University of Naples "Federico II"

Doctorship in Chemical Sciences

(XXV cycle)

Ph.D. Thesis

Structural and functional characterization of proteins involved in bacterial cell growth and division

Naples, Italy

2013

Tutor:

Prof. Gabriella D'Auria

Co-tutor:

Dr. Rita Berisio

Candidate:

Flavia Squeglia

I lovingly dedicate this thesis to my family,
especially to my grandfather,
a mentor, an example, a joy

"Read not to contradict and confute, not to believe and take for granted, not to find talk and discourse, but to weigh and consider"
Sir Francis Bacon

Index

An outline

In this section I am giving a brief account of thesis structure. This thesis represents a culmination of work and learning that has taken place over a period of three years at the Istitute of Biostructure and Bioimaging of Naples.

In my Ph.D. project I focus my attention on proteins involved in cell wall division and regulation of revival from dormancy in different kind of bacteria.

In the following section of this thesis there is a part dedicated to **A GENERAL OVERVIEW** on bacterial cell wall composition and biosyntesis. In this context, I also introduce the main molecular players in bacterial cell wall division and resuscitation from dormancy. Following the general overview, this thesis consists of two parts, PART I and PART II, describing two different aspects of cell division that have been investigated during my Ph.D. work.

PART I describes structural and functional studies of RipA, an enzyme essential for cell division in *Mycobacterium tuberculosis*, and of its inactive mutants.

PART II describes structural and functional investigations on two proteins involved in regulation of cell wall division and resuscitation form dormancy, the serine/threonine kinases PrkC from *Staphylococcus aureus* and *Bacillus subtilis* and the Penicillin Binding Protein PonA2 from *Mycobacterium tuberculosis*.

I would like to take this opportunity to thank my supervisors, Prof. Gabriella D'Auria and Dr. Rita Berisio and the other members of the Istitute of Biostructure and Bioimaging of Naples for providing an extremely stimulating atmosphere to work in.

A GENERAL OVERVIEW: Bacterial cell wall composition and modeling

A GENERAL OVERVIEW: Bacterial cell wall composition and modeling

Bacterial cell wall

The bacterial cell wall is an elastic macromolecule that defines the shape of bacteria and enables them to resist lysis (Johnson et al., 2013). Despite an essential structural role, the bacterial cell wall is highly dynamic and intimately involved in a broad range of cellular processes, including growth and division, chromosome segregation, morphogenesis, interaction between the bacterium and its environment, biofilm formation, pathogenesis, homeostasis of the membrane-proximal ionic environment and movement of material in and out of the cell (Archibald et al., 1993). These dynamic processes must occur without compromising cell wall integrity, which would lead to lysis and death.

Cell wall containing bacteria fall into two groups, Gram-positive and Gram-negative, based on fundamental difference in the cell envelope structure. A few species, such as Mycobacteria, are difficult to classify as they have a unique envelope composition. Peptidoglycan (PG) is a key structural constituent of bacterial cell walls. It is present in all bacteria with the exception of Mycoplasma, Planctomyces and a few other bacterial species that lack a cell wall (Moulder, 1993; Seltmann and Holst, 2002). The chemical composition of PG in Gram-negative and Gram-positive bacteria is similar, with a few differences which will be described below.

Gram-negative cell wall

In electron micrographs, the gram-negative cell wall appears multilayered. It consists of an outer membrane (OM) that surrounds a thin PG layer of a few nanometers (2-3nm) embedded in a periplasmic space between the plasma membrane (PM) and the OM (Figure 1). Chemically, only 10 to 20% of the gram-negative cell wall is PG. The PG layer is covalently attached to the OM by the lipoprotein called Lpp or Braun´s lipoprotein (Braun, 1975). The OM of the gram-negative cell wall appears as a lipid bilayer which is composed of phospholipids, lipoproteins, lipopolysaccharides (LPS), and proteins. In addition, pore-forming proteins called porins span the OM (Figure 1). Porins function as channels for the entry and exit of solutes through the outer membrane of the gram-negative cell wall.

Gram-positive cell wall

A gram-positive cell wall is primarily composed of a thick layer PG which accounts for 30-70% of the total cell wall mass, compared with just 10-20% in Gram-negatives (Shockman and Barrett, 1983). It does not possess an OM and it contains large amounts of anionic polymers called teichoic acids (TAs) (Figure 1), which confer them a negative charge. TAs are covalently attached to the PG layer (wall teichoic acids, WTAs) or anchored to the cytoplasmic membrane (lipoteichoic acids, LTAs). Additional surface structures may also be present in both Gram-positive and Gram-negative species. These structures include pili, capsulae, fimbriae and flagella (Beveridge and Graham, 1991).

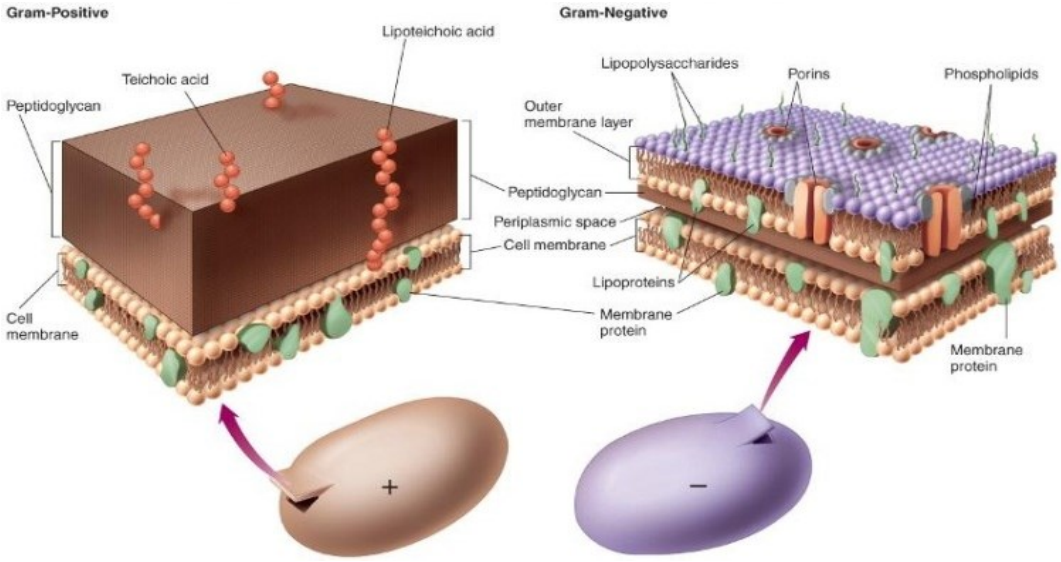


Figure 1. General structure of the Gram-positive and Gram-negative cell wall

Chemical composition of peptidoglycan

PG is a complex heteropolymer which consists of long glycan strands cross-linked by short peptide bridges, to form a large molecule of strong but elastic nature (Höltje, 1998; Park, 1996; Weidel and Pelzer, 1964). The glycan strands are formed by *N*-acetylmuramic acid (Mur*N*Ac) and *N*-acetylglucosamine (Glc*N*Ac) linked by β-1,4 glycosidic bonds. Glycan chain length varies considerably between organisms, from about 6 disaccharides (e.g. in *Staphylococcus aureus*) to a maximum length of 500 disaccharides (e.g. in *Bacillus subtilis*) (Ward, 1973; Hayhurst et al., 2008). The peptide stems are much less conserved than the glycan strands and are composed of tetrapeptides (L-Ala, γ D-Glu, diamino (DA), D-Ala) covalently attached to the lactyl group of Mur*N*Ac by L-alanine N-terminus (Figure 2). The presence of D-amino acids is

an almost unique feature of bacteria and plays a role in conferring resistance against host proteases.

PG tetrapeptides are linked together by covalent bonds between the DA group of one stem and the D-Ala residue of the crosslinking stem (Figure 2). The DA residue is commonly a *meso*-diaminopimelic acid (m-DAP), a dibasic amino acid, or an L-Lys residue. Most Gram-positive bacteria (e.g. *S. aureus*) contain an L-lysine residue at the third position of the stem peptide, whereas Gram-negative bacteria (e.g. *Bacillus* species) and most endospore formers have an m-DAP residue at this position (Shah et al., 2008).

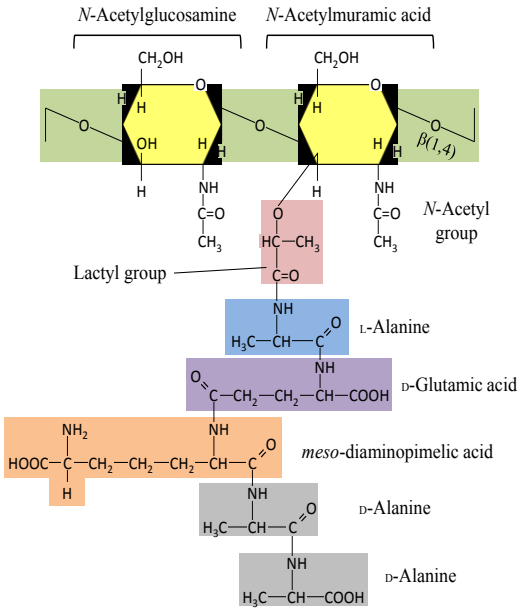


Figure 2. The chemical structure of peptidoglycan. A typical PG disaccharide pentapeptide. The glycan strand is composed of a repeating disaccharide of the amino sugars GlcNAc and MurNAc (Yellow) linked by β-1,4 glycosidic bonds (green). The pentapeptide is covalently linked to the glycan strand via the lactyl group of Mur*N*Ac (pink). A typical pentapeptide stem composition is shown, although the precise composition is variable.

Peptidoglycan turnover, bacterial cell wall growth and division

The process of PG cell wall turnover was discovered fifty years ago, when DAP was found to be released from the cell wall of the bacterium *Bacillus megaterium* during exponential growth (Chaloupka, 1962a, 1962b). Later on, pulse-chase experiments with radioactively labeled cell wall precursors such as GlcNAc or D-glutamic acid revealed that Gram-positive species turn over up to 50% of their PG in each generation during vegetative growth (Rogers 1967; Mauck et al., 1971; Boothby et al., 1973; Doyle et al., 1988). Based on the "inside-to-outside" growth model of Gram-positive bacteria, new PG cell wall are laid down along the cytoplasmic membrane and, at the end of the

turnover cycle, removed from the exterior by autolysins (Merad et al., 1989; Pooley, 1976). Therefore, inner regions of the cell wall contain newly synthesized, unstressed PG. As the cell elongates, this material passes outward and stretches, becoming the middle, stress-bearing zone. The outer zone may thus consist of old, partially hydrolysed PG awaiting solubilization (Graham & Beveridge, 1994). This "inside-to-outside" model of wall structural dynamics suggests that autolysins are necessary for hydrolysis of older PG to allow newer PG to expand and become stress-bearing as the cell elongates. The "inside-to-outside" model for cell-wall growth suggests that breakage of covalent bonds within the PG is essential for cell growth.

Bacterial cell growth and cell wall division are mediated by a collection of proteins whose action is tightly coordinated at the level of the septal ring (Nanninga 1998). In *E. coli*, cell division takes place at the mid-cell after the chromosomal replication and segregation into two daughter nucleoids. After the completion of chromosome segregation, the division process begins with the formation of the septal ring, called Z-ring, a polymer of the tubulin-like protein FtsZ (Dajkovic et al., 2006; Lutkenhaus et al., 1980). FtsZ is almost universally conserved and has also been identified in *Mycobacterium tuberculosis* (MTB) as one of the major cytoskeletal organizers of the mycobacterial divisome (Hett and Rubin, 2008; Dziadek et al., 2003). The depletion of FtsZ from bacteria results in long filamentous cells (Lutkenhaus et al., 1980). The ring formed by FtsZ involves the highly ordered recruitment of both structural and enzymatic proteins involved in PG synthesis and thus in the formation of the septum (Margolin, 2005).

Molecular players in bacterial cell division and exit from dormant state

Septal PG is initially shared between daughter cells and must be degraded by PG hydrolases to complete the division process. Although as many as 18 hydrolases are known to be involved in septum cleavage of *E. coli*, only few of them are known in mycobacteria, which possess a unique envelope structure with additional layers of arabinogalactan and mycolic acids (Hett et al., 2008; Brennan, 2003). Cell separation is mediated in MTB by the essential NlpC/P60 endopeptidase Resuscitation Promoting factor Interacting Protein A (RipA), which cleaves peptidoglycan peptide crosslinks (Hett et al., 2008), similar to other cell separating endopeptidases, like CwlT from *B. subtilis* (Fukushima et al., 2008) and Spr from *E. coli* (Aramini et al., 2008). RipA has a remarkable

effect on the bacterial phenotype, since *ripA* depletion strains in *M. smegmatis* exhibit a decreasing growth and an abnormal phenotype, consisting of branching and chaining bacteria (Hett et al., 2008). Interestingly, RipA co-localizes at bacterial septa with the Resuscitation Promoting Factor RpfB (Ruggiero et al., 2010), a key cell wall hydrolase involved in MTB resuscitation from a state of low metabolism denoted as dormancy (Mukamolova et al., 2002; Kaprelyants et al., 2012; Ruggiero et al., 2012; Ruggiero et al., 2011). The combined action of RpfB and RipA seems to enhance PG hydrolysis, this suggesting that an interaction between the two molecules may positively regulate their catalytic functions.

Bacterial Serine/Threonine Protein Kinases and Penicillin Binding Proteins: Regulation of cell division

Protein kinases are classified into two families based on their similarities and enzymatic specifications, including the histidine kinase superfamily (Stock et al., 1989), the serine/threonine and tyrosine kinase superfamilies (Hanks et al., 1988). Histidine kinases represent the classical prokaryotic mechanism for detection and response to environmental changes, whereas the serine/threonine and tyrosine protein kinases have recently emerged as important regulatory systems (Shah et al., 2010; Pereira, S.F., 2011; Wehenkel et al., 2008).

Serine/threonine protein kinases (STPKs) have been shown to regulate process of cell division (Shah et al., 2010); Their wide distribution in bacterial genomes suggests that STPKs regulatory function in cell shape and division is widely preserved among prokaryotes. Recent progress has been made to elucidate the biological function of StkP in *S. pneumoniae* and *B. subtilis* cell division. The kinase StkP from *S. pneumoniae* is essential for bacteria survival and virulence (Echenique et al., 2004) and it localizes at the cell division sites. In *B. subtilis*, the kinase PrkC is responsible for resuscitation from dormancy induced by the presence of PG fragments in the bacterial *milieu* (Shah et al., 2008). Similarly, bacterial resuscitation has also been shown to be regulated in MTB by PonA2, a protein belonging to the Penicillin Binding Protein (PBP) family and classically associated to PG synthesis (Patru et al., 2010).

PART I - Structure and Functional Regulation of RipA, a Mycobacterial enzyme essential for cell division

PART I - Structure and Functional Regulation of RipA, a Mycobacterial enzyme essential for cell division

I.1 AIM OF THE STUDY

The final stage of bacterial cell division depends on cell-wall hydrolases that cleave the PG layer connecting daughter cells. In MTB, this process is governed by the predicted endopeptidase RipA. In the absence of this enzyme, the bacterium is unable to divide and exhibits an abnormal phenotype. This makes RipA an excellent candidate as a drug target against Tuberculosis (TB). We here report the crystal structure of a relevant portion of RipA, containing its catalytic-domain and an extra-domain of hitherto unknown function. The structure clearly demonstrates that RipA is produced as a zymogen, which needs to be activated to achieve cell-division. Bacterial cell-wall degradation assays and proteolysis experiments strongly suggest that activation occurs via proteolytic processing of a fully solvent exposed loop identified in the crystal structure. Indeed, proteolytic cleavage at this loop produces an activated form, consisting of the sole catalytic domain. This work provides the first evidence of self-inhibition in cell-disconnecting enzymes and opens a field for the design of novel anti-tubercular therapeutics. The aim of this study was to clone, express, purify and characterize the predicted endopeptidase RipA from MTB. We have investigated RipA function both structurally and biochemically by combining atomic structure characterization, circular dichroism (CD), fluorescence spectroscopy and mass spectroscopy.

Broken down, the aims were:

- Identification of well-structured RipA variants by a combination of different bioinformatic investigations;
- Establishing purification protocols for different RipA variants;
- Identify of best crystallization conditions by robotic procedure;
- Optimizing the crystallization conditions by manual procedure;
- Optimizing cell wall degrading assay;

I.2 BACKGROUND - Tuberculosis, old disease, new danger

TB is more prevalent in the world today than at any other time in human history. This kind of infectious bacterial disease is caused by MTB. Each year, approximately 2 million persons worldwide die of TB and 9 million become infected. TB remains a leading cause of mortality worldwide due to the ability of MTB to survive in a dormant status for extended periods of time in humans without producing symptoms for subsequently reviving into an active state. (Yukari C. et al., 2000). This apparent dormancy can develop into active disease even decades after initial infection, when the immune response diminishes, as in the case of TB-HIV co-infection (Yukari C. et al., 2000, Nancy A., Knechel, 2009).

MTB is a rod-shaped, non-spore forming, aerobic bacterium which typically measure 0.5µm by 3µm (Figure 3). It is classified as acid-fast bacillus and have a unique cell wall structure crucial to its survival. The cell wall complex contains PG, but otherwise it is composed of complex lipids. Over 60% of the MTB cell wall is lipid. The cell wall contains also mycolic acid, covalently attached to the underlying PG-bound polysaccharide arabinogalactan, providing an extraordinary lipid barrier. This barrier is responsible for many of the medically challenging physiological characteristics of TB, including resistance to antibiotics and host defense mechanisms.

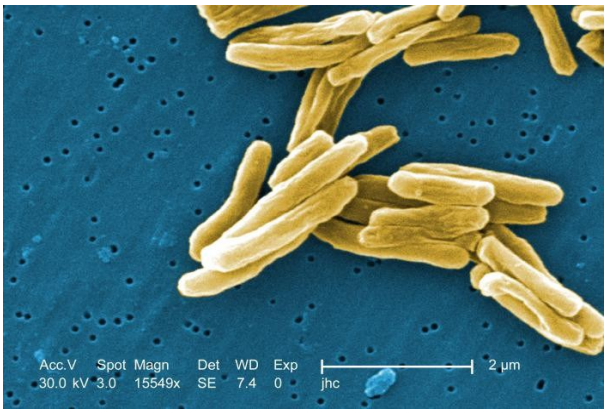


Figure 3. *Mycobacterium tuberculosis* scanning electron micrograph

Approximately 30% of exposed people become infected; 60-90% of these people will have an effective immune response allowing the successful containment of the infection. As more T cells, monocyte and macrophages are recruited to the area surrounding the bacilli, the bacilli slow replication and wait patiently, unable to transmit infection. During this phase the MTB survives in a state of minimum metabolism that allows it to escape the immune system

attacks and traditional drug therapy (Yukari C. et al., 2000). It is estimated that approximately one-third of the world's population is infected with MTB (WHO, Global Tuberculosis Control, 2010). TB in a latent state is very dangerous because persons with latent TB have no signs or symptoms and diagnosis is difficult. The reactivation of latent TB depends on many factors such as the virulence of the organism, the competence of host defenses and the presence of other diseases (diabetes, HIV infection) or immunosuppressive therapy (Kaufmann SH., 2008; Keep NH et al., 2006).

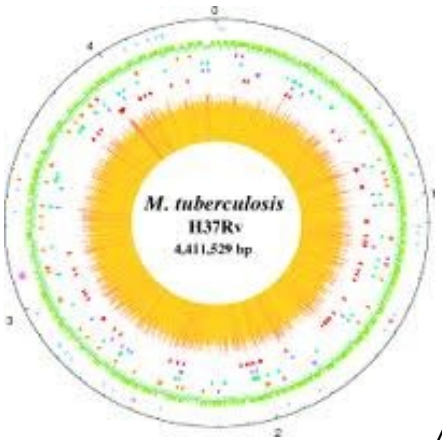
Treatment for active TB usually combines several different antibiotic drugs that are given for at least six months, sometimes for as long as 12 months. The principal drugs are: Rifampicin, Isoniazid, Pyrazinamide, Ethambutol. Associations of three drugs of first-line (Rifampicin, Isoniazid and Pyrazinamide) are recommended for the treatment of the intensive phase of TB, whereas associations of Rifampicin and Isoniazid are recommended for the continuation phase of the disease. Unfortunately, however, the treatments incomplete or inadequate has led to the emergence of strains resistant to antibiotics. Multidrug-resistant TB (MDR-TB) is defined by resistance to the two most commonly used drugs in the current four-drugs (or first-line) regimen, Isoniazid and Rifampin. MDR-TB must be treated necessarily with second-line drugs (Ethionamide, Cycloserine, Capreomycin). According to the WHO, Eastern Europe's rates of MDR-TB are the highest, where MDR-TB makes up 20 percent of all new TB cases. In some parts of the former Soviet Union, up to 28 percent of new TB cases are Multidrug-resistant (http://www.tballiance.org). Complications due to prolonged treatment with antibiotics, as well as the emergence of MDR-TB and the inability to treat latent TB gave a strong impetus to research and design of novel anti-tubercular drugs. In this contest, RipA would be an excellent new candidate as a drug target against TB.

I.3 EXPERIMENTAL PROCEDURES

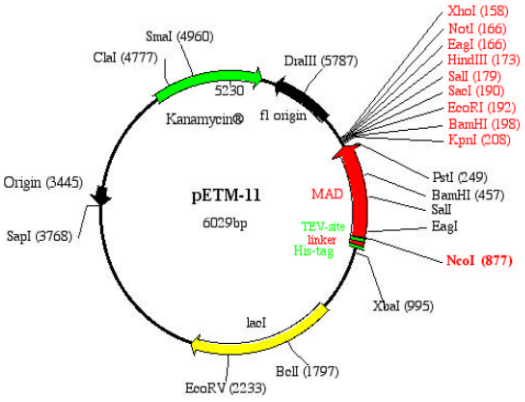
Here, we report the cloning, expression, purification, crystallization, crystallographic and functional investigations of a relevant portion of RipA (RipA$_{263-472}$), containing its predicted catalytic domain (residues 355-471) and the PB015164 domain (residues 265-354), as identified in the Protein Families (PFAM) database (Finn R.D. et al., 2008).

I.3.1 Cloning and expression of RipA variants

Oligonucleotide primers were synthesized by Primm (Milano, Italy) and were designed to amplify the nucleotide sequence corresponding to the three variants of RipA by Polymerase Chain Reaction (PCR) using genomic DNA of MTB as a template (H37Rv strain) (Figure 4A). NcoI/HindIII-digested fragments were cloned into the pETM-11 (Figure 4B), pETM-30 (Figure 4D) and pETM-20 (Figure 4C) expression vectors (Novagen, Wisconsin, USA) for $RipA_{40–472}$, $RipA_{263-472}$ and $RipA_{332–472}$, respectively. The pETM20 vector contains the stabilising TRX fusion tag, whereas the pETM30 vector allows the expression of GST-fusion proteins[1]. The resulting positive plasmids of $RipA_{40–472}$ and of $RipA_{263-472}$ were used to transform *E. coli* BL21(DE3) strain (Invitrogen, California, USA) while the resulting positive plasmid of $RipA_{332–472}$ was used to transform BL21 Star(DE3) strain[2]. The expression of RipA variants were carried out using the transformed cells grown overnight at 37°C in Luria-Bertani (LB) media containing the opportune antibiotics and then inducing them overnight with 1 mM Isopropil β-D-1-tiogalattopiranoside (IPTG) at 22°C.



