

RESEARCH ARTICLE

Subfamily-Specific Adaptations in the Structures of Two Penicillin-Binding Proteins from *Mycobacterium tuberculosis*

Daniil M. Prigozhin^{1*}, Inna V. Krieger^{2,3}, John P. Huizar¹, Daniela Mavrici¹, Geoffrey S. Waldo², Li-Wei Hung², James C. Sacchettini³, Thomas C. Terwilliger², Tom Alber¹

1. Department of Molecular and Cell Biology and California Institute for Quantitative Biosciences, University of California, Berkeley, California, 94720, United States of America, 2. Los Alamos National Laboratory, Los Alamos, New Mexico, 87545, United States of America, 3. Department of Biochemistry and Biophysics, Texas A & M University, College Station, Texas, 77843, United States of America

*daniil.prigozhin@gmail.com



CrossMark
click for updates

OPEN ACCESS

Citation: Prigozhin DM, Krieger IV, Huizar JP, Mavrici D, Waldo GS, et al. (2014) Subfamily-Specific Adaptations in the Structures of Two Penicillin-Binding Proteins from *Mycobacterium tuberculosis*. PLoS ONE 9(12): e116249. doi:10.1371/journal.pone.0116249

Editor: Claudine Mayer, Institut Pasteur, France

Received: August 17, 2014

Accepted: December 5, 2014

Published: December 31, 2014

This is an open-access article, free of all copyright, and may be freely reproduced, distributed, transmitted, modified, built upon, or otherwise used by anyone for any lawful purpose. The work is made available under the Creative Commons CC0 public domain dedication.

Data Availability: The authors confirm that all data underlying the findings are fully available without restriction. The atomic coordinates and structure factors for the Rv2911-PMSF and Rv3330-meropenem complexes have been deposited to the Protein Data Bank under accession codes 4P0M and 4PPR, respectively.

Funding: This work has been supported by the National Institute of Allergy and Infectious Diseases (NIAID)/National Institutes of Health (NIH) grant A1095208. The Advanced Light Source is supported by the Director, Office of Science, Office of Basic Energy Sciences, of the United States Department of Energy under Contract No. DE-AC02-05CH11231. The BCSB is supported in part by the NIH, National Institute of General Medical Sciences, and the Howard Hughes Medical Institute. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: GSW declares competing financial interests. The folding reporter GFP and split-GFP technologies are the subject of domestic and foreign patent applications by Los Alamos National Laboratories on behalf of the Department of Energy and LANS, L. L. C. Folding reporter GFP

Abstract

Beta-lactam antibiotics target penicillin-binding proteins including several enzyme classes essential for bacterial cell-wall homeostasis. To better understand the functional and inhibitor-binding specificities of penicillin-binding proteins from the pathogen, *Mycobacterium tuberculosis*, we carried out structural and phylogenetic analysis of two predicted D,D-carboxypeptidases, Rv2911 and Rv3330. Optimization of Rv2911 for crystallization using directed evolution and the GFP folding reporter method yielded a soluble quadruple mutant. Structures of optimized Rv2911 bound to phenylmethylsulfonyl fluoride and Rv3330 bound to meropenem show that, in contrast to the nonspecific inhibitor, meropenem forms an extended interaction with the enzyme along a conserved surface. Phylogenetic analysis shows that Rv2911 and Rv3330 belong to different clades that emerged in Actinobacteria and are not represented in model organisms such as *Escherichia coli* and *Bacillus subtilis*. Clade-specific adaptations allow these enzymes to fulfill distinct physiological roles despite strict conservation of core catalytic residues. The characteristic differences include potential protein-protein interaction surfaces and specificity-determining residues surrounding the catalytic site. Overall, these structural insights lay the groundwork to develop improved beta-lactam therapeutics for tuberculosis.

is available by MTA to non-profit academic requesters. GFP folding reporter vectors used in this article are covered under patent US6448087 (Method for determining and modifying protein/peptide solubility). Full list of patents related to split-GFP and GFP-reporter technologies held by Los Alamos National Lab can be found at (<http://www.lanl.gov/source/projects/gfp/licensing.shtml>). This does not alter the authors' adherence to the PLOS ONE policies regarding sharing of materials.

Introduction

Mycobacterium tuberculosis (*Mtb*), the etiologic agent of tuberculosis (TB), presents major threats to public health worldwide [1]. Emergence of *Mtb* strains resistant to the front-line and second-line drugs underscores the necessity of developing new therapeutics [2]. The mycobacterial cell wall, a complex structure responsible for mediating *Mtb* interactions with the environment, is a validated drug target [3]. The cell wall consists of several layers: peptidoglycan, arabinogalactan, mycolic acids, and the polysaccharide capsule [4]. Key anti-TB drugs target the biosynthesis of arabinogalactan and mycolic acids [3]. Beta-lactams, antibiotics that target peptidoglycan biosynthesis and are effective against diverse bacterial pathogens, have not yet been deployed against *Mtb* due to the presence of an efficient beta-lactamase. Nonetheless, a combination of meropenem, a beta-lactam, and clavulanate, a beta-lactamase inhibitor, is being tested as a treatment for both active and latent TB [5]. Genetic, biochemical, and structural characterization of enzymes involved in peptidoglycan homeostasis is critical for development of this and other new therapeutics.

Peptidoglycan, which forms the membrane-proximal layer of the mycobacterial cell wall, is composed of repeating disaccharide-pentapeptide units that form polysaccharide chains covalently crosslinked through the stem peptides [6]. In *Mtb*, the disaccharide is N-acetyl glucosamine N-glycolyl or N-acetyl muramic acid [7], and the pentapeptide attached to the lactate moiety of muramic acid is L-Ala-iso-D-Glu-meso-DAP-D-Ala-D-Ala (DAP stands for diaminopimelic acid) that is typically amidated on the free carboxyl groups of iso-D-Glu and DAP [8]. Beta-lactams inhibit several enzyme classes that catalyze different reactions on peptidoglycan substrates. These penicillin-binding proteins (PBPs) include D,D-transpeptidases, which crosslink the third and fourth amino acids in the pentapeptide stems (3–4 crosslinks) [9]. In addition to the synthetic enzymes, PBPs include hydrolytic D,D-carboxypeptidases and D,D-endopeptidases.

