myrna gp244 may have a role in lysis and peptidoglycan hydrolysis.

B
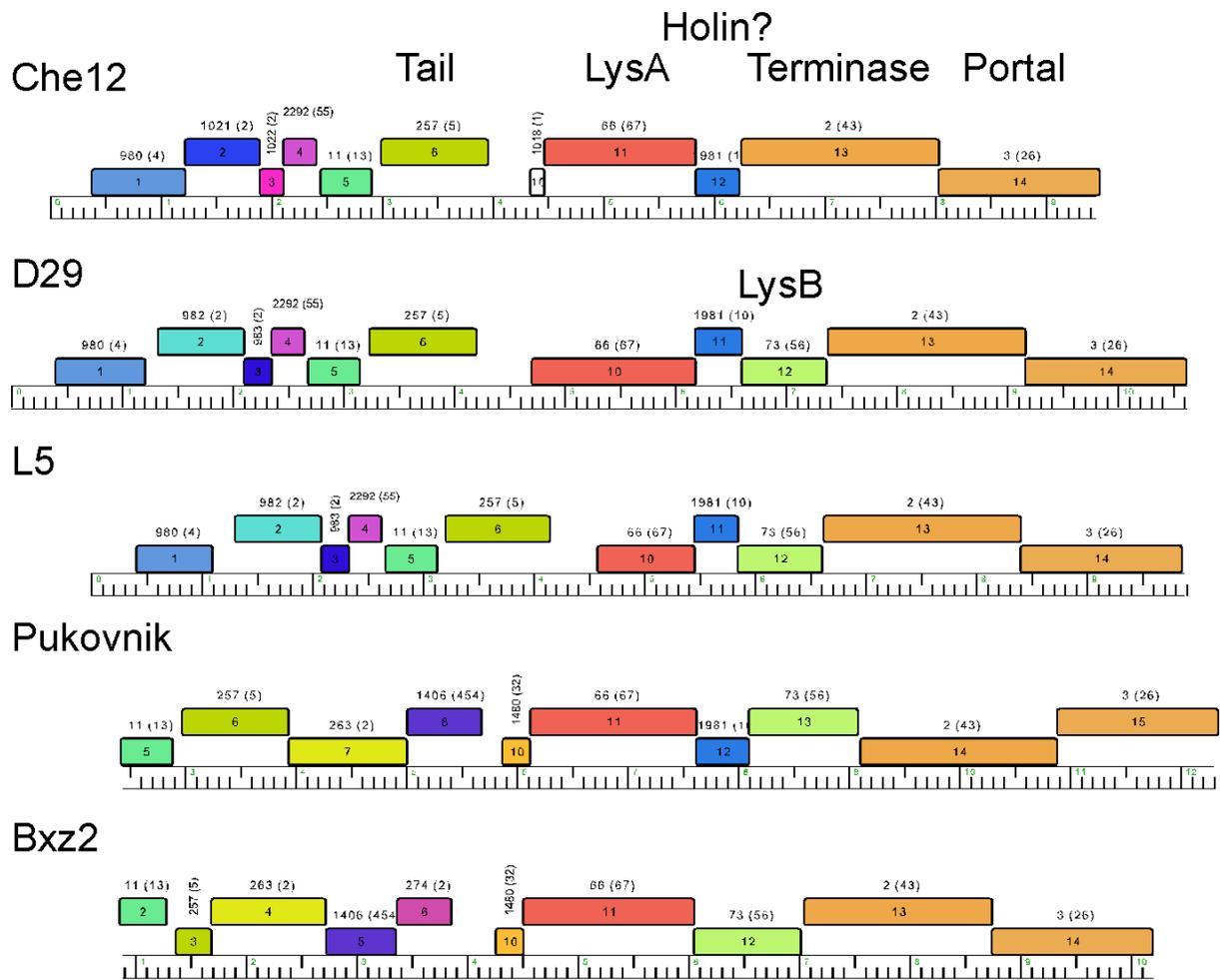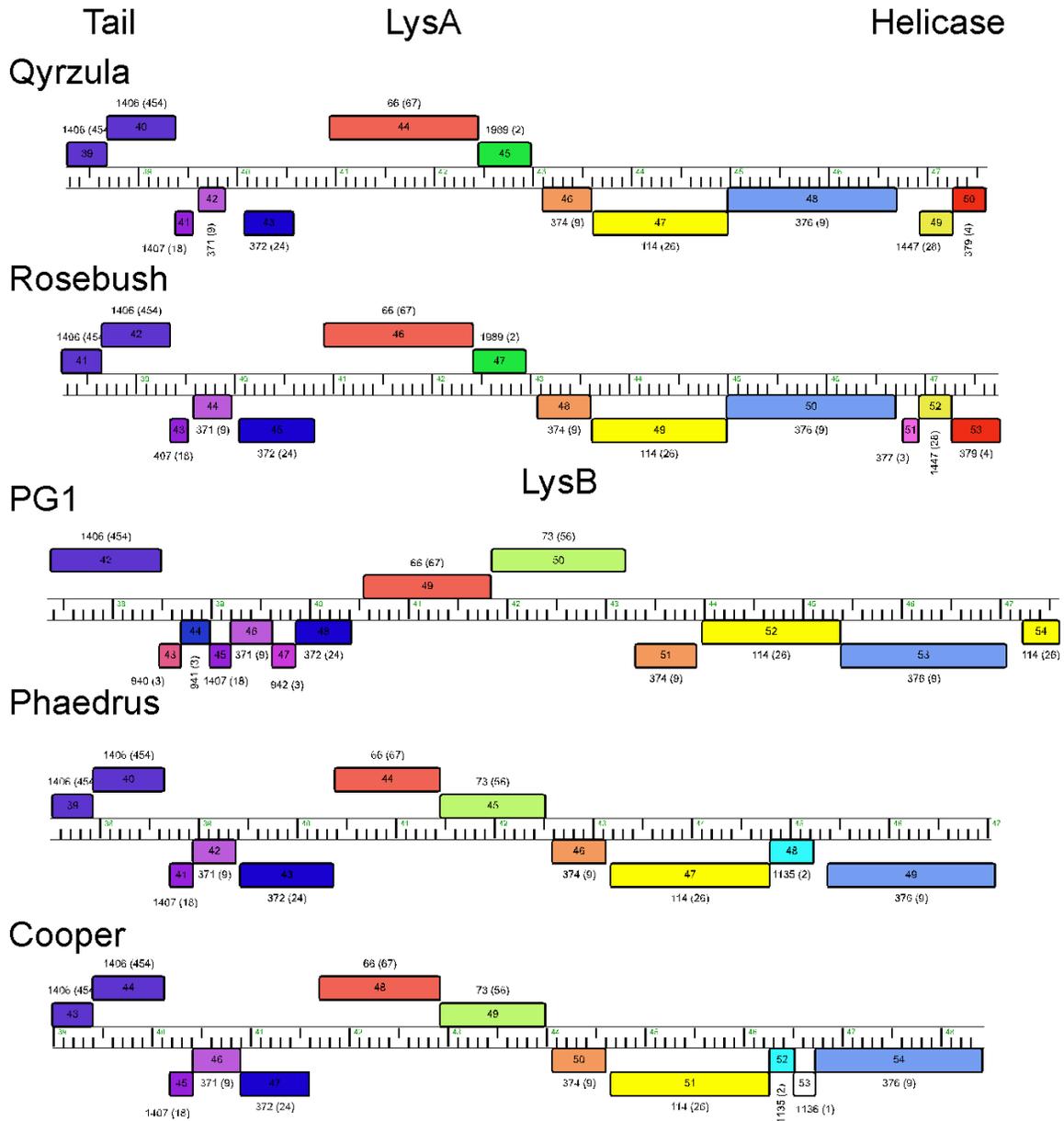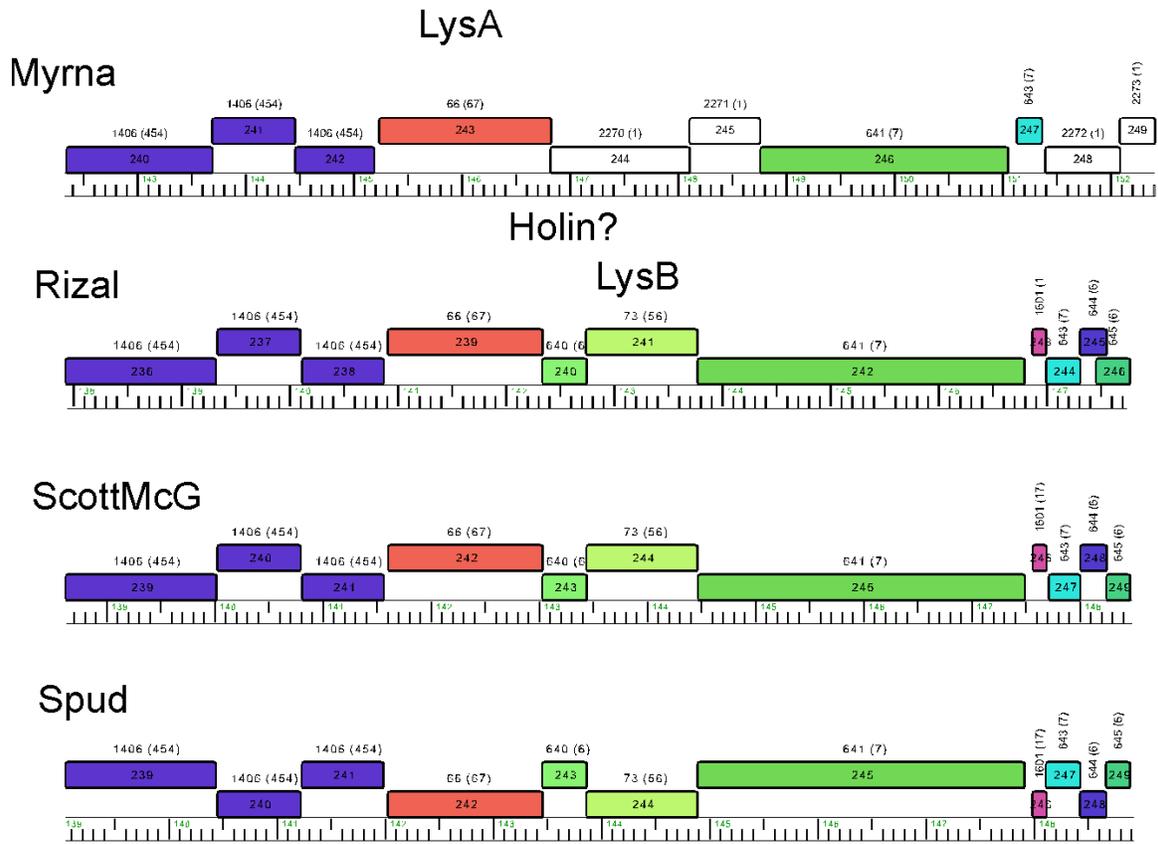
Figure 27. Genetic organization of mycobacteriophage lacking *lysB*

Figure 27:
The lysis cassettes and surrounding genes of the mycobacteriophages lacking a LysA are compared to those of related mycobacteriophages. ORFs are represented as boxes, with differing colors corresponding to membership in a specific Pham.

A. The lysis cassettes and surrounding genes of the five cluster A2 genomes are shown. All contain a *lysA* gene, but Che12 lacks *lysB*, seemingly through simple loss of the genes. Note that the putative holin is also apparently missing in Bxz2. The pham number with the number of phamily members in parentheses is shown above each gene.

B. The lysis cassettes of five cluster B genomes are shown; Qyrzula and Rosebush belong to cluster B2, PG1 to cluster B1, Phaedrus to cluster B3 and Cooper to cluster B4. Each genome contains a *lysA* gene, but both Qyrzula and Rosebush lack *lysB*. While both these genomes contain a different gene in the same location (Qyrzula *gene 45* and Rosebush *gene 47*), these genes constitute Pham1989, which PSI-Blast analysis suggests is distantly related to Pham 10, a group of predicted holins. The location of putative holin genes in PG1, Phaedrus and Cooper is not known.

C. The lysis cassettes of four phages of cluster C are shown; Rizal, ScottMcG and Spud all belong to subcluster C1, while Myrna belongs to subcluster C2. Each genome contains a *lysA* gene, but Myrna lacks *lysB*. The adjacent gene (*244*) is a candidate for involvement in lysis, but it is not related to the large family of cutinase-like proteins or other esterases. Myrna gene *245* is of unknown function.

3.3   ENZYMATIC ACTIVITY OF LYSB PROTEINS ON ARTIFICIAL

SUBSTRATES

The bioinformatic analyses suggest that mycobacteriophage LysB proteins comprise a novel family of lipolytic proteins that are most closely related to bacterial and fungal cutinases. Lipolytic activity has been shown with the LysB homolog of mycobacteriophage Ms6 using various artificial esterase and lipase substrates (Gil et al., 2008). Prior to studies with native mycobacterial substrates, we wanted to confirm this esterase activity in our cloned LysB, as well as engineer active site mutants to verify the classification of LysB proteins as serine esterases.

3.3.1   Cloning and expression of D29B and an active site mutant

The LysB (gp12) of mycobacteriophage D29 was chosen for *in vitro* assays to determine the structure and function of LysB proteins. This is the smallest LysB protein sequenced to date (254 aa) and shares only 40% amino acid sequence identity with the previously characterized Ms6 LysB protein (Gil et al., 2008), lacking 90 N-terminal residues that are present in Ms6 LysB (Figure 25). However, D29 is one of the best-studied mycobacteriophages (David et al., 1984; David et al., 1992; Ford et al., 1998; Rybniker et al., 2006; Schafer et al., 1977; Sellers et al., 1962), and the genes of the lysis cassette of D29 (*gene 10*, *lysA*; *gene 11*, putative holin; and *gene 12*, *lysB;* cloned as pLAM1, pLAM2, and pLAM3, respectively) had previously been cloned by L. Marinelli (unpublished data) into a pET21a vector for expression with a C-terminal His-tag for purification. While other phage endolysins have the cell wall binding domain at the

C-terminal end, most serine esterases and lipases contain secondary structure elements positioned above the active site that are involved in substrate binding (Gilham and Lehner, 2005). Since LysB does not appear to have a modular organization like other phage lysins it seemed reasonable to expect little or no interference on activity from a C-terminally-placed His tag. D29 LysB was induced with 1 mM IPTG and expressed in high amounts, the majority of which was present in the soluble fraction (Figure 28). The protein was purified to near homogeneity and readily concentrated to greater than 10 mg/ml (Figure 28). In addition, based on the prediction of Ser82, Asp166, and H240 as the active residues, two D29 LysB active site mutants were created. Using site-directed mutagenesis with the pLAM3 plasmid, the Ser82 codon (TCG) was mutated to an Ala (GCG) and the His240 codon (CAC) was mutated to Ala (GCC) to create the D29 LysB S82A (pKC20) and H240A (pKC23) mutants, respectively.
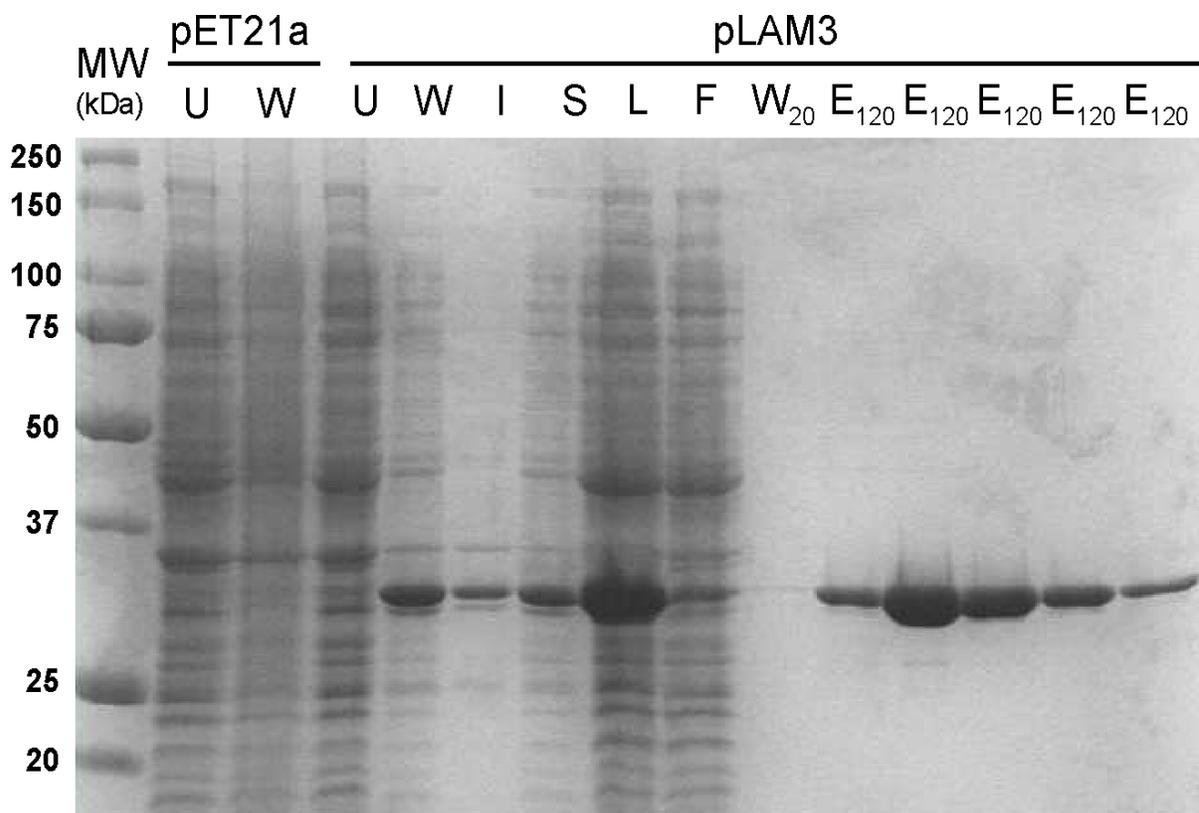
Figure 28. Purification of D29 LysB

Figure 28:

D29 LysB was expressed from plasmid pLAM3, a derivative of vector pET21a, in *E. coli* BL21(DE3) and purified to near-homogeneity. SDS-PAGE of un-induced cells (U), whole cell lysates of induced strains (W) separated into insoluble (I) and soluble (S) fractions, and a clarified soluble lysate (L) are shown. The 30 kDa His-tagged D29 LysB was bound to a cobalt-affinity matrix, and flow-through (F), a 20 mM imidazole wash (W20) sample, and fractions collected at 120 mM imidazole elutions (E120) are shown.

3.3.2 Assay of lipolytic activity with *p*-nitrophenyl esters

The activities of purified D29 LysB and the putative active site mutants were measured using a chromogenic assay involving hydrolysis of various *p*-nitrophenyl esters with acyl chains, which are composed of a carbon chain esterified to a *p*-nitrophenol group (Gilham and Lehner, 2005). Hydrolysis of the ester bond releases the *p*-nitrophenol, which can be detected as a yellow color that absorbs at 420 nm (Figure 29 A). The carbon chains assayed here varied in length from a single carbon (*p*-nitrophenyl acetate) to the 16-carbon hexadecanoic acid palmitate (*p*-nitrophenyl palmitate). In general, shorter chain length substrates are used to assay esterase activity, while longer chain lengths assay for lipase activity (Gilham and Lehner, 2005). Most studies, including those conducted on Ms6 LysB (Gil et al., 2008), utilize *p*-nitrophenyl butyrate (*p*NPB), which has a four-carbon fatty acid and is easily solubilized. Therefore, the assay conditions were optimized using *p*NPB as the substrate. During optimization, it was found that trace amounts of imidazole remaining after purification of the recombinant LysB proteins caused non-specific substrate hydrolysis, and therefore, <1 ml amounts of the purified proteins were dialyzed against 2 L of storage buffer instead of the standard 1 L. Additionally, when alternate substrates were tested, it was found that as the carbon chain length increased, the *p*NP esters became more insoluble, as was observed in similar assays (Gilham and Lehner, 2005). In order to increase solubility, an emulsifier was added to the reactions, which would allow *p*NP esters of up to 16 carbons to solubilize without greatly inhibiting activity (as measured with shorter substrates like *p*NPB). Several concentrations of Tween-20, Tween-80, and Triton X-100 were tested, and optimal results were obtained with 0.1% Triton X-100.

Using the assays described above, we obtained a specific activity for wild-type D29 LysB of 0.72 U/mg (Figure 29 B). This is somewhat higher than the 0.12 U/mg observed for Ms6 LysB (Gil et al., 2008) or the activity of the seven cutinase-like proteins found in *M. tuberculosis* (West et al., 2008), although slightly different assay conditions may have influence the resulting activities. The S82A and H240A LysB mutants were tested and found to be inactive (Figure 29 B and data not shown), consistent with these residues being part of the catalytic triad. As a positive control, a purified lipase from *Pseudomonas fluorescens* showed minimal activity on *p*NPB, but activity increased dramatically as the carbon chain length increased (Figure 29 B), which follows the known preferences of lipase enzymes (Gilham and Lehner, 2005). The D29 LysB, however, decreased in activity with longer substrates (Figure 29 B). This was also observed with Ms6 LysB (Gil et al., 2008), supporting a shared activity between the two mycobacteriophage LysB proteins, despite only 40% sequence identity. It was curious to see this decrease, considering that the predicted substrate of the LysBs is the exceptionally long (C_{40}-C_{90}) mycolic acids, but this could be explained by other differences in the substrate binding region that are specific to the structures found on the host. Indeed, the N-terminal region is positioned over the active site in many serine esterases, forming the lid structure that determines substrate specificity (Longhi and Cambillau, 1999). With the absence of a large portion of N-terminal residues of D29 LysB compared to the LysBs of Ms6 and other mycobacteriophages, it is possible that this may contribute to the increased specific activity observed for D29 LysB compared to Ms6 LysB, while substrate specificity remains similar to Ms6 LysB.
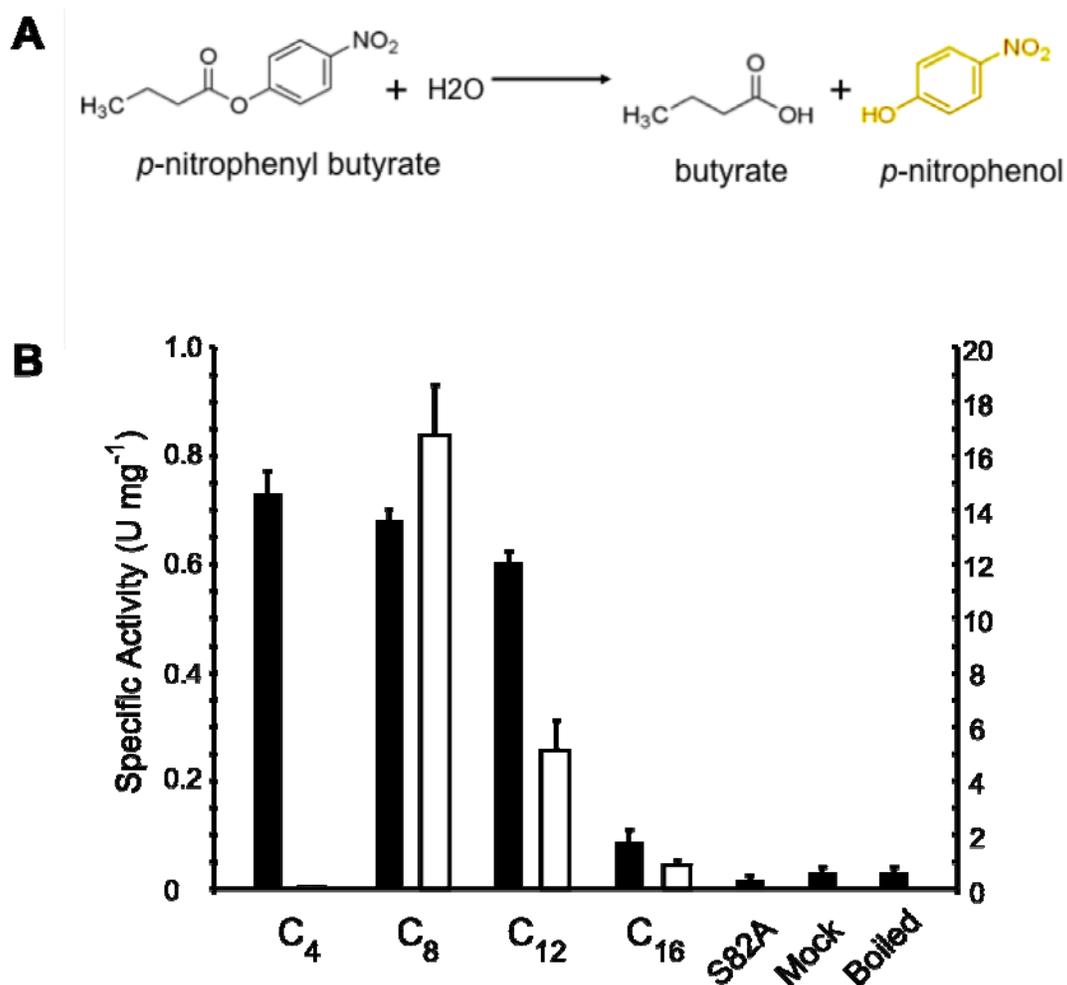
Figure 29. D29 LysB is a lipolytic enzyme

Figure 29:
A. Hydrolysis of the ester bond of *p*-nitrophenyl esters releases a fatty acid and *p*-nitrophenol, which absorbs light at 420 nm.
B. Lipolytic activity of D29 LysB (filled bars; left axis scale) is shown using *p*-nitrophenyl substrates containing different lipid chain lengths (C_4, C_8, C_{12} and C_{16}); lipase activities on the same substrate are also shown (open bars; right axis scale). Specific activities are shown as Units/mg protein, with 1 Unit corresponding to the release of 1 μmol *p*-nitrophenol min^{-1}. An active site mutant (S82A), boiled sample, and a mock-purified sample show little or no activity.

3.4   CRYSTAL STRUCTURE OF D29 LYSB (Q. SUN AND J. SACCHETTINI)

In collaboration with Qingan Sun and Dr. James Sacchettini (Payne et al., 2009), a crystal structure of D29 LysB was obtained (Figure 30 A) and has been deposited in the Protein Data Bank (PDB) with accession code 3HC7. A structural alignment (not shown) found similarity to members of the *α/β* hydrolase family, which includes cutinases and lipases (Holm et al., 2008; Masaki et al., 2005). The closest structural relative is a *Cryptococcus sp.* cutinase-like protein (PDB 2czq, superimposed in cyan over D29 LysB), although the sequence identity is only 21% (Table 4). However, D29 gp12 and other LysBs are larger than most cutinases (~225 residues), and this extra length occurs between the catalytic Asp and His residues, which are normally <20 residues apart (Carvalho et al., 1999; Longhi and Cambillau, 1999). The remainder of D29 LysB (Ala162-Asn243) forms an 81-residue linker region of four α-helices with the catalytic His240 at the end. This linker region is positioned over the active site and may be involved in substrate binding and positioning (Figure 30 A). Notably, the Gly-rich region seen in some LysBs (above) is located in this α-helix linker region. Further supporting the identification of LysB as a serine esterase, D29 LysB possesses the catalytic triad of Ser-Asp-His in the same place structurally as the cutinase-like protein, with Ser82, Asp166, and His240 (Figure 30 B). The conserved GXP motif is also located near the active site, suggesting a possible role in structural stability or substrate specificity.

176

Label in panel B: Ser 82

Label in panel B: His240

Label in panel B: Asp166

Panel labels: A, B
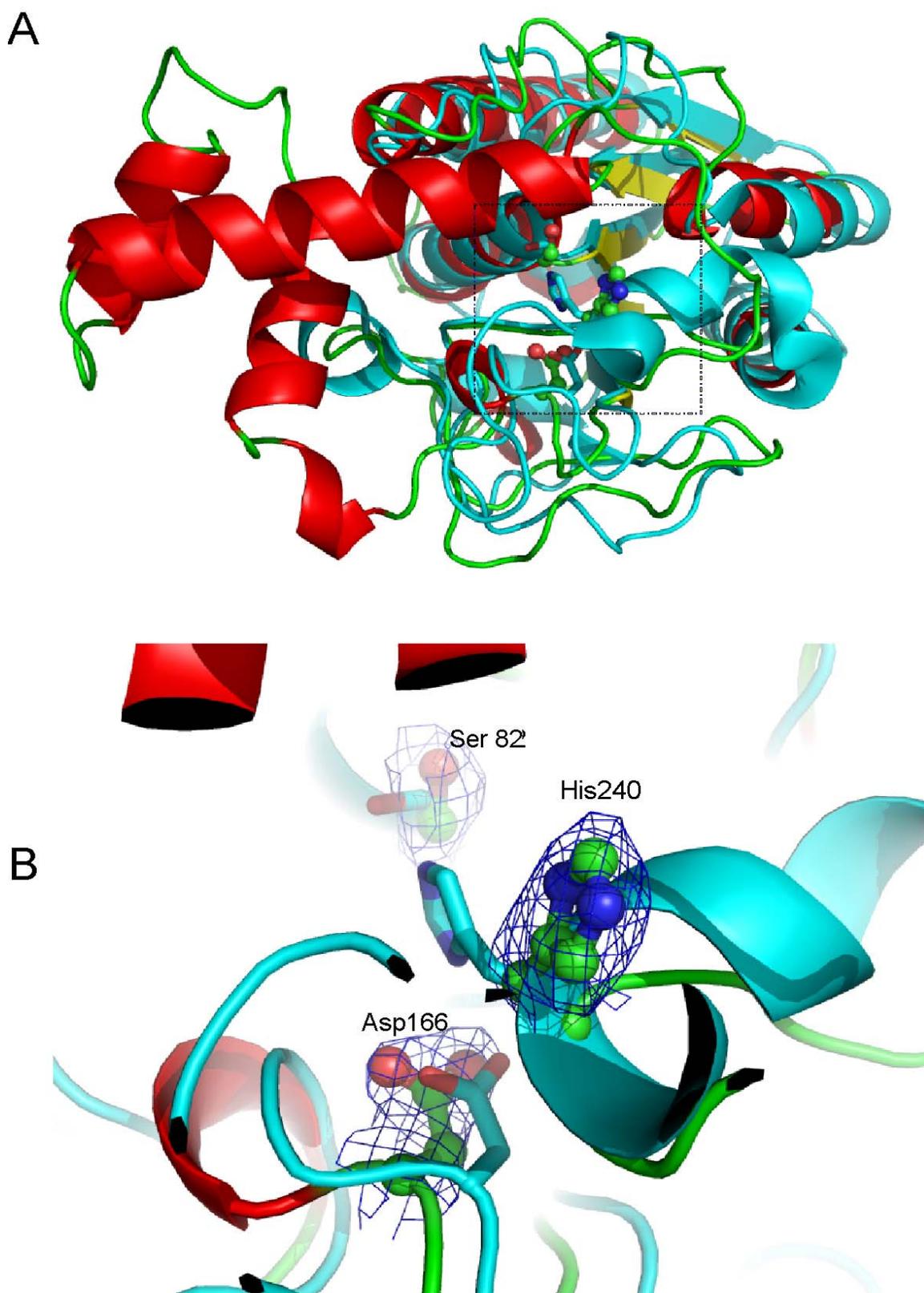
Figure 30. Crystal structure of D29 LysB (Q. Sun and J. Sacchettini)

177

Figure 30:
A. The D29 LysB structure is shown in alignment with the *Cryptococcus sp* cutinase-like protein (PDB: 2czq). D29 LysB is colored according to secondary structure (red: α-helix, yellow: β-strand, green: loops); 2czq is shown in cyan.
B. The catalytic triad of D29 LysB catalytic triad is composed of Ser82, Asp166 and His420 (ball-and-stick representations). Electron density map (2Fo-Fc) around the triad is shown in blue mesh at 1σ level. The triad of 2czq is in stick.

3.5   ACTIVITY OF LYSB ON PURIFIED MYCOLYL-ARABINOGALACTAN-PEPTIDOGLYAN

The D29 LysB structure, along with the observed lipolytic activity on artificial substrates, confirms its identification as a serine esterase.   While D29 LysB is most similar to a *Cryptococcus* cutinase-like protein, there is little reason to believe that mycobacteriophages would encounter cutin in their natural life cycles, much less to such an extent as to exert a positive selective pressure to retain a cutinase in their genome.   Indeed, while the structure is largely retained, the low sequence identity between the *Cryptococcus* cutinase and D29 LysB suggests a divergence far in the past.   Given the broad nature of serine esterase substrates, it would follow that LysB targets a substrate that is different from cutin.   Upon examining the mycobacterial cell wall structure, the only fatty acid ester bond seen is that connecting the mycolic acids to the arabinogalactan (Figure 6).   Thus, we reasoned that this is the likely target of the third member of the mycobacteriophage lysis system, the LysB protein.

3.5.1   Preparation of purified mycolyl-arabinogalactan-peptidoglycan

To determine if the mycolyl-arabinogalactan bond is indeed the target of LysB, the entire mycolyl-arabinogalactan peptidoglycan (mAGP) complex was purified from *M. smegmatis* and incubated with purified LysB protein.   As with the *p*NP assays, the hydrophilic AGP combined with the hydrophobic mycolic acids was highly insoluble and required a buffer containing an emulsifier (0.1% Triton X-100) to solubilize the substrate.   Despite this, we were unable to

completely resuspend the powdered mAGP in the reaction buffer, as evidenced by visible clumps of substrate in all buffer mixtures. Therefore, it was not possible to precisely determine the rate of hydrolysis or the specific activity, as these calculations require knowing the exact amount of substrate accessible to enzyme.

3.5.2 Hydrolysis of mAGP

Purified D29 LysB was incubated with mAGP at varying concentrations and for different lengths of time (Figure 31 A, B). As a positive control, mycolic acids were released from mAGP by alkaline hydrolysis using tetrabutylammonium (TBAH) (Besra, 1998; Hamid et al., 1993; Watanabe et al., 2001). The reaction products were phase-separated and developed by thin-layer chromatography (TLC) on silica plates using chloroform/methanol (97:3) and visualized by spraying with 5% molybdophosphoric acid in ethanol and charring for 15 minutes (Parish and Stoker, 1998). These analyses showed a time- and concentration-dependent release of mycolic acids by D29 LysB (Figure 31 A, B).

To further verify that the lipid products were mycolic acids, they were methyl-esterified with iodomethane during phase-separation, which yielded methyl esters of α, α' and epoxy mycolic acids similar to those from TBAH treatment and specific to *M. smegmatis* (Figure 31 C) (Watanabe et al., 2001). Preliminary analysis of the released lipids by mass spectrometry and NMR was consistent with identification as free mycolic acids (L. Kremer, X. Trivelli, and Y. Guerardel, data not shown).

Finally, we examined the ability of other proteins to release mycolic acids from mAGP. As expected, mycolic acid release was abolished when the catalytic serine or histidine was mutated to an alanine in the D29 LysB S82A mutant (Figure 31 D). Further, no activity was

seen on mAGP by the *P. fluorescens* lipase (Figure 31 D), implying a substrate specificity, although reaction conditions could have influenced this result.

Additionally, a second spot that traveled less distance than any free mycolates was observed for all reactions developed with chloroform/methanol except the TBAH alkaline hydrolysis control. While its exact nature is unknown, based on the consistent presence and constant density of this spot in both positive and negative reactions (Figure 31 A, B, and D) it may be a by-product of the purification of mAGP, the hydrolysis conditions, or the phase-separation process. The harsh alkaline hydrolysis treatment by TBAH may have prevented the survival of this compound in the control samples.
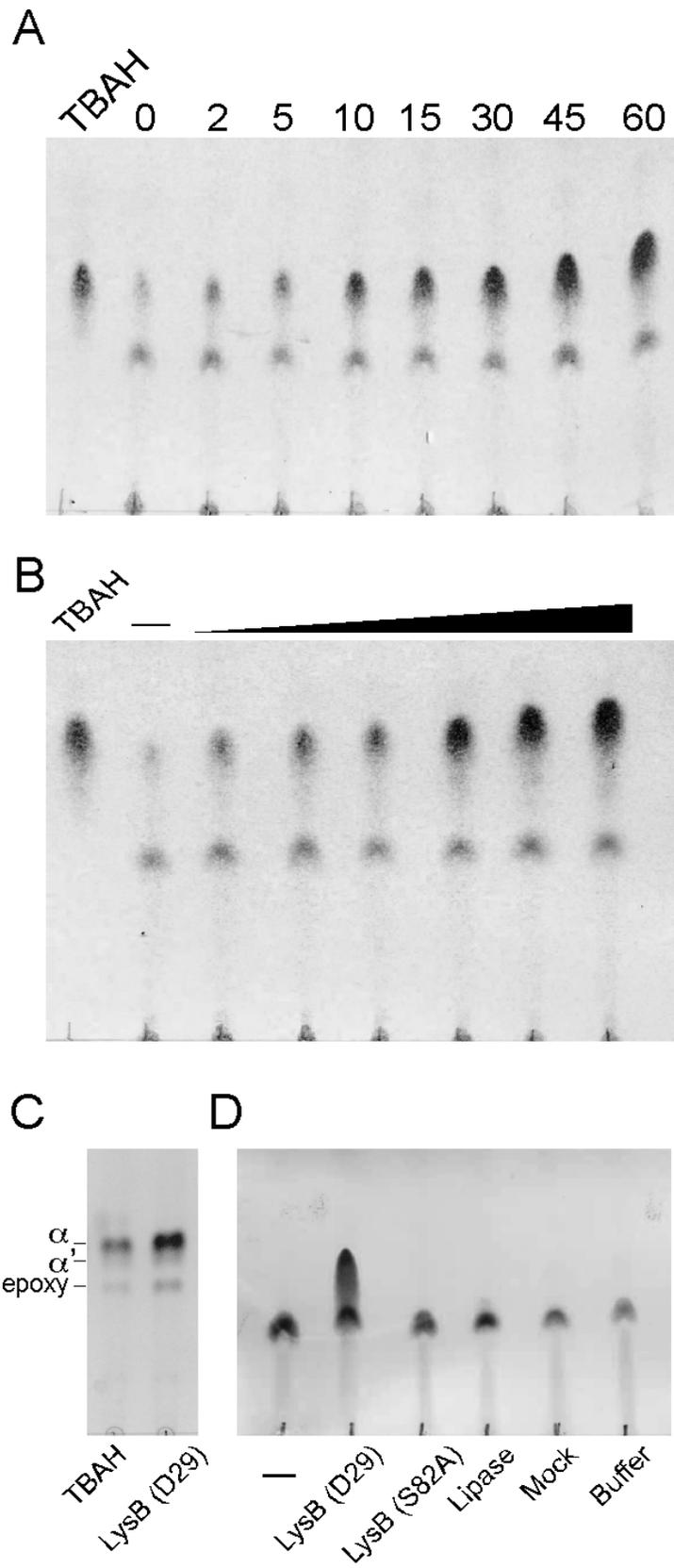
Figure 31. D29 LysB is a mycolyl-arabinogalactan esterase

Figure 31:
A and **B**. Time- and enzyme-concentration-dependence of D29 LysB-mediated hydrolysis of mycolyl-arabinogalactan-peptidoglycan (mAGP). mAGP was purified from *M. smegmatis*, incubated with D29 LysB, and the lipid products separated by thin layer chromatography. The reaction was monitored as a function of time (panel **A**; incubation times as shown, in minutes) or amount of protein (panel **B**; 0.1, 0.25, 0.5, 1, 5, 10 µg D29 LysB). D29 LysB catalyzes production of a species migrating in the same position as the free mycolic acids released by tetrabutylammonium (TBAH) treatment.
C. Methylesterification of the products generated by D29 LysB cleavage of mAGP generates epoxy, α and α' mycolates.
D. Neither a D29 LysB catalytic mutant (S82A), lipase purified from *P. fluorescens*, a mock purified sample, nor buffer release free mycolic acids from mAGP.

3.6   ROLE OF LYSB IN MYCOBACTERIOPHAGE LYSIS

Based on its prevalence in mycobacteriophages (56/60 published genomes), LysB appears to confer some selective benefit or advantage, and its association with the lysis cassette implies that this role is in phage lysis. The ability of LysB to hydrolyze mAGP, and conceivably to separate the mycobacterial outer membrane from the host bacterium, could arguably be a function that is advantageous to mycobacteriophages during lysis. However, there are four phages (Che12, Myrna, Qyrzula, and Rosebush) that lack a LysB homolog or any protein with a similar predicted function near their lysis cassettes, indicating that LysB may not be essential for mycobacteriophage infection and lysis. Therefore, we sought to test its essentiality and determine the role of LysB during infection by constructing a *lysB* deletion mutant and analyzing any resulting phenotypes for clues as to its function.

3.6.1   Construction of a LysB deletion mutant phage

To explore the role of LysB in mycobacteriophage lytic growth, we first asked whether it is essential for plaque formation by using the Bacteriophage Recombineering of Electroporated DNA (BRED) technology developed in our lab (Marinelli et al., 2008) to delete *lysB* (*gene 32*) from the mycobacteriophage Giles genome. We decided to target Giles *lysB* because this phage has been successfully targeted by BRED recombineering in numerous instances, and notably, was previously used to show that the Giles *lysA gene* (*gene 31*) is essential for plaque formation (Marinelli et al., 2008). In addition, we anticipated that it might be necessary to complement the

deletion with a *lysB* gene that is sufficiently different genetically from the one deleted by BRED, as to avoid recombination (i.e. D29 *gene 12*). It should be noted, however, that the Giles LysB protein was cloned and purified in the same manner as D29 LysB and was also able to hydrolyze purified mAGP (data not shown).