A



B

[1] Novagen pET expression system is the most popular commercial expression system for *E. coli* and is based on the T7 promoter (Yin et al., 2007). It represents more than 90% of the 2003 PDB protein preparation systems (Sørensen and Mortensen 2005). The pET system was first described in 1990, and has been developed for a variety of expression applications (Studier et al., 1990; Dubendorff and Studier 1991).

[2] *Escherichia coli* represents the ideal expression system of choice for the majority of laboratories engaged in high-throughput cloning, expression and purification of proteins for structural genomics (Goulding, C.W., Perry, L.J., 2003).

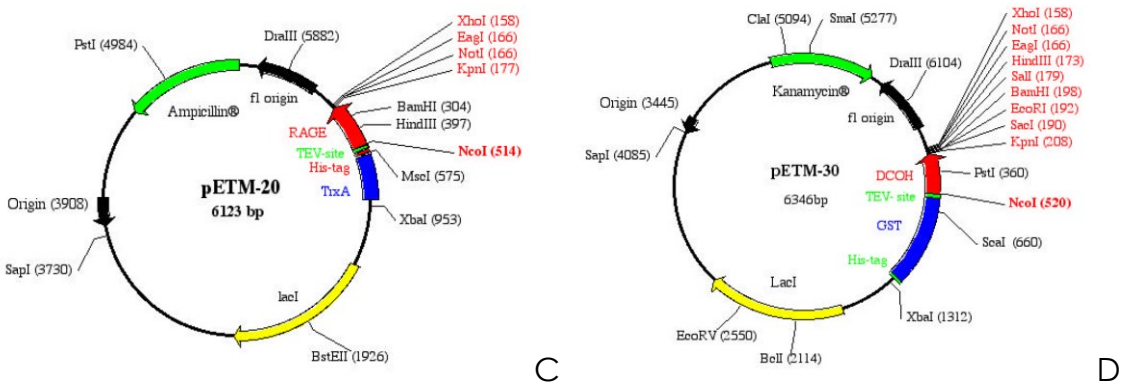


Figure 4. **Genome of H37Rv strain and pETM vectors maps.** A) Genomic DNA of MTB; B) pETM11 vector map; C) pETM20 vector map; D) pETM30 vector map;

I.3.2 Purification of RipA variants

The same procedure of purification was applied to all RipA variants. A 1 Liter expression cell pellet of each RipA variants was re-suspended in 30 mL of binding buffer (10 mM imidazole, 150 mM NaCl, 50 mM TrisHCl, pH 8.0), containing a protease-inhibitor cocktail (Roche Diagnostic), and sonicated for 10 minutes with intervals during the procedure in order to avoid protein degradation due to heat produced by the sonicator's sound waves. The lysates were cleared by centrifugation at 18,000 rpm for 40 minutes at 4°C. The supernatants were loaded on a 5-mL Ni^{2+}-nitriloacetic acid (NTA) column (Pharmacia), equilibrated with binding buffer, and purified by immobilized metal affinity chromatography (IMAC). After washing with 10 volumes of binding buffer, a linear gradient of imidazole (5-300 mM) was applied to elute the proteins (Figure 5).

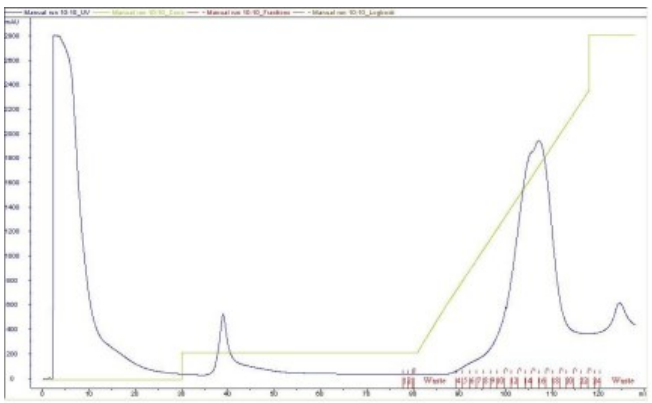


Figure 5. Affinity chromatography chromatogram of RipA

The 2 mL fractions containing RipA variants were brougth to a final volume of 20 mL with a buffer containing 50mM Tris-HCl, 150 mM NaCl (pH 8.0) and kept under TEV protease digestion at 20°C overnight. After removal of tag, $RipA_{40-472}$ and $RipA_{263-472}$ were further purified by size exclusion chromatography on Superdex 200 (Pharmacia) (Figure 6) with a buffer containing 50 mM TrisHCl, 150 mM NaCl, 5% glycerol (pH 8.0), whereas $RipA_{332-472}$ was purified on Superdex 75 (Pharmacia) with the same buffer. The proteins eluted in a single peak and were homogeneous as judged by SDS–PAGE analysis. The molecular mass were checked by mass-spectroscopy. Fresh concentrated proteins, usually 7-10 mg/mL, were used for crystallization experiments.

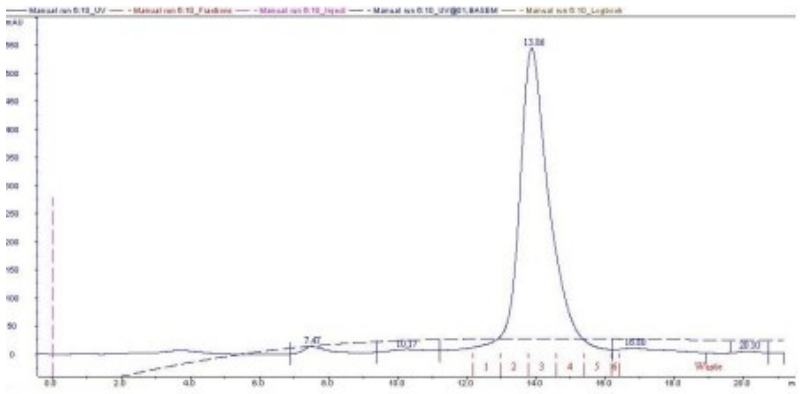


Figure 6. Gel filtration chromatogram of $RipA_{263-472}$

I.3.3 Crystallization experiments

We used high-throughput robotic systems to screen and optimise crystallisation conditions of RipA variants. The initial screening involves the high throughput crystallization screening robot (Hampton) and commercially available sparse-matrix kits (Crystal Screen kits I and II, Hampton Research) and utilizing the hanging-drop vapor diffusion technique. The Hampton robot was utilized to create the hanging drops containing the protein solution and crystallizing buffer in 1:1 ratio. After protein and buffer administration, the plates were sealed and stored at 20°C. Crystal formation was observed with a light microscope.

The initial screenings revealed several promising conditions for crystallization of $RipA_{263-472}$ variant. All favourable conditions were characterized by the presence of PEG4000 as precipitating agent. The quality of the crystals was improved by fine-tuning the concentration of the protein and of the precipitants.

I.3.4 Production of selenomethionine-labeled $RipA_{263-472}$

A SeMet derivative of the protein was prepared to perform Multi-wavelength Anomalous Diffraction (MAD) experiments. Seven methionines are present in the sequence, which consists of 210 residues. For the preparation of labeled protein of RipA ($SeMetRipA_{263-472}$), the *E. coli* BL21(DE3) cells expressing the recombinant enzyme were grown in 1 L of minimal media (M9) containing 0.4% Glucose, 1 mM $MgSO_4$, 0.1 mM $CaCl_2$, 50 ugL-1 kanamycin, 100 ugL-1 thiamine at 37°C. After reaching the OD600 of 0.7, an aminoacid mix (50 mg L-1 Ile, Leu and Val and 100 mg L-1 of Phe, Thr, and Lys) was added to the culture to inhibit methionine biosynthesis in *E. coli* then shifted to 22°C. After about 15 minutes, 60 mg L-1 of seleno-L-methionine were added and the induction was performed. The same purification protocol as for the native protein was used for the labeled protein. Crystals of $SeMetRipA_{263-472}$ grew in similar conditions as the native protein.

I.3.5 RipA mutants preparation

The expression vectors encoding the C383A, H432A, E444H mutations were generated by site-directed mutagenesis of wild-type plasmids pETM20-$RipA_{263-472}$ and pETM-20-$RipA_{332–472}$ using the Stratagene QuikChange kit and the mutagenic primers reported below.

C383A	5'-CCGTCGGCTTCGACGCCTCAGGCCTGGTGTTG-3' 5'-CAACACCAGGCCTGAGGCGTCGAAGCCGACGG-3'
H432A	5'-CCGAACGGTAGCCAGGCCGTGACGATCTACCTC-3' 5'-GAGGTAGATCGTCACGGCCTGGCTACCGTTCGG-3'
E444A	5'-CAACGGCCAGATGCTCGCGGCGCCCGACGTCGG-3' 5'-CCGACGTCGGGCGCCGCGAGCATCTGGCCGTTG-3'

Introduction of the expected mutations was confirmed by DNA sequencing. Expression and purification of the mutants were carried out in the same conditions as for their unmutated enzyme.

I.3.6 Data collection and processing of RipA$_{263-472}$

Preliminary diffraction data were collected in-house at 100K using a Rigaku Micromax 007 HF generator producing Cu Ka radiation and equipped with a Saturn944 CCD detector. Higher diffraction data for both native RipA$_{263-472}$ (Figure 7) and $_{SeMet}$RipA$_{263-472}$ were collected at the synchrotron beamline BM14 at the ESRF (Grenoble, France) at 100 K. Cryoprotection of the crystals was achieved by a fast soaking in a solution containing glycerol to a final concentration of 30% (v/v). MAD was carried out using three different wavelengths determined from the selenium absorption spectrum. A native dataset was also recorded. The data sets were scaled and merged using HKL2000 program package (Otwinowski, Z. et al., 1997).

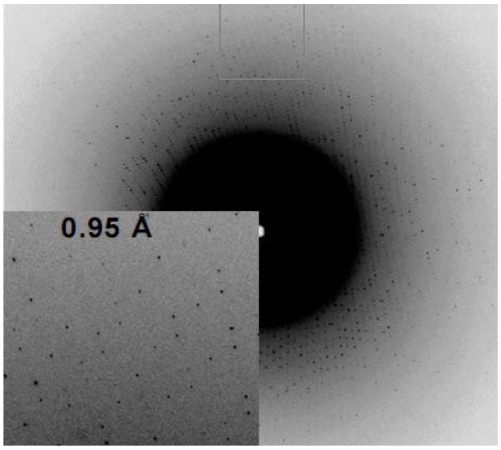


Figure 7. Diffraction pattern of native RipA crystal. Diffraction data are detectable to 0.95 resolution.

I.3.7 Crystallization, data collection and processing of RipA$_{263-472}$ mutants

Crystallization trials were performed at 293 K using the hanging-drop vapor-diffusion method. Best crystals were obtained using 5–10 mgxmL^{-1} protein solution and 8% (v/v) 2-Propanol, 16% (w/v) PEG4000 in 60 mM sodium citrate trihydrate buffer, pH 5.6. Diffraction data of the H432A mutant were recorded inhouse at 100K using a Rigaku Micromax 007 HF generator producing Cu *Ka* radiation and equipped with a Saturn944 CCD detector whereas diffraction data of the C383A mutant were collected at the synchrotron beamline BM14 at the ESRF (Grenoble, France). Cryoprotection of the crystals was achieved by a fast soaking in a solution containing glycerol to a final concentration of 30% (v/v). The data sets were scaled and merged using HKL2000 program package (Otwinowski and Minor, 1997).

I.3.8 Structure determination of $RipA_{263-472}$

The structure of the enzyme was solved by MAD methods using the anomalous signal from the Se atoms of selenomethionine-labeled protein. In order to determine the peak and the inflection wavelengths, a fluorescence scan was recorded on a single SeMet-labeled $RipA_{263-472}$ crystal. Using data sets collected at wavelengths optimized for SeMet, the program SOLVE (Terwilliger, T.C., Berendzen, 1999) identified five selenium sites in the asymmetric unit of the protein. The program SOLVE provided a set of initial phases, which were improved using the solvent flattening methods implemented in the program RESOLVE (Terwilliger, T.C., 2003) and wARP (Langer, G. et al., 2008). Using these phases and atomic resolution data from a native crystal (1.0 Å, Figure 7), nearly 90% of the residues present in the asymmetric unit could be automatically modeled using the program wARP. Model building was optimized using both automatic and manual approaches (Jones, T.A., 2004). Crystallographic refinement was first carried out against 95% of the measured data using the CCP4 program suite (Potterton et al., 2003). The remaining 5% of the observed data, which was randomly selected, was used in Rfree calculations to monitor the progress of refinement. The refinement in Refmac was started using data up to 1.7Ă resolution and gradually increased in subsequent rounds of refinement to the highest resolutions (Murshudov et al., 1997). At this stage, water molecules were incorporated into the structure in several rounds of successive refinement. This refined model was used to carry out CGLS refinement using SHELXL97 (Sheldrick, 2008), where X-ray intensities were used in refinement calculations. Structures were validated using the program PROCHECK (Laskowski et al., 1996).

I.3.9 Structure refinement of mutants

The structures of $RipA_{263-472}$C383A and $RipA_{263-472}$H432A were refined against the structure of $RipA_{263-472}$. Crystallographic refinement was carried out against 95% of the measured data using the CCP4 program suite (Potterton et al., 2003). The remaining 5% of the observed data, which was randomly selected, was used in Rfree calculations to monitor the progress of refinement. The refinement in Refmac was started with rigid body refinement, followed by restrained refinement (Murshudov et al., 1997). Water molecules were incorporated into the structure in several rounds of successive refinement. In the case of the C383A mutant, whose crystals diffracted at atomic resolution, CGLS refinement

cycles were performed using SHELXL97 (Sheldrick, 2008). Initial rounds of restrained CGLS refinement were carried out by keeping all atomic displacement parameters (ADPs) isotropic. Subsequently, the ADPs were converted to anisotropic values, leading to improved Fourier maps. The bulk solvent was modeled based on Babinet's principle, as implemented in the SWAT option in the SHELXL program. An approximate isotropic behaviour was attributed to solvent atoms (ISOR restraint). The final round of refinement was carried out with the inclusion of riding H atoms for protein residues. The positions of H atoms assigned based on the known geometrical criteria were not refined. Structures of C383A and H432A were validated using the program PROCHECK (Laskowski et al., 1996).

I.3.10 Circular Dichroism studies

All CD spectra were recorded with a Jasco J-810 spectropolarimeter equipped with a Peltier temperature control system (Model PTC-423-S). Molar ellipticity per mean residue, [θ] in deg cm^2 x $dmol^{-1}$, was calculated from the equation: $[\theta] = [\theta]obs \times mrw \times (10 \times l \times C)^{-1}$, where [θ]obs is the ellipticity measured in degrees, mrw is the mean residue molecular mass, C is the protein concentration in g x L^{-1}, and *l* is the optical path length of the cell in cm. Far-UV measurements (183–250 nm) were carried out at 20°C using a 0.1 cm optical path length cell and a protein concentration of 0.2 mgxmL^{-1}.

I.3.11 Limited proteolysis experiments

$RipA_{263–472}$ variant was incubated at room temperature with trypsin, thermolysin, and pronase at various protease/enzyme ratios (1:50, 1:100, 1:200, and 1:500) in a suitable buffer containing 50 mM Tris/HCl pH 8 and 10 mM $CaCl_2$. Samples were taken after 30 and 90 minutes and the reactions stopped by adding SDS-PAGE sample buffer. The reaction products were analyzed by SDS-PAGE. Samples were also analyzed by mass spectrometry, after trypsin (1:50 ratio), thermolysin (1:100 ratio), and pronase (1:500 ratio) digestion.

I.3.12 Mass spectrometry studies

Mass Spectrometry (MS) analysis was carried out on an LCQ DECA XP Ion Trap mass spectrometer (ThermoElectron, Milan, Italy) equipped with an OPTON ESI source (operating at 4.2-kV needle voltage and 320°C), and with a complete Surveyor HPLC system. Narrow bore 50 x 2 mm C18 BioBasic Liquid

Chromatography-Mass Spectrometry (LC-MS) columns from Thermo- Electron were used for the analyses. LC binary gradient was from 5% to 70% of B (where B was CH_3CN 0.05% TFA, and A H_2O 0.08% TFA) in 45 minutes. Initial protein concentration was 22 µM (0.5 mgxmL-1). Three microliters of this solution were injected for all analyses. Mass spectra were recorded continuously between the mass range 400–2000 Da in positive mode. Multicharge spectra were deconvoluted using the BioMass program implemented in the Bioworks 3.1 package provided by the manufacturer.

I.3.13 Cell wall degradation assays

Inactivated lyophilized cells of *M. lysodeikticus* (ATCC No. 4698) were labeled with fluorescein isothiocyanate (FITC) by covalently linking FITC to amine groups in the cell wall. In this reaction, 40 mg cells were incubated at 20°C with 20 mg FITC (50 mgxmL-1) in 100 mM $NaCO_3$, pH 8.5, protected from light. After 16 hours incubation, the reaction mixture was centrifuged; the insoluble material was washed until the supernatant was completely colorless, to eliminate unreacted fluorochrome. The labeled insoluble material was re-suspended and stored at -20°C in a 4X reaction buffer (50 mM Tris-HCl, 10 mM $MgCl_2$, 2 mM $MnCl_2$, 100 mM NaH_2PO_4, 50 mM KCl, 0.01% (v/v) CHAPS). $RipA_{263–472}$, $RipA_{332–472}$ and $RipA_{332–472}$C383A, $RipA_{332–472}$H432A, $RipA_{332–472}$E444H mutants were incubated with FITC-labeled cells at 30°C in the reaction buffer; the insoluble substrate was centrifuged (16,000 x g) and soluble FITC conjugates were measured in triplicate with filters for excitation at 492 nm and emission at 518 nm. The buffer alone was used to correct for background release of FITC.

I.3.14 Modeling studies and bioinformatics analyses

The conformation of the muropeptide GlcNAc-MurNAc-L-Ala-γ-D-Glu-mesoDAP-D-Ala was modeled in the RipA binding cleft using the program INSIGHT. The position of L-Ala-γ-D-Glu moieties of the muropeptide was locked to that observed in the structure of its complex with endopeptidase YkfC from *Bacillus cereus* (28.6% identity with RipA catalytic domain) (Xu, acta 2010). The resulting model was energy minimized using GROMACS (Lindahl et al., 2001).

Sequence conservation studies were carried out using the software ConSurf. The homologue search algorithm CSI-BLAST was used to retrieve sequences from the UNIREF-90 sequence database (150 sequences). Sequences were

aligned using MAFFT-L-INS-I alignment method (minimal and maximal sequence identities were 35 and 95%, respectively).

I.4 - RESULTS AND DISCUSSION

I.4.1 Structural studies

In RipA sequence, four distinct regions can be identified (Figure 8): a signal peptide (orange) at its N terminus, two domains of unknown function denoted in PFAM-B as PB07342 (red) and PB015164 (green), which are mainly distributed in mycobacteria, and a predicted catalytic domain of the NlpC/P60 family at its C-terminus (purple). RipA catalytic domain contains a core of about 70 amino-acid residues (residues 385-445) which shares a sequence identity of about 35% with cysteine proteases of the NlpC/P60 family. The PB010495 is instead both of unknown structure and function.

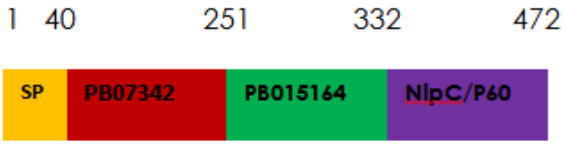
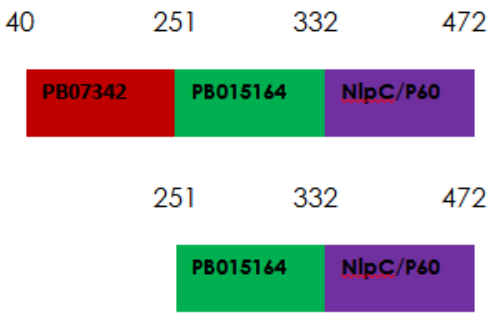


Figure 8. Modular structure of RipA: PFAM prediction

To reduce the overall complexity of the RipA protein, a systematic approach allowed the expression and purification of different RipA variants. In particular, selective domain truncation performed using a combination of different bioinformatic investigations like the prediction of secondary structure elements (PSIPRED, Jpred) (McGuffin et al., 2000; Cuff, J.A. et al., 1998), disordered regions (DISOPRED, PONDR) (Ward, J.J. et al., 2004; Xue, B. et al., 2010), coiled-coils (PCOILS) (Gruber M. et al., 2006) and transmembrane regions (TMMHM) (Krogh, A., et al., 2001) or signal peptides (Bendtsen, J.D. et al., 2004). Based on these sequence analysis we have cloned and expressed three variants of RipA including: RipA deprived of its signal peptide ($RipA_{40–472}$), a shorter variant containing both the PB015164 and the catalytic domains ($RipA_{263–472}$), and the sole C-terminal catalytic domain ($RipA_{332–472}$) (Figure 9).



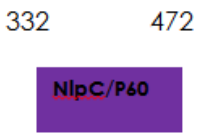


Figure 9. RipA variants. $RipA_{40–472}$ (red, green and purple domains), $RipA_{263-472)}$ (green and purple domains) and $RipA_{332-472}$ containing the sole catalytic domain (purple domain).

Crystals (Figure 10) suitable for X-ray diffraction data collection (0.2 x 0.2 x 0.4 mm) were obtained for $RipA_{263-472}$ and $_{SeMet}RipA_{263-472}$ using 7 mgxmL-1 protein solution and 8% (v/v) 2-Propanol, 16% (w/v) PEG4000, in 60 mM Sodium citrate trihydrate buffer, pH 5.6.

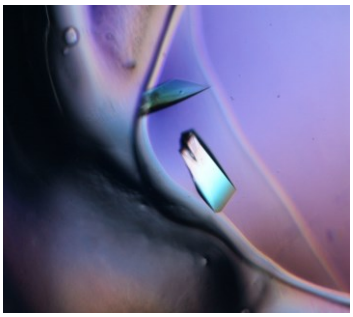


Figure 10. Image of typical RipA crystals

Statistics of data collection for $RipA_{263-472}$ and $_{SeMet}RipA_{263-472}$ are reported in Table 1(pag. 66). Matthews coefficient calculations (Matthews, B.W., 1968) suggested the presence of one molecule per asymmetric unit (VM =1.77, with 30.4% solvent content). The structure of the enzyme was solved by MAD methods using the anomalous signal from the Se atoms of selenomethionine-labeled and refined at 1.1 Å resolution (Table 2, pag. 67).