Beta-lactams, including meropenem, mimic the terminal D-Ala-D-Ala dipeptide of the stem peptide and act as suicide inhibitors of penicillin-binding proteins (Fig. 1A). The reaction mechanisms of biosynthetic and hydrolytic penicillin-binding proteins share the first step. Attack of a serine nucleophile on the peptide bond between the two D-alanines of the donor stem leads to acylation of the enzyme by the penultimate D-Ala (position 4 in the stem peptide) and the release of the terminal D-Ala (Fig. 1B). Transpeptidases use a D-Ala-like distal end of the DAP residue (position 3 in the stem peptide) of the acceptor stem as a nucleophile to resolve the covalent intermediate, creating a 4-3 crosslink. In contrast, carboxypeptidases use a water molecule to resolve the acyl-enzyme intermediate, converting a pentapeptide stem to a tetrapeptide. Tetrapeptide formation not only regulates the degree of peptidoglycan crosslinking, but also provides substrate for the synthesis of 3-3 crosslinks essential for various physiological adaptations [10]. Compared to many beta-lactams, carbapenems appear to have a broader specificity. Meropenem, for example, is effective against penicillin-binding proteins as well as L,D-transpeptidases [11], an unrelated family of enzymes that

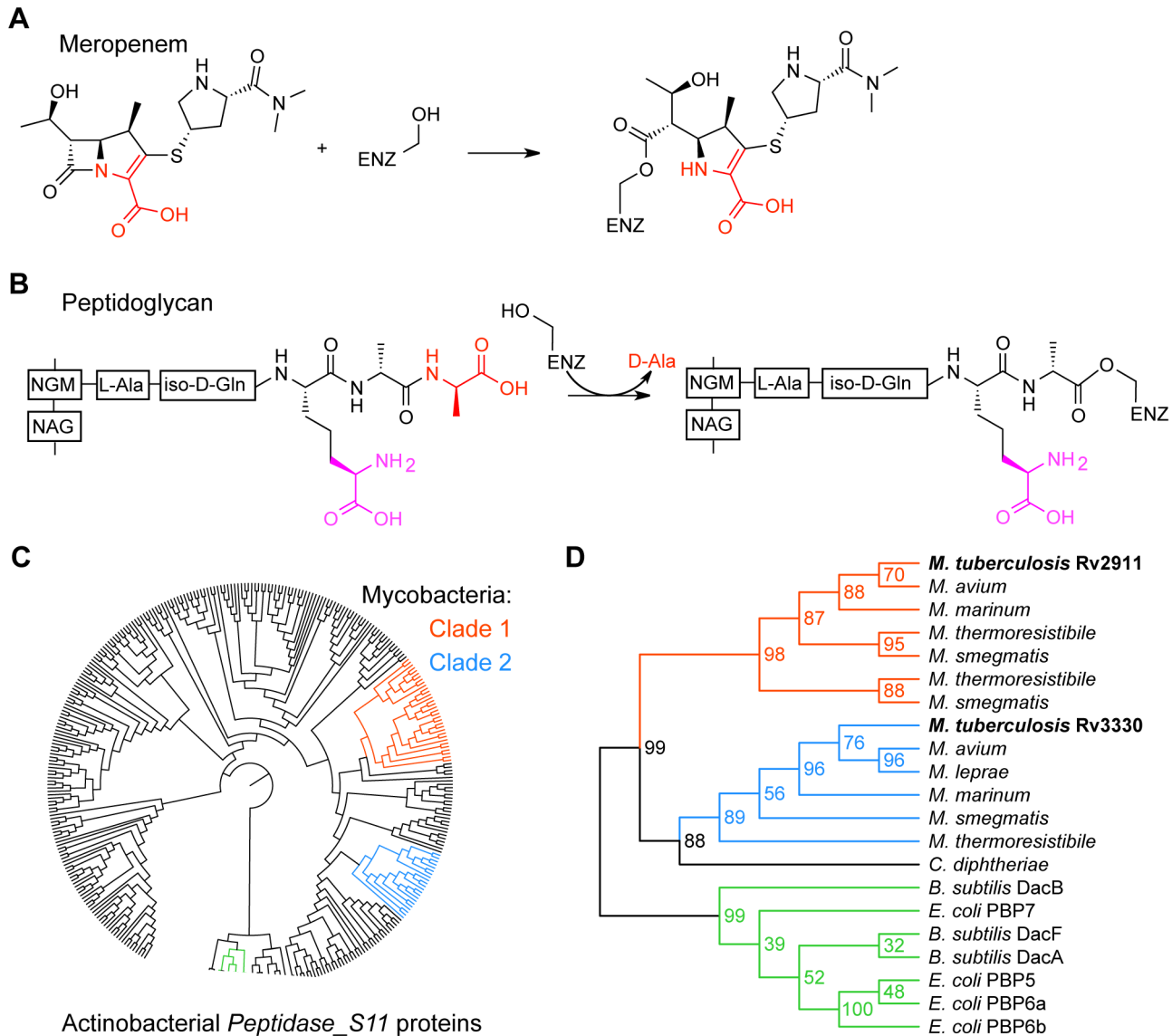


Fig. 1. Activity and phylogeny of Rv2911 and Rv3330. A. The beta-lactam, meropenem, contains a D-Ala-like group (red) and irreversibly acylates penicillin-binding proteins, including Rv2911 and Rv3330. B. Enzyme acylation by peptidoglycan stem peptide, the first step in peptidoglycan carboxy- and trans-peptidation, highlighting the similarity of the terminal D-Ala leaving group (red) and the D-Ala-like acceptor moiety of DAP (purple). C. Neighbor-joining tree of *Peptidase_S11* family proteins of Actinobacteria fall into several clades. Of the two mycobacterial clades, Clade 1 (orange) contains Rv2911, and Clade 2 (blue) contains Rv3330. *E. coli* and *B. subtilis* *Peptidase_S11* enzymes (green) form an out-group with several actinobacterial sequences. D. Maximum-likelihood tree of *Peptidase_S11* proteins from select mycobacterial species, *Corynebacterium diphtheriae*, *E. coli*, and *B. subtilis*. High bootstrapping values confirm the separation of mycobacterial sequences into two distinct clades.

doi:10.1371/journal.pone.0116249.g001

form the 3-3 crosslinks. Investigation of changes in *Mtb* peptidoglycan structure in response to meropenem treatment suggested that the inhibitor not only blocks the transpeptidases that synthesize the peptidoglycan crosslinks, but also the carboxypeptidases, which trim the terminal D-Ala to form tetrapeptides [12].

Of the 11 predicted *Mtb* penicillin-binding proteins [9], only the synthetic enzyme PbpA (Rv0016c) has been structurally characterized [13]. The two penicillin-binding proteins in the *Peptidase_S11* Pfam family, Rv2911 (DacB2, UniProtKB Q7D6F2) and Rv3330 (DacB1, UniProtKB O53380), have recently been investigated for their potential role as major D,D-carboxypeptidases of *Mtb* [12]. Rv2911 is detected in *Mtb* culture filtrates [14]. Deletion and overexpression of *rv2911* are tolerated in culture [15]. Rv3330 and Rv2911 both demonstrate specificity for DAP-containing stem peptides [12].