For the BRED recombineering, a 200 bp dsDNA substrate containing 100 bp flanking homology on either end was designed to introduce a 1,146 bp internal deletion in Giles *lysB*, fusing the 15 codons at the 5' and 3' ends of the gene (Figure 32 A). The 200 bp dsDNA substrate was constructed by PCR; a 100 bp oligo with 50 bp homology flanking each side of the deletion was extended using 75 bp extender primers. These primers had 25 bp homology to the 100 bp oligo and a further 50 bp of homology to the sequence 5' or 3' to the oligo, totaling 200 bp. The decision to retain a short sequence at the beginning and end of the gene was made after discovering that the ORFs of *lysA* (*gene 31*) and *lysB* (*gene 32*) overlap by one nucleotide, with the A nucleotide functioning both in the TG<u>A</u> of *lysA* and the <u>A</u>TG of *lysB*. This was done in order to minimize effects on expression of adjacent genes as well as to avoid genetic polarity.

After co-electroporation of the 200 bp deletion substrate and purified Giles genomic DNA into an *M. smegmatis* recombineering strain (i.e. containing a prior-induced pJV53 plasmid expressing Che9c gp60 and gp61) (van Kessel and Hatfull, 2007), plaques were recovered on a lawn of *M. smegmatis* mc²155:pKMC2 cells that were induced with acetamide to express D29 LysB. Not knowing if Giles *lysB* would be essential for plaque formation, we sought to plate the initial transformation on a putative complementing strain in the hopes of increasing the number of mixed plaques recovered. Of 22 primary plaques screened by Deletion Amplification Detection Assay (DADA) PCR (Marinelli *et al.*, 2008), two contained mixtures of wild-type and mutant phages (Figure 32 B).

In order to test for the presence of a viable Giles *ΔlysB* mutant, one of these mixed plaques was picked into phage buffer and plated to recover ~600 isolated plaques on both a putative complementing strain (*M. smegmatis* mc^{2}155:pKMC2) and a non-complementing control strain. At this stage, propagation of the Δ*lysB* mutant would only be possible if *lysB* was not essential for plaque formation (on the non-complementing strain) or if its function could be complemented by the D29 LysB protein expressed from pKMC2. Secondary lysates from each plate representing all of the recovered particles were harvested and tested by PCR for the presence of the deletion mutant (data not shown). The mutant genotype was present in approximately equivalent proportions in both the complementing and non-complementing strains. Thirty individual plaques from a secondary plating of the lysates were each screened by PCR, and a homogenous mutant plaque was identified (Figure 32 C). The Giles Δ*lysB* mutant underwent several rounds of purification, and the correct, in-frame deletion was confirmed by DNA sequencing (data not shown). As the Giles Δ*lysB* mutant proved to be viable without complementation, future experiments were done using wildtype *M. smegmatis*. This is in contrast to the Giles *lysA* mutant, which could only be propagated on *M. smegmatis* mc^{2}155:pKMC4, a complementing strain expressing the Corndog LysA (gp69) (Marinelli et al., 2008).
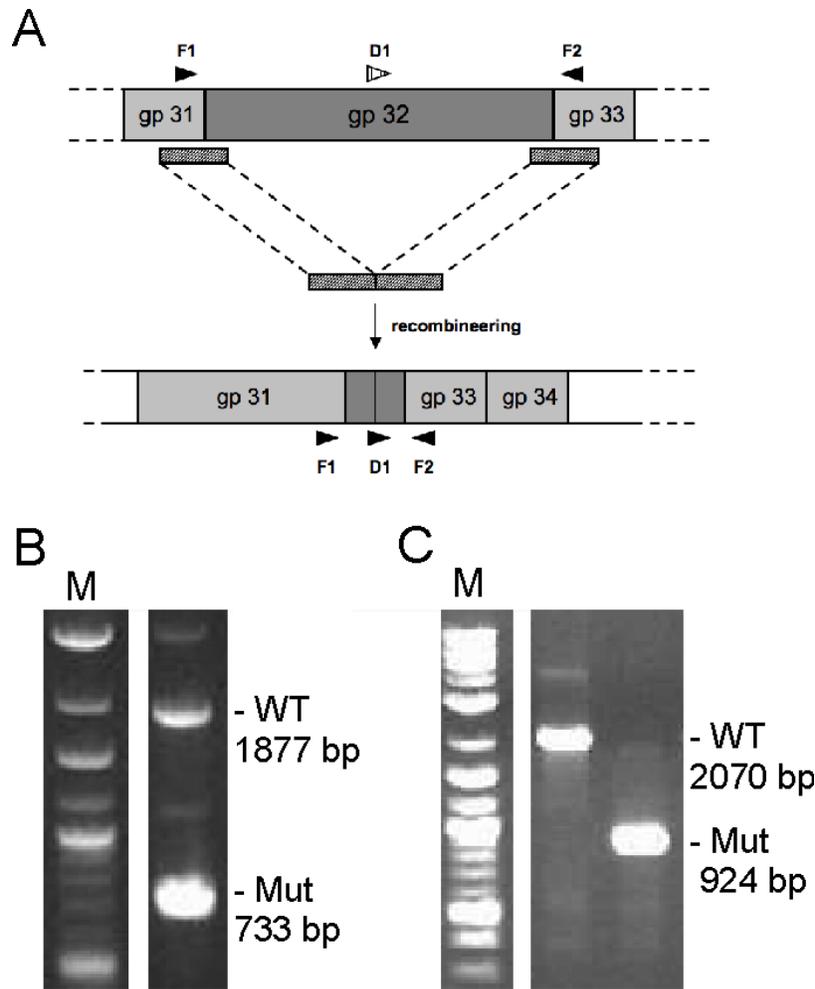
Figure 32. Construction of a Giles *lysB* deletion

Figure 32:
A. A 200 bp dsDNA substrate was designed that has 100 bp homology flanking a 1,146 deletion in Giles *gene 32*. The locations of primers F1, F2, and D1, used for PCR screening are shown.
B. Following co-electroporation of the 200 bp substrate and Giles genomic DNA, primary plaques were recovered and screened by PCR using primers D1 and F2, designed to preferentially amplify the deletion mutant. A mixed plaque containing wildtype and mutant DNA is shown.
C. A mixed primary plaque was diluted and plated, and isolated secondary plaques screened using primers F1 and F2. One homogenous wild-type Giles plaque (left) and one homogenous *lysB* deletion mutant plaque are shown (right).

3.7 CHARACTERIZATION OF GILES Δ*LYSB* MUTANT

Once the Giles Δ*lysB* mutant was constructed and found to be viable, characterization of the mutant phenotype was necessary to discern the role of LysB in phage infection. Since LysB was hypothesized to be involved in lysis, assays were chosen that examined this aspect of phage infection. Plaque formation was the first phenotype assessed, since the ability to form plaques is entirely dependent on lysis. Based on the plaque observations, other assays were conducted to determine if the mutant phenotype was the result of impaired ability to disrupt the cell wall, a delay in timing, or a decrease in the number of progeny phage produced.

3.7.1 Loss of *lysB* in Giles directly results in a small plaque phenotype

Since the Giles Δ*lysB* mutant was capable of forming plaques at equivalent efficiencies on complementing and non-complementing strains, and the titers of wild-type and mutant phage lysates prepared on a wild-type *M. smegmatis* host under standard conditions are similar, we could conclude that Giles *lysB* is not essential for lytic growth (Figure 33). However, we observed that the mutant forms somewhat smaller plaques on lawns of wild-type *M. smegmatis* when compared to the parental phage (Figure 33). This phenotype is exaggerated when higher densities of plating cells are used (conditions that favor fewer bacterial doublings and thus fewer rounds of phage infection). Under such conditions (2×10^8 cfu/plate), the average number of particles in each plaque is ~100-fold reduced in the Δ*lysB* mutant relative to wild-type Giles plaques (5×10^5 and 4×10^7 pfu/ml, respectively).

Plating on a complementing strain (*M. smegmatis* mc²155:pKMC9) expressing D29 LysB from the pNIT-1 vector (Pandey et al., 2009) restored the plaque size to that of the wildtype

Giles (Figure 33), indicating that the small plaque phenotype is a direct result of the loss of Giles *lysB*. To eliminate the possibility that this rescue was the result of a recombination event between the Giles genome and D29 *lysB*, several of the new plaques were picked and shown by PCR to contain only the mutant genotype (data not shown). Therefore, unlike the *lysA* mutant, which absolutely requires complementation to form plaques (Marinelli et al., 2008), *lysB* is not essential for plaque formation, but given the smaller plaque size, *lysB* does affect phage lytic infection.
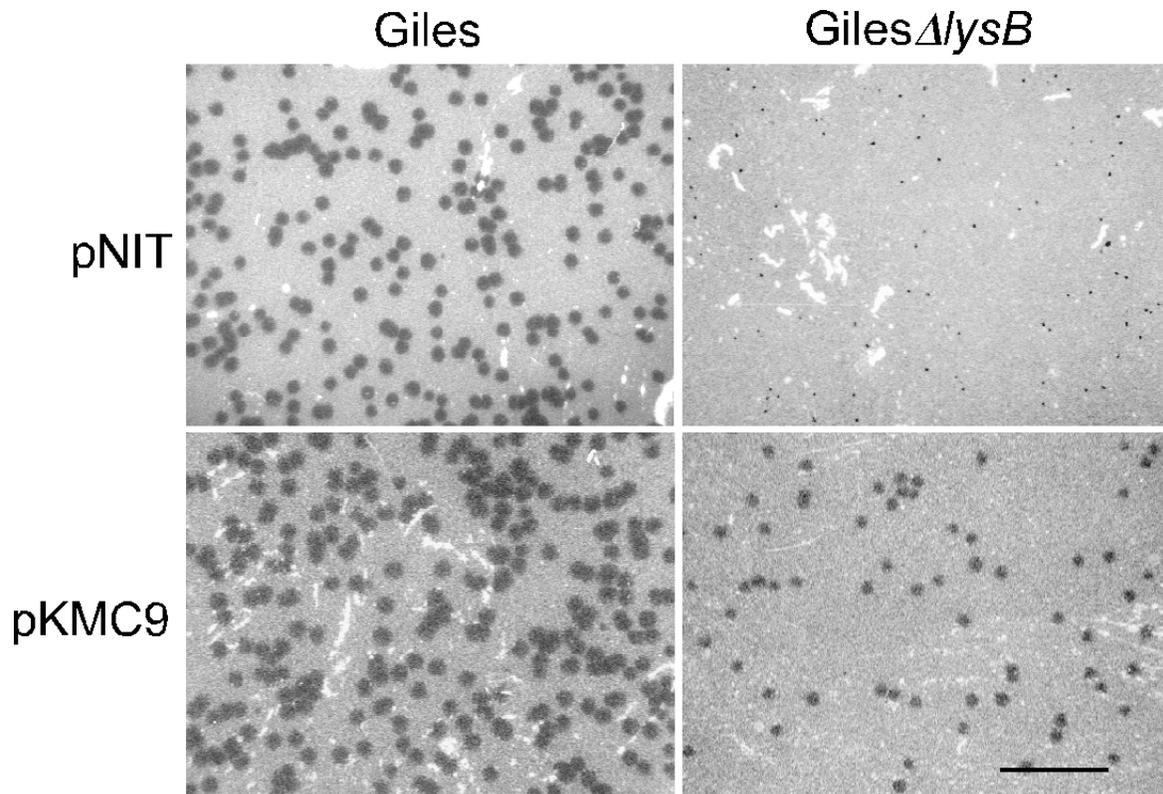
Figure 33. Giles*ΔlysB* has a small plaque phenotype that can be complemented by D29 LysB

Figure 33: *M. smegmatis* mc^{2}155 containing either plasmid vector pNIT-1 or plasmid pKMC9 carrying the D29 *lysB* gene was infected with wild-type or Giles *ΔlysB* phages, as indicated. The *ΔlysB* mutant forms smaller plaques than its wild-type parent, but plaque size is restored by complementation with D29 *lysB*. Scale bar indicates one cm.

3.7.2 Lysis defects in the Giles Δ*lysB* and Δ*lysA* mutants

In order to determine the specific effect of the lysB*Δ* on phage lysis, we needed to test whether the small plaque phenotype is the result of a change in the pattern or the timing of lysis. To do so we performed single-step infections wherein a sufficiently high multiplicity of infection (m.o.i.) was used to infect a culture of *M. smegmatis,* such that every cell would be infected simultaneously. After infection, the progression of phage lysis was monitored by following the decrease in OD at 600 nm as cells lysed, or the permeabilization of the cell membrane was tracked by measuring the increase in extracellular ATP.

After infection of wildtype *M. smegmatis*, the OD of wildtype Giles and the Δ*lysB* and Δ*lysA* mutants increased at the same rate as for the uninfected control culture for approximately 3 hours. Beginning at 3 hours post-infection, the OD of the wildtype-infected culture steadily declined before leveling after 5 hours, at which point all cells had been lysed (Figure 34 A). When infected with the Giles Δ*lysB* mutant, the OD did not begin to decline until 3.5 hours post-infection and was incomplete even 5.5 hours after infection, as indicated by the difference in OD between the Giles Δ*lysB*- and wildtype-infected cells (Figure 34 A). By comparison, cells infected with the Giles Δ*lysA* mutant cease to grow after 3-3.5 hours—the onset of lysis in the wildtype Δ*lysB* mutant—but show only a modest reduction in OD thereafter (Figure 34 A). This agrees with the inability of the Δ*lysA* mutant to form plaques and confirms the importance of having a functional lysin protein in order to effect lysis and phage release.

The same timeline was mirrored in the ATP assay (Figure 34 B). There is very little ATP release until three hours post-infection. The concentration of ATP in the media sharply increases

191

in the wildtype Giles-infected culture and continues to steadily increase until 4.5-5 hours post-infection, at which point the concentration levels off, presumably because all cells are lysed and thus there is no new ATP production (Figure 34 B). The Giles *ΔlysB* mutant is delayed in the onset of ATP release by about 30 minutes, and fails to achieve the wild-type level of extracellular ATP even 5.5 hours after infection. Interestingly, the *ΔlysA* mutant showed no defect in ATP release at all, and may even release more ATP than cells infected with wild-type Giles (Figure 34 B). This is in obvious contrast to the failure of cells to lyse as observed in the OD assay (Figure 34 A). However, this may reflect a key difference in the assays; the OD assay measures cell lysis, but the ATP assay measured the permeabilization of the membrane, which could still result in ATP release without cell lysis and the escape of progeny phages.

Figure 34. Giles *lysA* and *lysB* mutants show inefficient and incomplete lysis
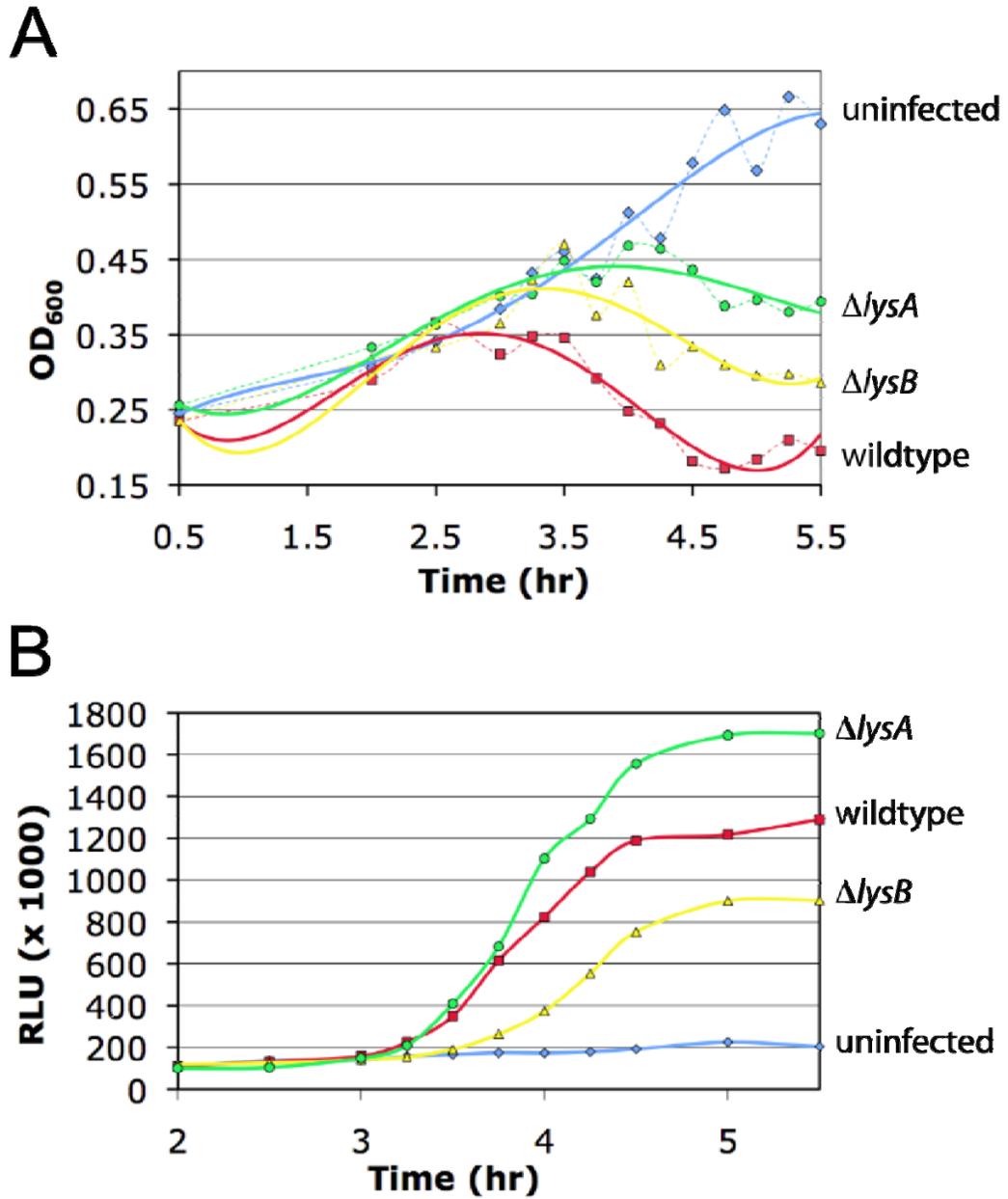
Figure 34:

A. Lysis of phage-infected *M. smegmatis* mc^{2}155 was monitored by measuring OD$_{600}$ of cultures infected with wild-type Giles (red), a *ΔlysA* mutant (green) or a *ΔlysB* mutant (yellow) at different times after phage addition; uninfected cells are shown in blue. Thicker lines correspond to trendlines.

B. *M. smegmatis* mc^{2}155 cells were infected with Giles (red), a *ΔlysA* mutant (green) or a *ΔlysB* mutant (yellow), and ATP release measured at different times after infection; uninfected cells are shown in blue. ATP was measured in a luciferase assay and reported as relative light units (RLU).

3.7.3 Production and distribution of new phage particles in Giles Δ*lysB* and Δ*lysA* mutants

It could be argued that the smaller plaque size and less efficient progeny release are somehow the result of lower phage production within the cell prior to lysis. We sought to test whether the Δ*lysB* and Δ*lysA* mutants produce the same yield of total phage particles as wildtype Giles, as well as to determine how those particles are distributed, that is, whether they are released into the supernatant or remain trapped in unlysed cells. This was assayed by infecting cultures at a high m.o.i. and taking samples at 3, 4, and 5 hours post-infection. At each time-point, cells were pelleted, and the supernatant was removed; phage were released from any unlysed cells by sonication, and each fraction was titered. The Δ*lysA* and Δ*lysB* mutants show no significant difference in the total yield of phage particles as compared to wildtype Giles (Figure 35 A). Immediately prior to the onset of lysis at 3 hours, the majority of new phages are in the cellular fraction for the wildtype and mutant Giles infections (Figure 35 B). However, by 4 hours post-infection, while >90% of wild-type particles are present in the culture supernatant, about 45% of the Δ*lysB* particles remain associated with the unlysed cells. After 5 hours nearly all phage particles have been released from both wildtype and Δ*lysB*-infected cells. In contrast and as expected, <10% of Δ*lysA* phage particles are released into the supernatant even 5 hours after infection (Figure 35 B). The disparity of particle distribution during lysis, and the lack of a significant difference in total phage yield, are consistent with the defects observed for the *lysB*Δ resulting from a defect in host cell lysis.
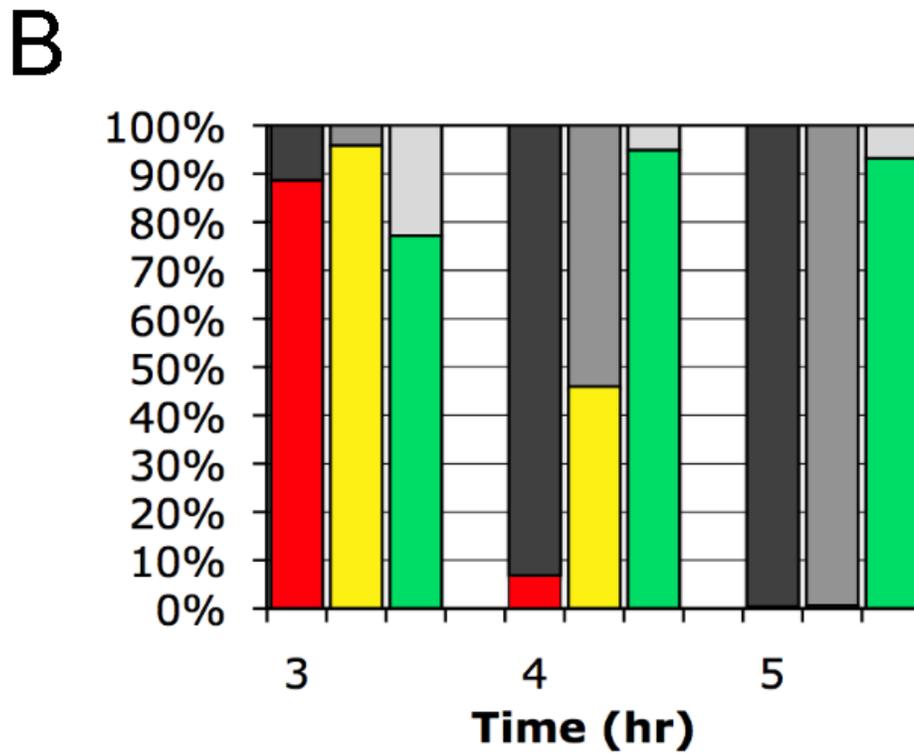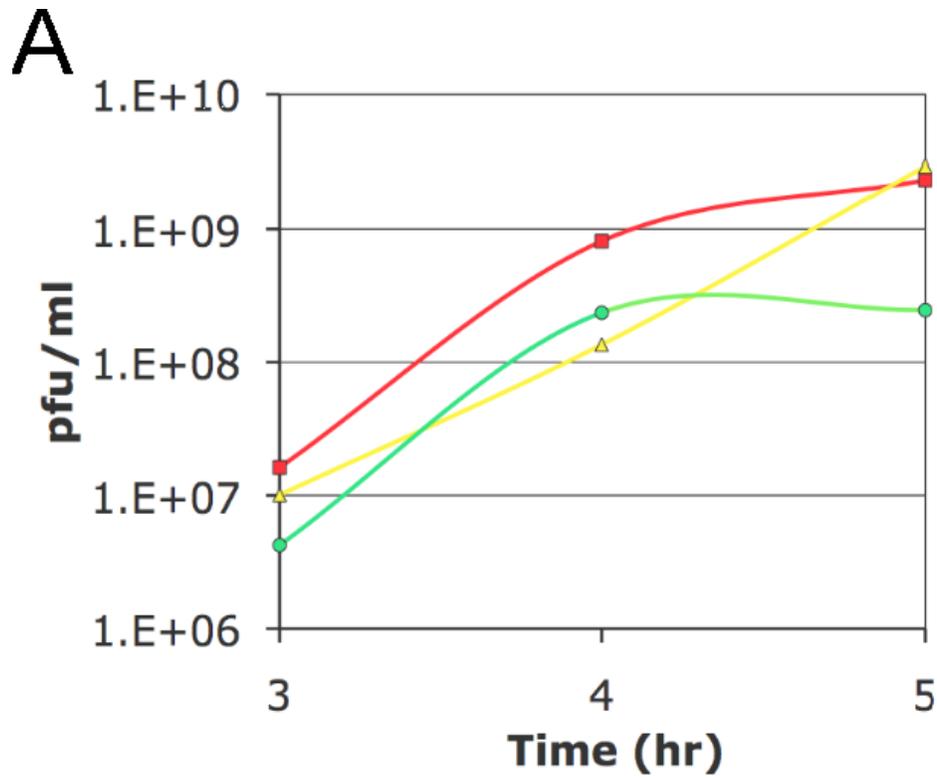
Figure 35. Phage particles are retained in unlysed cells in Giles *lysA* and *lysB* mutants

Figure 35:

A. *M. smegmatis* mc^{2}155 cells were infected with wild-type Giles, the *ΔlysA* mutant, or the *ΔlysB* mutant, and the total number of progeny phage particles (pfu/ml) for each was determined. Wild-type Giles is shown in red, *ΔlysA* in green, and the *ΔlysB* mutant in yellow.

B. The distribution (as percentage of total) of phage particles in the supernatant and cell pellet was determined. For wild-type Giles, the supernatant is shown in dark grey, and the cell pellet in red; the *ΔlysA* mutant supernatant is grey and the pellet is yellow; the *ΔlysB* mutant supernatant is shown in light grey and the pellet in green.

3.8 CONCLUSIONS

We have shown here that the mycobacteriophage D29 LysB protein—and by extension other mycobacteriophage LysB proteins—is a novel mycolylarabinogalactan esterase that efficiently completes the lysis of host mycobacterial cells. Figure 36 shows a model of the mycobacterial cell wall indicating the role of the LysA and LysB proteins, as well as the reaction catalyzed by LysB.

The efficient release of free mycolic acids from D29 LysB-treated cell walls, as observed by TLC, strongly suggests that mAGP is the substrate for the enzyme. While there are other mycolic acid-containing constituents in the mycobacterial cell envelope, only the mycolic acids esterified to the arabinogalactan are linked to the cell wall, and extractable lipids—such as trehalose dimycolate (TDM; cord factor) and trehalose monomycolate (TMM)—are not major components of the mAGP preparations (Besra, 1998; Brennan and Nikaido, 1995). Preliminary data suggest that D29 LysB can also hydrolyze TDM (A. Ojha, K. Payne and G. F. Hatfull, unpublished observations), but it seems unlikely that this has physiological relevance for lysis because TDM is not covalently attached to mycobacterial cells. We propose that cleavage of the mycolic acids to sever the mycobacterial outer membrane from the peptidoglycan-arabinogalactan layer is the primary role of LysB, which can be considered a mycolyl-arabinogalactan esterase.

A

B
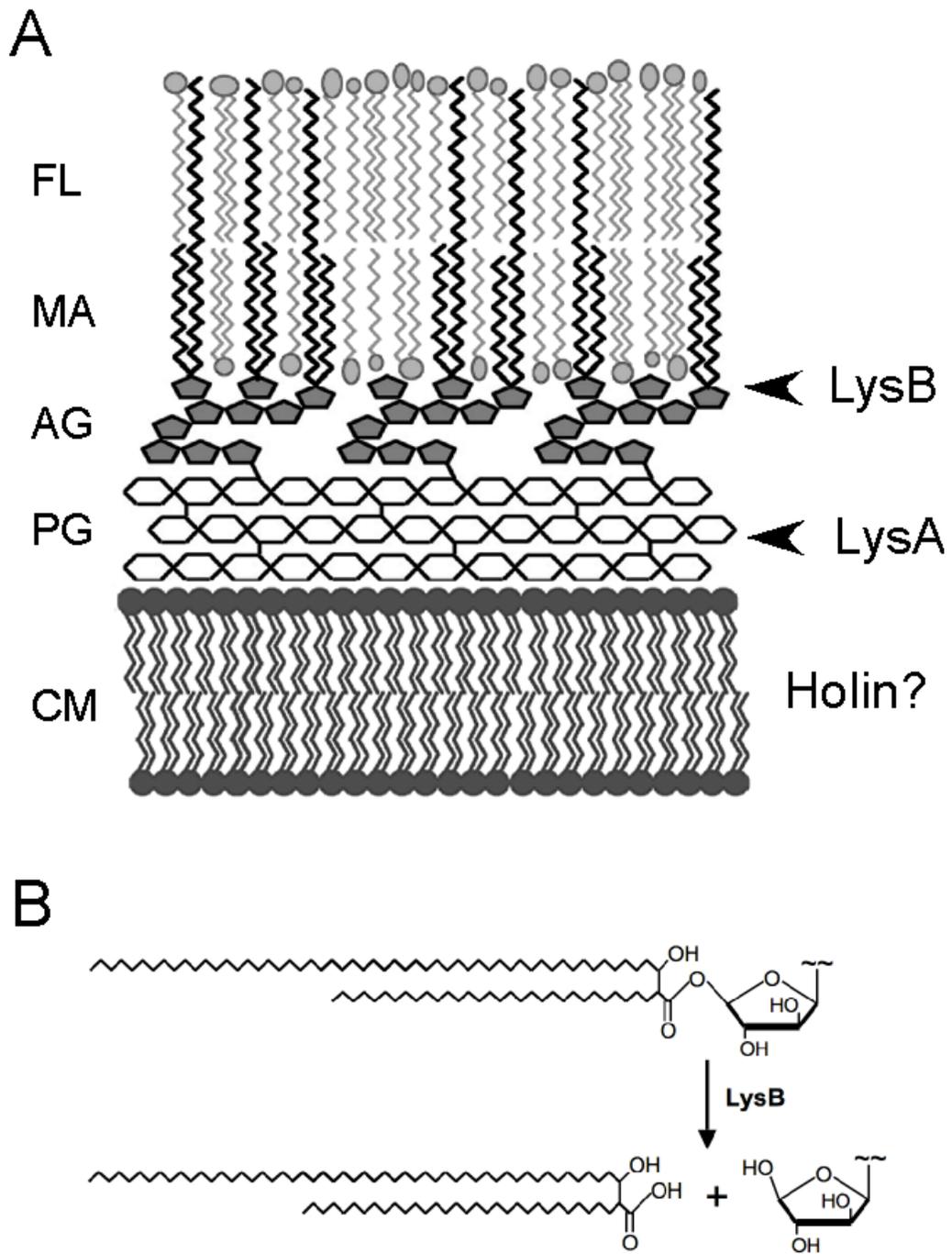
Figure 36. A model for mycobacteriophage lysis of mycobacteria

Figure 36: **A.** Mycobacterial cell walls are unusual in that the cytoplasmic membrane (CM) is surrounded by a peptidoglycan layer (PG) to which a network of arabinogalactan (AG) is covalently attached. A mycobacterial outer membrane consisting of mycolic acids (MA) and free lipids (FL) is covalently attached *via* an ester linkage of mycolic acids to arabinogalactan. We propose that LysA – assisted by holins encoded by at least some of mycobacteriophage – perform an essential step in lysis involving degradation of the peptidoglycan layer, and the lysis is completed through LysB-mediated cleavage of the outer membrane from arabinogalactan. **B**. Diagram illustrating LysB cleavage of the ester bond linking mycolic acids and arabinogalactan.

While removal of LysB function in mycobacteriophage Giles results in an apparently mild plaque phenotype that can be rescued by expression of D29 LysB (Figure 33), other problems in lysis are made more apparent by measuring changes in optical density of the culture, ATP release, and phage particle distribution. The phenotype observed in the Giles *ΔlysB* mutant is consistent with a mild lysis defect, similar to that seen in the coliphage λ Rz/Rz1 mutants (Zhang and Young, 1999), which are unable to efficiently achieve the final fusion of the *E. coli* membranes to induce lysis (Berry et al., 2008; Summer et al., 2007). The unexpected ATP release by the Giles *ΔlysA* mutant is in contrast to its lysis defect as seen in optical density measurements (Figure 34) and by the complete lack of plaque formation in the absence of complementation (Marinelli et al., 2008). A simple explanation for this observation is that the peptidoglycan layer, which most probably remains intact during infection with the *ΔlysA* mutant, provides no barrier to the release of hydrophilic ATP molecules, whereas both hydrophobic barriers—the cytoplasmic membrane and the mycobacterial outer membrane—must be compromised for complete ATP release. Presumably, the cytoplasmic membrane is permeabilized through the action of a Giles holin, and the differences in ATP release observed between the *ΔlysA* and *ΔlysB* mutants are based on the presence or absence of an attached mycobacterial outer membrane.

Phages infecting other Gram-positive bacteria typically complete lysis through simple endolysin-mediated degradation of the peptidoglycan layer, but the lipid-rich mycobacterial outer membrane presents a unique barrier. The mycolyl-arabinogalactan linkage is not common among bacteria, since mycolic acids are found primarily in the Corynebacterineae suborder of the Actinomycetales, which includes *Corynebacteria*, *Gordonia*, *Nocardia*, *Rhodoccocci* and *Mycobacteria*. Few phages infecting non-mycobacterial members of the Corynebacterineae have

been characterized, although these would be good candidates for also encoding mycolylarabinogalactan esterases. However, neither phage P2101 of *Corynebacterium glutamicum* (Chen *et al.,* 2008) nor BFK20 of *Brevibacterium flavum* (Bukovska *et al.,* 2006) encode a LysB relative. Unlike in *Mycobacteria* the mycolic acids of *Corynebacteria,* such as *C. glutamicum,* are not essential for growth (Portevin *et al.,* 2004), and therefore may not pose a significant barrier to efficient lysis, thus rendering a LysB-like protein unnecessary, or may not provide enough selective pressure to maintain such a protein. Alternatively, phages infecting these bacteria may address the mycolic acid barrier through another means, similar to the four mycobacteriophages that lack a LysB.

Currently, the only known physiological circumstances in which free mycolic acids are released from mycobacterial cells is during maturation of *M. tuberculosis* biofilms (Ojha *et al.*, 2008), although the source of biofilm mycolic acids is predominantly from TDM (Ojha et al., 2010). While seven cutinase-like proteins (culps1-7; Rv1984c, Rv2301, Rv3451, Rv3452, Tv1758, Rv3802c, and Rv3724) encoded by *M. tuberculosis* have been expressed and characterized (Schue et al., 2010; West et al., 2009), they are not evidently related to LysB and have not yet been tested for mAGP hydrolysis. Culp1 (Rv1984c) prefers medium-chain acylglycerols, while Culp4 (Rv3452) has phospholipase A_2 activity (Schue et al., 2010) and only Culp6 (Rv3802c) has significant activity on longer carbon chain substrates (West et al., 2009). It remains to be seen whether there are any host-encoded enzymes that share mycolylarabinogalactan esterase activity with LysB.

4.0   DISCUSSION

At the end of a lytic infection, all bacteriophages need to release the newly made phage particles into the environment to find and infect new hosts. All known dsDNA bacteriophages employ a minimum of two proteins to accomplish this: a holin and an endolysin (Young et al., 2000). The holin regulates the timing of lysis and permeabilizes the cell membrane, resulting in a depolarization of the membrane potential. The lysin is a peptidoglycan (PG) hydrolase, and it accesses its substrate either by passing through a holin lesion or by being secreted; in this case it is activated when a pinholin disrupts the membrane potential (Young, 2002). Most lysins consist of two domains, an N-terminal PG hydrolytic domain and a C-terminal cell wall-binding domain (Lopez et al., 1997). Each lysin possesses one or (rarely) two types of PG hydrolytic activity – amidase, glycosidase, or peptidase – and a binding domain that is highly specific to a component of the host's cell wall.

The holin-endolysin lysis system appears straightforward, but research continues to uncover new variations on the above basic mechanism. Two types of holins have been identified thus far: classic holins, which create large lesions in the cell membrane, allowing lysins to pass directly through; and pinholins, which make small holes in the membrane and merely disrupt the membrane potential (Dewey et al., 2010). In addition, the continual recombination of the lysins' lytic and binding domains with those of other phage and host lysins is constantly generating unique lysins (Diaz et al., 1990; Lopez et al., 1997). The resulting lysins are highly tailored to

their specific hosts, primarily due to the specificity of the cell wall-binding domain, but also due in part to selection for a lytic activity that is effective at degrading the host's PG.

In addition to this modular design, novel holin-independent methods of lysin transport across the host membrane have been discovered. Signal-anchor-release (SAR) lysins are secreted into the periplasm via the host's secretory pathway and are tethered to the cell membrane until depolarization – often caused by a pinholin – releases the lysins to act on the PG substrate (Xu et al., 2004). Another method was very recently identified in the mycobacteriophage Ms6; an accessory lysis protein, gp1, acts as a chaperone to export the LysA across the membrane, where it is also activated by membrane depolarization (Catalao et al., 2010).

The most intriguing variations on the holin-endolysin theme involve accessory lysis proteins like gp1. Often these proteins are encoded in alternate reading frames that overlap or are completely embedded within the reading frames of other lysis proteins (Sanger et al., 1982; Young, 1992). Two accessory protein classes were originally identified in the lysis cassettes of enterophages such as λ: the anti-holin and the Rz/Rz1 proteins (Figure 10) (Hanych et al., 1993; Young, 1992). The ratio of anti-holin to holin determines the triggering time of lysis, and these are often encoded within the holin gene and translated using a dual start motif (Barenboim et al., 1999; Blasi and Young, 1996; Graschopf and Blasi, 1999).

While the need for a means of regulating the timing of lysis is common to all phages, the specific type of accessory proteins serving this role are often dictated by the phage's host. The Rz/Rz1 proteins and their single-polypeptide functional analogs, the spanins, are encoded by phages infecting Gram-negative bacteria (Summer et al., 2007). These proteins perform the final step in lysis by fusing the inner and outer membranes of the Gram-negative host bacterium

(Berry et al., 2008). Lysis is still possible, albeit less efficient, in the absence of Rz/Rz1, but the evolutionary advantage of these proteins is evidenced by their widespread presence in the phages of Gram-negative species.

In summary, the holin-lysin system is the most basic unit required for bacteriophage lysis. However, phages have found ways to modify these two proteins and to incorporate additional proteins in an ongoing effort to optimize lysis of their specific hosts. While the PG backbone of all bacteria is largely uniform, cell walls can vary greatly, from the dramatically different organizations of Gram-negative and Gram-positive bacteria (Figure 3) to the single substitution of a different amino acid in the PG peptide cross-link (Figure 2). Evolutionary pressure to optimize lysis favors holins that time lysis for the most effective release of progeny and selects for lysins that can efficiently access and degrade the modified PG of their host. As more barriers are introduced, phages must adapt further to effect lysis, as illustrated by the incorporation of Rz/Rz1 proteins to eliminate the barrier of the outer membrane in Gram-negative bacteria.

4.1   A COMPLEX BACTERIUM REQUIRES A COMPLEX LYSIS SYSTEM.

While *Mycobacteria* have been classified as Gram-positive species, in some ways the cell wall organization is analogous to that of a Gram-negative bacterium (Minnikin et al., 1982). Beyond the thick layer of PG is an asymmetrical lipid bilayer that can account for up to 60% of the cell envelope's weight (Kremer and Besra, 2005). Cryo-electron microscopy has shown this mycobacterial outer membrane to be only slightly larger than the cell membrane (Hoffmann et al., 2008; Zuber et al., 2008). Its composition of mycolic acids and other extractable lipids

renders it highly hydrophobic and impermeable, which complicates chemotherapeutic treatments (Brennan and Nikaido, 1995). The most significant difference between the mycobacterial outer membrane and that of Gram-negative species is that the membrane of mycobacteria is covalently tethered to the cell wall complex (Mahapatra et al., 2005b). The mycolic acids that make up much of the membrane are esterified to the arabinogalactan chains, which are polysaccharides that substitute some of the MurNAc residues in the PG (McNeil et al., 1990). The result of this complex structure is a cell wall with three barriers: the cell membrane, the PG complex, and the mycobacterial outer membrane.

Mycobacteriophages have evolved mechanisms to address these barriers. Indeed, they are not only a barrier to progeny phage egress, but also to phage infection; mycobacteriophages must get their genomic DNA through these layers and into the cell. Research from our lab (Marinelli, 2008; Pedulla et al., 2003; Piuri and Hatfull, 2006) has revealed several features of the mycobacteriophage tape measure proteins that may aid in the injection of DNA into the cell during infection. Regions of many of the tape measure proteins are predicted to be hydrophobic, which may facilitate its insertion into the lipophilic membranes of mycobacteria, and other observations have noted possible changes in the nature of the bacterial cell surface upon expression of the Barnyard tape measure protein (Marinelli, 2008). Even more interesting, several PG hydrolytic motifs have been identified in the tape measure proteins (Marinelli, 2008; Piuri and Hatfull, 2006). Motif 1 shows similarity to Rpf lysozymes (Pedulla et al., 2003), and the Giles and Barnyard Rpf Motifs have been demonstrated to have hydrolytic activity when expressed into the *E. coli* periplasm (Marinelli, 2008). PG hydrolytic activity has also been demonstrated for Motif 3-containing tape measure proteins via zymography (Piuri and Hatfull,

2006). Intriguingly, the Rpf motif in the tape measure proteins of Cluster D phages shows high sequence similarity to the lytic domain of one of the LysA proteins.

My research has focused on the same obstacles presented by the mycobacterial cell wall, but from the opposite end of the phage infection cycle. I have characterized the diverse LysA proteins, which, while sharing many of the same PG hydrolytic activities as other phage lysins, are unique in terms of their modular organizations and the presence of several domains of completely unknown function. I have also examined the novel mycobacteriophage accessory lysis protein, LysB, and identified its most probable natural substrate in the mycobacterial cell wall and its role in mycobacteriophage infection.