The overall crystal structure shows that its two domains, the catalytic NlpC/P60 domain and the PB015164 domains identified in the PFAM database, are tightly bound (Figure 11). Surprisingly, we found that the catalytic site cleft of the enzyme is physically blocked by the PB015164 domain. This finding strongly suggests functional inactivity of the enzyme in this form and reveals a zymogenic nature for RipA.

The structure of RipA catalytic domain comprises a central β sheet of six antiparallel β strands, a small two-stranded β sheet and six helices, arranged in a αββαaββββββ topology. Its putative catalytic cysteine (Cys383) is located at the N-terminal end of a helix (α2) and is packed against the sixstranded β sheet core (Figure 12). At this location, Cys383 facets another conserved residue, His432, belonging to the β strand β3 (Figure 12). This histidine is

hydrogen bonded to the side chain of Glu444, which is in turn tightly anchored to Arg453 and Arg458 (Figure 12). Notably, the crystal structure shows that RipA contains a Cys-His-Glu catalytic triad, which is unusual in NlpC/P60 domains (Figure 13). A similar triad (Cys-His-Asp) characterizes another class of cysteine proteases involved in cell separation, denoted as Cysteine Histidine Amino Peptidase (CHAP) domain containing proteases (Bateman and Rawlings, 2003; Rossi et al., 2009).

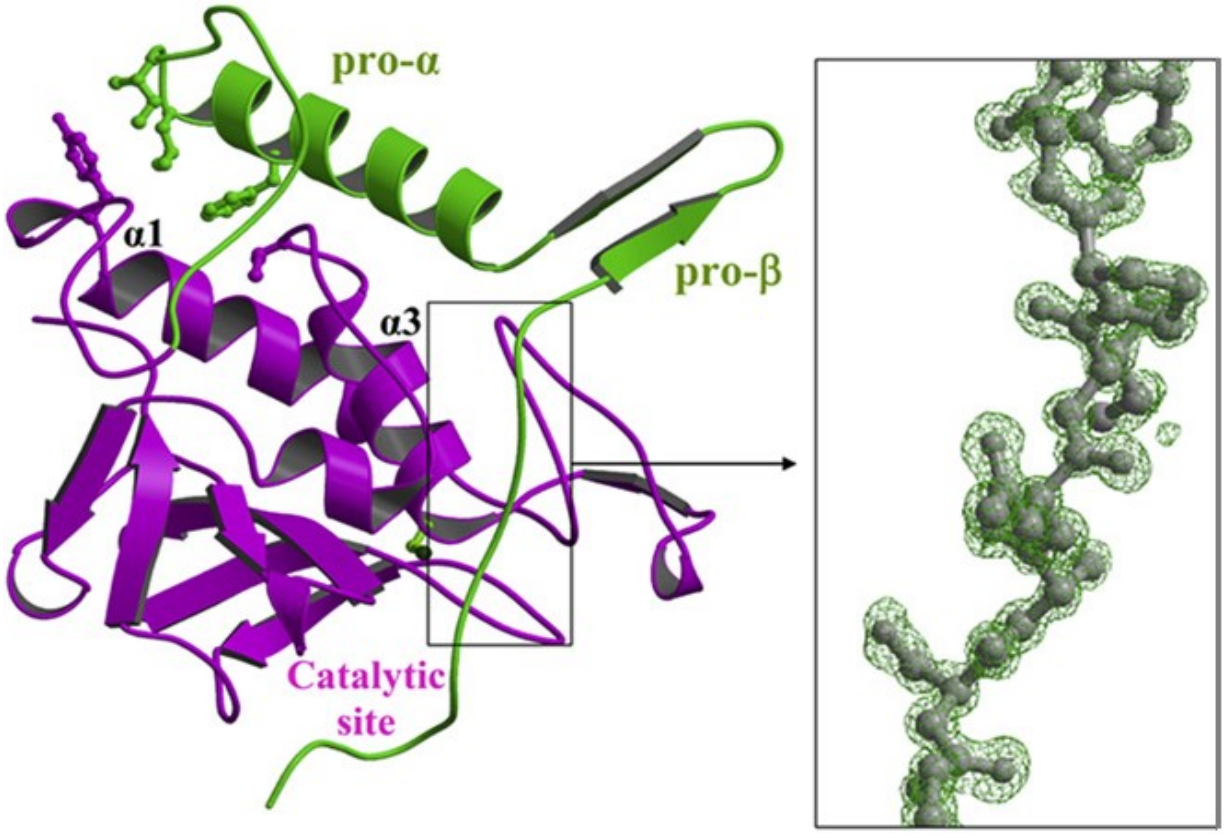


Figure 11. The crystal structure of RipA$_{263–472}$ shows self-inactivation

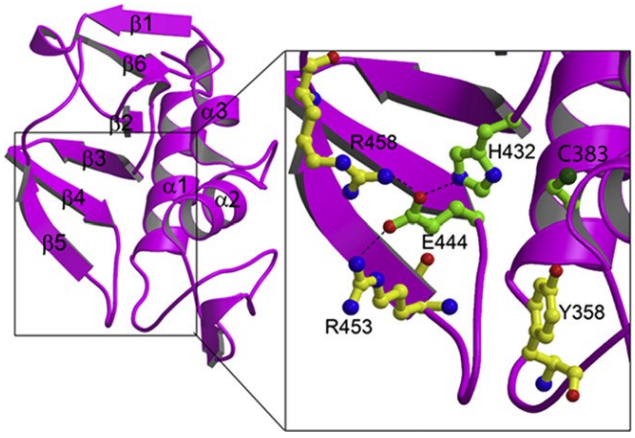


Figure 12. Topology of RipA catalytic domain and zoom of the catalytic site

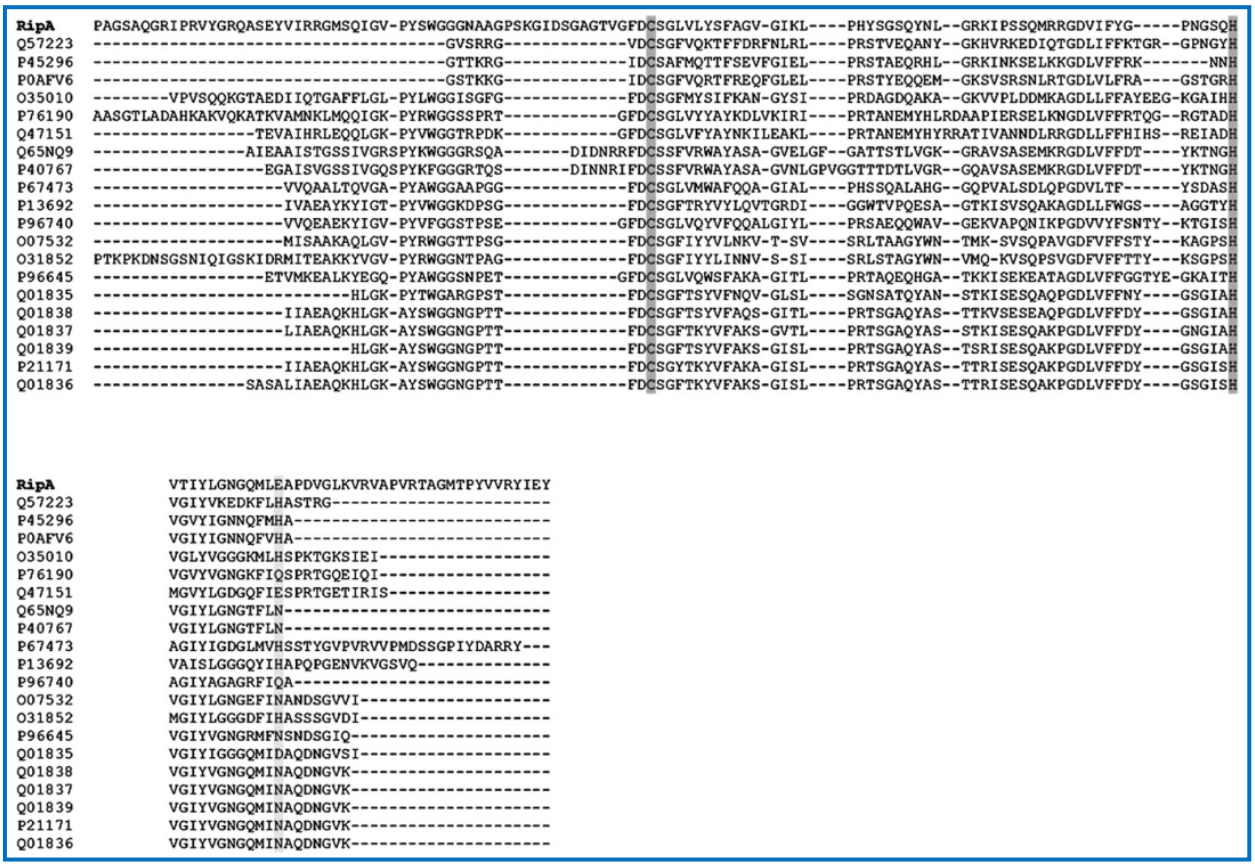


Figure 13. Multiple Sequence Alignment of NlpC/P60 Domains Identified by CONSURF Conserved Cys and His are highlighted in dark grey, the third catalytic residue is shown in light grey.

As mentioned above, the crystal structure shows that RipA catalytic residues are not accessible to potential substrates (Figure 11), as they are locked by the PB015164 domain, here denominated pro-domain. A tight interaction between the catalytic domain and the pro-domain is observed, with a total interaction surface area of 1809 Å. Whereas the β hairpin is nearly completely solvent exposed (Figure 11), both the catalytic-cleftblocking loop region and the a helix, hereafter denoted as pro-a, establish several interactions with the main core of the RipA mature enzyme (Figure 11). Besides providing the first structural description of the pro-domain, our results elucidate its functional role. Using CD spectroscopy, we observed that the recombinant $RipA_{332–472}$, lacking the pro-domain, is still able to fold (Figure 14).

This finding suggests that the pro-domain is not needed for enzyme folding, as observed for proteases of the cathepsin family (Menard et al.,1998). Instead, the X-ray structure reported here highlights a role of the pro-domain as a regulator of the enzyme catalytic activity. With this concept in mind, we explored the possible mechanism that activates the zymogenic form of RipA and unlocks the catalytic site. A suggestion comes from the X-ray structure,

which displays a peculiarly long and fully solvent exposed loop connecting the catalytic- and the pro-domain.

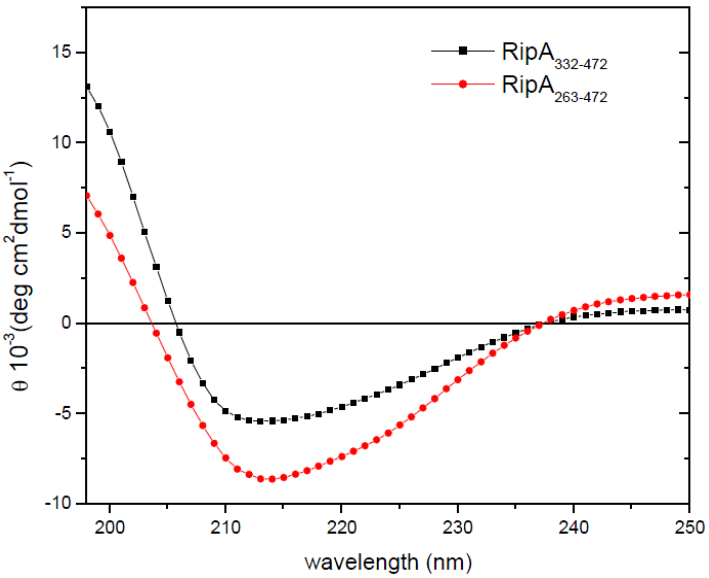


Figure 14. CD spectra of RipA$_{263-472}$ (red) and of RipA$_{332-472}$ (black)

To investigate RipA vulnerability to proteases, we carried out limited proteolysis experiments by trypsin, thermolysin, and pronase whereas to investigate RipA activity we carried out cell wall degradation experiments using fluorescence spectroscopy. To elucidate the role of the putative catalytic triad in catalysis, we produced several mutants of RipA to repeat functional studies on bacterial cell wall and to investigate their crystal structure.

I.4.2 Limited proteolysis experiments

Using protease/enzyme concentration ratios of 1:500, these proteases were able to degrade RipA$_{263–472}$ variant to fragments of about 15 kDa (Figure 15A) (see experimental procedures). This molecular weight, also checked by mass spectrometry, corresponded to the catalytic domain (RipA$_{332–472}$ variant). Cleavage sites by these proteases were identified using LC-MS analysis. As shown in Figure 15A, all cleavage sites are located in the large and fully solvent exposed loop, which connects the catalytic NlpC/P60 domain and the pro-domain. This shows that the connecting loop is a vulnerability point for the enzyme, as it is highly sensitive to proteolytic degradation (Figure 15B). In addition, LC-MS analysis shows that, after proteolysis at the connecting loop, the pro-domain is released and further proteolyzed.

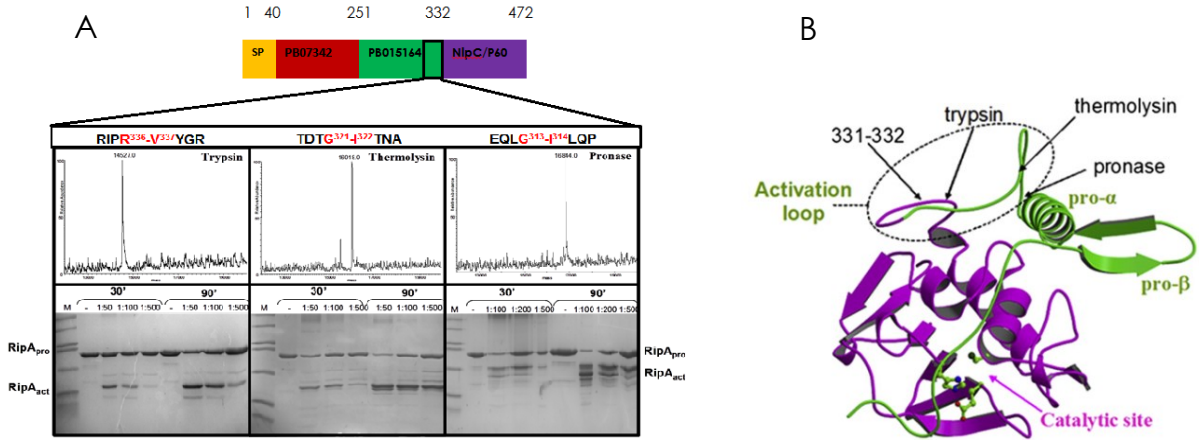


Figure 15. Limited proteolisys studies and location of proteolytic cleavage sites
A) Schematic representation of RipA domains, as defined by PFAM. The region that was found to be sensitive to degradation is indicated by a black box. The figure shows the mass spectra of main RP HPLC peaks after limited proteolysis of $RipA_{263–472}$ with trypsin, themolysin, and pronase (cleavage sites are indicated) and the coomassie stained SDS-PAGE after limited proteolysis of $RipA_{263–472}$ with trypsin, themolysin, and pronase. Molecular mass markers (Sigma, Low range) from the top to bottom: 45, 36, 29, 24, 20, 14.2, 6.5 kDa; B) Ribbon structure of $RipA_{263–472}$ with proteolytic cleavage sites highlighted.

I.4.3 Cell wall degradation experiments

Cell wall degrading experiments were carried out using lyophilized cells of *Micrococcus lysodeikticus* labeled with FITC (see experimental procedures). Both recombinant $RipA_{263–472}$ and $RipA_{332–472}$ were incubated with the FITC-labeled cells. PG hydrolyzing activity of RipA forms was evaluated by fluorescence spectroscopy as the amount of FITC-labeled cell wall released in solution. Fluorescence measurements of supernatants at time intervals allowed us to characterize the time course of cell degrading reaction. The most evident result is that $RipA_{332–472}$ is able to degrade FITC-labeled cell wall because early incubation time intervals and its kinetics reaches a plateau after 20 hours (Figure 16A). By contrast, $RipA_{263–472}$ displayed hardly any activity after 20 hours (Figure 16B). To investigate the role of the putative catalytic Cys383 in catalysis, we used mutant of $RipA_{332–472}$C383A in the cell degrading experiments and we observed that its hydrolytic activity was completely suppressed (Figure 16C). The same suppression of RipA activity was observed on treatment of the enzyme with the sulfhydryl-alkylating reagent iodoacetamide (data not shown).

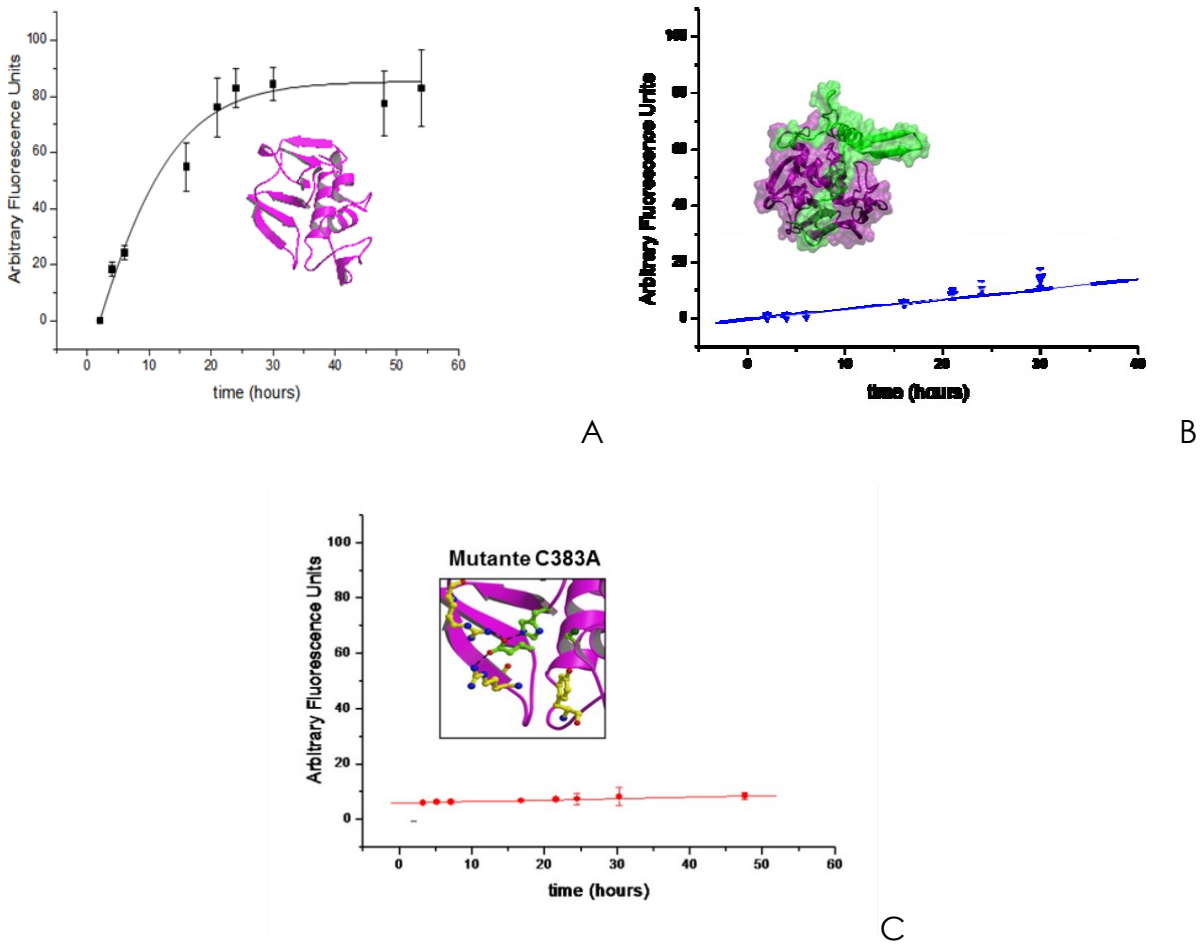


Figure 16. Cell wall degrading activity of: A)RipA$_{332–472}$, B) RipA$_{263–472}$ and C) RipA$_{332-472}$C383A. Values were measured as the amount of FITC-labeled cell wall released in solution on incubation with the enzymes.

These results experimentally prove that RipA is a cysteine protease, as they demonstrate the fundamental role of Cys383 in catalysis. Furthermore, they confirm the result we derived from the X-ray structure that RipA exists as an inactive zymogen, and that removal of the pro-domain activates the enzyme.

Because we observed some level of instability to proteolytic degradation of RipA variants produced, in particular on storage, we checked whether enzyme processing is an autocatalytic event, as observed for the lysosomal cysteine proteases cathepsins (Menard et al., 1998; Rozman et al., 1999; Wittlin et al., 1999). To this aim, we carried out a typical assay for the assessment of self-processing (Rozman et al., 1999), where the inactive mutant RipA$_{263–472}$ (C383A) was incubated with the activated Rip$_{332–472}$ form at room temperature. After 16 hours, SDS-PAGE analysis showed that RipA$_{332–472}$ is unable to process RipA$_{263–472}$C383A. This result unambiguously showed that RipA is not self-activated, but

it needs an activating mechanism. Thus far, it is unclear which mechanism is able to activate RipA. The strong susceptibility of the loop connecting the catalytic and prodomains to proteolytic degradation, as well as the tight interdomain interactions observed in the X-ray structure (Figure 11) strongly suggests that activation proceeds via proteolytic cleavage. However, because RipA is able to interact with RpfB (Hett et al., 2008; Hett et al. 2007), we cannot exclude a possible alternative activation mechanism during resuscitation from latency, which involves a conformational change of RipA on RpfB binding and a consequent unlock of RipA catalytic site. It should be noted, however, that RipA has a fundamental role in cell disconnection of actively growing bacteria. The collective dispensability of all five Rpfs in in vitro bacterial growth (Kana et al., 2008) makes an activation mechanism via proteolytic processing more plausible, at least during bacterial active growth. Consistent with this observation, activation via proteolytic maturation is the typical activation mechanism of cysteine proteases (Schroder and Tschopp, 2010; Sripa et al., 2010).

Our results provide a step forward in the molecular knowledge of the structural and functional properties of RipA in the PG degradation process. We demonstrate that this enzyme, vital to MTB, exists in a zymogenic form that needs to be activated. This evidences a regulation mechanism that is likely precious for a fundamental process like daughter cells separation. Finally, RipA is a secreted enzyme and easily accessible to drugs.

I.4.4 Functional studies of RipA$_{332-472}$ mutants

To evaluate the impact of the other two putative catalytic residues His432 and Glu444 on the enzyme catalytic function, we measured the cell wall degradation ability of these mutants. Results show that these mutations have the same dramatic effect on the enzyme functionality observed for C383A mutant, with a complete loss of PG degradation activity (Figure 17). Intriguingly, despite the lower sequence conservation of Glu444, we observe this residue is fundamental for RipA catalysis.

