Title
Peptidoglycan Degrading and Sensing Systems of Mycobacterium tuberculosis

Author
Prigozhin, Daniil Markovich

Publication Date
2012

Peer reviewed|Thesis/dissertation

Peptidoglycan Degrading and Sensing Systems of *Mycobacterium tuberculosis*

by

Daniil Markovich Prigozhin

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Molecular and Cell Biology

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:
Professor Tom Alber, Chair
Professor Carolyn Bertozzi
Professor Dan Portnoy
Professor Kathleen Ryan

Fall 2012

Peptidoglycan Degrading and Sensing Systems of *Mycobacterium tuberculosis*

Daniil Markovich Prigozhin

Abstract

Peptidoglycan Degrading and Sensing Systems of *Mycobacterium tuberculosis*

by

Daniil Markovich Prigozhin

Doctor of Philosophy in Molecular and Cell Biology

University of California, Berkeley

Professor Tom Alber, Chair

Mycobacterium tuberculosis (*Mtb*) cell wall, built on a cross-linked sugar-peptide polymer called peptidoglycan, protects the bacterial cell from adverse environments. Peptidoglycan homeostasis is maintained by extracellular peptidoglycan synthases and hydrolases. Intricate coordination of their activities is required to maintain structural integrity of the cell wall during growth, division, and response to stress. The sensor protein kinase B (PknB) is likely to play a critical role in monitoring the state of the peptidoglycan outside the cell and inducing subsequent metabolic changes inside. In this study, computational, biochemical, and structural approaches were used to characterize the peptidoglycan hydrolases of *Mtb* and to investigate the molecular mechanism of peptidoglycan signaling through PknB.

Peptidoglycan hydrolases are critical players in bacterial growth, division, cell shape determination, and peptidoglycan fragment-mediated communication. Computational analysis identified 22 mycobacterial peptidoglycan hydrolases based on homology to known enzymes from model organisms. The peptidoglycan degradation machinery of *Mtb* includes 4 N-acetylmuramoyl-L-alanine amidases, 8 lytic transglycosidases, and 10 peptidases of various specificities. Ten of these enzymes form a core set of mycobacterial peptidoglycan hydrolases, while four of them are essential for growth in *Mtb*. Comprehensive biochemical and structural investigation of the *Mtb* peptidoglycan hydrolases was initiated by cloning and heterologously expressing constructs representing all 22 *Mtb* peptidoglycan hydrolases in *Escherichia coli*. Robust expression was observed for all but one target protein. Twelve were successfully purified on large scale.

The peptidoglycan amidases Rv3717 and Rv3915 share similar catalytic cores yet have non-redundant functions in peptidoglycan turnover. Hydrolase activity assays using polymerized peptidoglycan sacculi and soluble peptidoglycan fragments elucidated contributions of individual amino acid residues, metal binding, and disulfide bond formation to catalysis. The structure of product-bound Rv3717 suggested a mechanism that limits this enzyme's activity on polymerized sacculi.

Peptidoglycan D,D-peptidases Rv2911, Rv3330 and Rv3627 are low molecular weight penicillin-binding proteins that participate in peptidoglycan maturation and degradation. Surprisingly, these three enzymes were inactive on peptidoglycan sacculi or peptidoglycan fragments, yet were active on beta-lactams meropenem and Bocillin. The structure of Rv3330 solved in complex with meropenem revealed a potential peptide-binging groove distant from the active site. The observed lack of activity of low molecular weight penicillin-binding proteins suggests a requirement for an activator. Discovery of such factors will significantly advance our understanding of *Mtb* peptidoglycan homeostasis.

PknB is an essential sensor kinase that controls cell wall biosynthesis. Its homologs in Gram-positive bacteria have been implicated in binding peptidoglycan fragments and in mediating bacterial responses to cell wall stress. To investigate the mechanism of peptidoglycan recognition by PknB, the structure of its extracellular sensor domain was solved. It consists of four 70-amino acid PASTA repeat domains that adopt an extended conformation. The last repeat domain contains a hydrophobic pocket with a conserved tryptophan, a signature of a ligand-binding site. The structure of PknB extracellular domain suggests that ligand-dependent localization and oligomerization control kinase activity.

Dedication

This thesis is dedicated to my teachers,
formal and informal, young and old,
past and present,
those still near me and those a world away.
I would not have been where I am without you.

Table of Contents

Acknowledgements

I thank my colleagues John Paul Huizar, Daniela Mavrici, and Hilary Vansell for their dedication to the peptidoglycan hydrolase project and for the many exciting discoveries and crushing disappointments that we had shared.

I thank Christine Gee for teaching me most of what I know about practical protein crystallography.

I thank Alejandra Cavazos and Roberta Cozzi for two most enjoyable and fruitful collaborations that were a welcome addition to thesis work.

I thank James Holton, George Meigs, and Jane Tanamachi for welcoming me to the wonderful Beamline 8.3.1.

I thank James Fraser for supervising me through a great rotation project and for getting me addicted to X-ray crystallography.

I thank Tony Chen, a rotation student, for his invaluable help with the PknB project and for allowing me to pretend I knew any better.

I am forever indebted to Christina Baer and Terry Lang for their advice and support.

I thank Ksenia Krasileva for setting an example of how much one can achieve in graduate school and for critically reading this manuscript.

Lastly, I thank my mentor Tom Alber for welcoming me to the lab and helping me freely explore and grow as a scientist. I could not be more grateful for the inspiration, the encouragement, and the wisdom so generously shared.

Chapter 1
Introduction

Along with HIV/AIDS and malaria, tuberculosis (TB) is among the three communicable diseases that cause the most morbidity and mortality worldwide. The current TB treatments take 6-9 months to complete, while emergence of strains with resistance to all front-line drugs puts further emphasis on developing new antimicrobials (Russell, Barry et al. 2010). Understanding the physiology of *Mycobacterium tuberculosis* (*Mtb*), the etiologic agent of TB, can help improve current therapies and lay the foundation for developing new treatments. The insights gained by studying model bacteria should be applied to *Mtb* research and expanded to elucidate the unique features of *Mtb* biology. The focus of this thesis is the cell wall homeostasis in *Mtb*, one of the most complex and exciting areas in bacterial physiology.

The cell wall is a composite structure that surrounds each bacterium and provides a selective barrier - letting nutrients flow across and keeping out harmful molecules; it is also a validated target of many antimicrobials (Koch 2003). In Gram-positive bacteria, the cell wall contains a thick layer of peptidoglycan linked to the plasma membrane by teichoic acids (Navarre and Schneewind 1999). The peptidoglycan is used for covalent attachment of cell wall-associated proteins and, in some species, a polysaccharide capsule. Gram-negative bacteria lack teichoic acids, their peptidoglycan layer is much thinner, but they elaborate an additional outer membrane that contains lipopolysaccharide and acts as an efficient barrier to protein toxins (Tokuda 2009). Mycobacteria are closely related to Gram-positive bacteria, yet are not stained with Gram stain; their cell walls are unique in both complexity and exceptionally low permeability (Hett and Rubin 2008). In mycobacterial cell wall, a thick peptidoglycan layer is covalently attached to an arabinogalactan layer linked to a mycolic acid layer. The mycolic acids and other non-covalently attached lipids form a mycolic outer membrane, in turn surrounded by a polysaccharide capsule. This assembly makes mycobacteria highly resistant to many water-soluble antibiotics. Nonetheless, the mycobacterial cell wall is a validated antibiotic target. Isoniazid, a first-line drug, inhibits the synthesis of mycolic acids, while ethambutol targets arabinogalactan. Peptidoglycan has not been considered a good target of anti-*Mtb* therapies due to the presence of a highly efficient beta-lactamase, but this view is currently being challenged (Hugonnet, Tremblay et al. 2009).

Peptidoglycan is a polymer common to all bacterial cell walls. It forms a single molecule that surrounds the entire cell, allowing it to resist turgor pressure (Crick, Mahapatra et al. 2001). Additionally, peptidoglycan serves as a basis for attachment of all other cell wall-associated polysaccharides, lipids, and proteins. Coincidentally, it is also the target of the first isolated antibiotic – penicillin. The initial steps in peptidoglycan biosynthesis are highly conserved across bacterial species. On the cytoplasmic side of the plasma membrane, multiple enzymes condense sugars and amino acids to create a disaccharide pentapeptide attached to a lipid carrier (van Heijenoort 2007). Following synthesis, the lipid-attached peptidoglycan monomer flips to the outside of the cell membrane for polymerization. Using the energy stored in the bonds of each monomer

building block, transglycosylases condense the disaccharides into long polysaccharide chains and transpeptidases crosslink these chains with peptide bridges. While all bacteria follow this general pathway, differences in the chemistry of the peptidoglycan monomer, in the length of polysaccharide chains, and the nature and abundance of crosslinks result in a wide diversity of bacterial peptidoglycans (Schleifer and Kandler 1972).

While the synthesis of the extracellular part of peptidoglycan depends on the enzymes with just two principal chemical specificities, bacteria require a much wider range of peptidoglycan degrading enzymes. These peptidoglycan hydrolases, often called autolysins, have multiple functions in shape maintenance, cell growth, daughter cell separation, peptidoglycan maturation, and fragment recycling. Autolysin chemical specificities are as diverse as their functions: in *Escherichia coli*, a model Gram-negative bacterium, at least one enzyme has been found for each glycosidic and peptide bond in the polymer, with most bonds targeted by multiple hydrolases (van Heijenoort 2011).

Based on their chemical specificity, the peptidoglycan hydrolases are grouped into glycosidases that cleave between the sugars, peptidases that target peptide bonds, and N-acetylmuramyl-L-alanine amidases that can separate peptides from sugar strands. While autolysins have been comparatively well studied in *E. coli* and *Bacillus subtilis* (Smith, Blackman et al. 2000, van Heijenoort 2011), a model Gram-positive bacterium, much less attention has been given to the peptidoglycan degradation machinery of *Mtb*. Nonetheless, key observations from model organisms can provide points of reference. First, elongation, septation, peptidoglycan maturation and recycling likely require dedicated hydrolases. Secondly, enzymes that act on the same bond should be differentially utilized via changes in localization, expression, regulation, and substrate selection. Third, most hydrolases should be active outside the cell and not in the cytoplasm where they could damage the monomer pool. Finally, peptidoglycan hydrolases should be tightly regulated in order to prevent autolysis – a process where hydrolytic enzyme activities overwhelm the synthetic machinery, resulting in peptidoglycan breakdown and bacterial lysis. Regulation can be achieved either by requiring an enzyme activator, or, in case of an active hydrolase, by making the target bonds accessible only in the sections of peptidoglycan targeted for degradation.

Despite many years of research, few examples of these kinds of regulation have emerged. Recently, Uehara and colleagues showed genetically that three *E. coli* peptidoglycan amidases require protein activators for full activity (Uehara, Parzych et al. 2010). Whether activation is a result of a physical interaction has not yet been confirmed. Hett and coworkers described an intriguing model where a mycobacterial peptidoglycan peptidase interacts alternatively with a synthetic or a hydrolytic enzyme, altering its activity (Hett, Chao et al. 2010). Protein complexes involving hydrolases and their putative regulators have been discovered in several systems, but the detailed molecular mechanisms and functional importance of these complexes are not well established (Romeis and Holtje 1994, Yang, Peters et al. 2011).

The centrality of cell wall homeostasis to bacterial survival warrants a comprehensive investigation of this complex system especially in pathogenic bacteria. Such studies should aim at uncovering the core enzymes of cell wall metabolism and at elucidating their regulatory/targeting protein complexes. Each new finding in this area will aid in developing new methods of disrupting cell wall metabolism, and can result in great therapeutic benefit. As part of such an investigation, I surveyed the *Mtb* genome for autolysins and partially characterized two autolysin families. Chapter 2 presents the computational analysis of the autolysin complement of *Mtb* including comparisons to model species and more closely related mycobacteria. Chapter 3 contains the biochemical and structural description of two peptidoglycan amidases, while Chapter 4 describes three D,D-peptidases that belong to the penicillin-binding protein superfamily.

The soluble fragments released by the hydrolases during the peptidoglycan turnover have emerged as potent signaling molecules. There are multiple receptors for peptidoglycan in higher organisms, including both animals and plants. NOD-like receptors detect intracellular peptidoglycan fragments and trigger immune responses within the cell (Chamaillard, Hashimoto et al. 2003), while peptidoglycan recognition proteins (PGRPs) exist as either membrane-associated or soluble circulating components of animal innate immunity (Royet, Gupta et al. 2011). Importantly, peptidoglycan sensors are also found in bacteria, where they can link the state of peptidoglycan to the intracellular responses or detect growth of nearby bacteria.

Protein kinase B (PknB) is a peptidoglycan sensor in *Mtb*. It is a trans-membrane, eukaryotic-like, Ser/Thr protein kinase (STPK) whose extracellular receptor domain is homologous to peptidoglycan binding domains of a subset of peptidoglycan synthetases. Its downstream targets include the putative *Mtb* flippase for peptidoglycan precursors, along with targets in central metabolism and transcriptional regulation of growth (Alber 2009). *Mtb* PknB is likely to play a critical role in monitoring the state of the peptidoglycan outside the cell and inducing subsequent metabolic changes within it. In Chapter 5 of this thesis, I determined the structure of the sensor domain of PknB to infer the structural basis for ligand recognition.

Chapter 2
Peptidoglycan Hydrolases of *Mycobacterium tuberculosis*

Introduction

Peptidoglycan hydrolases help bacteria grow, divide, and communicate using peptidoglycan fragments. Due to the complexities of working with mycobacteria, early efforts to biochemically identify and characterize *Mtb* autolysins were limited. Following the publication of the *Mtb* genome sequence, several hydrolases have been identified through homology to known enzymes in model systems and then assayed for their activity and contributions to cell biology. Systematically identifying mycobacterial peptidoglycan hydrolases and comparing this set to that of model bacteria is a first step towards understanding their contributions to cell wall homeostasis and peptidoglycan signaling. Surveying heterologous expression and purification of *Mtb* autolysins from *E. coli* defines the subset amenable to biochemical and structural characterization.

Peptidoglycan hydrolases were first identified as proteins responsible for autolysis, an inducible self-destruction of growing bacterial cultures (Conover, Thompson et al. 1966, Young 1966). Further investigation uncovered their many roles in normal cell growth, spore formation and germination, phage escape, and other processes (Wyckoff, Taylor et al. 2012). While many enzymes were identified via forward genetics and biochemical approaches and studied in detail using model bacterial species, the extension of this work to important pathogens such as *Mtb* has recently become a priority.

Few attempts have been made to directly isolate mycobacterial peptidoglycan hydrolases. In 1977 Kilburn and Best extracted *Mycobacterium smegmatis* cell walls with lithium chloride and used peptidoglycan fragment analysis to infer three distinct activities: an amidase, an L-Ala aminopeptidase, and an endopeptidase capable of cleaving iso-D-Gln-mDAP dipeptide (Kilburn and Best 1977). Nonetheless, most of the studies had identified mycobacterial autolysins through homology to known enzymes from other species.

Resuscitation-promoting factors A-E. Working with *Micrococcus luteus* Mukamolova and coworkers identified resuscitation-promoting factor (Rpf), a small protein capable of reactivating quiescent cells in starved cultures (Mukamolova, Kaprelyants et al. 1998). Rpf was subsequently shown to be a peptidoglycan glycosidase (Mukamolova, Murzin et al. 2006). The authors also noted the wide distribution of Rpfs among Gram-positive species and in mycobacteria. Indeed, later studies showed that five *Mtb* resuscitation-promoting factors (RpfA-E) were collectively required for resuscitation from dormancy while being dispensable for growth in culture (Kana, Gordhan et al. 2008).

Peptidoglycan amidase CwlM. A peptidoglycan amidase CwlM was discovered by homology to *B. subtilis* amidase CwlB, and the muralytic activity of CwlM was confirmed experimentally (Deng, Humphries et al. 2005).

NLPC/P60 endopeptidases RipA and RipB. Peptidase RipA was identified in a yeast two-hybrid screen for RpfB interactors, and its function as a peptidoglycan hydrolase was inferred through homology to NLPC/P60 family enzymes and then confirmed experimentally (Hett, Chao et al. 2007). RipB, also an NLPC/P60 family protein, is found next to RipA in the *Mtb* genome an is active on peptidoglycan fragments but not polymerized sacculi (Böth, Schneider et al. 2011).

This chapter defines the *Mtb* peptidoglycan hydrolase complement and describes experiments to comprehensively survey for the first time their heterologous expression.

Results

Bioinformatic analyses of Mtb autolysins

To find candidate *Mtb* autolysins, I consulted several review articles on bacterial autolysins (Layec, Decaris et al. 2008, Vollmer, Joris et al. 2008) and retrieved protein sequences from the NCBI Protein database. These primary sequences were searched against a Pfam domain database (Finn, Mistry et al. 2010) resulting in 18 Pfam Hidden Markov Models of the catalytic domains. Using these models, I scanned the *Mtb* genome with the program hmmsearch (Eddy 2011) to identify genes with significant matches (Table 2.1). A total of 22 putative hydrolase catalytic domains belonging to 8 different Pfam families were identified.

Similar results were obtained when this approach was applied to model bacterial species, yielding 31 and 39 hits in *E. coli* and *B. subtilis*, respectively (Table 2.1). Thus, both Gram-positive and Gram-negative bacteria share a limited set of autolysin catalytic domains highlighting the ancient evolutionary origin of this system. However, the precise composition of these sets is variable due to uneven expansion of some classes of enzymes as well as inclusion of rare domains.

The peptidoglycan degradation machinery of *Mtb* includes 4 N-acetylmuramoyl-L-alanine amidases, 8 lytic transglycosidases, and 10 peptidases of various predicted specificities (Table 2.2). This set is smaller than the one of either *E. coli* or *B. subtilis* with fewer hydrolase families present, as well as fewer representatives per family (Table 2.1). Unlike the model bacterial species, *Mtb* has 4 peptidoglycan hydrolases that are essential for growth in culture as defined by high density transposon mutagenesis (Griffin, Gawronski et al. 2011). Therefore, in *Mtb* peptidoglycan hydrolases and not just synthetases are potential drug targets.

