DISS. ETH Nr. 23401

Architecture, Assembly and Interactors of the M*ycobacterium tuberculosis* Clp chaperone-proteases

A thesis submitted to attain the degree of

Doctor of Sciences of ETH Zurich

(Dr. sc. ETH Zurich)

presented by

Julia Leodolter

Mag. rer. nat in Molecular Biology, University of Vienna

born on 23.07.1985

citizen of Austria

accepted on the recommendation of:

Prof. Eilika Weber-Ban

Prof. Rudolf Glockshuber

Prof. Markus Seeger

Acknowledgements

First of all, I would like to thank my supervisor Eilika Weber-Ban for giving me the opportunity to work in her lab and for her great scientific guidance throughout the ups and downs of the Clp project. Thank you for always having an open ear for discussions and always taking the time to help. I really enjoyed the experience and learned a lot these past years.

Furthermore, I would like to thank my committee members, Rudi Glockshuber and Markus Seeger, for taking the time to advise me and for their helpful input.

My special thanks go to my present and past lab colleagues of the "E-floor madness", who shared all the highs and lows of PhD life with me. Thank you for the great working atmosphere, all the constructive, creative, earnest, fun and random discussions about science, life and everything: Michał Ziemski, Marcel Bolten, Jürg Laederach, Cyrille Delley, Andreas Müller, Anne Kerschenmeyer, Dennis Özcelik, Jannis Warweg, Jonas Barandun, Lena Poplutz, Fabia Canonica, Max Sauer; as well as all former students of the Weber-Ban lab and the other people of the Glockshuber group. My special thanks go to Michał for his tremendous help in the collaboration with the bacterial two hybrid screen. Thanks also to Marcel and Marc Leibundgut for their advice in all things crystallographic; and to my master student Anne, for putting all the effort into the Clp project, it was a great time working together.

Thanks to my friends and family near and far for believing in me and helping me keep up my spirits.

Last, I would like to thank my partner Andi, for his never-ending love and support.

Abstract

The Clp chaperone-protease is an essential protein quality control machine in the pathogenic bacterium *Mycobacterium tuberculosis (Mtb)* that additionally performs controlled degradation of regulatory proteins. The Clp system features the modular architecture typical of chaperone-proteases, where hexameric chaperone rings of the ClpX and ClpC type associate with a tetradecameric ClpP double-ring proteolytic core via conserved interaction motifs. The chaperones recognize and unfold substrates, and translocate them into the proteolytic chamber for degradation. In recent years, the Clp system of *Mtb* arose as a promising drug target, as dysregulation of chaperone or protease components leads to detrimental effects in the cell. To gain a further understanding of this essential complex, this thesis investigates the Clp chaperone-protease system with respect to the assembly of the functional chaperone-protease complex and aims to identify further adaptor and substrate proteins.

The Clp system of *Mtb*, in contrast to well-studied systems like *E. coli*, features not only one, but two protease subunits. Homoheptameric ClpP1 and ClpP2 rings assemble to a hetero ClpP1P2 double ring complex. By using mutational studies of the chaperone interaction interfaces of ClpP1 and ClpP2 and subsequent chaperone-mediated degradation assays, I show that both chaperones, ClpX and ClpC1 interact with the ClpP2 side of the proteolytic core. This creates an unusual asymmetric Clp chaperone-protease assembly since for the well-studied *E. coli* system, chaperone binding to both sides of the proteolytic core is described as the more efficient assembly. It however provides the ClpP1 side of the ClpP1P2 complex with the opportunity for binding of novel interactors. Further asymmetric behavior was observed on the level of propeptide cleavage, where I show that ClpP1 is the main actor in this process, indicating that the ClpP1 and ClpP2 display divergent, but complementary functions.

The correct assembly of the chaperone-protease is critical for it to perform its function of substrate protein degradation. As only few *bona fide* substrates of the *Mtb* Clp chaperone-protease system were known, this thesis aimed to increase the known substrate clientele. I show that the protein Rv1331 is the adaptor protein ClpS that delivers an N-end rule model protein to the ClpC1P1P2 complex, demonstrating that *Mtb* is equipped to deal with the substrate class of N-end rule substrates. To further identify substrates of the ClpC1 chaperone, I used it as a bait in a

bacterial two hybrid screen of an *Mtb* ORF library. In this screen we found that a significant portion of hits belong to the functional class of toxin-antitoxin systems. These systems are bipartite complexes where the ribonuclease toxin is activated by removal of the antitoxin by degradation and is implicated in the persistence of *Mtb.* The degradation of two antitoxins of the functional Vap and Rel classes by the ClpC1P1P2 complex was confirmed by *in vitro* degradation assays, expanding the substrate spectrum of the *Mtb* Clp chaperone-protease by another functional class.

Taken together, this work provides insights into a novel asymmetric assembly behavior of the *Mtb* Clp chaperone-protease complex and increases the known substrates classes by N-end rule substrates and antitoxins. As in the library screen for ClpC1 more putative interactors were found, investigation of promising hits presents an avenue for future research.

Zusammenfassung

Clp Chaperon-Proteasen sind essentielle molekulare Maschinen im pathogenen Bakterium *Mycobacterium tuberculosis (Mtb),* die einerseits wichtige Funktionen der Proteinqualitätskontrolle erfüllen, aber auch für den kontrollierten Abbau von regulatorischen Proteinen verantwortlich sind. Das Clp System verfügt über die für Chaperon-Proteasen typische modulare Architektur, bei der die Chaperone ClpX und ClpC1 hexamere Ringe bilden, die über konservierte Motive mit dem tetradekameren Doppelring der Protease ClpP interagieren. Die Chaperone erkennen und entfalten Substrate, und translozieren sie in die proteolytische Kammer der Protease für den Abbau. In den letzten Jahren kam das Clp System von *Mtb* vermehrt als vielversprechendes Ziel für die Medikamentenentwicklung auf, da die Deregulation von sowohl Protease als auch Chaperon Komponenten zu schädlichen Effekten für die *Mtb* Zelle führt. Um ein besseres Verständnis für diesen wichtigen Komplex zu entwickeln, untersucht diese Arbeit das Clp Chaperon-Protease System in Hinblick auf den Aufbau des funktionellen Komplexes und zielt außerdem darauf ab, sowohl Adaptoren als auch Substrate dafür zu identifizieren.

Das Clp System von *Mtb,* im Gegensatz zum bereits gut erforschten System von *E. coli*, verfügt nicht nur über eine, sondern über zwei ClpP Proteaseuntereinheiten. Jeweils ein homoheptamerer Ring von ClpP1 und ClpP2 kommen zusammen, um einen Hetero-ClpP1P2 Doppelringkomplex zu bilden. Mittels Mutationsstudien von für die Chaperonbindung wichtigen ClpP1 und ClpP2 Interaktionsoberflächen und anschließendem Test von Chaperon-vermitteltem Substratabbau, zeige ich, dass beide Chaperone, ClpX und ClpC1, mit der ClpP2 Proteaseuntereinheit interagieren. Das resultiert in einem asymmetrischen Aufbau des Chaperon-Proteasekomples, der insofern ungewöhnlich ist als dass für das besser erforschte *E. coli* System gezeigt wurde, dass Chaperonbindung nicht nur an eine, sondern an beide Seiten der Protease effizienteren Substratabbau ermöglicht. Trotzdem könnte dieser asymmetrische Aufbau einen Vorteil für das *Mtb* Clp System darstellen, da dadurch die Interaktionsoberfläche der ClpP1 Protease frei wäre um unbekannte potentielle Interaktoren zu binden. Die *Mtb* ClpP Untereinheiten zeigen außerdem nicht nur asymmetrisches Verhalten bei der Chaperonbindung, sondern auch in ihrer Substratspezifität. Am Beispiel der Prozessierung der ClpP Propetide zeige ich, dass bei diesem Prozess ClpP1 die Hauptrolle spielt, währen ClpP2 kaum Aktivität zeigt, was

darauf hinweist, dass ClpP1 und ClpP2 unterschiedliche, aber sich ergänzende Rollen entwickelt haben.