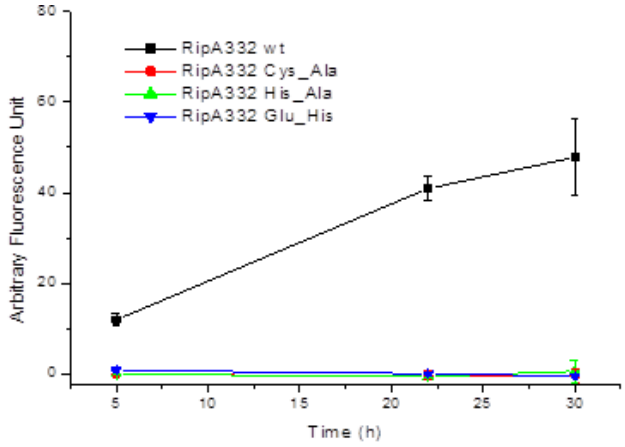


Figure 17. Cell wall degrading activity of RipA$_{332–472}$ and RipA$_{332–472}$C383A, H432A, E444H mutants.

Using CD spectroscopy, we checked conformational properties of all RipA$_{332-472}$ mutants, to exclude that the loss of activity was due to a decrease of the structural content of the enzyme upon mutation. Results show that all mutations do not affect RipA structural integrity, since CD spectra of mutants are superimposable to that of the un-mutated form (Figure 18).

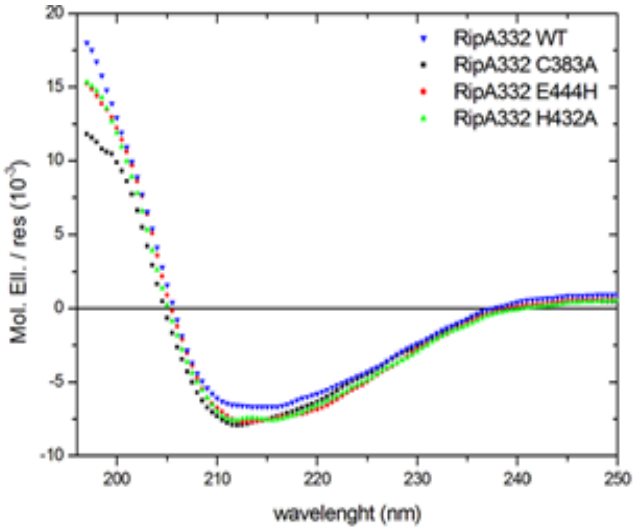


Figure 18. CD spectra of RipA$_{332-472}$ mutants

I.4.5 Crystal structures of RipA mutants

To evaluate the effect of catalytic site mutation on the architecture of RipA catalytic site, we determined crystal structures of the two mutants, RipA$_{263-472}$C383A and RipA$_{263-472}$H432A. Atomic resolution data were collected for both mutants (Table 3, pag.68). Differently, despite a complete conservation of protein secondary structure, we were unable to crystallize the E444H mutant,

likely due to a destructuration of the locking peptide induced by the +2 charge difference. An analysis of electron density maps for RipA$_{263-472}$C383A showed that the catalytic site of RipA remains virtually identical in the presence of the cysteine mutation. Indeed, the cavity formed by the lack of the Cys side chain is filled by a water molecule and the conformation of all residues of the catalytic side cleft is fully preserved (Figure 19A). Similarly, His mutation has marginal effect on RipA structure. In this case, the cavity formed by the lack of His side chain is filled by two water molecules (Figure 19B).

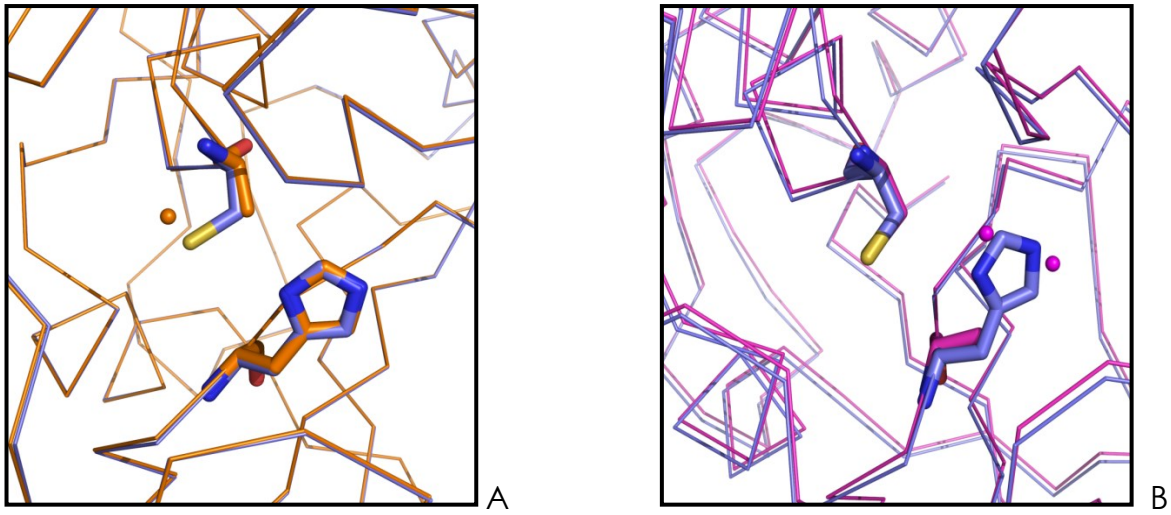


Figure 19. Comparison of the RipA crystal structure with mutants crystal structures. A) RipA catalytic site (indigo) vs RipA$_{263-472}$C383A catalytic site (orange). B) RipA catalytic site (indigo) vs RipA$_{263-472}$H432A (magenta).

I.4.6 Modeling of substrate in RipA catalytic site

Compared to NlpC/P60 containing enzymes, RipA structure is unique in possessing a domain (Aramini et al., 2008; Xu et al., 2009), the PB015164 domain, that physically blocks the enzyme substrate binding cleft. The inaccessibility of catalytic residues to potential substrates in RipA structure makes experimental techniques to investigate enzyme-substrate interactions difficult to be carried out. On the other hand, no information on either the cleavage site on the substrate or the set of interactions between RipA binding cleft and its substrate is hitherto known. Therefore, we adopted modeling techniques to investigate interactions of the muropeptide GlcNAc-MurNAc-L-Ala-γ-D-Glu-mesoDAP-D-Ala (Figure 20) with RipA catalytic domain.

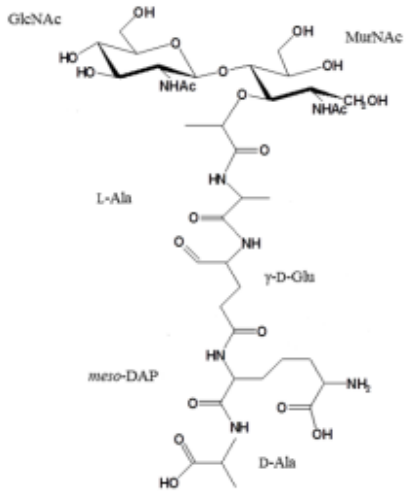


Figure 20. Muropeptide GlcNAc-MurNAc-L-Ala-γ-D-Glu-mesoDAP-D-Ala

Molecular modeling, followed by energy minimization with GROMACS (Lindahl et al., 2001), produced key information on the interaction mode of muropeptides in RipA binding cleft. Notably, the shape of RipA catalytic site cleft well accounts for the branched nature of PG. Consistently, the conformation of GlcNAc-MurNAc-L-Ala-γ-D-Glu-mesoDAP-D-Ala is well locked by the catalytic site cleft. Both the carbonyl group and the side chain of DAP are tightly bound to the enzyme, through hydrogen bonding interactions with Gln431 and His432 (Figure 21). Also, the side chain of γ-D-Glu is h-bonded to the side chain of Ser384. Notably, the peptide group between γ-D-Glu and DAP forms several h-bonding interactions with the enzyme and directly contacts the catalytic residues Cys383 and His432. Beside h-bonding interaction with Gln431, the backbone N of DAP forms an h-bond with the nd1 atom of His432 side chain (Figure 21). Also, the carbonyl group of γ-D-Glu forms h-bonds with both the backbone nitrogen and the side chain of Cys383. Differently, Glu444 does not establish interactions with the substrate, but it hydrogen bonds His432. This feature confirms, as previously proposed that the role of Glu444 is to the properly orient the side chain of His432 (Ruggiero et al., 2010). By h-bonding the Ne2 atom of His432, Glu444 locks His432 in an orientation suitable to form an h-bond, through its Nd1 atom, with the backbone nitrogen of the DAP moiety. The observed conformation brings the peptide group linking γ-D-Glu to DAP in close vicinity to Cys383/His432

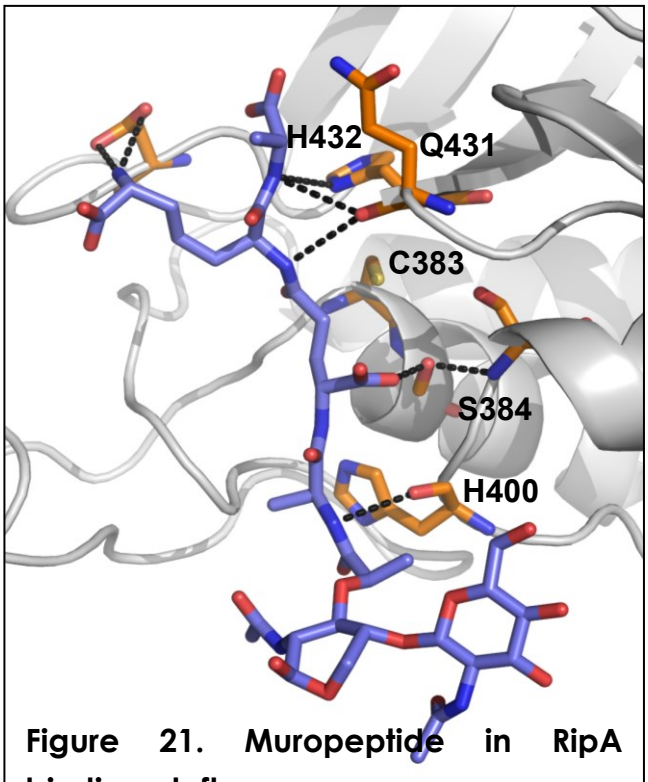


Figure 21. Muropeptide in RipA binding cleft

catalytic dyad, a finding which suggests that RipA acts by cleaving the peptide bond between γ-D-Glu and DAP. The identified set of interactions is currently being used for the design of RipA inhibitors.

I.5 CONCLUSION

The process of daughter cell separation requires a delicate balance of cell wall hydrolases that cleave the septa connecting daughter cells. Cell-separating enzymes usually contain endopeptidase domains like CHAP or NLPC/P60 domains, and/or Glucosaminidase domains (Layec et al., 2008; Layec et al., 2009; Rossi et al., 2009). RipA has a remarkable effect on the bacterial phenotype, since *ripA* depletion strains in *M. smegmatis* exhibit a decreasing growth and an abnormal phenotype, consisting in branching and chaining bacteria (Hett and Rubin, 2008). This enzyme is believed to cleave PG peptide crosslinks (Hett et al., 2008), similar to other cell separating endo-peptidases, like LytE, LytF, CwlS from *B. subtilis* (Fukushima et al., 2008; Margot et al., 1999; Margot et al., 1998) and Spr from *E. coli* (Aramini et al., 2008).

RipA was cloned, expressed and purified in native conditions. Crystals were obtained using vapor diffusion techniques. The structure of RipA was solved by MAD using the anomalous signal from the Se atoms of labeled enzyme. Crystallographic studies of RipA have yielded new insights in the functional regulation of this enzyme. Indeed, the crystal structure clearly reveals a zymogenic nature of RipA, a finding which is confirmed by cell wall degradation assays (Ruggiero et al., 2010). Furthermore bacterial cell-wall degradation assays with mutated forms of RipA clearly identify the key catalytic triad of this enzyme. Finally, modeling of a PG fragment in the catalytic site of RipA has provided a key set of interactions, which will be exploited for the design of novel anti-tubercular therapeutics.

PART II - Regulation of resuscitation from dormancy by STPK kinases and Penicillin Binding Proteins: the role of PASTA domain

PARTII - Regulation of resuscitation from dormancy by STPK kinases and Penicillin Binding Proteins: the role of PASTA domains

II.1 AIM OF THE STUDY

Bacterial resuscitation from dormancy is a complex mechanism associated to stress and nutritional limitation and involves the recognition by bacteria that external conditions are optimal for growth. As such, this phenomenon requires a fine regulation which involves several molecular players active in muropeptide modeling and recognition.

Among those, eukaryotic-type STPK PrkC have been shown to be essential for resuscitation of dormant *B. subtilis* (Shah and Dworkin, 2010; Shah et al., 2008). Generally, growing bacteria release muropeptides in the surrounding environment, due to cell wall PG remodeling associated to cell growth and division (Shah and Dworkin, 2010; Shah et al., 2008; Ruggiero et al., 2010). Therefore, the presence of muropeptides in the close environment of dormant spores is a clear signal that conditions are optimal for growth. Interestingly, *B. subtilis* spores germinate in response to m-DAP containing muropeptide, which constitutes *B. subtilis* cell wall (Shah et al., 2008).

In order to understand the molecular mechanism allows PrkC to sense muropeptides in the bacterial *milieu*, we undertook a structural study of the extracellular region (EC) of PrkC from *B. subtilis* and *S. aureus*. Furthermore, by coupling crystallographic studies with bioinformatic analyses, calorimetric and spectroscopic studies we exploited the structural requirements necessary for recognition and binding and proved that m-DAP containing muropeptide physically binds to a specific PASTA domain of EC-PrkC from *B. subtilis*.

Based on these results, we carried out structural studies of the PASTA domain of another protein from MTB, PonA2, which was shown to play an important role for the adaptation to dormancy of MTB (Patru et al., 2010). We hypothesized that the PASTA domain of PonA2 could act in a similar fashion as that of PrkC. Therefore, we applied NMR, calorimetric and computational methods to unveil structural and dynamic properties of PonA2 PASTA domain.

II.2 - BACKGROUND

Eukaryotic-type STPK are expressed in many prokaryotes including a broad range of pathogens. Proteins of this class are known to regulate various cellular functions, such as developmental processes, secondary metabolism, stress response, biofilm formation, antibiotic resistance, virulence, cell wall biogenesis, cell division, and central metabolism (Madec et al., 2002; Fiuza et al., 2008). The first reported eukaryotic-type STPK, Pkn1 from *Myxococcus xanthus*, was found to be required for normal bacterial development (Mu˜noz-Dorado et al., 1991). The advance of genome sequencing has prompted the identification of similar kinases in many bacteria.

Previous studies have found that the eukaryotic-type STPK PrkC from *Bacillus subtilis* is also involved in bacterial exit from dormancy (Shah, 2008; Shah and Dworkin, 2010). Under conditions of nutritional limitation, *B. subtilis* produces spores which are resistant to harsh environmental conditions and can survive in a dormant state for years (Shah, 2008; Kana et al., 2010; Keep et al., 2006). The process of resuscitation is called, in these sporulating bacteria, germination (Shah, 2008; Shah and Dworkin, 2010).Generally, growing bacteria release muropeptides into the surrounding environment, owing to cell wall PG remodeling associated with cell growth and division (Hett et al., 2008; Ruggiero et al., 2010). Therefore the presence of muropeptides in the close environment of dormant spores is a clear signal that the conditions are optimal for growth. Consistently, Shah et al. (2008) reported that m-DAP-containing muropeptides are powerful germinants of *B. subtilis* spores. Notably, other authors have proposed independently the idea that the activating ligand of PrkC could be a component or a degradation product of the cell wall PG (Absalon et al., 2009). These authors also showed that, once activated, PrkC is able to phosphorylate the small-ribosome-associated GTPase CpgA, the translation factor EF-Tu (elongation factor thermo-unstable) and a component of the bacterial stressosome, denoted YezB (Absalon et al., 2009). On the other hand, the possible involvement of EF-G (elongation factor G) as a substrate of PrkC is controversial (Absalon et al., 2009; Shah et al., 2008; Gaidenko et al., 2002). A close homologue of PrkC exists in *Staphylococcus aureus*, a significant human pathogen that causes a number of infections ranging from skin infections to toxic shock syndrome, osteomyelitis and myocarditis (McGahee et al., 2000; Ohlsen et al., 2010). PrkC from *S. aureus* is predicted to be a membrane protein, similar to its homologue from *B. subtilis*. The two kinases present the

same domain organizations, with an intracellular serine/threonine kinase domain, a transmembrane region and an extracellular portion that contains three domains known as Penicillin binding Associated and Serine/Threonine kinase-Associated (PASTA) domains. According to Shah et al. (2008), *B. subtilis* spores germinate in response to m-DAP-containing muropeptide, which constitutes the *B. subtilis* cell wall, but not in response to L-lysine-containing muropeptide. Moreover, when the PrkC from *B. subtilis* is replaced by PrkC from *S. aureus*, which is characterized by L-lysine-containing cell walls, germination is observed both in response to m-DAP and L-lysine-containing muropeptide. This finding has shown that the source of PrkC determines the bacterial ability to respond to muropeptides and has suggested that PrkC extracellular domains exhibit specificity of muropeptide binding (Shah et al. 2008).

PASTA domains also exist in PBP (Patru et al., 2010). The crystal structure of the PBP2x from *Streptococcus pneumoniae*, which contains two C-terminal PASTA domains, was solved in complex with cefuroxime, a β-lactam antibiotic mimicking the unlinked PG (Gordon et al., 2000; Dessen et al., 2001). Interestingly, the PBP PonA2 from MTB contains one PASTA domain and has been associated with regulation of bacterial resuscitation from dormancy. These findings suggest that PASTA domains can sense muropeptides. Despite these hypotheses, structural results are still needed to help the understanding of the function of extracellular PASTA domains.

II.3 - EXPERIMENTAL PROCEDURES

II.3.1 Expression and purification of Extra Cellular region of PrkC from *S. aureus* and *B. subtilis*

The plasmidic constructs corresponding to the EC region of *S. aureus* PrkC (EC-PrkC *S.a*, residues 378- 664) and of *B. subtilis* (EC-PrkC *B.s*, residues 356-648) have been prepared as previously described by Shah et al. (2008). The overexpression of EC-PrkC *S.a* and of EC-PrkC *B.s* , containing an N-terminal histidine tag (His6) were carried using *E. coli* DH5a cells (Invitrogen). Briefly, an overnight starting culture of 100 mL were prepared to grow 1 L of LB containing 100 µg/mL ampicillin. Induction was performed at OD 600=0.6 with 0.02% (w/v) L-arabinose and harvested after four hours. The proteins were isolated by sonicating resuspended cells in a binding buffer (10 mM imidazole, 200 mM NaCl, 20 mM TrisHCl, 5%(v/v) glycerol pH 8.0) containing a protease inhibitor

cocktail (Roche Diagnostic). The lysate were cleared by centrifugation at 18,000 rpm, and the supernatants loaded on a 5-mL Ni^{2+}-NTA column (GEHelthcare), equilibrated with binding buffer. After washing with 10 volumes of binding buffer, a linear gradient of imidazole (10-300 mM) was applied to elute the proteins. Fractions containing the proteins were pooled and loaded on a gel filtration column (Superdex 200, GeHelthcare), equilibrated in 150 mM NaCl, 20 mM TrisHCl buffer at pH 8.0 and 5% (v/v) glycerol, for a further purification step. The proteins eluted in a single peak and were homogeneous, as judged by SDS–PAGE analysis. All the purification steps were carried out at 4°C. The proteins were concentrated using a centrifugal filter device (Millipore), and the concentration were determined by Bradford protein assay (Biorad). Fresh concentrated proteins at 10 mgxmL-1, were used for crystallization experiments.

II.3.2 Cloning, expression and purification of $^{15}N/^{13}C$ labeled PONA2 PASTA domain from *M. tuberculosis*

The region encoding amino acids G700 to I764 of PONA2 was amplified by PCR using genomic DNA of MTB as template (H37Rv strain, Figure 4A) and the following oligonucleotides as primers (ponA2 forward: CATGCCATGGGCTCACGGGTACCAAGC; ponA2 reverse: CCCAAGCTTATCAGATGCCGTTGCTGATCTGG). NcoI/HindIII-digested fragments were cloned into the pETM-11 (Figure 4B) with N-Terminal Histag 6xHis tag for metal-affinity purification. The resulting positive plasmid was used to transform *E. coli* Star BL21(DE3) competent cells. Transformed cells were cultured overnight in LB broth with 50 µg ml^{-1} kanamycin at 37°C . For the production of isotope labeled sample ($^{15}N/^{13}C$) the culture was seeded in 1:100 volume ratio either in 1 L of minimal media (M9) containing (422mM of Na_2HPO_4, 220mM of KH_2PO_4, 85,5 mM NaCl, 186,7 mM of ^{15}N ammonium chloride, 1mM $MgSO_4$, 0.2mM $CaCl_2$, 1 mL of Thiamine 40 $mgxml^{-1}$ and 0.3% final of ^{13}C-glucose. Culture was grown at 37° C in a shacking incubator, induced with 0.7 mM IPTG, and further grown at 22°C for 18 hours for protein production. *E. coli* cells were harvested by centrifugation at 6000 rpm for 20 min, and the bacterial pellet re-suspended in a buffer containing 300 mM NaCl, 50mM Tris-HCl, 10mM m imidazole, 5% (v/v) glycerol and complete protease inhibitor cocktail (Roche) pH 8 and lysed by sonication on ice to release recombinant protein. The cell lysate was centrifuged at 16500 rpm at 4°C for 30 min and the surnatant loaded on Ni^{2+}-derivatized HisTrap columns

(GE Healthcare). After washing with 10 volumes of binding buffer, a linear gradient of imidazole (5-300 mM) was applied to elute the protein. Eluted protein was dialyzed against buffer without imidazole containing 150 mM NaCl, 50 mM Tris-HCl, 5% glycerol pH 8.0 at 4°C for 4 hours and after was digested with TEV pretease to remove the 6xHis tag. The protein was further purified by a second Ni^{2+} affinity chromatography and by a gel filtration on Superdex75 (GE Healthcare) with a buffer containing 150mM NaCl, 50mM Tris-HCl, pH 8. The protein, which eluted in a single peak, was concentrated using a centrifugal filter device (Millipore) with a cut-off of 3KDa, and the concentration was determined using the Pierce BCA Protein Assay Kit. The fresh concentrated protein (1mM) was dialyzed against the NMR buffer containing 30mM sodium phosphate buffer pH 6.5. $^{15}N/^{13}C$ labeled protein was prepared with 5% of sodium azide and 10% of deuterated water for NMR experiments.