In order to determine the core set of mycobacterial hydrolases, I surveyed the autolysin complements of four other mycobacterial species: *M. avium*, *M. abscessus*, *M. leprae*, and *M. smegmatis* (Table 2.3). The high degree of sequence conservation between the mycobacterial species allowed me to infer orthologous relationships based

on synteny – conservation of nearby genes. Thus, every row in Table 2.3 represents a group of orthologous genes. Ten of these groups have representatives in every organism investigated, forming the core mycobacterial autolysin set. It consists of two amidases, six peptidases, and two glycosidases. Predictably, all four essential *Mtb* autolysins belong to the core set.

Cloning and Expression of Mtb peptidoglycan hydrolases

To define a subset of autolysins amenable to biochemical and structural analysis, working together with Daniela Mavrici and John Huizar, I cloned over 50 constructs representing all 22 *Mtb* hydrolases and tested their heterologous expression in *E. coli*. Initial hydrolase constructs were cloned without the transmembrane and signal sequences, subsequent optimization often resulted in shorter constructs.

We observed expression of all but one target protein, nine of these were purified as soluble proteins, and four were refolded from solubilized inclusion bodies (Table 2.2). A common impediment to purification was the necessary oxidation of the disulfide bonds present in many of the *Mtb* hydrolases. For example, all the Rpf family proteins expressed as insoluble aggregates likely due to lack of formation of a conserved disulfide bridge that connects the ends of the catalytic domain. In such cases, refolding was preformed in a disulfide oxidizing buffer containing a 10:1 ratio of reduced to oxidized glutathione. Similarly, three of the putative peptidoglycan hydrolases, LpqF, LpqR, and LpqU are predicted to have lipid attachment sites and to be localized to membranes. All of these expressed as insoluble proteins and will likely require refolding in the presence of detergents to properly fold.

Discussion

Peptidoglycan is a shared constituent of most bacterial cell walls. The pathway for peptidoglycan monomer synthesis and the enzymes used for polymer condensation into the peptidoglycan sacculus are highly conserved. In comparison, the sets of peptidoglycan hydrolases found in model bacteria and in *Mtb* are more diverse. There are multiple enzymes in each class that target the same bond for hydrolysis but are likely differentially employed by the bacteria. In all, roughly twenty autolysin families have been identified to date. This limited number of catalytic domains is diversified by mutation and addition of accessory domains and motifs capable of regulating these enzymes via targeting to different locations in the cell wall and formation of protein complexes. This diversification limits inferences of enzyme function to their chemical specificities, and makes it necessary to investigate individual hydrolases, their regulation, interacting proteins, and their contributions to bacterial cell wall homeostasis within the species of interest.

In this chapter, I used HMM-based search algorithms to predict the autolysin complement of *Mtb*. This approach is both more sensitive and more accurate than the alignment-based approaches, such as BLAST, resulting in a target set not likely to contain false-positive hits. However, this method is limited by the number of search models used and, therefore, cannot identify a hydrolase with a catalytic domain not

previously observed. In this regard, it is important to note that two of the three autolysin specificities inferred from biochemical analysis of LiCl-extracted *M. smegmatis* cell walls (Kilburn and Best 1977) have candidate enzymes on the list. Specifically, there are four candidate amidases and five NLPC/P60 family enzymes that can cleave the D-iso-Gln-mDAP bond but no aminopeptidase that can cleave L-Ala away from the above dipeptide.

Compared to model organisms, *Mtb* has a reduced set of peptidoglycan degrading enzymes. This observation makes it likely that assignment of functions of individual hydrolases will be easier in mycobacteria than in the model organisms where the autolysin sets are highly redundant. Indeed, unlike either *E. coli* or *B. subtilis* where each peptidoglycan hydrolase can be deleted without arresting growth, four *Mtb* autolysins: Rv3915, Rv3627, Rv3953, and Rv1477 are essential for growth. Of these, only Rv1477 has an identified biological role in cell division (Hett, Chao et al. 2008). The uniqueness of these enzymes to bacteria along with their essentiality to *Mtb* growth makes these proteins useful drug targets.

The bacterial autolysin complements are constantly evolving by acquisition of new genes, duplication, and gene loss. In this chapter I compared the *Mtb* autolysin set to those of four other mycobacteria and found 21 group of orthologous enzymes, ten of which had representatives in every species. While a wider analysis of mycobacterial genomes is required to elucidate the mechanisms of autolysin evolution, the present work identified the core set of ten enzymes. These enzymes likely have conserved functions in mycobacteria and should be the focus of the investigation of *Mtb* autolysins. Unsurprisingly, the four essential *Mtb* autolysins fall into this core set.

The key aspects of autolysin biology include their biological functions, their substrate ranges, and their modes of regulation. To find out which autolysins are amenable to biochemical and structural studies, we have cloned all of the predicted *Mtb* autolysins and tested their recombinant expression. To overcome the complications in recombinant protein purification, we have employed multiple expression tags and optimized individual purification strategies to produce high purity soluble *Mtb* peptidoglycan hydrolases. Design of optimal protein constructs, oxidation of the disulfide bonds, and insoluble protein denaturation and refolding have emerged as the key strategies enabling *Mtb* autolysin purification.

Identification, cloning, and purification of *Mtb* peptidoglycan hydrolases are the first steps towards analysis of individual proteins' activity, structure, and regulation. The next two chapters present the results of biochemical and structural analysis of Amidase_3, Peptidase_S11, and Peptidase_S13 family proteins. Together with the study of knock out and overexpression phenotypes, autolysin localization, and partner discovery, this work will lead to a comprehensive understanding of the peptidoglycan homeostasis in *Mtb*.

Methods

Bioinformatic analyses

The genomes for all bacterial species surveyed were downloaded from NCBI (ftp.ncbi.nih.gov). The Pfam domain database (Finn, Mistry et al. 2010) and the HMMER software (Eddy 2011) were obtained from their respective sites (http://pfam.sanger.ac.uk/ and http://hmmer.janelia.org/). Supplying known autolysin protein sequences to the Pfam web server identified the catalytic peptidoglycan hydrolase domains. The resulting list of Pfam domain names was used to retrieve and concatenate models using *hmmfetch* program and bacterial genomes were surveyed for presence of these domains using *hmmscan*, both subprograms in the HMMER package. The resulting data tables were analyzed and annotated in Excel. Signal peptide identification was obtained from SignalP server version 3 (Bendtsen, Nielsen et al. 2004) and trans-membrane region identification was performed using TMHMM (Krogh, Larsson et al. 2001). In order to determine synteny relationships between pairs of mycobacterial autolysins, GenoList multiple genome browser was used (Lechat, Hummel et al. 2008).

Construct design and cloning

For the initial survey of recombinant peptidoglycan hydrolase expression I designed constructs that lacked predicted trans-membrane and signal peptide sequences. These were cloned from *Mtb* genomic DNA of strain H37Rv using two-step PCR protocol resulting in the addition of 5' attR recombination site and an in-frame TEV protease cleavage site and a 3' attR site. The amplified fragments were recombined into Gateway entry vector pDONR221, and further recombined into expression vectors pDEST15 and pHGWA (Busso, Delagoutte-Busso et al. 2005) adding an N-terminal GST or 6X-His tags.

Protein expression and purification

For the initial expression screen the proteins were expressed in *E. coli* BL21 CodonPlus (Stratagene) using auto-induction (Studier 2005). Cultures were grown at 37 °C for 8 hours and shifted to 18-22 °C for 18-24 hours. The pellets were harvested by centrifugation and stored at -80 °C. For purification, each pellet was resuspended in buffer A (300 mM NaCl, 25 mM HEPES pH 8.0, 0.5 mM TCEP, and 5% glycerol), additionally supplemented with 25 mM imidazole for His-tagged proteins. The cells were lysed by sonication and the resulting suspension clarified by centrifugation. The soluble portion was run on a 5 mL glutathione affinity or nickel HiTrap Sepharose column and each column was washed extensively with buffer A prior to elution with 300 mM imidazole in Buffer A or 100 mM reduced L-glutathione for Ni-affinity and GST-affinity columns, respectively. The soluble, insoluble and affinity-purified fractions were analyzed by SDS-PAGE. The optimized purification protocols for individual hydrolases are presented in the following chapters.

Table 2.1 Distribution of Autolysin Pfam Domains

Pfam Domain	Example	*Mtb*	*E. coli*	*B. subtilis*
3D	*B. subtilis* YocH	-	-	4
Amidase_2	*E. coli* AmiD, *B. subtilis* CwlA	2	2	5
Amidase_3	*E. coli* AmiA, *B. subtilis* CwlD	2	3	5
Amidase_5	Gram-positive phage endolysins	-	-	-
CHAP	Staphylococcus aureus Sle1	-	-	1
MltA	*E. coli* MltA	-	1	-
NLPC_P60	*E. coli* Spr, *B. subtilis* CwlS	5	5	7
Glucosaminidase	*E. coli* FlaZ	-	2	2
Glyco_hydro_25	*Bacillus anthracis* lysozyme	-	1	-
Peptidase_M15	*Enterococcus* VanX	1	1	-
Peptidase_M74	*E. coli* MepA	-	1	-
Peptidase_S11	*E. coli* PBP7	3	4	4
Peptidase_S**13**	*E. coli* PBP4, *B. subtilis* PBP4a	1	1	1
Peptidase_S66	*E. coli* LdcA	-	1	2
Phage_lysozyme	Bacteriophage P1 lysozyme	-	3	-
SLT	*E. coli* Slt70	3	6	-
Transglycosylas	*M. luteus* Rpf	5	-	6
VanY	*B. subtilis* CwlK	-	-	2
		22	31	39

Table 2.2 Putative *Mtb* Peptidoglycan Hydrolases

Pfam Domain	Gene Name	Gene Number	Essential	Expressed	Soluble	Purified
I. N-acetylmuramyl-L-alanine amidases						
Amidase_3		Rv3717		Y	S	Y
Amidase_3	CwlM	Rv3915	Y	Y	I	Y
Amidase_2		Rv3811		Y	I	
Amidase_2		Rv3594		Y	I	
II. Peptidases						
Peptidase_S13		Rv3627c	Y	Y	S	Y
Peptidase_S11	DacB2	Rv2911		Y	S	Y
Peptidase_S11	DacB1	Rv3330		Y	S	Y
Peptidase_M15	LpqR	Rv0838		Y	I	
Peptidase_S11	LpqF	Rv3593	Y	Y	I	
NLPC_P60	RipA	Rv1477	Y	Y	S	Y
NLPC_P60	RipB	Rv1478		Y	S	Y
NLPC_P60		Rv1566c		Y	S	Y
NLPC_P60		Rv2190c		Y	S	Y
NLPC_P60		Rv0024		Y	S	Y
III. Glycosydases						
SLT		Rv1230c		N		
SLT		Rv3896c		Y	I	
SLT	LpqU	Rv1022		Y	I	
Transglycosylas	RpfA	Rv0867c		Y	I	
Transglycosylas	RpfB	Rv1009		Y	I	Y
Transglycosylas	RpfC	Rv1884c		Y	I	
Transglycosylas	RpfD	Rv2389c		Y	I	Y
Transglycosylas	RpfE	Rv2450c		Y	I	Y

Table 2.3 Orhtologs of *Mtb* Peptidoglycan Hydrolases

Pfam Domain	*Mtb*	*M. avium*	*M. abscessus*	*M. leprae*	*M. smegmatis*
Amidase_2	Rv3811	MAP_0209c	MAB_0168c	-	MSMEG_6406
Amidase_2	Rv3594	-	-	-	-
Amidase_2	-	-	-	-	MSMEG_5315
Amidase_2	-	-	-	-	MSMEG_0533
Amidase_2	-	-	MAB_4807	-	-
Amidase_3	Rv3717	MAP_0318	MAB_0318c	ML2331	MSMEG_6281
Amidase_3	Rv3915	MAP_4341	MAB_4942	ML2704	MSMEG_6935
Amidase_3	-	MAP_2021c	-	-	-
NLPC_P60	Rv1477	MAP_1203	MAB_2728c	ML1812	MSMEG_3145
NLPC_P60	Rv1478	MAP_1204	MAB_2727c	ML1811	MSMEG_3146
NLPC_P60	Rv1566c	MAP_1272c	-	ML1214	-
NLPC_P60	Rv2190c	MAP_1928c	MAB_1974	ML0885	MSMEG_4256
NLPC_P60	Rv0024	MAP_0036	-	-	-
NLPC_P60	-	-	MAB_2474	-	-
NLPC_P60	-	-	MAB_3664	-	MSMEG_1686
NLPC_P60	-	-	-	-	MSMEG_3477
Peptidase_M15	Rv0838	MAP_0670	-	-	-
Peptidase_M15	-	-	MAB_1843	-	-
Peptidase_M15	-	-	-	-	MSMEG_5879
Peptidase_S11	Rv2911	MAP_2979	MAB_3234	-	MSMEG_2432 MSMEG_2433
Peptidase_S11	Rv3330	MAP_3448	MAB_3681	ML0691	MSMEG_1661
Peptidase_S11	Rv3593	MAP_0466c	MAB_0552c	ML1923	MSMEG_6087
Peptidase_S13	Rv3627c	MAP_0436	MAB_0519	ML0211	MSMEG_6113
SLT	Rv1230c	MAP_2552	MAB_1367c	-	MSMEG_5067
SLT	Rv3896c	MAP_4324c	-	-	-
SLT	Rv1022	MAP_0989	MAB_1164	-	MSMEG_5416
SLT	-	MAP_0586c	-	-	-
SLT	-	MAP_3011c	-	-	-
Transglycosylas	Rv0867c	MAP_0805c	MAB_0869c	ML2151	MSMEG_5700
Transglycosylas	Rv1009	MAP_0974	MAB_1130	ML0240	MSMEG_5439
Transglycosylas	Rv1884c	MAP_1607c	-	ML2030	-
Transglycosylas	Rv2389c	-	-	-	-
Transglycosylas	Rv2450c	MAP_2273c	MAB_1597	-	MSMEG_4643
Transglycosylas	-	-	MAB_4080c	-	-
Transglycosylas	-	-	-	-	MSMEG_4640
VanY	-	-	MAB_2019	-	-
VanY	-	-	-	-	MSMEG_1900

Chapter 3
Peptidoglycan Amidases Rv3717 and Rv3915

Introduction

L-Alanine-muramic acid amidases are ubiquitous peptidoglycan degrading enzymes with diverse roles in bacterial homeostasis, spore germination, and phage escape among others. In Chapter 2 of this thesis, I identified four putative peptidoglycan amidases in the genome of *Mtb* strain H37Rv: Rv3717 and Rv3915 both belong to Amidase_3 Pfam family while Rv3811 and Rv3594 contain Amidase_2 catalytic domains. Interestingly, the two Amidase_3 proteins have orthologs in the reduced genome of *M. leprae* and three other mycobacteria, suggesting they belong to the core set of mycobacterial autolysins with conserved functions.

Both amidase families contain well-studied representative enzymes. Of these, the better-studied examples are found in the Pfam family Amidase_2 that in addition to bacterial and phage amidases includes eukaryotic proteins with roles in anti-bacterial immunity such as fly and human peptidoglycan recognition proteins (Royet, Gupta et al. 2011). In contrast, members of Pfam Amidase_3 family are mostly found among bacteria and include the *E. coli* AmiA, AmiB, AmiC and *B. subtilis* CwlB and CwlC enzymes. CwlB, previously called LytC, is the founding member of this family and was discovered as one of two major *B. subtilis* autolysins (Young 1966, Rogers, Taylor et al. 1984). In the *E. coli* system all three enzymes require protein activators: AmiA and AmiB share an activator protein EnvC, and AmiC requires an EnvC homolog, NlpD (Uehara, Parzych et al. 2010). The *Mtb* system is different from the model organisms because it has fewer predicted amidases, and the Rv3915 amidase fulfills an unknown vital function in *Mtb* life cycle and is required for mycobacterial survival.

The peptidoglycan amidases Rv3717 and Rv3915 present a paradox common to peptidoglycan hydrolases: they share similar catalytic cores, yet have non-redundant functions in peptidoglycan turnover. This is a common feature of autolysin gene families. In the previous chapter, I identified twenty two putative *Mtb* autolysins that belong to eight families, with six families represented by at least two members. These families are diversified by duplication of the original gene, acquisition of mutations by the regulatory and coding sequences, amino acid substitutions to the catalytic domains and addition of targeting sequences, such as lipid attachment, transmembrane helices, and cell wall-binding domains. As a result, autolysins differ in timing and intensity of gene expression, protein localization, specificity, and processivity.

To answer the question what makes Rv3717 and Rv3915 gene products non-redundant, I produced them in *E. coli* and assayed them for peptidoglycan turnover. To determine the mode of substrate binding, I solved the structure of Rv3717 with bound reaction product. Structural analyses identified key catalytic residues and a subset of residues involved in substrate coordination.

Results

Sequence analysis, cloning, and expression of Mtb Amidase_3 family proteins

Pfam search for peptidoglycan-degrading enzymes identified two Amidase_3 domain-containing proteins, Rv3717 and Rv3915. In addition to the catalytic domain, Rv3717 but not Rv3915 contains a predicted signal peptide, while Rv3915 contains two N-terminal cell wall-binding domains. Rv3717 and its ortholog in *M. smegmatis*, MSMEG_6281, also contain a 20 amino acid insertion in the catalytic domain (Figure 3.1A). Sequence alignment of the catalytic domains of these proteins along with their *M. smegmatis*, *E. coli*, and *B. subtilis* homologs allowed construction of a maximum-likelihood tree (Figure 3.1B). This phylogenetic tree showed that while all of *E. coli and Mtb* amidases are closely related to each other, the *B. subtilis* clade is more diverse.

Rv3717 was cloned and expressed with different tags in *E. coli*. The Rv3717 constructs lacked the N-terminal 22 residues, a predicted signal peptide. The GST-tagged construct was expressed by autoinduction, purified over a glutathione affinity column, and cleaved with TEV protease to separate the GST tag. Tag removal was performed using a glutathione affinity column with an additional anion exchange step. Following disulfide oxidation, the protein was further purified by size exclusion chromatography that showed it was a monomer in solution.