Die korrekte Bildung des Chaperon-Protease Komplexes ist essentiell, um kontrollierten Substratproteinabbau durchführen zu können. Da zur Zeit nur wenige Substrate für den Komplex bekannt sind, hat diese Arbeit zum Ziel, das bekannte Substratspektrum zu erweitern. Hier zeige ich, dass das Protein Rv1331 das Adaptorprotein ClpS ist, da es ein Modellsubstrat der „N-end Regel“ Klasse zum ClpC1P1P2 Komplex zur Degradation liefert, was wiederum bedeutet, dass *Mtb* dafür ausgestattet ist die N-end Substratklasse zu verdauen. Um weitere Substrate des ClpC1 Chaperons zu identifieren, wurde es als Köder in einem Bakterien-Zwei-Hybrid Experiment eingesetzt und in einer Bibliothek aus *Mtb* Proteinen nach Bindern gefischt. In den Resultaten konnte eine signifikante Menge von möglichen Interaktoren der Funktionsklasse der Toxin-Antitoxin Systeme zugeordnet werden. Toxin-Antitoxin Systeme sind zweiteilige Komplexe, bei denen das Ribonuklease Toxin durch die degradative Entfernung des Antitoxins aktiviert wird, und Funktionen erfüllt, die mit der Persistenz von *Mtb* in Verbindung gebracht werden. Der Abbau zweier Antitoxine der funktionellen Klassen Vap und Rel durch den ClpC1P1P2 Komplex wurde in dieser Arbeit durch *in vitro* Experimente gezeigt und damit das Substratspektrum der *Mtb* Clp Chaperon-Protease durch eine neue funktionelle Klasse erweitert.

Diese Arbeit präsentiert neue Einblicke in den asymmetrischen Aufbau des *Mtb* Clp Chaperon-Protease Komplexes und beschreibt neue Substratklassen des Systems. In den Experimenten mit dem Bakterien-Zwei-Hybrid System wurden außerdem weitere potenziell interessante Substrate gefunden, die es sich lohnen würde ebenfalls zu erforschen.

Contents

Acknowledgements ..... **i**

Abstract ..... **iii**

Zusammenfassung ..... **v**

Contents ..... **vii**

Chapter 1 Introduction **1**

1.1 Energy-dependent protein degradation in bacteria 1

1.2 Architecture and occurrence of bacterial chaperone-protease complexes 2

1.2.1 Occurrence of bacterial chaperone-proteases 5

1.3 The Clp chaperone-protease system 6

1.3.1 Architecture and assembly of Clp chaperone and protease complexes 7

1.3.2 Degradation mechanism 13

1.4 Substrate classes of the Clp chaperone-protease system 16

1.4.1 Substrates of protein quality control 16

1.4.2 Substrates of regulative degradation 17

1.5 The Clp system of *Mtb* 18

1.6 The *Mtb* Clp system as a drug target 21

1.6.1 Mycobacterium tuberculosis 21

1.6.2 Antibacterial compounds acting against components of the *Mtb* Clp system 22

1.7 Aim of the thesis 24

Chapter 2 The *Mycobacterium tuberculosis* ClpP1P2 Protease Interacts Asymmetrically with its ATPase Partners ClpX and ClpC1 ..... **27**

2.1 Introductory statement 27

2.2 Abstract 27

2.3 Introduction 28

2.4 Materials and Methods 30

2.5 Results 32

2.5.1 The ClpP1P2 double-ring complex can be assembled without activator peptide and is processed in a chaperone-dependent manner 32

2.5.2 ClpP1 is the main actor in ClpP1P2 double-ring processing 39

2.5.3 ClpP2 is the main interaction platform for the ATPase rings 40

2.6 Discussion 45

2.7 Acknowledgment 49

2.8 Supplementary Information 50

Chapter 3 Interaction partners of the chaperone ClpC1 and its N-terminal domain homologue ClpC2 59

3.1 Introduction 59

3.2 Results 59

3.2.1 Sequence relationship between ClpC1 and ClpC2 59

3.2.2 Screening for interactors of ClpC1 in an *Mtb* ORF library using the BACTH system. 62

3.2.3 Proof of principle tests for the BACTH system show expected interactions between the Clp chaperone and protease subunits 63

3.2.4 Screening of ClpC1 and ClpC2 against an *Mtb* ORF library 64

3.2.5 ClpS as a binder of ClpC1 68

3.2.6 BACTH co-transformation confirms library screen hits of toxins and antitoxins as binders of ClpC1 and ClpC2 70

3.2.7 Antitoxins of the Vap and Rel class are degraded by the ClpC1P1P2 chaperone-protease 74

3.2.8 The putative adaptor ClpC2 does not significantly influence antitoxin degradation. 77

3.3 Materials and Methods 80

3.4 Discussion 86

3.4.1 BACTH library screens for ClpC1 and ClpC2 reveal toxin-antitoxin systems as binders 86

3.4.2 The putative adaptor protein ClpC2 90

3.4.3 Rv1331 is the N-end rule adaptor ClpS in *Mtb* 91

3.4.4 The chaperone ClpC1 autodegrades in absence of substrate 92

3.4.5 Notes on the activator peptide 93

Chapter 4 Conclusion 95

References 98

Appendix 111

Abbreviatons 123

Curriculum Vitae 124

Chapter 1 Introduction

1.1 Energy-dependent protein degradation in bacteria

Bacterial cells live in all but the harshest milieus on earth, where they are challenged by a multitude of factors, ranging from hostile, rapidly changing environments to competition by other organisms. To survive in a given setting, cells have to adapt to their surroundings. One important cornerstone of adaptation is protein quality control. Stressful conditions damage proteins or cause misfolding, resulting in detrimental effects to the cell. Therefore, cells employ a network of protein quality control systems composed of molecular chaperones and proteases to deal with protein damage. Molecular chaperones alleviate protein damage by refolding of damaged proteins as well as assist in folding of newly synthesized proteins. Proteases on the other hand degrade damaged/unwanted proteins before they can form toxic aggregates.

Proteolysis in the cell must be performed in a highly controlled manner and the proteases involved are tightly regulated to prevent unspecific degradation. The key to controlled degradation lies in compartmentalization – chaperone-proteases form multi-subunit complexes of roughly cylindrical shape, where the active sites of the protease are located inside a proteolytic chamber and away from the cellular environment. A narrow pore allows only unstructured substrates to enter this chamber, where they are processively degraded into peptides. The access to the proteolytic chamber itself is guarded by regulatory ATPases of the AAA+ (ATPase associated with various cellular activities) type which associate to the proteolytic core and are responsible for substrate recognition. They use ATP turnover to unfold their protein substrates and thread them into the proteolytic chamber for degradation (Gur et al, 2013; Sauer & Baker, 2011).