4.1.1 The Diversity of LysA Proteins

Bioinformatic analyses of more than 60 mycobacteriophage LysAs (Hatfull et al., 2010) revealed that, similar to all other endolysins, the LysA proteins are modular, containing PG hydrolytic domains and predicted cell wall-binding domains. However, these analyses also revealed an amazing degree of diversity and recombination within these proteins. The variety of domains found in LysAs is unusually high relative to the lysins of other phages; this family of proteins contains 13 shared domains that are combined in various ways to form 19 unique organizations. Six different PG hydrolytic domains (Ami1, Ami2, GH19, GH25, M23, and TG) were identified, as well as three C-terminal domains (C1-3), but there were also four N-terminal domains (N1-4) and three domain-sized regions of unknown function. This unusual diversity of domains may

207

reflect the large number of sequenced mycobacteriophages, relative to phages of other bacteria, and possibly the variable sites of their origin; the mycobacteriophages were isolated from soil samples taken from around the world (Hatfull et al., 2006).

The LysA proteins have also diverged from the standard two-domain lysin structure and predominantly consist of three-domain architectures, in which the predicted lytic domain is in the center, most often preceded by one of the unknown N-terminal domains. The second most-commonly observed organization is a two-domain structure, which lacks any of the predicted lytic domains, containing only one of the four unknown N-terminal domains and one of three C-terminal domains. In fact, only one LysA, Barnyard gp39, showed an organization similar to other phage lysins with an N-terminal amidase domain and a C-terminal domain with LGFP binding motifs. To date, only a few lysins of other phages have been identified that have three domains, and in all cases, the third domain is an extra lytic domain at the N-terminus (Navarre et al., 1999; O'Flaherty et al., 2005; Oshida et al., 1995; Pritchard et al., 2004). The two-lytic domain organization is seen in a few LysAs (*e.g.* Che9d gp35, Giles gp31, Halo gp27, and Wildcat gp49); however, the three-domain architecture with the unknown N-terminal domain seems to be a hallmark of the mycobacteriophage LysAs. The reason for this is unknown, but the most frequently identified LysA organizations contain both an N-terminal domain and a central PG hydrolytic domain, so the combination of the two may be important for function. However, the existence of two-domain LysAs with no predicted PG hydrolytic activity suggests that the unknown N-terminal domains may have lytic activity.

The majority of the PG hydrolytic activities identified in the LysAs (amidase, muramidase, and peptidase) have also been observed in other phage lysins, although there may be some modifications to accommodate the slight differences in mycobacterial PG (Figure 5).

There were, however, some interesting observations related to the origins of the different domains. The amidase domains were divided into two groups, Ami1 and Ami2, based on a complete lack of sequence similarity between the two; the Ami1 domain shows weak resemblance to *Corynebacterium* and *Propionibacterium* amidases, while the Ami2 domain is similar to mycobacterial proteins. This suggests two separate acquisition events in the past, similar to what has been proposed for the acquisition of the GH19 family chitinase genes in purple bacteria; based on a phylogenetic analysis, the GH19 genes in purple bacteria form two clusters that are independently traced back to ancestral plant GH19 chitinases (Udaya Prakash et al., 2010). Regardless of the dissimilarities in sequence, both groups appear to have all of the conserved catalytic residues, and preliminary zymography studies have shown PG hydrolytic activity by Bxz1 gp236 (Ami2), Che8 gp32 (Ami1), and Corndog gp69 (Ami2) (Appendix A.1). Two other PG hydrolytic domains, the glycoside hydrolase family 25 (GH25) and M23 family metallopeptidase (M23), have been identified in both phage and host PG hydrolytic proteins. Alignments for each of these domains show a conservation of the catalytic and other important residues. However, experimental evidence for their PG hydrolytic ability has not yet been obtained.

Transglycosylase (TG) activity is rarely seen in phage lysins outside of the λ R lysin and a few other Gram-negative lysins (Bienkowska-Szewczyk et al., 1981; Briers et al., 2007). This makes Hammer gp13 rather unique in its possession of a TG domain in place of the GH19 domain found in other Cluster A phages. Transglycosylases are muramidases, but are distinct in that they cleave the glycosidic bond between MurNAc and GlcNAc to form a 1,6-anhydro-MurNAc residue (Figure 6). Conversely, lysozome-type muramidases hydrolyze the same bond but produce MurNAc residues with a reducing end. TG motifs, also referred to as Rpf motifs,

due to the similarity in sequence to resuscitation-promoting factors (Skeiky and Sadoff, 2006), are often found in the phage tape measure proteins (Marinelli, 2008; Moak and Molineux, 2000; Pedulla et al., 2003; Piuri and Hatfull, 2006). Hammer gp13 is the first mycobacteriophage LysA identified as having a TG domain. Mycobacteriophage Hammer is a member of the larger Cluster A, whose lysins consist of N1, N4, GH19, and C1 domains. The TG domain of Hammer gp13 shows 71% sequence similarity with regions of the tape measure proteins of Cluster D phages. This level of similarity argues for a recombination event between proteins with two different functions from mycobacteriophages in two different clusters.

The N-terminal domains of LysA proteins appear to be completely novel and are thus far unique to phages infecting mycobacteria. The four domains are found in different combinations with other domains, either with glycosidases (N1 and N4) or amidases (N2 and N3); peptidases are always found N-terminal to another activity, so they are not associated with any N-terminal domains. In Chapter 3, I speculated on many possible functions for the N-terminal domains; here, I will focus on those possibilities with the most experimental support.

There are very few two-domain LysAs, and the majority of these consist of an N1 or N4 domain with a C1 or C3 domain, or with an unmatching C-terminal region as seen in Brujita gp29. The absence of an identifiable PG hydrolase domain in these LysAs suggests that the N1 or N4 domains may in fact be novel PG hydrolases. Indeed, the Cluster D and Sub-cluster H1 LysAs, as well as Brujita gp29, all possess PG-binding motifs; the sole exception is L5 gp10 (Cluster A), which has an N4-C1 organization and is the only N4 LysA that lacks a PG hydrolytic domain. The presence of a PG-binding motif implies that the activity of the LysA is targeted to the PG. Indeed, some form of lytic activity has been observed for three of the LysAs possessing N4 domains: D29 gp10 (N4-GH19-C1), Kostya gp33 (N4-GH25-C1), and L5 gp10

(N4-C1). Upon endogenous expression of these LysAs in *M. smegmatis* (Appendix A.3), L5 gp10 showed the most activity, efficiently lysing all cells, followed by the slightly less effective D29 gp10 and the moderately lytic Kostya gp33. However, endogenous expression of Plot gp36 (N1-C3) and Brujita gp29 (N4-PGBD, unmatched) did not result in lysis. D29 gp10 has shown some PG hydrolytic activity on zymograms (Appendix A.1), but this could be attributable to the GH19 domain. An *in vitro* assay for PG hydrolysis by purified L5 gp10, Plot gp36, and Brujita gp29 would be very informative, and to date, similar studies with several LysAs have used zymography for this purpose (Appendix A.1). However, a negative result would be uninformative, since this could result from the dissimilarities between the types of PG used in the assay (*Micrococcus luteus* and, to a lesser extent, *Bacillus subtilis*), the uncertainty of a successful renaturation, or to the fact that conditions for optimal hydrolysis are unknown.

The N2 and N3 domains are always followed by an amidase domain, making it unlikely that the N2 or N3 domains themselves possess PG hydrolytic activity. A putative role may have been identified for the N3 and possibly the N2 domain in the recent characterization of the accessory lysis protein gp1 of mycobacteriophage Ms6 (Catalao et al., 2010). It was found that gp1 acts as a chaperone that binds to the Ms6 gp2 LysA and aids in its export into the periplasm through the sec system. The LysA appears to become active upon membrane depolarization, which is similar to the SAR lysins, except that in this case, an accessory protein, rather than a signal peptide, mediates translocation (Xu et al., 2004). Catalao *et al*. (2010) have also found that the N-terminal 60 amino acids are necessary and sufficient for export. The Ms6 gp2 LysA is highly similar to the Fruitloop gp29 LysA (with a N3-A1-C2 orgamization), and homologs of Ms6 gp1 are seen in the lysis cassettes of other Cluster F1 phages, all of which encode LysAs containing both N2 and N3 domains. These observations suggest the N2 and N3 domains may

have a role in the export of the LysA by binding gp1-like chaperones. Experimental evidence may further support the idea of a non-PG hydrolytic function for N3 and possibly N2; zymography with purified Corndog gp69 (N3-Ami2-C2) revealed highly active degradation products that were discovered to have lost most of the N3 domain (Appendix A.1). Similar fragments were found in zymograms with Bxz1 gp236 (N2-Ami2-C1), but the loss of the N2 domain in this case has not been verified.

There is significantly less diversity in the C-terminal domains of the LysAs, but this is not unexpected. C1, C2, and C3 dominate the organizations found, although these groups are very broad; the C1 domains from two different LysA proteins could be completely unrelated as membership in a particular group only requires identity to at least one other member of the group, similar to the relationships observed in the numerous members of some phamilies (Hatfull et al., 2008). The decreased variety of C-terminal domains is likley a result of evolution directed to bind a select set of cell wall components. These domains must bind structures that are not only specific to their host, but also that cannot easily be altered in any way that would disrupt binding or remove the substrate altogether. Because of this, lysins have evolved to target cell wall components that are essential for the organism's growth; for example, many lysins of Gram-positive phages target the polysaccharide-based teichoic acids (Garcia et al., 1988; Giudicelli and Tomasz, 1984; Usobiaga et al., 1996). Further, some *Streptococcus* phage lysins have evolved to bind only to choline-containing teichoic acids, which restricts their range to specific *Streptococcus* species (Hermoso et al., 2003). While several LysAs contain a more widely recognized PG-binding motif in the C3 domain, the motifs identified in the C1 and C2 domains could represent novel binding motifs that are specific to essential mycobacterial cell wall components. Two likely candidates for binding substrates are the arabinogalacatan and the

lipoarabinomannan, two polysaccharides that are attached to the PG and cell membrane, respectively (Brennan and Nikaido, 1995).

4.1.2   The Novelty of the LysB Proteins and Comparison to Other Host-Specific Lysis Proteins

As discussed above, a more complex barrier to progeny phage release may necessitate additional lysis proteins. The identification of *lysB* as a second putative lysis protein by Garcia *et al.* (2002) was originally based solely on its consistent location downstream of *lysA* and the prevalence of homologs in the majority of mycobacteriophage genomes. A later study identified the Gly-X-Ser-X-Gly motif characteristic of lipolytic enzymes and showed enzymatic activity on lipase and esterase substrates (Gil et al., 2008). However, this artificial activity did not explain LysB's presence in the lysis cassette or identify its natural substrate. Still, both a substrate and role for LysB could be hypothesized merely by looking at the mycobacterial cell wall. A LysA is identifiable in all genomes, and a holin protein can be predicted in many, but these proteins only compromise the cell membrane and PG complex, leaving the lipid-rich mycobacterial outer membrane largely intact. This outer membrane is reminiscent of that in Gram-negative bacteria, and the phages infecting these species are known to encode proteins – Rz/Rz1 or spanins – that remove this barrier by fusing it with the inner membrane (Berry et al., 2008; Summer et al., 2007).   It would be reasonable to predict that an analogous system should exist in mycobacteriophage to address the mycobacterial outer membrane.

Thus, the predicted role of LysB is to remove this final barrier to lysis, similar to the role of Rz/Rz1 proteins, and this has been verified experimentally. LysB has been shown to be a serine esterase capable of hydrolyzing purified *M. smegmatis* cell wall to release mycolic acids,

which constitute much of the outer membrane. The removal of this barrier greatly increases the efficiency of lysis, as is demonstrated by the impaired lysis and phage release observed in a GilesΔ*lysB* mutant. However, as with the Rz/Rz1 proteins (Berry et al., 2008), LysB is not essential for plaque formation but merely enhances it. This is logical, as the outer membrane is connected to the PG complex through the mycolylarabinogalactan, and so as LysAs degrade the PG, they are also weakening the attachment of the outer membrane to the cell wall and compromising the entire structure.

The evolutionary advantage likely granted by LysB has resulted in its inclusion in the lysis cassette of 56 of the 60 published mycobacteriophage genomes; however, its absence in the remaining four (Che12, Myrna, Qyrzula, and Rosebush) is curious. It is possible that these phages have simply not needed the additional help, perhaps due to the lack of competition by other phages that would select for such an advantage in quick progeny release. Alternatively, these phages may infect a different host in which mycolic acids are either not present or not essential for growth, such as in *Corynebacterium* (Portevin *et al.,* 2004). No LysB-homologs have been found in the few phages characterized as infecting non-*Mycobacterium* bacteria that possess mycolic acids (Bukovska et al., 2006; Chen et al., 2008). It is unclear whether these phages encode other proteins with analogous functions; no other predicted lipolytic enzymes have been identified. The circumstances by which these mycobacteriophages came to be without a LysB may differ. The area surrounding the lysis cassette of Che12 is syntenic with other Cluster A phages except for the absence of LysB, and there is no alternate protein in its place, suggesting that Che12 may simply have lost its *lysB* gene. Rosebush and Qyrzula both possess an additional gene downstream of *lysA* that is not present in fellow Cluster B phages, all of which encode a LysB. This small additional protein (178 aa) is predicted to have two

transmembrane domains, similar to class II holins (Wang et al., 2000); no putative holins are found in the lysis cassettes of the other Cluster B phages, which is in itself, a perplexing observation. Lastly, in Myrna, there are two proteins, gp244 and gp245, immediately downstream of the LysA gene with no similarity to other mycobacteriophage proteins. However, gp244 appears to share some similarity with unclassified regions in several bacterial PG hydrolases, so it is a good candidate for another lysis protein, perhaps replacing the role of LysB.

We have observed that most mycobacteriophage lysis cassettes contain at least three proteins: a LysA endolysin, a putative holin, and a LysB protein, which has an important role in removing the final barrier to lysis. However, there are many other predicted ORFs surrounding *lysA* and *lysB* that may play a role in mycobacteriophage lysis. Most recently, Ms6 gp1 was identified as a chaperone for the Ms6 LysA protein (Catalao et al., 2010). Although gp1 homologs have only been identified in some phages of Sub-Cluster F1, small ORFs are found upstream of many *lysA* genes in other mycobacteriophages, while others have intergenic regions upstream of the *lysA* that may contain promoters or unrecognized ORFs. Most mycobacteriophage lysis cassettes contain a LysA followed by a LysB and putative holin, all of which are adjacent, but several have two or three small genes intervening between the *lysA* and *lysB* genes, the majority of which are unique to mycobacteriophage and have no currently assigned function. The variety and complexity of the mycobacteriophage lysis cassettes suggest that we have only just begun to understand the strategies that these phages have evolved to solve the problem of progeny phage release in *Mycobacterium* species.

4.2 A FUTURE FOR MYCOBACTERIOPHAGE LYSINS IN RESEARCH AND THERAPY?

The research on mycobacteriophage lysins reaches beyond the objective of increasing understanding of mycobacteriophage lysis mechanisms. There are several potential applications for mycobacteriophage lysins, both as tools for research and as therapeutics.

4.2.1 Applications of Mycobacteriophage Lysins in Research

Much of the recent work from our lab has taken advantage of mycobacteriophages and the genetic tools they provide to augment current research and clinical methods. Several recent examples include the design of a mycobacterial recombineering system using gp60 and gp61 from mycobacteriophage Che9c (van Kessel and Hatfull, 2007), the adaptation of that system to engineer mutations and deletions in mycobacteriophages (Marinelli et al., 2008), and the creation of a fluorescently labeled mycobacteriophage for detection of *M. tuberculosis* in clinical samples and for the assessment of antibiotic sensitivities (Piuri et al., 2009). The mycobacteriophage lysins have the potential to aid in cell wall studies by selectively hydrolyzing certain fragments. For example, mycolic acids are currently removed from purified mAGP by harsh alkaline hydrolysis with tetrabutyl ammonium hydroxide overnight at 100°C (Parish and Stoker, 1998); LysB is capable of hydrolyzing mycolic acids at a much faster rate at room temperature in a basic buffer. LysA proteins can also be employed to selectively hydrolyze portions of the PG to isolate specific fractions for further study. Purified mAGP − or AGP if the mycolic acids are removed by LysB − could be incubated with an amidase LysA such as Bxz1 gp236 or Corndog gp69, both of which have demonstrated PG hydrolytic activity (Appendix A.1), in order to

separate the glycan strands from the peptide cross-links, which could then be isolated and analyzed separately.

Another interesting application may be derived from the recently observed, and as yet not understood, ability of L5 gp10 to rapidly lyse *M. smegmatis* cells when expressed endogenously. Currently L5 gp10 is on the pLAM12-derived plasmid pKMC11 under the control of an acetamidase promoter, which is well-repressed in the absence of acetamide in *M. smegmatis*. However, when induced with 0.2% acetamide, a log-phase culture begins to lyse after 2 hours and is nearly completely lysed within 1-2 hours of induction (Appendix A.3), similar to the lysis observed during a phage infection. Interestingly, this occurs in the absence of a holin, or any other apparent mechanism for transport across the cytoplasmic membrane. Currently mycobacteria can be sonicated or put through a French pressure cell press in order to rupture the cells and gain access to intracellular materials, but these methods can be inadequate in breaking the strong cell walls or separating clumped bacteria. Mycobacteria with L5 gp10 on a plasmid or integrated into the genome can be efficiently lysed on command by the addition of 0.2% acetamide. In fact, while most proteins cloned into pLAM12 require induction in succinate media to prevent repression of the acetamide promoter by the ADC present in normal mycobacterial media, induced L5 gp10 is capable of lysing the culture even in the presence of ADC (data not shown). However, since the mechanism of lysis by L5 gp10 and any other effects on the cell are not yet fully understood, this method should be used with caution, although this is another interesting potential future area of investigation.

4.2.2   Mycobacteriophage Lysins as Anti-mycobacterial Agents

Tuberculosis (TB), caused by *M. tuberculosis,* is the number one cause of death due to an infectious agent; more than 1.3 million deaths were reported in 2008 (W.H.O., 2009).   The bacterium is present in more than one-third of the world's population, although only 5-10% of those who are infected with *M. tuberculosis* develop active TB, while the remaining 90-95% maintain an asymptomatic latent infection (W.H.O., 2010).   However, these numbers do not account for people co-infected with human immunodeficiency virus (HIV).   TB is the leading cause of death among HIV-positive individuals, whose immune systems weaken until the bacterium can no longer be suppressed (W.H.O., 2010).   Currently the World Health Organization (WHO) is conducting a program focused in Sub-Saharan Africa called The Global Plan to Stop TB, in an effort to reduce the occurrence of TB and curb the development of drug-resistant *M. tuberculosis* (W.H.O., 2009).   The strategy centers on providing all populations with access to quality health care and TB treatments, as well as promoting research into new diagnostic methods, drugs, and vaccines.

TB is primarily a pulmonary disease; aerosolized *M. tuberculosis* particles are inhaled and multiply, eventually activating the immune system, which sends macrophages to the site of infection (Tripathi et al., 2005).   Many bacteria are eliminated by the macrophages, but others survive phagocytosis and replicate within the macrophages.   As cells die, more macrophages aggregate around the site to suppress the infection.   This results in the formation of a granuloma, a collection of macrophages and other immune cells that mediate containment of the mycobacteria (Saunders and Cooper, 2000).   The mycobacteria may ultimately form a caseating granuloma in which they lie dormant within necrotic tissue in the center of the granuloma.   This is the basis for a latent infection and the reason that 6 to 9 months of antibiotic therapy is needed

to eradicate these hard to reach dormant or persister cells (Connolly et al., 2007). The granulomas are not absolute in their containment of the bacterium, and as a result, bacteria-laced caseous material can be expelled from the lungs during coughing (Ducati et al., 2006), thereby disseminating *M. tuberculosis* bacteria to new hosts. This can be especially problematic in crowded populations (Tripathi et al., 2005).

Current methods for the control of TB focus on vaccination and chemotherapeutic treatment. The Bacilli Calmette-Guerin (BCG) vaccine is currently used, but its efficacy at preventing pulmonary TB varies, especially in adults (Tripathi et al., 2005). Much research is currently being conducted to improve the BCG vaccine and develop new vaccines (Skeiky and Sadoff, 2006). Currently, most TB infections are treated with antibiotics, primarily rifampicin and isoniazid. Rifampicin interacts with the β subunit of the prokaryotic RNA polymerase, and isoniazid inhibits mycolic acid synthesis, leading to rapid lysis of mycobacteria (Di Perri and Bonora, 2004). These are considered first-line drugs, and resistance to both defines multidrug-resistant TB (MDR-TB) (Telenti and Iseman, 2000). Resistance requires extending treatment or proceeding to second- and third-line drugs, which are increasingly toxic to humans. A few years ago, the Centers for Disease Control defined a new group, the extensively drug-resistant TB (XDR-TB), which are resistant to isoniazid, rifampicin, and at least three of the six classes of second-line drugs (CDC, 2009).

The spread of antibiotic resistance, the general phenotypic resistance of latent *M. tuberculosis*, and the insufficiency of current vaccines make the need for novel treatment options for TB all the more pressing. Given the success of lysins in treating other diseases in animal models of infection (Borysowski et al., 2006; Fischetti, 2008; Hermoso et al., 2007), mycobacteriophage lysins should also be investigated as possible therapeutic agents. While new

chemotherapeutics and drug cocktails are being developed, the threat of antibiotic resistance is ever-present, and, as noted in Chapter 1, lysins have many advantages over antibiotics. Foremost of these advantages is the inability of bacteria to develop resistance to lysins. While antibiotics bind to a single protein that is capable of mutation to prevent the binding, lysins target a complex structure – the cell wall – that is the result of over 20 enzymes, and thus far no attempts to generate resistance have been successful (Loeffler et al., 2001) making mutation to prevent its degradation by lysins nearly impossible. However, not all resistance to antimicrobial treatments is due to genetic mutations; phenotypic changes in bacteria, such as the formation of a biofilm or entry into a non-replicating state, also increase resistance to treatment by antibiotics and likely lysins as well (Russell, 2001).

In addition, lysins are a more focused tool than some antibiotics. That is, although they target the PG found in all bacteria, lysins show a high degree of specificity, a result of their cell wall-binding domains that bind unique molecules on the cell surface. In contrast, the spectrum of antibiotics varies; while some are targeted – isoniazid, for example, is very specific to mycobacteria, since it targets InhA, an enzyme in the biosynthesis pathway of mycolic acids (Quemard et al., 1995) – others function more broadly. Rifampicin, for example, is most often used to treat *Mycobacterium* infections, but it also affects other Gram-positive bacteria like *Staphylococcus aureus* (Wehrli, 1983). Moreover, use of antibiotics, as with many small compounds, poses the risk of drug interactions, allergic responses, and non-specific interactions. For example, many people are allergic to penicillin (Solensky, 2003), and the antibiotic chloramphenicol targets the 50S ribosomal subunit to inhibit protein synthesis in Gram-negative and Gram-positive bacteria (Weber and DeMoss, 1969), but it can also inhibit protein synthesis in human mitochondria (Yunis et al., 1970). As substantially larger proteins that target

structures not found in eukaryotes, the risk of side effects for lysins is considerably lower, and indeed few adverse effects have been observed (Fischetti, 2008).

Additionally, the effectiveness of lysins is less influenced by the growth stage of the bacteria. Many antibiotics, especially those targeting cell wall synthesis, require the bacteria to be actively replicating to be effective, since stationary phase cells are not dividing and therefore are not synthesizing new cell wall material (Gilbert et al., 1990). Lysins do not require active cell growth to hydrolyze the PG, although alterations in the PG in different growth phases, such as increased cross-linking, can affect the activity of a lysin (Vollmer et al., 2008a).

Finally, more phage genomes are constantly being sequenced, ever expanding the genetic repertoire from which one can find new lytic genes, including traditional endolysins, new lysins such as LysB, and other antimicrobial proteins. Due to the modular nature of endolysins, an endless number of chimeric proteins can also be engineered in order to create a lysin with the optimal level of activity and specificity. These artificially-generated anti-bacterial proteins could also be altered to contain signal peptides or localization sequences and targeting domains, or to have different biophysical properties such as an altered overall charge or hydrophobicity. Small molecule compounds are less malleable, with more limitations on the possible alterations in structure.

Despite the advantages of lysins discussed above, many of which could apply to mycobacteriophage lysins, there are still factors that must be considered for every potential lysin-based therapeutic, such as the possibility of the protein eliciting an immune response, being inactivated *in vivo*, or releasing toxic cell wall or cytosolic compounds upon cell lysis. Moreover, there are several additional considerations particular to treatment of TB, and in these circumstances, antibiotics have several key advantages over lysins. Previous lysin studies have

focused on treating diseases with bacteria that are relatively exposed and thus accessible to lysins – either on mucosal surfaces (Fischetti, 2003) or in septic models (Rashel et al., 2007). However, much of the *M. tuberculosis* bacteria reside in masses of necrotic tissue encased by macrophages, *i.e.* granulomas. This is not as much of a problem for antibiotics, but it will likely present a barrier to larger proteins.

Not only have lysin therapies focused on extracellular bacteria, but additionally, so far only Gram-positive pathogens have been successfully treated (Borysowski et al., 2006; Hermoso et al., 2007; Loessner, 2005). While technically *Mycobacterium* species are Gram-positive bacteria that lack a traditional outer membrane and contain a relatively thick layer of PG, they also have a unique outer membrane structure that, in some ways, is analogous to the outer membrane of Gram-negative species. For a small molecule compound with the proper physiochemical properties, diffusing past these barriers and into the cell is not a problem. The mycobacterial outer membrane is highly impermeable and hydrophobic, but the anti-mycobacterial antibiotics rifampicin and isoniazid are both moderately lipophilic and thus capable of passing into the cell (Tripathi et al., 2005). Having evolved to lyse the cell from within during phage progeny release, the substrates of both LysA and LysB are on the opposite side of the mycobacterial outer membrane, and the passage of large proteins through a densely packed lipid membrane is problematic.

Thus the immediate problem with any treatment of TB using mycobacteriophage lysins is access, both to the cell through the granuloma and then to the substrate underneath the membrane. However, the latter problem may already be solved – by the mycobacteriophage tape measure protein. Phages need to effectively inject their DNA into the cell, which requires the insertion of the phage tail through the cell envelope, even one as complex as that of

Mycobacterium. An analysis of the mycobacteriophage tape measure proteins revealed several features consistent with this objective (Marinelli, 2008; Pedulla et al., 2003; Piuri and Hatfull, 2006). Despite their enormous diversity, mycobacteriophage tape measure proteins have a characteristic amino acid composition, which in some ways (*e.g.* a high Ala-Gly content) resembles that of eukaryotic viral proteins involved in membrane fusion and penetration (Del Angel et al., 2002). Tape measures are also predicted to possess largely α-helical secondary structures, and many are predicted to contain transmembrane domains, which would better enable passage through the outer membrane (Marinelli, 2008). Additionally, there are several tape measure-associated PG hydrolytic motifs, some of which have demonstrated hydrolytic activity (Marinelli, 2008; Piuri and Hatfull, 2006). Of special interest is the Rpf Motif (Pedulla et al., 2003), which is similar to resuscitation promoting factors, signaling proteins that have been shown to stimulate the growth of dormant bacteria, likely through cleavage of inert PG on the surface of non-growing cells and/or at the septum (Kana and Mizrahi, 2010; Keep et al., 2006; Kell and Young, 2000). The presence of either this motif or Motif 3, a putative peptidase, appears to enhance the infection of stationary phase cells (Marinelli, 2008; Piuri and Hatfull, 2006), which may be due to localized PG hydrolysis and/or to a stimulatory effect on cell growth.

Exogenous applications of several purified LysAs and the D29 LysB to a freshly inoculated culture of *M. smegmatis* did result in some inhibition of growth, especially by the LysB (Appendix A.3), although whether this was due to a bactericidal or bacteriostatic mechanism is unknown. However, direct addition of purified LysA or LysB protein to grown cultures of *M. smegmatis* showed no effect (data not shown). If the reason for this lack of lysis is an inability to access the substrates, mycobacteriophage tape measure proteins may be able to

remedy this situation. While the tape measure proteins can have their own PG hydrolytic activity, naturally it is not sufficient to lyse the cell, since lysis prior to phage replication would not be evolutionarily advantageous. However, they are theoretically capable of penetrating the outer membrane to reach the PG. Therefore, adding purified tape measure protein in addition to LysA and/or LysB protein could create an effective cocktail that can lyse mycobacterial cells exogenously. Another possibility is the engineering of recombinant lysins that possess specific regions of the tape measure protein important for passing through the outer membrane that would then deliver catalytically active domains from LysA and/or LysB proteins to their respective substrates. There is also the potential for synergistic killing of mycobacteria by combining lysin treatment with antibiotics. In this scenario, including the Rpf motif of the tape measure (or other motifs that function similarly) could prove beneficial. Dormant *M. tuberculosis* are more resistant to antibiotics, especially those that rely on active cell growth. These motifs could serve to stimulate growth of the cells, while exogenously-applied lysins could make the cell more permeable to all compounds, thus enhancing the effects of antibiotic treatments.

While this theoretical therapeutic strategy addresses the problem of allowing lysins access to their substrates on the other side of the outer membrane, it still offers no solution for the problem of getting through the granuloma to the true persister cells. Therefore, mycobacteriophage lysins may not be capable of directly treating tuberculosis. However, *M. tuberculosis* is transmitted by coughing or sneezing, which expels aerosolized caseous material containing the bacteria. Along with its ability to cause a latent infection, the ease of dissemination of *M. tuberculosis* – especially in densely populated cities – is one of the reasons that it is so difficult to eradicate. It is estimated that, if not treated, an individual with active TB can infect 10-15 people each year (WHO, 2010). Preventing the transmission of *M. tuberculosis*

is one step of the epidemic that mycobacteriophage lysins could address; by administering lysins directly into the lungs of infected individuals, exposed mycobacteria that might have been spread to others could be killed. Naturally, such an approach would face different obstacles such as the mucosal barriers and secreted proteases in the lungs (Sheehan et al., 2006). While this would not cure the individual of TB and may not fully prevent transmission, a significant reduction in risk of transmission, especially in cities or health care settings, could go far in fighting TB.

5.0   MATERIALS AND METHODS

5.1   BACTERIAL STRAINS AND MEDIA

5.1.1   *Mycobacterium smegmatis*

M. smegmatis mc^{2}155, a high efficiency transformation strain (Snapper et al., 1990) was grown in 7H9 broth (Difco) supplemented with 10% albumin dextrose complex (ADC), 0.2% glycerol and 0.05% Tween-80 and on 7H10 agar (Difco) supplemented with 10% ADC and 0.5% glycerol.   Unless stated otherwise, the antibiotics carbenicillin (CB, 50 μg/ml) and cyclohexamide (CHX, 10 μg/ml) were included.   When required, kanamycin (KAN) was added to a concentration of 20 μg/ml.   For phage infections, 1 mM CaCl$_2$ was added and Tween-80 was omitted.   During inductions using pLAM12-derived plasmids, ADC was replaced with 0.2% succinate and acetamide was added to a final concentration of 0.2% for induced samples. Growth inhibition assays used biofilm minimal media (see 5.9.6).   Strains were stored in 20% glycerol at -80°C.

5.1.2   *Escherichia coli*

E. coli strains were grown in Luria broth (LB) or on LB agar (Difco) supplemented as needed with CB (50 μg/ml) and KAN (20 μg/ml).   High transformation efficiency GC5 (Gene Choice,

Inc.) or NEB5alpha (New England Biolabs) competent cells were used to transform ligation

reactions, and plasmids were propagated in these cells or DH5α (Sambrook and Maniatis, 1989).

BL21(DE3) cells (Novagen) were used for protein over-expression from pET21a-derived

plasmids. Strains were stored in 20% glycerol at -80°C.

5.2   GENERAL CLONING AND DNA MANIPULATIONS

5.2.1   Plasmids

Table 4. Plasmids cloned by others

Plasmid	Gene of Interest	Parental Plasmid	Insert Source	Antibiotic Marker	Origin(s)	Source	Features
pBAD/ gIIIA,B,C	None	pBR322	None	Cb	OriE	Invitrogen	Arabinose-inducible expression with secretion into the *E. coli* periplasm, optional C-terminal His-tag and c-myc
pET21a	None	pET21a	None	Cb	OriE	Novagen	IPTG-inducible T7-expression vector with optional C-terminal His-tag in *E.*
pET28b	None	pET28b	None	Kan	OriE	Novagen	IPTG-inducible T7-expression vector with optional N-terminal His-tag
pLAM3	D29 gp12	pET21a	D29	Cb	OriE	Payne and Hatfull, 2008	Derivative of pET21a
pTH01	Bxz1 gp236	pET21a	Bxz1	Cb	OriE	T. Huang, unpublished	Derivative of pET21a
pTH08	Che8 gp32	pET21a	Che8	Cb	OriE	T. Huang, unpublished	Derivative of pET21a
pLAM12	None	pJL37	None	Kan	OriE/ OriM	van Kessel and Hatfull, 2007	Acetamide-inducible expression from the P*acetamidase* in *Mycobacteria*
pJV53	Che9c gp60 and gp61	pJV53	Che9c	Kan	OriE/ OriM	van Kessel and Hatfull, 2007	Derivative of pLAM12
pNIT	None	pNIT	None	Kan	OriE/ OriM	Pandey et al., 2009	e-caprolactam-inducible expression from P*nitA* promoter in *Mycobacteria*

Cb, carbenicillin
Kan, kanamycin
OriE, *E. coli* origin of replication
OriM, *Mycobacterium* origin of replication

227

Table 5. Plasmids cloned by KP

Plasmid	Gene of Interest	Parental Plasmid	Insert Source	Antibiotic Marker	Origin(s)
pKC1	Corndog gp69	pET21a	Corndog	Cb	OriE
pKC2	D29 gp10	pBAD/gIIIC	pLAM1	Cb	OriE
pKC3	D29 gp10	pBAD/gIIIC	pLAM1	Cb	OriE
pKC4	D29 gp12	pBAD/gIIIA	pLAM3	Cb	OriE
pKC5	D29 gp12	pBAD/gIIIA	pLAM3	Cb	OriE
pKC6	Corndog gp69	pBAD/gIIIC	pKC1	Cb	OriE
pKC10	Che8 gp32	pBAD/gIIIA	pTH08	Cb	OriE
pKC11	Che8 gp32	pBAD/gIIIA	pTH08	Cb	OriE
pKC12	Corndog gp70	pET21a	Corndog	Cb	OriE
pKC13	CwlM	pET21a	*M. smegmatis*	Cb	OriE
pKC17	D29 gp12	pET21a	pLAM3	Cb	OriE
pKC20	D29 gp12 S82A mutant	pET21a	pLAM3	Cb	OriE
pKC21	Giles gp32	pET21a	Giles	Cb	OriE
pKC22	D29 gp12 H240A mutant	pET21a	pLAM3	Cb	OriE
pKC23	D29 gp12 S82A H240A mutant	pET21a	pKC20	Cb	OriE
pKC24	D29 gp12 190-233 del-mutant	pET21a	pLAM3	Cb	OriE
pKC25	L5 gp10	pET21a	L5	Cb	OriE
pKMC1	D29 gp10	pLAM12	pLAM1	Kan	OriE/OriM
pKMC2	D29 gp12	pLAM12	pLAM3	Kan	OriE/OriM
pKMC3	Bxz1 gp236	pLAM12	PTH01	Kan	OriE/OriM
pKMC4	Corndog gp69	pLAM12	pKC1	Kan	OriE/OriM
pKMC5	Che8 gp32	pLAM12	pTH08	Kan	OriE/OriM
pKMC6	D29 gp12 S82A mutant	pLAM12	pKC20	Kan	OriE/OriM
pKMC7	D29 gp12 H240A mutant	pLAM12	pKC22	Kan	OriE/OriM
pKMC8	D29 gp12 S82A H240A mutant	pLAM12	pKC23	Kan	OriE/OriM
pKMC9	D29 gp12	pNIT	pLAM3	Kan	OriE/OriM

(Table 6 Continued)

Plasmid	Gene of Interest	Parental Plasmid	Insert Source	Antibiotic Marker	Origin(s)
pKMC10	D29 gp12 S82A	pNIT	pKC20	Kan	OriE/OriM
pKMC11	L5 gp10	pLAM12	L5	Kan	OriE/OriM
pKMC12	Barnyard gp39	pLAM12	Barnyard	Kan	OriE/OriM
pKMC13	Brujita29	pLAM12	Brujita	Kan	OriE/OriM
pKMC14	Bxz2 gp11	pLAM12	Bxz2	Kan	OriE/OriM
pKMC15	Kostya gp33	pLAM12	Kostya	Kan	OriE/OriM
pKMC16	Giles gp31	pLAM12	Giles	Kan	OriE/OriM
pKMC17	PLot gp36	pLAM12	Plot	Kan	OriE/OriM
pKMC18	Rosebush gp46	pLAM12	Rosebush	Kan	OriE/OriM
pKMC19	Bxb1 gp8	pLAM12	Bxb1	Kan	OriE/OriM
pKMC20	Halo gp27	pLAM12	Halo	Kan	OriE/OriM
pKMC21	Wildcat gp49	pLAM12	Wildcat	Kan	OriE/OriM
pKMC22	Myrna gp240	pLAM12	Myrna	Kan	OriE/OriM

Cb, carbenicillin
Kan, kanamycin
OriE, *E. coli* origin of replication
OriM, *Mycobacterium* origin of replication

5.2.2 Primers and Oligonucleotides

Table 6. Cloning primers

Primer	Sequence	Restriction Site	Description
Barnyard-NheIR	TGT TGG TTG CTA GCT CAT ACG AAT TCG TTG TCC	Nhe I	RCP for Barnyard gp39
Barnyard-XhoIR	TGT TGG TGC TCG AGT ACG AAT TCG TTG TCC	Xho I	RCP for Barnyard gp39
BarnyardA-NdeIF	TGG TGG TTC ATA TGT CTC GAC GTG GTG ATG TAG	Nde I	FCP for Barnyard gp39
BrujitaA-HindIIIR	TGT GTG GTA AGC TTC TTC ATC ACC ACA CCG	Hind III	RCP for Bxb1 gp8
BrujitaA-NdeIF	TTT TTT TTC ATA TGC CGA TTC TTC GCG CGA ACG	Nde I	FCP for Bxb1 gp8
BrujitaA-NheIR	TTT GTT TTG CTA GCT CAC TTC ATC ACC ACA CCG	Nhe I	RCP for Bxb1 gp8
Bxb1A-NdeIF	TTT TTT TTC ATA TGC CGA GGG TCG TCT ACG GGC	Nde I	FCP for Bxb1 gp8
Bxb1A-NheIR	TTT TTG TTG CTA GCT TAA GCG GCA TGG AAC TTC	Nhe I	RCP for Bxb1 gp8
Bxb1A-XhoIR	TTT TTG TTC TCG AGA GCG GCA TGG AAC TTC	Xho I	RCP for Bxb1 gp8
Bxz1A-HindIIIR*	TGT GTG TTA AGC TTT CAA CGG TTC TTC TCC TGC CAG	Hind III	RCP Bxz1A w/ TGA
Bxz1A-NdeIF	TTG TTG TTC ATA TGC CCG CAT ACA CAC AGG ACG G	Nde I	FCP Bxz1A gp236
Bxz1gp236-XbaIR	TTG GTG TCT AGA GTC GGT TCT TCT CCT GC	Xba I	RCP Bxz1gp236 into pBAD/gIIIA w/o His
Bxz1gp236-XbaIRHis	TTG GTG TCT AGA AGT CGG TTC TTC TCC TGC	Xba I	RCP Bxz1gp236 into pBAD/gIIIA w/ His Tag
Bxz1gp236-XhoIF	TTT TTT CTC GAG CCC GCA TAC ACA CAG GAC	Xho I	FCP Bxz1gp236 into pBAD/gIIIA
Bxz1gp236F	GTT TTG GTC ATA TGC CCG CAT ACA CAC AGG A	Nde I	FCP Bxz1gp236
Bxz1gp236R	TGG TGT GCT AGC TCA TCG GTT CTT CTC C	Nhe I	RCP Bxz1gp236
Bxz2A-NdeIF	TGG TGT TTC ATA TGA CGG AGA AGG TAC TTC CCT	Nde I	FCP for Bxz2 gp11
Bxz2A-NheIR	TTT TTT TTG CTA GCT CAG TTC TCC CCG TAC GG	Nhe I	RCP for Bxz2 gp11
Bxz2A-XhoIR	TTG TTT TTC TCG AGG TTC TCC CCG TAC GG	Xho I	RCP for Bxz2 gp11
CdogA-HindIIIR	GGT GGT CCA AGC TTT CAC TTG GTC TTG GCC TTC	Hind III	RCP for CorndogA
CdogA-XhoIR*	TTT TCT CGA GTC ACT CCT TGG TCT TGG CCT TC	Xho I	RCP CorndogA w/ TGA
CdogB-XhoIR*	TTT TTT TTC TCG AGT CAT CAG GAC CGG AGG TAT TCG ATG GC	Xho I	RCP CorndogB w/ TGA
Cgp69-NcoIF	TTT TTT TTC ATG GCG GCC AC CAA AGA CC	Nco I	FCP Corndoggp69 into pBAD/gIIIC
Cgp69-NheIR	TGT GGT TTG CTA GCT CAC TCC TTG GTC TTG	Nhe I	RCP Corndog gp69
Cgp69-XbaIR	TTG TTG TTC TAG AGC TCC TTG GTC TTG GCC	Xba I	RCP Corndog gp69 into pBAD/gIIIC w/o His
Cgp69-XbaIRHis	TTG TGG TTC TAG AAG CTC CTT GGT CTT GGC	Xba I	RCP Corndoggp69 into pBAD/gIIIC w/ His Tag
Cgp69F2	TTT GTT CTC ATA TGG CGG CCA CCA AAG ACC AG	Nde I	FCP corndog gp69
Cgp69R	TTT TCT CGA GCT CCT TGG TCT TGG CCT TC	Xho I	RCP corndog gp69
Cgp70-NdeIF	TTG TTT GTC ATA TGC GGA TCG GCG GCG AGT ATG TA	Nde I	FCP Corndog gp70
Cgp70-XhoIR	TTT TTT TTC TCG AGT CAG GAC CGG AGG TAT TCG ATG GC	Xho I	RCP Corndog gp70
Che8A-HindIIIR*	TGT TTT GTA AGC TTT CAG GAC TTG CCC TTG ACG TCG	Hind III	RCP Che8A w/ TGA
Che8A-NdeIF	TTG TGG TTC ATA TGG TGA GCT TCA CGT GGT CC	Nde I	FCP Che8A gp32
Che8gp32-NcoIF	TGT TTT TTC ATG GCA AGC TTC ACG TGG TT	Nco I	FCP Che8gp32 into pBAD A
Che8gp32-NdeIF	TGG TGG TTC ATA TGA GCT TCA CGT GGT CC	Nde I	FCP Che8gp32