GST-Rv3915 fusion protein was soluble, but it aggregated upon storage and was not cleavable using TEV protease despite the presence of a TEV cleavage site. In contrast, His-Trx-Rv3915 was insoluble upon expression. Isolation of insoluble inclusion bodies followed by solubilization in guanidine hydrochloride, a denaturant, allowed purification of 6X-His and thioredoxin-tagged Rv3915 by Ni affinity chromatography. A refolding screen condition containing 800 mM arginine and pH 8.3 BTP buffer allowed refolding of His-Trx-Rv3915 fusion protein. However, the protein remained prone to aggregation and conditions for tag removal could not be worked out. Consequently, the protein was not used in crystallization trials.

The two amidases are differentially active on polymerized peptidoglycan

The two amidases were used for activity tests on polymerized *B. subtilis* peptidoglycan. While Rv3717 did not clear the peptidoglycan suspension, refolded Rv3915 was almost as active as mutanolysin, a highly active peptidoglycan hydrolase used as a positive control. The difference between Rv3717 and Rv3915 in the catalytic activities on polymerized peptidoglycan could be attributed to either enzyme specificity or processivity.

Purified recombinant Rv3717 was active on muramyl dipeptide (MDP), a peptidoglycan fragment, but did not solubilize *Mtb*, *B. subtilis*, or *E. coli* peptidoglycan (Figure 3.2). The activity on MDP had a pH optimum of 6.5 and was metal-dependent, since addition of EDTA abolished the activity. HPLC separation and mass spectroscopy of reaction products confirmed the specificity of Rv3717 protein for the muramic acid - alanine bond. Since enzymes with the same target bond specificity are capable of clearing peptidoglycan suspensions, we concluded that Rv3717 might be limited by

either its rate of catalysis or by the context in which the target bond is found. In contrast, refolded recombinant Rv3915 cleared a suspension of *B. subtilis* peptidoglycan.

Crystallographic analysis of Rv3717

Rv3717 crystallized in presence of MDP by vapor diffusion. X-ray data were collected at cryogenic temperature using synchrotron radiation to 2.1 Å maximum resolution (Table 3.1). The structure was solved with molecular replacement using *Bartonella henselae* AmiB (PDB: 3NE8) trimmed down to the Rv3717-aligning sequences. The Rv3717 structure contains the typical features of Amidase_3 fold: a central six-stranded beta sheet, five surrounding alpha helices, and a Zn atom coordinated by two histidines and one glutamate in the active site (Figure 3.3A). In addition to these characteristic features, it has a 20-amino-acid insertion that forms a short alpha helix and a beta-hairpin that is attached to the core enzyme with a disulfide bond, as seen in comparison to *Paenibacillus polymyxa* CwlV (Figure 3.3B).

A clear density envelope remained above the catalytic Zn ion, its shape perfectly matched one of the products – the dipeptide L-Ala-iso-D-Gln (Figure 3.4A). Building this reaction product into the density completed the model building and refinement. This ligand refined with 100% occupancy and an average B-factor of 31 $Å^2$, equal to that of the protein atoms (Table 3.1). While MDP was added to the crystallization, it is clear from the density that only the dipeptide was bound, implying that the enzyme processed the substrate during the experiment and further suggesting that there is not a strong binding site for the second product, the muramic acid. The orientation of the ligand in the active site suggested that Glu200 contributes to Zn-mediated hydrolysis. This residue is positioned in the back of the active site at a distance of 5.1 Å from the metal. This position could allow it to activate a water molecule for hydrolysis of the muramic acid-L-alanine amide bond in a mechanism common to metal-dependent hydrolases (Figure 3.4B).

Contributions of Zn, Glu200, and the disulfide bridge to folding and catalysis

To further investigate contributions of the key structural elements to catalysis, I tested Rv3717 metal specificity. After washing with EDTA, the amidase lost activity. Following reconstitution with various divalent cations, only Zn^{2+}- and Co^{2+}-bound enzymes showed hydrolytic activity, while enzymes reconstituted with Mn^{2+}, Mg^{2+}, or Ca^{2+} ions did not (Figure 3.5A). The E200A mutant was inactive, supporting the role of Glu200 in catalysis (Figure 3.5B).

The 20-residue-long insertion that forms a beta hairpin positioned close to the active site is a characteristic feature of Rv3717 and its mycobacterial orthologs. Since it is attached to the core enzyme by a disulfide bridge, I hypothesized that disulfide oxidation will affect enzyme folding and catalysis. To address this question, I solved a crystal structure of reduced (disulfide free) and Zn-free Rv3717. Compared to the fully assembled product complex, this structure showed differences in the Zn^{2+}-coordinating residues, and the beta hairpin insertion was disordered (Figure 3.6A). The crystals of the reduced Rv3717 were soaked with Zn^{2+}-containing cryoprotectant solution and crystallographic analysis revealed Zn^{2+} binding to the catalytic center residues (Figure

3.6B). Thus, disulfide oxidation is dispensable for both folding of the core of the enzyme and of its catalytic center.

To investigate the effect of beta hairpin folding on catalysis, I compared MDP breakdown by Rv3717 under oxidizing and reducing conditions. The enzyme was minimally affected by disulfide reduction (Figure 3.6C). Taken together, these data suggest that disulfide oxidation is necessary for completion of folding, yet dispensable for catalytic center formation and substrate turnover.

Rv3717 substrate-binding site could be selective for short peptides

Since Rv3717 is active on MDP but incapable of solubilizing polymerized peptidoglycan, I looked at the possible sources of Rv3717 substrate selectivity. Figure 3.7 shows the dipeptide sitting at the mouth of a deep, water-filled tunnel. The sides of this tunnel are lined with hydrophobic residues, suggesting it could accommodate the long aliphatic chain of diaminopimelic acid found at position 3 of the peptide stem in *Mtb* peptidoglycan (Figure 3.7). Since the substrates are not readily available to test whether this is possible, we built models containing longer stem peptides. Without rearranging the protein, a larger substrate containing an extra residue can be accommodated. However, a full crosslink cannot be fit into the tunnel, suggesting the shape of the binding site may limit substrate range. This feature of the protein, not universally conserved among its homologs, may limit the activity of Rv3717 to a subset of peptidoglycan substructures.

Discussion

The diversification of peptidoglycan hydrolase families is responsible for the appearance of non-redundant enzymes that nonetheless cleave the same chemical bond. Duplication of the ancestral genes allows specialization and contributes unique features to peptidoglycan biology of various organisms while limiting functional inference to closely related species. Functional annotation by homology is therefore limited to orthologous pairs of enzymes such as Rv3717 and MSMEG_6281. In contrast, no functional inference can be made between paralogs such as Rv3717 and Rv3915. Likewise, little can be inferred about Rv3717 or Rv3915 function in *Mtb* by studying their distant homologs *E. coli*.

In this chapter, I used biochemical and structural analyses to determine conserved features of the Amidase_3 family proteins and their shared mechanism of substrate hydrolysis. Using activity assays with polymerized peptidoglycan and soluble peptidoglycan fragments, I confirmed that both proteins are functional amidases. However, Rv3915, but not Rv3717, could clear a suspension polymerized peptidoglycan sacculi. Structural analysis highlights several features of these enzymes that could explain the difference in activity. These differences include substitutions in the catalytic core, insertion of a beta hairpin adjacent to the active site in the Rv3717 catalytic domain, and addition of two N-terminal cell wall-binding domains in Rv3915.

The conserved features of Amidase_3 family proteins include the central beta sheet, the surrounding alpha helices, and the conserved catalytic center with three Zn^{2+}-coordinating residues and an extra acidic residue that could act as a proton acceptor for the water molecule used to cleave the target bond. The Zn^{2+} is necessary for catalysis and cannot be replaced by other divalent cations with the exception of Co^{2+}. The product-bound structure of Rv3717 identifies the shared substrate binding mode of Amidase_3 proteins and helps infer the residues used in substrate coordination and catalysis. Glutamic acid (Glu200) in the active site is required for catalysis and likely serves as a proton acceptor necessary to activate a water molecule for the attack on the amide bond.

The insertion in the Amidase_3 core domain, which distinguishes Rv3717 from other members of its family, forms the long beta hairpin whose folding depends on disulfide oxidation. By looking at the structure and activity of the reduced form, I concluded that this insertion is not required for folding of the core of the enzyme, and it contributes minimally to the *in vitro* hydrolysis of MDP. These observations do not rule out potential contributions of this novel feature of the enzyme to substrate selection and *in vivo* activity. Disulfide oxidation necessary to lock down the beta hairpin does not occur in the reducing cytoplasmic environment and can only happen after protein secretion. It can therefore serve as a localization control, allowing completion of folding only in the periplasmic environment. Additionally, a deep substrate-binding tunnel formed with the help of this unique addition could limit the selection of stem peptides available for hydrolysis by Rv3717. The inability to bind stem peptides whose third residue, diaminopimelic acid, is engaged in a crosslink with another stem peptide could limit the role of Rv3717 to monomer recycling. In this case, substrate processing could depend on a peptidase autolysin making the first cut.

Future work on the comparison of fragments released by Rv3915 and Rv3717 amidases from polymerized peptidoglycan will determine the enzyme specificities. Identification of enzyme products will help infer the roles of these enzymes *in vivo*. The higher activity of the Rv3915 amidase on polymerized peptidoglycan correlates with its essentiality to *Mtb*. Based on our observations, we hypothesize that Rv3717 acts as a recycling enzyme capable of cleaving short peptides but not full crosslinks, while Rv3915 is a processive enzyme with an essential role in either cell growth or division.

Methods

Sequence analysis, cloning, and expression of Rv3717 and Rv3915

HMM-based search of *Mtb* genome for Amidase_3 Pfam family proteins identified putative peptidoglycan amidases Rv3717 and Rv3915. Further sequence analysis using the SignalP web service revealed that Rv3717 had a signal peptide at its N-terminus, while Rv3915 did not. Consequently, Rv3915 was cloned full length into the Gateway family of cloning vectors while an N-terminal truncation of Rv3717 matching the sequence of the predicted mature protein was cloned. Multiple sequence alignment of catalytic domains was produced in *hmmalign* (Eddy 2011). Maximum likelihood

phylogenetic tree construction and bootstrapping were performed using the Phylip package (Felsenstein 2005).

GST-fused Rv3717 was expressed by auto-induction in *E. coli* BL-21 Codon Plus cells. Cell pellets were harvested by centrifugation and stored at -80 °C. Cells were resuspended in buffer A (300 mM NaCl, 20 mM HEPES pH 7.5, 5% glycerol, 0.5 mM TCEP) supplemented with protease inhibitors AEBSF and E64. Following cell lysis by sonication and pelleting of the insoluble fraction, glutathione affinity purification was carried out at room temperature using 5 mL GST-affinity columns. After elution with 30 mM glutathione in buffer A, protein was simultaneously cleaved with 0.1 mg TEV protease per liter of culture and dialyzed against buffer B (30 mM NaCl, 20 mM HEPES, 5% glycerol, 0.5 mM TCEP). The sample was passed through GST-affinity and anion exchange columns attached in tandem to achieve complete removal of the GST tag. The flow-through fraction was oxidized by addition of one-tenth final volume of oxidizing buffer C (100 mM reduced glutathione, 10 mM oxidized glutathione, 300 mM bis-tris propane pH 9, 10% glycerol, 300 mM NaCl, 10 mM zinc acetate). The sample was filtered over 0.2 um syringe filter and concentrated using Amicon centrifugal filters with 10 KDa MW cutoff. Concentrated oxidized protein was applied to Superdex-75 column mounted on an Acta FPLC instrument (General Electric) and preparative size-exclusion chromatography was performed against a non-reducing buffer containing 100 mM NaCl, 20 mM HEPES pH 7.5, and 10% glycerol.

His-Trx-tagged Rv3915 was expressed as described above. However, since the protein was found in the insoluble fraction, following lysis in buffer A and centrifugation, the pellet fraction was washed twice with 0.2% Triton X-100 in buffer A. The pellet was then resuspended by pipetting in denaturing buffer (6 M guanidine in buffer A) and incubated with constant stirring overnight. Following centrifugation to remove undisolved components, Ni-affinity chromatography was performed by adding 2 mL of equilibrated Ni resin to the sample and incubating 2 hours at 4 °C. Samples were added to gravity flow column and beads washed with 4 mL of denaturing buffer followed by elution with 100 mM imidazole in denaturing buffer. Following sample concentration and removal of imidazole, renaturation was achieved by rapid drop-by-drop dilution into renaturing buffer (800 mM arginine, 100 mM bis-tris propane pH 8.3, supplemented with one tenth final volume buffer C).

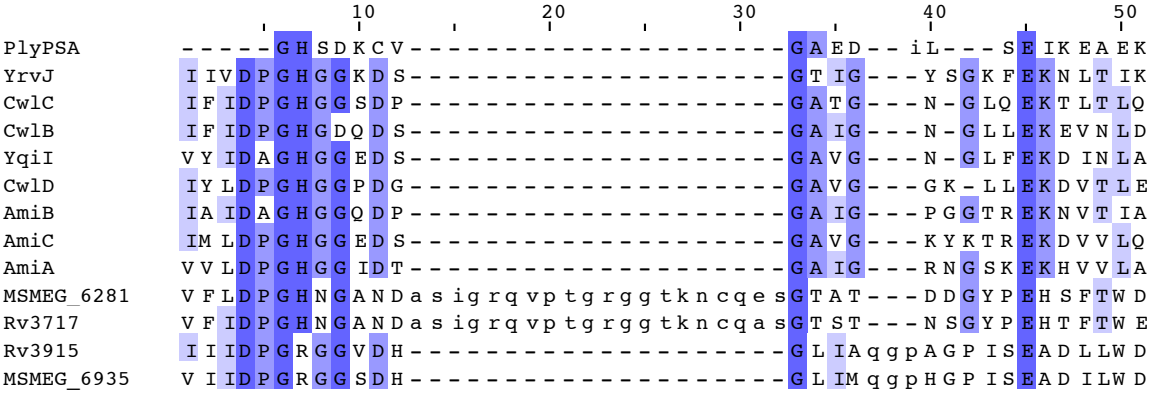
MDP degradation by Rv3717

Reaction buffers were prepared 5x with PIPES buffering agent unless stated otherwise (100 mM PIPES pH 6.5, 100 mM NaCl, 5% glycerol). MDP was used at 500 uM and Rv3717 at 5 uM. Reactions were mixed and incubated at room temperature for 40 minutes and stopped by centrifugation through a 10 KDa cutoff filter. Sample aliquots of 20uL were mixed with 100uL o-phthalaldehyde developing solution (1.5 mM o-phthalaldehyde, 350 mM sodium borate pH 9.5, 1% v/v beta-mercaptoethanol). Following 15 minute incubation, absorbance reading at 340 nm were taken and results analyzed in Excel.

Crystallographic analysis of Rv3717

All forms of Rv3717 were crystallized in 200 mM NaCl, 100 mM TRIS pH 8.5, 25% PEG 3350. The crystals were harvested using mother liquor as cryoprotectant and frozen in liquid nitrogen. Data were collected at Beamline 8.3.1 at the Advanced Light Source (MacDowell, Celestre et al. 2004) and reduced in HKL2000 (www.hkl-xray.com). Molecular replacement, using using *Bartonella henselae* AmiB (PDB: 3NE8) was performed using Phenix (Zwart, Afonine et al. 2008) and model building and refinement were carried out in Phenix and Coot (Emsley and Cowtan 2004). The data collection and model refinement statistics are listed in Table 3.1. Molecular images were generated using Chimera (Pettersen, Goddard et al. 2004).

A.



B.

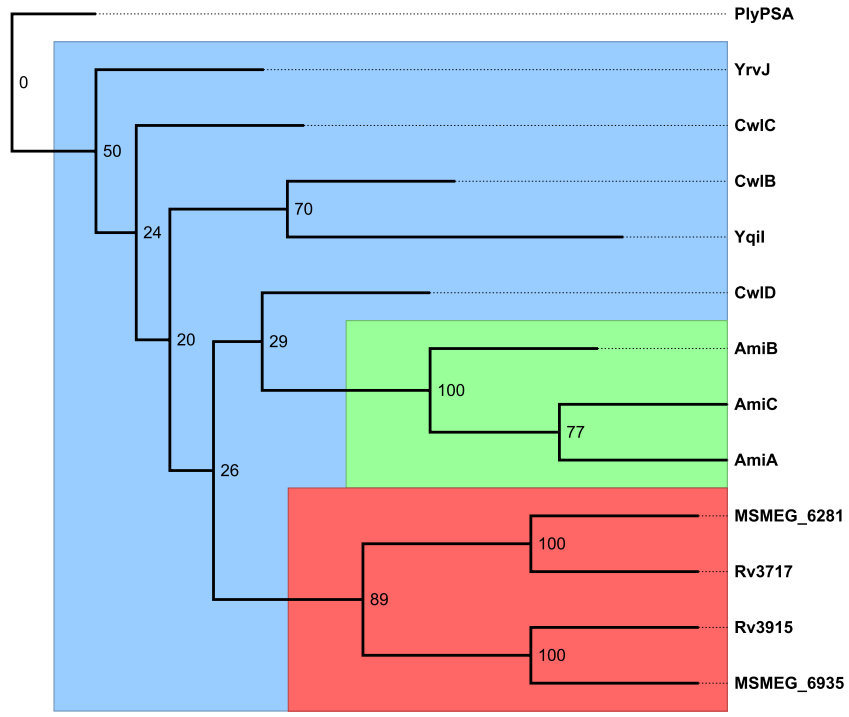
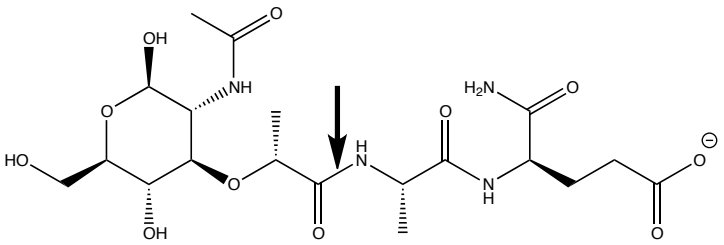


Figure 3.1. Phylogenetic relationships among peptidoglycan amidases.
A. Section of the sequence alignment of Amidase_3 family proteins. Protein sequences were trimmed to their catalytic domains and aligned with *hmmalign* program, part of the HMMER package (Eddy 2011). Residues 13-32 in the alignment are a unique insertion in Rv3717 and its mycobacterial orthologs **B.** Maximum-likelihood tree of Amidase_3 family proteins. Amidase_3 protein alignment was analyzed using the Phylip package (evolution.genetics.washington.edu/phylip). Bootstrapping values at nodes show high confidence in the mycobacterial (red) and *E. coli* (green) clades; *B. subtilis* clade (blue) is more diverse.