In bacteria, degradation by chaperone-proteases is not only responsible for stress relief of damaged proteins. Under nutrient-limiting or rapidly changing conditions, controlled removal of old or unnecessary proteins provides resources for building fresh proteins suitable to the new situation. Furthermore, chaperone-protease complexes play an important role in specifically degrading regulatory proteins under certain stress conditions. Lastly, controlled degradation is involved in housekeeping proteostasis by for example removing proteins that are incompletely synthesized. Chaperone-proteases work together with small adaptor proteins that expand and/or modulate their substrate spectrum (Gur et al, 2013; Sauer & Baker, 2011).

1.2 Architecture and occurrence of bacterial chaperone-protease complexes

Unlike eukaryotes with the 26S proteasome, bacteria have not only one main protein degradation machine, but employ a number of partially redundant degradation complexes. The usual bacterial set of ATP-dependent chaperone-proteases contains the protease Lon, the Clp system (ClpXP, ClpAP, ClpCP, ClpEP), HslUV (ClpYQ) and the membrane associated FtsH (Figure 1.1). The general architecture of chaperone-proteases is conserved throughout bacteria as well as archaea and eukaryotes. Protease and chaperone subunits assemble into ring-shaped oligomeric structures that stack onto one another to form barrel-shaped complexes. While the AAA+ chaperone components always form hexameric rings, the proteases form either hexameric or heptameric rings, resulting in some of the AAA+ protease complexes to be symmetric, and others to be asymmetric assemblies (Figure 1.1). In the functional assembled state the proteolytic core rings usually form the center of the cylinder, and are sandwiched on one or both sides by a hexameric chaperone (unfoldase) ring. The chaperone unfoldases all belong to the AAA+ type and contain one or two AAA+ modules. The AAA+ modules are composed of a P-loop ATPase domain and a subsequent smaller α-helical domain that are used to convert ATP hydrolysis to unfolding and translocation action (Snider et al, 2008). While the AAA+ domains are highly conserved between AAA+ ATPases, they in addition contain unique, non-homologous N-terminal domains which are not required for basic unfolding and threading activity, but allow for selective adaptor and substrate binding (Sauer & Baker, 2011).

AAA+ proteases can be categorized into two groups, depending on whether the AAA+ domains are on the same polypeptide as the protease or whether they are expressed as separate proteins (Gur et al, 2013) (Figure 1.1). The Lon and FtsH proteases belong to the first group where the AAA+ unfoldase and the proteolytic domain are on the same polypeptide. In the assembled, functional state they form homo-hexameric barrel shaped complexes. The second group contains modular two-component systems, where chaperones and proteases are produced as separate proteins that associate to perform their function, which is the case for HslUV and the Clp system. The Clp system demonstrates the advantage of a modular architecture, as different regulatory partners can interact with the same proteolytic core (Imkamp et al, 2015; Laederach et al, 2014).

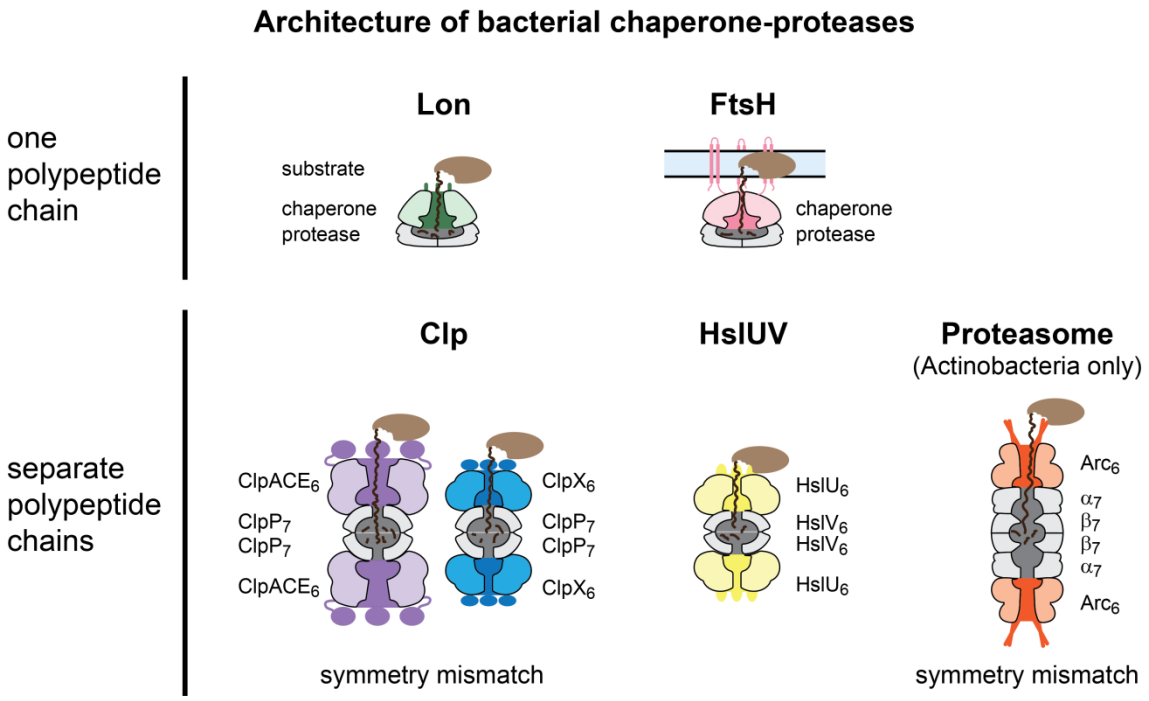


Figure 1.1: Architecture of bacterial chaperone-protease complexes. Cartoon depiction of the architecture of bacterial chaperone-protease-complexes. The protease core is always shown in grey, and the chaperone components in colors. Unique N-terminal regions are indicated as different shapes on top of the chaperone subunits. The substrate (brown) is translocated through the axial pore and degraded inside the proteolytic chamber. The Lon and FtsH protease and chaperone subunits are encoded on one polypeptide chain and they form symmetric hexameric assemblies. The Lon N-terminal domains can interact to form a double complex. FtsH features transmembrane segments that anchor it in the cell membrane (light blue). In the Clp system, HslUV and the Proteasome system, chaperone and protease subunits are encoded on separate polypeptide chains. HslUV forms a symmetric particle with hexameric protease ($HslV_6$) rings associating with hexameric chaperone rings ($HslU_6$). The Clp and proteasome architecture is characterized by a symmetry mismatch where heptameric protease particles ($ClpP_7$ or the proteasomal α_7 and β_7 subunits) associate with hexameric chaperone rings (ClpA/C/E_6, $ClpX_6$, Arc_6).

The following sections give a brief overview over the individual chaperone-protease systems, which are depicted in Figure 1.1.

Lon

The Lon assembled complex appears as a double-ring with AAA+ and protease domain appearing as a separate ring structure in electron microscopy images (Park et al, 2006). Lon features in addition to the AAA+ and the protease an N-terminal domain which was recently shown to be capable of forming head-to-head double-Lon complexes, thereby regulating substrate access to the protease (Vieux et al, 2013). Lon features a Ser-Lys catalytic dyad and is a broad specificity protease that recognizes and cleaves hydrophobic regions containing aromatic residues exposed only in unfolded proteins (Gur et al, 2013).