FCP, forward cloning primer
RCP, reverse cloning primer

230

(Table 7 Continued)

Primer	Sequence	Restriction Site	Description
Che8gp32-NheIR	TTG TGT TTG CTA GCT CAG GAC TTG CCC TTG	Nhe I	RCP Che8gp32
Che8gp32-XbaIR	TGT TGG TTT CTA GAG GGA CTT GCC CTT GAC	Xba I	RCP Che8gp32 in pBAD/gIIIA w/o His
Che8gp32-XbaIRHis	TGT GGT TTT CTA GAA GGG ACT TGC CCT TGA C	Xba I	RCP Che8gp32 in pBAD/gIIIA w/ His tag
CJW1A-NdeIF	TGT GTT TTC ATA TGA GTG TAA CTC GCG CGA ACG	Nde I	FCP for CJW1 gp32
CJW1A-NheIR	TTT GGT TTG CTA GCT TAT GCC TTC TGT GAG TCG	Nhe I	RCP for CJW1 gp32
CJW1A-XhoIR	TTT GGT TTC TCG AGT GCC TTC TGT GAG TCG	Xho I	RCP for CJW1 gp32
CwlM-EcoRIR2	GAA GGG CGA ATT CCA GGC GCC GCG G	Eco RI	RCP CwlM
CwlM-NdeIF	TTG TTG TTC ATA TGT CGA GTC TGC GTC GCG GTG ATC	Nde I	FCP CwlM
D2910-12-PmeIR	GGG TCG CGG TTT AAA CTC AGA TCT GTC GTA G	Pme I	RCP D29 gp10-12
D2910-12-PmlIF	TGG TGG TTC ACG TGA TGA CGC TCA TAG TC	Pml I	FCP D29 gp10-12
D29A-BstBIR	TGG TGT GGT TCG AAT CAT AGG GCT CCA TTC CTG	BstB I	RCP D29A
D29A-HindIIIR	TGG TGG TAA GCT TTC ATA GGG CTC CAT TCC	Hind III	RCP for D29A pNIT
D29A-NdeIF	GGT GTG GTC ATA TGA CGC TCA TAG TCA CA	Nde I	FCP D29A
D29A-XhoRI*	TTG TGT CTC GAG TCA TAG GGC TCC ATT CCT G	Xho I	RCP D29A w/ TGA
D29B-BstBIR	GGG TGG TGT TCG AAT CAG ATC TGT CGT A	BstB I	RCP D29B BstBI site
D29B-HindIIIR*	TGG TGT AAG CTT TCA GAT CTG TCG TAG GAA CTC GAC	Hind III	RCP D29B w/ TGA
D29B-NdeIF	TGT TTT GTC ATA TGA GCA AGC CTG GCT GTT CA C	Nde I	FCP D29B
D29B-NotIR	TTT TTT TTG CGG CCG CGA TCT GTC GTA GGA ACT C	Not I	RCP D29B
D29B-S82AF	GAT GGC GGG TTA CGC GCA GGG AGC CAT CG		FSDMP for S82A
D29B-S82AR	CGA TGG CTC CCT GCG CGT AAC CCG CCA TC		RSDMP for S82A
D29gp10-NcoIF	TTG TTG TTC CAT GGC GCT CAT AGT CAC ACG	Nco I	FCP D29gp10 into pBAD/gIIIC
D29gp10-XbaIR	TTG TTT TTC TAG AGG GCT CCA TTC CTG GCG	Xba I	RCP D29gp10 into pBAD/gIIIC w/o His Tag
D29gp10-XbaIRHis	TTG TTT TTC TAG AAG GGC TCC ATT CCT GGC G	Xba I	RCP D29gp10 into pBAD/gIIIC w/ His Tag
D29gp10F	GTT TGG GTC ATA TGA CGC TCA TAG TCA CAC GC	Nde I	FCP D29gp10 w/ Nde I
D29gp10R	TTG TTG TGC TAG CTC ATA GGG CTC CAT TCC	Nhe I	RCP D29gp10 w/ Nhe I
D29gp12-NcoIF	TTT TTT TTC CAT GGC AAG CAA GCC CTG GCT G	Nco I	FCP D29gp12 in pBAD A
D29gp12-XbaIR	GTG GGT TGG TCT AGA GGA TCT GTC GTA GGA A	Xba I	RCP D29gp12 in pBADgIIIA w/o His
D29gp12-XbaIRHis	GTT GGG TGG TCT AGA AGG ATC TGT CGT AGG	Xba I	RCP D29gp12 in pBAD/gIIIA w/ His tag
D29gp12F	TGG TTT TTC ATA TGA GCA AGC CTG GCT G	Nde I	FCP D29gp12
D29gp12R	GTT GGG TGT GCT AGC TCA GAT CTG TCG TA	Nhe I	RCP D29gp1
D29H-BstBIR	TGG TTG TTT TCG AAT CGG TTC CAG GGC TG	BstB I	RCP D29H
D29H-NdeIF	GTG GTG TTC ATA TGA GCC CCA AGA TCC GTG AA	Nde I	FCP D29H
D29H-NheIR	TTG TTT TTG CTA GCT CGG TTC CAG GGC TG	Nhe I	RCP D29H
GilesA-NdeIF	TTT GTT TTC ATA TGC CGG TGT ACG CCA TCC AAT	Nde I	FCP for Giles gp31
GilesA-NheIR	TTT TTT TTG CTA GCT CAG AGC ACC AAC CCA ATC	Nhe I	RCP for Giles gp31
GilesA-XhoIR	TTT TTG TTC TCG AGG AGC ACC AAC CCA ATC	Xho I	RCP for Giles gp31
GilesB-HindIIIR	TTT TGT TTA AGC TTT CAT GCG GTG ACG GCC G	Hind III	RCP for GilesB pNIT

FCP, forward cloning primer
RCP, reverse cloning primer
FSDMP, forward site-directed mutagenesis primer
RSDMP, reverse site-directed mutagenesis primer

231

(Table 7 Continued)

Primer	Sequence	Restriction Site	Description
GilesB-NdeIF	TGT TGT TGC ATA TGG CGT GGA AAC CTA CCG AA	Nde I	FCP GilesB
GilesB-XhoIR	TTT TTT TTC TCG AGT GCG GTG ACG GCC G	Xho I	RCP GilesB
HaloA-NdeIF	TTT TGT TTC ATA TGG CCG ATC GTT TCT TCC CGA	Nde I	FCP for Halo gp27
HaloA-NheIR	TTT TTT TTG CTA GCT CAG AAC ACC AGC CCG A	Nhe I	RCP for Halo gp27
HaloA-XhoIR	TTT TTT TTC TCG AGG AAC ACC AGC CCG A	Xho I	RCP for Halo gp27
L5A-NdeIF	TTT GTT TTC ATA TGA CCT TCA CAG TCA CCC GCG	Nde I	FCP for L5 gp10
L5A-NheIR	TTG TTT TTG CTA GCT CAT AGG CCA CCT CTT TCT	Nhe I	RCP for L5 gp10
L5A-XhoIR	TTG GTT GTC TCG AGT AGG CCA CCT CTT TCT	Xho I	RCP for L5 gp10
MSMEG0194-NdeIF	GGG TGG TGC ATA TGG TGA AAT GGA TTA GAG CTC	Nde I	FCP MSMEG0194
MSMEG0194-XhoIR	TTG TTT TTC TCG AGG TTG GTC GCC GTG ACA	Xho I	RCP MSMEG0194
MyrnaA-NdeIF	GTG TGT GGC ATA TGG CGT TGA TTA CCG AGA	Nde I	FCP for Myrna gp243
MyrnaA-NheIR	TTT TGT TTG CTA GCC TAG GCA CCG AAC TGC	Nhe I	RCP for Myrna gp243
MyrnaA-XhoIR	TTT TGT TTC TCG AGG GCA CCG AAC TGC	Xho I	RCP for Myrna gp243
pKP01-ACYC-ClaIR	TGG TGG TTA TCG ATG ATA AGC TGT CAA ACA TGA GA	Cla I	RCP for fragment of pACYC184
pKP01-ACYC-DrdIF	TTG TTG TTC CGG GTC GAA TTT GCT TTC GA	Drd I	FCP for fragment of pACYC184
pKP01-ClaIF	CGG CGG CGA TCG ATA CTT ACA TTA ATT GCG	Cla I	FCP pET21a region to pACYC184
pKP01-ClaIF	CGG CGG CGA TCG ATA CTT ACA TTA ATT GCG	Cla I	FCP for pET21a promoter (into pACYC184)
pKP01-DrdIR	TTT TTT GAC CCG GTC GTC ATC CGG ATA TAG TTC	Drd I	RCP pET21a region to pACYC184
pKP01-DrdIR	TTT TTT GAC CCG GTC GTC ATC CGG ATA TAG TTC	Drd I	RCP for pET21a promoter (into pACYC184)
PLotA-NdeIF	TGT TGT GTC ATA TGG CAT TCA TCC AAA AGC AGG	Nde I	FCP for PLot gp36
PLotA-NheIR	TTT TTT TTG CTA GCT CAC TGC GAC AGC CTT C	Nhe I	RCP for PLot gp36
PLotA-XhoIR	TTG TTT TTC TCG AGC TGC GAC AGC CTT C	Xho I	RCP for PLot gp36
RosebushA-NdeIF	TTG TTT GTC ATA TGG CGC TAG GAA TGA CTC TGG	Nde I	FCP for Rosebush gp46
RosebushA-NheIR	TTT TTT TTG CTA GCT CAG TTG CTG CCT TTC CG	Nhe I	RCP for Rosebush gp46
RosebushA-XhoIR	TTT GTT TTC TCG AGG TTG CTG CCT TTC CG	Xho I	RCP for Rosebush gp46
WildcatA-NdeIF	TGG TTT GTC ATA TGG ATG TAG GGA CTT GCG TC	Nde I	FCP for Wildcat gp49
WildcatA-NheIR	TTG TTT TTG CTA GCT CAC GAT GGG TCC CAA TTC	Nhe I	RCP for Wildcat gp49
WildcatA-XhoIR	TTG TTT TTC TCG AGC GAT GGG TCC CAA TTC	Xho I	RCP for Wildcat gp49

FCP, forward cloning primer
RCP, reverse cloning primer

232

Table 7. Sequencing and verification primers

Primer	Sequence	Description
GilesBHis-SEQF	GGT TGA ACT GTC CGA CTT TCT CGG CAC	FSP for GilesBHis
GilesBHis-SEQR	CAA ACC GAA GGC ATA CAG GGC GAT CTG	RSP for GilesBHis
MBP-SEQF	TGC TGA CTG ATG AAG GTC TGG AAG C	FSP beginning in MBP tag
pET21aSEQF	TCA GCA AAA AAC CCC TCA AGA CCC G	FSP for pET21a-based sequencing
pET21aSEQR2	ATC TTC CCC ATC GGT GAT GTC GGC	RSP for pET21a (second)
pKC10SEQMF	CTC AAC GAA ACC ACA CCG GAG GAA C	FSP for middle of pKC10
pKC10SEQR	CGA GCG GTT CGA AGC CGT ACT TCT TC	RSP for middle of pKC10
pKC11SEQF	CCT ATA ACG AGT TTC CGA TCT GGT CGG	FSP for middle pKC11
pKC11SEQF	CCT ATA ACG AGT TTC CGA TCT GGT CGG	FSP for middle of pKC11
pKC2-3SEQF	TTC TGG TAC CTG GAC GCG AAG CTT GAA	FSP for middle of pKC2 and pKC3
pKC5SEQR	CAG CTT CAG CTC GAT CTG CAG GAT CA	RSP for middle of pKC5
pKC6SEQR	GAC GTC GCG CAT CGA GAT GAT CTG	RSP for middle of pKC6
pKMC1SEQMF	TCT CGC GTA CAA CCG AGC GGT AGC ACT	FSP for middle of pKMC1
pKMC3SEQMF	GTC TGG CTC AAG GAG GTC CTC GAA G	FSP for middle of pKMC3
pKP01-SEQverifF	CGC GCA GAC CAA AAC GAT CTC AAG AAG	FSP to verify pKP01
pKP01-SEQverifR	TGG CCA ATA TGG ACA ACT TCT TCG CC	RSP to verify pKP01
pLAM12SEQF (LIM20)	CCGCAGTTGTTCTCGCATACCCCATC	FSP pLAM12-based sequencing
pLAM12SEQR	GGCTCATAACACCCCTTGTATTACTG	RSP pLAM12-based sequencing
LIM75	CCA TAA GAT TAG CGG ATC CTA CCT GAC GC	FSP for pBAD (Marinelli)
LIM76	GCT ACT GCC GCC AGG CAA ATT CTG	RSP for pBAD (Marinelli)

FSP, forward sequencing primer
RSP, reverse sequencing primer

Table 8. Recombineering primers

Primer	Sequence	Description
Bxb1B-DADAPCR	AAC CCT ACG GTG CGC CGC TGG TA	DADA primer for Bxb1B deletion
Bxb1B-DiagF	AAG CGC ATC TTG AAG CGC GTC AAG C	FDP for Bxb1B deletion
Bxb1B-DiagR	GTG TTC TTC GTC TGG TCC TGG GAG ACC	RDP for Bxb1B deletion
Bxb1B-DO	GTT CCA TGC CGC TTA AGC TCG GCG ACC GGA ACC CTA CGG TGC GCC GCT GGT ACA ACA TCG ACC CTG CGA CGG ACT TTC TGC GCT CTG TGA CTT GAT ACG T	DO for Bxb1B
Bxb1B-EP1	ATC ACC GCT GCT CAG GTC CAG ATC CAG AAG TGG CTG GCT GCC GAG CAG AAG TTC CAT GCC GCT TAA GCT CGG CGA	FEP for Bxb1B recombineering
Bxb1B-EP2	GGC GGG GAG TAG CGA CTC CGG GTA GTG AAC CGC CAC TCC ACC TCC TCG TTA CGT ATC AAG TCA CAG AGC GCA GAA	REP for Bxb1B recombineering
GilesA-DiagF (LIM133)	CCA GCT TCT GTA CCG CGA CTA GCG CGT TC	FDP for GilesA
GilesB-DiagF	GAA CGT CGT GAA GCT GCA AAC GAC GTT G	FDP for GilesB deletion
GilesB-DiagR	CAA TCT TCG AGC GTG CGT AGC GGA TTG	RDP for GilesB deletion
GilesB-DiagR2	AGG ATC GTC GAA TCG GCG ACG ATG T	RDP for Giles recombineering
GilesB-DiagR3	CGT CAC CTC GTC CTC ACG TTT CCG TTT G	RDP for GilesB
GilesB-DO2	CTC TGA TGG CGT GGA AAC CTA CCG AAT ACC AAA TCG GAG ATC GGC ACC CCC ACG TCA GAG ACT GGG CCG GGC GCA CCC CGG CCG TCA CCG CAT AAC ACA C	DO for GilesB
GilesB-EP1.2	ATG TCG ACG GCA TCG TGG GTC CGC TCA CGG CCG CGA AGA TTG GGT TGG TGC TCT GAT GGC GTG GAA ACC TAC CGA	FEP for GilesB
GilesB-EP2.2	TGG CCG GGT TCG GAT TGG GCG GAT TCG TGC TCA CTG TTG CGA TCC CTT CGG TGT GTT ATG CGG TGA CGG CCG GGG	REP for GilesB
GilesB-MAMAPCR1	GCG AAG ATT GGG TTG GTG CTC TGC A	DADA-PCR primer for Giles recombineering
GilesB-MAMAPCR2	AAT ACC AAA TCG GAG ATC GGC ACC CCC A	DADA-PCR primer for Giles recombineering
GilesB-MAMAPCR3	CGA ATA CCA AAT CGG AGA TCG GCA CCC AC	DADA-PCR primer for GilesB
WildcatB-DADAPCR	TTG GCT GGG GCC TTG GGG ATA TCA AC	DADA PCR primer for WildcatB deletion
WildcatB-DiagF	GAA GGC TAT CAC TGC CTT CAT CAC CAC T	FDP for WildcatB deletion
WildcatB-DiagR	CGT TCC TGA TAT TCC TCA GGA AGC TCG TC	RDP for WildcatB deletion
WildcatB-DO	AAG TAA TGC GTG TAG CAG GTC AGT GGG TTG GCT GGG GCC TTG GGG ATA TCA ACT ACA ACG TGG TCC CCG CAG TCG ATT TTC TCG CAG GAC TGT AGG GTA T	DO for WildcatB
WildcatB-EP1	CCG TGG ATT CTG TTG TGA ACC AGG TTA TTC AGG AAT ACC AGT CGA AGG TTA AGT AAT GCG TGT AGC AGG TCA GTG	FEP for WildcatB deletion
WildcatB-EP2	CAG AGA CGG CAG CCG GGA TTG CTG AAA ATC GTT ACG TCC ATA CGA CAA TCA TAC CCT ACA GTC CTG CGA GAA AAT	REP for WildcatB deletion

DADA, deletion amplification detection assay primer
DO, deletion oligo
FDP, forward diagnostic primer
RDP, reverse diagnostic primer
FEP, forward extender primer
REP, reverse extender primer

5.2.3　General Cloning Procedures

Cloning of plasmids used DNA inserts generated by PCR of genomic DNA or a parental insert. PCR reactions were cleaned and concentrated using the Qiaquick PCR purification kit or Qiagen gel extraction kit. DNA inserts and vector were digested with restriction enzymes (NEB) in 20-30 μl volumes and incubated at the required temperature for approximately 3 hr. Restriction enzymes were heat-killed and 1 μl of calf intestinal phosphatase (NEB) added to the vector and incubated at 37°C for 1 hr to remove the 5' phosphate and prevent re-ligation of any partially-digested vector. Digestion products were gel-purified with the Qiagen kit and quantified using an ND-1000 Spectrophotometer nanodrop.

5.2.3.1 Ligation and Transformation

Ligations were performed with Fast-link DNA ligase (Epicentre) in 15 μl reactions containing 1 μl ligase, 1.5 μl ligation buffer and 1.5 μl ATP that were incubated at room temperature for 1 hr, followed by heat-killing by incubation at 75°C for 20 min. Ligation reactions were transformed into GC5 or NEB5α high transformation efficiency chemically-competent cells, or in some instances XL-1 Blue electrocompetent *E. coli* cells. Cells were thawed on ice and 5 μl of ligation reaction were added to 50 μl of cells and incubated on ice for 20 min.

Transformation into E. coli

Chemically-competent *E. coli* were transformed by heat-shocking for 30 sec at 42°C. Transformation into electrocompetent *E. coli* was done using an electroporator set to 2.5 kV, 25

uF, and 200 Ω. Following transformation, 450 μl of TSB was added and cells were incubated at 37°C for 30-60 min. Recovered transformations were then plated on selective media.

Transformation into M. smegmatis

Electrocompetent *M. smegmatis* mc^{2}155 cells were made as previously described (Bibb and Hatfull, 2002). Briefly, cells were grown to an OD$_{600}$ between 0.8 – 1.0 and then rested on ice for 30 min to 2 hr. Cells were then washed repeatedly with ice-cold 10% glycerol in shrinking volumes down to between 1/25 – 1/50 of the original culture volume. Cells were aliquoted, frozen on dry ice, and stored at -80°C.

5.2.3.2 Sequencing

Sequencing was done through GeneWiz, Inc. using approximately 500-800 ng of plasmid DNA and 8 pmol of a primer that was 100-200 bp upstream or downstream of the region to be sequenced.

5.2.4 PCR

5.2.4.1 Primer Design

Cloning primers were designed to be between 25 – 35 nt in length with a melting temperature around 62°C. Each has 18-25 bp homology flanking the PCR target and, if required, a restriction enzyme sequence preceded by 8 nucleotides intended to aid in restriction endonuclease digestion by enzymes requiring a significant number of nucleotides to either side of the cleavage site. The 8 nucleotides were modified to obtain the desired melting temperature. Sequencing primers were designed to have sequence homology along the full length and were 23 – 30 nt in length

with melting temperatures around 62°C. Specific primers for recombineering or Deletion Amplification Detection Assay (DADA) PCR are described below. Primers and other oligonucleotides were purchased from Integrated DNA Technologies (IDT) and resuspended with dH$_2$O to a concentration of 100 μM (stock solution) and diluted with dH$_2$O to 10 μM (working solution) unless otherwise specified. All were stored at -20°C.

5.2.4.2 Standard and Diagnostic PCR

PCR for the generation of cloning inserts and for the detection of a DNA fragment used the same protocol. PCR reactions were performed using 1U of *Pfu* polymerase (Stratagene) with 1X Pfu Buffer, 5% DMSO, 0.2 mM dNTPs, 0.5 μM primers, and varying concentrations of DNA template. The thermocycler conditions were 94°C for 5 min, 25x (94°C for 30 sec, varying annealing temperature for 30 sec, 72°C for 1 min/kb of desired product), 72°C for 7 min, and cooling at 4°C.

5.2.4.3 Colony and Plaque PCR

Colony PCR was performed as stated for the standard PCR except that the template was obtained by picking a single colony into 100 μl of dH$_2$O and boiling for 5 min, and 5 μl was used in the reaction. Plaque PCR was similarly performed except that the plaque was picked into 100 μl phage buffer and did not require boiling.

5.2.4.4 DADA-PCR

Deletion Amplification Detection Assay (DADA) PCR (Marinelli et al., 2008) used primers with homology to the region adjacent to the deletion, but the final 3' base matched the sequence of the deletion mutant and not the wildtype, thus preferentially amplifying the mutant. DADA-PCR

was performed as described for standard PCR except for the use of 1 μl of Platinum Taq High-Fidelity DNA polymerase and the addition of 2 mM $MgSO_4$.

5.2.4.5 Site-directed mutagenesis

The QuikChange XL site-directed mutagenesis kit (Stratagene) was used according to manufacturer's instructions. Briefly, the plasmid target was amplified via PCR with *Pfu* Turbo using PAGE-purified primers containing the mutation. The PCR reaction was digested with Dpn I to degrade template plasmid and then transformed into GC5 *E. coli* cells.

5.3   PROTEIN EXPRESSION AND PURIFICATION

Escherichia coli BL21(DE3) cells carrying the different *lysA* or *lysB* genes cloned into pET21a were grown to OD_{600} between 0.4 and 0.6 at $37^{o}C$ in LB containing carbenicillin (50 μg/ml), followed by induction with 1 mM IPTG for 4 h at $30^{o}C$. Cells were pelleted and resuspended in 10 ml TWEB (50 mM Tris-HCl pH 8.0, 300 mM NaCl) per 500 ml original culture volume and then frozen at $-80^{o}C$. Pellets were thawed, sonicated 6x for 30s bursts on ice, and pelleted by centrifugation at 12,500 rpm for 25 min at $4^{o}C$. The clarified lysate was added to 3 bed-volumes (BV) per original 500 ml culture of TALON Co^{2+} resin (Clonetech) pre-equilibrated with TWEB and rocked at $4^{o}C$ for 1 h. The resin was washed sequentially by rocking for 10 min at $4^{o}C$ with 10 BV each of TWEB (2x), TWEB with 10 mM imidazole (1x), and TWEB with 20 mM imidazole (2x). The resin was allowed to settle in the column and bound protein was eluted with five BV of 120 mM imidazole in TWEB. Samples of each elution were analyzed on an SDS PAGE gel and factions containing a significant amount of protein were concentrated using

Vivaspin concentration columns (molecular weight cut-off 10 kDa; Sartorius) followed by dialysis against storage buffer (50 mM Tris pH 8.0, 50 mM NaCl, 50% glycerol) and stored at -20°C. Concentrations were determined using the ND-1000 Spectrophotometer nanodrop and adjusted based on the extinction coefficient of each protein sequence as calculated by ProtParam (ExPasy).

5.4 PHAGE MANIPULATIONS

5.4.1 Lysate production and titering

Serially-diluted phage was added to 300 μl of a saturated culture of *M. smegmatis* mc^{2}155 that contained 1 mM CaCl$_2$ and no Tween. Cells were infected at room temperature standing for 30 minutes for adsorption and then added to 1.2 ml 7H9/ADC/CaCl$_2$ and 1.5 ml MBTA and poured as top agar lawn onto a 7H10 plate. After incubation at 37°C, plates with a web-like pattern of infection were flooded with 5 ml of phage buffer and placed at 37°C for 30 min. The lysate was removed and cellular and other debris pelleted by centrifugation at 5000 rpm for 5 min. The remaining lysate was filtered (0.22 μm filters), titered, and stored at 4°C.

The titer of the lysate or other samples was determined by spotting serially diluted samples on *M. smegmatis* lawns. A top agar lawn was created by adding 300 μl of a saturated culture of *M. smegmatis* mc^{2}155 to 1.5 ml 7H9/ADC/CaCl$_2$ and 1.5 ml MBTA and pouring the mixture onto a 7H10 plate with appropriate antibiotics. Phage solutions were serially diluted with phage buffer and 5 or 10 μl amounts were spotted onto the *M. smegmatis* lawn. The plates were incubated face-up at 37°C.

5.4.2   Single-step infections

The intent of the single-step infection for analysis of lysis defects required the simultaneous infection of all cells in a culture to create a synchronized infection. A calculated m.o.i. (usually 10) was added to a culture of log-phage *M. smegmatis* mc^{2}155 cells (OD$_{600}$ between 0.3 – 0.8) and allowed to adsorb at room temperature for 30 min. If necessary, viable phages remaining in the media were inactivated by the addition of 10 mM virucide (ferrous ammonium sulfate) for 5 min at room temperature (McNerney et al., 1998), followed by gently pelleting the cells at 5000 rpm for 5 min and resuspending in fresh media. Infected cultures were then incubated shaking at 37°C while measurements were taken.

5.4.3   DNA Isolation

Phage genomic DNA was isolated as previously described (Sarkis and Hatfull, 1998). Briefly, high-titer dialyzed phage was extracted with equal volumes of buffer-equilibrated phenol to remove protein. The aqueous phase was removed and repeatedly extracted with the phenol while the phenolic phases were back-extracted with TE (Tris-EDTA) buffer to increase the yield. Once the interface containing the protein was no longer visible, the final extraction used phenol:chloroform:isoamyl alcohol (25:24:1) followed by an additional extraction with chloroform. The DNA was ethanol-precipitated with 95% ethanol and pelleted at 13,500 x g for 30 min. After a final wash with 70% ethanol, the pellet was air-dried and then resuspended in TE.

5.5   BACTERIOPHAGE RECOMBINEERING

5.5.1   Deletion Mutant Construction

The phage deletion mutant was constructed as described previously (Marinelli et al., 2008). The 200 bp dsDNA substrate was generated using a PAGE-purified 100 nt deletion oligo (IDT DNA) with 50 bp of homology upstream and downstream of the deletion site, which was then expanded by PCR using two PAGE-purified 75 nt extender primers. Each extender primer was designed with 25 nt of homology to either end of the 100 nt deletion oligo and a further 50 nt homology upstream or downstream of the deletion oligo region. The 200 bp deletion substrate was verified on an agarose gel and cleaned with the QIAquick PCR purification kit (Qiagen).

5.5.2   Competent cell preparation

Electrocompetent recombineering cells were made with *M. smegmatis* mc^{2}155:pJV53 cells. The cells were grown in an inducing media (7H9, 0.2% succinate, Kan, and Tween) to an OD$_{600}$ between 0.4 – 0.6 and then induced with 0.2% acetamide for 3 hr with shaking at 37°C. After induction, the competent cells were prepared as described above (1.2.3.2).

5.5.3   Transformation and isolation of mutant phage

Purified phage DNA (350 ng) was co-electroporated with 200 ng of the 200 bp substrate into induced electrocompetent *M. smegmatis* mc^{2}155 pJV53 cells, recovered at 37°C for 2 hr in 7H9 with ADC and Tween and plated on top agar lawns with *M. smegmatis* mc^{2}155. Individual

recovered plaques were picked into 100 µl phage buffer and analyzed by DADA-PCR (Marinelli *et al.*, 2008) using 1 µl sample. Primary plaques containing both wild-type and mutant alleles were diluted and plated with 300 µl of *M. smegmatis* mc^{2}155 cells. To test for viability of the mutant, lysates from plates containing approximately 600 plaques were harvested and tested by diagnostic PCR. Individual plaques from the secondary plating were picked and also tested with diagnostic PCR. A mutant derivative was plaque purified and the deletion confirmed by DNA sequencing (GENEWIZ).

5.6 *IN VITRO* ASSAYS

5.6.1 Zymography

Zymograms were performed as described previously (Piuri and Hatfull, 2006) by incorporation of 0.2% lyophilized *Micrococcus luteus* cells as a source of peptidoglycan into the polyacrylamide gel matrix. After separation at 200V, the gel was rinsed with dH$_2$O for 15 min and then renatured overnight at 37°C in renaturation buffer (25 mM Tris pH 7.5, 1% Triton X-100 and 0.1 mM ZnSO4), stained with 0.5% methylene blue with 0.01% KOH and destained with dH$_2$O.

5.6.2 *p*-Nitrophenol Ester Assay

Enzymatic assays for lipolytic activity were adapted from those described previously (Gilham and Lehner, 2005). pNPX substrates were purchased from Sigma and Spectrum (X represents

the carbon chain length: A, acetate (1 carbon); R, propionate (3); B, butyrate (4); V, valerate (5); O, octoanate (8); C, caprate (10); L, laurate (12); M, myristate (14); P, palmitate (16). The compounds were diluted or dissolved into pure dichloromethane at a final concentration of 250 mM or 100 mM based on solubility requirements and stored at -20°C. Immediately prior to the assay, pNPX solutions were diluted into 2X pNP buffer; at higher volumes or concentrations, the dichloromethane interacted with the Triton X-100 in pNP buffer to form a white precipitate that disappeared upon further dilutions. 1 ml of pNPX substrates (1 mM) (Sigma) in 1X PNP buffer was incubated with 1 μg of D29 LysB, D29 LysB S82A, lipase (*Pseudomonas fluorescens*, Sigma), or 5 μl of a mock purified sample (derived from *E. coli* BL21(DE3) cells containing pET21a) in pNP assay buffer at room temperature for 30 min. Release of *p*-nitrophenol was determined by measuring absorbance at 420 nm (A420) using a Beckman Coulter DU 530 Spectrophotometer.

A standard curve was generated by measuring the A_{420} of different concentrations of *p*-nitrophenol (1, 5, 10, 25, 50, 75, 100, 250 uM) and determining that 140.8 x Abs_{420} = 1 uM of pNP. The reaction rate (product formed over time) was calculated as uM min^{-1}. Enzyme activity (U) was calculated by multiplying the reaction rate by the volume of the reaction to obtain umoles min^{-1}. Lastly, the specific activity was calculated as the moles of product formed per mg of enzyme, or umoles min^{-1} mg^{-1}.

5.6.3   mAGP Hydrolysis

5.6.3.1 Purification of mAGP

Mycolylarabinogalactan–peptidoglycan was isolated from *M. smegmatis* as described previously (Besra, 1998). Briefly, *M. smegmatis* cells were grown, collected and washed three times in

phosphate buffered saline (PBS; pH 7.4), and resuspended in PBS + 2% Triton X-100 (PBSX).

Cells were disrupted by extensive sonication, centrifuged to collect the insoluble cell wall

fraction, resuspended in PBSX and agitated overnight at 4°C. After centrifugation the pellet was

resuspended in PBS + 2% SDS and incubated at 100°C for 60 min; this was done three times.

After three rounds of extraction, the pellet was washed once each with H2O, 80% acetone in

H2O, and acetone. Following evaporation of the acetone, the mAGP-enriched cell wall material

(mAGP) was resuspended in PBS + 0.1% Triton X-100 (final concentration 10 mg ml-1) and

frozen aliquots stored at -80°C.

5.6.3.2 Chemical and Enzymatic Hydrolysis of mAGP

For a control sample, mAGP material was chemically hydrolyzed by addition of an equal volume

of 15% tetrabutyl ammonium hydroxide (TBAH) to 1 mg mAGP resuspended in PBS + 0.1%

Triton X-100 and incubated at 100°C overnight. Enzymatic assays were performed by

incubation of 1 mg mAGP in 100 μl with varying concentrations of protein at 37°C. TBAH- and

enzyme-treated samples were prepared for analysis by thin-layer chromatography (TLC) by

addition of an equal volume of dichloromethane and incubation for 15 min at room temperature.

The lipid-rich lower dichloromethane layer was removed, extracted once with 0.25 M HCl and

once with water, and lipids collected by evaporation.

5.6.3.3 Thin Layer Chromatography

Dried reaction products were resuspended in dichloromethane, spotted onto silica-aluminium

TLC plates, and separated once by chromatography in chloroform/methanol (97:3); lipids were

identified by spraying with 5% molybdophosphoric acid (in ethanol) and charring for 15 min at

110°C.

Methyl-esterification of lipids in order to identify specific types of mycolic acids was performed by resuspension of enzyme reaction products in 15% TBAH, addition of an equal volume of dicholoromethane and 1/10 volume iodomethane. Reactions were incubated with shaking at room temperature for 15 min and the lower dichloromethane layer recovered and extracted with HCl and water as described above. Lipids were separated by six developments with 95:5 petroleum ether/ethyl ether and recognized with 5% molybdophosphoric acid and charring as above.

5.7   *IN VIVO* ASSAYS

5.7.1   OD Assay

Mycobacterium smegmatis cells grown in 7H9 supplemented with ADC, carbenicillin, cyclohexamide and calcium were grown to log-phase, as indicated by an OD_{600} 0.3–1.0. Single-step infections were set up with a starting culture OD_{600} of 0. 3 and m.o.i. of 10, and no virucide was used. 1 ml samples were removed at different times and the OD_{600} was measured using a Beckman Coulter DU 530 Spectrophotometer.

5.7.2   ATP Release Assay

Mycobacterium smegmatis cells were grown to log-phase as described above. Single-step infections were set up with a starting culture OD_{600} of 0.03 and m.o.i. of 10, and no virucide was used. 100 μl samples were periodically taken from the infected culture. ATP release was

measured by addition of 100 ml of ENLITEN rLuciferase/Luciferin reagent (Promega), and luminescence recorded for a 10 s interval in a Monolight 2010 luminometer. The results were reported as the fold-difference in ATP-release compared to an uninfected control culture.

5.7.3 Total Phage and Distribution Assay

To determine the number of phage released into the supernatant or retained in unlysed cells, *M. smegmatis* cells were grown as above and diluted to an OD_{600} of 0.25. These cells were infected at a m.o.i. of 0.1 with adsorption for 30 min followed by inactivation of any viable phage in the media using virucide and then incubation with shaking at 37°C. 1 ml of samples was removed at different times and separated by centrifugation at 5000 rpm for 5 min into supernatant and pellet fractions. The pellet was resuspended in 1 ml phage buffer and sonicated. Both pellet and supernatant were serially diluted in phage buffer and 5 μl samples were spotted onto top agar lawns containing *M. smegmatis* in 0.35% MBTA with 1 mM $CaCl_2$ on 7H10 plates and incubated at 37°C.

5.7.4 Growth Inhibition Assay

The assay for growth inhibition in detergent-free media is adapted from the biofilm assays (Ojha et al., 2005) that use minimal media. Media is dispensed into small petri dishes or multiple-well plates and inoculated 1:1000 with *M. smegmatis*. An amount of purified lysin or control is added to the media and the plate is incubated without motion at 30°C and monitored for 5-8 days. If endogenous expression is being tested, the media was supplemented with succinate instead of

glucose and acetamide was added to a final concentration of 0.2% to induce expression. Images were taken with a Canon Powershot A530.

5.8   BIOINFORMATIC ANALYSIS

5.8.1   Programs

5.8.1.1 Sequence Comparisons

BLASTp and PSI-BLAST: alignment of two or more protein sequences; Position-Specific-Iterated (PSI) Blast allows the construction of a position-specific scoring matrix that can be used in successive alignments. (Altschul et al., 1997; Schaffer et al., 2001)

ClustalW: multiple sequence alignments of proteins, default settings used (Thompson et al., 1994)

NJPlot: phylogenetic tree drawing program (Perriere and Gouy, 1996)

PHYML 3.0: generation of maximum likelihood phylogenetic trees based on ClustalW multiple alignments with bootstrap and SH tests (Felsenstein, 1993; Guindon and Gascuel, 2003)

5.8.1.2 Protein Analysis and Pattern Recognition

MEME (v. 4.4.0): Multiple Em for Motif Elicitation, pattern discovery based on alignment of multiple protein sequences (Bailey et al., 2006)

ProtParam (ExPASy): calculation of general protein properties including molecular weight based on primary amino acid sequences (Gasteiger et al., 2005)

SignalP 3.0 Server: prediction of signal peptides based on primary amino acid sequences (Bendtsen et al., 2004)

TMHMM Sever (v. 2.0): prediction of transmembrane helices based on primary amino acid sequences(Krogh et al., 2001)

5.8.1.3 Others

CLC Bio Workbench (v. 4.1.2): organization of sequences, cloning and primer design, ClustalW and other alignment programs, maximum likelihood phylogeny (neighbor joining), pairwise comparisons

MyDomains (Prosite): image creator for protein domain maps

5.8.2   Databases

5.8.2.1 Annotated and Classified Protein Sequences

InterPro: database of predictive protein signatures for classification and automatic annotation of proteins (Hunter et al., 2009)

GenBank: NIH gene and protein sequence database (Benson et al., 2009)

Pfam: database of annotated protein families (Finn et al., 2010)

5.8.2.2 Enzymes

BRENDA: general enzyme database (Chang et al., 2009)

CAZy: Carbohydrate-Active enZYmes database (Henrissat and Davies, 1997)

MEROPS: peptidase database (Rawlings et al., 2010)

5.9 REAGENTS AND BUFFERS

5.9.1 Protein purification buffers

TEWB (Tris Elution Wash Buffer):

50 mM Tris Hcl, 300 mM NaCl, pH 8.0

(optionally added 10-120 mM imidazole from 1 M imidazole stock)

Filter, store at room temperature or 4°C

Dialysis/storage buffer:

50 mM Tris HCl, pH 8.0, 50 mM NaCl, 50% glycerol

(optionally added: 0.1 M EDTA, 0.1 mM DTT)

Store at 4°C or -20°C

5.9.2 SDS-PAGE

Separating gel (10%, 10 mL for two small gels): 5 mL dH$_2$O, 2.5 mL acrylamide:bis-acrylamide (29:1), 2.5 mL 4X separating buffer, 50 µl 10% ammonium persulfate (APS), 10 µl tetramethylethylenediamine (TEMED)

Make fresh

Stacking gel (4.5%, 5 mL for two small gels): 3 mL dH$_2$O, 1.25 mL 4X stacking

buffer, 0.625 mL acrylamide:bis-acrylamide (29:1), 25 µl 10% APS, 5 µl TEMED

Make fresh

4X Separating Buffer:

1.5M Tris HCl pH 8.8, 0.4% (w/v) SDS

Filter, store at room temperature

4X Stacking Buffer:

0.5 M Tris HCl pH 6.8, 0.4% (w/v) SDS

Filter, store at room temperature

4X Sample Buffer (40 mL):

62.5 mM Tris HCl pH 6.8, 40% glycerol, 2% (w/v) SDS, 5% 2-mercaptoethanol, 0.2%

(w/v) bromophenol blue

Filter, store at room temperature

Protein Gel Running Buffer:

25 mM Tris base, 192 mM glycine, 0.035% (w/v) SDS, pH 8.3

Store at room temperature

Commassie Blue Stain:

1.25 g Coomassie blue, 225 ml methanol, 45 ml acetic acid, dH$_2$O to 500 ml

Store at room temperature

5.9.3   Zymograms

Zymogram Separating gel (10%, 10 mL for two small gels):   5 mL dH$_2$O, 2.5 mL acrylamide:bis-acrylamide (29:1), 2.5 mL 4X Zymogram separating buffer, 0.01 g peptidoglycan or 0.02 g lyophilized cells, 50 µl 10% ammonium persulfate (APS), 10 µl tetramethylethylenediamine (TEMED)

Make fresh.