II.3.3 Circular dichroism experiments

CD spectra were recorded with a Jasco J-810 spectropolarimeter equipped with a Peltier temperature control system (Model PTC-423-S). Molar ellipticity per mean residue, $[\theta]$ in deg $cm^2 \cdot dmol^{-1}$, was calculated from the equation: $[\theta] = [\theta]obs \cdot mrw \cdot (10 \cdot l \cdot C)^{-1}$, where $[\theta]obs$ is the ellipticity measured in degrees, mrw is the mean residue molecular mass (111.17 Da for EC-PrkC *S.a*, 110.11 Da for EC-PrkC *B.s* and 102 Da for PonA2-PASTA, respectively), *C* is the protein concentration in $g \cdot L^{-1}$ and *l* is the optical path length of the cell in cm. Far-UV measurements (190-260 nm) were carried out at 20 °C using a 0.1 cm optical path length cell and protein concentration of 0.2 $mg x mL^{-1}$.

II.3.4 Crystallization experiments

Crystallization trials were performed at 293 K using the hanging-drop vapor-diffusion method. Preliminary crystallization conditions were set up using a robot station for high throughput crystallization screening (Hamilton STARlet NanoJet 8+1) and commercially available sparse-matrix kits (Crystal Screen kits I and II, Hampton Research, Index).

The initial screenings revealed several promising conditions for crystallization of EC-PrkC *S.a*. Optimization of the crystallization conditions was performed manually both using the vapor-diffusion and the macroseeding technique. After optimization of the crystallization conditions, only tiny needles could be obtained. Best crystals of EC-PrkC *S.a* were obtained using the macro-seeding

technique. Needles were partially dissolved in their mother liquor and transferred to solutions containing mixtures, at 1:1 ratio, of 10 mgxmL-1 EC-PrkC *S.a* and dilutions of commercially available solutions (Hampton Research). $LuCl_3$-derivative crystals were prepared by soaking a native crystal in a solution containing 1–3 mM $EuCl_3$, 25% (w/v) MPEG 2000 [methoxypoly(ethylene glycol) 2000 kDa], 160 mM $(NH_4)2SO_4$ and 60 mM sodium acetate trihydrate buffer for 3 hours at pH 4.6.

II.3.5 Data collection and processing of EC-PrkC from *S. aureus*

A MAD experiment was carried out on a crystal derivatized with 2 mM $LuCl_3$ at the X12 synchrotron beamline, DORIS storage ring, DESY (Deutsches Elektronen Synchrotron; Hamburg, Germany), at 100 K (Table 4, pag 69). Cryoprotection of the crystals was achieved by a fast soaking in a solution containing glycerol to a final concentration of 10% (v/v). These data, extending to 3.0 Å (1Å=0.1 nm) resolution, allowed us to build a part of the molecule. Derivatization by overnight soaking in a solution containing 2 mM $EuCl_3$ provided higher resolution X-ray data, at 2.15 Å (Table 3, pag. 68). Data collection was performed in-house at 100 K using a Rigaku Micromax 007 HF generator producing Cu Ka radiation and equipped with a Saturn944 CCD (charge-coupled device) detector. The datasets were scaled and merged using the HKL2000 program package (Otwinowski et al., 1997)(Table 3, pag. 68).

II.3.6 Structure determination of EC-PrkC from *S. aureus*

Phasing was achieved using in-house SAD (Single-wavelength Anomalous Diffraction) data, collected on a crystal soaked overnight in a solution containing 2 mM $EuCl_3$. These data provided structural factor phases to 2.2Å resolution. Both SHELXD (Sheldrick et al., 2008) and SOLVE (Terwilliger, 2004) identified five Eu^{3+} ions. Phases, improved by phase extension and density modification by RESOLVE (Terwilliger, 2004) and wARP (Langer e al., 2008), allowed us to trace nearly the entire molecule structure. Crystallographic refinement was carried out against 95% of the measured data using the CCP4 program suite (Potterton et al., 2003) . The remaining 5% of the observed data, which was randomly selected, was used in Rfree calculations to monitor the progress of refinement. Structures were validated using the program PROCHECK (Laskowski et al., 1996).

II.3.7 Homology modeling of EC-PrkC from *B.subtilis*

The three-dimensional structure of EC-PrkC was modeled by comparative protein modeling methods using the MODELLER workspace. The model was built using the 2.2 Å resolution structure of the EC region of PrkC from *S. aureus* (Protein Data Bank (PDB) entry 3py9) as a template. The degree of sequence identity between the template (residues 377-660) and EC-PrkC (residues 361-643) is 28.7%, which enabled a preliminary model to be generated. This model was energetically minimized using GROMACS (Lindahl E. et al., 2001).

II.3.8 Sequence alignments and mutants design

For sequence conservativeness studies, residue conservation scores were obtained using the program CONSURF. Homologous sequences were collected using PSI-BLAST heuristic algorithm with default parameters (Kaushik S., et al., 2013). The search was carried out in the full UNI-PROT knowledgebase (SWISS-PROT + TrEMBL) database, using an E-value cutoff of 0.001. The program MUSCLE was used to align the homologues extracted from the PSI-BLAST output file.

For the identification of the DAP interaction site on the protein, a statistical survey was carried out by searching all structures containing DAP in the PDB. This analysis was conducted using the Application Programming Interface implemented in the PDB server.

II.3.9 EC-PrkC from *B. subtilis* mutants production

R500E and R500A mutations were generated by site-directed mutagenesis of wild-type recombinant plasmid pBAD24/EC-PrkC using the Stratagene QuikChange kit. Mutagenic primers for R500E were 5'-CCGAAGACATTACGCTTGAAGACTTGAAAACCTACAG-3' and 5'-CTGTAGGT TTTCAAGTCTTCAAGCGTAATGTCTTCGG-3'. Primers for R500A were 5'-CCGAAGACATTACGCTTGCAGACTTGAAAACCTACAG-3' and 5'-CTGTAGGTTTTCAAGTCTGCAAGCGTAATGTCTTCGG-3'. Introduction of the expected mutations was confirmed by DNA sequencing. Expression and purification of the two mutants were performed using the same conditions used for the wild-type form.

II.3.10 Isotermal titration calorimetry experiments

The interactions of EC-PrkC from *B. subtilis* with natural muropeptides from *Bacillus subtilis* were investigated at 298K by Isothermal Titration Calorimetry

(ITC) using a MicroCal ITC200 calorimeter (GEHelthcare Milan) calibrated with standard electrical pulses. All solutions were degassed by stirring under vacuum before use. 18 consecutive injections of 2 µL aliquots of a 20 mM solution of muropeptides (the molecular weight was calculated based on the 3:1 molar ratio between muropeptides B and A, Figure 27) were added to the calorimeter cell (0.280 mL) containing 0.05 mM of EC-PrkC at intervals of 150.

Similarly, interactions of ponA2-PASTA with synthetic mur-TriDAP (InvivoGen) were checked using a protein concentration of 0.15 mM, which were titrated with 15 mM mur-TriDAP.

In all experiments, proteins and titrants were prepared in the same buffer (PBS 1X) to minimize the contribution of the dilution heat. In order to ensure proper mixing after each injection, a constant stirring speed of 1000 rpm was maintained during the experiments. Data were analyzed using a 'one set of sites' binding model.

II.3.11 NMR experiments of ponA2-PASTA domain from *M. tuberculosis*

All NMR experiments were recorded at 298 K on Inova 600 MHz spectrometer (^{1}H-^{15}N experiments), equipped with a cryogenic probe optimized for 1H detection, or Varian Inova 500 MHz (^{1}H-^{13}C-^{15}N triple resonance experiments), using the standard pulse sequences. NMR samples consisted of approximately 1.4-1.8 mM unlabeled or uniformly ^{15}N or ^{15}N-^{13}C doubly labeled protein dissolved in 30 mM sodium phosphate (pH 6.4), 0.02% sodium azide and 10% 2H2O. ^{1}H chemical shifts were directly referenced to the methyl resonance of TSP, while ^{13}C and ^{15}N chemical shifts were referenced indirectly to the absolute $^{13}C/^{1}H$ or $^{15}N/^{1}H$ frequency ratios.

The ^{1}H-^{15}N- Heteronuclear Single Quantum Coherence (HSQC) spectra were recorded with a number of complex points (cp) and acquisition times equal to 128 cp and 64 ms for ^{15}N (F1 dimension) and 1024 cp and 146 ms for ^{1}H (F2 dimension). The assignment of ^{1}H and ^{15}N resonances was achieved using a suite of heteronuclear 2D and 3D spectra: ^{1}H-^{15}N-HSQC, 3D [^{1}H, ^{15}NH]-TOCSY-HSQC and, 3D [^{1}H, ^{15}NH]-NOESY-HSQC. The assignment was extended to ^{13}C resonances and globally assessed by analyzing HCCH-TOCSY and triple resonance (HNCO, HNCACN, CBCA(CO)NH) experiments. NMR experiments were processed using the software Varian (VNMR 6.1B). The program CARA was used to analyze and assign the spectra. The assignments of the aliphatic

side-chain ^{1}H resonances remain in some cases ambiguous due to cross peaks overlapping. Globally, the NMR analysis of the double label-ed (^{15}N-^{13}C) PonA2-PASTA allowed us to confidently assign 98 % of the resonances.

The structure calculation was performed by program CYANA 2.1 starting from 150 random conformers. 3D structure of PonA2-PASTA domain was determined based on 465 experimental NOE constraints and 87 constraints on Φ and φ torsion angles as obtained on the basis of backbone (HN, ^{15}N, ^{13}Ca, $^{13}C'$, $^{13}C\beta$) chemical shifts using TALOS+. The 40 conformers with the lowest final CYANA target function values (TF average value= 1.68 ± 0.19 Å2) showed an average backbone RMSD = 0.62 ± 0.31 Å and an average heavy atom 1.02 ± 0.29 Å. Validation procedure using PROCHECK-NMR (Laskowskiet al., 1993) program demonstrated that the final family of 3D structures is in agreement with the distance restraints. The molecular graphics program MOLMOL (Koradi et al., 1996). was employed to perform the structural statistics analysis.

II.3.12 Relaxation measurements

Backbone ^{15}N relaxation parameters (R_1, R_2 and hetero-nuclear NOE) were obtained for ^{15}N-labeled PASTA sample at the concentration of 200 µM by NMR experiments recorded at 298 K and 600 MHz. Seven relaxation delays were used both for R_1 (0.01, 0.05, 0.1, 0.3, 0.6, 0.8, 1s) and R_2 (0.01, 0.03, 0.07, 0.09, 0.13, 0.15, 0.19 s) measurements. ^{15}N-{1H} hetero-nuclear steady state NOEs were measured with recycling time of 5 s and 3 s of proton saturation period. The rotational correlation time was determined by using the average R_2/R_1 values.

II.4 RESUSTS

II.4.1 Structural studies of EC-PrkC from *S. aureus*

PrkC from *S.a* is a 664-residue-long multidomain protein. Using sequence analysis tools, the PrkC sequence is predicted to embed a serine/threonine kinase domain (residues 1–270), a region of unknown structure and function (residues 271–377) that includes a transmembrane helix (residues 349–373) and an extracellular region, denoted EC-PrkC (residues 378–664) (Figure 22).

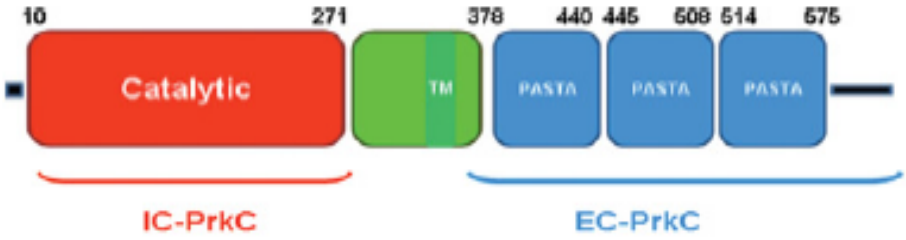


Figure 22. Domain prediction of PrkC, according to the Pfam database

We determined the crystal structure of EC-PrkC *S.a*, which is predicted to contain three successive PASTA domains (Figure 22). We overexpressed this fragment of PkrC in *E. coli* and obtained a soluble form that was amenable for crystallization. Since a sole methionine residue is present in the sequence of EC-PrkC, we tackled structure factor phasing by preparing derivatives of EC-PrkC crystals with several heavy metals. The best results were obtained with lanthanides. Indeed, MAD data collection using a $LuCl_3$-derivatized crystal was performed at DESY. Diffraction data, at 2.9Å resolution, allowed us to trace a significant part of the molecule. Higher diffraction data (2.2Å resolution) were collected in-house on a $EuCl_3$-derivatized crystal. Using the SAD method, these data produced a readily interpretable electron density throughout the entire structure. The final model contained 285 residues, five Eu^{3+} ions and 125 water molecules.

Data collection, refinement and model statistics are summarized in the Table 4 (pag. 69).

II.4.2 EC-PrkC from *S. aureus* has the shape of a golf club

The crystal structure of EC-PrkC revealed that it consists of four consecutive domains. The four domains are arranged sequentially in a golf-club shape, such that only neighbouring domains interact with each other (Figure 23). Three of the four domains are, as predicted, PASTA domains (Figure 22). The structure shows that the three PASTA domains display a linear and regular organization. Indeed, each domain exhibits a two-fold symmetry with respect to its neighbouring domains (Figure 23). In this organization, the sole a-helix of each domain is alternatively located on the two sides of the golf club. Interestingly, the structure reveals the existence of a fourth domain, at the C-terminal end of the molecule, not predicted by searches in the PFAM database (Finn et al., 2008). Furthermore, sequence analyses against the PDB do not identify any significant homologue for this domain.

Figure 23. Ribbon representation of EC-PrkC S.a. Helices and β-strands are shown in orange and blue respectively.

II.4.3 Features of PASTA domains

The present study provides the highest-resolution study of PASTA domains, which are arranged in βαββββ motifs (Figure 23). Despite moderate sequence identity between EC-PrkC PASTA domains (from 21 to 27%), the structures of these domains are well conserved. PASTA domains are also constituents of the PBP2x protein from *S. pneumoniae* (PDB code 1QME) (Gordon et al., 2000). The crystal structure of PBP2x shows that its two PASTA domains form a compact structure with a pronounced interdomain bending (Figure 24A). In this

arrangement, each PASTA domain is involved in interactions with both the other PASTA domain and with the dimerization and transpeptidase domains of PBP2x (Gordon et al., 2000). In contrast, in the structure of EC-PrkC, the three PASTA domains are arranged linearly, with a limited number of interdomain interactions. This finding suggests a high flexibility of the EC-PrkC. The structural comparison of PrkC PASTA domains with those of PBP2x shows a significant conservation of the PASTA fold, despite the low sequence identities, ranging from 5 to 28%, between domains. All secondary-structure elements are conserved, although β-strands β4 and β5 do not appear in lower resolution PBP2x structures (Figure 24B) (Pares et al.,1996).

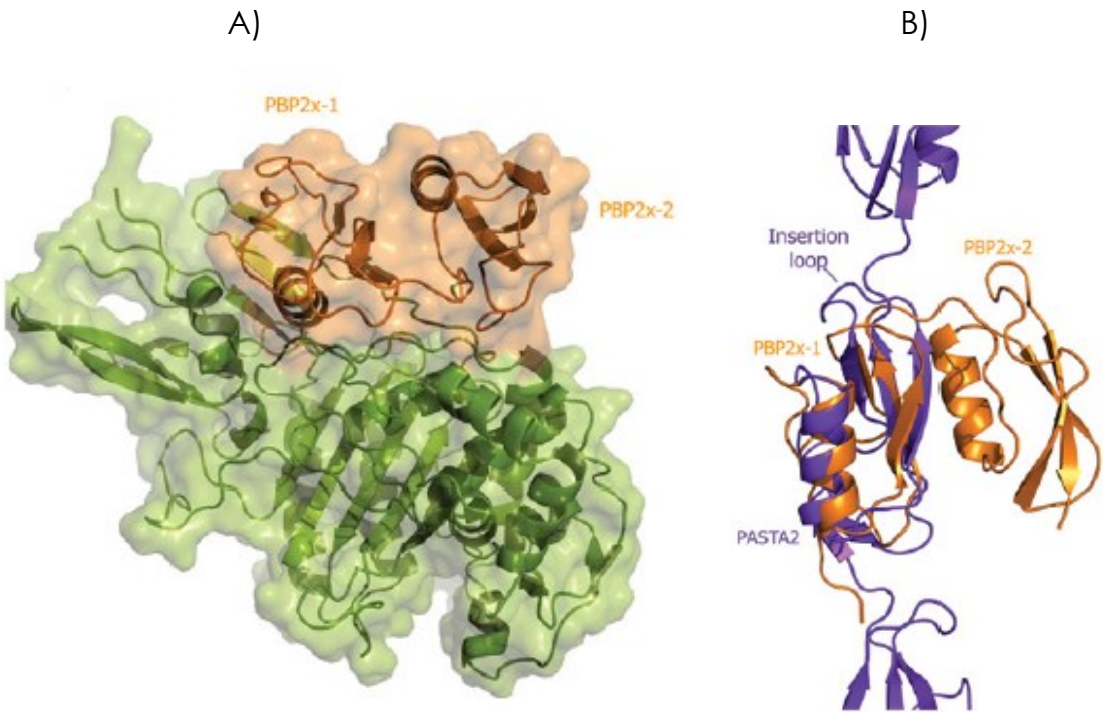


Figure24. A) Superposition of the PASTA domain 1 of PBP2x (PBP2x-1, orange) with the PASTA domain 2 of PrkC (purple). B) Top view of PBP2x structure. The ribbon and surface representations of PASTA domains are coloured orange, whereas representations of the rest of the molecule are shown in green.

Superpositon of EC-PrkC structure with the Small Angle X-ray Scattering (SAXS) structure of the extracellular portion of its homologue PKnB from MTB (four PASTA domains) shows that, similar to PrkC, the organization of PASTA domains in the PknB is nearly linear (Figure 25B) (Gordon et al., 2000; Barthe et al., 2010). However, the overall Root Mean Square Deviation (RMSD) calculated on Ca atoms after superposition of the three PASTA domains of EC-PrkC from *S.a* with those of PknB is as high as 9.5Å (after superposition with PknB PASTA domains 2, 3 and 4) and 14.4Å (after superposition with PknB PASTA domains 1, 2 and 3).

This is mainly due to different arrangements of PASTA domains in PrkC and PknB (Figure 25A, B). Indeed, when single PASTA domains of EC-PrkC and PknB are superposed, a strong conservation of the domain fold is observed, even in cases when sequence identities are as low as 2–5%. On the basis of the observation that the two PASTA domains of PBP2x interact with each other (Figure 25A) (Gordon et al., 2000), the authors proposed that extracellular regions of PknB dimerize to activate the kinase, since only dimerization would bring two PASTA domains in close contact (Barthe et al., 2010). However, it should be noted that the two PASTA domains of PBP2x are oriented in opposite directions (Figure 25B). Therefore similar interactions would produce PrkC dimers that would not allow the catalytic domains to interact.

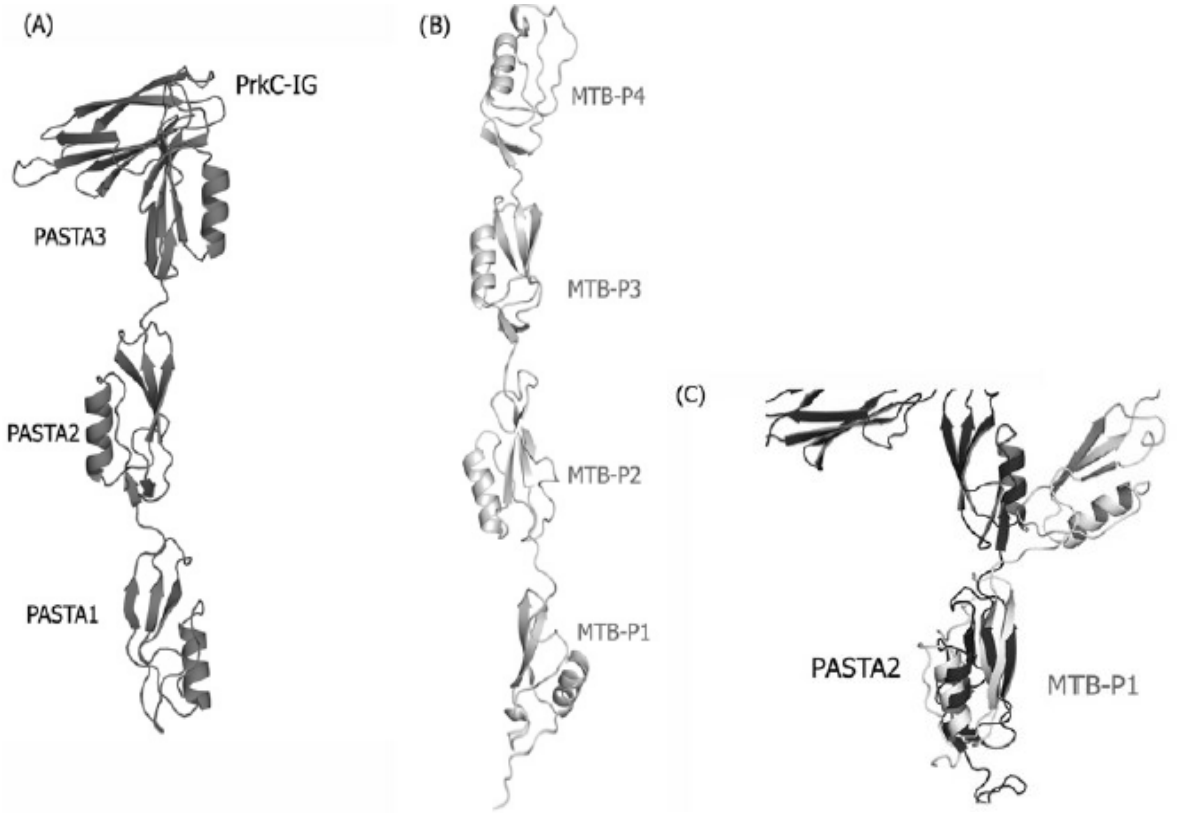


Figure 25. Comparison of EC-PrkC *S.a* with PknB from MTB. A) Ribbon representation of the X-ray structure of EC-PrkC. B) Ribbon representation of PASTA domains arrangement in PknB, as derived by SAXS studies; PDB code 1KUI. C) Superposition of PknB PASTA domain 1 (light grey) with PrkC PASTA domain 2 (dark grey).

II.4.4 Features of the Immunoglobulin (IG)-like domain

The crystal structure shows of EC-PrkC *S.a* contains a C-terminal domain which was not predicted by searches in the PFAM database (Figure 22). This latter domain (residues 577–664), exhibits an all-β structure, which contains six β-strands arranged in a β-sandwich. Sequence alignment analyses do not

identify any structure in the PDB with significant sequence identity. However, a DALI (Holm and Sander, 1995) search of the PDB reveals that PrkC C-terminal domain structurally resembles a set of IG domains. The canonical structure of IG domains is formed by a three-stranded and a four-stranded β-sheet packed in β-sandwich arrangement (Bork et al., 1994). Like canonical IG domains, PrkC-IG domain (PrkC-IG) presents a hydrophobic core formed by both β-sheets, whereas residues pointing to the solvent are mainly hydrophilic; this results in alternating polar/non-polar sequence patterns of the β-strands. The topology of PrkC-IG can be classified as belonging to the s-type IG domains (Bork et al., 1994). However, superposition of PrkC-IG to those identified with DALI (Table 5) shows that this domain lacks the N-terminal strand (β-strand β* of the canonical IG domain; Figure 26).