A.



B.

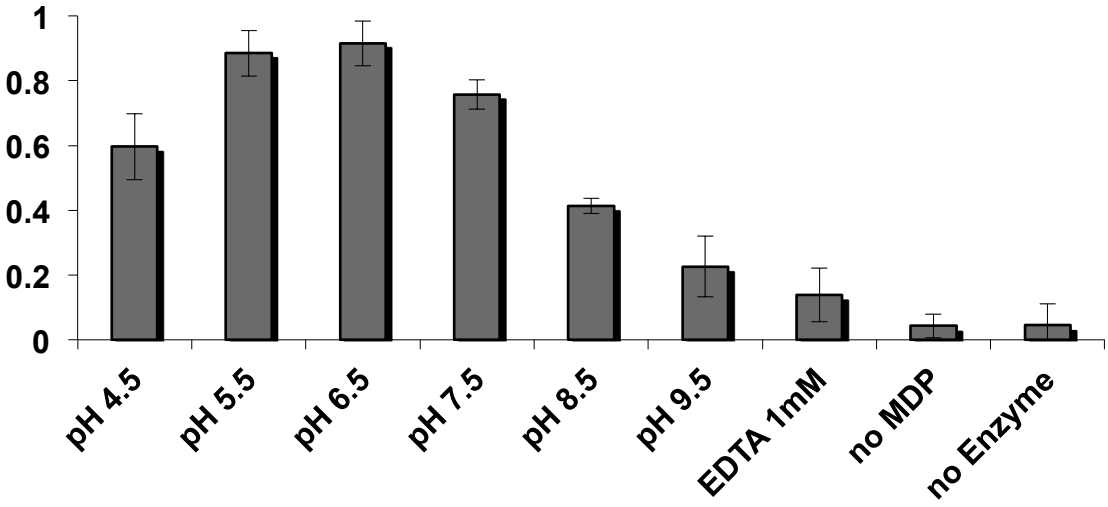
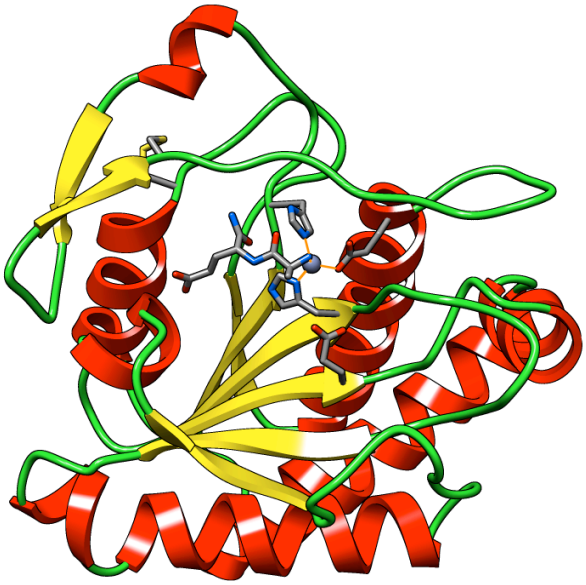


Figure 3.2. Activity of Rv3717.
A. Muramyl dipeptide substrate. Rv3717 hydrolyzes the amide bond between muramic acid and L-alanine (arrow), releasing a dipeptide with a primary amine group. **B.** Rv3717 cleaves MDP in a pH- and metal-dependent manner (Y-axis in arbitrary units; error bars represent the standard deviation of three measurements). Reactions were carried out in different primary amine-free buffers (acetate pH4.5, MES pH 5.5, PIPES pH 6.5, HEPES pH 7.5, BICINE pH 8.5, and CHES pH 9.5). Derivatization of the primary amine of the dipeptide product by o-phthalaldehyde allows product detection by absorbance at 340 nm. Addition of EDTA (1 mM) to the optimal pH 6.5 reaction removes protein-bound metals and prevents catalysis.

A.



B.

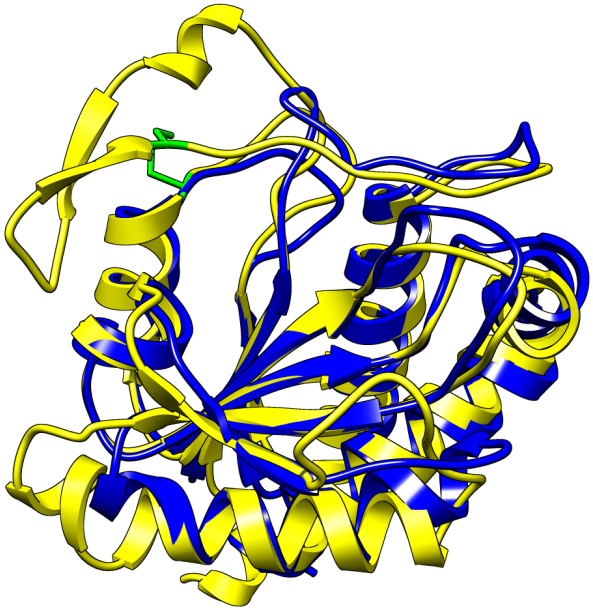
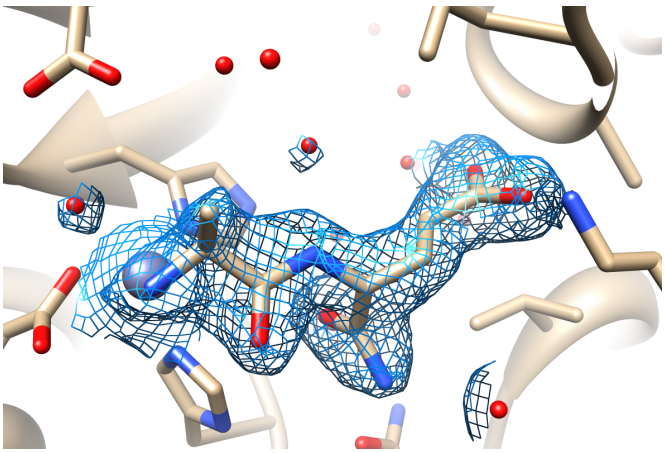


Figure 3.3. Crystal structure of Rv3717.
A. Ribbon diagram of Rv3717 in complex with dipeptide L-Ala-iso-D-Gln, one of two products formed during MDP hydrolysis. The amidase contains a Zn^{2+} cation in the catalytic center that is coordinated by two histidines and one glutamate residue. **B.** Comparison of Rv3717 structure (yellow) to another Amidase_3 protein, CwlV from *Paenibacillus polymyxa* (PDB: 1JWQ, blue) highlights the intra-domain insertion of a short alpha helix and a beta-hairpin characteristic of Rv3717 and its mycobacterial orthologs. The added features are held close to the enzyme core by two disulfide-bonded cysteine residues (green).

A.



B.

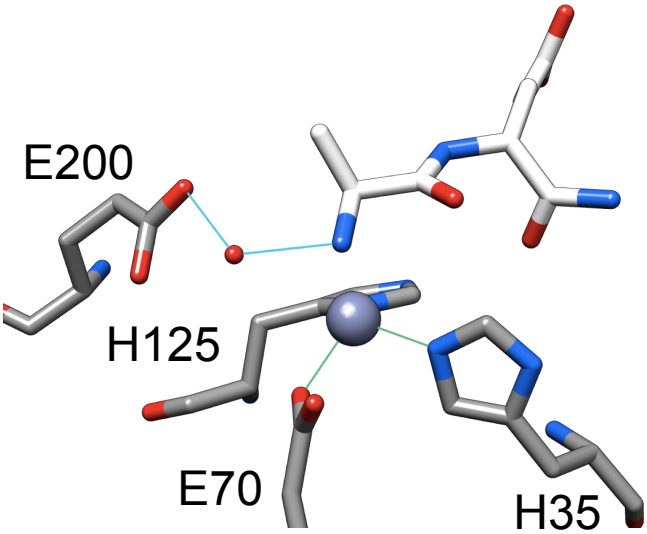
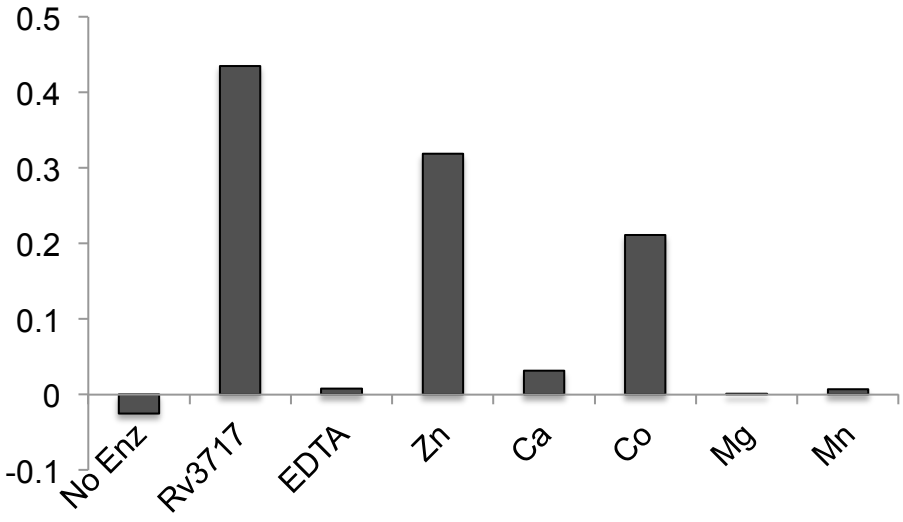


Figure 3.4. Rv3717 catalytic center.
A. Electron density map and the model of L-Ala-iso-D-Gln dipeptide bound to the Rv3717 active site. **B.** Rv3717 catalytic center showing catalytic residues (gray), the central Zn^{2+} atom (blue sphere), and the dipeptide product. Glutamate 200 is positioned close to the Zn-coordinated water molecule (red sphere) and could activate it for a hydrolytic attack on the muramic acid-L-alanine amide bond.

A.



B.

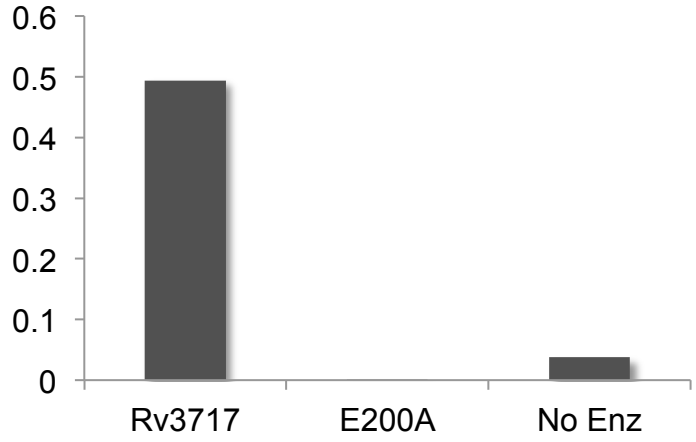


Figure 3.5. Rv3717 enzyme activity depends on Zn^{2+} and Glu200.
A. Rv3717 is a Zn^{2+}-dependent amidase. Enzyme activity on MDP was lost following EDTA wash and regained following addition of Zn^{2+} or Co^{2+} (Y-axis in arbitrary units). All reactions were performed in PIPES pH 6.5 buffer. **B.** Glu200 is required at the Rv3717 catalytic center. Wild-type and mutant enzymes were purified in parallel and tested for amidase activity on MDP. E200A mutant protein was inactive, demonstrating the importance of this residue to catalysis (Y-axis in arbitrary units).

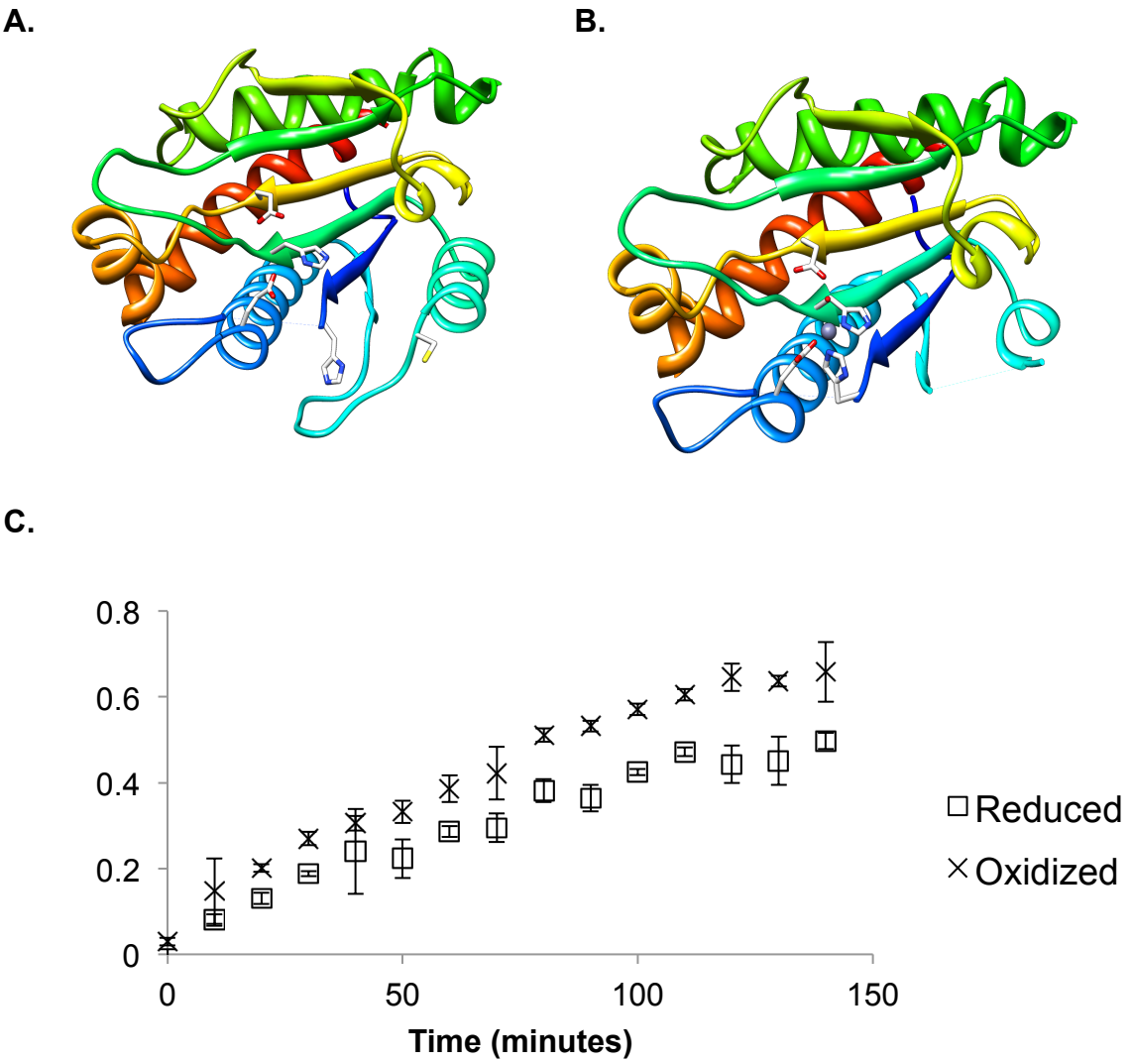
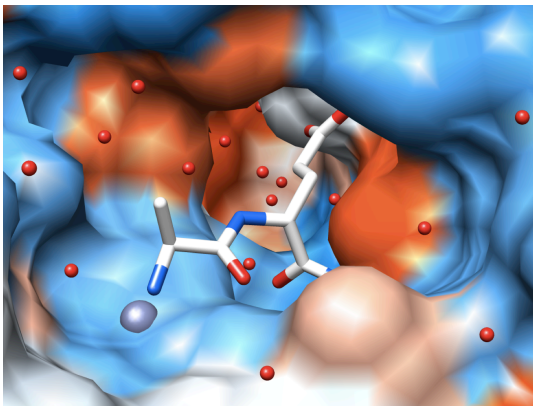


Figure 3.6. Zn^{2+}-binding and disulfide oxidation complete Rv3717 folding.
A. Ribbon diagram of the reduced and metal-free Rv3717 contains the core beta sheet and five helices, while the Rv3717-specific insertion and the catalytic site are disordered. **B.** Structure of reduced, Zn^{2+}-bound Rv3717 shows that the amidase is capable of binging Zn^{2+} and forming the catalytic center before folding is completed via disulfide oxidation. **C.** Disulfide reduction via addition of 0.5mM TCEP to the reaction buffer does not stop catalysis (Y-axis in arbitrary units).

A.



B.