To define the features that mediate inhibitor binding, we determined the crystal structures of a quadruple-mutant of Rv2911 in complex with a non-specific inhibitor phenylmethylsulfonyl fluoride (PMSF) and Rv3330 in complex with meropenem. The Rv3330-meropenem complex highlights an extended interaction surface that is distinct from that engaged by a donor peptide. Neither enzyme shows a strong binding affinity for peptidoglycan. In addition, the Rv2911 and Rv3330 structures reveal subfamily-specific conserved surfaces that distinguish these enzymes and are consistent with distinct functions in peptidoglycan homeostasis.

Results

A search of the *Mtb* genome for *Peptidase_S11* Pfam family proteins (PF00768) yielded two hits: Rv2911 and Rv3330. Similar searches identified penicillin-binding proteins 5, 6a, 6b, and 7 in the model Gram-negative bacterium, *E. coli*, and DacA, DacB, and DacF in the model Gram-positive bacterium, *B. subtilis*. The neighbor-joining tree of unique actinobacterial *Peptidase_S11* sequences from the Pfam database showed at least five separate clades, one of which contained the *E. coli* and *B. subtilis* proteins included in the analysis. All mycobacterial sequences split into two clades, Clade 1 containing Rv2911 and Clade 2 containing Rv3330 peptidase (Fig. 1C). Bootstrapping analysis of the maximum likelihood tree of select mycobacterial species showed high confidence values for the two clades (Fig. 1D). Orthologs of both proteins were found in most sequenced mycobacteria. One exception was *M. leprae*, where the *rv2911* gene is inactivated. *M. smegmatis* and several other mycobacteria not only encode an *rv3330* ortholog, but also contain duplications of the *rv2911* gene. Based on this analysis, we concluded that the duplication event that allowed for diversification and specialization of Rv2911 and Rv3330 likely occurred in a common ancestor of all mycobacteria.

To investigate the biochemical and structural features of these proteins, we cloned, expressed in *E. coli*, and purified Rv2911 and Rv3330 peptidoglycan peptidases. Sequence analysis indicated that the two proteins share a similar architecture, containing predicted signal peptides at their N-termini and lacking the characteristic C-terminal domain of *E. coli* PBP5 (Fig. 2). A unique feature of Rv3330 is the presence of a single transmembrane helix at the C-terminus. The Rv3330 peptidase is thus predicted to be membrane associated, while Rv2911 is a

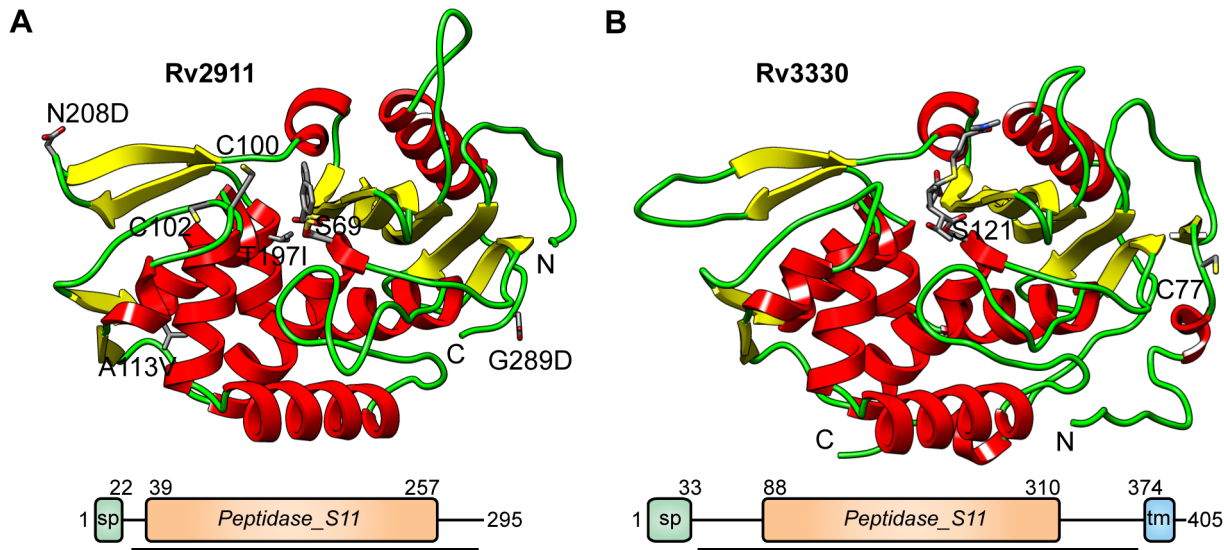


Fig. 2. Structures and domain diagrams of Rv2911_{opt} and Rv3330 enzymes. A. Rv2911_{opt} ribbon diagram showing the inhibitor, PMSF, bound to the catalytic Ser69 (sticks). Four mutations obtained during split-GFP optimization are highlighted along with two cysteines that participate in intermolecular disulfide bonds in the crystal. B. Ribbon diagram of Rv3330 structure showing meropenem bound to the catalytic Ser121 (sticks). Cys77, which forms an intermolecular disulfide bond in the crystals occurs in an N-terminal extension absent from Rv2911. In the domain diagrams, SP denotes the signal peptide, TM represents the trans-membrane helix, and numbers indicate sequence positions of domain borders. Sequences present in the crystallization constructs are underlined.

doi:10.1371/journal.pone.0116249.g002

periplasmic protein, consistent with its documented presence in *Mtb* culture filtrates [14].

Because the wild-type Rv2911 was poorly soluble and failed to crystallize, we undertook extensive construct optimization using the GFP folding reporter method [16]. Several rounds of directed evolution to improve solubility yielded an optimized gene sequence carrying six point mutations. Two of these mutations were in the predicted signal peptide region, which was truncated in further optimization. The resulting A113V, T197I, N208D, G289D quadruple-mutant protein was named Rv2911_{opt}.

The Rv2911_{opt}-PMSF structure was determined at 2.0 Å resolution using multiwavelength anomalous dispersion of Se-labeled enzyme. This structure was used as the search model to determine the Rv3330-meropenem structure, also at 2.0 Å resolution (Table 1). The structures showed similar overall folds containing two subdomains (Fig. 2). In the first subdomain, the β-sheet core is formed by strands from both N- and C-terminal sequences. The second subdomain is α-helical and contains two peripheral two-stranded β-sheets. Structural alignment of the two peptidases revealed a root-mean-square deviation (RMSD) of 0.76 Å over 235 Cα atoms, despite just 46% residue identity over the aligning region. The extended N-terminal sequences of Rv3330 add an extra strand to the β-sheet and a short peripheral helix together with a long coil. The Rv3330 C-terminal extension contains another short helix and extends away from the protein to make a crystal contact. Remarkably, crystals of both proteins were stabilized by disulfides across

Table 1. Data collection and refinement statistics.