FtsH

FtsH is a membrane-anchored protease, located at the cytoplasmic side of the inner membrane. The structure of the soluble part of FtsH shows the hexameric assembly with protease and chaperone domains arranged in separate rings (Bieniossek et al, 2006; Suno et al, 2006). The protease active site is a zinc metalloprotease, and the substrate spectrum of FtsH includes mostly unfolded cytoplasmic and membrane proteins, as its ATPase domain has weaker unfolding activity than related AAA+ unfoldases (Ogura et al, 2013).

Clp

In the Clp chaperone-protease system the protease subunit ClpP forms heptameric rings that stack back-to-back to form a double-ring proteolytic core with a chymotrypsin like catalytic triad (Ser-His-Asp). The hexameric ATPases (ClpX, ClpA, ClpC, ClpE) associate to both sides of the proteolytic core and can be grouped according to the number of AAA+ modules they harbor. ClpX contains one AAA+ module, while ClpA, ClpC and ClpE harbor two. Structurally, the AAA+ modules arrange in a separate ring each, giving these chaperones a double-ring appearance. The chaperones differ in their N-terminal domains that confer substrate selectivity and recognize broad substrate classes such as proteins tagged with the ssrA peptide or with destabilizing N-terminal residues (N-end rule pathway) as well as specific regulatory proteins (Kress et al, 2009a).

HslUV

HslUV (also termed ClpYQ) particles are composed of hexameric HslV (ClpQ) protease double-rings, that associate with the hexameric HslU (ClpY) chaperone rings binding to both sides of the proteolytic double-ring core (Sousa et al, 2000). The protease active residue is a threonine, and the substrate spectrum of HslUV overlaps for example with the Lon protease (Wu et al, 1999). The chaperone does not feature a unique domain on the N-terminus as described for the other AAA+ chaperones, but instead features an insertion within the AAA+ module. This insertion is referred to as the intermediate domain and has a function in substrate recognition similar to the N-domains, but also in degradation (Sundar et al, 2012).

Proteasome

The bacterial proteasome is a special case, since it does not belong to the canonical set of bacterial chaperone-proteases. It is homologous to the eukaryotic and archaeal proteasomes and was likely obtained by horizontal gene transfer. It is restricted to the actinobacterial phylum, where it occurs alongside the typical bacterial chaperone-protease systems such as the Clp system and

FtsH (Laederach et al, 2014). The proteasome core particle (20S proteasome) is composed of two α- and two β-subunits that form heptameric rings each, stacking in a α_7-β_7-β_7-α_7 fashion, where the β-subunits carry the active site Thr residue. A specific AAA+ ATPase referred to as ARC (ATPase forming ring-shaped complexes) has been shown to interact with the 20S core to bring about the degradation of proteins covalently modified with the prokaryotic ubiquitin-like protein Pup. ARC features one AAA+ module for unfolding and an N-terminal coiled-coil region responsible for recognition of the Pup tag. Recently, also a non-ATPase binding partner for the proteasome was described that promotes the degradation of unfolded protein substrates (Delley et al, 2014).

1.2.1 Occurrence of bacterial chaperone-proteases

As mentioned above, energy-dependent protein degradation in bacteria can be mediated by a number of canonical bacteria-typic complexes, namely Lon, the Clp system (ClpXP, ClpAP, ClpCP, ClpEP), HslUV (ClpYQ) and FtsH, and exceptionally in the Actinobacteria, be augmented by a eukaryotic-like proteasome complex.

The combinations of these degradation systems co-occurring in the various bacterial groups vary. While FtsH, the membrane-associated member, occurs in all bacteria, the soluble complexes are never present all at the same time. With the exception of the genus mycoplasma that features an extremely reduced genome and only encodes Lon and FtsH, other bacteria usually feature some subset of Clp protease complexes (Gur et al, 2013; Staats et al, 2007). Gram negative bacteria, like for example *E. coli*, harbor the ClpAP and ClpXP protease, while Gram positive bacteria contain ClpCP and ClpXP complexes. Lon protease, the degradation complex for removal of damaged proteins, and the HslUV protease are featured in most bacteria alongside the Clp complexes. Mycobacteria are an exception to this, they do not encode HslUV, but on the other hand they have added the proteasome to their repertoire. The minimal set of chaperone-proteases in mycobacteria is FtsH, the Clp system (ClpXP, ClpCP) and the proteasome. Some mycobacterial members, e.g. *Mycobacterium smegmatis*, contain in addition to the minimal set the Lon protease (Laederach et al, 2014; Ribeiro-Guimaraes & Pessolani, 2007). In *Mtb*, only the Clp system and FtsH are essential for the bacterium, while the proteasome is required under nitric oxide stress and for persistence of the bacterium in the host cells (Darwin et al, 2003; Gandotra et al, 2007).

As the topic of this thesis is the *Mtb* Clp system, the following chapters focus on the Clp system, with specific attention to the *Mtb* proteins.

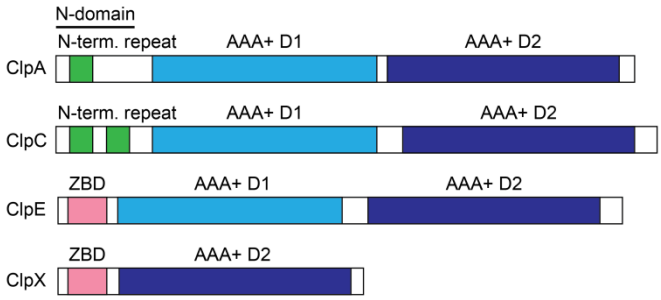


Figure 1.2: Domain organization of ClpP-associating Clp chaperone ATPases. ClpA, ClpC and ClpE contain two AAA+ modules each (light and dark blue), and ClpX only one (dark blue). The ClpA and ClpC N-terminal domains (N-domains) are composed of one or two N-terminal repeats (N-term. repeat, green), while ClpE and ClpX carry an N-terminal Zinc binding domain (ZBD, pink).

1.3 The Clp chaperone-protease system

The Clp chaperone-protease system, named for its ability to degrade the model substrate casein (caseinolytic protease) is composed of the proteolytic ClpP subunits and several alternative AAA+ chaperone binding partners. The proteolytic core is composed heptameric ClpP rings that in turn assemble to tetradecameric double-ring proteolytic core complexes (Laederach et al, 2014; Sauer & Baker, 2011).

Amongst the AAA+ chaperones, ClpX, which has only one AAA module, exists ubiquitously in all bacteria and co-occurs with a double-AAA module chaperone, either ClpA (in gram-negative bacteria) or ClpC (in gram-positive bacteria and cyanobacteria) as well as ClpE (in Firmicutes). The different chaperones carry N-terminal domains that confer substrate selectivity: ClpA and ClpC have one or two N-terminal repeat sequences, while ClpX and ClpE have a zinc binding domain (Figure 1.2).