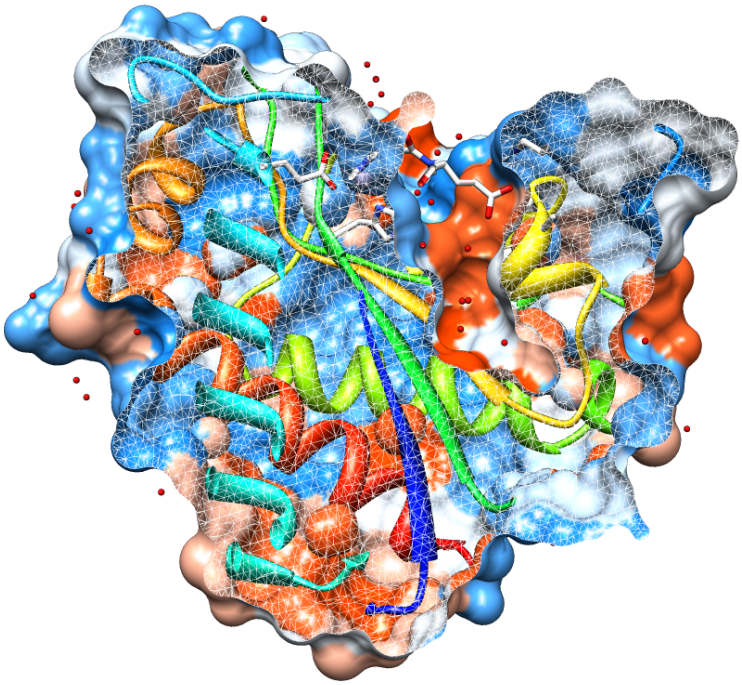


Figure 3.7. The dimensions of the substrate-binding site may determine Rv3717 specificity.
A. View into the Rv3717 active site showing catalytic Zn^{2+} and the dipeptide product on top a deep water-filled tunnel. **B.** Cross section shows the tunnel is large enough to accommodate a longer stem peptide, yet is closed at the bottom. The surface is colored by hydrophobicity from most (orange) to least hydrophobic (blue).

Table 3.1 Data collection and refinement statistics

	Rv3717 L-Ala-iso-D-Gln complex	Rv3717 reduced, metal-free form	Rv3717 reduced, Zn^{2+}-bound form
Data Collection			
Wavelength (Å)	1.12	1.12	1.28
Temperature (K)	100	100	100
Space group	P 21 21 2	C 2	C 2
Cell parameters			
a b c (Å)	55.9/76.0/49.5	125.4/45.7/67.8	125.4/45.9/68.0
α β γ (^{o})	90/90/90	90/116.0/90	90/116.3/90
Copies per a.s.u.	1	2	2
Resolution (Å) [a]	45.06 – 2.10 (2.16 – 2.10)	37.80 – 2.20 (2.24 –2.20)	42.52 – 2.67 (2.72 – 2.67)
R_{sym} (%)	17.3 (62.0)	9.1 (51.9)	14.4 (77.7)
I/σI	12.6 (3.36)	9.8 (1.4)	13.1 (1.5)
Completeness (%)	99.9 (99.4)	90.4 (87.9)	98.2 (71.5)
Redundancy	10.9 (9.7)	1.3 (1.3)	7.5 (5.1)
Refinement			
Resolution (Å)	45.07 – 2.08	37.8 – 2.19	42.52 – 2.67
Number of reflections	13,045	17,421	9,721
R_{work}/R_{free} (%)	17.8/23.0	19.1/25.2	17.7/23.8
Number of atoms			
Protein	1565	2563	2654
Solvent	90	125	35
B factors			
Protein ($Å^2$)	31	33	44
Solvent ($Å^2$)	34	36	44
Rmsd			
Bond lengths (Å)	0.014	0.015	0.011
Bond angles (°)	1.270	1.273	1.438
Ramachandran plot			
Favored (%)	97	98	97
Allowed (%)	100	99	99

[a] Values in parentheses are for the highest resolution shell.

Chapter 4
Peptidoglycan Peptidases Rv2911, Rv3330, and Rv3627

Introduction

Peptidoglycan peptidases Rv2911, Rv3330 and Rv3627 belong to a class of low-molecular-mass penicillin-binding proteins (LMM PBPs). While the high-molecular-mass PBPs are necessary for the last step of peptidoglycan crosslinking, LMM PBPs participate in peptidoglycan maturation and degradation. The LMM PBPs have been studied extensively and are classified based on substrate specificity as carboxy-, endo-, or trans-peptidases (van Heijenoort 2011). The last group does not include peptidoglycan hydrolases because trans-peptidases do not use water to resolve the covalent enzyme-substrate intermediate. The other two groups differ in the context of the bond they hydrolyze: the carboxypeptidases require a terminal alanine, while endopeptidases do not have this restriction. While conceptually simple, this distinction remains mysterious, since the structural basis for selection of the substrate or the nucleophile for the second step of the reaction have not been defined (Josephine, Charlier et al. 2006).

The LMM PBPs of *E. coli* are among the best studied. *E. coli* PBP5, one of two major LMM PBPs, is a D,D-carboxypeptidase responsible for maturation of peptidoglycan and trimming of excess D-Ala-D-Ala peptide termini in both peptidoglycan monomers and dimers (van Heijenoort 2011). While knockout of PBP5 has minor abnormalities, its overexpression is highly deleterious, likely due to decreased crosslinking that compromises peptidoglycan integrity. PBP7, the second most abundant LMM PBP, is an endopeptidase that only cleaves crosslinks. Both the knockout and overexpression of PBP7 are well tolerated.

In Chapter 2 of this work, I identified four putative LMM PBPs in *Mtb*, three of which, Rv2911, Rv3330, and Rv3627, were successfully expressed in *E. coli* and purified to homogeneity. The first two enzymes belong to the Peptidase_S11 family and are more closely related to E. coli PBP5, while Rv3627 belongs to Peptidase_S13 family and is more closely related to PBP7. Interestingly, Rv3627 is essential for mycobacterial growth, unlike its *E. coli* homolog.

In this chapter, I used biochemical and structural approaches to investigate the three peptidoglycan peptidases. I found that all three were capable of binding beta-lactams but were inactive on polymerized peptidoglycan or on soluble peptidoglycan fragments. A crystal structure of Rv3330 in complex with meropenem is presented that contains a peptide-binding groove away from the catalytic site.

Results

Sequence analysis, cloning, and expression of Mtb D,D-peptidases

Rv2911, Rv3330, and Rv3627 were cloned from genomic *Mtb* DNA with their predicted signal peptides removed. In addition, Rv3330 contains a C-terminal hydrophobic sequence predicted to form a transmembrane helix that tethers the enzyme on the periplasmic side of the plasma membrane. This helix was also removed to produce a soluble Rv3330 construct. All three proteins expressed as soluble enzymes and were purified using either glutathione or Ni affinity chromatography.

The Mtb LMM PBPs are active on penicillin but not peptidoglycan

Working with John Huizar, a junior specialist in the Alber lab, we first tested binding of the three LMM PBPs to polymerized peptidoglycan sacculi. None of the enzymes bound peptidoglycan of *B. subtilis* (Figure 4.1A) or *E. coli* (not shown).

We assayed release of soluble peptifoglycan fragments by these enzymes using o-phthalaldehyde derivatization that detects primary amino groups produced during peptide bond hydrolysis. All three enzymes were inactive on peptidoglycan preparations (Figure 4.1B). Finally, we solubilized *E. coli* peptidoglycan with hen egg white lysozyme and subjected the fragments to each of the three peptidases, but no activity was detected (Figure 4.1C).

As an additional test of catalytic activity, we tested binding of the three LMM PBPs to the mechanism-based, covalent inhibitor, Bocillin, a fluorescent penicillin analog (Zhao, Meier et al. 1999). Following a brief incubation with Bocillin, the proteins were subjected to denaturing SDS-PAGE, and the gels were scanned for characteristic fluorescence emission at 530 nm. This analysis revealed all three proteins stained with Bocillin, suggesting they were active. In this assay the S114A mutant of Rv3627 that lacks the catalytic serine did not stain with Bocillin, confirming the specificity of binding and nucleophilic attack (Figure 4.2).

Crystallographic analysis of Rv3330

The structure of Rv2911 was determined previously by Krieger and coworkers (PDB: 2BCF), so we focused our efforts on crystallizing the Rv3330 and Rv3627 enzymes. While the latter protein failed to crystallize in sparse matrix screens, the former yielded diffraction quality crystals from an ammonium sulfate condition that also included a two-fold molar excess of meropenem, a beta-lactam antibiotic. The data were collected at cryogenic temperature using synchrotron radiation to 2.0 Å maximum resolution (Table 4.1), and the structure was solved by molecular replacement using the Rv2911 coordinates as the search model.

The Rv3330 structure is a typical Peptidase_S11 two-domain structure where one domain has a beta-sheet core formed by strands from both N- and C-termini of the protein, and the second domain is alpha-helical. The refined structure also included a covalently linked molecule of meropenem, its beta-lactam ring open, covalently bonded to the catalytic nucleophile, Ser114 (Figure 4.3A).

The structure of Rv3330 is similar to that of Rv2911, with an rmsd of 0.76 Å over 235 alpha-carbons, although the proteins share just 46% of residues over the aligning region (Figure 4.3B). It therefore appears that the key difference between the two enzymes is the presence in Rv3330 of a C-terminal transmembrane domain. Further localization studies could reveal specialization of Rv2911 and Rv3330 via targeting to different regions of the sacculus as a result of this modification.

A single molecule of meropenem was bonded to the active-site serine and appears clearly in the electron density (Figure 4.4A). While we failed to crystallize Rv3330 without meropenem, comparison to ligand-free Rv2911 suggests that no major rearrangements occur in response to beta-lactam binding. The meropenem molecule extends away from the active site with the carboxyl group that mimics the terminal D-alanine bound to the enzyme. No other interactions that could be informative of the substrate selection mechanisms are apparent in the Rv3330-meropenem complex (Figure 4.4B).

Investigation of the crystal contacts in the Rv3330 structure suggests there is potential for protein-protein interactions and participation of Rv3330 in cell wall-remodeling complexes. First, the N-terminal region contains a single cysteine residue that formed a disulfide bond with a symmetry-related Rv3330 monomer in the crystal (Figure 4.5A). Whether this cysteine is part of the mature protein and is used as an attachment site *in vivo* will require further investigation. However, its conservation in other mycobacteria as part of an LXXC lipobox-like motif (LGGC in *Mtb*) further suggests that this could be a site of lipid attachment. Second, the face of the enzyme opposite the active site contains a crystal contact with the C-terminal most sequence of another Rv3330 monomer (Figure 4.5B). The peptide binding in this region resulted in burial of hydrophobic surfaces covering an area of 250 $Å^2$. This potential peptide-binding pocket is lined with residues that are conserved among Rv3330 orthologs.

Discussion

Peptidoglycan D,D-peptidases act on the terminal amino acid residues in the stem peptides to create or break crosslinks and to trim unused pentapeptide stems as part of the maturation process. These enzymes play important roles in shape maintenance, antibiotic resistance, and monomer recycling.

Mtb contains five predicted D,D-peptidases, including four penicillin-sensitive ones, three of which are described here. Rv2911 and Rv3330 are Peptidase_S11 family members and paralogs of *E. coli* PBP5, while Rv3627 is a paralog of *E. coli* PBP7 and a Peptidase_S13 family protein. While the former two enzymes are dispensable for mycobacterial growth *in vitro*, Rv3627 is an essential enzyme. In view of this observation, it is puzzling that the only activity we observed for these enzymes was binding to penicillin analogs.

There are two hypotheses that can explain the negative results from our activity assays. First, these enzymes may be inactive decoys that soak up beta-lactam antibiotics and protect the critical synthetic PBPs. Given that Rv3627 is essential, and that similar lack of activity has been observed for many soluble PBPs (Josephine, Charlier et al. 2006), we favor the second hypothesis that additional factors are required for proper activation of these potentially harmful enzymes. Activity on penicillin analogs could be explained by noting that penicillin is a chemically reactive, mechanism-based suicide inhibitor that is more likely to bind to these enzymes in absence of an activator protein. Discovery of such regulatory factors will present an important advance in the study of peptidoglycan biology.

The structure of the meropenem complex of Rv3330 shows that this enzyme is similar to its Rv2911 paralog. The two should, however, differ in localization, because Rv3330 contains a potential N-terminal lipid attachment site and a C-terminal transmembrane helix, while Rv2911 is predicted to be a soluble periplasmic protein. The structure also suggests that Rv3330 contains a peptide-binding site. While the binding of Rv3330 C-terminal sequence to this site is likely a byproduct of high protein concentration used in crystallization, the burial of hydrophobic area as a result of binding, as well as conservation of pocket-lining residues, hint that this is a possible protein-protein binding site. Further studies should test the hypothesis that protein activators are required for *Mtb* LMM PBPs.

Methods

Protein Purification

Proteins were purified with high yield using an N-terminal GST tag (Rv2911, Rv3330) or an N-terminal His-Trx tag (Rv3627) following heterologous expression in *E. coli* via either autoinduction (Studier 2005) or IPTG induction. Following elution from the primary GST or IMAC column, the protein was simultaneously cleaved with TEV protease and dialyzed to remove the eluent. The secondary GST or IMAC was used to remove the tag, and a Superdex 75 size-exclusion step completed the purification.

Peptidoglycan binding assay

B. subtilis peptidoglycan (1 mg) was twice washed in 1 mL of utra-pure water and resuspended in 100 µl of binding buffer (0.15 M NaCl, 25 mM HEPES, pH 7). 10 µl of protein at 0.4 mg/mL to 10 µl of peptidoglycan suspension were mixed and incubated for 5 minutes at room temperature with nutation. Samples were centrifuged for two minutes at 11,500 g, and bound and free protein fractions were analyzed by SDS-PAGE and Coomassie staining along with an input control.

Detection of PBP-released peptide N-termini

Hydrolysis of the peptide bond releases an N-terminal primary amine. We used 5 µM enzyme to digest either intact or lysozyme-solubilized *E. coli* peptidoglycan, and detected primary amino groups using o-pthalaldehyde (Böth, Schneider et al. 2011). Following overnight incubation with the hydrolase, the samples were filtered through a concentrator (MW cutoff 3,000 Da) to remove the protein and mixed with fresh o-

phthalaldehyde derivatization buffer (1.5 mM o-phthalaldehyde, 0.35 M sodium borate, 1% v/v beta-mercaptoethanol). The absorbance at 340 nm was measured using a plate reader, and the data were analyzed in Excel.

Bocillin binding to Mtb LMM PBPs

Mtb enzymes (0.25 mg/reaction) were labeled with 25 μM Bocillin for 30 minutes at 37 °C in labeling buffer (25 mM NaCl, 5 mM HEPES pH 7, 5% glycerol). SDS-PAGE separation and imaging on a Typhoon fluorescent scanner (excitation at 488 nm, emission at 530 nm) were followed by Coomassie staining.

Structure of Rv3330

Rv3330 recombinant protein lacking the signal peptide and the transmembrane helix gave large crystals in the following condition: 0.1 M NaCl, 0.1 M HEPES pH 7.23, 1.7 M ammonium sulfate, 10 mg/mL protein, and a 2:1 molar excess of meropenem. The X-ray diffraction data were collected at Beamline 8.3.1 at the Advanced Light Source (MacDowell, Celestre et al. 2004) and processed in HKL2000 (www.hkl-xray.com). The structure was determined by molecular replacement using the Rv2911 coordinates as the search model. Molecular replacement, model building and refinement were performed using Phenix (Zwart, Afonine et al. 2008) and Coot (Emsley and Cowtan 2004). The data and model statistics are listed in Table 4.1. Molecular images were generated using Chimera (Pettersen, Goddard et al. 2004).

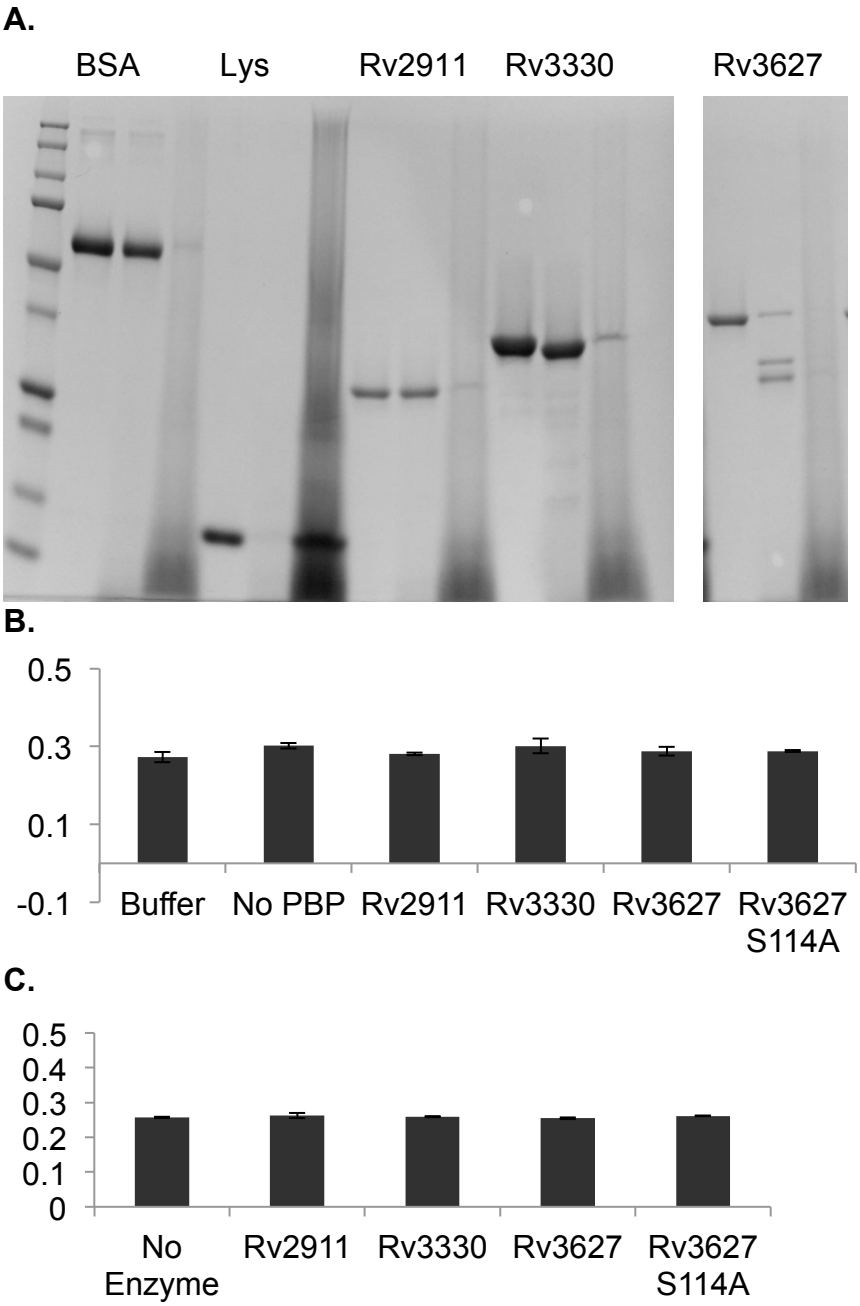


Figure 4.1. Lack of peptidoglycan binding or hydrolysis by purified *Mtb* D,D-peptidases *in vitro.*

A. *Mtb* LMM PBPs do not bind to polymerized peptidoglycan. Bovine serum albumin (BSA), hen egg white lysozyme (Lys) and three *Mtb* LMM PBPs were incubated with *B. subtilis* peptidoglycan, and input, free, and bound fractions were analyzed by SDS-PAGE. **B.** *Mtb* LMM PBPs did not release any soluble reaction products with free amino termini from polymerized *E. coli* peptidoglycan (Y-axis in arbitrary units). **C.** Lysozyme-solubilized *E. coli* peptidoglycan fragments were not hydrolyzed by the *Mtb* LMM PBPs (Y-axis in arbitrary units).