	Rv2911-PMSF	Rv3330-Meropenem
Data Collection		
Wavelength (Å)	0.980	1.116
Temperature (K)	100	100
Space group	H 32	P 41 2 2
Cell parameters		
<i>a b c</i> (Å)	75.13 75.13 230.44	76.291 76.291 99.978
$\alpha \beta \gamma$ (°)	90 90 120	90 90 90
Copies per a.s.u.	1	1
Resolution (Å) ^a	43.13-2.00 (2.07-2.00)	50.0 – 2.00 (2.07-2.00)
R _{sym} (%)	16 (56.7)	9.5 (80.4)
I/σI	7.7 (3.8)	19.9 (3.4)
Completeness (%)	99.8 (100)	100 (100)
Redundancy	14.9 (11.2)	14.0 (13.9)
Refinement		
Resolution (Å)	43.13 – 2.00	47.48-2.00
Number of reflections	32,821	20,595
R _{work} /R _{free} (%)	17.3/19.9	18.35/21.04
Number of atoms		
Protein	1929	2214
Ligand	16	26
Solvent	150	144
B factors		
Protein (Å ²)	23.14	33.10
Ligand (Å ²)	22.7	43.10
Solvent (Å ²)	29.74	37.70
RMSD		
Bond lengths (Å)	0.007	0.012
Bond angles (°)	1.1	0.99
Ramachandran plot		
Favored (%)	97	100
Outliers (%)	0	0
PDB ID	4P0M	4PPR

^aValues in parentheses refer to the high-resolution shell.

doi:10.1371/journal.pone.0116249.t001

two-fold axes. Cys100 of Rv2911_{opt} was found in a disulfide with Cys102 from a symmetry-related molecule. In the Rv3330 crystals, Cys77 residues from adjacent molecules formed a disulfide across a two-fold axis. Of the four mutations in Rv2911_{opt}, two, A113V and T197I, were buried, while the other two, N208D and G289D, were exposed and participated in a crystal contacts.

Both enzymes crystallized with covalently bound active-site inhibitors. Rv2911_{opt} was bound to a molecule of the nonspecific inhibitor, PMSF, carried over from the purification. Rv3330 was bound to meropenem, a beta-lactam

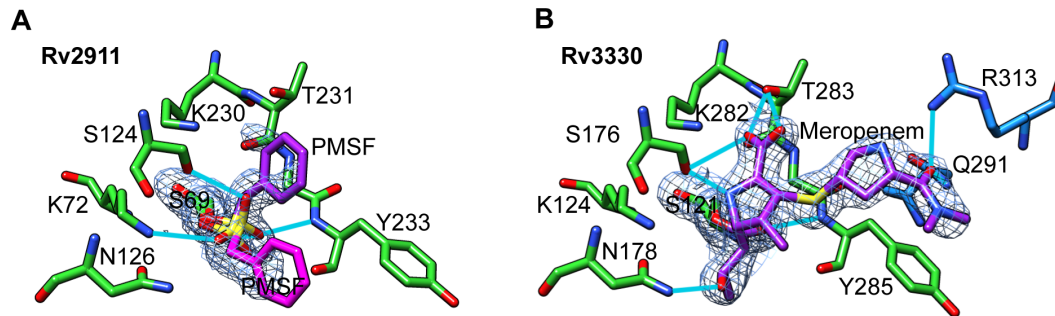


Fig. 3. Conserved arrangement of catalytic-site residues of Rv2911 and Rv3330 peptidases. A. Rv2911_{opt}-PMSF. Electron-density (2Fo-Fc, 2.0-Å resolution, 1σ) shows the inhibitor adopts two major conformations (purple and magenta). B. Rv3330-meropenem. Electron-density (2Fo-Fc, 2.0-Å resolution, 1σ) supports a unique model with extended interactions between the inhibitor (purple) and the enzyme. Common catalytic residues are colored green and Clade 2 conserved residues in Rv3330 are colored blue.

doi:10.1371/journal.pone.0116249.g003

antibiotic that was added prior to crystallization (Fig. 3). The active sites of both proteins contain the three typical penicillin-binding protein motifs: SXXK, SXN, and KTG. In both structures, the inhibitors were covalently linked to the serine of the first motif (STIK in Rv2911, SVIK in Rv3330). Residues of the other two motifs (SGN and KTG_Y in both proteins) interacted with the bound ligands. A common feature of both enzyme-inhibitor complexes is the interaction of either the carboxyl oxygen of meropenem or one of the sulfonyl group oxygens of PMSF with the oxyanion hole formed by the main-chain amides of the catalytic serine and the tyrosine of the third motif.

In the Rv2911_{opt}-PMSF structure the conformation of the inhibitor was not unique, and the electron density for the phenyl group of the ligand suggested at least two alternate conformations. Neither conformation appeared stabilized by additional enzyme-inhibitor interactions (Fig. 3A). In contrast, upon binding to Rv3330, the meropenem molecule assumed a single conformation, as demonstrated by a well-defined electron density envelope extending away from the active site (Fig. 3B). This unique conformation is stabilized by several key interactions. First, the alanine-like moiety within meropenem is hydrogen bonded to Thr283 through the carboxyl and to Ser176 through both the carboxyl and amide groups. This interaction mimics the contacts made by the D-alanine leaving group of the natural pentapeptide substrate. Second, the sulfur of the thioester group packs against the main-chain atoms of Tyr284, with distances of approximately 3.7 Å to the nitrogen, oxygen, and the C β atoms. Third, the π -system of the distal carbamoyl moiety participates in a stacking interaction with the phenyl ring of Tyr284. Finally, the oxygen of the carbamoyl group is hydrogen bonded to Asn291 and Arg313, two residues outside of the canonical active site motifs that are nonetheless universally conserved among mycobacterial orthologs of Rv3330.