In addition to the basic chaperone-protease core, the Clp system employs adaptor proteins that increase or change the substrate recognition capabilities of the chaperone. Usually, adaptors are small proteins that are specific for a certain chaperone. In *E. coli*, both ClpA and ClpX have the ability to degrade proteins featuring a specific C-terminal 11 amino acid long tail, called the ssrA tag. However, in case of ClpX the adaptor SspB increases the affinity for ssrA-tagged substrates by delivering them to the chaperone. In case of ClpA on the other hand, the adaptor ClpS switches the substrate specificity of ClpA away from ssrA-tagged substrates to substrates of the so-called N-end rule class. This way substrate flow to one or the other complex can be flexibly controlled via the adapters. An adaptor protein for *B. subtilis* ClpC is for example the protein

MecA which triggers ClpC oligomerization and ATPase activity, as well as aiding in degradation of substrates. Apart from these general adaptor proteins, adaptors or even anti-adaptors exist that are involved in regulative degradation of specific substrates (Battesti & Gottesman, 2013; Kirstein et al, 2009b)

1.3.1 Architecture and assembly of Clp chaperone and protease complexes

The protease core

The ClpP protease core is a tetradecameric double-ring complex, composed of 14 ClpP subunits. The ClpP subunits are produced as proenzymes, which become mature ClpP subunits once an N-terminal stretch of aminoacids is autocatalytically removed in the assembled complex (Maurizi et al, 1990). The molecular mass weight of ClpP subunits is around 22 kDa, resulting in a tetradecameric, fully assembled complex of about 300 kDa, with dimensions of approximately 90-100 Å in both diameter and height (Figure 1.3 A) (Schmitz et al, 2014; Wang et al, 1997). The enclosed proteolytic chamber is a spherical space of around 50 Å diameter, suitable to fit several hundred amino acids of unfolded substrates. The 14 active sites of the serine-histidine-aspartate catalytic triad line the walls of the chamber near the equatorial plane of the complex (Liu et al, 2014).

The ClpP monomer has been likened to a hatchet, composed of a large, globular head-domain and smaller handle domain plus an N-terminal loop region near the axial pore (Figure 1.3 B). Intra-ring interactions in the heptamer are stabilized mainly by hydrophobic interactions between the head domains (Wang et al, 1997), and inter-ring interactions are mediated also by the more flexible handle domains (Liu et al, 2014). While the intra-ring interactions keep the heptameric rings very stable, the interactions between the two heptamers in the double-ring are less stable and more susceptible to various buffer conditions, e.g. *E. coli* ClpP double-rings can be dissociated into single rings by changes in ionic strength and glycerol content (Maglica et al, 2009; Maurizi et al, 1998). At the axial pore of the ClpP rings where substrates enter the proteolytic chamber, the N-terminal loop is located. It is a stretch of about 15 amino acids that can form a β-hairpin (β-turn-β) structure (Figure 1.3 C, right panel). In crystal structures, the N-terminal loop was observed to either adopt an ordered “up” conformation, where the N-terminal loops reach upwards from the surface of the molecule either as loops or as a β-hairpin structure (Figure 1.3 C, left and right panel) or a partly disorderd “down” conformation where they point inwards to the proteolytic core, possibly obscuring the entry through the pore (Figure 1.3 C, middle panel) (Bewley et al, 2006; Liu et al, 2014).

Crystallization studies of ClpP double-rings of various organisms described three conformational states of the cylinder: extended, compact and compressed (Figure 1.3 D).

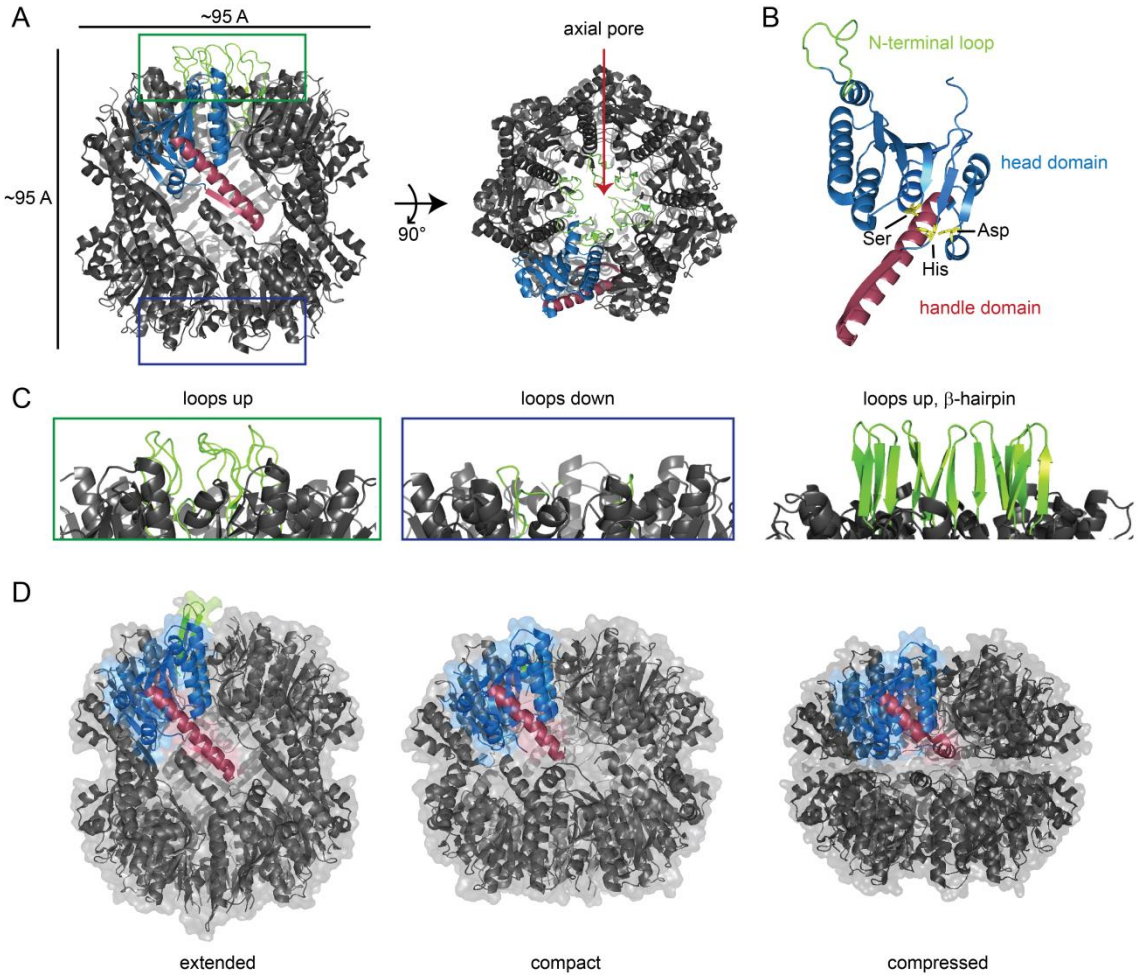


Figure 1.3: The ClpP protease core. A. Side view and top view of the *E. coli* ClpP double-ring (1YG6.pdb). One subunit is colored according to its domain composition (see B). The N-terminal loop regions are found in two different conformations in this particle that are shown in a close up in panel C (green and blue box). **B**. A ClpP subunit (*E. coli,* 1YG6.pdb; grey) is composed of the head domain (blue; residues 17-124; 158-193), the handle domain (red; 125-157) and an N-terminal loop (green; residues 1-16). The active side residues are shown in stick representation in yellow. **C.** Conformations of the ClpP N-terminal loop in close up view. The loops can be in the 'up' or 'down' conformation (1YG6.pdb; *E. coli* ClpP) and also adopt a β-hairpin conformation (4U0G.pdb; *Mtb* ClpP2). **D.** The extended (*S. aureus*, 3STA.pdb), compact (*S. aureus*, 4EMM.pdb) and compressed (*S. aureus*, 3ST9.pdb) conformations of ClpP.