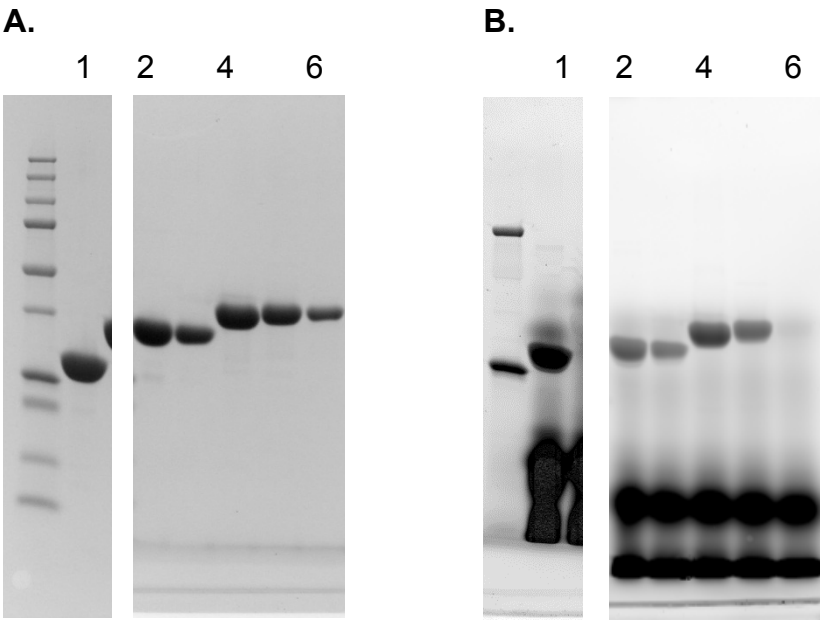
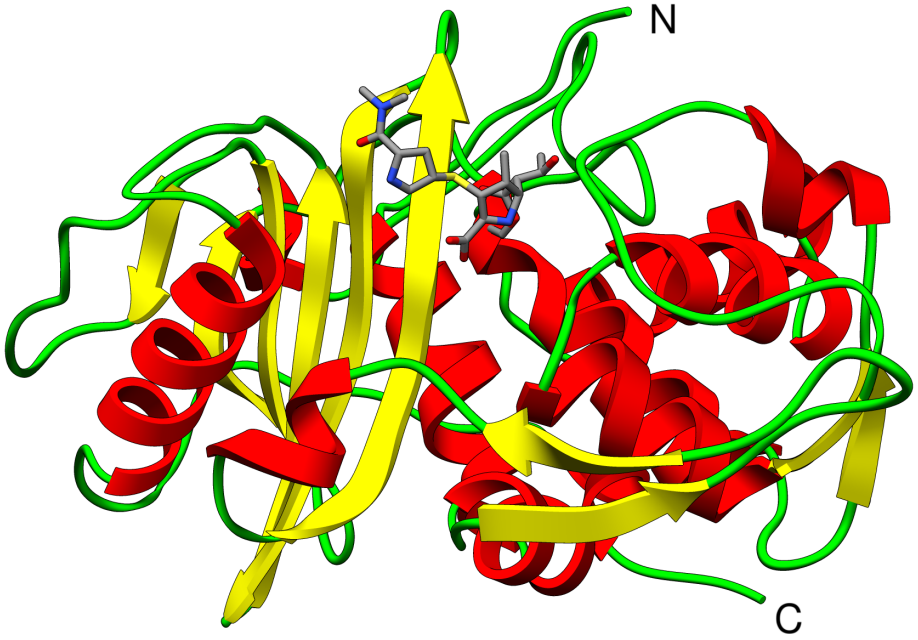


Figure 4.2. Purified *Mtb* D,D-peptidases react with a suicide inhibitor *in vitro.*
A. Coomassie staining and **B.** Fluorescence scan of Bocillin-labeled *Mtb* LMM PBPs: Rv2911 (Lane 1), Rv3330 (Lanes 2, 3), Rv3627 (Lanes 4, 5), and S114A Rv3627 (Lane 6). Lack of Bocillin staining of catalytically inactive S114A Rv3627 mutant confirms that the interaction is specific for active enzymes.

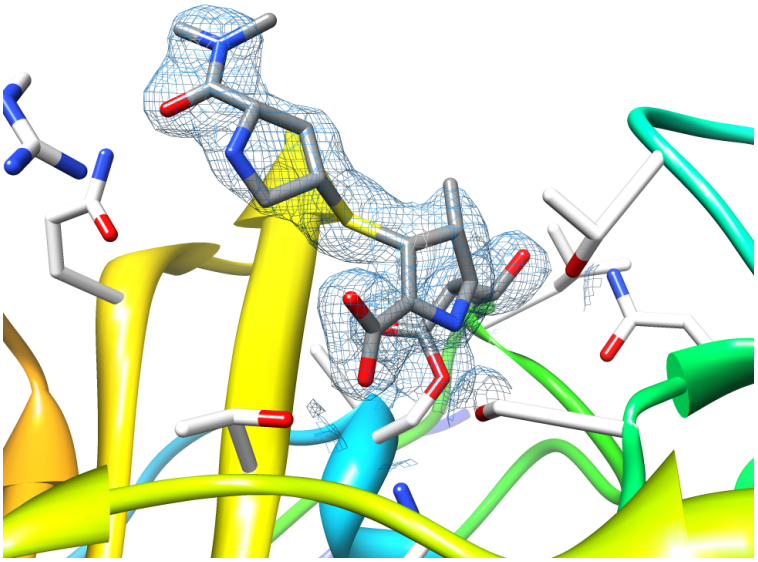
A.



B.

Figure 4.3. Structure of the meropenem adduct of Rv3330.
A. The enzyme folds into two domains. The beta-sheet in the first domain has strands from the N- and C-termini. **B.** Rv3330 (yellow) is structurally similar to its paralog, Rv2911 (blue).

A.



B.

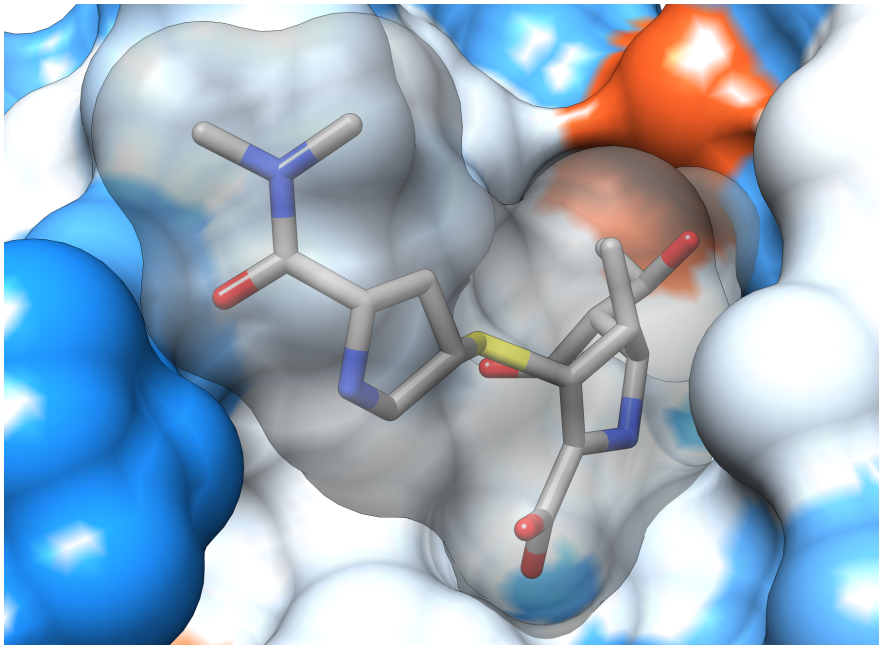
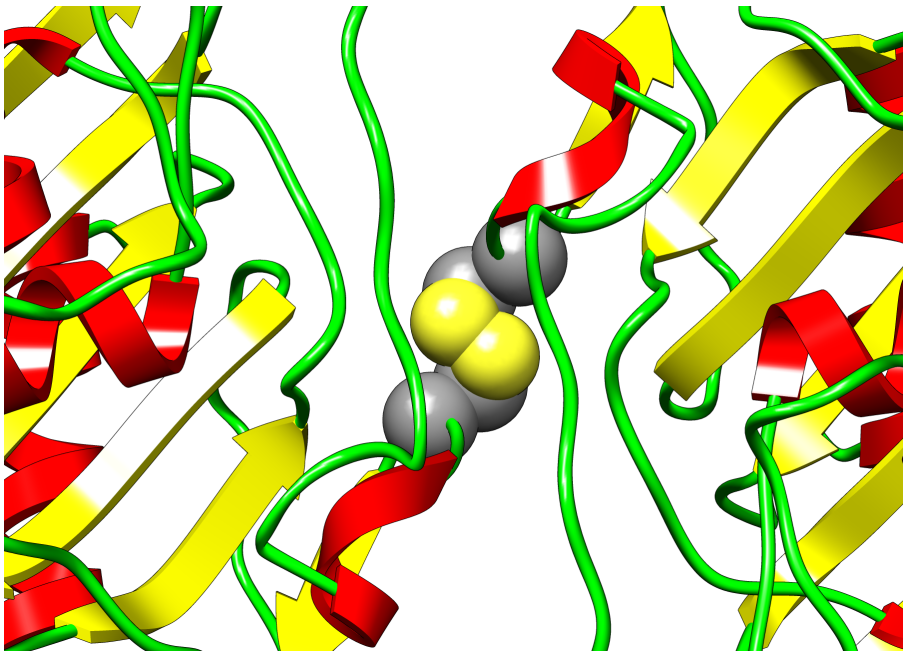


Figure 4.4. Rv3330 active site with bound meropenem.
A. Meropenem electron density at 1 sigma level (2.0 Å resolution) was visible in the Rv3330 active site. **B.** The molecular surfaces of meropenem (stick representation, gray translucent surface) and Rv3330 active site (surface colored by hydrophobicity) highlight interaction between the D-Ala-mimicking carboxylate of the drug with the enzyme active site.

A.



B.

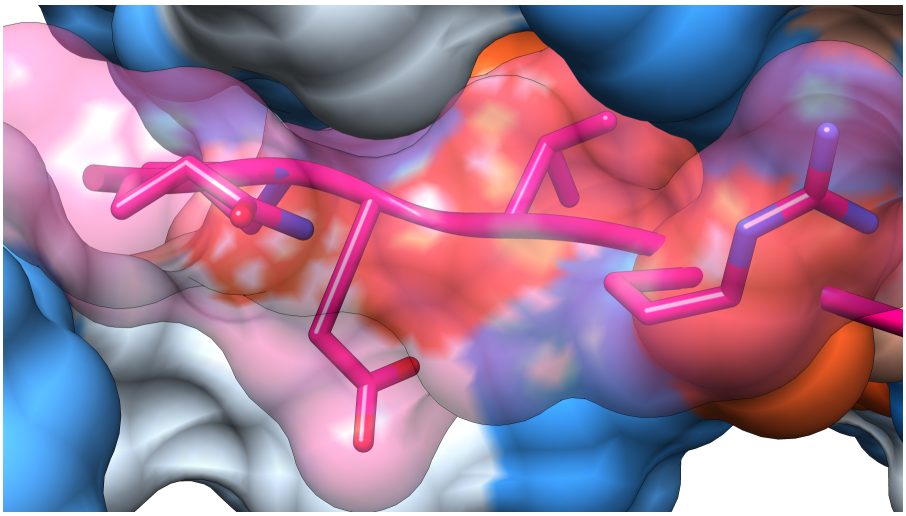


Figure 4.5. Crystal contacts observed in the Rv3330 structure.
A. Disulfide bond connecting two symmetry-related molecules of Rv3330. Participating cysteines depicted as spheres. **B.** Peptide-binding groove on the surface of Rv3330 accommodated an ordered C-terminal peptide from another monomer. Interaction between a hydrophobic groove (solid surface) and a C-terminal peptide from another monomer (pink ribbon, pink translucent surface) resulted in burial of hydrophobic surface (orange).

Table 4.1 Data collection and refinement statistics

	Rv3330
Data Collection	
Wavelength (Å)	1.12
Temperature (K)	100
Space group	P 4_1 2 2
Cell parameters	
a *b* *c* (Å)	76.29 76.29 99.98
α β γ (^{o})	90 90 90
Copies per a.s.u.	1
Resolution (Å) [a]	47.49- 2.00 (2.07 – 2.00)
R_{sym} (%)	9.5 (80.4)
I/σI	31.2 (3.43)
Completeness (%)	100 (100)
Redundancy	14.0 (13.9)
Refinement	
Resolution (Å)	47.49- 2.00
Number of reflections	20,594
R_{work}/R_{free} (%)	18.8/22.3
Number of atoms	
Protein	2205
Ligand	26
Solvent	85
B factors	
Protein (Å^2)	33
Ligand (Å^2)	43
Solvent (Å^2)	34
Rmsd from ideal values	
Bond lengths (Å)	0.010
Bond angles (°)	1.071
Ramachandran plot	
Favored (%)	99
Allowed (%)	100

[a] Values in parentheses are for the highest resolution shell.

Chapter 5
Structure of the Peptidoglycan Sensor Domain of Protein Kinase B

Introduction

Cell wall homeostasis involves biochemical reactions that take place both inside and outside of cells. Peptidoglycan monomers are synthetized intracellularly and polymerized in the periplasm, peptidoglycan degradation and recycling occur in both compartments. Given the necessity of coordinating the state of the cell wall to the growth status of the cell, there is considerable interest in uncovering the various pathways that make this coordination possible. One such pathway that has been investigated in *Mtb* involves the peptidoglycan-sensing, transmembrane protein kinase, PknB.

Unlike the classical bacterial histidine kinases, PknB is a eukaryotic-like Ser/Thr protein kinase (STPK). The *Mtb* genome encodes 11 such STPKs with functions in growth, virulence, and persistence. While over 250 STPK phosphorylation targets have been identified, the environmental cues that regulate STPKs have remained largely unknown (Alber 2009). Nonetheless, PknB has emerged as a strong candidate for coordination of cell-wall remodeling with multiple intracellular pathways. PknB is an essential mycobacterial protein containing an intracellular N-terminal kinase domain followed by a 50-residue linker, a single transmembrane helix and an extracellular domain that consists of four PASTA repeats. Interestingly, in many Gram-positive bacteria with only one STPK, the single STPK has the same architecture as PknB, although the number of PASTA domains can vary.

In mycobacteria, PknB deletion is lethal, knockdown results in elongated cells, and overexpression produces round bulbous cells (Kang, Abbott et al. 2005). PknB phosphorylates proteins involved in peptidoglycan synthesis, the tricarboxylic acid cycle, cell division, and the extracellular stress response, consistent with its proposed role in cell-wall homeostasis (Alber 2009). The extracellular sensor domain of PknB regulates the intracellular kinase activity, but the mechanism of this regulation remains unclear. Based on homology to PASTA domains of high-molecular-weight, type II penicillin-binding proteins and the sequence analysis of multiple PknB homologs, Jones and Dyson hypothesized that individual PASTA domains bind chemically different peptidoglycan stem peptides (Jones and Dyson 2006). Determining the ligands of the PknB sensor domain will provide important clues to the function of this widely conserved kinase and will provide the basis for investigating its regulation.

The best experimental evidence supporting the idea that peptidoglycan fragments directly activate PknB comes from work on the *Bacillus subtilis* homolog, PrkC (Protein Kinase C). In elegant experiments, Shah et al. first demonstrated that *Bacillus* spore reactivation induced by peptidoglycan fragments is PrkC dependent. Complementing a PrkC deletion mutant with a chimeric receptor kinase containing the

PASTA domains of the *Staphylococcus* homolog produced mutant spores that responded to staphylococcal, but not to *Bacillus*, peptidoglycan (Shah, Laaberki et al. 2008). Interestingly, PrkC also regulates a peptidoglycan hydrolase, YocH (Shah and Dworkin 2010). Most recently, Mir et al. demonstrated direct binding of muropeptides to *Mtb* PknB extracellular domain (Mir, Asong et al. 2011).

Because of the preliminary evidence that the PknB sensor domain is regulated by peptidoglycan fragments, the identity of its ligands and the structural basis of ligand binding emerge as the key unanswered questions in this field. Likewise, the roles of individual PASTA domains are unclear. With just 30% average pairwise identity, the four PASTA repeats of PknB are hypothesized to bind different peptidoglycan-derived substrates giving the bacteria ability to respond to peptidoglycan fragments from different species as well as their own. Elucidation of the collective binding repertoire of PASTA repeats and the structural basis of peptidoglycan recognition will define the range of chemical stimuli that can regulate intracellular phosphorylation by the conserved mycobacterial sensor kinase, PknB.