To determine whether the catalytic domains bind polymerized peptidoglycan, we performed pull-down experiments with peptidoglycan from *B. subtilis*, which has a chemotype similar to peptidoglycan from *Mtb*. In contrast to hen egg-white lysozyme, a positive control that was retained in the peptidoglycan fraction, both

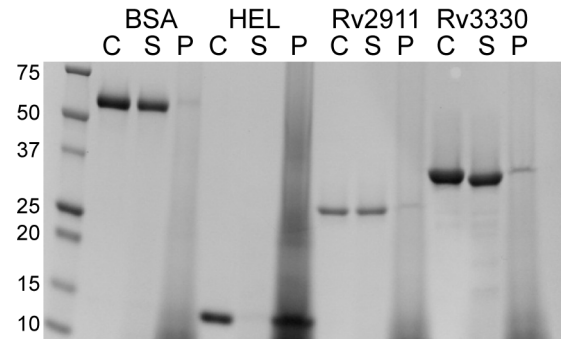


Fig. 4. Rv2911 and Rv3330 catalytic domains bind weakly to peptidoglycan *in vitro*. SDS-PAGE of control (C), supernatant (S), and peptidoglycan fractions (P) following pull-down with *B. subtilis* peptidoglycan. Hen egg lysozyme (HEL) and bovine serum albumin (BSA) were used as positive and negative controls, respectively. The left lane contains molecular marker in kDa units.

doi:10.1371/journal.pone.0116249.g004

Rv2911 and Rv3330 catalytic domains were found primarily in the soluble fraction (Fig. 4). This weak binding to peptidoglycan may reflect a requirement for additional protein-protein or protein-membrane interactions to properly localize and orient these peptidases relative to the substrate.

To determine the subset of residues that underpin penicillin-binding protein specialization, we constructed multiple sequence alignments of Clade 1 and Clade 2 proteins and mapped the conservation scores onto the protein surfaces (Fig. 5, A and C). To aid the interpretation of the observed differences, we overlaid peptidoglycan-mimicking ligands from *E. coli* PBP6 (PDB 3ITB) and *B. subtilis* PBP4a (DacC, PDB 2J9P) structures onto the *Mtb* peptidase surfaces (Fig. 5, B and D). Residues predicted to distinguish the stem-peptide meso-DAP, including Thr234, Leu166, and Asp99 (Rv2911 sequence numbering), are conserved across the two clades along with the catalytic center residues.

In contrast, clade-specific surface residues surround the catalytic centers of both enzymes. In Rv2911, these characteristic residues include the 100-CysAsnCys-102 triad on one edge of the catalytic site, as well as Gln216 on an adjacent side. The equivalent motifs in Rv3330 are 152-GlyThr-153 and Asp268, respectively. In addition, the extended meropenem-binding site of Rv3330 is distinct from that occupied by the donor peptide (Fig. 5D). Away from the active site of Rv3330, a surface region that centers on the loop containing 116-ArgHisArg-118 is conserved only in Clade 2 (Fig. 5C). This conserved surface measures approximately 300 \AA^2 and participates in a crystal contact that includes the intermolecular disulfide bond. The conservation of this area could indicate an interaction site with other proteins or with cell-wall constituents.

Discussion

Investigation of mycobacterial peptidoglycan biology can potentially lead to improvements to current anti-TB therapies. The recent suggestion that a

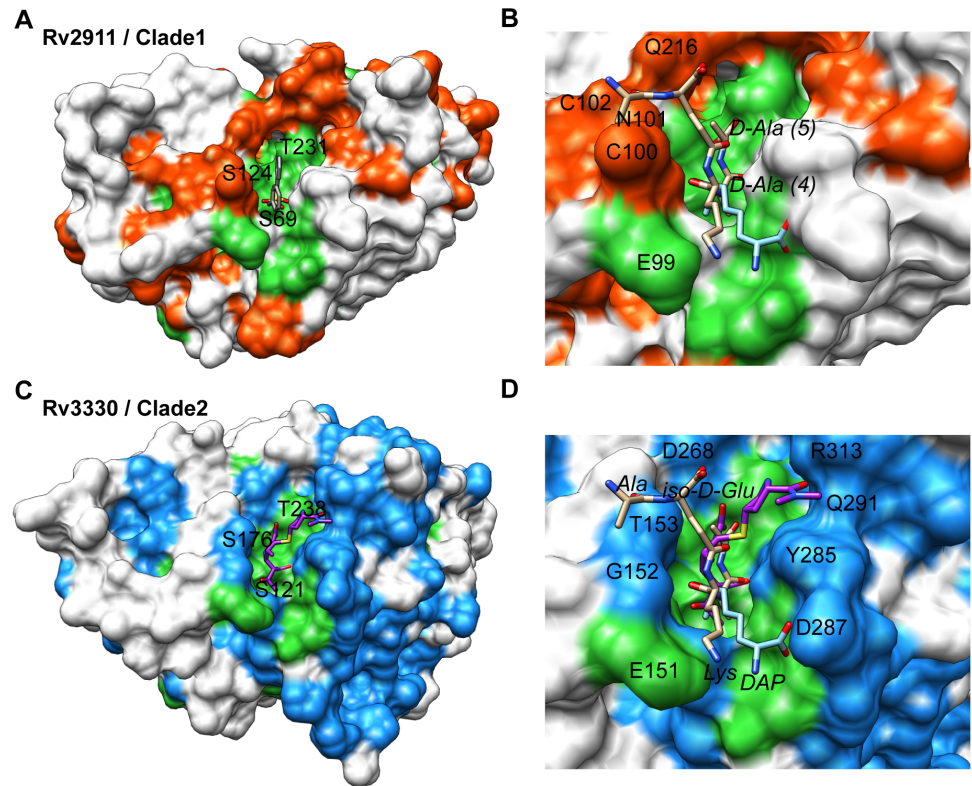


Fig. 5. Different conserved surfaces surround the catalytic site in Clade 1 and Clade 2 D,D-carboxypeptidases. A. Surface representation of Rv2911_{opt}-PMSF showing regions conserved across the two clades (green, 100% identity) as well as clade-specific conserved surfaces (orange, at least 90% identity within Clade 1). Surfaces corresponding to residues in the three conserved active site motifs are labeled to establish orientation. B. Catalytic site of Rv2911 (colored as in A) showing the superposition of stem-peptide mimicking ligands from *E. coli* PBP6 (tan, PDB 3ITB) and *B. subtilis* PBP4a (light blue, PDB 2J9P). The site that distinguishes lysine (tan) from DAP (light blue) is conserved at the bottom of the figure. PMSF was omitted for clarity. C. Rv3330-meropenem surface representation showing surfaces conserved across the two clades (green) and across Clade 2 (blue, at least 90% identity within Clade 2). A large surface (300 Å²) on the lower right in the figure is conserved in Clade 2 enzymes. D. The catalytic site of Rv3330 bound to meropenem (purple) with stem-peptide mimics from the superimposed complexes of *E. coli* PBP6 (tan, PDB 3ITB) and *B. subtilis* PBP4a (light blue, PDB 2J9P). The donor-peptide mimic (tan) exits the active site to the upper left of the figure. In contrast, meropenem interacts with a surface conserved in Clade 2 (upper right).

doi:10.1371/journal.pone.0116249.g005

combination of meropenem with clavulanate can be effective at treating drug-resistant TB [5] has spurred interest in *Mtb* penicillin-binding proteins and L,D-transpeptidases. The latter enzymes were previously thought to be insensitive to beta-lactams, but Cordillot et al. showed robust inhibition of four out of five *Mtb* L,D-transpeptidases by four carbapenems, including meropenem [11]. The realization that broad target specificity can mediate drug potency [17] intensifies the need to understand carbapenem recognition by different target families. Here, we presented structural and phylogenetic analyses of two meropenem-sensitive *Mtb* penicillin-binding proteins, Rv2911 and Rv3330.