While in all three states a large part of the head domain stays unchanged, the more flexible handle domains and some head domain residues involved in the inter-ring interactions undergo conformational changes (Liu et al, 2014). In the extended conformation, the handle region is well structured, the proteolytic chamber is closed off from the environment and the active site is in an ordered, active conformation. In the compact form of ClpP, the handle regions are partly unstructured and the active site is in a disordered, inactive conformation. Similarly, in the compressed form the handle region is partly unstructured, the active site distorted, but pores of around 6 Å in diameter appear in the equatorial plane of the molecules, that could be involved in product release (Liu et al, 2014). Molecular dynamic simulations suggest dynamic switching of

ClpP between these three states (Ye et al, 2013).

Assemblies of ClpP in various organisms differ depending on the number of ClpP subunits encoded and expressed in the respective organism. The gram-negative and gram-positive model organisms where the Clp system was extensively studied, *E. coli* and *B. subtilis*, harbor only one type of ClpP subunit. The resultant proteolytic core complexes are therefore homo-oligomeric tetradecamers. Other organisms however, for example *Mtb*, carry two copies in their genome, *clpP1* and *clpP2,* where one homo-heptameric ring of each subunit assemble to the functional double-ring complex (Akopian et al, 2012; Schmitz et al, 2014). In contrast, cyanobacteria and chloroplasts also feature several ClpP isoforms, which however do not oligomerize into homo-heptameric rings, but form mixed assemblies where one heptameric ClpP ring is composed of different subunits. (Andersson et al, 2009; Nishimura & van Wijk, 2015).

The Clp chaperone components

The Clp chaperone subunits are composed of a unique N-terminal domain and one (ClpX) or two (ClpA, ClpC) conserved AAA+ modules (Figure 1.2). AAA+ modules are around 200-250 amino acids in size and composed of a large core domain and a smaller C-terminal helical domain (Figure 1.4 A). The large domain has a P-loop NTPase fold and contains two nucleotide binding motifs, termed Walker A and Walker B (Figure 1.4 A). Walker A (the P-loop) is important for binding the β- and γ- phosphates of ATP, and Walker B for nucleotide hydrolysis. In addition, AAA+ proteins contain additional conserved motifs, termed sensor and Box motifs, that are involved in ATP hydrolysis and subunit communication and movement (Snider et al, 2008).

The crystal structure of ClpX shows that large and small AAA+ domains are connected by a flexible hinge region (Figure 1.4 A and B). Interestingly, the small AAA+ domain of one subunit packs tightly to the large domain of the adjacent subunit forming a rigid body, and leaving the hinge region as the flexible connector (Glynn et al, 2009; Stinson et al, 2013). ATP is bound in a cleft between the large and small domains of one subunit and the large subunit of the adjacent ClpX subunit (Figure 1.4 A and B). The conformation of the hinge is related to the ability of the subunit to bind nucleotides and provides the conformational flexibility needed for coupling ATP hydrolysis to the substrate unfolding function (Glynn et al, 2009; Stinson et al, 2013). While ClpX features only one AAA+ module, ClpA/ClpC harbour two modules, termed D1 and D2 (Figure 1.4 D). The D1 module was shown to be necessary for ATP dependent oligomerization, D2 carries the bulk of the ATPase activity (Seol et al, 1995; Singh & Maurizi, 1994).