Results

Cloning, expression, and protein purification

Multiple fragments of PknB sensor domain were cloned into Gateway expression vectors as GST-tagged, TEV-cleavable fusions. These constructs included the full sensor domain, each PASTA domain individually, and PASTA 1-2, PASTA 2-4 and PASTA 3-4 truncations. All proteins expressed well, and they were purified by glutathione affinity chromatography.

Structural analysis of ligand-free PknB sensor domain

I first crystallized PASTA domain 4 of PknB by vapor diffusion and collected high-resolution data on these crystals (Table 5.1). However, attempts to solve the structure via molecular replacement failed. Since the crystallization condition contained Zn^{2+} cations, I collected anomalous data and succeeded in phasing the structure using Zn^{2+} anomalous scattering. The subsequent solutions of other PASTA domains were achieved via molecular replacement. First, PASTA 3 was solved in collaboration with Tony Chen, a rotation student, then PASTA 3-4, and then PASTA 2-4. The latter constructs required trimming of the N-terminal sequences to remove all unstructured residues to allow crystallization. Finally, the two-domain, PASTA 1-2 structure completed the coverage of the entire sensor domain (Figures 5.1 and 5.2).

Crystals of PASTA 1-4 were also obtained, but these never diffracted to higher than 8 Å resolution, making their solution impractical. Instead, combination of the PASTA 1-2 structure and the PASTA 2-4 structure aligned on their shared PASTA 2 domain allowed the modeling of the full extracellular domain structure (Figure 5.3A). Mapping of percent conservation scores to the surface of the resulting model highlighted two regions with highest conservation, one around the hinge between domains 1 and 2, and the other on the surface of domain 4 (Figure 5.3B). The conserved region in domain 4 was bound to a molecule of citric acid in the PASTA 2-4

structure (Figure 5.3C), suggesting it is a potential binding site. The bottom of the groove in which the citrate is bound contains partially exposed tryptophan and phenylalanine residues, also typical of protein binding sites (Figure 5.3D). The presence of ordered citrate was not induced by binding of divalent cations, which were not included in crystallization.

Binding and co-crystallization studies

Since the ultimate goal of the project was to determine the binding mode of the PknB sensor domain ligands, crystallization trials and crystal soaking experiments with the available peptidoglycan fragments including MDP, tracheal cytotoxin (TCT), and muramic pentapeptide were performed. However, diffraction data collected on substrate-soaked crystals failed to reveal definitive ligand density. Additionally, since binding to D-Ala-D-Ala and penicillins was proposed as a possible PknB function, binding to Bocillin was tested in a fluorescence polarization assay. These measurements revealed no binding when compared to *E. coli* PBP4, a positive control.

Discussion

In this chapter, I described the atomic structure of the extracellular sensor domain of PknB, an essential mycobacterial signaling kinase. I found that the four PASTA repeats adopt an extended conformation, unlike that found in the PBPs, where the PASTA domains lie side by side. These findings are consistent with the observations made by Barthe et al., who used NMR to interrogate the structures of three two-domain constructs (Barthe, Mukamolova et al. 2010). X-ray structures of the *Staphylococcus aureus* PrkC extracellular domain also show a similar extended architecture (Ruggiero, Squeglia et al. 2011). The mechanism of peptidoglycan binding is yet to be fully elucidated in any of the PknB homologs.

Eukaryotic-like STPKs with extracellular PASTA domains are widespread among Gram-positive bacterial species. The presence of diverse PASTA domains in the individual kinases led to the hypothesis that each repeat contains a unique binding site that recognizes different peptidoglycan fragments (Jones and Dyson 2006). Such diversification could allow the bacterium to respond to a variety of other species growing in the same environment. Sensing that competitors are growing could reactivate cells and stimulate growth.

Structural analysis reported in this chapter points towards a different view of PASTA domains. It may be that a single site in the extracellular domain is sufficient to produce a fully active PknB, the other domains remaining as spacers. Diversification of the other domains could be due to lack of selection rather than the presence of a diversifying evolutionary pressure. This hypothesis is based on the analysis of mycobacterial PknB conservation: there appears to be a single conserved region in the protein in PASTA 4 that carries a signature of a binding site – exposed hydrophobic residues. This idea is further supported by the observation by Sassetti and colleagues that the C-terminal region of the PknB gene does not tolerate transposon insertions (Griffin, Gawronski et al. 2011). Many genes can tolerate transposon insertions near

their C-termini because the resulting deletion of a few amino acids would leave their function intact. Instead, as observed in the structures presented in this chapter, in PknB the C-terminal residues are part of the core of the folded PASTA 4 domain, their essentiality suggests that all of PASTA 4 is essential for PknB function, as well. This idea should be more rigorously tested genetically.

Further evidence for a single binding site in the PknB sensor domain comes from the work of Husson and colleagues, who used chemically-synthesized peptidoglycan fragments to test PASTA 1-4 binding repertoire in a surface plasmon resonance essay (Mir, Asong et al. 2011). In their work, all peptidoglycan fragments that bound to GST-tagged PASTA 1-4 did so with single-step kinetics, consistent with a unique binding site. This study could also explain why we failed to observe co-crystallization of the various peptidoglycan fragments with the PASTA domains. The authors observed a very narrow specificity range where the muramic acid, a tripeptide from L-Ala through meso-DAP, and amidation of carboxyl groups on iso-D-Glu and meso-DAP are each necessary for optimal binding. These requirements collectively rule out TCT (lack of amidation), MDP (lack of DAP), and muramic pentapeptide (lack of muramic acid) as potential binders. Once again, these narrow requirements argue against the hypothesis of a broad PknB ligand range.

The PknB sensor domain consists of four repeats, the last of which contains a hydrophobic groove with a conserved partially exposed tryptophan, a signature of a ligand-binding site. We hypothesize that this is the binding site for peptidoglycan recognition. This hypothesis is consistent with published reports showing that this region of PknB is necessary for function and that small peptidoglycan fragments bind with single-step kinetics. The extended conformation of the PASTA domains, which appears common to all PknB orthologs, suggests that the role of additional PASTA domains may be to determine the distance from the cell surface to the site of PknB-peptidoglycan interaction. The number of PASTA domains found in PknB homologs varies from one in *Streptomyces clavuligerus* to seven in *Eubacterium siraeum*. It is tempting to speculate that difference in the number of PASTA domains may have an effect on bacterial cell shape. Still, the potential contribution of these domains to cell shape remains to be investigated.

Our work defining the first N-terminal folded residues of PknB extracellular domain suggests that this domain is unlikely to be closely tethered to the membrane on its own. Instead, it forms an extended structure that is tethered loosely to the membrane by a six-residue peptide with amino acid sequence G-G-I-T-R-D. This peptide was not included in our crystallographic analysis of PASTA1-2 construct; its boundaries are inferred from transmembrane helix prediction and the first residue of the PASTA1 domain. This peptide was included in the NMR structural investigation of PASTA1-2 domain by Barthe et al. where it extended away from PASTA1 domains and did not appear well-structured (Barthe, Mukamolova et al. 2010). Similar structural features were found during the our lab's investigation of PknD and PknH extracellular domains (Good, Greenstein et al. 2004, Cavazos, Prigozhin et al. 2012), suggesting that ligand binding controls localization and oligomerization of the kinase domain in the membrane

rather than directly signaling via induction of conformational changes across the cytoplasmic membrane.

Methods

Cloning, Expression, and Purification

The original set of PASTA-domain clones was designed by Alisa Moskaleva, an undergraduate reasearcher in the Alber lab. To allow efficient crystallization of longer constructs, the N-termini were later trimmed to only include well-structured residues seen in the single-domain structures. GST-tagged, TEV-cleavable fusion constructs were cloned from *Mtb* genomic DNA via the Gateway system and heterologously expressed in *E. coli* BL-21 Codon Plus cells by autoinduction. Glutathione affinity purification was carried out as described in the previous chapters.

Crystallization and Data Collection and Analysis

PASTA 4 protein was crystallized by vapor diffusion in a condition containing 0.2 M zinc acetate, 0.1 M imidazole pH 8, and 20% PEG 3000. Crystals were cryoprotected with 20% xylitol in mother liquor, mounted, and frozen in liquid nitrogen. Data were collected at cryogenic temperature on Beamline 8.3.1 of the Advanced Light Source (MacDowell, Celestre et al. 2004). Data were reduced using HKL2000 (www.hkl-xray.com), and structure solution, building, and refinement carried out in Phenix (Zwart, Afonine et al. 2008) and Coot (Emsley and Cowtan 2004). Data collection and refinement statistics for all structure presented in this chapter are listed in Table 5.1.

The L512M PASTA 3 domain was constructed by site-directed mutagenesis to introduce a methionine residue for phasing. A leucine residue was chosen because it is nearly isosteric with methionine. Selenomethionine labeling was performed via the dedicated autoinduction protocol (Studier 2005). The resulting protein crystallized from 0.2 M lithium sulfate, 0.1 M sodium acetate pH 4.5, and 50% PEG 400. However, due to prior solution of PASTA 4 structure, the PASTA 3 structure was solved by molecular replacement and not via selenium anomalous diffraction-based methods.

PASTA 3-4 crystallized from a condition containing 0.1 M sodium acetate pH 4.5, 2 M ammonium sulfate. These crystals were cryoprotected in 3 M sodium malonate pH 4.5. Data were collected and processed as described above. PASTA 3 and 4 domains were used as search models in molecular replacement. The PASTA 2-4 construct crystallized in 0.1 M citrate pH 3.5, 25% PEG 3350. The structure, however, was not determined using molecular replacement with the PASTA 3-4 search model due to variability in inter-domain angles. Instead, multiple rounds of molecular replacement with individual PASTA domains as search models yielded phases that eventually produced an interpretable electron density map. Selection of partial solutions based on log likelihood gain scores allowed sequential identification of the six PASTA domains found in the asymmetric unit. The PASTA 1-2 construct was crystallized from 0.1 M CHES pH 9.5, 30% isopropanol, 30% PEG 3350 condition. This structure was solved without complication via molecular replacement.

Soaking and co-crystallization experiments were performed by mixing a 2:1 or 3:1 molar excess of potential ligands into the standard crystallization experiment or by addition of the same amount to the cryoprotectant solution prior to crystal harvesting.

Analysis of the surface conservation was performed in Chimera (Pettersen, Goddard et al. 2004), where the Multalign viewer allowed mapping of conservation scores from protein alignment of 16 orthologous mycobacterial PknB sequences onto the surface of the PASTA 1-4 model. All figures in this chapter were generated in Chimera.

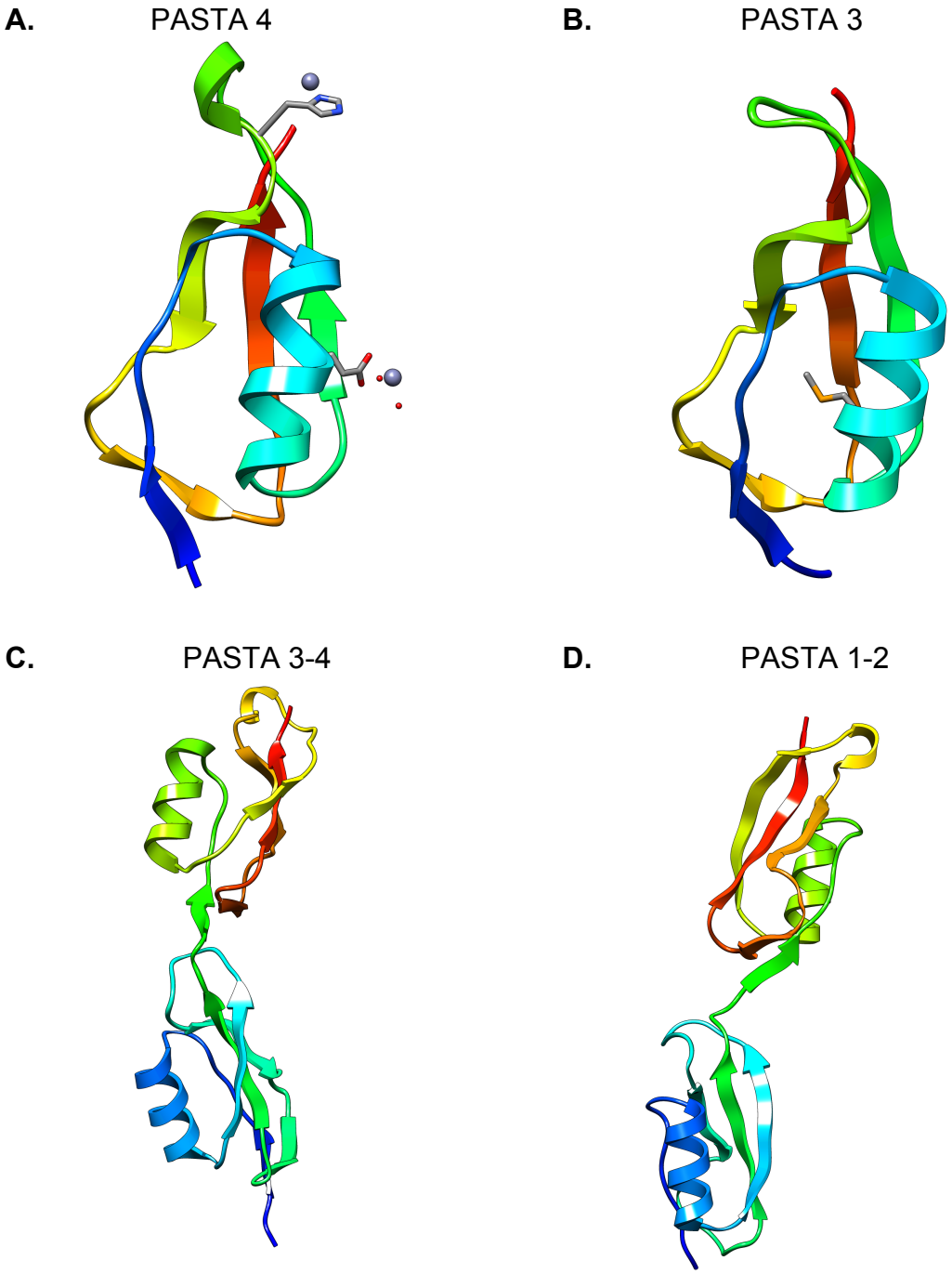


Figure 5.1. Structures of modules of the PknB extracellular domain.
A. PASTA 4 repeat in complex with Zn^{2+} (gray spheres). Zn-bound residues His600 and Glu575 and Zn-coordinating waters (red spheres) are highlighted. The structure was phased using anomalous scattering from the bound Zn^{2+} ions. This and other structures are colored blue to red from N- to C-termini. **B.** PASTA3 domain of PknB. The selenomethionine derivative of the L512M mutant was solved by molecular replacement with the PASTA 4 model; selenomethionine 512 is highlighted. **C.** PASTA 3-4 two-domain structure. **D.** PASTA 1-2 two-domain structure.

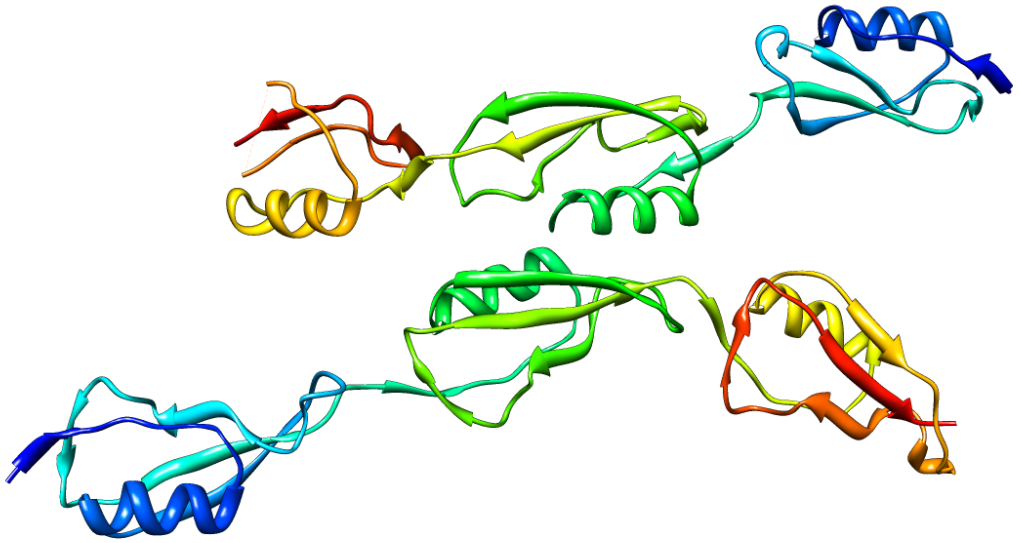


Figure 5.2. Crystal structure of the last three PASTA repeats of the PknB sensor domain.
PASTA 2-4 crystallized with 2 molecules in the asymmetric unit. The structure was determined by multiple rounds of molecular replacement. A loop in PASTA 4 of the molecule shown at the top was disordered due to a close crystal contact. The independent molecules formed extended arrays of repeats.