Poor solubility often constitutes a major challenge to structural and biochemical characterization of cell wall-associated enzymes. In this study,

structure determination of wild-type Rv2911 protein could not be achieved. To overcome this challenge, we utilized an unbiased directed-evolution approach. Four point mutations discovered using directed evolution and the GFP folding reporter method resulted in the Rv2911_{opt} protein, which showed improved solubility and allowed crystallographic structure determination. Application of directed evolution methods to cell-wall associated proteins can aid research progress in this important area.

Structural analysis can be used to interpret biochemical observations such as Rv2911 and Rv3330 specificity for DAP-containing stem peptides. Based on superposition with structures of *E. coli* [18] and *B. subtilis* [19] penicillin-binding proteins bound to stem-peptide mimics we infer that the conserved SGN motif of the *Mtb* enzymes forms hydrogen bonds to the proximal carboxyl of DAP. Binding of the aliphatic middle part of the DAP residue occurs in a narrow groove that extends away from the catalytic serine between the side chains of conserved Thr234/286 (Rv2911/Rv3330) and Leu166/218. It is likely that a conserved Glu99/151 binds to the DAP amide group. The main-chain amides of residues in the loop that follows the KTG β -strand are positioned to interact with the distal DAP carboxyl. In contrast to these shared features of the DAP-binding site, the L-Ala and iso-D-Gln residues of the superimposed stem peptide are positioned to interact with clade-specific 100-CysAsnCys-102 in Rv2911 and 152-GlyThr-153 in Rv3330.

The structure of the meropenem adduct of Rv3330 provides insights into the interactions of this inhibitor with the target enzyme. The beta-lactam core of meropenem interacts with the three characteristic penicillin-binding protein active-site sequence motifs. In contrast, the distal end of this inhibitor, which is variable in carbapenems, points away from the stem peptide-binding site (Fig. 5). The meropenem distal end binds a surface that is partially varied between Rv2911 and Rv3330. This binding mode explains the meropenem inhibition of these enzymes and suggests that the distal interactions of the drug could be exploited to mediate selectivity.

The phylogenetic analysis of the *Peptidase_S11* family proteins shows not only that Actinobacteria contain a set of penicillin-binding proteins different from that of Proteobacteria or Firmicutes, but also that gene duplication and specialization in mycobacteria produced two clades represented by Rv2911 and Rv3330. These proteins lack the characteristic PBP5 C-terminal domain that is widespread in both Proteobacteria and Firmicutes, supporting their classification as class 7 penicillin-binding proteins. This analysis contradicts the view that Rv3330 is a class 5 penicillin-binding protein [9]. The clade-specific residues in Rv2911 and Rv3330 define sequences that can mediate functional specializations.

The Rv2911_{opt} and Rv3330 inhibitor-complex structures provide a basis to interpret adaptations that occurred in the two peptidase subfamilies. Importantly, unlike comparisons of penicillin-binding proteins from different species, the Rv2911 and Rv3330 comparison excludes the possible effects of differences in peptidoglycan chemotype or other cell-wall features. Key subfamily adaptations include clade-specific residues surrounding the catalytic center. In addition,

Rv3330 contains an extended conserved surface located away from the active site, and a conserved trans-membrane helix. These features suggest that the two peptidases could have distinct catalytic activities, different protein partners, or different modes of interaction with cell-wall polymers. In model organisms, penicillin-binding proteins interact with other peptidoglycan hydrolases and with dedicated regulatory proteins [20]. These interactions are critical to the functions of the cell wall remodeling complexes, including divisome and elongasome complexes that are responsible, respectively, for daughter cell separation and cell growth [21]. Discovery of peptidoglycan-processing complexes of *Mtb* as well as detailed biochemical analysis of substrate specificity should help reveal the overall logic of *Mtb* cell-wall homeostasis and point out the vulnerabilities that can be attacked by novel and updated drugs.

Materials and Methods

Bioinformatic analyses

Sequences of *Peptidase_S11* (PF00768) family proteins from Actinobacteria (taxid:1760) and Mycobacteria (taxid:1763) were downloaded from the Pfam server [22]. Sequences of homologous proteins from *E. coli* and *B. subtilis* were obtained from the UniProt database. Alignment of protein sequences truncated to the *Peptidase_S11* catalytic domain were generated with the *hmmalign* program, part of the HMMER package [23]. Belvu multiple-alignment editor was used to discard columns in the alignment that contained greater than 50% gap characters. Thereafter, redundant rows and rows containing more than 30% gaps were also removed. The edited alignment was analyzed using the PHYLIP package. The neighbor-joining tree of all actinobacterial and the maximum-likelihood tree of the selected mycobacterial proteins were generated using default parameters with 100 bootstrap replicates. The trees were analyzed and rendered using the FigTree program.

Clade 1 and Clade 2 proteins were gathered using BLAST search against mycobacterial genomes (taxid:1763) with either Rv2911 (Clade 1) or Rv3330 (Clade 2) as query. The BLAST results were first filtered based on percent query coverage (95% minimum) and then based on presence of Clade 2 characteristic transmembrane domains using TMHMM [24]. N-terminal signal peptide sequences were removed to generate predicted mature forms using the Signal-P web server [25], keeping the trans-membrane sequences intact. Alignments of clade-assigned sequences were generated using the MUSCLE program [26].