Table 5 Structural comparison of the PrkC Ig domain with those, identified using DALI, with the highest structural similarity (*Z*-score>5)

Description	PDB code	*Z*-score	RMSD (Å)	Sequence identity (%)
Intercellular adhesion molecule 1 (ICAM-1)	1MQ8	5.1	2.6	13
Sulfur covalently binding protein	2NNF	5.3	2.4	21
Neural cell wall adhesion molecule	2VKX	5.3	3.2	6
Tenascin	3B83	5.2	2.6	9
Fibronectin	2H41	5.1	2.6	9

The sole identified protein that shares a similar feature is a type III domain of human fibronectin (PDB code 2H41; Table 5). In EC-PrkC, the N-terminal lacking strand of PrkC-IG forms the linker with the PASTA3 domain (Figure 23). Different from typical IG folds, where hydrophobic residues in the β* strand pack against the hydrophobic core, the PrkC linker is highly hydrophilic. The volume occupied by hydrophobic residues of the β* strand in canonical IG domains is filled by bulky hydrophilic residues of β-strands β1 and β6 in EC-PrkC *S.a* (results not shown). It is worth noting that IG domains similar to PrkC-IG are usually involved in cell–cell interactions and cell signalling (Lemmon and Schlessinger, 2010). Their adhesive properties are exploited both by bacterial (Bowden et al., 2008; Ganesh et al., 2008) and eukaryotic cells (Carafoli et al., 2008) and are typically due to the exposure of an anomalously large number of backbone β-

strand hydrogenbond donors and acceptors (Richardson et al., 2002). This is particularly true when IG folds are not complete and lack one β-strand (as in PrkC-IG), since this feature leads in extreme cases to IG polymerization (Kline et al., 2010; Vitagliano et al., 2007). These properties of PrkC-IG *S.a* suggest that this domain may also play a role in peptidoglycan binding, in a fashion similar to the *E. coli* adhesin PapG with host glycolipids (Dodson et al., 2001; Imberty et al., 2005). Consistently, we noted that PrkC from *B. subtilis*, which shares similar properties in inducing bacterial sporulation, also contains a similar domain at its C-terminus.

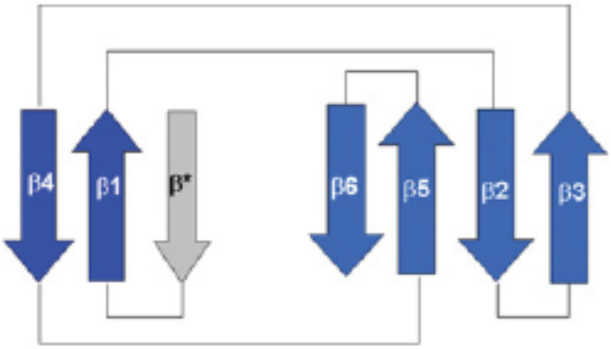


Figure 26. Schematic diagram of canonical IG domain topology. The grey strand in the topology sketch (strand β*) is lacking in PrkC-IG.

II.4.5 Binding studies of PrkC from *B. subtilis* to muropeptides

B. subtilis spores germinate in response to DAP-type muropeptides but not in response to L-Lys-type muropeptides. This finding suggests that PrkC extracellular domains exhibit specificity of muropeptide sensing. However, whether muropeptides are able to bind the protein physically and how the extracellular region is able to distinguish the two types of muropeptides was hitherto unknown.

To explore the structural requirements necessary for recognition and binding of muropeptides to PrkC from *B. subtilis*, we expressed and purified the extracellular portion of the protein and performed binding studies with m-DAP containing muropeptide blend (GlcNAcMurNAcAla$_2$GluDAP, GlcNAc$_2$MurNAc$_2$Ala$_4$Glu$_2$DAP$_2$: Figure 27), purified in the group of Prof. Molinaro from the University of Naples.

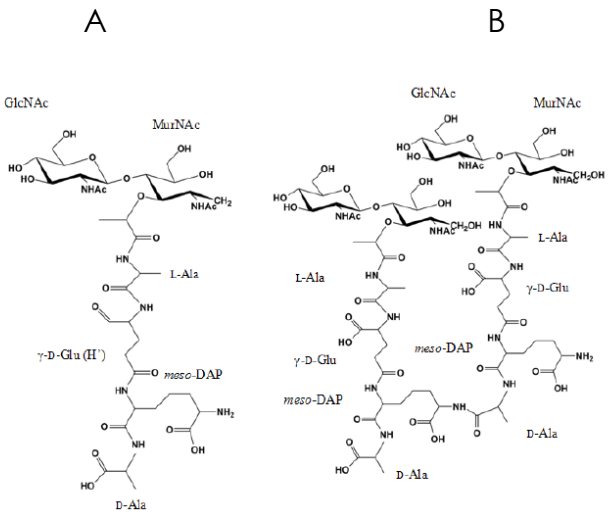


Figure 27. Chemical structures of typical constituents of a muropeptide blend, composed of monomers A) and dimers B).

Using ITC technique we proved that muropeptides are able to bind EC-PrkC from *B. subtilis* and measured their binding affinity to the protein. Using the typical blend of muropeptides released during cell wall remodeling (Figure 27), we observed a low-millimolar protein−ligand affinity, with K_d = 1.2 mM (K_a = 801.0 ± 33.3 M−1, ΔH = −14.7 ± 0.4 kcal ·mol−1) (Figure 28).

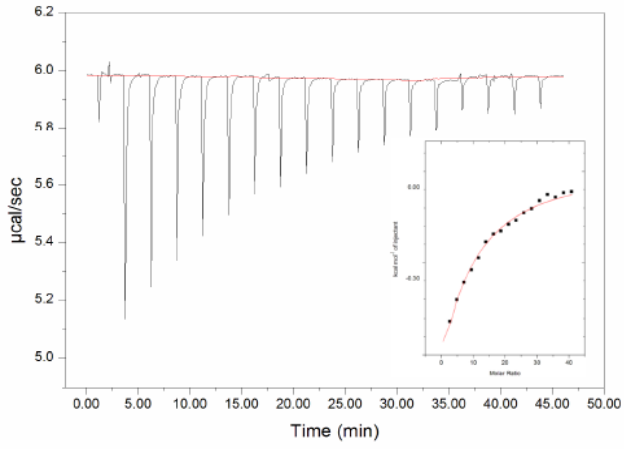


Figure 28. Muropeptide binding with EC-PrkC. Raw data for the titration of 50 µM protein with 20 mM muropeptide blend (muropeptides B and A in 3:1 ratio) at 25°C. Integrated heats of binding obtained from the raw data, after subtracting the heat of dilution, are shown in the inset. The red line represents the best curve fit to the experimental data, using the 'one set of sites' model from MicroCal Origin Software.

Saturation Transfer NMR data obtained in the group of Prof. Molinaro from the University of Naples, revealed the binding epitope of the muropeptide, by measuring relative STD effects from the STD amplification factors. These studies showed that strongest STD NMR signals involve the DAP residue whereas low-

intensity STD signals were recorded for the carbohydrate moieties. The key involvement of the DAP residue in protein recognition well agrees with the previous finding that only muropeptides containing DAP in their peptide stem resuscitate *B. subtilis*, whereas L-Lys-type muropeptides do not (Shah et al., 2008). The NMR binding data also showed that EC-PrkC *B.s* is unable to bind Lys-type peptides.

II.4.6 Bioinformatic studies: Statistical survey and homology modeling

Prompted by the observed binding of muropeptides to EC-PrkC *B.s*, we focused on the identification of their interaction site on the protein surface. We attempted crystallization of EC-PrkC *B.s* with no success. Therefore, we determined the three-dimensional model of EC-PrkC from *B.subtilis* using homology modeling (Figure 29). Modeling was carried out using the software MODELLER and the structure of EC-PrkC from *S.aureus* as a template (28.7% sequence identity).

Analysis of sequence conservation on the model surface did not suggest any obvious conserved patch of residues which could play a role in muropeptide binding. Therefore, the key role of DAP in protein recognition revealed by STD NMR studies prompted us to perform a statistical survey PDB to identify the structural determinants responsible for DAP binding to proteins. Notably, in all structures of protein complexes with DAP, the carboxylate end of DAP forms a salt bridge with an arginine side chain (Figure 29). This finding suggested us that muropeptide binding occurs mainly through the interaction of DAP with an arginine residue of EC-PrkC from *B. subtilis*.

The homology model of EC-PrkC from *B. subtilis* shows the location of the two arginine residues in the protein structure (Figure 29). Arg614 is located in the IG-like domain and is involved in a salt bridge with Glu604 (Figure 29A). On the other hand, Arg500 in located the PASTA3 domain and is fully solvent-exposed (Figure 29B). These considerations led us to hypothesize that Arg500 is involved in muropeptide binding in a fashion similar to those observed in the PDB (Figure 30). To corroborate this hypothesis, we produced mutants of EC-PrkC *B.s* in which Arg500 was changed to alanine (R500A) and to glutamic acid (R500E), two residues which we predicted to be more disruptive of the Arg−DAP interaction.

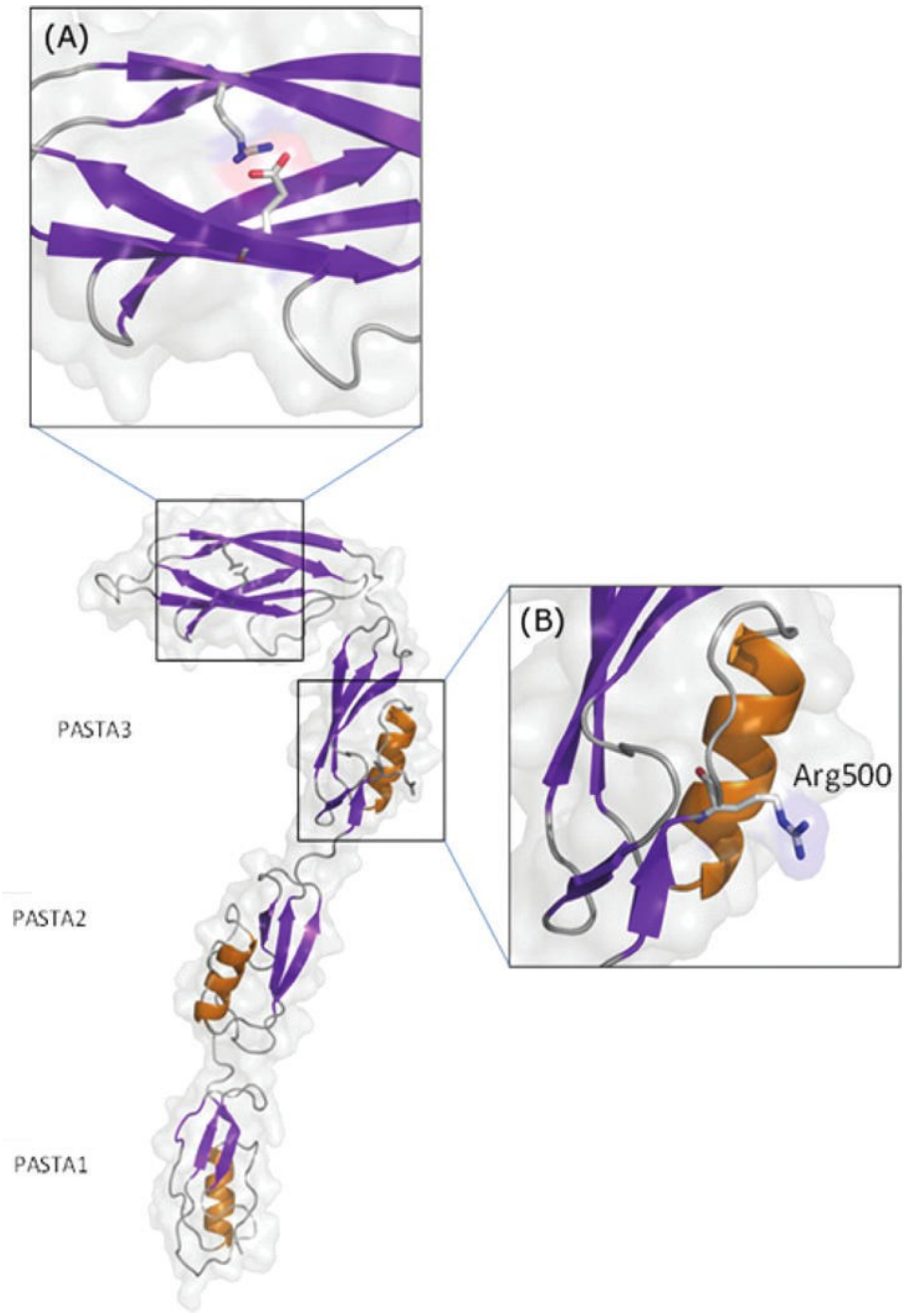


Figure 29. Structure of EC-PrkC from *B. subtilis* derived by homology modeling. The two insets show enlargements of A) the salt bridge between Arg614 and Glu604 in the IG-like domain and B) the location of Arg500 in PASTA3.

STD experiments unequivocally showed that neither R500E nor R500A are able to bind muropeptides, as no STD signals were observed in either case (data not shown). To exclude that the inability of mutants to bind muropeptides was due to alterations of their structure, we carried out CD spectroscopy studies. These studies showed CD spectra of mutants were superimposable to that of the unmutated structure and excluded the possibility of defects in protein folding

induced by the mutations (Figure 31). The inability of the R500E and R500A mutants to bind muropeptides shows that Arg500 is the primary site of interaction with muropeptides. Therefore, we were able to map the site of interaction both on the muropeptides and on EC-PrkC *B.s* and could conclude that PrkC senses muropeptides through interactions of the negatively charged DAP side chain of the muropeptide with the side chain of Arg500.

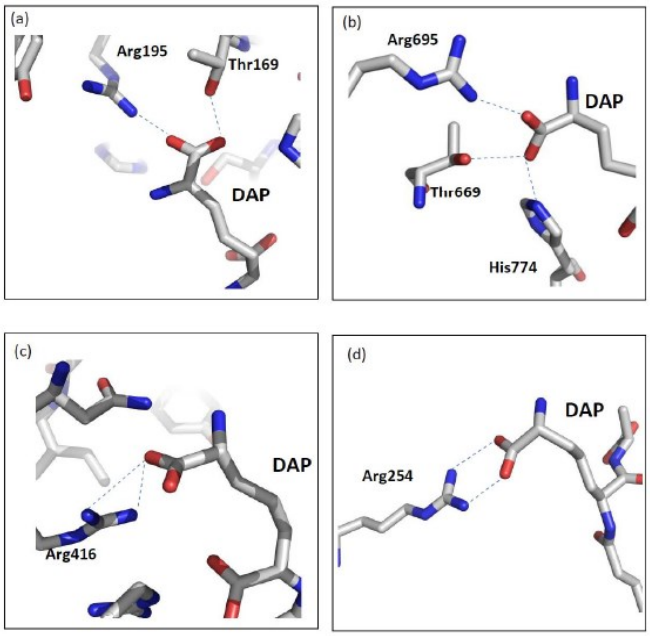


Figure 30. Interaction of DAP with protein structures available in the Protein Data Bank. The four panels represent the binding of DAP to (a,b) Corynebacterium glutamicum diaminopimelate dehydrogenase (PDB codes 2DAP and 1F06, respectively), (c) UDP-N-acetylmuramoyl-L-alanyl-D-glutamate: meso-diaminopimelate ligase from *Escherichia coli* (PDB code 1E8C), (d) Drosophila peptidoglycan recognition protein (PGRP)-LE (PDB code 2CB3).

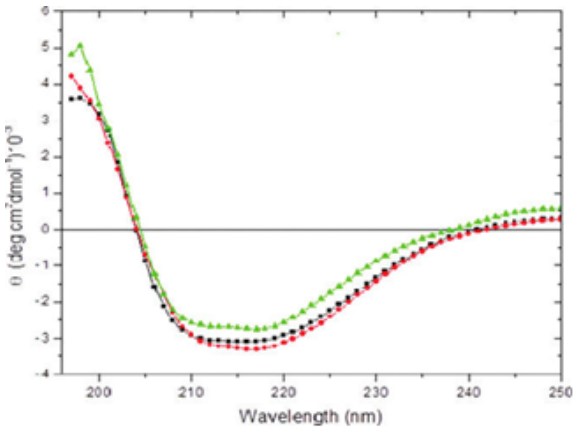


Figure 31. CD spectra of R500A (black) and R500E (red) mutants of EC-PrkC from *B.subtilis* compared to the CD spectrum of EC-PrkC not mutated (green).

II.4.7 Structural features of the PASTA domain of PonA2 from *M. tuberculosis*

We coupled CD and NMR spectroscopy studies with computational approaches to gather insights into the structure and binding properties of the PASTA domain of the PBP PonA2 from *M. tuberculosis*.

The PASTA domain (PonA2-PASTA, residues 700-764) of PonA2 was produced as isotope labeled ($^{15}N/^{13}C$) for NMR experiments. Prior to these studies, PonA2-PASTA fold and stability were evaluated using UV circular dichroism (CD) spectroscopy (Figure 32). CD spectra are characterised by two minima at 208 nm and 222 nm and witness a good degree of structural integrity of the domain (Figure 32). The analysis of thermal denaturation curves evidences that PonA2-PASTA is a stable module with a melting temperature (Tm) of 42°C (Figure 32) and a fully reversible denaturation process (Figure 32).

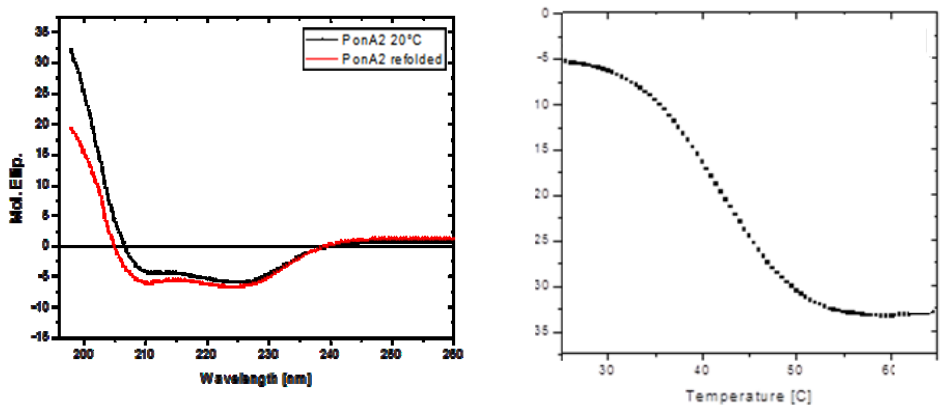


Figure 32. CD studies of ponA2-PASTA. Left) Superposition of CD spectra at 20°C and after refolding from 70°C to 20°C. Right) Thermal denaturation curve. A red arrow locates the melting T.

The $^{1}H–^{15}N$ heteronuclear single quantum coherence (HSQC) spectrum of ^{15}N–labeled PonA2-PASTA domain shows a good dispersion of signals indicative of a well-folded structure (Figure 33). The 3D model (Figure 34), calculated by CYANA program, shows a global topology consisting of a α-helical segment approaching a three stranded β-sheet. The N-terminal helix, which involves about 10 residues, from Val12 to Ala22, is well defined by a dense set of NOE effects (RMSD= 0.05 ± 0.02 Å). Extended conformation is found in the traits Ala27-Ser32 (β1), Gly40-Thr45 (β2) and, Val57-Ser62 (β3). The RMSD values measured on backbone atoms (20 structures) of β1, β2 and β3 strands are all < 0.11 ± 0.06 Å. The inter-strands NOEs (particularly dense between the anti-parallel β2 and β3) used as constraints in the structure calculation gave back

the β1-β3-β2 expected arrangement for PASTA PonA2-PASTA sheet. The helix orientation is defined by NOE contacts between side chains from pre-helix residues (Ala8-Gly9) and β2 ones (Val42-Val43). Phe24 and Tyr39, the only two aromatic residues own by the sequence, are both localized in loop regions. Phe24 side chain, in particular, participates to the hydrophobic core. The numerous unambiguous long range NOEs involving Phe24 aromatic protons and Val4, Leu19, Gly39 and Val57 side chains define the structure of the hydrophobic core. N- and C-terminus residues Ser2 and Asn63 participate to the long range NOEs network with loop nuclei. The backbone flexibility was tested by analyzing relaxation measurements performed on ^{15}N backbone nuclei. The values of R_2/R_1 relaxation rate ratios as well as ^{15}N-{^{1}H} hetero-nuclear NOEs are quite uniform along the sequence except for the more flexible very N-terminal residues. Globally, the domain appears compact. Relevant structural parameters and statistics are given in Table 6 (pag. 70).

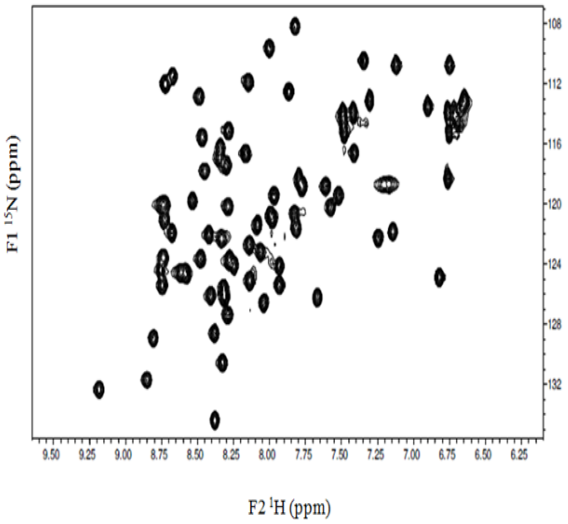


Figure 33. [^{1}H, ^{15}N]-HSQC spectrum of ponA2-PASTA

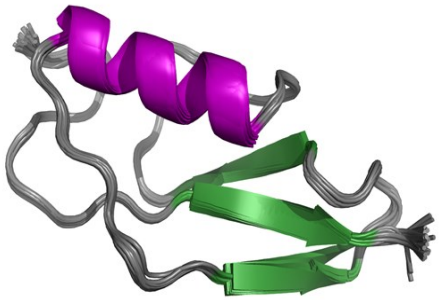


Figure 34. Solution structure of ponA2-PASTA (20 structures)

The 3D structure of PonA2-PASTA exhibits the typical PASTA topology, consisting of a N-terminal a-helix followed by a three strand β-sheet with β1β2β3 arrangement. Interestingly, superposition of the NMR structure of PonA2-PASTA with other available structures shows a higher structural similarity with PASTA domains of STPKs than to that of PBP2x. This finding well correlates with the lower sequence identity with PBP2x than with PASTA domains belonging to STPKs. Despite this, PonA2-PASTA holds a characteristic feature of PBPs. Namely, like in PBPs the loop region between β1 and β2 strands is 4-5 residue shorter than in STPKs.