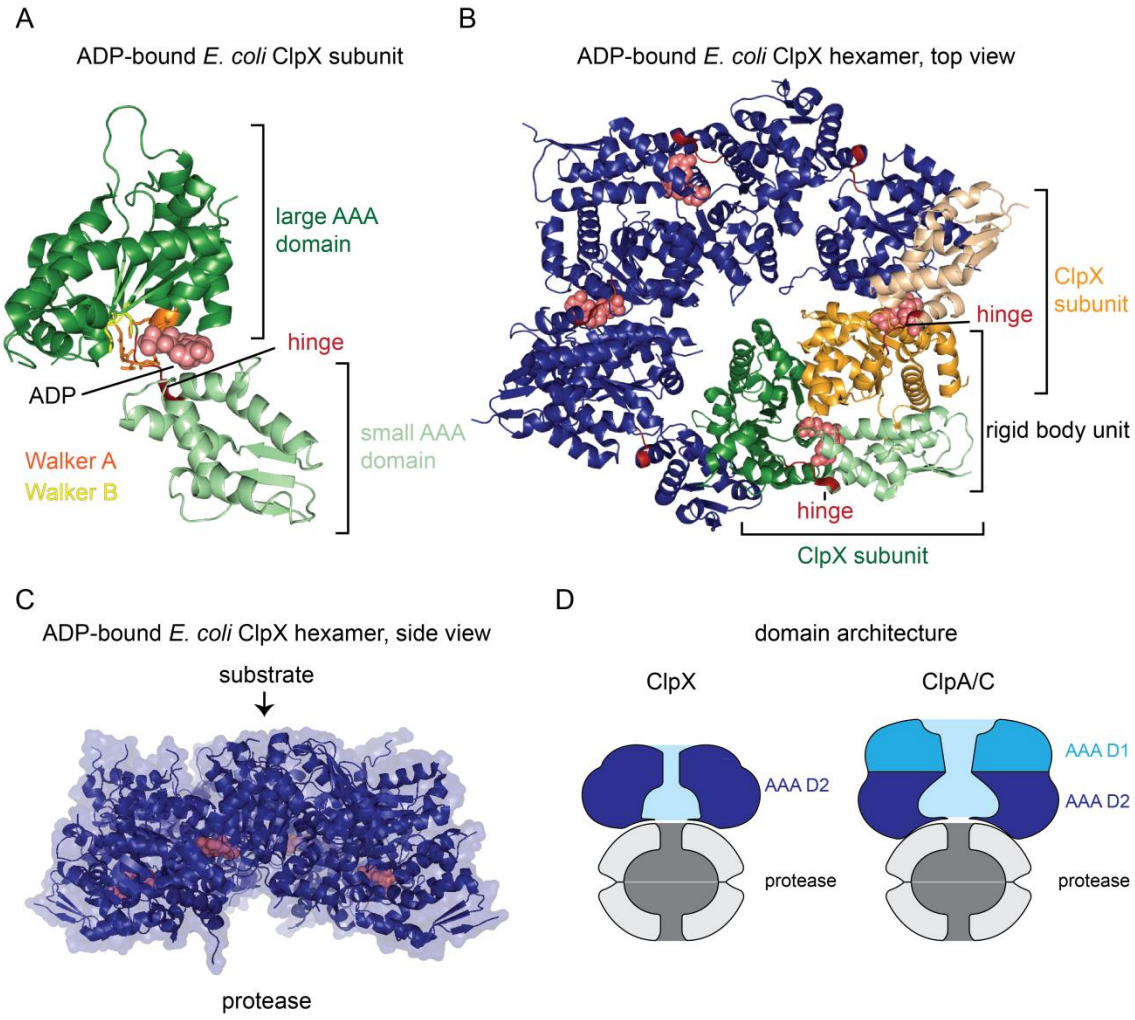


Figure 1.4: Clp chaperone rings. A. Cartoon representation of a subunit of the ClpX hexamer without the N-terminal Zinc binding domain (E. coli, 3HWS.pdb). The nucleotide ADP (sphere representation) is bound in a cleft between the large AAA domain (dark green; residues 61-314) and the small AAA domain (light green; residues 320-416). The residues of the Walker A (residues 119-126) and Walker B (residues 184-185) motifs are shown as sticks in orange and yellow, respectively. **B.** Cartoon representation of the ADP-bound ClpX hexamer in top view (E. coli, 3HWS.pdb). Four ADP molecules (sphere representation, light red) are bound in the structure. One ClpX subunit is colored in green (large AAA domain: dark green; small AAA domain: light green), an adjacent subunit is colored in orange (large AAA domain: orange; small AAA domain: light orange), the remainder of the ClpX subunits are depicted in blue. The hinge regions (red) connect the large and small domains of one ClpX subunit. A rigid body unit is formed by the large domain of one subunit (orange) and the small domain of the adjacent subunit (light green). **C.** Side view of a ClpX hexamer (E. coli, 3HWS.pdb) in blue with bound ADP in light red spheres. **D.** Schematic depiction of the domain architecture of ClpX containing one AAA module (dark blue; AAA D2) and ClpA/ClpC which are composed of two AAA modules (dark blue: AAA D2; light blue: AAA D1) associated to the protease core (grey).

Analysis of ClpA variants that were able to bind but unable to hydrolyse ATP in either one or the other AAA+ module showed that ATP hydrolysis occurs independently in D1 and D2. While the D2 module alone can support degradation of all but the most stable proteins, the D1 module barely shows any substrate unfolding capability. However, for optimal degradation of highly stable proteins both modules are necessary (Kress et al, 2009b).

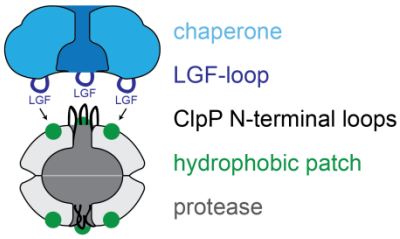


Figure 1.5: Clp chaperone and protease assembly. The LGF-loops (dark blue) of the chaperone (light blue) bind to hydrophobic patches (green dots) on the ClpP protease face (grey). Furthermore, the ClpP harbours N-terminal loops (black) in its axial pore that can adopt an "up" conformation for associating with the chaperone (upper protease ring), and a "down" conformation where the axial pore is blocked (lower protease ring).

In solution, ClpA exisits in a mixture of monomers and dimers (Maurizi et al, 1998). Only upon addition of ATP and Mg^{2+} it assembles to the active hexameric form via a tetrameric intermediate (Kress et al, 2007). The chaperone ClpX can hexamerize in the absence of nucleotide (Singh et al, 2001), nevertheless ATP stabilizes the interactions between the subunits and thereby promotes hexamer formation (Kim & Kim, 2003).

Assembly of chaperone-protease complexes

For successful assembly of protease and chaperones, both partners carry conserved interaction motifs (Figure 1.5). The chaperones carry a flexible loop with a conserved tripeptide sequence of two hydrophobic residues flanking a glycine ([LIV]-G-[FL]) (Kim et al, 2001). As in *Mtb* this loop has the LGF-motif on its chaperones, it is termed "LGF-loop" in this thesis (Figure 1.5). Mutation of the chaperones "LGF"-loops abolishes chaperone binding to the protease, and all six loops are necessary for stable interaction (Kim et al, 2001) (Martin et al, 2007). All Clp chaperones capable of interacting with the protease carry this loop. In addition, ClpX features another loop at the bottom of its pore (pore-2 loop) that is also involved in protease interaction (Martin et al, 2007).

The protease ClpP has two conserved interaction elements responsible for binding to the chaperone: a hydrophobic patch located on the face of the ClpP ring and the N-terminal β-hairpin loop at the central pore (Figure 1.5). The hydrophobic patch is formed by residues from adjacent ClpP subunits in the ring and allows for binding of the LGF-loops of the chaperones (Joshi et al, 2004; Kim et al, 2001; Wang et al, 1997). The N-terminal loop of ClpP is likewise required for successful interaction of the protease with the chaperone partner, and mutation of conserved residues in this region prevent stable chaperone binding (Bewley et al, 2006; Gribun et al, 2005; Kang et al, 2004). As mentioned above, the N-terminal loop regions can adopt an ordered "up"-conformation and a less ordered "down" conformation. While the "down" conformation is asso-

ciated with restricted access to the degradation chamber, the formation of the β-hairpin structure in the “up” conformation is necessary for efficient substrate translocation (Alexopoulos et al, 2013; Gribun et al, 2005). Cryo-electron microscopy studies show that the N-loops switch from a closed “down” conformation to the open “up” conformation upon binding of the chaperone ClpA (Effantin et al, 2010a; Effantin et al, 2010b).

The ClpP double-ring presents two chaperone-binding interfaces, one at each end of the cylinder. Cryo-electron micrographs of the *E. coli* Clp system show that single- as well as double-capped Clp complexes exist, where the chaperones bind to one or both sides of the ClpP core cylinder in a 1:1 or 2:1 stoichiometry. Additionally, mixed complexes of ClpP with the chaperone ClpX bound on one side and the chaperone ClpA on the other were observed (Grimaud et al, 1998). While single-capped ClpXP and ClpAP complexes are possible, double-capped complexes form more efficient assemblies by allowing for twice the degradation (Maglica et al, 2009).

When the hexameric rings of Clp chaperones associate with the heptameric rings of the proteolytic core, a symmetry mismatch occurs whose function is not clear. For chaperone-protease assembly, the chaperone rings assemble in the presence of ATP and the fully formed rings then associate with the preformed double-ring proteolytic core (Kress et al, 2007). The mitochondrial ClpXP complex forms an exception here, as there the ClpP double-ring formation from inactive single-rings is only induced by association of the ATP-bound ClpX hexamer (Kang et al, 2005).

1.3.2 Degradation mechanism

Recognition

Recognition of substrates for Clp chaperone-proteases is mediated by degradation motifs on the substrates. These degradation motifs are either directly recognized by dedicated loops in the central pore of the chaperone, or by the N-terminal domains featured by the chaperones. Additionally, adaptor proteins recognize specific substrates and deliver them to the chaperone, often by binding to the unique N-terminal domains (Figure 1.6, subtrate recognition) (Sauer & Baker, 2011).