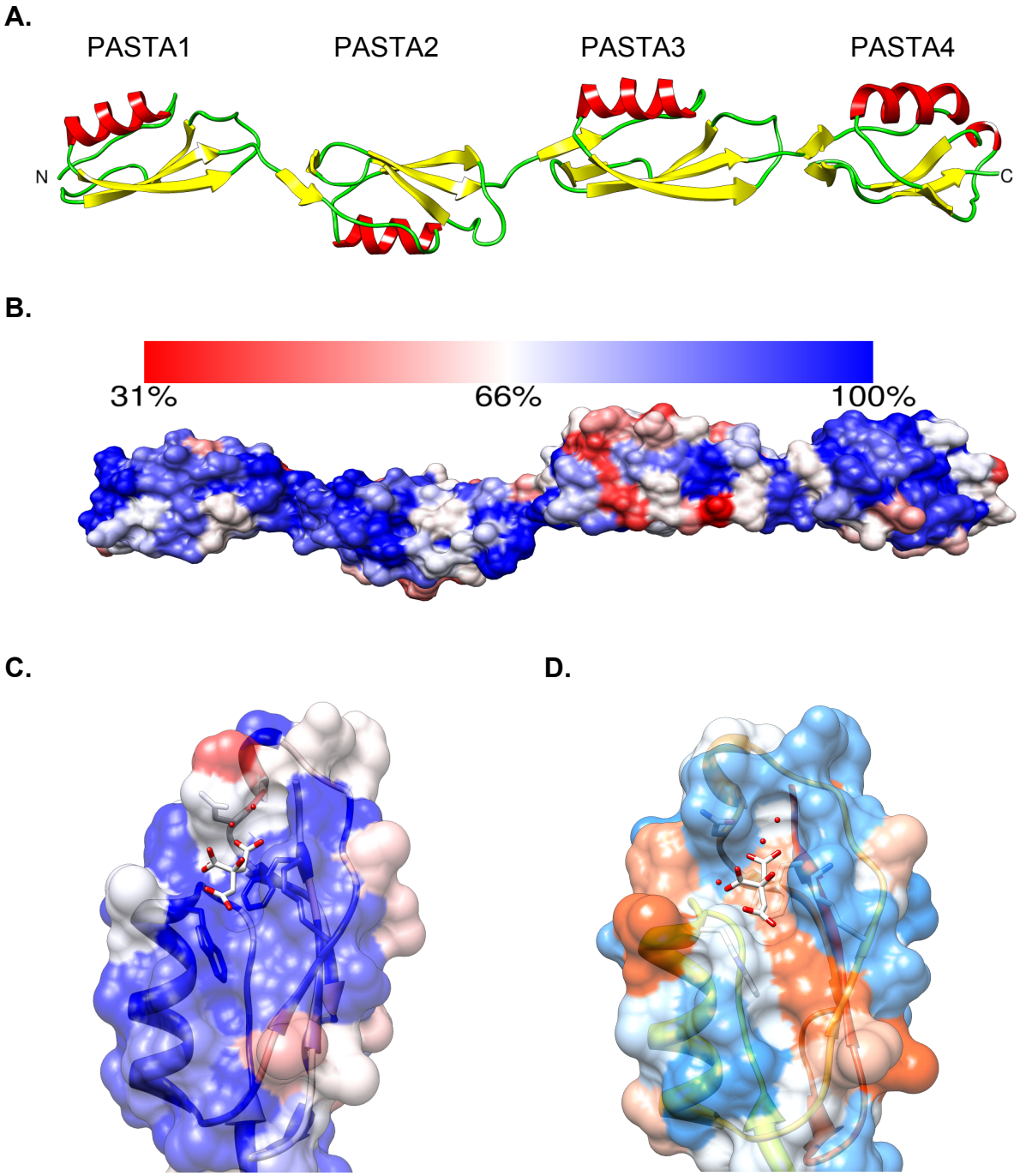


Figure 5.3. The PknB extracellular sensor domain forms a specific arrangement of PASTA repeats.
A. Ribbon diagram of a composite model created by superimposing PASTA 2 of the PASTA 1-2 and PASTA 2-4 structures. **B.** Surface residue conservation in mycobacterial PknB orthologs from least conserved (red) to most conserved (blue). **C.** A conserved region in the PASTA 4 domain bound to a molecule of citric acid (surface colored as in **B.**) **D.** Same surface of PASTA 4 colored by hydrophobicity.

Table 5.1. Data collection and refinement statistics

	PASTA4	PASTA3	PASTA3-4	PASTA1-2	PASTA3-4
Data Collection					
Wavelength (Å)	1.12[a]	1.12	1.12	1.12	1.12
Temp (K)	100	100	100	100	100
Space group	P 6_5 2 2	P 2_1 2_1 2	P 2_1	C 2	P 2_1
Cell parameters					
a / *b* / *c* (Å)	41.6/41.6/122.5	24.4/73.4/32.6	40.4/24.5/64.7	80.7/43.1/43.7	101.3/24.5/73.3
α / β / γ (^{o})	90/90/120	90/90/90	90/98.56/90	90/106.3/90	90/89.35/90
Prot. per a.s.u.	1	1	1	1	2
Resolution (Å)[b]	34.61-1.80 (1.83-1.80)	36.71-2.00 (2.07-2.00)	27.02-1.50 (1.55-1.50)	42.00– 1.80 (1.86-1.80)	41.90 – 2.20 (2.28- 2.20)
R_{sym} (%)	10 (47)	12 (46)	5.9 (25.6)	9.0 (57.1)	13.8 (65.5)
I/σI	12.92 (1.71)	8.04 (2.05)	15.5 (2.8)	14.8 (2.1)	7.8 (1.9)
Refinement					
Resolution (Å)	34.61-1.80	36.71-2.00	27.02-1.50	42.00– 1.80	41.90 – 2.20
R_{work}/R_{free} (%)	23.09/29.08	19.00/25.36	17.07/20.41	18.08/22.33	19.68/25.79
Rms bonds	0.007	0.006	0.017	0.007	0.008
Rms angles	1.022	1.029	1.428	1.131	1.074

[a] Native data set used in refinement. Structure solution achieved from data sets collected at inflection, peak, and high remote wavelength of 1.2829 Å, 1.2824 Å, and 1.2194 Å, respectively.
[b] Values in parentheses are for the highest resolution shell.

References

Alber, T. (2009). Signaling mechanisms of the Mycobacterium tuberculosis receptor Ser/Thr protein kinases. Current Opinion in Structural Biology.

Barthe, P., G. V. Mukamolova, C. Roumestand and M. Cohen-Gonsaud (2010). The structure of PknB extracellular PASTA domain from mycobacterium tuberculosis suggests a ligand-dependent kinase activation. Structure. **18:** 606-615.

Bendtsen, J. D., H. Nielsen, G. von Heijne and S. Brunak (2004). "Improved prediction of signal peptides: SignalP 3.0." J Mol Biol **340**(4): 783-795.

Böth, D., G. Schneider and R. Schnell (2011). Peptidoglycan Remodeling in Mycobacterium tuberculosis: Comparison of Structures and Catalytic Activities of RipA and RipB. Journal of Molecular Biology.

Busso, D., B. Delagoutte-Busso and D. Moras (2005). Construction of a set Gateway-based destination vectors for high-throughput cloning and expression screening in Escherichia coli. Anal Biochem. **343:** 313-321.

Cavazos, A., D. M. Prigozhin and T. Alber (2012). Structure of the Sensor Domain of Mycobacterium tuberculosis PknH Receptor Kinase Reveals a Conserved Binding Cleft. J Mol Biol. **422:** 488-494.

Chamaillard, M., M. Hashimoto, Y. Horie, J. Masumoto, S. Qiu, L. Saab, Y. Ogura, A. Kawasaki, K. Fukase, S. Kusumoto, M. A. Valvano, S. J. Foster, T. W. Mak, G. Nunez and N. Inohara (2003). "An essential role for NOD1 in host recognition of bacterial peptidoglycan containing diaminopimelic acid." Nat Immunol **4**(7): 702-707.

Conover, M. J., J. S. Thompson and G. D. Shockman (1966). Autolytic enzyme of Streptococcus faecalis: release of soluble enzyme from cell walls. Biochemical and Biophysical Research Communications. **23:** 713-719.

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

Deng, L. L., D. E. Humphries, R. D. Arbeit, L. E. Carlton, S. C. Smole and J. D. Carroll (2005). Identification of a novel peptidoglycan hydrolase CwlM in Mycobacterium tuberculosis. Biochim Biophys Acta. **1747:** 57-66.

Eddy, S. R. (2011). "Accelerated Profile HMM Searches." PLoS Comput Biol **7**(10): e1002195.

Emsley, P. and K. Cowtan (2004). Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr. **60:** 2126-2132.

Felsenstein, J. (2005). "PHYLIP (Phylogeny Inference Package) version 3.6." Distributed by the author. Department of Genome Sciences, University of Washington, Seattle.

Finn, R. D., J. Mistry, J. Tate, P. Coggill, A. Heger, J. E. Pollington, O. L. Gavin, P. Gunasekaran, G. Ceric, K. Forslund, L. Holm, E. L. Sonnhammer, S. R. Eddy and A. Bateman (2010). "The Pfam protein families database." Nucleic Acids Res **38**(Database issue): D211-222.

Good, M. C., A. E. Greenstein, T. A. Young, H. L. Ng and T. Alber (2004). "Sensor domain of the Mycobacterium tuberculosis receptor Ser/Thr protein kinase, PknD, forms a highly symmetric beta propeller." J Mol Biol **339**(2): 459-469.

Griffin, J. E., J. D. Gawronski, M. A. Dejesus, T. R. Ioerger, B. J. Akerley and C. M. Sassetti (2011). High-resolution phenotypic profiling defines genes essential for mycobacterial growth and cholesterol catabolism. PLoS Pathog. **7:** e1002251.

Hett, E. C., M. C. Chao, L. L. Deng and E. J. Rubin (2008). A mycobacterial enzyme essential for cell division synergizes with resuscitation-promoting factor. PLoS Pathog. **4:** e1000001.

Hett, E. C., M. C. Chao and E. J. Rubin (2010). Interaction and modulation of two antagonistic cell wall enzymes of mycobacteria. PLoS Pathog. **6:** e1001020.

Hett, E. C., M. C. Chao, A. J. Steyn, S. M. Fortune, L. L. Deng and E. J. Rubin (2007). A partner for the resuscitation-promoting factors of Mycobacterium tuberculosis. Mol Microbiol. **66:** 658-668.

Hett, E. C. and E. J. Rubin (2008). Bacterial growth and cell division: a mycobacterial perspective. Microbiol Mol Biol Rev. **72:** 126-156, table of contents.

Hugonnet, J.-E., L. W. Tremblay, H. I. Boshoff, C. E. Barry and J. S. Blanchard (2009). Meropenem-clavulanate is effective against extensively drug-resistant Mycobacterium tuberculosis. Science. **323:** 1215-1218.

Jones, G. and P. Dyson (2006). Evolution of transmembrane protein kinases implicated in coordinating remodeling of gram-positive peptidoglycan: inside versus outside. J Bacteriol. **188:** 7470-7476.

Josephine, H. R., P. Charlier, C. Davies, R. A. Nicholas and R. F. Pratt (2006). "Reactivity of penicillin-binding proteins with peptidoglycan-mimetic beta-lactams: what's wrong with these enzymes?" Biochemistry **45**(51): 15873-15883.

Kana, B. D., B. G. Gordhan, K. J. Downing, N. Sung, G. Vostroktunova, E. E. Machowski, L. Tsenova, M. Young, A. Kaprelyants, G. Kaplan and V. Mizrahi (2008). The resuscitation-promoting factors of Mycobacterium tuberculosis are required for virulence and resuscitation from dormancy but are collectively dispensable for growth in vitro. Mol Microbiol. **67:** 672-684.

Kang, C.-M., D. W. Abbott, S. T. Park, C. C. Dascher, L. C. Cantley and R. N. Husson (2005). The Mycobacterium tuberculosis serine/threonine kinases PknA and PknB: substrate identification and regulation of cell shape. Genes & Development. **19:** 1692-1704.

Kilburn, J. O. and G. K. Best (1977). Characterization of autolysins from Mycobacterium smegmatis. J Bacteriol. **129:** 750-755.

Koch, A. L. (2003). "Bacterial wall as target for attack: past, present, and future research." Clin Microbiol Rev **16**(4): 673-687.

Krogh, A., B. Larsson, G. von Heijne and E. L. Sonnhammer (2001). "Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes." J Mol Biol **305**(3): 567-580.

Layec, S., B. Decaris and N. Leblond-Bourget (2008). Diversity of Firmicutes peptidoglycan hydrolases and specificities of those involved in daughter cell separation. Res Microbiol. **159:** 507-515.

Lechat, P., L. Hummel, S. Rousseau and I. Moszer (2008). "GenoList: an integrated environment for comparative analysis of microbial genomes." Nucleic Acids Res **36**(Database issue): D469-474.

MacDowell, A. A., R. S. Celestre, M. Howells, W. McKinney, J. Krupnick, D. Cambie, E. E. Domning, R. M. Duarte, N. Kelez, D. W. Plate, C. W. Cork, T. N. Earnest, J. Dickert, G. Meigs, C. Ralston, J. M. Holton, T. Alber, J. M. Berger, D. A. Agard and H. A. Padmore (2004). Suite of three protein crystallography beamlines with single superconducting bend magnet as the source. J Synchrotron Radiat. **11:** 447-455.

Mir, M., J. Asong, X. Li, J. Cardot, G.-J. Boons and R. N. Husson (2011). The extracytoplasmic domain of the Mycobacterium tuberculosis Ser/Thr kinase PknB binds specific muropeptides and is required for PknB localization. PLoS Pathog. **7:** e1002182.

Mukamolova, G. V., A. S. Kaprelyants, D. I. Young, M. Young and D. B. Kell (1998). A bacterial cytokine. Proc Natl Acad Sci USA. **95:** 8916-8921.

Mukamolova, G. V., A. G. Murzin, E. G. Salina, G. R. Demina, D. B. Kell, A. S. Kaprelyants and M. Young (2006). Muralytic activity of Micrococcus luteus Rpf and its relationship to physiological activity in promoting bacterial growth and resuscitation. Mol Microbiol. **59:** 84-98.

Navarre, W. W. and O. Schneewind (1999). Surface proteins of gram-positive bacteria and mechanisms of their targeting to the cell wall envelope. Microbiol Mol Biol Rev. **63:** 174-229.

Pettersen, E. F., T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng and T. E. Ferrin (2004). UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem. **25:** 1605-1612.

Rogers, H. J., C. Taylor, S. Rayter and J. B. Ward (1984). "Purification and properties of autolytic endo-beta-N-acetylglucosaminidase and the N-acetylmuramyl-L-alanine amidase from Bacillus subtilis strain 168." J Gen Microbiol **130**(9): 2395-2402.

Romeis, T. and J. V. Holtje (1994). "Specific interaction of penicillin-binding proteins 3 and 7/8 with soluble lytic transglycosylase in Escherichia coli." J Biol Chem **269**(34): 21603-21607.

Royet, J., D. Gupta and R. Dziarski (2011). "Peptidoglycan recognition proteins: modulators of the microbiome and inflammation." Nat Rev Immunol **11**(12): 837-851.

Ruggiero, A., F. Squeglia, D. Marasco, R. Marchetti, A. Molinaro and R. Berisio (2011). "X-ray structural studies of the entire extracellular region of the serine/threonine kinase PrkC from Staphylococcus aureus." Biochem J **435**(1): 33-41.

Russell, D. G., C. E. Barry and J. L. Flynn (2010). Tuberculosis: what we don't know can, and does, hurt us. Science. **328:** 852-856.

Schleifer, K. H. and O. Kandler (1972). Peptidoglycan types of bacterial cell walls and their taxonomic implications. Bacteriol Rev. **36:** 407-477.

Shah, I. M. and J. Dworkin (2010). Induction and regulation of a secreted peptidoglycan hydrolase by a membrane Ser/Thr kinase that detects muropeptides. Mol Microbiol.

Shah, I. M., M.-H. Laaberki, D. L. Popham and J. Dworkin (2008). A eukaryotic-like Ser/Thr kinase signals bacteria to exit dormancy in response to peptidoglycan fragments. Cell. **135:** 486-496.

Smith, T. J., S. A. Blackman and S. J. Foster (2000). "Autolysins of Bacillus subtilis: multiple enzymes with multiple functions." Microbiology **146 (Pt 2)**: 249-262.

Studier, F. W. (2005). Protein production by auto-induction in high density shaking cultures. Protein Expr Purif. **41:** 207-234.

Tokuda, H. (2009). "Biogenesis of outer membranes in Gram-negative bacteria." Biosci Biotechnol Biochem **73**(3): 465-473.

Uehara, T., K. R. Parzych, T. Dinh and T. G. Bernhardt (2010). Daughter cell separation is controlled by cytokinetic ring-activated cell wall hydrolysis. EMBO J. **29:** 1412-1422.

van Heijenoort, J. (2007). "Lipid intermediates in the biosynthesis of bacterial peptidoglycan." Microbiol Mol Biol Rev **71**(4): 620-635.

van Heijenoort, J. (2011). Peptidoglycan Hydrolases of Escherichia coli. Microbiol Mol Biol Rev. **75:** 636-663.

Vollmer, W., B. Joris, P. Charlier and S. Foster (2008). Bacterial peptidoglycan (murein) hydrolases. FEMS Microbiology Reviews. **32:** 259-286.

Wyckoff, T. J., J. A. Taylor and N. R. Salama (2012). Beyond growth: novel functions for bacterial cell wall hydrolases. Trends Microbiol.

Yang, D. C., N. T. Peters, K. R. Parzych, T. Uehara, M. Markovski and T. G. Bernhardt (2011). "An ATP-binding cassette transporter-like complex governs cell-wall hydrolysis at the bacterial cytokinetic ring." Proc Natl Acad Sci U S A **108**(45): E1052-1060.

Young, F. E. (1966). "Autolytic enzyme associated with cell walls of Bacillus subtilis." J Biol Chem **241**(15): 3462-3467.

Zhao, G., T. I. Meier, S. D. Kahl, K. R. Gee and L. C. Blaszczak (1999). BOCILLIN FL, a sensitive and commercially available reagent for detection of penicillin-binding proteins. Antimicrob Agents Chemother. **43:** 1124-1128.

Zwart, P. H., P. V. Afonine, R. W. Grosse-Kunstleve, L.-W. Hung, T. R. Ioerger, A. J. McCoy, E. McKee, N. W. Moriarty, R. J. Read, J. C. Sacchettini, N. K. Sauter, L. C. Storoni, T. C. Terwilliger and P. D. Adams (2008). Automated structure solution with the PHENIX suite. Methods Mol Biol. **426:** 419-435.