II.4.8 Binding studies of ponA2-PASTA to muropeptides and β-lactam antibiotics

It is well known that the cross peaks distribution in protein [^{1}H-^{15}N] HSQC maps (Figure 32) represents an efficient way to detect structural changes in response to various phenomena such as molecular binding. In order to test the possible ability of PonA2-PASTA to interact with antibiotic molecules and/or PG fragments, we performed titrations with the two β-lactam antibiotics cefuroxime and cefotaxime (Sigma-Aldrich) and with synthetic Tri-DAP (InvivoGen), a molecule typically used to mimic the PG peptide stem. In all cases, we observed no significant chemical shift perturbations in the [^{1}H-^{15}N]-HSQC spectra upon ligand addition, this indicating that PonA2-PASTA is unable to bind either β-lactam antibiotics, or Tri-DAP (data not shown). The same conclusions were drawn using ITC experiments. An example is reported in Figure 35, which shows that no heat changes are observed upon titration of ponA2-PASTA with mur-TriDAP.

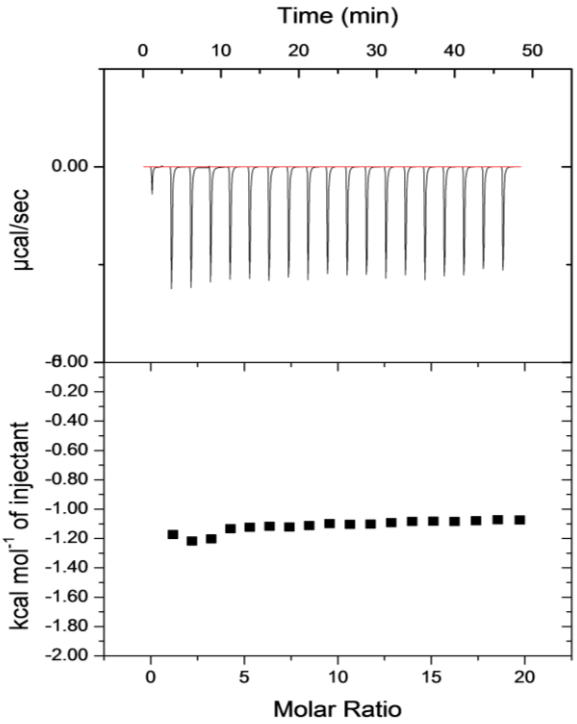


Figure 35. ITC experiment of ponA2-PASTA (0.15mM) with mur-TriDAP (15mM)

II.5 CONCLUSION

Muropeptides are a clear molecular signal that growing conditions are promising, since they are produced during cell wall PG remodeling associated with cell growth and division of neighboring bacteria. However, whether muropeptides are able to bind the protein physically and how the extracellular region is able to distinguish between different muropeptide types was unclear. We focused our attention on STPK PrkC from *Staphylococcus aureus* and from *Bacillus subtilis* and on a PBP from MTB, denominated ponA2. The *fil rouge* between these proteins is that they are all involved in regulation of resuscitation from bacterial dormancy and they all contain PASTA domains.

To this end, we conducted structural studies on PrkC from *Staphylococcus aureus* and from its homologue *Bacillus subtilis*. Furthermore, here we tackled the important question of how the extracellular region of PrkC from *B. subtilis* senses muropeptides. By coupling crystallography, bioinformatic, protein mutagenesis and ITC we proved that EC-PrkC *B.s* is able to bind DAP-type muropeptides and investigated the structural requirements necessary for recognition and binding. We have also shown that this recognition occurs through interactions of DAP with Arg500, as a mutation of this amino acid in EC-PrkC completely impaired muropeptide binding. This finding well agrees with the key role played by arginine in the specific recognition of DAP-type muropeptides by PG recognition proteins (Shah and Dworkin J., 2010; Shah et al., 2008). In this scenario, the key role of Arg500 in binding provides a clear explanation for the ability of EC-PrkC from *B. subtilis* to discriminate between DAP- and Lys-type muropeptides in bacterial revival. Using this mechanism, *B. subtilis* bacteria, which possess a DAP-type PG, can cross-talk and trigger resuscitation by its own cell wall turnover. Our data provide the first molecular clues into the mechanism of sensing of muropeptides by PrkC. These findings opened a novel and more complex scenario in the interpretation of the mechanism of bacterial revival from dormancy. In this scenario, hydrolysis of the cell wall by hydrolases (like RipA in MTB) is only the first step, that has the twofold effect to alter the mechanical properties of the cell wall and produce specific PG fragments.

With the aim of investigating whether a similar muropeptide sensing mechanism is mediated by the PBP PonA2, we determined the NMR structure of its PASTA domain. Beside elucidating the structural properties of this domain, our main interest was to understand if this domain may be a possible target for

muropeptides or β-lactam antibiotics. Indeed, it was previously shown that the the β-lactam antibiotic cefuxomine binds a PASTA domain of the PBP2x from *S. pneumoniae*, which contains two C-terminal PASTA domains (Gordon et al., 2000; Dessen et al., 2001). Using NMR titration and ITC experiments, we show that PonA2-PASTA is unable to bind the PG fragment mur-TriDAP, different than what we observed for PrkC. Similarly, we found that PonA2-PASTA does not bind the β-lactam antibiotics cefuxomine and cefatoxime. Although the importance of this domain in PonA2 function is still to be clarified, our results point to a structural role and not necessarily a muropeptide sensing role. Consistent with our findings, we observed that only one PASTA domain of PrkC is involved in muropeptide binding.

PASTA-containing proteins have been often proposed as interesting targets for β-lactam antibiotics, to be considered in the future for the design of new antimicrobials as well as anchors for muropeptides. Our work shows that this role of PASTA domains, which were considered as 'β-lactam binding domains' (Yeats et al., 2002) cannot be generalized. Although our NMR structure confirms that the PASTA fold is highly conserved (Figure 33), sequence identities among PASTA domains can be as low as 5%. Therefore, specific properties of PASTA domains as β-lactam or muropeptide binders strongly depend on surface residues, which are widely variable.

Table 1. Statistics of data collection for $RipA_{263-472}$ and $_{SeMet}RipA_{263-472}$

	SeMet derivative			Native
	Peak	**Inflection point**	**Remote**	
Beamline	BM14	BM14	BM14	BM14
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Unit cell parameters (Å)				
a	36.83	36.83	36.83	36.77
b	65.84	65.76	65.77	65.41
c	67.84	67.86	67.87	67.79
Resolution (Å)	1.35	1.35	1.52	1.00
Wavelength (Å)	0.979	0.979	0.946	0.974
Average	11.6 (4.0)	10.7 (7.0)	4.6 (4.3)	12.2
Redundancy				
Unique reflections	35146	35658	26106	85882
Completeness	95.3 (73.0)	96.4 (75.9)	99.9 (100.0)	96.5 (72.6)
Rmerge (%)	7.5 (17.1)	5.9 (16.3)	5.0 (10.1)	8.0 (29.2)
Average I/σ(I)	58.7 (7.4)	54.5 (9.2)	55.8 (21.3)	19.7 (3.0)

Table 2. Data Collection and Rifinement statistics of RipA$_{263-472}$

Data Collection and Refinement Statistics	
Data collection	
Space group	$P2_12_12_1$
Unit-cell parameters *a*,*b*,*c* (Å)	36.77, 65.41, 67.79
Resolution shell (Å)	30.0–1.0
Unique reflections (n)	85,882
Average redundancy	12.2 (3.1)
R_{merge} (%)	8.0 (29.2)
Completeness (%)	96.5 (72.6)
Mean I/σ (I)	19.7 (3.0)
Refinement	
Resolution range (Å)	15.0–1.0
Rwork (%)	10.7
Rfree (%)	13.7
Atoms (n)	
Protein	1552
Water	390
Average atomic displacement parameters (Å^2)	
Protein main chain, side chain	9.8, 13.8
Water	28.3
Root mean square deviations	
Bond lengths (Å)	0.019
Bond angle distance (Å)	0.030
Mean coordinate error (Å)	0.021

Table 3. Data collection and refinement statistics. Values in parentheses are for the highest resolution shells (1.00-0.97 Å and 1.43-1.39 Å for C383A and H432A mutants, respectively)

	$RipA_{263-472}C383A$	$RipA_{263-472}H432A$
A. *Data collection*		
Space group	$P2_12_12_1$	$P2_12_12_1$
Cell parameters *a,b,c* (Å);	36.76, 65.51, 67.98	36.70, 65.49, 68.33
Resolution range (Å)	30-0.97	30-1.39
N. of unique reflections	97211	32925
Average redundancy	8.9 (6.4)	5.0 (3.5)
R_{merge} (%)	5.4 (45.0)	4.8 (9.0)
Completeness (%)	99.4 (99.9)	98.8 (90.0)
Mean $I/\sigma(I)$	43.8 (2.9)	61.0 (15.1)
B. *Refinement*		
Rwork / Rfree (%)	14.4/18.2	16.9/18.5
Bond lengths (Å)	0.01	0.01

Table 4. Data collection and refinement statistic of EC-PrkC from *S.aureus*

	Lutetium chloride derivative			
	Peak	Inflection point	Remote	
Beamline	X12	X12	X12	In-house
Space group	$P2_12_12$	$P2_12_12$	$P2_12_12$	$P2_12_12$
Unit cell parameters (Å)				
a	55.50	55.48	55.98	55.60
b	68.75	68.75	68.28	69.28
c	88.49	88.47	87.92	88.43
Resolution range (Å)	20.00-2.95	20.00-2.95	20.00-3.20	20.00-2.15
Wavelength (Å)	1.3387	1.3393	1.0000	1.5418
Average redundancy	4.7 (3.2)	6.1 (3.0)	4.8 (2.3)	7.1 (4.8)
Unique reflections	6892	7250	5517	19217
Completeness	92.6 (75.9)	97.4 (82.0)	88.3 (70.1)	99.4 (94.7)
Rmerge (%)	9.1 (45.0)	9.2 (45.8)	9.0 (28.2)	6.2 (44.3)
Average I/σ(I)	9.3 (3.0)	10.8 (2.0)	19.9 (2.7)	22.6 (2.6)

Values in parentheses are for highest resolution shells: 3.06-2.95 Å for peak and inflection point data collections, 3.31-3.20 Å for the remote data collection and 2.23-2.15 Å for in-house data collection.

Values in parentheses are for the highest-resolution shell (2.28–2.20 Å)

Parameter	Value
(a) Data collection	
Space group	$P2_12_12$
Unit cell parameters *a*, *b*, *c* (Å)	55.64, 69.40, 88.62
Resolution shell (Å)	30.0–2.2
Number of unique reflections	18045
Average redundancy	8.0 (8.0)
R_{merge}	0.08 (0.44)
Completeness (%)	99.9 (99.9)
Mean $I/\sigma(I)$	32.4 (3.8)
(b) Refinement	
Resolution range (Å)	15.0–2.2
R_{work}	0.210
R_{free}	0.257
Number of atoms	
Protein	2195
Water	125
Average atomic displacement parameters (Å^2)	
Protein main chain, side chain	38.2, 41.9
Water	43.6
RMSDs	
Bond lengths (Å)	0.01
Bond angles (°)	1.56

Table 6. Protein Structure Violation Suite Statistics for the ensemble of 10 structures calculated for ponA2-PASTA (residues 703-764)

NOE based distance constraints	
Total[a]	465
intra-residue [i = j]	189
sequential [|i - j| = 1]	133
medium range [1 < |i - j| < 5]	38
long range [|i - j| ≥ 5]	105
Backbone dihedral angle constraints	87
Total number of constraints	552
Residual constraint violations	
Distance violations/structure	
0.1-0.2 Å	26.4
0.2-0.5 Å	7.8
> 0.5 Å	13.3
RMS of distance violation/constraint	0.13 Å
Maximum distance violation	1.17 Å
Dihedral angle violations/structure	
1-10°	10.8
>10°	0
RMS of dihedral angle violation/constraint	0.74°
Maximum dihedral angle violation	4.70°
RMSD Values	
All backbone atoms	0.4 Å
All heavy atoms	0.8 Å
Ramachandran plot(%)	
Most favored region	89
Additionally allowed region	11
Generously allowed region	0
Disallowed region	0

[a]Distance constraints which excluded fixed intra-residue distances.

Abbreviations

CD Circular Dichroism

CHAP Cysteine Histidine Amino Peptidase

DA Diamino

EC Extra Cellular

EF-G Elongation Factor G

EF-Tu Elongation Factor Thermo-unstable

ESRF European Synchrotron Radiation Facility

FITC fluorescein isothiocyanate

GlcNAc *N*-acetylglucosamine

GST Glutatione S-Transferase

HSQC Heteronuclear Single Quantum Coherence

IG Immunoglobulin

IMAC Immobilized Metal Affinity Chromatography

IPTG Isopropil β-D-1-tiogalattopiranoside

ITC Isotermal Titration Calorimetry

KDa KiloDalton

LB Luria Broth

LC-MS Liquid Chromatography-Mass Spectrometry

LPS Lipopolysaccharides

LTAs Lipoteichoic acids

m/z Mass-to-charge ratio

MAD Multi-wavelength Anomalous Diffraction

m-DAP *meso*-diaminopimelic acid

MDR-TB Multidrug-resistant TB

MS Mass Spectrometry

MTB *Mycobacterium tuberculosis*

MurNAc *N*-acetylmuramic acid

NMR Nuclear Magnetic Resonance

NOE Nuclear Overhauser Effect

NOESY Nuclear Overhauser Effect Spectroscopy

NTA Nitriloacetic Acid

OD Optical Density

OM Outer Membrane

PASTA PBP and Serine/Threonine kinase Associated

PBP Penicillin Binding Protein

PCR Polymerase Chain Reaction

PDB Protein Data Bank

PEG Polyethylene glycol

PFAM Protein Families

PG Peptidoglycan

RipA Resuscitation promoting factor Interacting Protein A

RMSD Root Mean Square Deviation

RpfB Resuscitation Promoting Factor

rpm revolutions per minute

SAD Single-wavelength Anomalous Diffraction

SAXS Small Angle X-ray Scattering

SDS Sodium Dodeciyl Sulfate

STD Saturation Tranfer Difference

STPKs Ser/Thr protein kinases

TAs Teichoic Acids

TB Tuberculosis

TEV protease Tobacco Etch virus protease

TFA Trifluoroacetic Acid

Tm melting temperature

TOCSY Total Correlation Spectroscopy

TRX Tioredossina

WHO World Health Organization

WTAs Wall Teichoic Acids

Standard amino acid abbreviations

Alanine Ala (A)

Arginine Arg (R)

Asparagine Asn (N)

Aspartic Acid Asp (D)

Cysteine Cys (C)

Glutamine Gln (Q)

Glutamic Acid Glu (E)

Glycine Gly (G)

Histidine His (H)

Isoleucine Ile (I)

Leucine Leu (L)

Lysine Lys (K)

Methionine Met (M)

Phenylalanine Phe (F)

Proline Pro (P)

Serine Ser (S)

Threonine Thr (T)

Tryptophan Trp (W)

Tyrosine Tyr (Y)

Valine Val (V)

Bibliography

- Absalon, C., Obuchowski, M., Madec, E., Delattre, D., Holland, I. B. and Seror, S. J., (2009), "CpgA, EF-Tu and the stressosome protein YezB are substrates of the Ser/Thr kinase/phosphatase couple, PrkC/PrpC, in *Bacillus subtilis", Microbiology* 155, 932–943;
- Aramini, J.M., Rossi, P., Huang, Y.J., Zhao, L., Jiang, M., Maglaqui, M., Xiao, R., Locke, J., Nair, R., Rost, B., et al., (2008), "Solution NMR structure of the NlpC/P60 domain of lipoprotein Spr from *Escherichia coli*: structural evidence for a novel cysteine peptidase catalytic triad"*, Biochemistry;*
- Archibald, A.R., I.C. Hancock and C.R. Harwood, (1993), "Cell wall structure, synthesis and turnover. In: Sonnenshein A.L., Hoch J.A., Losick R. (eds) *Bacillus subtilis* and other Gram-positive bacteria: biochemistry, physiology, and molecular genetics", *ASM Press, Washington*, DC.391-410;
- Barthe, P., Mukamolova, G. V., Roumestand, C. and Cohen-Gonsaud, M., (2010), "The structure of PknB extracellular PASTA domain from *Mycobacterium tuberculosis* suggests a ligand-dependent kinase activation", *Structure*, 18, 606–615;
- Bateman, A., and Rawlings, N.D., (2003), "The CHAP domain: a large family of amidases including GSP amidase and peptidoglycan hydrolases", *Trends Biochem. Sci.;*
- Bendtsen, J.D., Nielsen, H., von Heijne, G., Brunak, S., (2004), "Improved prediction of signal peptides: SignalP 3.0", *J. Mol. Biol.*, *340*, 783-795;
- Beveridge TJ, Graham LL., (1991), "Surface layers of bacteria", *Microbiol Rev.;*
- Boothby, D., L. Daneo-Moore, M.L. Higgins, J. Coyette and G.D. Shockman, (1973), "Turnover of bacterial cell wall peptidoglycans", *J Biol Chem.* 248:2161-2169;
- Bork, P., Holm, L. and Sander, C., (1994), "The immunoglobulin fold: structural classification, sequence patterns and common core", *J. Mol. Biol.* 242, 309–320;
- Bowden, M. G., Heuck, A. P., Ponnuraj, K., Kolosova, E., Choe, D., Gurusiddappa, S., Narayana, S. V., Johnson, A. E. and Hook, M., (2008), "Evidence for the "dock, lock, and latch" ligand binding mechanism of the staphylococcal microbial surface component recognizing adhesive matrix molecules (MSCRAMM) SdrG", J. *Biol. Chem*. 283, 638–647;

- Braun, V., (1975), "Covalent lipoprotein from the outer membrane of *Escherichia coli*", Biochim Biophys Acta. 415:335-377;
- Brennan, P.J., (2003), "Structure, function, and biogenesis of the cell wall of *Mycobacterium tuberculosis", Tuberculosis*, *83*, 91-97;
- Carafoli, F., Saffell, J. L. and Hohenester, E., (2008), "Structure of the tandem fibronectin type 3 domains of neural cell adhesion molecule", *J. Mol. Biol*. 377, 524-534;
- Chaloupka, J., P. Křečková and L. Řihová., (1962a), "The mucopeptide turnover in the cell walls of growing cultures of *Bacillus megaterium*", *KM. Experientia,* 18:362-363;
- Chaloupka, J., P. Křečková and L. Říhová., (1962b), "Changes in the character of the cell wall in growth of *Bacillus megaterium* cultures", *Folia Microbiol.* 7:269-274;
- Cuff, J.A., Clamp, M.E., Siddiqui, A.S., Finlay, M., Barton, G.J., (1998), "Jpred: a consensus secondary structure prediction server", *Bioinformatics*, 14, 892-893;
- Dajkovic, A.; Lutkenhaus, J., (2006), "Z ring as executor of bacterial cell division." *J. Mol. Microbiol. Biotechnol., 11, 140-151;*
- Dessen, A., Mouz, N., Gordon, E., Hopkins, J. and Dideberg, O., (2001), "Crystal structure of PBP2x from a highly penicillin-resistant *Streptococcus pneumoniae* clinical isolate: a mosaic framework containing 83 mutations", *J. Biol. Chem*, 276, 45106 45112;
- Dodson, K. W., Pinkner, J. S., Rose, T., Magnusson, G., Hultgren, S. J. and Waksman, G., (2001), "Structural basis of the interaction of the pyelonephritic *E. coli* adhesin to its human kidney receptor", *Cell*, 105, 733–743;
- Doyle, R.J., J. Chaloupka and V. Vinter., (1988), "Turnover of cell walls in microorganisms", *Microbiol Rev*, 52:554–567;
- Dubendorff, J. W. and F. W. Studier,(1991), "Controlling basal expression in an inducible T7 expression system by blocking the target T7 promoter with lac repressor", *J. Mol. Biol*. 219(1): 45-59;
- Dziadek, J.; Rutherford, S.A.; Madiraju, M.V.; Atkinson, M.A.; Rajagopalan, M., (2003), "Conditional expression of *mycobacterium smegmatis* ftsz, an essential cell division gene", *Microbiology*, *149*, 1593-1603;
- Echenique, J., Kadioglu, A., Romao, S., Andrew, P.W., Trombe, M.C., (2004), "Protein serine/threonine kinase stkp positively controls virulence and competence in *Streptococcus pneumonia" Infect. Immun.*, 2004, 72, 2434-2437;

- Finn, R.D., Tate, J., Mistry, J., Coggill, P.C., Sammut, S.J., Hotz, H.R., Ceric, G., Forslund, K., Eddy, S.R., Sonnhammer, E.L., Bateman, A., (2008) "The pfam protein families database", *Nucleic Acids Res*, 36, D281-288;
- Fiuza, M., Canova, M. J., Zanella-Cleon, I., Becchi, M., Cozzone, A. J., Mateos, L. M., Kremer, L., Gil, J. A. and Molle, V., (2008), "From the characterization of the four serine/threonine protein kinases (PknA/B/G/L) of *Corynebacterium glutamicum* toward the role of PknA and PknB in cell division", *J. Biol. Chem.*, 283, 18099–18112;
- Fukushima, T., Kitajima, T.,Yamaguchi, H.,Ouyang, Q.,Furuhata, K.,Yamamoto, H., Shida, T., and Sekiguchi, J., (2008), "Identification and characterization of novel cell wall hydrolase CwlT: a two-domain autolysin exhibiting n-acetylmuramidase and DL-endopeptidase activities", *J. Biol. Chem.* 283, 11117–11125;
- Gaidenko, T. A., Kim, T. J. and Price, C. W., (2002), "The PrpC serine-threonine phosphatase and PrkC kinase have opposing physiological roles in stationary-phase *Bacillus subtilis* cells", *J. Bacteriol*, 184, 6109–6114;
- Ganesh, V. K., Rivera, J. J., Smeds, E., Ko, Y. P., Bowden, M. G., Wann, E. R., Gurusiddappa, S., Fitzgerald, J. R. and Hook, M., (2008), "A structural model of the *Staphylococcus aureus* ClfA–fibrinogen interaction opens new avenues for the design of anti-staphylococcal therapeutics", *PLoS Pathog*, 4, e1000226;
- Gordon, E., Mouz, N., Duee, E. and Dideberg, O., (2000), "The crystal structure of the Penicillin-Binding Protein2x from *Streptococcus pneumoniae* and its acyl-enzyme form: implication in drug resistance", *J. Mol. Biol*. 299, 477–485;
- Goulding, C.W., Perry, L.J., 2003, "Protein production in *Escherichia coli* for structural studies by X-ray crystallography", *J. Struct. Biol.*, 142, 133-143;
- Graham LL, Beveridge TJ., (1994), "Structural differentiation of *the Bacillus subtilis* 168 cell wall", *J Bacteriol*, Mar;176(5):1413-21;
- Gruber, M., Soding, J., Lupas, A.N., (2006), "Comparative analysis of coiled-coil prediction methods", *J. Struct. Biol.*, *155*, 140-145;
- Hanks S.K.,Quinn A.M., Hunter T., (1988), "The protein kinase family: conserved features and deduced phylogeny of the catalytic domains", Jul 1;241(4861):42-52;
- Hayhurst, E.J., L. Kailas, J.K. Hobbs and S.J. Foster., (2008), "Cell wall peptidoglycan architecture in *Bacillus subtilis*", *Proc Nat Acad Sci*, 105:14603-14608;