Zymogram Stacking gel (4.5%, 5 mL for two small gels):   3 mL dH$_2$O, 1.25 mL 4X Zymogram stacking buffer, 0.625 mL acrylamide:bis-acrylamide (29:1), 25 µl 10% APS, 5 µl TEMED

Make fresh

4X Zymogram Separating Buffer:

1.5M Tris HCl pH 8.8, 0.04% (w/v) SDS

Filter, store at room temperature

4X Zymogram Stacking Buffer:

0.5 M Tris HCl pH 6.8, 0.04% (w/v) SDS

Filter, store at room temperature

2X Sample Buffer (20 mL):

125 mM Tris, pH 6.8, 4% SDS, 40% glycerol, 10% β-ME, 0.01% bromophenol blue

Filter, store at room temperature

Renaturing Solution:

25 mM Tris HCl, 1% Triton X-100, pH as needed

(optionally added: 0.1 mM DTT, various cations)

Make fresh

Methylene Blue Staining Stain:

0.25% methylene blue, 0.01% KOH

Make fresh

5.9.4 *p*-Nitrophenol Ester Assay

pNPX Stock Solutions

Substrates were dissolved to a final concentration of 250 mM or 100 mM in dichloromethane

Store at -20°C, dilute fresh at room temperature for assay

pNP Assay Buffer:

20 mM Tris-HCl pH 8.0, 100 mM NaCl, 0.1% Triton X-100

Filter, store at room temperature

5.9.5 mAGP hydrolysis and Thin Layer Chromatography

mAGP hydrolysis buffer:

PBS pH 7.4, 0.1% Triton X-100

Store at room temperature

Chloroform : Methanol:

97 ml chloroform, 3 ml methanol

Make fresh

Petroleum Ether : Ethyl Ether:

95 ml petroleum ether, 5 ml ethyl ether

Make fresh

Molybdophosphoric acid spray:

1 g molybdophosphoric acid, 20 ml 100% ethanol

Store at room temperature

5.9.6 Others

Phage buffer:

10 mM Tris-HCl pH 7.5,10 mM $MgSO_4$, 68.5 mM NaCl, 1 mM $CaCl_2$

Filter, store at room temperature

Virucide:

10 mM ferrous ammonium sulfate in dH_2O

Make fresh

Biofilm Base Minimal Media

In 900 ml dH_2O, add 13.6 g KH_2PO_4, 2.0 g $(NH_4)_2SO_4$, adjust the pH to 7.2 with KPH

pellets, add 0.5 mg $FeSO_4 \bullet 7H_2O$, 0.5 g casamino acid

Autoclave 20 min at 121 ^{o}C, store at room temperature

Reconstituted Biofilm Minimal Media

Prior to assay, reconstitute 100 ml of media: 94 ml of base media, 5 ml 40% glucose, 1

ml $CaCl_2$, 100 ul $MgSO_4$ (if inducing, replace 5 ml 40% glucose with 1 ml 20% succinate

and 4 ml dH_2O and add 1 ml/100 ml 20% acetamide to induce; add 40 ul/100 ml of Kan

if necessary)

Use fresh, usually at 30°C

APPENDIX

DEMONSTRATED LYTIC ACTIVITY & GROWTH INHIBITION OF *M. SMEGMATIS* BY MYCOBACTERIOPHAGE LYSINS

The body of this dissertation has focused on the process of bacteriophage lysis, the mechanisms of cell wall degradation and host-specific substrate binding, and the complex evolution between phage and host. However, much current lysin research focuses on the therapeutic potential of lysins in preventing and treating disease (Borysowski et al., 2006; Fischetti, 2008; Loessner, 2005). Extensive research has demonstrated the ability of lysins to kill the host of their phage with incredible speed and specificity, both *in vitro* and more recently *in vivo* using animal models. In particular, lysins have proven efficacious against *Streptococcus pneumoniae* and other pathogenic streptococci, *Listeria monocytogenes*, *Bacillus anthracis*, and *Staphylococcus aureus* (Cheng et al., 2005; Djurkovic et al., 2005; Donovan et al., 2006b; Entenza et al., 2005; Rashel et al., 2007; Schuch et al., 2002; Turner et al., 2007).

Development of new therapeutic agents is a long-term goal of research on mycobacteriophage lysins as well. Accordingly, several individual LysA proteins were cloned and purified for empirical study. The ability of these proteins to hydrolyze cell walls was tested using zymography (Leclerc and Asselin, 1989), and the effects of both LysA and LysB on the

growth of mycobacteria was studied with exogenous application of purified protein and endogenous expression within *M. smegmatis*. The preliminary results of these experiments are described below.

A.1 *IN VITRO* HYDROLYTIC ACTIVITY OF LYSA PROTEINS

Five LysA proteins were cloned into the pET21a vector for IPTG-induced expression with a C-terminal His tag. These were Bxz1 gp236, Che8 gp32, Corndog gp69, D29 gp10, and L5 gp10 (Figure 37). Three of these LysAs were used because they had already been cloned for expression prior to my arrival. Corndog gp69 was chosen based on its domain organization: an N3 domain (Bxz1 gp236 and Che8 gp32 have N2; D29 gp10 has N4); a central Ami2 domain similar to Bxz1 gp236 and having the same predicted activity as the Ami1 of Che8 gp32; and a C2 domain similar to Che8 gp32 but distinct from the C1 domains in Bxz1 gp236 and D29 gp10. L5 gp10 was later added because it is similar to D29 gp10, yet lacks a GH19 domain (Figure 37). It was thought that by comparing activities and domains through recombination the functions of the distinct domains could be interpreted. However, the *in vitro* analysis of these proteins has not yet proceeded beyond an examination of their general ability to hydrolyze cell wall material using zymography.

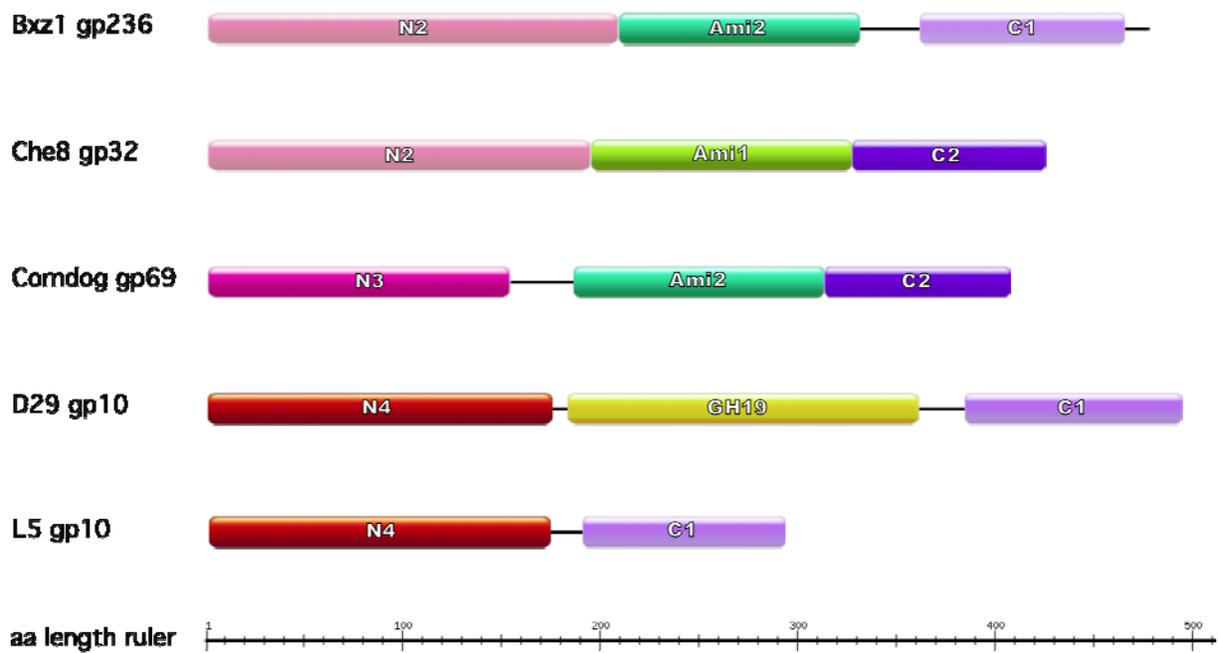**Figure 37. Domain organization of LysAs cloned and purified for empirical study**

Figure 37: Five purified LysA proteins were used for studies of lytic activity. Not all identified domains are represented; this set contains the N-terminal domains N2, N3, and N4 (pink and red shades); the lytic domains Ami1, Ami2, and GH19 (greens and yellow, respectively); and the C-terminal domains C1 and C2 (purple shades).

257

A.1.1 Zymography of LysA proteins with *Micrococcus luteus* peptidoglycan

Initial experiments used lyophilized *Micrococcus luteus* cells, which are a standard substrate for zymography for peptidoglycan (PG) hydrolytic activity (Leclerc and Asselin, 1989; Piuri and Hatfull, 2006). The LysAs Bxz1 gp236, Che8 gp32, and Corndog gp69 showed high activity on the *M. luteus* zymogram (Figure 38). However, there was little activity seen from D29 gp10 (not shown), and no activity by D29 gp12, the LysB protein (not shown). Knowing that LysB proteins target the mycolyl-arabinogalactan ester bond – a substrate unique to mycobacteria and related genera – it is logical that no activity would be seen on the *M. luteus* zymogram. The addition of Zn_2SO_4 to the renaturation buffer did not appear to have a significant effect, which seemed odd considering Bxz1 gp236, Che8 gp32, and Corndog gp69 are all predicted Zn^{2+} amidases. The controls used were ovalbumin (non-hydrolytic) and lysozyme (hydrolytic), although the quantity of lysozyme used made the clearing difficult to discern (Figure 38, red arrows).

Several interesting phenomena were observed in these zymograms, in that numerous other bands were visible in addition to the primary band corresponding to the protein (Figure 38). It is possible that these were contamination, but the bands between the proteins were not consistent, and they were not observed on zymograms with mock purifications of induced cells containing the empty pET21a vector (not shown).

First, several bands (Figure 38, green arrows) traveled more slowly than the larger band or clearing that corresponds in size to the LysA. Che8 gp32 especially had numerous bands of high molecular weight as seen at the top of the SDS PAGE gel, and there was activity in the zymogram corresponding to these bands. The other bands seen for Corndog gp69 could also

correspond to dimers or other complexes, and at least one showed activity on the zymogram. This is not entirely unprecedented among lysins; PlyC forms a multimeric holoenzyme based on two lysis proteins that assemble in an 8:1 ratio (Nelson et al., 2006). This has not been recorded for any other lysin, nor are the genes encoding other lysins similar in organization to that of PlyC; however, while the *lysA* genes do not resemble PlyC, this does not preclude the possibility of multimers.

Second, numerous other bands were seen below the expected protein sizes, several of which seemed to correspond to smaller clearings on the zymogram (Figure 38, brackets). These were hypothesized to be degradation fragments that retained activity; the modular nature of LysAs presumes short linker sequences between the domains, which are more susceptible to cleavage. When comparing the amount of protein as seen on the SDS PAGE gel to the size of the clearing on the zymogram, several of these fragments had incredible hydrolytic activity given the amount. The best example of this was seen with Corndog gp69 (Figure 39). Two very small fragments around that were barely visible on SDS PAGE around 30 kDa showed levels of activity similar to the full-sized Corndog gp69 protein (Figure 39 A). The larger of these bands was purified and sequenced by Edman degradation to determine the amino acid sequence at the N-terminus (Konigsberg, 1967). The N-terminal sequence corresponded to amino acids 153-160 of Corndog gp69 (Figure 39 B), which results in a 28 kDa, 253 aa polypeptide that has lost nearly its entire N-terminal, or N3, domain. This indicates that, in these *in vitro* circumstances using *M. luteus* PG substrate, the N3 domain is dispensable for hydrolytic activity and may in fact inhibit activity.

Protein	kDa
Lysozyme	10
Ovalbumin	47
Corndog gp69	45
Bxz1 gp236	52
Che8 gp32	48
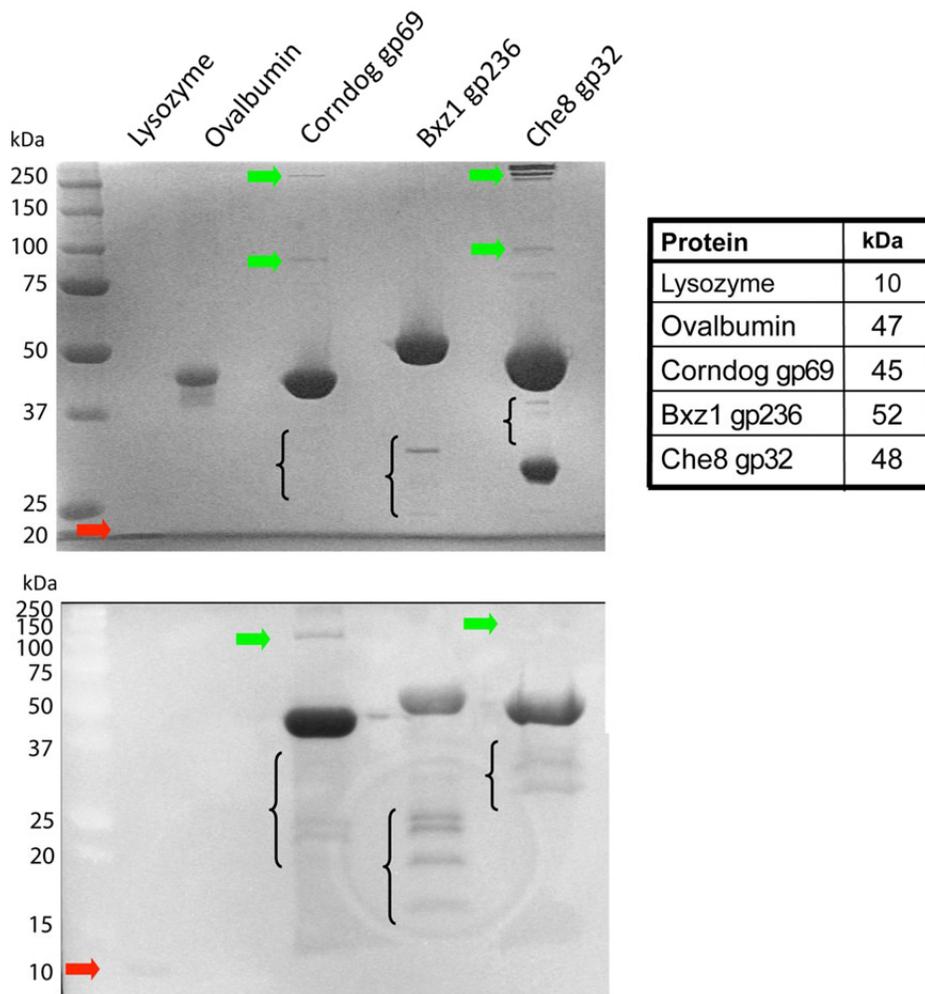
Figure 38. SDS PAGE and corresponding zymogram of three LysA proteins

Figure 38: The upper gel is a standard SDS PAGE and shows all protein present in each sample, while the lower gel is a zymogram that has been inverted so that the clearings appear as dark bands. Clearings in the zymogram indicate hydrolytic activity, and the size of the bands is correlated to those in the SDS PAGE gel. It should be noted that, due to the addition of an enzymatic substrate, protein separation is altered in the zymogram compared to the SDS PAGE gel, as shown by the ladder sizes. The red arrow indicates the lysozyme positive control. Green arrows highlight protein larger than the expected protein size showing hydrolytic activity. Brackets highlight smaller protein fragments showing activity in the zymogram. The table lists the sizes of the various proteins.
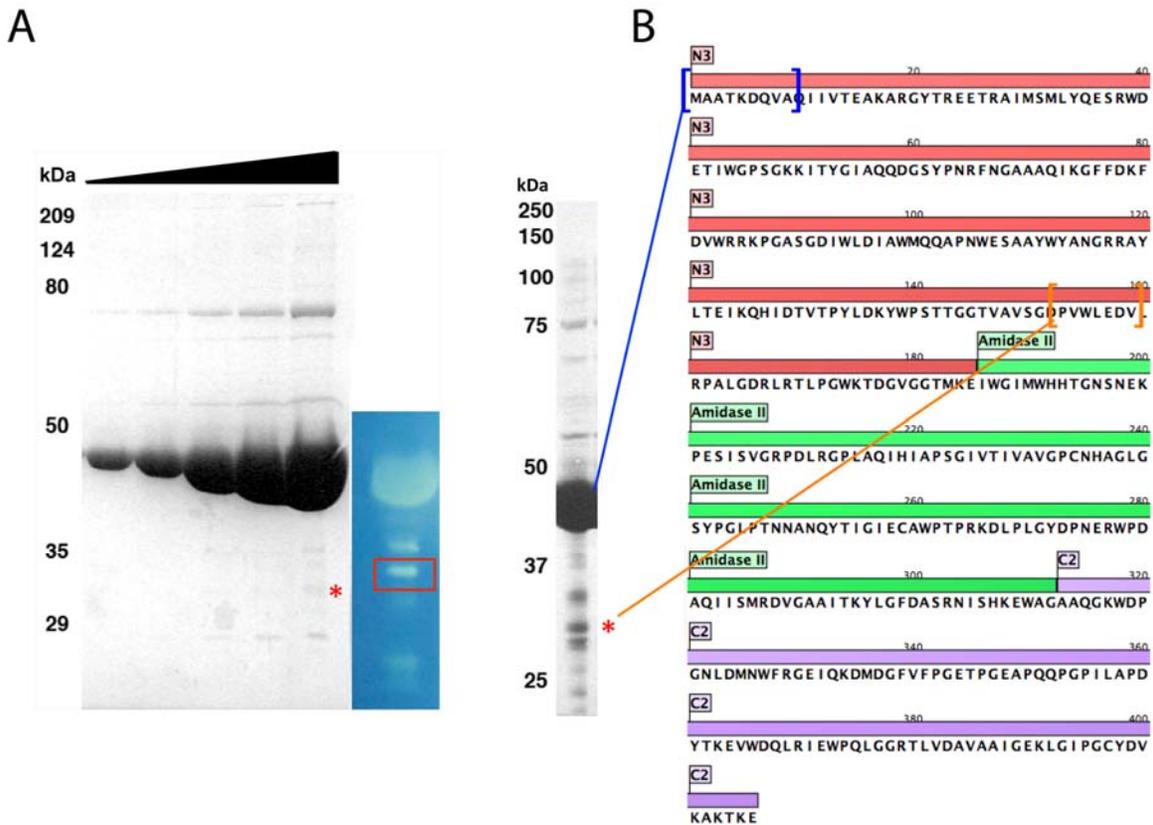
Figure 39. Identification of Corndog gp69 fragment with increased lytic activity

Figure 39: **A.** Left, SDS PAGE with increasing amounts of Corndog gp69 sample illustrating the minute amount of protein (red asterisk) corresponding to one of the bands of high activity seen in the zymogram on the right (red box). **B.** Highly concentrated sample of Corndog gp69 from which a sample was taken (red asterisk) and submitted to Edman degradation. The control sequencing of the large band matched the N-terminal sequence of Corndog gp69 (blue brackets), while the smaller fragment corresponded to a sequence 153 aa C-terminal to the start of Corndog gp69 (orange brackets).

262

A.1.2 Zymography of LysA proteins with *B. subtilis* and *M. smegmatis*

Considering the high specificity of lysins, there was some concern that the differences between mycobacterial PG and that of *M. luteus* may inhibit the activity of the LysAs. Specifically, the type A2 *M. luteus* PG has an interpeptide bridge in its PG cross-links, and the type Alγ of mycobacteria has direct m-DAP-D-Ala cross-links (Figure 5). Later experiments tried *Bacillus* cells or purified PG, which also have A1γ PG but are less amenable to zymography. *M. smegmatis* PG was more difficult to use as a substrate due to the small amount available, its uneven distribution throughout the gel, and problems retaining the methylene blue dye. However, preliminary results were obtained with both of these substrates alongside *M. luteus* zymograms acting as a control.

D29 gp10 and L5 gp10 were also used in these studies. Both proved difficult to purify; additional bands were occasionally seen in SDS PAGE gels, and the proteins had a consistent tendency to precipitate during concentration and dialysis or upon addition to other solutions, including buffers used for assays. Despite these difficulties, adequate amounts of protein were purified to use in zymography and other studies. The following results are the accumulation of data from three separate zymography studies, although only one example is shown.

All five LysA proteins were separated on both an SDS PAGE gel and zymograms containing *M. luteus*, *B. subtilis,* or *M. smegmatis* PG. After renaturation overnight the gels were stained with methylene blue and destained to observe any clearings formed by hydrolysis by the proteins. *B. subtilis* did not stain as well as *M. luteus*, and the *M. smegmatis* zymograms did not retain the stain well nor have strong contrast between the stain and any clearings (Figure 40).

Lysozyme was added as a positive control, and D29 gp12 (LysB) was used as a possible negative control.

Interestingly, as shown on the SDS PAGE gel, Che8 gp32 again did not separate but remained at the top of the separating gel (Figure 40 A). In addition, L5 gp10 appeared as two bands (Figure 40 A); these bands are approximately the correct size, but in alternate purifications a single band has been observed.

Clearings were distinctly visible on the *M. luteus* zymogram by Bxz1 gp236, Che8 gp32, and Corndog gp69 (Figure 40 B). In this instance, a Bxz1 gp236 degradation product of around 30 kDa – which is a plausible size if Bxz1 gp236 were to lose a large portion of its N2 domain – was seen to have strong activity. Also, the aggregate of Che8 gp32 at the top of the gel again showed activity (Figure 40 B). Faint activity was also shown by D29 gp10 (Figure 40 B, red arrow), but there was no activity from either L5 gp10 or D29 gp12. A small band around 48 kDa appeared in all lanes except Corndog gp69 and lysozyme (Figure 40 B); this may be a contaminant.

Activity by the first three LysAs was also apparent on the *B. subtilis* zymogram (Figure 40 C). Interestingly, there did not appear to be an active degradation product for Bxz1 gp236; this could indicate a lack of degradation during the separation through *B. subtilis* PG or, more intriguingly, the possibility that the N2 domain is required for activity on A1γ-type PG. Again, weak activity was shown by D29 gp10 (Figure 40, red arrow), although activity was stronger on other zymograms (not shown). Oddly, there were bands in the lanes of both L5 gp10 and D29 gp12. The clearing for L5 gp10 was at approximately the correct location for a protein of 32 kDa (Figure 40 C, yellow arrow); based on size alone, the band seen for D29 gp12 (Figure 40 C, green arrowhead) was not likely to be the 29 kDa LysB protein. Weak activity by L5 gp10 was

also seen in a second *B. subtilis* zymogram (not shown) but was indeterminant in a third one (data not shown).

The *M. smegmatis* zymogram was difficult to discern due to the granular nature of the substrate and a gradient of staining that rendered the left half of the gel largely unreadable (Figure 40 D). Difficult to see in this image were a smear of potential clearing in the lane of Che8 gp32 and a possible clearing in the lane of Corndog gp69, although better activity was seen by Bxz1 gp236 and Corndog gp29 in a later zymogram (not shown). There did appear to be clearings present in the lanes of L5 gp10 (Figure 40 D, yellow arrow) and D29 gp10 (Figure 40 D, red arrow). The band for L5 gp10 seemed to be a little smaller in comparison to the ladder and to the protein seen in the SDS PAGE gel (Figure 40 A). Still, protein separation varies in each zymogram, so it is plausible that this band was the L5 gp10 protein seen in the SDS PAGE. Since the dyeing properties of *B. subtilis* and *M. smegmatis* zymograms are uncharacterized, it is also possible that clearings on these gels were due to the protein excluding the dye; however, this was not a concern for *M. luteus* zymograms.
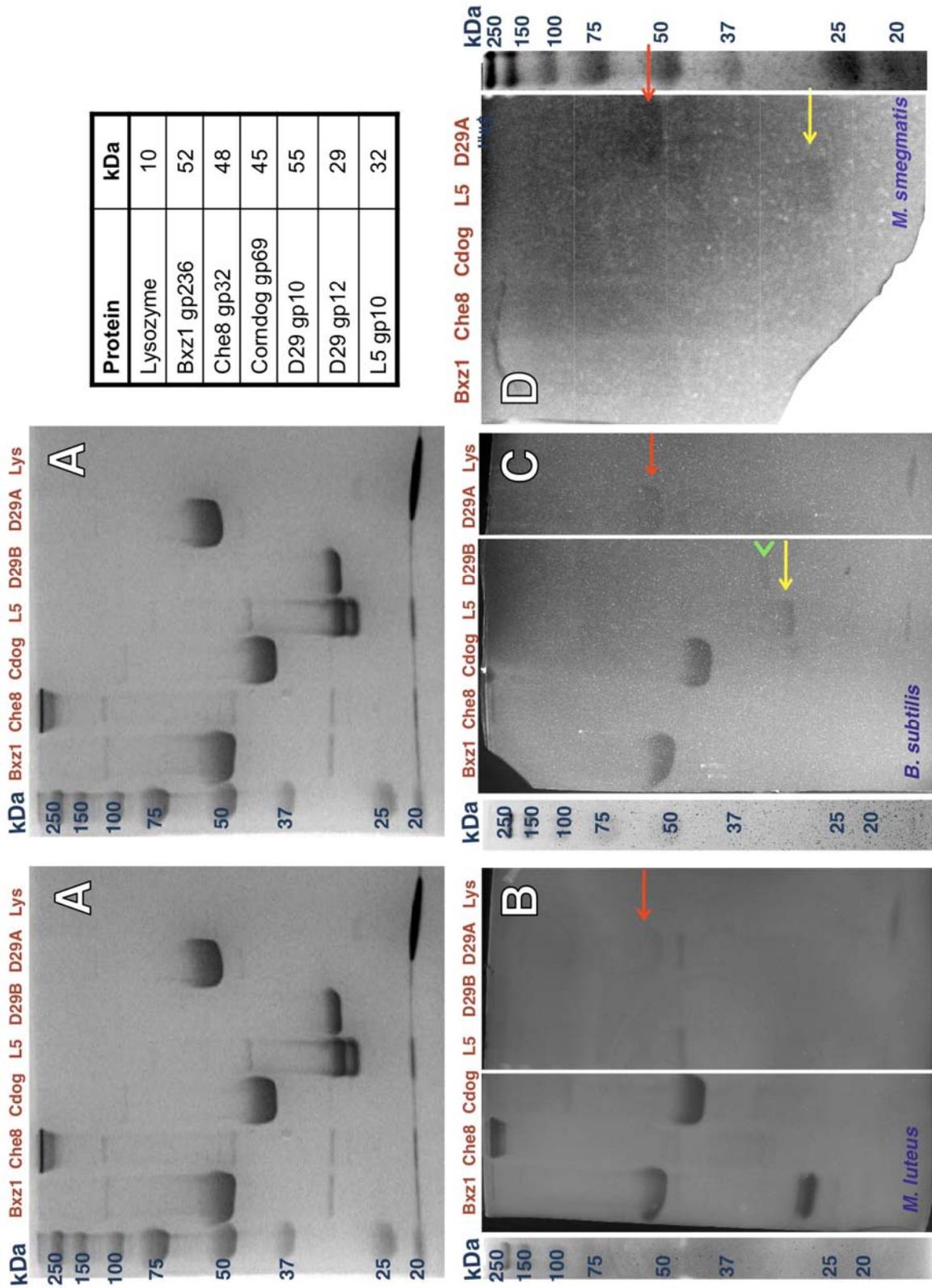
Protein	kDa
Lysozyme	10
Bxz1 gp236	52
Che8 gp32	48
Corndog gp69	45
D29 gp10	55
D29 gp12	29
L5 gp10	32

Figure 40. Zymography with *M. luteus*, *B. subtilis*, and *M. smegmatis* substrates

Figure 40: Several zymograms with different embedded substrates were run simultaneously. The zymogram images have been inverted so that the clearings appear as dark bands. **A.** SDS PAGE gel of the protein loaded, shown twice so that it can be aligned above both B and C. **B**. *M. luteus* zymogram. **C.** *B. subtilis* zymogram. **D.** *M. smegmatis* zymogram. The red arrows point to clearings by D29 gp10, the yellow arrows point to possible clearings by L5 gp10, and the green arrowhead points to an unknown band in the lane of D29 gp12. The table indicates the protein sizes in kDa, and ladders are found to the side of every zymogram.

A.1.3   Conclusions

Several purified LysAs were shown to have PG hydrolytic activity on zymograms containing *M. luteus* cells and, to varying extents, on zymograms with *B. subtilis* PG or *M. smegmatis* PG. The three LysAs with predicted amidase activity, Bxz1 gp236, Che8 gp32, and Corndog gp69, all showed strong activity on *M. luteus* and *B. subtilis* substrates, and possible activity on the *M. smegmatis* zymogram. D29 gp10, a predicted GH19 lysozyme, showed weak activity on the *M. luteus* and *B. subtilis* substrates, but appeared more active on the *M. smegmatis* zymogram. L5 gp10 did not show any activity on *M. luteus*, but may have been active on *B. subtilis* and *M. smegmatis* zymograms.

The results of these zymogram experiments may provide some clues to the functions of the N-terminal domains. When the N3 domain was lost from Corndog gp69, there was an increase in Corndog gp69's hydrolytic ability on *M. luteus* PG (Figure 39). A similar circumstance may also have occurred with Bxz1 gp236, where a smaller fragment showed greatly increased activity on a later *M. luteus* zymogram (Figure 38 B); however, in this instance we cannot be confident that it was the N2 domain lost, since we did not perform Edman degradation to identify the N-terminal sequence. Nonetheless, this smaller band of activity was not seen on the *B. subtilis* zymogram that was run simultaneously (Figure 40 C). This raises the possibility that the portion of Bxz1 gp236 that was lost is important in the hydrolysis of *B. subtilis* substrate, which varies in PG structure from *M. luteus*. Therefore, the function of the N2 and N3 domains may not be conducive to *in vitro* hydrolysis of *M. luteus* PG, but at the same time may not inhibit or could even enhance activity on *B. subtilis* PG, which is highly similar in composition to that of mycobacteria. The majority of Bxz1 gp236 and Corndog gp69 was still in

the full-length form, which argues against proteolytic processing to create a truncated protein, unless this is executed by a protease not present in *E. coli*. These zymograms do not provide any overt evidence that the N2 or N3 domains alone have PG hydrolytic activity; however, the smaller hydrolytic bands have not been sequenced and could be fragments of the N-terminal domain. The increase in activity seen upon the loss of the N3 domain of Corndog gp69 and the presumed loss of the N2 domain of Bxz1 gp236 is consistent with a role in protein export as was observed for the N-terminus of the homologous Ms6 LysA (Catalao). It is reasonable that the loss of a N-terminal domain that has no function in PG hydrolysis could result in an increase in PG hydrolytic ability

Considering the similarity between D29 gp10 and L5 gp10, these results may support a hydrolytic role for the N4 domain that is specific to A1γ-type PG. The GH19 lysozyme domain in D29 gp10 alone would be expected to have activity on most PG types, and indeed there was some activity by D29 gp10 seen on all zymograms (Figure 40 B-D). L5 gp10, on the other hand, lacks this domain (Figure 37) and showed no activity on *M. luteus* (FIG B). However, L5 gp10 may have been active on the A1γ-type PG that is shared between *B. subtilis* and *M. smegmatis* (Figure 40 C, D). This could mean that the N4 domain is an uncharacterized PG hydrolase, which would then identify D29 gp10 and other N4-containing LysAs (Table 1) as having two catalytic domains. Meanwhile, the N4 activity of L5 gp10 alone may be sufficient for the lysis of mycobacteria by phage L5, although this introduces the question of why the two-domain architecture is so common if the single N4 domain can fulfill the role of endolysin. Mutagenesis studies to elucidate the mechanism of lysis by L5 gp10 have been started.

A.2 INHIBITION OF MYCOBACTERIAL GROWTH BY EXOGENOUS APPLICATION OF LYSINS

The ability of lysins of Gram-positive-infecting phage to rapidly lyse their hosts when externally applied is well documented (Borysowski et al., 2006; Hermoso et al., 2007). This is attributed to the exposure of the PG to the external environment (Figure 3 B). In contrast, it has been significantly more difficult to effect lysis of Gram-negative cells in the same manner due to the presence of the outer membrane, a hydrophobic permeability barrier that prevents access to the underlying PG (Figure 3 A). The mycobacterial cell wall shares characteristics of both Gram-negative and Gram-positive cell walls: it has a thick layer of PG, but this is surrounded by a hydrophobic bilayer of mycolic acids and other lipid derivatives (Figure 4 A) similar to the outer membrane. This waxy coating presents a formidable barrier to any enzyme trying to access the PG, as well as any molecule trying to gain entry to mycobacteria in general, such as antibiotics (Brennan and Nikaido, 1995; Karakousis et al., 2004; McNeil and Brennan, 1991). However, unique to mycobacteriophage is the second lysin, LysB, which targets the mycolic acid ester bond. Not only is this bond found in the mycolyl-arabinogalactan, but also many of the glycolipids in the mycobacterial outer membrane contain sugars esterified to mycolic acids, most notably trehalose dimycolate (TDM, or cord factor). They are also important in the formation of biofilms, which are surface-bound communities of bacteria within a secreted extracellular matrix made of polysaccharides, lipids, protein, or DNA (Hall-Stoodley and Stoodley, 2005). It has recently been shown that *M. smegmatis* encodes a cutinase-like serine esterase that acts similarly to LysB to liberate the mycolate from TDM (Ojha et al., 2010). Mycobacterial biofilms are

composed primarily of free mycolic acids, and the hydrolysis of mycolic acids from TDM by this esterase is implicated in biofilm maturation in *M. smegmatis* and likely *M.tuberculosis* (Ojha et al., 2010). Mycobacteria within the biofilm are drug-tolerant persisters (Ojha et al., 2008), and such formations may contribute to the difficulty in treating *M. tuberculosis* infections, which require 6-9 months of antibiotic treatment (Connolly et al., 2007).

A.2.1   Addition of purified LysA and LysB to *M. smegmatis*

The complex mycobacterial cell wall augmented by the formation of lipid-based biofilm formation presents a formidable barrier to exogenous lysis by mycobacteriophage lysins. Still, while there is little hope of LysA proteins reaching the PG, there is some evidence for LysB activity on TDM (Ojha et al., 2010), and the external lipid-rich cell wall structures are more likely to contain substrates for LysB. As a note, the addition of lysins to shaking cultures of *M. smegmatis* had no visible effect on growth (data not shown), so these experiments were conducted in detergent-free standing cultures similar to the assays for biofilm formation. It was not investigated at this time whether the lysins had a bactericidal or bacteriostatic effect on *M. smegmatis*, and so all observations are in terms of "growth inhibition." Tested substances included lysins, other enzymes such as lysozyme and lipase, non-enzymatic proteins like bovine serum albumin (BSA), and enzyme storage buffer. Growth was monitored for 5-8 days, and pictures were taken with a hand-held digital camera. In addition to overall growth, formation of a mature biofilm (similar to negative control) was also examined.

Figure 41 shows 5 day-old cultures to which 1, 10, or 100 µg/ml each of D29 gp10 (D29A, Figure 41 A), D29 gp12 (D29B, Figure 41 B), or the D29 gp12 Ser82Ala mutant (S82A, Figure 41 C). In addition to a negative control, one well had 50 µl of the protein storage buffer

added, and a second had 100 ug/ml of BSA (Figure 41 D). D29A had very little effect on growth

at a concentration of 1 µg/ml, but appeared to inhibit biofilm maturation at 10 µg/ml and

inhibited the formation of a full film at 100 µg/ml (Figure 41 A). D29B had an even stronger

effect, inhibiting maturation at only 1 µg/ml, significantly inhibiting growth at 10 µg/ml, and

almost completely preventing growth at 100 µg/ml (Figure 41 B). The D29B active site mutant,

S82A, had no obvious effect at 1 µg/ml, but seemed to impact maturation of the biofilm at higher

concentrations (Figure 41 C). This could indicate that enzymatic activity is not absolutely

necessary; perhaps the binding of the substrate may be sufficient for some inhibition of

stationary growth. All negative controls appeared normal (Figure 41 D).

It is possible that the proteins were being degraded by extracellular proteases or

eliminated in some other way. To examine this, a controlled amount of sample was carefully

taken from the media beneath the growth at the interface every day for 7 days. These were

separated on an SDS PAGE gel (Figure 42). The results showed no degradation of D29B over

the entire time. However, the bands for D29A, S82A, and BSA all decreased, especially after

day 4. It is therefore possible that the difference in the levels of inhibition was due not to the

proteins' activity, but to whether they remained intact throughout the experiment. Considerable

changes in protein expression occur over the course of biofilm growth and maturation (Ojha et

al., 2005; Zambrano and Kolter, 2005); some protein may have been produced and secreted

starting around day 4 or 5 that can degrade the lysins and BSA. The sustained presence of D29B

could be due to a resistance to degradation, or to the failure of the culture to reach the point in

time where the degradative protein is expressed.
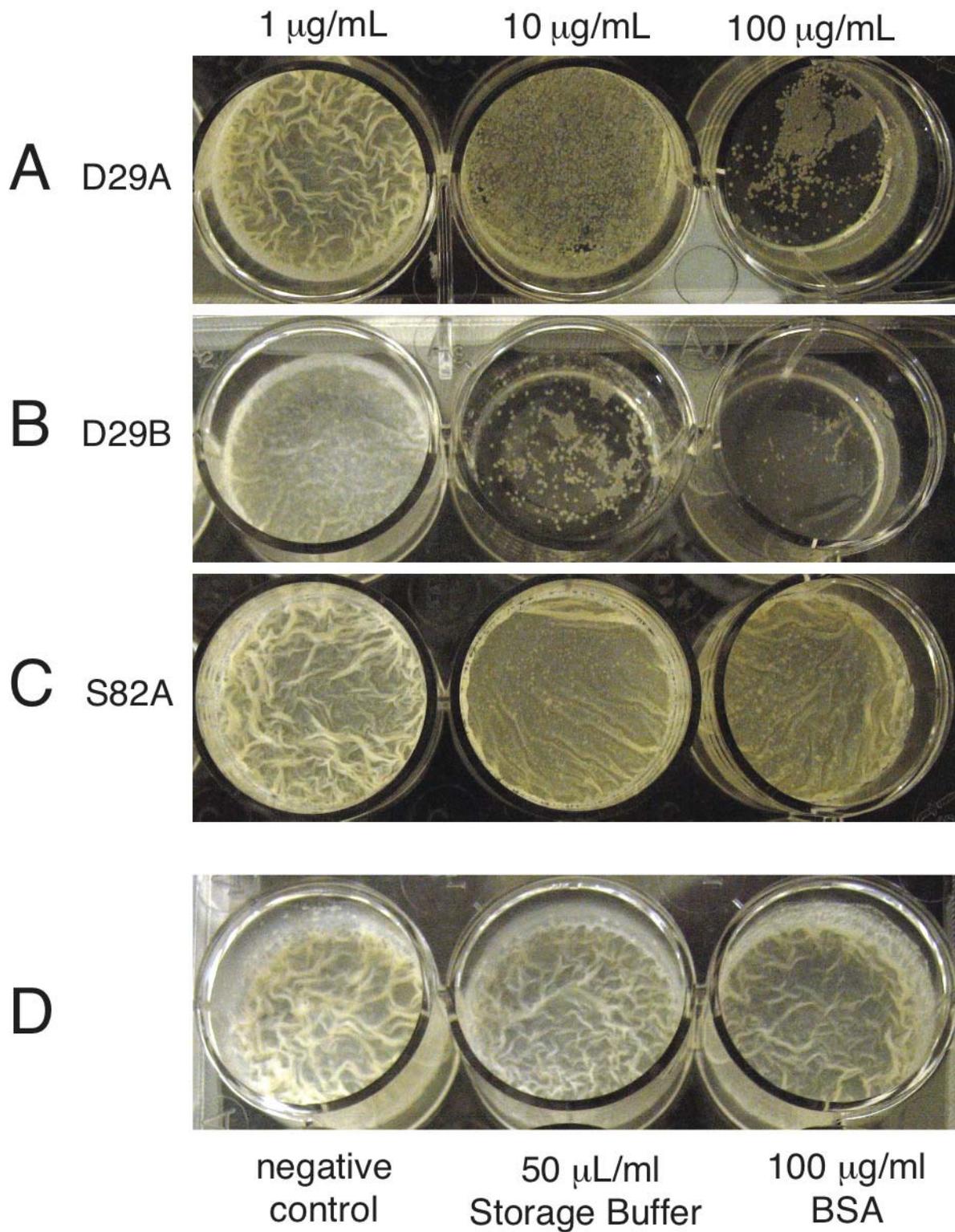
Figure 41. Growth inhibition by D29 LysA, LysB, and LysB S82A mutant

Figure 41: Biofilm pictures were taken after 5 days of growth at 30°C. Increasing amounts of each protein were added, including 1 µg/ml, 10 µg/ml, and 100 µg/ml. The negative control has nothing added, the storage buffer control was a solution of 50 mM Tris pH 8.0, 50 mM NaCl, 50% glycerol, and the BSA was a non-reactive protein.
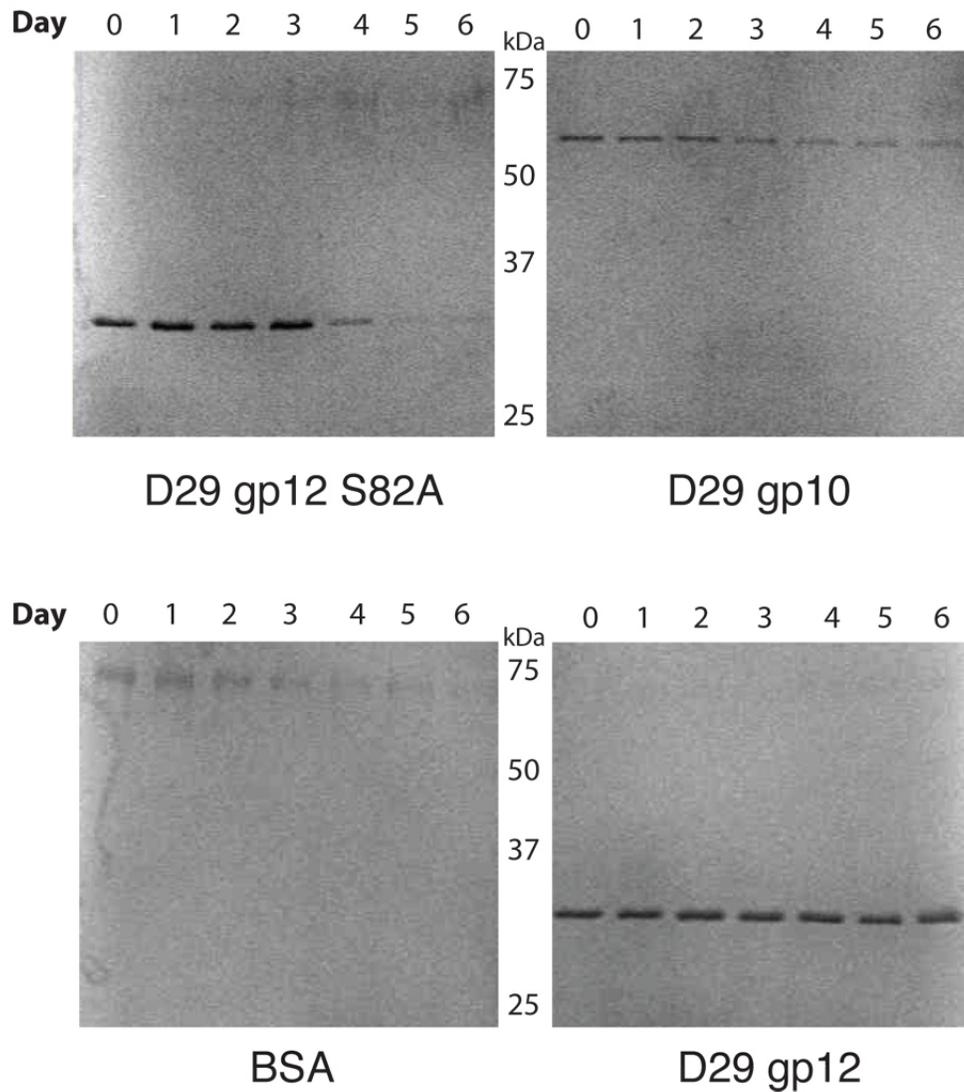
Figure 42. Lysin protein present in growth media over time

Figure 42: Samples from the growth media with 100 µg/ml of each protein added were taken daily for 6 days. The debris was pelleted and an equal amount was loaded and separated on an SDS PAGE gel and stained with Coomassie blue dye. A ladder in kDa is in the center of each pair of gels.