Construct design, cloning, and protein purification

Maximal constructs matching the predicted mature forms were used: 26–291 for Rv2911 and 38–368 for Rv3330. This Rv3330 construct stopped at the beginning of the predicted C-terminal trans-membrane domain. The two genes were cloned from *Mtb* strain H37Rv genomic DNA into pDEST15 vector using the Gateway

system (Invitrogen). Rosetta 2(DE3) *E. coli* expression cells (Novagen) transformed with GST-tagged TEV-cleavable constructs were inoculated into Terrific Broth, grown to OD₆₀₀ of 0.6, and induced with 0.2 mM IPTG. Cells were grown with shaking at room temperature for 18 hours and harvested by centrifugation. Cell pellets were resuspended in 300 mM NaCl, 20 mM HEPES, pH 7.5, 0.5 mM TCEP, 10% glycerol and lysed by sonication at 4°C. Lysates were cleared by centrifugation and filtration using 0.22 μm filters. Samples were applied to equilibrated glutathione-affinity columns (GE), washed and then eluted with 20 mM reduced glutathione (Sigma) in the above buffer. Overnight dialysis into the glutathione-free buffer was performed following addition of 0.5 mg of TEV protease. Glutathione-affinity chromatography was repeated to remove the GST tag. Size-exclusion chromatography on a Superdex-75 column (GE) was used to complete purification. The proteins were used in peptidoglycan-binding assays and to obtain crystals of Rv3330. The Rv2911 construct used for crystallization, however, required extensive optimization as described below.

Rv2911 cloning and construct optimization by directed evolution and the GFP folding reporter method

Full-length wild-type Rv2911, including a predicted leader sequence, was expressed from pET28 vector in *E. coli* BL21(DE3) cells at either 27°C or 37°C and was insoluble. The wild-type Rv2911 was subjected to forward evolution using the GFP folding reporter method [16]. During each round of evolution, 40 brightest colonies out of approximately 40,000 were picked for subsequent recombination by gene shuffling [16, 27]. After four rounds of forward evolution, there was no further increase in GFP fusion fluorescence. Following two rounds of back-crossing, the optimized sequence contained C12R, V18D, A113V, T197I, N208D, and G289D mutations. Because the soluble full-length protein failed to yield diffraction-quality crystals, an N-terminal truncation lacking 16-residue leader sequence was used.

Rv2911_{opt} protein purification

E. coli BL21(DE3) cells transformed with C-terminally His-tagged Rv2911_{opt} expression plasmid were grown to OD₆₀₀ of 0.6 and induced with 1 mM IPTG. Growth was continued at 22°C for 20 hours. The cell pellets were resuspended in buffer A (20 mM Tris, pH 8.0, 100 mM NaCl, protease inhibitor cocktail) and lysed by sonication. After ultracentrifugation for 30 min at 38,000 rpm, the supernatant was loaded on a 5 mL Talon superflow affinity column (Clontech) equilibrated with buffer A. After washing with 30 mL buffer A, the His-tagged Rv2911_{opt} was eluted from the Talon affinity column using buffer B (20 mM Tris, pH 8.0, 100 mM NaCl, 300 mM imidazole). The eluted fraction was dialyzed against buffer A and purified by gel filtration on a Superdex-75 column (Amersham Pharmacia Biotech) using buffer A with 10 mM β-mercaptoethanol for equilibration and elution. Selenomethionyl Rv2911 was prepared using

methionine pathway inhibition method [28] by growing cells in presence of 60 mg/L L-Se-Met with additional 100 mg/L of threonine, lysine and phenylalanine each, and 50 mg/L of isoleucine, leucine and valine each. Mass spectrometry (Applied Biosystems) was used to confirm, that all 7 Met residues were replaced by Se-Met.

Structure of Rv2911_{opt}

Rv2911_{opt} crystals were obtained by mixing protein (15 mg/ml) 1:1 with the mother liquor (17% PEG monomethyl ester 5000, 0.15 M Ammonium sulfate, 0.1 M Citrate buffer pH 5.6) in hanging-drop format at 22°C. The three-wavelength multiwavelength anomalous diffraction data of Rv2911_{opt} SeMet crystal were collected at Beamline 5.0.2 of the Advanced Light Source (ALS) at Lawrence Berkeley National Laboratory (LBNL), and processed with the HKL2000 package [29]. The structure was determined with SHELX [30] and the model was partially traced by RESOLVE [31] after solvent flattening with DM [32] in CCP4 package [33]. Tracing was completed with Arp/wArp [34]. PMSF moiety covalently linked to Ser69 was traced manually using SEB (o-benzylsulfonyl-serine) residue from PDB entry 1ZIIY as a starting model. The program Coot [35] was used for manual rebuilding. Refinement was carried out using the program Phenix [36]. The structure has been deposited to the Protein Data Bank under accession number 4P0M.

Structure of Rv3330

Rv3330 (10 mg/ml) was incubated with a two-fold molar excess of meropenem (Sigma) for 30 min and mixed 1:1 with 0.1 M NaCl, 0.1 M HEPES, pH 7.23, 1.7 M ammonium sulfate well solution in a hanging-drop format. The crystals were cryopreserved in 20% xylitol in well solution and frozen in liquid nitrogen. X-ray diffraction data were collected under cryogenic conditions at Beamline 8.3.1 at the ALS at LBNL [37] and processed using HKL2000 [29]. The structure was determined by molecular replacement using the Rv2911_{opt} coordinates as the search model. Molecular replacement, model building, refinement, and validation were performed using Phenix [36] and Coot [35]. The structure has been deposited to the Protein Data Bank under accession number 4PPR. Molecular images were generated using Chimera [38].

Peptidoglycan binding assay

One milligram of *B. subtilis* peptidoglycan (Sigma) was washed twice in 1 ml of pure water and resuspended in 100 µl of binding buffer (150 mM NaCl, 25 mM HEPES, pH 7). Protein samples (10 µl at 0.4 mg/ml) were added to 10 µl of peptidoglycan suspension, mixed, and incubated for 5 min at room temperature with nutation. Samples were centrifuged for two minutes at 11,500 g. Input control, bound, and free protein fractions were analyzed by SDS-PAGE and stained with Coomassie Brilliant Blue (Bio-Rad).

Acknowledgments

We thank James Holton, George Meigs, Jane Tanamachi, and Berkeley Center for Structural Biology at LBNL for help with Beamlines 8.3.1 and 5.0.2 of the ALS. We thank the TB Structural Genomics Consortium for helpful discussions.

Author Contributions

Conceived and designed the experiments: DMP IVK JPH DM GSW LWH JCS TCT TA. Performed the experiments: DMP IVK JPH DM GSW. Analyzed the data: DMP IVK JPH. Wrote the paper: DMP IVK JPH DM GSW LWH JCS TCT TA.