- Hett, E. C. and Rubin, E. J., (2008), "Bacterial growth and cell division: a mycobacterial perspective", *Microbiol. Mol. Biol, Rev*. 72, 126–156;
- Hett, E.C., Chao, M.C., Deng, L.L., and Rubin, E.J., (2008), "A mycobacterial enzyme essential for cell division synergizes with Resuscitation-Promoting Factor", *PLoS Pathog*:
- Hett, E.C., Chao, M.C., Steyn, A.J., Fortune, S.M., Deng, L.L., and Rubin, E.J., (2007), "A partner for the Resuscitation-Promoting Factors of *Mycobacterium tuberculosis", Mol. Microbiol*;
- Holm, L. and Sander, C., (1995), "Dali: a network tool for protein structure comparison", *Trends Biochem*. Sci. 20, 478–480;
- Höltje, J.-V., (1998), "Growth of the stress-bearing and shape-maintaining murein sacculus of *Escherichia coli*", *Microbiol Mol Biol Rev*., 62:181-203;
- Imberty, A., Mitchell, E. P. and Wimmerova, M., (2005), "Structural basis of high-affinity glycan recognition by bacterial and fungal lectins", *Curr. Opin. Struct. Biol*, 15, 525–534;
- Johnson J.W., Fisher J.F., and Mobashery S., (2013),"Bacterial cell-wall recycling", *Antimicrobial Therapeutics Reviews;*
- Jones, T.A., (2004) "Interactive electron-density map interpretation: From inter to O", *Acta Crystallogr D Biol Crystallogr*, 60, 2115-2125;
- Kana, B.D., and Mizrahi, V, (2010), "Resuscitation-Promoting Factors as lytic enzymes for bacterial growth and signaling", *FEMS Immunol. Med. Microbiol*.;
- Kana, B.D., Gordhan, B.G., Downing, K.J., Sung, N., Vostroktunova, G., Machowski, E.E., Tsenova, L., Young, M., Kaprelyants, A., Kaplan, G., and Mizrahi, V., (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*.;
- Kaprelyants, A.S.; Mukamolova, G.V.; Ruggiero, A.; Makarov, V.A.; Demina, G.R.; Shleeva, M.O.; Potapov, V.D.; Shramko, P.A., (2012), "Resuscitation-Promoting Factors (Rpf): In search of inhibitors", *Pro. Pept. Lett*.;
- Kaufmann S.H., (2008), "Tubercolosis: Deadly combination", *Nature,* 453: 295-296;
- Keep N.H., Ward J.M., Cohen-Gonsaud M., Henderson B., (2006), "Wake up! Peptidoglycan lysis and bacterial non-growth states", *Trends Microbiol,* 14: 271-276;
- Kline, K. A., Dodson, K. W., Caparon, M. G. and Hultgren, S. J., (2010), "A tale of two pili: assembly and function of pili in bacteria" *Trends Microbiol*. 18, 224–232;

- Koradi R., Billeter M., Wüthrich K., (1996), "MOLMOL: a program for display and analysis of macromolecular structures", *J Mol Graph*, Feb; 14(1):51-5, 29-32;
- Krogh, A., Larsson, B., von Heijne, G., Sonnhammer, E.L., (2001), "Predicting transmembrane protein topology with a hidden markov model: Application to complete genomes", *J. Mol. Biol.*, *305*, 567- 580;
- Langer, G., Cohen, S.X., Lamzin, V.S., Perrakis, A., (2008), "Automated macromolecular model building for x-ray crystallography using ARP/wARP version 7", *Nat Protoc*, 3, 1171-1179;
- Laskowski, R.A., Rullmannn, J.A., MacArthur, M.W., Kaptein, R., and Thornton, J.M., (1996)," AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR" *J. Biomol.*, NMR 8, 477 486;
- Layec, S., Decaris, B., and Leblond-Bourget, N., (2008), "Diversity of Firmicutes peptidoglycan hydrolases and specificities of those involved in daughter cell separation", *Res. Microbiol*. 159, 507–515;
- Layec, S., Gerard, J., Legue, V., Chapot-Chartier, M.P., Courtin, P., Borges, F., Decaris, B., and Leblond-Bourget, N., (2009), "The CHAP domain of Cse functions as an endopeptidase that acts at mature septa to promote *Streptococcus thermophilus* cell separation", *Mol. Microbiol*. 71, 1205–121;
- Lemmon, M. A. and Schlessinger, J., (2010), "Cell signaling by receptor tyrosine kinases", *Cell,* 141, 1117–1134;
- Lindahl E., Hess B., Van der Spoel D, (2001), "Gromacs 3.0: A package for molecular simulation and trajectory analysis", *J. Mol. Modell*., 7, 306;
- Lutkenhaus, J.F.; Wolf-Watz, H.; Donachie, W.D., (1980), "Organization of genes in the ftsa-enva region of the *Escherichia coli* genetic map and identification of a new fts locus (ftsz)", *J. Bacteriol*., 142, 615-620;
- Madec, E., Laszkiewicz, A., Iwanicki, A., Obuchowski, M. and Seror, S. ,(2002), "Characterization of a membrane-linked Ser/Thr protein kinase in *Bacillus subtilis*, implicated in developmental processes", Mol. *Microbiol*, 46, 571–586;
- Margolin, W., (2005), "Ftsz and the division of prokaryotic cells and organelles", *Nat. Rev. Mol. Biol.*, 6, 862-871;
- Margot P, Pagni M, Karamata D., (1999), "*Bacillus subtilis* 168 gene lytF encodes a gamma-D-glutamate-meso-diaminopimelate muropeptidase expressed by the alternative vegetative sigma factor, sigmaD" *Microbiology*;
- Margot P, Wahlen M, Gholamhoseinian A, Piggot P, Karamata D., (1998), "The lytE gene of *Bacillus subtilis* 168 encodes a cell wall hydrolase", *J Bacteriol;*

- Matthews, B.W., (1968), "Solvent content of protein crystals", *J Mol Biol*, 33, 491-497;
- Mauck, J., L. Chan and L. Glaser, (1971), "Turnover of the cell wall of Gram-positive bacteria", *J Biol Chem*. 246:1820-1827;
- McGahee, W. and Lowy, F. D., (2000), "Staphylococcal infections in the intensive care unit", *Semin. Respir. Infect*. 15, 308–313;
- McGuffin, L.J., Bryson, K., Jones, D.T., (2000), "The Psipred protein structure prediction server", *Bioinformatics*, *16*, 404-405;
- Menard, R., Carmona, E., Takebe, S., Dufour, E., Plouffe, C., Mason, P., and Mort, J.S., (1998), "Autocatalytic processing of recombinant human procathepsin L. Contribution of both intermolecular and unimolecular events in the processing of procathepsin L in vitro", *J. Biol. Chem.;*
- Merad T., Archibald A.R., Hancock I.C., Harwood C.R., Hobot J.A.,(1989), "Cell wall assembly in *Bacillus subtilis*: visualization of old and new wall material by electron microscopic examination of samples stained selectively for teichoic acid and teichuronic acid", *Journal of' General Microbiology;*
- Moulder, J.W., (1993), "Why is *Chlamydia* sensitive to penicillin in the absence of peptidoglycan?", *Infect Agents*, Dis. 2:87-99;
- Mu˜noz-Dorado, J., Inouye, S. and Inouye, M., (1991), "A gene encoding a protein serine/threonine kinase is required for normal development of *M. xanthus,* a Gram-negative bacterium", *Cell* 67, 995–1006;
- Mukamolova, G.V.; Turapov, O.A.; Young, D.I.; Kaprelyants, A.S.; Kell, D.B.; Young, M., (2002), "A family of autocrine growth factors in *Mycobacterium tuberculosis*" *Mol. Microbiol.*, *46*, 623-635;
- Murshudov, G.N., Vagin, A.A., and Dodson, E.J., (1997), "Refinement of macromolecular structures by the maximum-likelihood method", *Acta Crystallogr. D Biol. Crystallogr.*, 53, 240–255;
- Nancy A. Knechel, (2009), "Tuberculosis: Pathophysiology, Clinical Features, and Diagnosis", *American Association of Critical-Care Nurses*;
- Nanninga, N., (1998), "Morphogenesis of *Escherichia coli*", *Microbiol. Mol. Biol. Rev., : MMBR*, *62*, 110-129;
- Ohlsen, K. and Donat, S., (2010), "The impact of serine/threonine phosphorylation in *Staphylococcus aureus", Int. J. Med. Microbiol*. 300, 137–141;
- Otwinowski, Z., and Minor, W., (1997), "Processing of X-ray diffraction data collected in oscillation mode", *Methods Enzymol*;

- Pares, S., Mouz, N., Petillot, Y., Hakenbeck, R. and Dideberg, O., (1996), "X-ray structure of *Streptococcus pneumoniae* PBP2x, a primary penicillin target enzyme" *Nat. Struct. Biol*. 3, 284–289;
- Park, J.T., (1996), "The murein sacculus. *In*: Neidhardt F.C., Curtiss R. III, Ingraham J.L., Lin E.C.C., Low K.B., Magasanik B., Reznikoff W.S., Riley M., Schaechter M., and Umbarger H.E. (eds) *Escherichia coli* and Salmonella: cellular and molecular biology", *ASM Press, Washington*, DC:48-57;
- Patru MM, Pavelka MS Jr., (2010), "A role for the class A Penicillin-Binding Protein PonA2 in the survival of *Mycobacterium smegmatis* under conditions of nonreplication", *J Bacteriol.,* Jun, 192(12): 3043-54;
- Pereira, S.F.; Goss, L.; Dworkin, J., (2011), "Eukaryote-like serine/threonine kinases and phosphatases in bacteria", *Microbiol. Mol. Biol. Rev.,* MMBR, 75, 192-212;
- Pooley HM., (1976), "Turnover and spreading of old wall during surface growth of *Bacillus subtilis*", *J Bacteriol*;
- Potterton, E., Briggs, P., Turkenburg, M., and Dodson, E., (2003), "A graphical user interface to the CCP4 program suite", *Acta Crystallogr. D Biol. Crystallogr.*, 59, 1131–1137;
- Richardson, J. S. and Richardson, D. C., (2002), "Natural β-sheet proteins use negative design to avoid edge-to-edge aggregation", *Proc. Natl. Acad. Sci. U.S.A*. 99, 2754–2759;
- Rogers, H., (1967), "The structure and biosynthesis of the components of the cell walls of Gram-positive bacteria", *Folia Microbiol*. 12:191-200;
- Rossi, P., Aramini, J.M., Xiao, R., Chen, C.X.,Nwosu, C., Owens, L.A.,Maglaqui, M., Nair, R., Fischer, M., Acton, T.B., et al., (2009), "Structural elucidation of the of the Cys-His-Glu-Asn proteolytic relay in the secreted CHAP domain enzyme from the human pathogen *Staphylococcus saprophyticus*", *Proteins*;
- Rozman, J., Stojan, J., Kuhelj, R., Turk, V., and Turk, B., (1999), "Autocatalytic processing of recombinant human procathepsin B is a bimolecular process", *FEBS Lett;*
- Ruggiero A.*, Squeglia F.*, Pirone L., Correale S., Berisio R., (2011), "Expression, purification, crystallization and preliminary X-ray crystallographic analysis of a major fragment of the Resuscitation-Promoting Factor RpfB from *Mycobacterium tuberculosis", Acta Crystallogr Sect F Struct Biol Cryst Commun*, 67, 164-8, (*Shared first author);
- Ruggiero A., Marchant J., Squeglia F., Makarov V., De Simone A., Berisio R., (2012), "Molecular determinants of inactivation of the resuscitation

Promoting Factor B from *Mycobacterium tuberculosis*", *J Biomol Struct,* July 25;

- Ruggiero, A.; Marasco, D.; Squeglia, F.; Soldini, S.; Pedone, E.; Pedone, C.; Berisio, R., (2010), "Structure and functional regulation of ripa, a mycobacterial enzyme essential for daughter cell separation", *Structure*, 18, 1184-1190;
- Schroder, K., and Tschopp, J., (2010), "The inflammasomes", *Cell,* 140, 821–832;
- Seltmann, G. and O. Holst., (2002), "Periplasmic space and rigid layer. In: Seltmann G., Holst O. (eds), The bacterial cell wall", *Springer, Berlin*:105 132;
- Shah I.M., Dworkin J., (2010), "Induction and regulation of a secreted peptidoglycan hydrolase by a membrane Ser/Thr kinase that detects muropeptides", *Mol Microbiol.*, Mar;75(5):1232-43;
- Shah, I.M.; Laaberki, M.H.; Popham, D.L.; Dworkin, J., (2008), "A eukaryotic-like ser/thr kinase signals bacteria to exit dormancy in response to peptidoglycan fragments", *Cell*, 135, 486-496;
- Sheldrick, G.M., (2008), "A short history of SHELX" *Acta Crystallogr.,* A 64, 112–122;
- Shockman, G.D. and J.F. Barrett., (1983), "Structure, function, and assembly of cell walls of Gram-positive bacteria", *Annu Rev Microbiol*., 37:501-527;
- Sørensen, H. P., and K. K. Mortensen., (2005) "Advanced genetic strategies for recombinant protein expression in *Escherichia coli*" *J. Biotechnol*, 115:113–128;
- Sripa, J., Laha, T., To, J., Brindley, P.J., Sripa, B., Kaewkes, S., Dalton, J.P., and Robinson, M.W., (2010), "Secreted cysteine proteases of the carcinogenic liver fluke, Opisthorchis viverrini: regulation of cathepsin F activation by autocatalysis and trans-processing by cathepsin B", *Cell Microbiol*., 12, 781–795;
- Stock J.B., Ninfa A.J., Stock A.M., (1989),"Protein phosphorylation and regulation of adaptive responses in bacteria", *Microbiol Rev.*, Dec;53(4):450-90;
- Studier, F., A. Rosenberg, J. Dunn and J. Dubendorff, (1990), "Use of T7 RNA polymerase to direct expression of cloned genes." *Methods Enzymol*, 185: 60-89;
- Terwilliger, T., (2004), "SOLVE and RESOLVE: automated structure solution, density modification and model building", *J. Synchrotron Radiat*., 11, 49–52;
- Terwilliger, T.C., (2003), "Solve and resolve: Automated structure solution and density modification", *Methods Enzymol*, 374, 22-37;

- Terwilliger, T.C., Berendzen, (1999), "J. Automated mad and mir structure solution", *Acta Crystallogr D Biol Crystallogr,* 55, 849-861;
- Vitagliano, L., Ruggiero, A., Pedone, C. and Berisio, R., (2007,) "A molecular dynamics study of pilus subunits: insights into pilus biogenesis", *J. Mol. Biol.* 367, 935–941;
- Ward, J., (1973), "The chain length of the glycans in bacterial cell walls.", *Biochem J*. 133:395-398;
- Ward, J.J., McGuffin, L.J., Bryson, K., Buxton, B.F., Jones, D.T., (2004), "The Disopred server for the prediction of protein disorder", *Bioinformatics*, *20*, 2138-2139;
- Wehenkel, A., Bellinzoni, M., Grana, M., Duran, R., Villarino, A., Fernandez, P., Andre-Leroux, G., England, P., Takiff, H., Cervenansky, C., Cole, S.T., Alzari, P.M., (2008), "Mycobacterial ser/thr protein kinases and phosphatases: Physiological roles and therapeutic potential", *Biochim. et biophys. acta*., 1784, 193- 202;
- Weidel, W. and H. Pelzer., (1964), "Bag-shaped macromolecules - a new outlook on bacterial cell walls", *Adv Enzymol Relat Areas Mol Biol*. 26:193-232;
- Wittlin, S., Rosel, J., Hofmann, F., and Stover, D.R., (1999), "Mechanisms and kinetics of procathepsin D activation. Eur*", J. Biochem;*
- Xu Q, Abdubek P, Astakhova T, Axelrod HL, Bakolitsa C, Cai X, Carlton D, Chen C, Chiu HJ, Chiu M, Clayton T, Das D, Deller MC, Duan L, Ellrott K, Farr CL, Feuerhelm J, Grant JC, Grzechnik A, Han GW, Jaroszewski L, Jin KK, Klock HE, Knuth MW, Kozbial P, Krishna SS, Kumar A, Lam WW, Marciano D, Miller MD, Morse AT, Nigoghossian E, Nopakun A, Okach L, Puckett C, Reyes R, Tien HJ, Trame CB, van den Bedem H, Weekes D, Wooten T, Yeh A, Hodgson KO, Wooley J, Elsliger MA, Deacon AM, Godzik A, Lesley SA, Wilson IA, (2010), "Structure of the γ-D-glutamyl-L-diamino acid endopeptidase YkfC from *Bacillus cereus* in complex with L-Ala-γ-D-Glu: insights into substrate recognition by NlpC/P60 cysteine peptidases", *Acta Crystallogr Sect F Struct Biol Cryst Commun;*
- Xue, B., Dunbrack, R.L., Williams, R.W., Dunker, A.K., Uversky, V.N., (2010), "PONDR-FIT: A meta-predictor of intrinsically disordered amino acids", *Biochim. Biophys. Acta, 1804*, 996-1010;
- Yeats C, Finn RD, Bateman A., (2002), "The PASTA domain: a beta-lactam-binding domain", *Trends Biochem Sci,* Sep;27(9):438;
- Yin, J., G. Li, X. Ren and G. Herrler, (2007), "Select what you need: a comparative evaluation of the advantages and limitations of frequently used expression systems for foreign genes" *J. Biotechnol*, 127(3): 335-347;

• Yukari C. Manabe & William R. Bishai, (2000), "Latent *Mycobacterium tuberculosis*–persistence, patience, and winning by waiting", *Nature Medicine*;

Outcome of the Ph.D. work

The results of my Ph.D. work were published in several research articles and disseminated through posters at conferences and one oral communication.

Publications

1. Calvanese L., Falcigno L., Maglione C., **Squeglia F.**, Berisio R. and D'Auria G., (2013), "Structural and binding properties of PASTA domain of ponA2, a Penicillin Binding Protein from *Mycobacterium tuberculosis*", *In preparation;*

2. Ruggiero A., De Simone P., Smaldone G., **Squeglia F.**, Berisio R., (2012), "Bacterial Cell Division Regulation by Ser/Thr Kinases: A Structural Perspective", *Curr Protein Pept Sci*;

3. Ruggiero A., Marchant J., **Squeglia F.**, Makarov V., De Simone A., Berisio R., (2012), "Molecular determinants of inactivation of the resuscitation promoting factor B from *Mycobacterium tuberculosis*", *J Biomol Struct Dyn;*

4. Berisio R., Ciccarelli L., **Squeglia F.**, De Simone A., Vitagliano L., (2012), "Structural and dynamic properties of incomplete immunoglobulin-like fold domains", *Protein Pept Lett*.;

5. Ruggiero A., Smaldone G., **Squeglia F.**, Berisio R., (2012), "Enhanced crystallizability by protein engineering approaches: a general overview", *Protein Pept Lett*.;

6. **Squeglia F.***, Marchetti R.*, Ruggiero A.*, Lanzetta R., Marasco D., Dworkin J.E., Petoukhov M., Molinaro A., Berisio R., Silipo A., (2011), "Chemical Basis of Peptidoglycan Discrimination by PrkC, a Key Kinase Involved in Bacterial Resuscitation from Dormancy", *J Am Chem Soc*. (* Shared first author);

7. Ruggiero A.*, **Squeglia F.*,** Marasco D., Marchetti R., Molinaro A., Berisio R., (2011), "X-ray structural studies of the entire extra-cellular region of the Ser/Thr kinase PrkC from *Staphylococcus aureus*", *Biochem J*. (*Shared first author);

8. Ruggiero A.*, **Squeglia F.***, Pirone L., Correale S., Berisio R., (2011), "Expression, purification, crystallization and preliminary X-ray crystallographic analysis of a major fragment of the Resuscitation-Promoting Factor RpfB from *Mycobacterium tuberculosis*", *Acta Crystallogr Sect F Struct Biol Cryst Commun*. (*Shared first author);

9. Ruggiero A., Marasco D., **Squeglia F.**, Soldini S., Pedone E., Pedone C. and Berisio R., (2010), "Structure and functional regulation of RipA, a mycobacterial enzyme essential for daughter cell separation", *Structure;*

10. Ruggiero A., **Squeglia F.**, Izzo V., Silipo A., Vitagliano L., Molinaro A., Berisio R., (2010), "Expression, purification, crystallization and preliminary X-ray crystallographic analysis of the peptidoglycan binding region of the Ser/Thr kinase PrkC from *Staphylococcus aureus*", *Protein & Peptide Letters*;

11. Ruggiero A., **Squeglia F.**, Esposito C., Marasco D., Pedone E., Pedone C., and Berisio R., (2010), "Expression, purification, crystallization and preliminary X-ray crystallographic analysis of the Resuscitation Promoting factor Interacting Protein RipA from *M. tuberculosis*", *Protein Pept Lett.;*

Oral presentation

"Bacterial Ser/Thr kinases and exit from dormancy: a structural perspective", XLI National Congress of the Italian Crystallographic Association, Verona, Italy, (11th-14th September 2012);

Posters presentation

- **Squeglia F.**, Ruggiero A., Chirafisi P., De Simone P., Berisio R., "Muropeptide-driven revival from dormancy in bacterial pathogens: structural studies", 45th Course of the International School of Crystallography "Present and Future Methods for Biomolecular Crystallography", Erice, Italy. (31th May-10th June 2012);

- **Squeglia F.**, Ruggiero A., Silipo A., Molinaro A., Berisio R., "Cell growth regulation in pathogenic bacteria by structural studies", International Year of Chemistry 2011, Monte Sant'Angelo, Naples, Italy. (13th December 2011);

- **Squeglia F.**, Ruggiero A., Marasco D., Berisio R., "Muropeptide-driven revival from dormancy in bacterial pathogens: a structural perspective", XII International Congress and General Assembly, International Union of Crystallography, Madrid, Spain. (August 22nd-30th 2011);

- **Squeglia F.**, Ruggiero A., Palamara L., Marasco D., Pedone E., Pedone C., and Berisio R., "Structural bases of *M. tuberculosis* daughter cell separation", 12th Naples Workshop on Bioactive Peptides and 2nd Italy-Korea symposium on anti-microbial peptides, Naples, Italy. (June 4th-7th, 2010);