A.2.2 Inhibitory activity of lysin combinations

We predicted that D29A and gp12 might work synergistically to lyse bacteria since they are from the same mycobacteriophage. We tested varying combinations of D29A and D29B in the growth inhibition assay. Having observed the strength of the response in previous assays, the base amounts for D29A and D29B were 5 µg/ml and 1 µg/ml, respectively. These amounts showed an inhibitory effect on *M. smegmatis* but were not so great as to mask any increases in efficacy (*e.g.* 10 µg/ml of D29B was sufficient to prevent most but not all growth; Figure 41 B). The growth was assessed after 5 days.

The addition of increasing amounts of D29A to 1 µg/ml D29B showed varying increases in growth inhibition (Figure 42 A, B). A significant increase in inhibition and prevention of biofilm maturation was seen at 10 µg/ml of D29A, but the inhibition seemed to decrease from this peak when the concentration was increased to 25 µg/ml and 50 µg/ml (Figure 42 A). Additionally, 5 µg/ml of D29B was tested with 5 or 10 ug/ml of D29A (Figure 42 B). Compared to 5 µg/ml of D29A alone, addition of 5 µg/ml of D29B showed a great increase in inhibition, and this was even higher when 10 µg/ml of D29A was added to 5 µg/ml D29B. Unfortunately a control with only 5 µg/ml of D29B was not prepared, so it is difficult to say with certainty that the effect was truly synergistic. In sum, there is potentially a synergistic interaction between D29A and D29B in growth inhibition of *M. smegmatis*, but it should be explored further.

Another interaction that was tested is the combination of D29B with the active site mutant D29B S82A. Increasing amounts of S82A were added to cultures with 10 µg/ml D29B and assessed after 7 days of growth (Figure 42 C). Interestingly, S82A amounts of 25 µg/ml or more limited D29B's inhibition of *M. smegmatis* growth. When 100 µg/ml of S82A was added

with 10 µg/ml D29B, there was cell growth across the entire well (Figure 42 C). However, there was still no maturation of the biofilm, which seems to happen in any situation where a lysin is added. From these results, it appears that the addition of S82A interferes with the effect of D29B on the growth inhibition. Indeed, S82A alone slightly inhibited growth, although there were still some signs of biofilm maturation (Figure 41 C). S82A may have a sort of "dominant negative" effect in its ability to decrease the effectiveness of D29B.
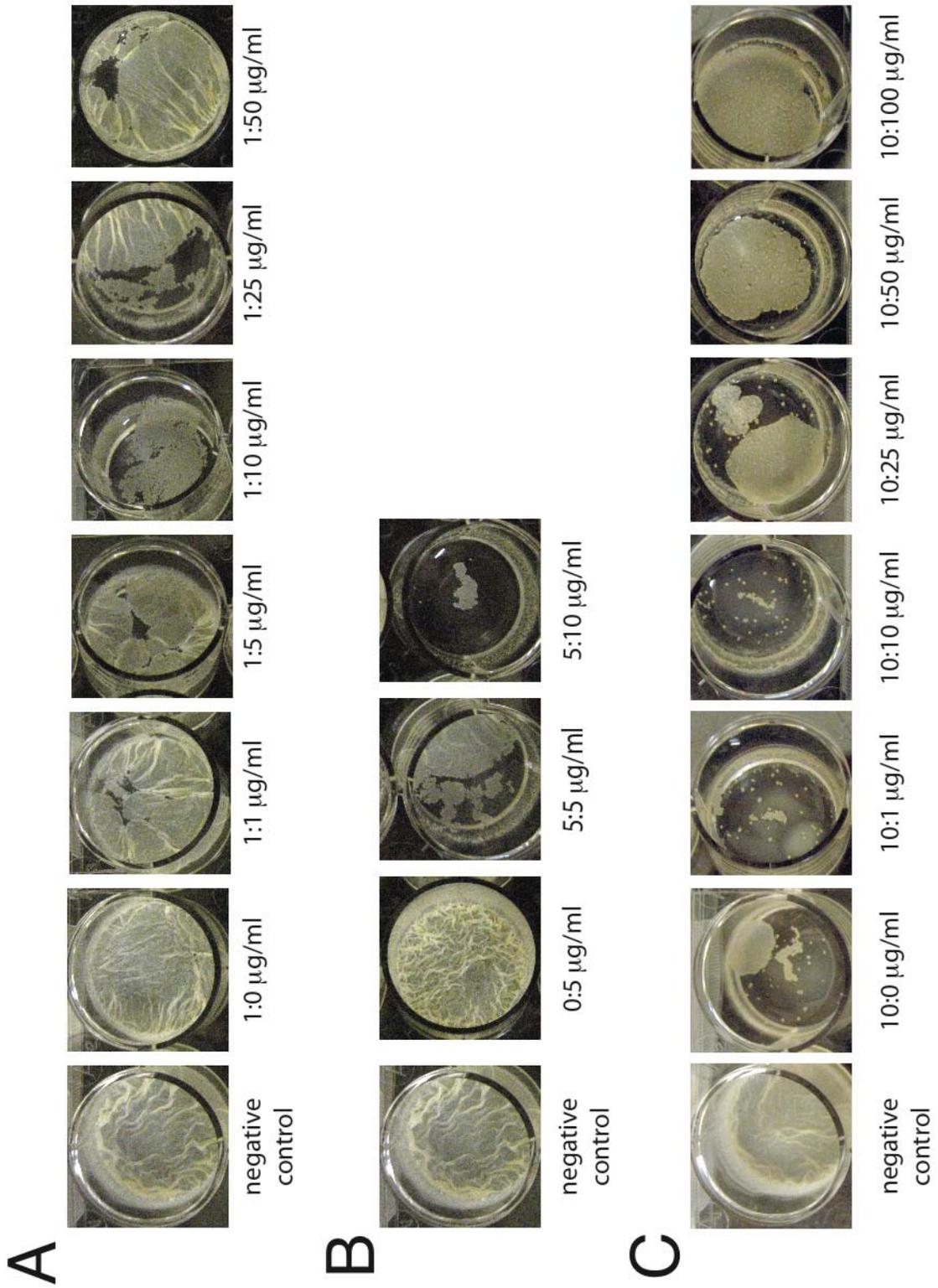
Figure 43. The effect of combinations of lysins on growth of *M. smegmatis*

278

Figure 43: Biofilm pictures were taken after 7 days of growth at 30°C. A. and B. Combinations of D29B and D29A. The amount of D29B precedes that of D29A (*e.g.* D29B µg/ml : D29A µg/ml). C. Increasing amounts of the D29B mutant S82A was added to cultures with 10 µg/ml of D29B. The amount of D29B precedes that of S82A (*e.g.* D29B µg/ml : S82A µg/ml).

A.2.3 Lytic activity of lysins on *Propionibacterium acnes*

The lytic activity of exogenously-applied lysins was also tested on *Propionibacterium acnes*, another member of the order Actinomycetales that is a Gram-positive bacteria but lacks mycolylarabinogalactan or the complex hydrophobic outer membrane of mycobacteria (Kamisango et al., 1982). Purified lysins were added to cultures of *P. acnes*, and after 30-60 min clumping and lysis were visible. The OD_{600} of the cultures decreased over time (data not shown), and samples with the most rapid decrease in OD corresponded to the more extreme lytic activity seen below. Samples of the cultures were mounted on a slide and observed using bright-field microscopy (Figure 44). The strongest lytic activity was observed with Bxz1 gp236 (Figure 44 B), followed by D29 gp10 (Figure 44 E). Significant lysis was seen with Che8 gp32 (Figure 44 C), although it required more time to lyse the culture. Corndog gp69 (Figure 44 D) and the control lysozyme (Figure 44 G) showed a little activity; the cultures slowly lysed, but it was not very apparent under the microscope. Surprisingly, the D29 gp12 LysB also caused the culture to slowly lyse, although it was also not evident when a sample was examined with the microscope (Figure 44 F). Nonetheless, the strong activity of Bxz1 gp236 and D29 gp10 demonstrates that these LysAs can effect lysis of bacteria when applied exogenously.

Figure 44. Addition of lysins to *Propionibacterium acnes*

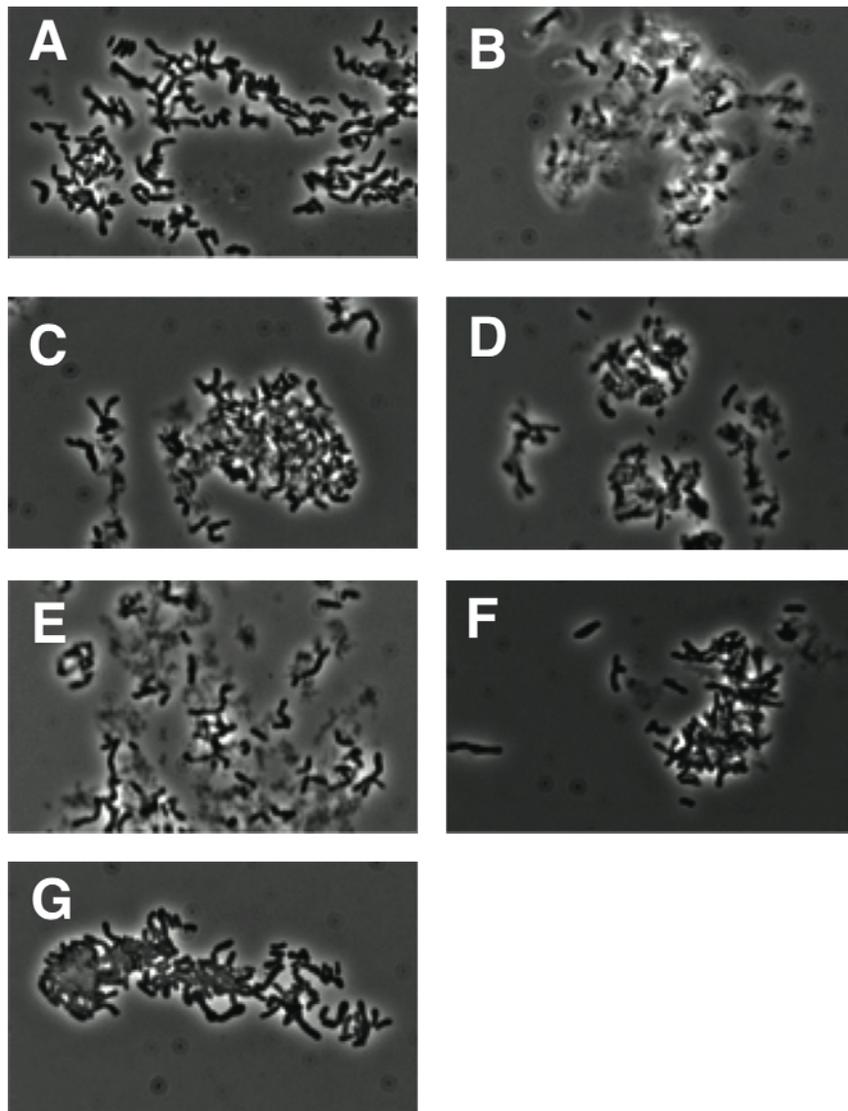**Figure 44:** Purified lysins were added exogenously to cultures of *P. acnes* and, after 30-60 min of incubation at room temperature, samples were mounted on a slide and examined using bright-field microscopy. **A.** Negative control; **B.** Bxz1 gp236; **C.** Che8 gp32; **D.** Corndog gp69; **E.** D29 gp10; **F.** D29 gp12; **G.** lysozyme.

A.2.4 Conclusions

The external application of purified lysins to *M. smegmatis* does not have the dramatic effect that other lysins have shown on Gram-positive bacteria. This is not surprising considering the exceptionally effective barrier of lipids mycobacteria have assembled around the cell, which can elaborate into a complex biofilm under the right conditions (Zambrano and Kolter, 2005). In spite of this, exogenously-applied LysA and LysB proteins did appear to inhibit the growth and biofilm formation by *M. smegmatis* in detergent-free cultures. D29B was especially effective (Figure 41 B), although this may be influenced by the stability of the protein or its rate of action (Figure 42). Surprisingly, the S82A active site mutant of D29B also inhibited biofilm maturation to some extent (Figure 41 C). The potency of D29B and the effect of S82A may be tied to the composition of mycobacterial cell walls and the extracellular matrix (Brennan, 2003; Ojha et al., 2008; Ortalo-Magne et al., 1996). Numerous glycolipids, including TDM, are in the mycobacterial outer membrane, and the biofilm matrix is comprised of free mycolic acids that have been released from TDM by a mycobacterial esterase. It has been shown that D29B is also able to hydrolyze TDM to release mycolic acids (Ojha et al., 2010), and so the LysB may interfere with the formation of biofilms by binding substrates, unregulated hydrolysis, or some other mechanism. Interference by binding the substrates or products may explain why the S82A mutant is able to affect the maturation of the biofilm but does not appear to inhibit growth significantly (Figure 41 C). Competition for substrates may also explain why S82A is able to decrease the inhibitory effects of D29B (Figure 43 C). However, the difference in the extent of inhibition between S82A and D29B shows that the inhibition of growth by D29B is connected to its enzymatic activity.

The mechanism of inhibition by LysA proteins is unclear, since it is highly unlikely that they have access to the PG from outside of the cells. It would be interesting if the inhibitory activity was associated with specific N-terminal domains or other activities, but that possibility was not addressed in these studies. There may also have been some synergy in growth inhibition when D29A and D29B were added simultaneously to a culture, but the majority of the activity still appears to be the result of D29B (Figure 32 A). However, the degradation of D29A in the growth media may complicate these results (Figure 42).

The activity of the lysins on *P. acnes* demonstrates their ability to effect lysis exogenously. It is somewhat surprising to see such strong activity by Bxz1 gp236 and D29 gp10, since lysins are thought to be highly specific to their phage's host, but *P. acnes* is significantly different; it lacks arabinogalactan and mycolic acids, and does not have type A1γ PG (Kamisango et al., 1982). Intriguingly, the only domain Bxz1 gp236 and D29 gp10 share is C1 (Figure 37), which is presumed to be involved in cell wall binding. Che8 gp32 and Corndog gp69 both have C2 domains and both show less effective lysis of *P. acnes*. Therefore, the C1 domain may represent a less-specific cell wall binding domain that confers a broader specificity on some LysAs.

These results warrant further assessment of the activity of purified lysins on mycobacteria. Studies involving alternate lysins need to be conducted, and it would be very interesting to identify any synergistic activity when combining lysins with antibiotics such as isoniazid and rifampicin. Many of these experiments should also be conducted with *M. tuberculosis*; in fact, several of the experiments discussed in this appendix were also attempted with *M. tuberculosis*, but either failed due to technical difficulties or were not completed as of this writing.

A.3   ENDOGENOUS EXPRESSION OF LYSINS IN *M. SMEGMATIS*

Outside of the context of phage infection, lysin research is primarily concerned with exogenous application and resulting lysis (Borysowski et al., 2006; Fischetti, 2008; Hermoso et al., 2007; Loessner, 2005). The canonical holin-endolysin system emphasizes the requirement of a holin for the hydrolytic activity of a lysin (Young, 2002), whether by allowing it access to the PG or activation by depolarization of the membrane. Studies expressing lysins within cells have required a means of permeabilizing the membrane, usually with chloroform, in order to effect lysis (Ames et al., 1984). Therefore, lysins are not expected to have significant activity when expressed from within cells in the absence of a holin or a compound serving an analogous function.

Lysins are known to be essential for plaque formation (Young, 1992), and so as a proof of principle experiment for mycobacteriophage recombineering, Marinelli *et al.* (2008) created a deletion mutant of Giles *lysA*. The resulting mutant phage was unable to form plaques unless infecting a complementing strain, one that was expressing an alternate LysA. This requirement proved the essentiality of LysA in mycobacteriophage Giles, and the same method demonstrated that LysB was not essential, since it could plaque without complementation (Chapter 3). The complementing strain was *M. smegmatis* mc^{2}155:pKMC4, which expresses Corndog gp69 under an acetamidase promoter (Marinelli et al., 2008). However, the first attempt to complement the Giles∆*lysA* mutant used pKMC1 instead, which expressed the D29 gp10 LysA. When plated on inducing media (plates including acetamide), no growth was seen. This bactericidal phenotype from pKMC1 expression was intriguing, especially when pKMC4 expression had no effect on

growth.   We decided to examine the effects of endogenous expression of other mycobacteriophage lysins on mycobacterial growth.

A.3.1   Cloning and expression of a diverse set of LysA proteins into pLAM12

In addition to Bxz1 gp236, Che8 gp32, Corndog gp69, and D29 gp10, 8 additional LysA proteins were cloned to create a set of 12 LysAs that included at least one representative of each domain (Figure 45), excluding the transglycosylase domain of Hammer gp13. All were cloned into the acetamide-inducible pLAM12 vector and transformed into *M. smegmatis* mc^{2}155.

We attempted to assess the expression of the protein upon induction and the distribution (pellet or supernatant) by separating induced and uninduced culture samples on SDS PAGE gels (Figure 46). Larger amounts of sample were loaded for D29 gp10, Kostya gp33, and L5 gp10 because of the decreased density in culture (see A.3.3). Expression of protein was visible for almost all LysAs. Several showed low or uncertain expression (Figure 46, question marks). The presence of protein was difficult to verify with L5 gp10, although based on later results it is assumed that protein was expressed. For all of the LysAs, the majority of the protein was consistently found in the pellet; this could reflect binding of the PG, which would be associated with the rest of the cell wall in the pellet. There does not appear to be any association of protein in the supernatant with a specific C-terminal domain, nor with any other domains. The LysAs with the strongest expression included Brujita gp29, Bxz1 gp236, and Rosebush gp46 (Figure 46).
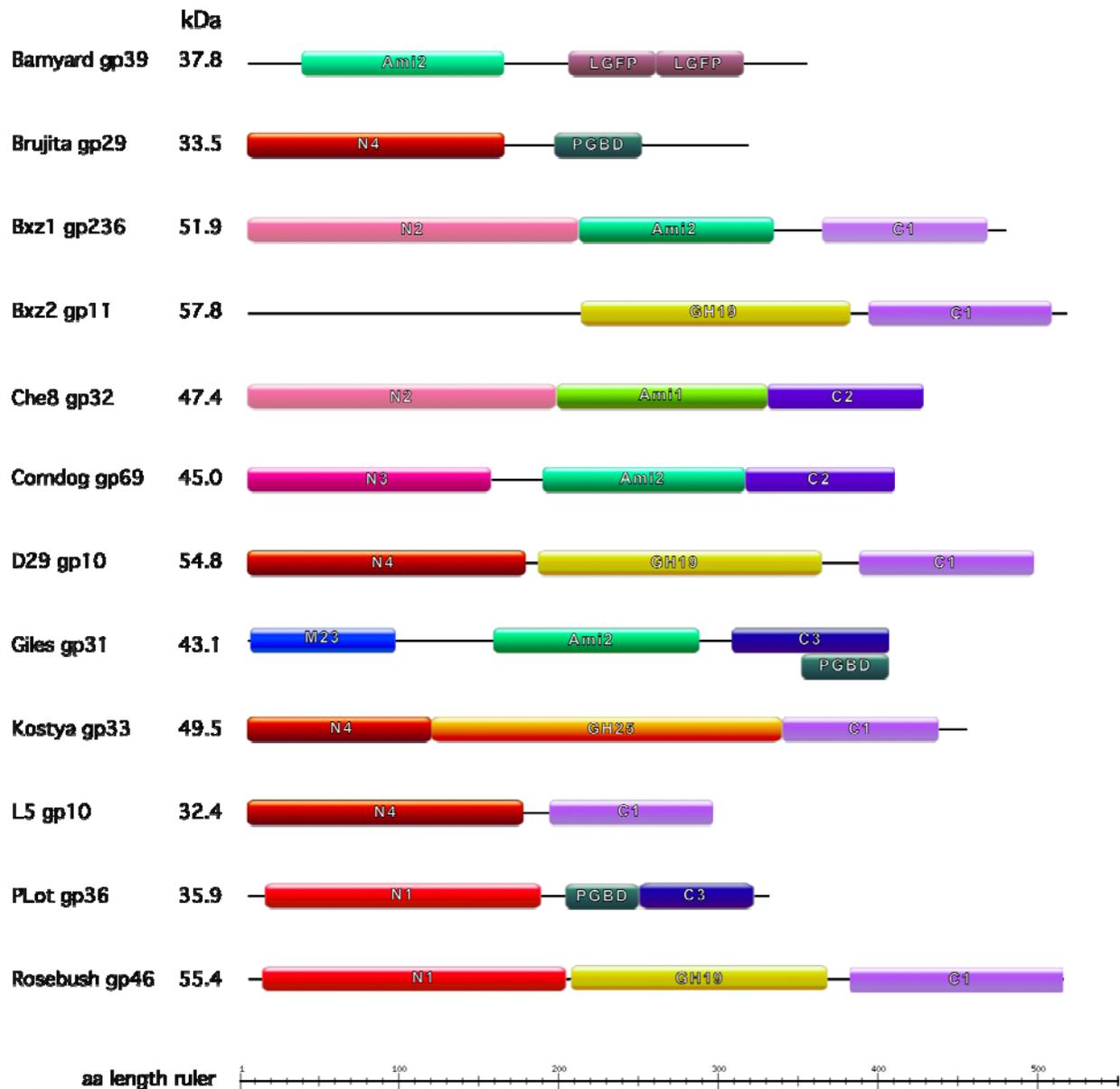
Figure 45. Cloned LysAs for expression in *M. smegmatis*

Figure 45: The above 12 LysAs were cloned into the acetamide-inducible pLAM12 vector. They were chosen to represent the diversity of domains, with each domain represented at least once and all three unmatching regions present (C-terminus of Barnyard gp39, C-terminus of Brujita gp29, and N-terminus of Bxz2 gp11). Molecular weights are listed to the right of the LysA name, and an amino acid length ruler is below.
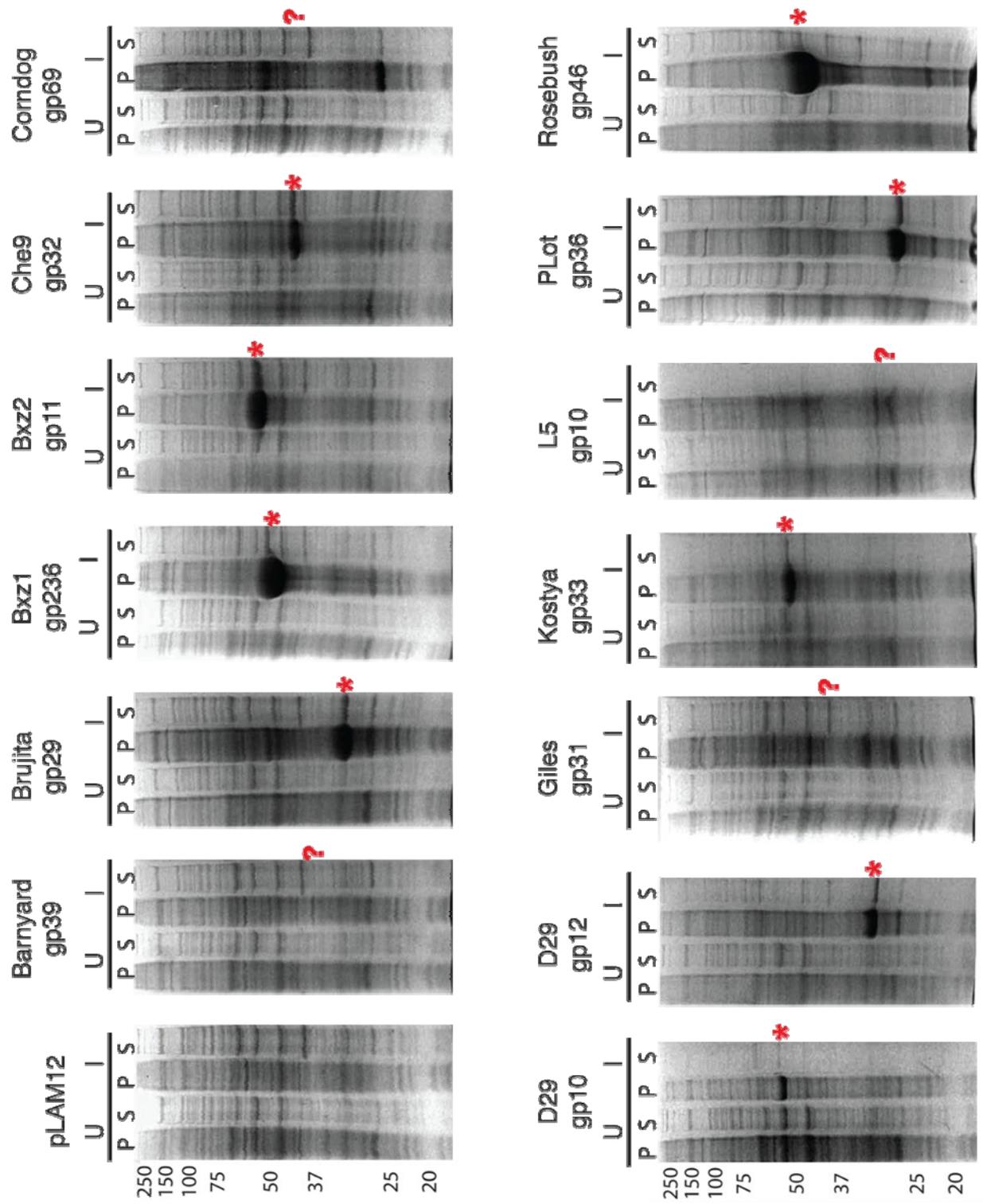
Figure 46. Expression profiles of lysins induced in *M. smegmatis*

Figure 46: The above samples are from the cultures used in the ATP assay described in A.3.2. One milliliter samples were taken from induced and uninduced cultures with pLAM12 vector control, one of 12 LysAs, or one LysB (D29 gp10). These were sonicated, centrifuged at 14,000 rpm to produce pellet and supernatant fractions, and separated on SDS PAGE gels. Red asterisks mark identifiable expressed protein. Red question marks indicate instances of uncertain protein expression.

A.3.2   ATP release from *M. smegmatis* endogenously expressing lysins

Upon reaching log-phase growth, all 12 LysAs and the D29 LysB cultures were split, and half were induced with 0.2% acetamide. Induced and uninduced cultures were sampled hourly for 7 hours, and the release of ATP was measured using a luminometer and luciferase reagents (see Chapter 3). The fold-difference in ATP release between induced and uninduced strains was plotted over time (Figure 47). The results showed an increase in ATP release by three LysAs: D29 gp10, Kostya gp33, and L5 gp10. There was also a vast difference in the degree of release between these three LysAs; D29 gp10 was about 2-fold greater than Kostya gp33, but L5 gp10 released nearly 3-fold more ATP than D29 gp10 and 4.5-fold more than Kostya gp33.
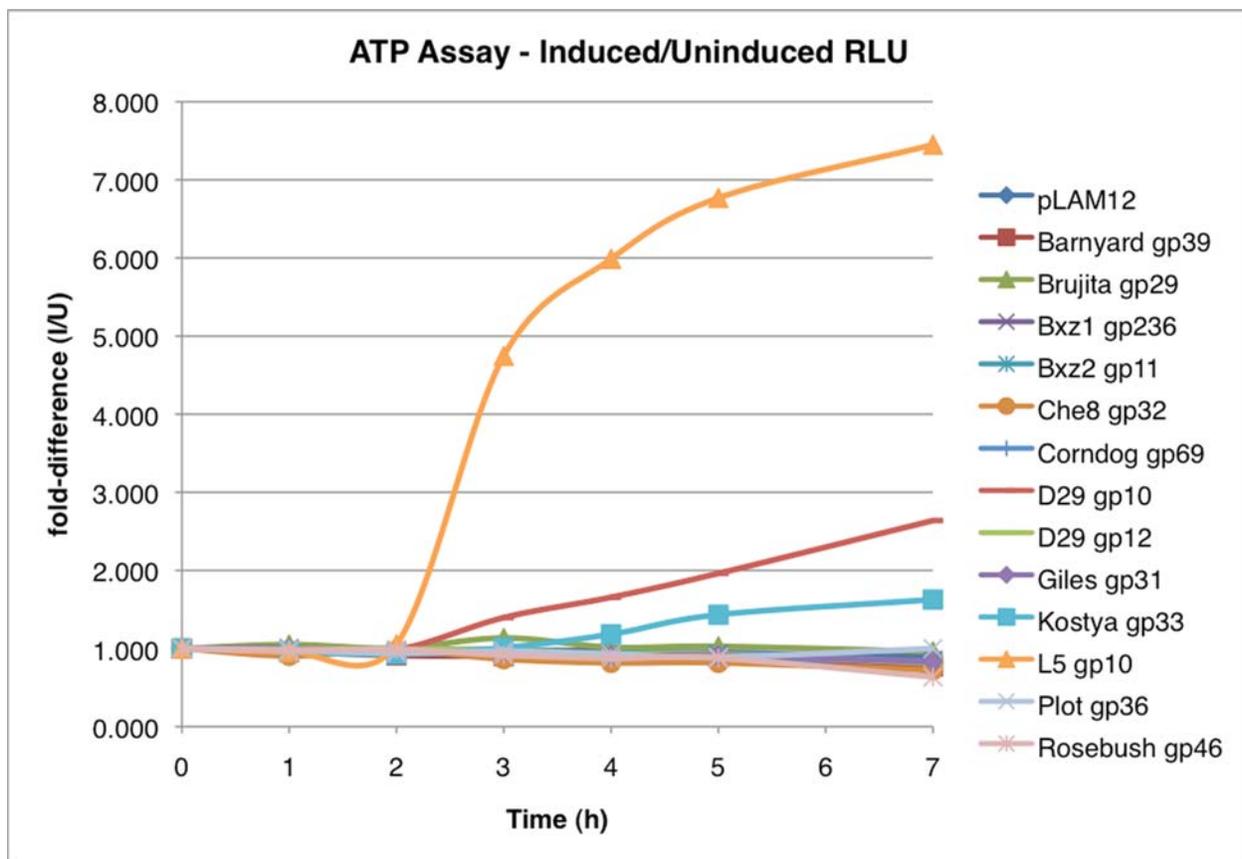
Figure 47. ATP release upon induction of lysin expression

Figure 47: Cultures of *M. smegmatis* mc²155 carrying pLAM12-based plasmids with different lysins were split and half were induced with 0.2% acetamide. ATP release was measured for 7 hours and the fold-difference between induced and uninduced ATP release calculated for each lysin and plotted versus the time.

A.3.3 Visible lysis and growth inhibition of lysin-expressing strains

The induced cultures from A.3.2 were allowed to grow overnight at 37°C, and the turbidity was examined the next day. Most cultures did not vary significantly in cell density from the negative control (pLAM12 empty vector), but significant differences were observed for D29 gp10, Kostya gp33, and L5 gp10 (Figure 48). Kostya gp33 had a noticeable decrease in turbidity compared to pLAM12; however, the cultures expressing D29 gp10 and L5 gp12 were completely lysed (Figure 48).

To analyze the effects of lysin induction on *M. smegmatis* growth on solid media, cultures were plated on media with or without 0.2% acetamide. Plates were placed at 37°C overnight and colony size was assessed. Four outcomes were observed when comparing induced to uninduced colonies: no change in colony size, a moderate decrease in colony size, a large reduction in size (pinprick colonies), and no growth. The negative control strain carrying pLAM12 showed no difference in colony size (Figure 49 A), as did Barnyard gp39, Bxz2 gp11, and Giles gp31. This could be due to the poor expression seen when inducing Barnyard gp39 and Giles gp31 (Figure 46), but Bxz2 gp11 showed good protein expression (Figure 46). Brujita gp29, Che8 gp32, Corndog gp69, Plot gp36, and Rosebush gp46 had only a moderate reduction in size (Figure 49 B). Bxz1 gp236, D29 gp12, and Kostya gp33 had tiny colonies (Figure 49 C). Only D29 gp10 and L5 gp12 showed no growth (not shown).
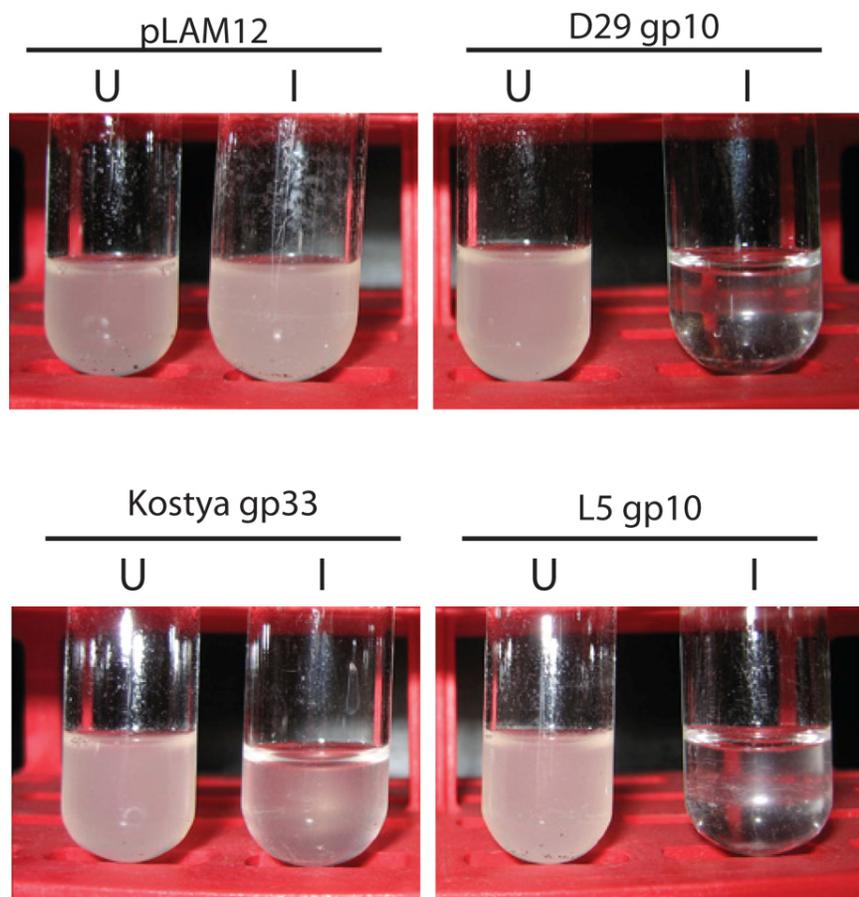
Figure 48. Turbidity of induced and uninduced *M. smegmatis* cultures expressing lysin

Figure 48: Induced and uninduced *M. smegmatis* cultures used for the ATP assay were allowed to grow overnight at 37°C, and the turbidity was compared the following day. pLAM12 is the vector negative control. U is uninduced; I is induced.
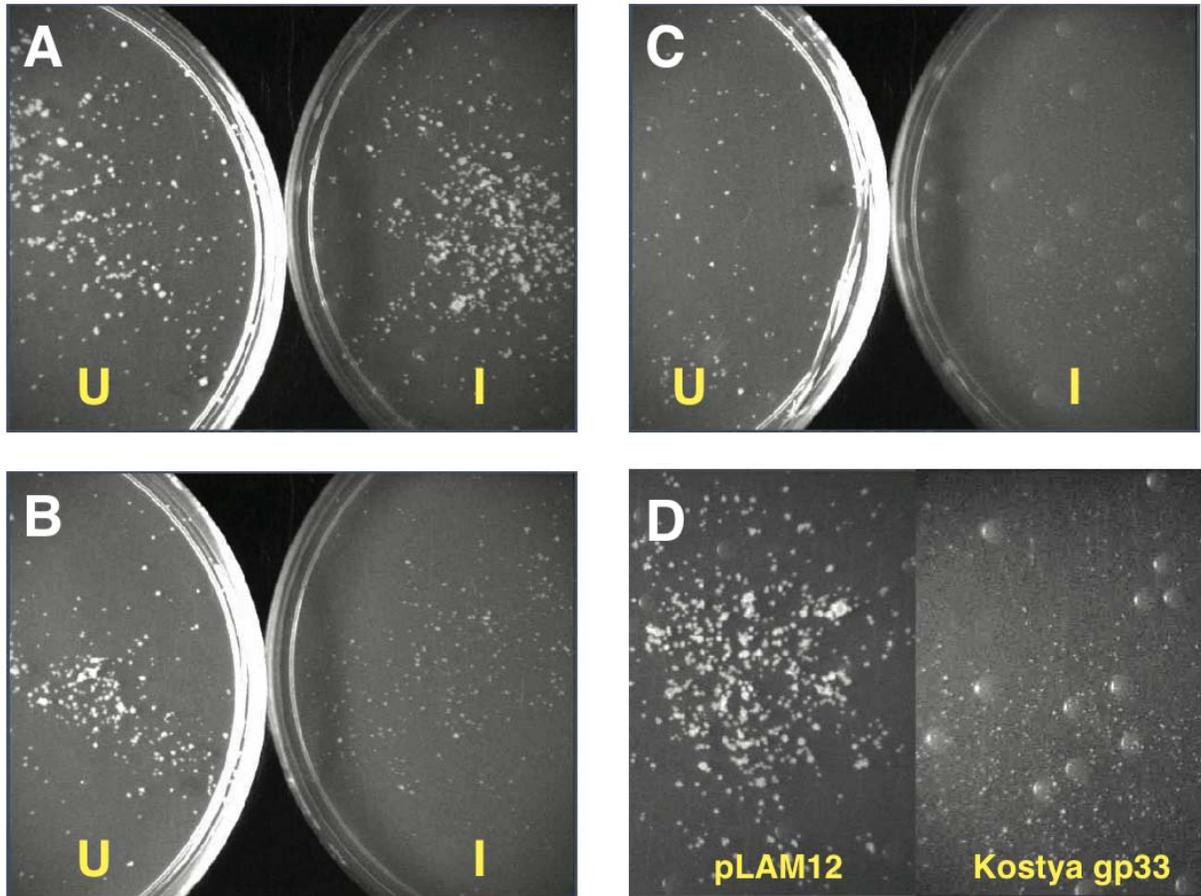
Figure 49. Effect of endogenous lysin expression on *M. smegmatis* growth on solid media

Figure 49: The colony size of *M. smegmatis* containing acetamide-inducible lysins was compared between uninduced and induced. Representative plates are shown for three of the four phenotypes scored for growth based on colony size. **A.** No significant difference in size between uninduced and induced. The empty vector pLAM12 plates are shown and represent results for Barnyard gp39, Bxz2 gp11, and Giles gp11. **B.** Moderate reduction in colony size when lysin expression was induced. Colonies for Plot gp36 are shown and represent the results for Brujita gp29, Che8 gp32, Corndog gp69, and Rosebush gp46. **C.** Induction that produced pinprick-sized colonies. Kostya gp33 represents the results for Bxz1 gp236 and D29 gp12 (LysB). **D.** Closeup of colonies for induced pLAM12 and Kostya gp33. The fourth phenotype – no visible growth – was observed for D29 gp10 and L5 gp10 and is not shown here.