References

1. Russell DG, Barry CE, Flynn JL (2010) Tuberculosis: what we don't know can, and does, hurt us. *Science* 328: 852–856.
2. Koul A, Arnoult E, Lounis N, Guillemont J, Andries K (2011) The challenge of new drug discovery for tuberculosis. *Nature* 469: 483–490.
3. Janin YL (2007) Antituberculosis drugs: ten years of research. *Bioorg Med Chem* 15: 2479–2513.
4. Crick DC, Mahapatra S, Brennan PJ (2001) Biosynthesis of the arabinogalactan-peptidoglycan complex of *Mycobacterium tuberculosis*. *Glycobiology* 11: 107R–118R.
5. Hugonnet JE, Tremblay LW, Boshoff HI, Barry CE, 3rd, Blanchard JS (2009) Meropenem-clavulanate is effective against extensively drug-resistant *Mycobacterium tuberculosis*. *Science* 323: 1215–1218.
6. Schleifer KH, Kandler O (1972) Peptidoglycan types of bacterial cell walls and their taxonomic implications. *Bacteriol Rev* 36: 407–477.
7. Mahapatra S, Scherman H, Brennan PJ, Crick DC (2005) N Glycosylation of the nucleotide precursors of peptidoglycan biosynthesis of *Mycobacterium* spp. is altered by drug treatment. *J Bacteriol* 187: 2341–2347.
8. Lederer E, Adam A, Ciorbaru R, Petit JF, Wietzerbin J (1975) Cell walls of *Mycobacteria* and related organisms; chemistry and immunostimulant properties. *Mol Cell Biochem* 7: 87–104.
9. Sauvage E, Kerff F, Terrak M, Ayala JA, Charlier P (2008) The penicillin-binding proteins: structure and role in peptidoglycan biosynthesis. *FEMS Microbiol Rev* 32: 234–258.
10. Lavollay M, Arthur M, Fourgeaud M, Dubost L, Marie A, et al. (2008) The peptidoglycan of stationary-phase *Mycobacterium tuberculosis* predominantly contains cross-links generated by L,D-transpeptidation. *J Bacteriol* 190: 4360–4366.
11. Cordillot M, Dubee V, Triboulet S, Dubost L, Marie A, et al. (2013) In vitro cross-linking of *Mycobacterium tuberculosis* peptidoglycan by L,D-transpeptidases and inactivation of these enzymes by carbapenems. *Antimicrob Agents Chemother* 57: 5940–5945.
12. Kumar P, Arora K, Lloyd JR, Lee IY, Nair V, et al. (2012) Meropenem inhibits D,D-carboxypeptidase activity in *Mycobacterium tuberculosis*. *Mol Microbiol* 86: 367–381.
13. Fedarovich A, Nicholas RA, Davies C (2012) The role of the beta5-alpha11 loop in the active-site dynamics of acylated penicillin-binding protein A from *Mycobacterium tuberculosis*. *J Mol Biol* 418: 316–330.
14. Malen H, Berven FS, Fladmark KE, Wiker HG (2007) Comprehensive analysis of exported proteins from *Mycobacterium tuberculosis* H37Rv. *Proteomics* 7: 1702–1718.
15. Bourai N, Jacobs WR, Jr., Narayanan S (2012) Deletion and overexpression studies on DacB2, a putative low molecular mass penicillin binding protein from *Mycobacterium tuberculosis* H(37)Rv. *Microb Pathog* 52: 109–116.

16. **Waldo GS, Standish BM, Berendzen J, Terwilliger TC** (1999) Rapid protein-folding assay using green fluorescent protein. *Nat Biotechnol* 17: 691–695.
17. **Wong S, McLaughlin J, Cheng D, Zhang C, Shokat KM, et al.** (2004) Sole BCR-ABL inhibition is insufficient to eliminate all myeloproliferative disorder cell populations. *Proc Natl Acad Sci U S A* 101: 17456–17461.
18. **Chen Y, Zhang W, Shi Q, Hesek D, Lee M, et al.** (2009) Crystal structures of penicillin-binding protein 6 from *Escherichia coli*. *J Am Chem Soc* 131: 14345–14354.
19. **Sauvage E, Duez C, Herman R, Kerff F, Petrella S, et al.** (2007) Crystal structure of the *Bacillus subtilis* penicillin-binding protein 4a, and its complex with a peptidoglycan mimetic peptide. *J Mol Biol* 371: 528–539.
20. **Lovering AL, Safadi SS, Strynadka NC** (2012) Structural perspective of peptidoglycan biosynthesis and assembly. *Annu Rev Biochem* 81: 451–478.
21. **Szwedziak P, Lowe J** (2013) Do the divisome and elongasome share a common evolutionary past? *Curr Opin Microbiol* 16: 745–751.
22. **Finn RD, Mistry J, Tate J, Coggill P, Heger A, et al.** (2010) The Pfam protein families database. *Nucleic Acids Res* 38: D211–222.
23. **Eddy SR** (2011) Accelerated Profile HMM Searches. *PLoS Comput Biol* 7: e1002195.
24. **Krogh A, Larsson B, von Heijne G, Sonnhammer EL** (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol* 305: 567–580.
25. **Bendtsen JD, Nielsen H, von Heijne G, Brunak S** (2004) Improved prediction of signal peptides: SignalP 3.0. *J Mol Biol* 340: 783–795.
26. **Edgar RC** (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32: 1792–1797.
27. **Stemmer WP** (1994) Rapid evolution of a protein in vitro by DNA shuffling. *Nature* 370: 389–391.
28. **Doublet S** (1997) Preparation of selenomethionyl proteins for phase determination. *Methods Enzymol* 276: 523–530.
29. **Otwinowski Z, Minor W** (1997) Processing of X-ray diffraction data collected in oscillation mode. *Macromolecular Crystallography, Pt A* 276: 307–326.
30. **Sheldrick GM, Schneider TR** (1997) SHELXL: high-resolution refinement. *Methods Enzymol* 277: 319–343.
31. **Terwilliger TC** (2003) Automated main-chain model building by template matching and iterative fragment extension. *Acta Crystallogr D Biol Crystallogr* 59: 38–44.
32. **Wang BC** (1985) Resolution of phase ambiguity in macromolecular crystallography. *Methods Enzymol* 115: 90–112.
33. **Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, et al.** (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* 67: 235–242.
34. **Perrakis A, Morris R, Lamzin VS** (1999) Automated protein model building combined with iterative structure refinement. *Nat Struct Biol* 6: 458–463.
35. **Emsley P, Lohkamp B, Scott WG, Cowtan K** (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66: 486–501.
36. **Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, et al.** (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66: 213–221.
37. **MacDowell AA, Celestre RS, Howells M, McKinney W, Krupnick J, et al.** (2004) Suite of three protein crystallography beamlines with single superconducting bend magnet as the source. *J Synchrotron Radiat* 11: 447–455.
38. **Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, et al.** (2004) UCSF Chimera - a visualization system for exploratory research and analysis. *J Comput Chem* 25: 1605–1612.