293

A.3.4 Conclusions

It is not difficult to imagine that endogenous expression of toxic proteins such as lysins in *M. smegmatis* could have an impact on growth; however, it is surprising to see actual lysis of *M. smegmatis* by D29 gp10, L5 gp10, and to a lesser extent, Kostya gp33, considering these proteins were expressed in the absence of a holin. These three LysAs share N4 and C1 domains (Figure 37). D29 gp10 and Kostya gp33 are predicted to have different lysozyme activities – GH19 and GH25, respectively – but L5 gp10 is comprised solely of an N4 and C1 domain, with no visible lytic domain (see Chapter 2). The C1 domains were found in many of the non-active LysAs (Figure 37) and so are unlikely to be responsible for the lytic activity. The strong bactericidal effect of endogenous L5 gp10 expression strongly supports a lytic activity for N4. However, this raises two questions: if the N4 domain is so lethal, why do D29 gp10 and Kostya gp33 contain other PG hydrolytic activities? And what is the substrate of N4?

Somewhat complicating the attribution of the lytic activity to N4 is Brujita gp29, which also has an N4 domain and no other predicted activity (Figure 37), similar in organization to L5 gp10 except for differing C-termini. It may be that both the N4 and C1 domains are required, since Brujita differs in its C-terminus (Figure 37). Brujita gp29 had strong protein expression (Figure 46) but shows no lysis as measured by ATP release (Figure 47) or a decrease in turbidity (not shown). In comparison with N4 of L5 gp10, the N4 of D29 gp10 is 81.0% similar, of Kostya gp33 is 14.4% similar, and Brujita gp29 is 29.9% similar; thus, decreasing similarity to the N4 of L5 gp10 does not appear to be the reason for a lack of lysis by Brujita gp29.

All known dsDNA phage lysins are PG hydrolases (Hermoso et al., 2007), and only a few examples of non-PG hydrolytic lysis systems have been characterized (Chapter 1). The latter

consist of single proteins that bind to enzymes involved in the biosynthesis of PG (Bernhardt et al., 2002), and so the focus is still on compromising the PG of the host to induce lysis. It would be easy to hypothesize that the N4 domain represents a new PG hydrolytic activity, but its holin-independent lysis brings this into question. Many of the targeted PG bonds are not made until the PG precursors have been assembled outside the cell membrane (Vollmer and Holtje, 2001), so either N4 has a different target that is accessible inside the cell or it is capable of accessing the PG without a holin. It remains to be determined by what mechanism L5 gp10 lyses *M. smegmatis*. Future studies may employ random mutagenesis to identify important residues, possible generation of resistant host mutants to identify important lysin-host interactions, and chemical analysis of lysis products by mass-spectroscopy or NMR to determine the molecular target. In addition, L5 gp10 and all lysins studied should be examined for activity on *M. tuberculosis*.

BIBLIOGRAPHY

1. Adam, A., Petit, J.F., Wietzerbin-Falszpan, J., Sinay, P., Thomas, D.W., and Lederer, E. (1969). FEBS Lett *4*, 87-92.

2. Aderem, A., and Ulevitch, R.J. (2000). Toll-like receptors in the induction of the innate immune response. Nature *406*, 782-787.

3. Adindla, S., Inampudi, K.K., Guruprasad, K., and Guruprasad, L. (2004). Identification and analysis of novel tandem repeats in the cell surface proteins of archaeal and bacterial genomes using computational tools. Comp Funct Genomics *5*, 2-16.

4. Aitken, A., and Stanier, R.Y. (1979). Characterization of peptidoglycan from the cyanelles of Cyanophora paradoxa. J Gen Microbiol *112*, 219-223.

5. Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res *25*, 3389-3402.

6. Ames, G.F., Prody, C., and Kustu, S. (1984). Simple, rapid, and quantitative release of periplasmic proteins by chloroform. J Bacteriol *160*, 1181-1183.

7. Arpigny, J.L., and Jaeger, K.E. (1999). Bacterial lipolytic enzymes: classification and properties. Biochem J *343 Pt 1*, 177-183.

8. Bailey, T.L., Williams, N., Misleh, C., and Li, W.W. (2006). MEME: discovering and analyzing DNA and protein sequence motifs. Nucleic Acids Res *34*, W369-373.

9. Barenboim, M., Chang, C.Y., dib Hajj, F., and Young, R. (1999). Characterization of the dual start motif of a class II holin gene. Mol Microbiol *32*, 715-727.

10. Bartel, P.L., Roecklein, J.A., SenGupta, D., and Fields, S. (1996). A protein linkage map of Escherichia coli bacteriophage T7. Nat Genet *12*, 72-77.

11. Bateman, A., and Rawlings, N.D. (2003). The CHAP domain: a large family of amidases including GSP amidase and peptidoglycan hydrolases. Trends Biochem Sci *28*, 234-237.

12. Becker, S.C., Dong, S., Baker, J.R., Foster-Frey, J., Pritchard, D.G., and Donovan, D.M. (2009). LysK CHAP endopeptidase domain is required for lysis of live staphylococcal cells. FEMS Microbiol Lett *294*, 52-60.

13. Bendtsen, J.D., Nielsen, H., von Heijne, G., and Brunak, S. (2004). Improved prediction of signal peptides: SignalP 3.0. J Mol Biol *340*, 783-795.

14. Benson, D.A., Karsch-Mizrachi, I., Lipman, D.J., Ostell, J., and Sayers, E.W. (2009). GenBank. Nucleic Acids Res *37*, D26-31.

15. Bernhardt, T.G., Roof, W.D., and Young, R. (2000). Genetic evidence that the bacteriophage phi X174 lysis protein inhibits cell wall synthesis. Proc Natl Acad Sci U S A *97*, 4297-4302.

16. Bernhardt, T.G., Struck, D.K., and Young, R. (2001a). The lysis protein E of phi X174 is a specific inhibitor of the MraY-catalyzed step in peptidoglycan synthesis. J Biol Chem *276*, 6093-6097.

17. Bernhardt, T.G., Wang, I.N., Struck, D.K., and Young, R. (2001b). A protein antibiotic in the phage Qbeta virion: diversity in lysis targets. Science *292*, 2326-2329.

18. Bernhardt, T.G., Wang, I.N., Struck, D.K., and Young, R. (2002). Breaking free: "protein antibiotics" and phage lysis. Res Microbiol *153*, 493-501.

19. Berry, J., Summer, E.J., Struck, D.K., and Young, R. (2008). The final step in the phage infection cycle: the Rz and Rz1 lysis proteins link the inner and outer membranes. Mol Microbiol *70*, 341-351.

20. Besra, G.S. (1998). Preparations of Cell Wall Fractions of Mycobacteria. In Mycobacteria Protocols, T. Parish, and N.G. Stoker, eds. (Towata, NJ, Humana Press Inc.), pp. 91-107.

21. Bibb, L.A., and Hatfull, G.F. (2002). Integration and excision of the Mycobacterium tuberculosis prophage-like element, phiRv1. Mol Microbiol *45*, 1515-1526.

22. Bienkowska-Szewczyk, K., Lipinska, B., and Taylor, A. (1981). The R gene product of bacteriophage lambda is the murein transglycosylase. Mol Gen Genet *184*, 111-114.

23. Blackburn, N.T., and Clarke, A.J. (2001). Identification of four families of peptidoglycan lytic transglycosylases. J Mol Evol *52*, 78-84.

24. Blasi, U., and Young, R. (1996). Two beginnings for a single purpose: the dual-start holins in the regulation of phage lysis. Mol Microbiol *21*, 675-682.

25. Bochtler, M., Odintsov, S.G., Marcyjaniak, M., and Sabala, I. (2004). Similar active sites in lysostaphins and D-Ala-D-Ala metallopeptidases. Protein Sci *13*, 854-861.

26.   Bogaert, D., De Groot, R., and Hermans, P.W. (2004). Streptococcus pneumoniae colonisation: the key to pneumococcal disease. Lancet Infect Dis *4*, 144-154.

27. Borysowski, J., Weber-Dabrowska, B., and Gorski, A. (2006). Bacteriophage endolysins as a novel class of antibacterial agents. Exp Biol Med (Maywood) *231*, 366-377.

28. Brennan, P.J. (2003). Structure, function, and biogenesis of the cell wall of Mycobacterium tuberculosis. Tuberculosis (Edinb) *83*, 91-97.

29.   Brennan, P.J., and Besra, G.S. (1997). Structure, function and biogenesis of the mycobacterial cell wall. Biochem Soc Trans *25*, 188-194.

30. Brennan, P.J., and Nikaido, H. (1995). The envelope of mycobacteria. Annu Rev Biochem *64*, 29-63.

31. Briers, Y., Volckaert, G., Cornelissen, A., Lagaert, S., Michiels, C.W., Hertveldt, K., and Lavigne, R. (2007). Muralytic activity and modular structure of the endolysins of Pseudomonas aeruginosa bacteriophages phiKZ and EL. Mol Microbiol *65*, 1334-1344.

32.   Briken, V., Porcelli, S.A., Besra, G.S., and Kremer, L. (2004). Mycobacterial lipoarabinomannan and related lipoglycans: from biogenesis to modulation of the immune response. Mol Microbiol *53*, 391-403.

33. Bryl, K., Kedzierska, S., Laskowska, M., and Taylor, A. (2000). Membrane fusion by proline-rich Rz1 lipoprotein, the bacteriophage lambda Rz1 gene product. Eur J Biochem *267*, 794-799.

34. Bukovska, G., Klucar, L., Vlcek, C., Adamovic, J., Turna, J., and Timko, J. (2006). Complete nucleotide sequence and genome analysis of bacteriophage BFK20--a lytic phage of the industrial producer *Brevibacterium flavum*. Virology *348*, 57-71.

35. Bull, J.J., Pfennig, D.W., and Wang, I.N. (2004). Genetic details, optimization and phage life histories. Trends Ecol Evol *19*, 76-82.

36.   Burman, L.G., Reichler, J., and Park, J.T. (1983). Evidence for multisite growth of Escherichia coli murein involving concomitant endopeptidase and transpeptidase activities. J Bacteriol *156*, 386-392.

37. Calendar, R. (2006). The Bacteriophages (New York, Oxford University Press).

38. Carvalho, C.M., Aires-Barros, M.R., and Cabral, J.M. (1999). Cutinase: from molecular level to bioprocess development. Biotechnol Bioeng *66*, 17-34.

39. Catalao, M.J., Gil, F., Moniz-Pereira, J., and Pimentel, M. (2010). The mycobacteriophage Ms6 encodes a chaperone-like protein involved in the endolysin delivery to the peptidoglycan. Mol Microbiol *77*, 672-686.

40. CDC (2009). Extensively Drug-Resistant Tuberculosis (XDR TB). In wwwcdcgov/tb, C.f.D.C.a. Prevention, ed. (Atlanta, GA).

41. Chang, A., Scheer, M., Grote, A., Schomburg, I., and Schomburg, D. (2009). BRENDA, AMENDA and FRENDA the enzyme information system: new content and tools in 2009. Nucleic Acids Res *37*, D588-592.

42. Chatterjee, D., Bozic, C.M., McNeil, M., and Brennan, P.J. (1991). Structural features of the arabinan component of the lipoarabinomannan of Mycobacterium tuberculosis. J Biol Chem *266*, 9652-9660.

43. Chen, C.L., Pan, T.Y., Kan, S.C., Kuan, Y.C., Hong, L.Y., Chiu, K.R., Sheu, C.S., Yang, J.S., Hsu, W.H., and Hu, H.Y. (2008). Genome sequence of the lytic bacteriophage P1201 from *Corynebacterium glutamicum* NCHU 87078: evolutionary relationships to phages from Corynebacterineae. Virology *378*, 226-232.

44. Cheng, Q., and Fischetti, V.A. (2007). Mutagenesis of a bacteriophage lytic enzyme PlyGBS significantly increases its antibacterial activity against group B streptococci. Appl Microbiol Biotechnol *74*, 1284-1291.

45. Cheng, Q., Nelson, D., Zhu, S., and Fischetti, V.A. (2005). Removal of group B streptococci colonizing the vagina and oropharynx of mice with a bacteriophage lytic enzyme. Antimicrob Agents Chemother *49*, 111-117.

46. Cheng, X., Zhang, X., Pflugrath, J.W., and Studier, F.W. (1994). The structure of bacteriophage T7 lysozyme, a zinc amidase and an inhibitor of T7 RNA polymerase. Proc Natl Acad Sci U S A *91*, 4034-4038.

47. Climo, M.W., Patron, R.L., Goldstein, B.P., and Archer, G.L. (1998). Lysostaphin treatment of experimental methicillin-resistant Staphylococcus aureus aortic valve endocarditis. Antimicrob Agents Chemother *42*, 1355-1360.

48. Cohen-Kupiec, R., and Chet, I. (1998). The molecular biology of chitin digestion. Curr Opin Biotechnol *9*, 270-277.

49. Connolly, L.E., Edelstein, P.H., and Ramakrishnan, L. (2007). Why Is Long-Term Therapy Required to Cure Tuberculosis? PLoS Med *4*, e120.

50. Costa, K., Bacher, G., Allmaier, G., Dominguez-Bello, M.G., Engstrand, L., Falk, P., de Pedro, M.A., and Garcia-del Portillo, F. (1999). The morphological transition of Helicobacter pylori cells from spiral to coccoid is preceded by a substantial modification of the cell wall. J Bacteriol *181*, 3710-3715.

51. Crick, D.C., Mahapatra, S., and Brennan, P.J. (2001). Biosynthesis of the arabinogalactan-peptidoglycan complex of Mycobacterium tuberculosis. Glycobiology *11*, 107R-118R.

52. Croux, C., Ronda, C., Lopez, R., and Garcia, J.L. (1993a). Interchange of functional domains switches enzyme specificity: construction of a chimeric pneumococcal-clostridial cell wall lytic enzyme. Mol Microbiol *9*, 1019-1025.

53. Croux, C., Ronda, C., Lopez, R., and Garcia, J.L. (1993b). Role of the C-terminal domain of the lysozyme of Clostridium acetobutylicum ATCC 824 in a chimeric pneumococcal-clostridial cell wall lytic enzyme. FEBS Lett *336*, 111-114.

54. Cummins, C.S., and Harris, H. (1956). The chemical composition of the cell wall in some gram-positive bacteria and its possible value as a taxonomic character. J Gen Microbiol *14*, 583-600.

55. D'Herelle, F. (2007). On an invisible microbe antagonistic toward dysenteric bacilli: brief note by Mr. F. D'Herelle, presented by Mr. Roux. 1917. Res Microbiol *158*, 553-554.

56. Daffe, M., Brennan, P.J., and McNeil, M. (1990). Predominant structural features of the cell wall arabinogalactan of Mycobacterium tuberculosis as revealed through characterization of oligoglycosyl alditol fragments by gas chromatography/mass spectrometry and by 1H and 13C NMR analyses. J Biol Chem *265*, 6734-6743.

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

58. Dajcs, J.J., Hume, E.B., Moreau, J.M., Caballero, A.R., Cannon, B.M., and O'Callaghan, R.J. (2000). Lysostaphin treatment of methicillin-resistant Staphylococcus aureus keratitis in the rabbit. Invest Ophthalmol Vis Sci *41*, 1432-1437.

59. David, H.L., Seres-Clavel, S., Clement, F., and Rastogi, N. (1984). Further observations on the mycobacteriophage D29-mycobacterial interactions. Acta Leprol *2*, 359-367.

60. David, M., Lubinsky-Mink, S., Ben-Zvi, A., Ulitzur, S., Kuhn, J., and Suissa, M. (1992). A stable Escherichia coli-Mycobacterium smegmatis plasmid shuttle vector containing the mycobacteriophage D29 origin. Plasmid *28*, 267-271.

61. de Ruyter, P.G., Kuipers, O.P., Meijer, W.C., and de Vos, W.M. (1997). Food-grade controlled lysis of Lactococcus lactis for accelerated cheese ripening. Nat Biotechnol *15*, 976-979.

62. de Vries, J., Harms, J., Broer, I., Kriete, G., Mahn, A., Duering, K., and Wackernagel, W. (1999). The bacteriolytic activity in transgenic potatoes expressing a chimeric T4 lysozyme gene and the effect of T4 lysozyme on soil-and phytopathogenic bacteria. Systematic and Applied Microbiology *22*, 280-286.

63. Del Angel, V.D., Dupuis, F., Mornon, J.P., and Callebaut, I. (2002). Viral fusion peptides and identification of membrane-interacting segments. Biochem Biophys Res Commun *293*, 1153-1160.

64. Delgado, C., Francis, G.E., and Fisher, D. (1992). The uses and properties of PEG-linked proteins. Crit Rev Ther Drug Carrier Syst *9*, 249-304.

65. Dewey, J.S., Savva, C.G., White, R.L., Vitha, S., Holzenburg, A., and Young, R. (2010). Micron-scale holes terminate the phage infection cycle. Proc Natl Acad Sci U S A *107*, 2219-2223.

66. Di Perri, G., and Bonora, S. (2004). Which agents should we use for the treatment of multidrug-resistant Mycobacterium tuberculosis? J Antimicrob Chemother *54*, 593-602.

67. Diaz, E., Lopez, R., and Garcia, J.L. (1990). Chimeric phage-bacterial enzymes: a clue to the modular evolution of genes. Proc Natl Acad Sci U S A *87*, 8125-8129.

68. Diaz, E., Lopez, R., and Garcia, J.L. (1991). Chimeric pneumococcal cell wall lytic enzymes reveal important physiological and evolutionary traits. J Biol Chem *266*, 5464-5471.

69. Dijkstra, A.J., and Keck, W. (1996). Peptidoglycan as a barrier to transenvelope transport. J Bacteriol *178*, 5555-5562.

70. Djurkovic, S., Loeffler, J.M., and Fischetti, V.A. (2005). Synergistic killing of Streptococcus pneumoniae with the bacteriophage lytic enzyme Cpl-1 and penicillin or gentamicin depends on the level of penicillin resistance. Antimicrob Agents Chemother *49*, 1225-1228.

71. Doermann, A.H. (1948). Lysis and Lysis Inhibition with Escherichia coli Bacteriophage. J Bacteriol *55*, 257-276.

72. Donovan, D.M., and Foster-Frey, J. (2008). LambdaSa2 prophage endolysin requires Cpl-7-binding domains and amidase-5 domain for antimicrobial lysis of streptococci. FEMS Microbiol Lett *287*, 22-33.

73. Donovan, D.M., Foster-Frey, J., Dong, S., Rousseau, G.M., Moineau, S., and Pritchard, D.G. (2006a). The cell lysis activity of the Streptococcus agalactiae bacteriophage B30 endolysin relies on the cysteine, histidine-dependent amidohydrolase/peptidase domain. Appl Environ Microbiol *72*, 5108-5112.

74. Donovan, D.M., Lardeo, M., and Foster-Frey, J. (2006b). Lysis of staphylococcal mastitis pathogens by bacteriophage phi11 endolysin. FEMS Microbiol Lett *265*, 133-139.

75. Ducati, R.G., Ruffino-Netto, A., Basso, L.A., and Santos, D.S. (2006). The resumption of consumption-- a review on tuberculosis. Mem Inst Oswaldo Cruz *101*, 697-714.

76. During, K., Porsch, P., Mahn, A., Brinkmann, O., and Gieffers, W. (1999). The non-enzymatic microbicidal activity of lysozymes. FEBS Lett *449*, 93-100.

77. Embley, T.M., Goodfellow, M., G., O.D., D., R., and Minnikin, E. (1986). Chemical criteria in the classification of some mycolateless wall chemotypeIV actinomycetes. In Biological, Biochemical and Biomedical Aspects of Actinomycetes, G. Szabo, S. Biro, and M. Goodfellow, eds. (Budapest: Akademiai Kiado).

78. Entenza, J.M., Loeffler, J.M., Grandgirard, D., Fischetti, V.A., and Moreillon, P. (2005). Therapeutic effects of bacteriophage Cpl-1 lysin against Streptococcus pneumoniae endocarditis in rats. Antimicrob Agents Chemother *49*, 4789-4792.

79. Epand, R.M., and Vogel, H.J. (1999). Diversity of antimicrobial peptides and their mechanisms of action. Biochim Biophys Acta *1462*, 11-28.

80. Felsenstein, J. (1985). Confidence limits on phylogenies: an approach using the bootstrap. Evolution *39*, 783-791.

81. Felsenstein, J. (1993). PHYLIP (Phylogeny Inference Package) version 3.5c. (Department of Genetics, University of Washington, Seattle., Distributed by the author.).

82. Finn, R.D., Mistry, J., Tate, J., Coggill, P., Heger, A., Pollington, J.E., Gavin, O.L., Gunasekaran, P., Ceric, G., Forslund, K.*, et al.* (2010). The Pfam protein families database. Nucleic Acids Res *38*, D211-222.

83. Fischetti, V.A. (2003). Novel method to control pathogenic bacteria on human mucous membranes. Ann N Y Acad Sci *987*, 207-214.

84. Fischetti, V.A. (2005). Bacteriophage lytic enzymes: novel anti-infectives. Trends Microbiol *13*, 491-496.

85. Fischetti, V.A. (2006). Using phage lytic enzymes to control pathogenic bacteria. BMC Oral Health *6 Suppl 1*, S16.

86. Fischetti, V.A. (2008). Bacteriophage lysins as effective antibacterials. Curr Opin Microbiol *11*, 393-400.

87. Flach, J., Pilet, P.E., and Jolles, P. (1992). What's new in chitinase research? Experientia *48*, 701-716.

88. Ford, M.E., Sarkis, G.J., Belanger, A.E., Hendrix, R.W., and Hatfull, G.F. (1998). Genome structure of mycobacteriophage D29: implications for phage evolution. J Mol Biol *279*, 143-164.

89. Fordham, W.D., and Gilvarg, C. (1974). Kinetics of cross-linking of peptidoglycan in Bacillus megaterium. J Biol Chem *249*, 2478-2482.

90. Frias, M.J., Melo-Cristino, J., and Ramirez, M. (2009). The autolysin LytA contributes to efficient bacteriophage progeny release in Streptococcus pneumoniae. J Bacteriol *191*, 5428-5440.

91. Froman, S., Will, D.W., and Bogen, E. (1954). Bacteriophage active against virulent Mycobacterium tuberculosis. I. Isolation and activity. Am J Public Health Nations Health *44*, 1326-1333.

92. Fujinami, Y., Hirai, Y., Sakai, I., Yoshino, M., and Yasuda, J. (2007). Sensitive detection of Bacillus anthracis using a binding protein originating from gamma-phage. Microbiol Immunol *51*, 163-169.

93. Gaeng, S., Scherer, S., Neve, H., and Loessner, M.J. (2000). Gene cloning and expression and secretion of Listeria monocytogenes bacteriophage-lytic enzymes in Lactococcus lactis. Appl Environ Microbiol *66*, 2951-2958.

94. Garcia, E., Garcia, J.L., Garcia, P., Arraras, A., Sanchez-Puelles, J.M., and Lopez, R. (1988). Molecular evolution of lytic enzymes of Streptococcus pneumoniae and its bacteriophages. Proc Natl Acad Sci U S A *85*, 914-918.

95. Garcia, J.L., Garcia, E., Arraras, A., Garcia, P., Ronda, C., and Lopez, R. (1987). Cloning, purification, and biochemical characterization of the pneumococcal bacteriophage Cp-1 lysin. J Virol *61*, 2573-2580.

96. Garcia, M., Pimentel, M., and Moniz-Pereira, J. (2002). Expression of Mycobacteriophage Ms6 lysis genes is driven by two sigma(70)-like promoters and is dependent on a transcription termination signal present in the leader RNA. J Bacteriol *184*, 3034-3043.

97. Garcia, P., Garcia, J.L., Garcia, E., Sanchez-Puelles, J.M., and Lopez, R. (1990). Modular organization of the lytic enzymes of Streptococcus pneumoniae and its bacteriophages. Gene *86*, 81-88.

98. Garcia, P., Martin, A.C., and Lopez, R. (1997). Bacteriophages of Streptococcus pneumoniae: a molecular approach. Microb Drug Resist *3*, 165-176.

99. Gardner, G.M., and Weiser, R.S. (1947). Discovery of a bacteriophage for Mycobacterium smegmatis. J Bacteriol *54*, 272.

100. Garrett, J., Fusselman, R., Hise, J., Chiou, L., Smith-Grillo, D., Schulz, J., and Young, R. (1981). Cell lysis by induction of cloned lambda lysis genes. Mol Gen Genet *182*, 326-331.

101. Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M.R., Appel, R.D., and Bairoch, A. (2005). Protein Identification and Analysis Tools on the ExPASy Server. In The Proteomics Protocols Handbook, J. Walker, ed. (Humana Press), pp. 571-607.

102. Ghuysen, J.M. (1968). Use of bacteriolytic enzymes in determination of wall structure and their role in cell metabolism. Bacteriol Rev *32*, 425-464.

103. Gil, F., Catalao, M.J., Moniz-Pereira, J., Leandro, P., McNeil, M., and Pimentel, M. (2008). The lytic cassette of mycobacteriophage Ms6 encodes an enzyme with lipolytic activity. Microbiology *154*, 1364-1371.

104. Gilbert, P., Collier, P.J., and Brown, M.R. (1990). Influence of growth rate on susceptibility to antimicrobial agents: biofilms, cell cycle, dormancy, and stringent response. Antimicrob Agents Chemother *34*, 1865-1868.

105. Gilham, D., and Lehner, R. (2005). Techniques to measure lipase and esterase activity in vitro. Methods *36*, 139-147.

106. Giudicelli, S., and Tomasz, A. (1984). Attachment of pneumococcal autolysin to wall teichoic acids, an essential step in enzymatic wall degradation. J Bacteriol *158*, 1188-1190.

107. Goodell, E.W. (1985). Recycling of murein by Escherichia coli. J Bacteriol *163*, 305-310.

108. Grandgirard, D., Loeffler, J.M., Fischetti, V.A., and Leib, S.L. (2008). Phage lytic enzyme Cpl-1 for antibacterial therapy in experimental pneumococcal meningitis. J Infect Dis *197*, 1519-1522.

109. Graschopf, A., and Blasi, U. (1999). Molecular function of the dual-start motif in the lambda S holin. Mol Microbiol *33*, 569-582.

110. Gronow, S., and Brade, H. (2001). Lipopolysaccharide biosynthesis: which steps do bacteria need to survive? J Endotoxin Res *7*, 3-23.

111. Grundling, A., Manson, M.D., and Young, R. (2001). Holins kill without warning. Proc Natl Acad Sci U S A *98*, 9348-9352.

112. Guindon, S., and Gascuel, O. (2003). A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol *52*, 696-704.

113. Gupta, R., Gupta, N., and Rathi, P. (2004). Bacterial lipases: an overview of production, purification and biochemical properties. Appl Microbiol Biotechnol *64*, 763-781.

114. Haidinger, W., Mayr, U.B., Szostak, M.P., Resch, S., and Lubitz, W. (2003). Escherichia coli ghost production by expression of lysis gene E and Staphylococcal nuclease. Appl Environ Microbiol *69*, 6106-6113.

115. Hall-Stoodley, L., and Stoodley, P. (2005). Biofilm formation and dispersal and the transmission of human pathogens. Trends Microbiol *13*, 7-10.

116. Hamid, M.E., Minnikin, D.E., and Goodfellow, M. (1993). A simple chemical test to distinguish mycobacteria from other mycolic-acid-containing actinomycetes. J Gen Microbiol *139*, 2203-2213.

117. Hanych, B., Kedzierska, S., Walderich, B., Uznanski, B., and Taylor, A. (1993). Expression of the Rz gene and the overlapping Rz1 reading frame present at the right end of the bacteriophage lambda genome. Gene *129*, 1-8.

118. Hatfull, G.F. (2010). Mycobacteriophages: Genes and Genomes. Annu Rev Microbiol.

119. Hatfull, G.F., Cresawn, S.G., and Hendrix, R.W. (2008). Comparative genomics of the mycobacteriophages: insights into bacteriophage evolution. Res Microbiol *159*, 332-339.

120. Hatfull, G.F., Jacobs-Sera, D., Lawrence, J.G., Pope, W.H., Russell, D.A., Ko, C.C., Weber, R.J., Patel, M.C., Germane, K.L., Edgar, R.H*., et al.* (2010). Comparative genomic analysis of 60 Mycobacteriophage genomes: genome clustering, gene acquisition, and gene size. J Mol Biol *397*, 119-143.

121. Hatfull, G.F., Pedulla, M.L., Jacobs-Sera, D., Cichon, P.M., Foley, A., Ford, M.E., Gonda, R.M., Houtz, J.M., Hryckowian, A.J., Kelchner, V.A*., et al.* (2006). Exploring the mycobacteriophage metaproteome: phage genomics as an educational platform. PLoS Genet *2*, e92.

122. Heidrich, C., Templin, M.F., Ursinus, A., Merdanovic, M., Berger, J., Schwarz, H., de Pedro, M.A., and Holtje, J.V. (2001). Involvement of N-acetylmuramyl-L-alanine amidases in cell separation and antibiotic-induced autolysis of Escherichia coli. Mol Microbiol *41*, 167-178.

123. Heinrich, P., Rosenstein, R., Bohmer, M., Sonner, P., and Gotz, F. (1987). The molecular organization of the lysostaphin gene and its sequences repeated in tandem. Mol Gen Genet *209*, 563-569.

124. Hendrix, R.W. (2003). Bacteriophage genomics. Curr Opin Microbiol *6*, 506-511.

125. Hendrix, R.W., Lawrence, J.G., Hatfull, G.F., and Casjens, S. (2000). The origins and ongoing evolution of viruses. Trends Microbiol *8*, 504-508.

126. Henrissat, B., Callebaut, I., Fabrega, S., Lehn, P., Mornon, J.P., and Davies, G. (1995). Conserved catalytic machinery and the prediction of a common fold for several families of glycosyl hydrolases. Proc Natl Acad Sci U S A *92*, 7090-7094.

127. Henrissat, B., and Davies, G. (1997). Structural and sequence-based classification of glycoside hydrolases. Curr Opin Struct Biol *7*, 637-644.

128. Henry, M., Begley, M., Neve, H., Maher, F., Ross, R.P., McAuliffe, O., Coffey, A., and O'Mahony, J.M. (2010). Cloning and expression of a mureinolytic enzyme from the mycobacteriophage TM4. FEMS Microbiol Lett.

129. Hermoso, J.A., Garcia, J.L., and Garcia, P. (2007). Taking aim on bacterial pathogens: from phage therapy to enzybiotics. Curr Opin Microbiol *10*, 461-472.

130. Hermoso, J.A., Monterroso, B., Albert, A., Galan, B., Ahrazem, O., Garcia, P., Martinez-Ripoll, M., Garcia, J.L., and Menendez, M. (2003). Structural basis for selective recognition of pneumococcal cell wall by modular endolysin from phage Cp-1. Structure *11*, 1239-1249.

131. Hoffmann, C., Leis, A., Niederweis, M., Plitzko, J.M., and Engelhardt, H. (2008). Disclosure of the mycobacterial outer membrane: cryo-electron tomography and vitreous sections reveal the lipid bilayer structure. Proc Natl Acad Sci U S A *105*, 3963-3967.

132. Holm, L., Kaariainen, S., Rosenstrom, P., and Schenkel, A. (2008). Searching protein structure databases with DaliLite v.3. Bioinformatics *24*, 2780.

133. Holm, L., and Sander, C. (1994). Structural similarity of plant chitinase and lysozymes from animals and phage. An evolutionary connection. FEBS Lett *340*, 129-132.

134. Holtje, J.V. (1995). From growth to autolysis: the murein hydrolases in Escherichia coli. Arch Microbiol *164*, 243-254.

135. Holtje, J.V. (1998). Growth of the stress-bearing and shape-maintaining murein sacculus of Escherichia coli. Microbiol Mol Biol Rev *62*, 181-203.

136. Holtje, J.V., and Fastrez, J. (1996). In Lysozymes: Model Enzymes in Biochemistry and Biology, J. P., ed. (Birkhäuser, Basel), pp. 35–64.

137. Holtje, J.V., Mirelman, D., Sharon, N., and Schwarz, U. (1975). Novel type of murein transglycosylase in Escherichia coli. J Bacteriol *124*, 1067-1076.

138. Hooper, N.M. (1994). Families of zinc metalloproteases. FEBS Lett *354*, 1-6.

139. Horgan, M., O'Flynn, G., Garry, J., Cooney, J., Coffey, A., Fitzgerald, G.F., Ross, R.P., and McAuliffe, O. (2009). Phage lysin LysK can be truncated to its CHAP domain and retain lytic activity against live antibiotic-resistant staphylococci. Appl Environ Microbiol *75*, 872-874.

140. Horsburgh, G.J., Atrih, A., and Foster, S.J. (2003). Characterization of LytH, a differentiation-associated peptidoglycan hydrolase of Bacillus subtilis involved in endospore cortex maturation. J Bacteriol *185*, 3813-3820.

141. Huet, J., Rucktooa, P., Clantin, B., Azarkan, M., Looze, Y., Villeret, V., and Wintjens, R. (2008). X-ray structure of papaya chitinase reveals the substrate binding mode of glycosyl hydrolase family 19 chitinases. Biochemistry *47*, 8283-8291.

142. Hunter, S., Apweiler, R., Attwood, T.K., Bairoch, A., Bateman, A., Binns, D., Bork, P., Das, U., Daugherty, L., Duquenne, L.*, et al.* (2009). InterPro: the integrative protein signature database. Nucleic Acids Res *37*, D211-215.

143. Ibrahim, H.R., Yamada, M., Matsushita, K., Kobayashi, K., and Kato, A. (1994). Enhanced bactericidal action of lysozyme to Escherichia coli by inserting a hydrophobic pentapeptide into its C terminus. J Biol Chem *269*, 5059-5063.

144. Ikeda, M., Wachi, M., Jung, H.K., Ishino, F., and Matsuhashi, M. (1991). The Escherichia coli mraY gene encoding UDP-N-acetylmuramoyl-pentapeptide: undecaprenyl-phosphate phospho-N-acetylmuramoyl-pentapeptide transferase. J Bacteriol *173*, 1021-1026.

145. Jacobs, C., Huang, L.J., Bartowsky, E., Normark, S., and Park, J.T. (1994). Bacterial cell wall recycling provides cytosolic muropeptides as effectors for beta-lactamase induction. EMBO J *13*, 4684-4694.

146. Jacobs, W.R. (2000). *Mycobacterium tuberculosis:* a Once Genetically Intractable Organism. In Molecular Genetics of *Mycobacteria*, G.F. Hatfull, and W.R. Jacobs, eds. (Washington DC, ASM Press), pp. 4-6.

147. Jacobs, W.R., Jr., Snapper, S.B., Lugosi, L., Jekkel, A., Melton, R.E., Kieser, T., and Bloom, B.R. (1989). Development of genetic systems for the mycobacteria. Acta Leprol *7 Suppl 1*, 203-207.

148. Jado, I., Lopez, R., Garcia, E., Fenoll, A., Casal, J., and Garcia, P. (2003). Phage lytic enzymes as therapy for antibiotic-resistant Streptococcus pneumoniae infection in a murine sepsis model. J Antimicrob Chemother *52*, 967-973.

149. Jain, V., and Mekalanos, J.J. (2000). Use of lambda phage S and R gene products in an inducible lysis system for Vibrio cholerae- and Salmonella enterica serovar typhimurium-based DNA vaccine delivery systems. Infect Immun *68*, 986-989.

150. Jalava, K., Eko, F.O., Riedmann, E., and Lubitz, W. (2003). Bacterial ghosts as carrier and targeting systems for mucosal antigen delivery. Expert Rev Vaccines *2*, 45-51.

151. Jardine, P., and Anderson, D. (2006). DNA Packaging in Double-stranded DNA Phages. In The Bacteriophages, R. Calendar, ed. (New York, Oxford University Press, Inc.), pp. 49-66.

152. Jolles, P., and Jolles, J. (1984). What's new in lysozyme research? Always a model system, today as yesterday. Mol Cell Biochem *63*, 165-189.

153. Jones, W.D., Jr. (1975). Phage typing report of 125 strains of "Mycobacterium tuberculosis". Ann Sclavo *17*, 599-604.

154. Kamisango, K., Saiki, I., Tanio, Y., Okumura, H., Araki, Y., Sekikawa, I., Azuma, I., and Yamamura, Y. (1982). Structures and biological activities of peptidoglycans of Listeria monocytogenes and Propionibacterium acnes. J Biochem *92*, 23-33.

155. Kana, B.D., and Mizrahi, V. (2010). Resuscitation-promoting factors as lytic enzymes for bacterial growth and signaling. FEMS Immunol Med Microbiol *58*, 39-50.

156. Karakousis, P.C., Bishai, W.R., and Dorman, S.E. (2004). Mycobacterium tuberculosis cell envelope lipids and the host immune response. Cell Microbiol *6*, 105-116.

157. Kawase, T., Saito, A., Sato, T., Kanai, R., Fujii, T., Nikaidou, N., Miyashita, K., and Watanabe, T. (2004). Distribution and phylogenetic analysis of family 19 chitinases in Actinobacteria. Appl Environ Microbiol *70*, 1135-1144.

158. Kedzierska, S., Wawrzynow, A., and Taylor, A. (1996). The Rz1 gene product of bacteriophage lambda is a lipoprotein localized in the outer membrane of Escherichia coli. Gene *168*, 1-8.

159. Keep, N.H., Ward, J.M., Cohen-Gonsaud, M., and Henderson, B. (2006). Wake up! Peptidoglycan lysis and bacterial non-growth states. Trends Microbiol *14*, 271-276.

160. Kell, D.B., and Young, M. (2000). Bacterial dormancy and culturability: the role of autocrine growth factors. Curr Opin Microbiol *3*, 238-243.

161. Kerff, F., Petrella, S., Mercier, F., Sauvage, E., Herman, R., Pennartz, A., Zervosen, A., Luxen, A., Frere, J.M., Joris, B.*, et al.* (2010). Specific structural features of the N-acetylmuramoyl-L-alanine amidase AmiD from Escherichia coli and mechanistic implications for enzymes of this family. J Mol Biol *397*, 249-259.

162. Kikkawa, H.S., Ueda, T., Suzuki, S., and Yasuda, J. (2008). Characterization of the catalytic activity of the gamma-phage lysin, PlyG, specific for Bacillus anthracis. FEMS Microbiol Lett *286*, 236-240.

163. Koch, A.L., and Doyle, R.J. (1985). Inside-to-outside growth and turnover of the wall of gram-positive rods. J Theor Biol *117*, 137-157.

164. Koch, A.L., and Woeste, S. (1992). Elasticity of the sacculus of Escherichia coli. J Bacteriol *174*, 4811-4819.

165. Konigsberg, W. (1967). In Methods in Enzymology, C.H.W. Hirs, ed. (New York, Academic Press,), p. 461.

166. Korndorfer, I.P., Danzer, J., Schmelcher, M., Zimmer, M., Skerra, A., and Loessner, M.J. (2006). The crystal structure of the bacteriophage PSA endolysin reveals a unique fold responsible for specific recognition of Listeria cell walls. J Mol Biol *364*, 678-689.

167. Krause, R.M. (1957). Studies on bacteriophages of hemolytic streptococci. I. Factors influencing the interaction of phage and susceptible host cell. J Exp Med *106*, 365-384.

168. Kremer, L., and Besra, G.S. (2005). A Waxy Tale, by *Mycobacterium tuberculosis*. In Tuberculosis and the Tubercle Bacillus, S.T. Cole, ed. (Washington. D.C.), pp. 287-305.

169. Kretzer, J.W., Lehmann, R., Schmelcher, M., Banz, M., Kim, K.P., Korn, C., and Loessner, M.J. (2007). Use of high-affinity cell wall-binding domains of bacteriophage endolysins for immobilization and separation of bacterial cells. Appl Environ Microbiol *73*, 1992-2000.

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

171. Krupovic, M., Cvirkaite-Krupovic, V., and Bamford, D.H. (2008). Identification and functional analysis of the Rz/Rz1-like accessory lysis genes in the membrane-containing bacteriophage PRD1. Mol Microbiol *68*, 492-503.

172. Labischinski, H., Goodell, E.W., Goodell, A., and Hochberg, M.L. (1991). Direct proof of a "more-than-single-layered" peptidoglycan architecture of Escherichia coli W7: a neutron small-angle scattering study. J Bacteriol *173*, 751-756.

173. Lawrence, J.G., Hatfull, G.F., and Hendrix, R.W. (2002). Imbroglios of viral taxonomy: genetic exchange and failings of phenetic approaches. J Bacteriol *184*, 4891-4905.

174. Leclerc, D., and Asselin, A. (1989). Detection of bacterial cell wall hydrolases after denaturing polyacrylamide gel electrophoresis. Can J Microbiol *35*, 749-753.

175. Lee, S., Kriakov, J., Vilcheze, C., Dai, Z., Hatfull, G.F., and Jacobs, W.R., Jr. (2004). Bxz1, a new generalized transducing phage for mycobacteria. FEMS Microbiol Lett *241*, 271-276.

176. Lehnherr, H., Hansen, A.M., and Ilyina, T. (1998). Penetration of the bacterial cell wall: a family of lytic transglycosylases in bacteriophages and conjugative plasmids. Mol Microbiol *30*, 454-457.

177. Levin, B.R., and Bull, J.J. (2004). Population and evolutionary dynamics of phage therapy. Nat Rev Microbiol *2*, 166-173.

178. Little, J. (2006). Gene Regulatory Circuitry of Phage Lambda. In The Bacteriophages, R. Calendar, ed. (New York, Oxford University Press, Inc.), pp. 74-82.

179. Loeffler, J.M., Djurkovic, S., and Fischetti, V.A. (2003). Phage lytic enzyme Cpl-1 as a novel antimicrobial for pneumococcal bacteremia. Infect Immun *71*, 6199-6204.

180. Loeffler, J.M., and Fischetti, V.A. (2003). Synergistic lethal effect of a combination of phage lytic enzymes with different activities on penicillin-sensitive and -resistant Streptococcus pneumoniae strains. Antimicrob Agents Chemother *47*, 375-377.

181. Loeffler, J.M., Nelson, D., and Fischetti, V.A. (2001). Rapid killing of Streptococcus pneumoniae with a bacteriophage cell wall hydrolase. Science *294*, 2170-2172.

182. Loessner, M.J. (2005). Bacteriophage endolysins--current state of research and applications. Curr Opin Microbiol *8*, 480-487.

183. Loessner, M.J., Gaeng, S., and Scherer, S. (1999). Evidence for a holin-like protein gene fully embedded out of frame in the endolysin gene of Staphylococcus aureus bacteriophage 187. J Bacteriol *181*, 4452-4460.

184. Loessner, M.J., Kramer, K., Ebel, F., and Scherer, S. (2002). C-terminal domains of Listeria monocytogenes bacteriophage murein hydrolases determine specific recognition and high-affinity binding to bacterial cell wall carbohydrates. Mol Microbiol *44*, 335-349.

185. Loessner, M.J., Maier, S.K., Daubek-Puza, H., Wendlinger, G., and Scherer, S. (1997). Three Bacillus cereus bacteriophage endolysins are unrelated but reveal high homology to cell wall hydrolases from different bacilli. J Bacteriol *179*, 2845-2851.

186. Loessner, M.J., Wendlinger, G., and Scherer, S. (1995). Heterogeneous endolysins in Listeria monocytogenes bacteriophages: a new class of enzymes and evidence for conserved holin genes within the siphoviral lysis cassettes. Mol Microbiol *16*, 1231-1241.

187. Longhi, S., and Cambillau, C. (1999). Structure-activity of cutinase, a small lipolytic enzyme. Biochim Biophys Acta *1441*, 185-196.

188. Lopez, R., and Garcia, E. (2004). Recent trends on the molecular biology of pneumococcal capsules, lytic enzymes, and bacteriophage. FEMS Microbiol Rev *28*, 553-580.

189. Lopez, R., Garcia, E., Garcia, P., and Garcia, J.L. (1997). The pneumococcal cell wall degrading enzymes: a modular design to create new lysins? Microb Drug Resist *3*, 199-211.

190. Low, L.Y., Yang, C., Perego, M., Osterman, A., and Liddington, R.C. (2005). Structure and lytic activity of a Bacillus anthracis prophage endolysin. J Biol Chem *280*, 35433-35439.

191. Lowy, F.D. (2003). Antimicrobial resistance: the example of Staphylococcus aureus. J Clin Invest *111*, 1265-1273.

192. Lu, J.Z., Fujiwara, T., Komatsuzawa, H., Sugai, M., and Sakon, J. (2006). Cell wall-targeting domain of glycylglycine endopeptidase distinguishes among peptidoglycan cross-bridges. J Biol Chem *281*, 549-558.

193. Mahapatra, S., Basu, J., Brennan, P.J., and Crick, D.C. (2005a). Structure, Biosynthesis, and Genetics of the Mycolic Acid-Arabinogalactan-Peptidoglycan Complex (Washington, DC, ASM Press).

194. Mahapatra, S., Basu, J., Brennan, P.J., and Crick, D.C. (2005b). Structure, Biosynthesis, and Genetics of the Mycolic Acid-Arabinogalactan-Peptidoglycan Complex. In Tuberculosis and the Tubercle Bacillus, S.T. Cole, ed. (Washington. D.C.), pp. 275-284.

195. Mahapatra, S., Yagi, T., Belisle, J.T., Espinosa, B.J., Hill, P.J., McNeil, M.R., Brennan, P.J., and Crick, D.C. (2005c). Mycobacterial lipid II is composed of a complex mixture of modified muramyl and peptide moieties linked to decaprenyl phosphate. J Bacteriol *187*, 2747-2757.

196. Marinelli, L.J. (2008). Characterization of the role of the Rpf motif in mycobacteriophage tape measure proteins. In Department of Biological Sciences (Pittsburgh, University of Pittsburgh).

197. Marinelli, L.J., Piuri, M., Swigonova, Z., Balachandran, A., Oldfield, L.M., van Kessel, J.C., and Hatfull, G.F. (2008). BRED: a simple and powerful tool for constructing mutant and recombinant bacteriophage genomes. PLoS One *3*, e3957.

198. Markov, D., Christie, G.E., Sauer, B., Calendar, R., Park, T., Young, R., and Severinov, K. (2004). P2 growth restriction on an rpoC mutant is suppressed by alleles of the Rz1 homolog lysC. J Bacteriol *186*, 4628-4637.

199. Masaki, K., Kamini, N.R., Ikeda, H., and Iefuji, H. (2005). Cutinase-like enzyme from the yeast *Cryptococcus* sp. strain S-2 hydrolyzes polylactic acid and other biodegradable plastics. Appl Environ Microbiol *71*, 7548-7550.

200. Matias, V.R., Al-Amoudi, A., Dubochet, J., and Beveridge, T.J. (2003). Cryo-transmission electron microscopy of frozen-hydrated sections of Escherichia coli and Pseudomonas aeruginosa. J Bacteriol *185*, 6112-6118.

201. Matias, V.R., and Beveridge, T.J. (2005). Cryo-electron microscopy reveals native polymeric cell wall structure in Bacillus subtilis 168 and the existence of a periplasmic space. Mol Microbiol *56*, 240-251.

202. Matias, V.R., and Beveridge, T.J. (2006). Native cell wall organization shown by cryo-electron microscopy confirms the existence of a periplasmic space in Staphylococcus aureus. J Bacteriol *188*, 1011-1021.

203. McCullers, J.A., Karlstrom, A., Iverson, A.R., Loeffler, J.M., and Fischetti, V.A. (2007). Novel strategy to prevent otitis media caused by colonizing Streptococcus pneumoniae. PLoS Pathog *3*, e28.

204. McLauchlin, J., Mitchell, R.T., Smerdon, W.J., and Jewell, K. (2004). Listeria monocytogenes and listeriosis: a review of hazard characterisation for use in microbiological risk assessment of foods. Int J Food Microbiol *92*, 15-33.

205. McNeil, M., Daffe, M., and Brennan, P.J. (1990). Evidence for the nature of the link between the arabinogalactan and peptidoglycan of mycobacterial cell walls. J Biol Chem *265*, 18200-18206.

206. McNeil, M., Daffe, M., and Brennan, P.J. (1991). Location of the mycolyl ester substituents in the cell walls of mycobacteria. J Biol Chem *266*, 13217-13223.

207. McNeil, M.R., and Brennan, P.J. (1991). Structure, function and biogenesis of the cell envelope of mycobacteria in relation to bacterial physiology, pathogenesis and drug resistance; some thoughts and possibilities arising from recent structural information. Res Microbiol *142*, 451-463.

208. McNerney, R., Wilson, S.M., Sidhu, A.M., Harley, V.S., al Suwaidi, Z., Nye, P.M., Parish, T., and Stoker, N.G. (1998). Inactivation of mycobacteriophage D29 using ferrous ammonium sulphate as a tool for the detection of viable Mycobacterium smegmatis and M. tuberculosis. Res Microbiol *149*, 487-495.

209. Mediavilla, J., Jain, S., Kriakov, J., Ford, M.E., Duda, R.L., Jacobs, W.R., Jr., Hendrix, R.W., and Hatfull, G.F. (2000). Genome organization and characterization of mycobacteriophage Bxb1. Mol Microbiol *38*, 955-970.

210. Minnikin, D.E., Minnikin, S.M., Goodfellow, M., and Stanford, J.L. (1982). The mycolic acids of Mycobacterium chelonei. J Gen Microbiol *128*, 817-822.

211. Moak, M., and Molineux, I.J. (2000). Role of the Gp16 lytic transglycosylase motif in bacteriophage T7 virions at the initiation of infection. Mol Microbiol *37*, 345-355.

212. Moak, M., and Molineux, I.J. (2004). Peptidoglycan hydrolytic activities associated with bacteriophage virions. Mol Microbiol *51*, 1169-1183.

213. Moffatt, B.A., and Studier, F.W. (1987). T7 lysozyme inhibits transcription by T7 RNA polymerase. Cell *49*, 221-227.

214. Monzingo, A.F., Marcotte, E.M., Hart, P.J., and Robertus, J.D. (1996). Chitinases, chitosanases, and lysozymes can be divided into procaryotic and eucaryotic families sharing a conserved core. Nat Struct Biol *3*, 133-140.

215. Morita, M., Tanji, Y., Mizoguchi, K., Soejima, A., Orito, Y., and Unno, H. (2001a). Antibacterial activity of Bacillus amyloliquefaciens phage endolysin without holin conjugation. J Biosci Bioeng *91*, 469-473.

216. Morita, M., Tanji, Y., Orito, Y., Mizoguchi, K., Soejima, A., and Unno, H. (2001b). Functional analysis of antibacterial activity of Bacillus amyloliquefaciens phage endolysin against Gram-negative bacteria. FEBS Lett *500*, 56-59.

217. Moulder, J.W. (1993). Why is Chlamydia sensitive to penicillin in the absence of peptidoglycan? Infect Agents Dis *2*, 87-99.

218. Nau, R., and Eiffert, H. (2002). Modulation of release of proinflammatory bacterial compounds by antibacterials: potential impact on course of inflammation and outcome in sepsis and meningitis. Clin Microbiol Rev *15*, 95-110.

219. Navarre, W.W., Ton-That, H., Faull, K.F., and Schneewind, O. (1999). Multiple enzymatic activities of the murein hydrolase from staphylococcal phage phi11. Identification of a D-alanyl-glycine endopeptidase activity. J Biol Chem *274*, 15847-15856.

220. Nelson, D., Loomis, L., and Fischetti, V.A. (2001). Prevention and elimination of upper respiratory colonization of mice by group A streptococci by using a bacteriophage lytic enzyme. Proc Natl Acad Sci U S A *98*, 4107-4112.

221. Nelson, D., Schuch, R., Chahales, P., Zhu, S., and Fischetti, V.A. (2006). PlyC: a multimeric bacteriophage lysin. Proc Natl Acad Sci U S A *103*, 10765-10770.

222. Nishihara, T., Morisawa, H., Ohta, N., Atkins, J.F., and Nishimura, Y. (2004). A cryptic lysis gene near the start of the Qbeta replicase gene in the +1 frame. Genes Cells *9*, 877-889.

223. O'Flaherty, S., Coffey, A., Meaney, W., Fitzgerald, G.F., and Ross, R.P. (2005). The recombinant phage lysin LysK has a broad spectrum of lytic activity against clinically relevant staphylococci, including methicillin-resistant Staphylococcus aureus. J Bacteriol *187*, 7161-7164.

224. O'Flaherty, S., Ross, R.P., and Coffey, A. (2009). Bacteriophage and their lysins for elimination of infectious bacteria. FEMS Microbiol Rev *33*, 801-819.

225. Obeso, J.M., Martinez, B., Rodriguez, A., and Garcia, P. (2008). Lytic activity of the recombinant staphylococcal bacteriophage PhiH5 endolysin active against Staphylococcus aureus in milk. Int J Food Microbiol *128*, 212-218.

226. Odintsov, S.G., Sabala, I., Marcyjaniak, M., and Bochtler, M. (2004). Latent LytM at 1.3A resolution. J Mol Biol *335*, 775-785.

227. Ojha, A., Anand, M., Bhatt, A., Kremer, L., Jacobs, W.R., Jr., and Hatfull, G.F. (2005). GroEL1: a dedicated chaperone involved in mycolic acid biosynthesis during biofilm formation in mycobacteria. Cell *123*, 861-873.

228. Ojha, A.K., Baughn, A.D., Sambandan, D., Hsu, T., Trivelli, X., Guerardel, Y., Alahari, A., Kremer, L., Jacobs, W.R., Jr., and Hatfull, G.F. (2008). Growth of Mycobacterium tuberculosis biofilms containing free mycolic acids and harbouring drug-tolerant bacteria. Mol Microbiol *69*, 164-174.

229. Ojha, A.K., Trivelli, X., Guerardel, Y., Kremer, L., and Hatfull, G.F. (2010). Enzymatic hydrolysis of trehalose dimycolate releases free mycolic acids during mycobacterial growth in biofilms. J Biol Chem *285*, 17380-17389.

230. Ollis, D.L., Cheah, E., Cygler, M., Dijkstra, B., Frolow, F., Franken, S.M., Harel, M., Remington, S.J., Silman, I., Schrag, J.*, et al.* (1992). The alpha/beta hydrolase fold. Protein Eng *5*, 197-211.

231. Ong, C.T., Kuti, J.L., Nightingale, C.H., and Nicolau, D.P. (2004). Emerging Pseudomonas aeruginosa resistance: implications in clinical practice. Conn Med *68*, 11-15.

232. Oppenheim, A.B., Kobiler, O., Stavans, J., Court, D.L., and Adhya, S. (2005). Switches in bacteriophage lambda development. Annu Rev Genet *39*, 409-429.

233. Orito, Y., Morita, M., Hori, K., Unno, H., and Tanji, Y. (2004). Bacillus amyloliquefaciens phage endolysin can enhance permeability of Pseudomonas aeruginosa outer membrane and induce cell lysis. Appl Microbiol Biotechnol *65*, 105-109.

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

235. Oshida, T., Sugai, M., Komatsuzawa, H., Hong, Y.M., Suginaka, H., and Tomasz, A. (1995). A Staphylococcus aureus autolysin that has an N-acetylmuramoyl-L-alanine amidase domain and an endo-beta-N-acetylglucosaminidase domain: cloning, sequence analysis, and characterization. Proc Natl Acad Sci U S A *92*, 285-289.

236. Pandey, A.K., Raman, S., Proff, R., Joshi, S., Kang, C.M., Rubin, E.J., Husson, R.N., and Sassetti, C.M. (2009). Nitrile-inducible gene expression in mycobacteria. Tuberculosis (Edinb) *89*, 12-16.

237. Pang, T., Savva, C.G., Fleming, K.G., Struck, D.K., and Young, R. (2009). Structure of the lethal phage pinhole. Proc Natl Acad Sci U S A *106*, 18966-18971.

238. Panthel, K., Jechlinger, W., Matis, A., Rohde, M., Szostak, M., Lubitz, W., and Haas, R. (2003). Generation of Helicobacter pylori ghosts by PhiX protein E-mediated inactivation and their evaluation as vaccine candidates. Infect Immun *71*, 109-116.

239. Parish, T., and Stoker, N.G. (1998). Mycobacteria protocols (Totowa, N.J., Humana Press).

240. Park, P.W., Senior, R.M., Griffin, G.L., Broekelmann, T.J., Mudd, M.S., and Mecham, R.P. (1995). Binding and degradation of elastin by the staphylolytic enzyme lysostaphin. Int J Biochem Cell Biol *27*, 139-146.

241. Park, T., Struck, D.K., Deaton, J.F., and Young, R. (2006). Topological dynamics of holins in programmed bacterial lysis. Proc Natl Acad Sci U S A *103*, 19713-19718.

242. Payne, K., Sun, Q., Sacchettini, J., and Hatfull, G.F. (2009). Mycobacteriophage Lysin B is a novel mycolylarabinogalactan esterase. Mol Microbiol *73*, 367-381.

243. Pedulla, M.L., Ford, M.E., Houtz, J.M., Karthikeyan, T., Wadsworth, C., Lewis, J.A., Jacobs-Sera, D., Falbo, J., Gross, J., Pannunzio, N.R.*, et al.* (2003). Origins of highly mosaic mycobacteriophage genomes. Cell *113*, 171-182.

244. Perez-Dorado, I., Campillo, N.E., Monterroso, B., Hesek, D., Lee, M., Paez, J.A., Garcia, P., Martinez-Ripoll, M., Garcia, J.L., Mobashery, S.*, et al.* (2007). Elucidation of the molecular recognition of bacterial cell wall by modular pneumococcal phage endolysin CPL-1. J Biol Chem *282*, 24990-24999.

245. Perriere, G., and Gouy, M. (1996). WWW-query: an on-line retrieval system for biological sequence banks. Biochimie *78*, 364-369.

246. Petit, J.F., and Lederer, E. (2000). The Structure of the Mycobacterial Cell Wall. In Molecular Genetics of Mycobacteria, G.F. Hatfull, and W.R. Jacobs, eds. (Washington DC, ASM Press), pp. 301-313.

247. Pilgrim, S., Stritzker, J., Schoen, C., Kolb-Maurer, A., Geginat, G., Loessner, M.J., Gentschev, I., and Goebel, W. (2003). Bactofection of mammalian cells by Listeria monocytogenes: improvement and mechanism of DNA delivery. Gene Ther *10*, 2036-2045.

248. Pisabarro, A.G., de Pedro, M.A., and Vazquez, D. (1985). Structural modifications in the peptidoglycan of Escherichia coli associated with changes in the state of growth of the culture. J Bacteriol *161*, 238-242.

249. Piuri, M., and Hatfull, G.F. (2006). A peptidoglycan hydrolase motif within the mycobacteriophage TM4 tape measure protein promotes efficient infection of stationary phase cells. Mol Microbiol *62*, 1569-1585.

250. Piuri, M., Jacobs, W.R., Jr., and Hatfull, G.F. (2009). Fluoromycobacteriophages for rapid, specific, and sensitive antibiotic susceptibility testing of Mycobacterium tuberculosis. PLoS One *4*, e4870.

251. Porter, C.J., Schuch, R., Pelzek, A.J., Buckle, A.M., McGowan, S., Wilce, M.C., Rossjohn, J., Russell, R., Nelson, D., Fischetti, V.A.*, et al.* (2007). The 1.6 A crystal structure of the catalytic domain of PlyB, a bacteriophage lysin active against Bacillus anthracis. J Mol Biol *366*, 540-550.

252. Portevin, D., De Sousa-D'Auria, C., Houssin, C., Grimaldi, C., Chami, M., Daffe, M., and Guilhot, C. (2004). A polyketide synthase catalyzes the last condensation step of mycolic acid biosynthesis in mycobacteria and related organisms. Proc Natl Acad Sci U S A *101*, 314-319.

253. Portevin, D., de Sousa-D'Auria, C., Montrozier, H., Houssin, C., Stella, A., Laneelle, M.A., Bardou, F., Guilhot, C., and Daffe, M. (2005). The acyl-AMP ligase FadD32 and AccD4-containing acyl-CoA carboxylase are required for the synthesis of mycolic acids and essential for mycobacterial growth: identification of the carboxylation product and determination of the acyl-CoA carboxylase components. J Biol Chem *280*, 8862-8874.

254. Pritchard, D.G., Dong, S., Baker, J.R., and Engler, J.A. (2004). The bifunctional peptidoglycan lysin of Streptococcus agalactiae bacteriophage B30. Microbiology *150*, 2079-2087.

255. Pritchard, D.G., Dong, S., Kirk, M.C., Cartee, R.T., and Baker, J.R. (2007). LambdaSa1 and LambdaSa2 prophage lysins of Streptococcus agalactiae. Appl Environ Microbiol *73*, 7150-7154.

256. Puech, V., Bayan, N., Salim, K., Leblon, G., and Daffe, M. (2000). Characterization of the in vivo acceptors of the mycoloyl residues transferred by the corynebacterial PS1 and the related mycobacterial antigens 85. Mol Microbiol *35*, 1026-1041.

257. Quemard, A., Sacchettini, J.C., Dessen, A., Vilcheze, C., Bittman, R., Jacobs, W.R., Jr., and Blanchard, J.S. (1995). Enzymatic characterization of the target for isoniazid in Mycobacterium tuberculosis. Biochemistry *34*, 8235-8241.

258. Quintela, J.C., Caparros, M., and de Pedro, M.A. (1995). Variability of peptidoglycan structural parameters in gram-negative bacteria. FEMS Microbiol Lett *125*, 95-100.

259. Raab, R., Neal, G., Garrett, J., Grimaila, R., Fusselman, R., and Young, R. (1986). Mutational analysis of bacteriophage lambda lysis gene S. J Bacteriol *167*, 1035-1042.

260. Raab, R., Neal, G., Sohaskey, C., Smith, J., and Young, R. (1988). Dominance in lambda S mutations and evidence for translational control. J Mol Biol *199*, 95-105.

261. Raetz, C.R., and Whitfield, C. (2002). Lipopolysaccharide endotoxins. Annu Rev Biochem *71*, 635-700.

262. Ramanculov, E., and Young, R. (2001). An ancient player unmasked: T4 rI encodes a t-specific antiholin. Mol Microbiol *41*, 575-583.

263. Rashel, M., Uchiyama, J., Ujihara, T., Uehara, Y., Kuramoto, S., Sugihara, S., Yagyu, K., Muraoka, A., Sugai, M., Hiramatsu, K.*, et al.* (2007). Efficient elimination of multidrug-resistant Staphylococcus aureus by cloned lysin derived from bacteriophage phi MR11. J Infect Dis *196*, 1237-1247.

264. Rau, A., Hogg, T., Marquardt, R., and Hilgenfeld, R. (2001). A new lysozyme fold. Crystal structure of the muramidase from Streptomyces coelicolor at 1.65 A resolution. J Biol Chem *276*, 31994-31999.

265. Rawlings, N.D., and Barrett, A.J. (1993). Evolutionary families of peptidases. Biochem J *290 (Pt 1)*, 205-218.

266. Rawlings, N.D., Barrett, A.J., and Bateman, A. (2010). MEROPS: the peptidase database. Nucleic Acids Res *38*, D227-233.

267. Raymond, J.B., Mahapatra, S., Crick, D.C., and Pavelka, M.S., Jr. (2005). Identification of the namH gene, encoding the hydroxylase responsible for the N-glycolylation of the mycobacterial peptidoglycan. J Biol Chem *280*, 326-333.

268. Razavi, B., Apisarnthanarak, A., and Mundy, L.M. (2007). Clostridium difficile: emergence of hypervirulence and fluoroquinolone resistance. Infection *35*, 300-307.

269. Russel, M. (1995). Moving through the membrane with filamentous phages. Trends Microbiol *3*, 223-228.

270. Russell, A.D. (2001). Mechanisms of bacterial insusceptibility to biocides. Am J Infect Control *29*, 259-261.

271. Rybniker, J., Kramme, S., and Small, P.L. (2006). Host range of 14 mycobacteriophages in Mycobacterium ulcerans and seven other mycobacteria including Mycobacterium tuberculosis--application for identification and susceptibility testing. J Med Microbiol *55*, 37-42.

272. Rybniker, J., Plum, G., Robinson, N., Small, P.L., and Hartmann, P. (2008). Identification of three cytotoxic early proteins of mycobacteriophage L5 leading to growth inhibition in Mycobacterium smegmatis. Microbiology *154*, 2304-2314.

273. Saenz, Y., Brinas, L., Dominguez, E., Ruiz, J., Zarazaga, M., Vila, J., and Torres, C. (2004). Mechanisms of resistance in multiple-antibiotic-resistant Escherichia coli strains of human, animal, and food origins. Antimicrob Agents Chemother *48*, 3996-4001.

274. Sambrook, J., and Maniatis, T. (1989). **Molecular cloning: a laboratory manual, Volume 1** (Cold Springs Harbor Laboratory).

275. Sanger, F., Coulson, A.R., Hong, G.F., Hill, D.F., and Petersen, G.B. (1982). Nucleotide sequence of bacteriophage lambda DNA. J Mol Biol *162*, 729-773.

276. Sanz, J.M., Garcia, P., and Garcia, J.L. (1996). Construction of a multifunctional pneumococcal murein hydrolase by module assembly. Eur J Biochem *235*, 601-605.

277. Sao-Jose, C., Parreira, R., Vieira, G., and Santos, M.A. (2000). The N-terminal region of the Oenococcus oeni bacteriophage fOg44 lysin behaves as a bona fide signal peptide in Escherichia coli and as a cis-inhibitory element, preventing lytic activity on oenococcal cells. J Bacteriol *182*, 5823-5831.

278. Sarkis, G.J., and Hatfull, G.F. (1998). Mycobacteriophages. Methods Mol Biol *101*, 145-173.

279. Saunders, B.M., and Cooper, A.M. (2000). Restraining mycobacteria: role of granulomas in mycobacterial infections. Immunol Cell Biol *78*, 334-341.

280. Savva, C.G., Dewey, J.S., Deaton, J., White, R.L., Struck, D.K., Holzenburg, A., and Young, R. (2008). The holin of bacteriophage lambda forms rings with large diameter. Mol Microbiol *69*, 784-793.

281. Schafer, R., Huber, U., and Franklin, R.M. (1977). Chemical and physical properties of mycobacteriophage D29. Eur J Biochem *73*, 239-246.

282. Schaffer, A.A., Aravind, L., Madden, T.L., Shavirin, S., Spouge, J.L., Wolf, Y.I., Koonin, E.V., and Altschul, S.F. (2001). Improving the accuracy of PSI-BLAST protein database searches with composition-based statistics and other refinements. Nucleic Acids Res *29*, 2994-3005.

283. Schaffer, C., and Messner, P. (2005). The structure of secondary cell wall polymers: how Gram-positive bacteria stick their cell walls together. Microbiology *151*, 643-651.

284. Scheurwater, E., Reid, C.W., and Clarke, A.J. (2008). Lytic transglycosylases: bacterial space-making autolysins. Int J Biochem Cell Biol *40*, 586-591.

285. Schleifer, K.H., and Kandler, O. (1972). Peptidoglycan types of bacterial cell walls and their taxonomic implications. Bacteriol Rev *36*, 407-477.

286. Schmidt, C., Velleman, M., and Arber, W. (1996). Three functions of bacteriophage P1 involved in cell lysis. J Bacteriol *178*, 1099-1104.

287. Schuch, R., Nelson, D., and Fischetti, V.A. (2002). A bacteriolytic agent that detects and kills Bacillus anthracis. Nature *418*, 884-889.

288. Schue, M., Maurin, D., Dhouib, R., N'Goma, J.C., Delorme, V., Lambeau, G., Carriere, F., and Canaan, S. (2010). Two cutinase-like proteins secreted by Mycobacterium tuberculosis show very different lipolytic activities reflecting their physiological function. FASEB J *24*, 1893-1903.

289. Sellers, M.I., Baxter, W.L., and Runnals, H.R. (1962). Growth characteristics of mycobacteriophages D28 and D29. Can J Microbiol *8*, 389-399.

290. Sheehan, J.K., Kesimer, M., and Pickles, R. (2006). Innate immunity and mucus structure and function. Novartis Found Symp *279*, 155-166; discussion 167-159, 216-159.

291. Sheehan, M.M., Garcia, J.L., Lopez, R., and Garcia, P. (1997). The lytic enzyme of the pneumococcal phage Dp-1: a chimeric lysin of intergeneric origin. Mol Microbiol *25*, 717-725.

292. Shimodaira, H., and Hasegawa, M. (1999). Multiple Comparisons of Log-Likelihoods with Applications to Phylogenetic Inference. Mol Biol Evol *16*, 1114-1116.

293. Shockman, G.D., and Barrett, J.F. (1983). Structure, function, and assembly of cell walls of gram-positive bacteria. Annu Rev Microbiol *37*, 501-527.

294. Skeiky, Y.A., and Sadoff, J.C. (2006). Advances in tuberculosis vaccine strategies. Nat Rev Microbiol *4*, 469-476.

295. Smith, T.J., Blackman, S.A., and Foster, S.J. (2000). Autolysins of Bacillus subtilis: multiple enzymes with multiple functions. Microbiology *146 (Pt 2)*, 249-262.

296. Snapper, S.B., Melton, R.E., Mustafa, S., Kieser, T., and Jacobs, W.R., Jr. (1990). Isolation and characterization of efficient plasmid transformation mutants of Mycobacterium smegmatis. Mol Microbiol *4*, 1911-1919.

297. Snider, D.E., Jr., Jones, W.D., and Good, R.C. (1984). The usefulness of phage typing Mycobacterium tuberculosis isolates. Am Rev Respir Dis *130*, 1095-1099.

298. Solensky, R. (2003). Hypersensitivity reactions to beta-lactam antibiotics. Clin Rev Allergy Immunol *24*, 201-220.

299. Spencer, J., Murphy, L.M., Conners, R., Sessions, R.B., and Gamblin, S.J. (2010). Crystal structure of the LasA virulence factor from Pseudomonas aeruginosa: substrate specificity and mechanism of M23 metallopeptidases. J Mol Biol *396*, 908-923.

300. Stahl, C., Kubetzko, S., Kaps, I., Seeber, S., Engelhardt, H., and Niederweis, M. (2001). MspA provides the main hydrophilic pathway through the cell wall of Mycobacterium smegmatis. Mol Microbiol *40*, 451-464.

301. Summer, E.J., Berry, J., Tran, T.A., Niu, L., Struck, D.K., and Young, R. (2007). Rz/Rz1 lysis gene equivalents in phages of Gram-negative hosts. J Mol Biol *373*, 1098-1112.

302. Summers, W.C. (2001). Bacteriophage therapy. Annu Rev Microbiol *55*, 437-451.

303. Suttle, C.A. (2005). Viruses in the sea. Nature *437*, 356-361.

304. Tamura, A., Ohashi, N., Urakami, H., and Miyamura, S. (1995). Classification of Rickettsia tsutsugamushi in a new genus, Orientia gen. nov., as Orientia tsutsugamushi comb. nov. Int J Syst Bacteriol *45*, 589-591.

305. Telenti, A., and Iseman, M. (2000). Drug-resistant tuberculosis: what do we do now? Drugs *59*, 171-179.

306. Thompson, J.D., Higgins, D.G., and Gibson, T.J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res *22*, 4673-4680.

307. Thunnissen, A.M., Dijkstra, A.J., Kalk, K.H., Rozeboom, H.J., Engel, H., Keck, W., and Dijkstra, B.W. (1994). Doughnut-shaped structure of a bacterial muramidase revealed by X-ray crystallography. Nature *367*, 750-753.

308. Thunnissen, A.M., Isaacs, N.W., and Dijkstra, B.W. (1995). The catalytic domain of a bacterial lytic transglycosylase defines a novel class of lysozymes. Proteins *22*, 245-258.

309. Tripathi, R.P., Tewari, N., Dwivedi, N., and Tiwari, V.K. (2005). Fighting tuberculosis: an old disease with new challenges. Med Res Rev *25*, 93-131.

310. Tuomanen, E., Cozens, R., Tosch, W., Zak, O., and Tomasz, A. (1986). The rate of killing of Escherichia coli by beta-lactam antibiotics is strictly proportional to the rate of bacterial growth. J Gen Microbiol *132*, 1297-1304.

311. Turner, M.S., Waldherr, F., Loessner, M.J., and Giffard, P.M. (2007). Antimicrobial activity of lysostaphin and a Listeria monocytogenes bacteriophage endolysin produced and secreted by lactic acid bacteria. Syst Appl Microbiol *30*, 58-67.

312. Udaya Prakash, N.A., Jayanthi, M., Sabarinathan, R., Kangueane, P., Mathew, L., and Sekar, K. (2010). Evolution, homology conservation, and identification of unique sequence signatures in GH19 family chitinases. J Mol Evol *70*, 466-478.

313. Usobiaga, P., Medrano, F.J., Gasset, M., Garcia, J.L., Saiz, J.L., Rivas, G., Laynez, J., and Menendez, M. (1996). Structural organization of the major autolysin from Streptococcus pneumoniae. J Biol Chem *271*, 6832-6838.

314. van Kessel, J.C., and Hatfull, G.F. (2007). Recombineering in Mycobacterium tuberculosis. Nat Methods *4*, 147-152.

315. Vollmer, W. (2008). Structural variation in the glycan strands of bacterial peptidoglycan. FEMS Microbiol Rev *32*, 287-306.

316. Vollmer, W., Blanot, D., and de Pedro, M.A. (2008a). Peptidoglycan structure and architecture. FEMS Microbiol Rev *32*, 149-167.

317. Vollmer, W., and Holtje, J.V. (2001). Morphogenesis of Escherichia coli. Curr Opin Microbiol *4*, 625-633.

318. Vollmer, W., Joris, B., Charlier, P., and Foster, S. (2008b). Bacterial peptidoglycan (murein) hydrolases. FEMS Microbiol Rev *32*, 259-286.

319. von Eiff, C., Becker, K., Machka, K., Stammer, H., and Peters, G. (2001). Nasal carriage as a source of Staphylococcus aureus bacteremia. Study Group. N Engl J Med *344*, 11-16.

320. W.H.O. (2009). Global Tuberculosis Control, W.H. Organization, ed. (Geneva, Switzerland, WHO Press).

321. W.H.O. (2010). Tuberculosis Fact Sheet, W.H. Organization, ed. (Geneva, Switzerland, WHO Press).

322. Wall, R.J., Powell, A.M., Paape, M.J., Kerr, D.E., Bannerman, D.D., Pursel, V.G., Wells, K.D., Talbot, N., and Hawk, H.W. (2005). Genetically enhanced cows resist intramammary Staphylococcus aureus infection. Nat Biotechnol *23*, 445-451.

323. Walsh, S., Shah, A., and Mond, J. (2003). Improved pharmacokinetics and reduced antibody reactivity of lysostaphin conjugated to polyethylene glycol. Antimicrob Agents Chemother *47*, 554-558.

324. Wang, I.N. (2006). Lysis timing and bacteriophage fitness. Genetics *172*, 17-26.

325. Wang, I.N., Deaton, J., and Young, R. (2003). Sizing the holin lesion with an endolysin-beta-galactosidase fusion. J Bacteriol *185*, 779-787.

326. Wang, I.N., Smith, D.L., and Young, R. (2000). Holins: the protein clocks of bacteriophage infections. Annu Rev Microbiol *54*, 799-825.

327. Ward, J.B. (1973). The chain length of the glycans in bacterial cell walls. Biochem J *133*, 395-398.

328. Warren, R.J., and Bose, S.K. (1968). Bacteriophage-induced inhibition of host functions. I. Degradation of Escherichia coli deoxyribonucleic acid after T4 infection. J Virol *2*, 327-334.

329. Watanabe, M., Aoyagi, Y., Ridell, M., and Minnikin, D.E. (2001). Separation and characterization of individual mycolic acids in representative mycobacteria. Microbiology *147*, 1825-1837.

330. Watanabe, T., Kanai, R., Kawase, T., Tanabe, T., Mitsutomi, M., Sakuda, S., and Miyashita, K. (1999). Family 19 chitinases of Streptomyces species: characterization and distribution. Microbiology *145 (Pt 12)*, 3353-3363.

331. Weaver, L.H., Grutter, M.G., Remington, S.J., Gray, T.M., Isaacs, N.W., and Matthews, B.W. (1984). Comparison of goose-type, chicken-type, and phage-type lysozymes illustrates the changes that occur in both amino acid sequence and three-dimensional structure during evolution. J Mol Evol *21*, 97-111.

332. Weber, M.J., and DeMoss, J.A. (1969). Inhibition of the peptide bond synthesizing cycle by chloramphenicol. J Bacteriol *97*, 1099-1105.

333. Wehrli, W. (1983). Rifampin: mechanisms of action and resistance. Rev Infect Dis *5 Suppl 3*, S407-411.

334. West, N.P., Chow, F.M., Randall, E.J., Wu, J., Chen, J., Ribeiro, J.M., and Britton, W.J. (2009). Cutinase-like proteins of *Mycobacterium tuberculosis*: characterization of their variable enzymatic functions and active site identification. Faseb J *23*, 1694-1704.

335. West, N.P., Wozniak, T.M., Valenzuela, J., Feng, C.G., Sher, A., Ribeiro, J.M., and Britton, W.J. (2008). Immunological diversity within a family of cutinase-like proteins of Mycobacterium tuberculosis. Vaccine *26*, 3853-3859.

336. Whisstock, J.C., and Lesk, A.M. (1999). SH3 domains in prokaryotes. Trends Biochem Sci *24*, 132-133.

337. Whittaker, E. (1950). Two bacteriophages for Mycobacterium smegmatis. Can J Public Health *41*, 431-436.

338. WHO (2009). Global Tuberculosis Control, W.H. Organization, ed. (Geneva, Switzerland, WHO Press).

339. WHO (2010). Tuberculosis Fact Sheet, W.H. Organization, ed. (Geneva, Switzerland, WHO Press).

340. Wilhelm, S.W., Brigden, S.M., and Suttle, C.A. (2002). A dilution technique for the direct measurement of viral production: a comparison in stratified and tidally mixed coastal waters. Microb Ecol *43*, 168-173.

341. Witte, A., Brand, E., Mayrhofer, P., Narendja, F., and Lubitz, W. (1998). Mutations in cell division proteins FtsZ and FtsA inhibit phiX174 protein-E-mediated lysis of Escherichia coli. Arch Microbiol *170*, 259-268.

342. Wommack, K.E., and Colwell, R.R. (2000). Virioplankton: viruses in aquatic ecosystems. Microbiol Mol Biol Rev *64*, 69-114.

343. Xu, M., Arulandu, A., Struck, D.K., Swanson, S., Sacchettini, J.C., and Young, R. (2005). Disulfide isomerization after membrane release of its SAR domain activates P1 lysozyme. Science *307*, 113-117.

344. Xu, M., Struck, D.K., Deaton, J., Wang, I.N., and Young, R. (2004). A signal-arrest-release sequence mediates export and control of the phage P1 endolysin. Proc Natl Acad Sci U S A *101*, 6415-6420.

345. Yagi, T., Mahapatra, S., Mikusova, K., Crick, D.C., and Brennan, P.J. (2003). Polymerization of mycobacterial arabinogalactan and ligation to peptidoglycan. J Biol Chem *278*, 26497-26504.

346. Yamada, T., Satoh, S., Ishikawa, H., Fujiwara, A., Kawasaki, T., Fujie, M., and Ogata, H. (2010). A jumbo phage infecting the phytopathogen Ralstonia solanacearum defines a new lineage of the Myoviridae family. Virology *398*, 135-147.

347. Yoong, P., Schuch, R., Nelson, D., and Fischetti, V.A. (2004). Identification of a broadly active phage lytic enzyme with lethal activity against antibiotic-resistant Enterococcus faecalis and Enterococcus faecium. J Bacteriol *186*, 4808-4812.

348. Young, I., Wang, I., and Roof, W.D. (2000). Phages will out: strategies of host cell lysis. Trends Microbiol *8*, 120-128.

349. Young, R. (1992). Bacteriophage lysis: mechanism and regulation. Microbiol Rev *56*, 430-481.

350. Young, R. (2002). Bacteriophage holins: deadly diversity. J Mol Microbiol Biotechnol *4*, 21-36.

351. Young, R., Way, J., Way, S., Yin, J., and Syvanen, M. (1979). Transposition mutagenesis of bacteriophage lambda: a new gene affecting cell lysis. J Mol Biol *132*, 307-322.

352. Yunis, A.A., Smith, U.S., and Restrepo, A. (1970). Reversible bone marrow suppression from chloramphenicol. A consequence of mitochondrial injury. Arch Intern Med *126*, 272-275.

353. Zambrano, M.M., and Kolter, R. (2005). Mycobacterial biofilms: a greasy way to hold it together. Cell *123*, 762-764.

354. Zhang, N., and Young, R. (1999). Complementation and characterization of the nested Rz and Rz1 reading frames in the genome of bacteriophage lambda. Mol Gen Genet *262*, 659-667.

355. Zheng, Y., Struck, D.K., Bernhardt, T.G., and Young, R. (2008a). Genetic analysis of MraY inhibition by the phiX174 protein E. Genetics *180*, 1459-1466.

356. Zheng, Y., Struck, D.K., Dankenbring, C.A., and Young, R. (2008b). Evolutionary dominance of holin lysis systems derives from superior genetic malleability. Microbiology *154*, 1710-1718.

357. Zheng, Y., Struck, D.K., and Young, R. (2009). Purification and functional characterization of phiX174 lysis protein E. Biochemistry *48*, 4999-5006.

358. Zimmer, M., Vukov, N., Scherer, S., and Loessner, M.J. (2002). The murein hydrolase of the bacteriophage phi3626 dual lysis system is active against all tested Clostridium perfringens strains. Appl Environ Microbiol *68*, 5311-5317.

359. Zuber, B., Chami, M., Houssin, C., Dubochet, J., Griffiths, G., and Daffe, M. (2008). Direct visualization of the outer membrane of mycobacteria and corynebacteria in their native state. J Bacteriol *190*, 5672-5680.

360. Zuber, B., Haenni, M., Ribeiro, T., Minnig, K., Lopes, F., Moreillon, P., and Dubochet, J. (2006). Granular layer in the periplasmic space of gram-positive bacteria and fine structures of Enterococcus gallinarum and Streptococcus gordonii septa revealed by cryo-electron microscopy of vitreous sections. J Bacteriol *188*, 6652-6660.