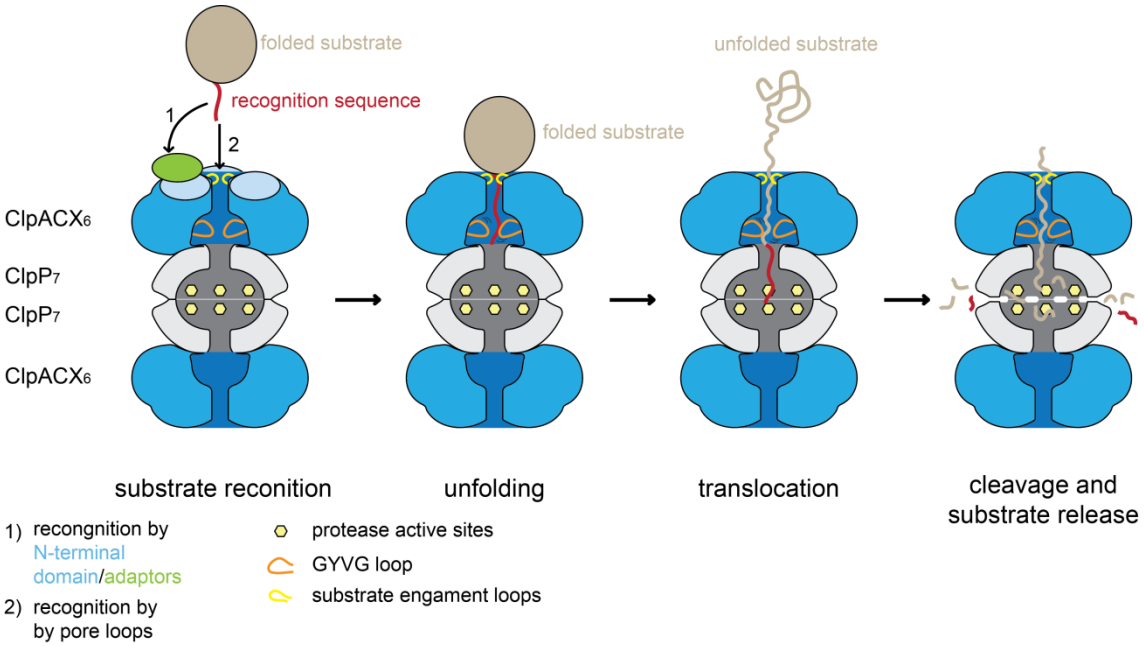


Figure 1.6: Degradation of substrates by the Clp chaperone-protease complex. Cartoon depiction of substrate degradation by the Clp chaperone-protease system. The chaperone ClpA/C/X hexamer is depicted in blue color, stacking to the protease core (grey). The protease active sites (yellow spots) are located inside the proteolytic chamber. Substrate (brown) recognition occurs via a recognition sequence (in red) that either interacts with the Clp chaperone N-terminal domain (light blue) directly or via an adaptor protein (green; 1). Alternatively, the recognition sequence interacts directly with loops in the axial pore of the chaperone (yellow; 2). After substrate engagement, GYVG loops (orange) grip the substrate and translate conformational changes caused by ATP hydrolysis into a mechanical pulling force. Once the substrate is unfolded, the chaperone translocates the polypeptide into the proteolytic chamber. Release of the cleaved peptides likely occurs through equatorial side pores in the ClpP double ring.

For example, ClpX binding of the ssrA-tag (*E. coli* sequence: AANDENYALAA) is dependent on binding of its two C-terminal residues to three loops in the ClpX pore (Farrell et al, 2007; Martin et al, 2007; Siddiqui et al, 2004). Substrate binding efficiency can additionally be mediated by the adaptor protein SspB, which binds to the N-terminal part of the ssrA-tag as well as the N-terminal domain of ClpX and effectively delivers ssrA-tagged substrates to ClpX (Kirstein et al, 2009b).

For ClpA-dependent degradation, the ssrA-tag binds similarly to loops in the ClpA pore, while another substrate, the protein RepA is recognized by the ClpA N-terminal domain (Hinnerwisch et al, 2005b). The N-terminal domain is attached via a flexible linker to the AAA+ ring and allows movement of the domain to the side of the ring, for example when ssrA-tagged substrates need to access the pore, and movement to the pore, when the adaptor protein ClpS binds to the N-terminal domain to deliver substrates (Cranz-Mileva et al, 2008; Ishikawa et al, 2004).

Unfolding and translocation

After engagement of the degradation tag, conserved loops in the axial channel of the chaperone pull on the substrates in an ATPase-mediated motion to unfold them and translocate them into the protease core (Figure 1.6, unfolding and translocation). Translocation of substrates is initiated either at the N- or the C-terminus of substrates. It starts with the translocation of the degradation tag, as was shown for both the ClpA and the ClpX chaperones (Kress et al, 2009b; Lee et al, 2001; Reid et al, 2001).

Unfolding and translocation of substrates by ClpXP has been extensively studied by single-molecule force spectroscopy using titin domains as substrate. After initial binding of the degradation tag, ClpX tries to unfold the substrate, resulting in a so-called dwell time, where the substrate is bound, but not unfolded or translocated. During the dwell time, ClpX expends ATP in attempts to unfold the substrate. Once the titin domain is unraveled, translocation of the unfolded polypeptide chain can occur (Olivares et al, 2016). The length of the dwell time depends on the stability of the local secondary structure near the degradation tag. Studies on a stable titin domain show that unfolding and translocation by ClpXP consume around 600 ATPs per molecule, while titin domains containing single point mutations that reduce the local stability of secondary structure near the engagement tag consume 100-200 ATPs. In both cases the bulk of ATP is used up in unfolding attempts of the folded domain, and only a small part in translocation of unfolded polypeptide chains (Kenniston et al, 2003; Olivares et al, 2016).

For substrate translocation a highly conserved loop in the pore of the chaperones is responsible, which in most AAA+ chaperones carries a glycine-aromatic-hydrophobic amino acid motif that is GYVG in Clp chaperones (Figure 1.6) (Kress et al, 2009a; Sauer & Baker, 2011). This loop, more specifically the tyrosine, is responsible for gripping substrates and translocating them through the pore, probably by adopting “up” and “down” states induced by conformation changes due to ATP hydrolysis (Hinnerwisch et al, 2005a; Martin et al, 2008b; Park et al, 2005).

Degradation and peptide release

The insides of the ClpP proteolytic chamber are mostly hydrophobic to provide an interaction interface for the unfolded substrates and to keep them in an unstructured state for processive degradation (Wang et al, 1997). The catalytic site serines within one heptamer are connected by a hydrophobic peptide binding groove and are spaced approximately 25 Å apart which fits to an average created peptide size of 7-8 aminoacids (Choi & Licht, 2005; Jennings et al, 2008; Wang

et al, 1997). The created peptide size is only determined by ClpP action, and not depending on ATPase binding or substrate translocation (Bewley et al, 2009; Jennings et al, 2008).

Peptide release from the proteolytic chamber occurs most likely through side pores formed in the compressed ClpP conformation due the dynamic nature of the ClpP handle domain (Figure 1.6, cleavage and substrate release) (Kimber et al, 2010; Sprangers et al, 2005).

1.4 Substrate classes of the Clp chaperone-protease system

The Clp chaperone-proteases are processive machines that recognize and degrade a broad range of substrates. On the one hand, they perform housekeeping roles involved in general protein quality control under standard conditions as well as under stress. On the other hand, they are involved in the regulatory degradation of specific substrates. The substrate clientele reflects these different roles.

1.4.1 Substrates of protein quality control

The Clp system is amongst others responsible for general protein quality control. Unfolded or partially unfolded proteins with low amounts of tertiary structure and/or a high amount of exposed hydrophobic stretches are recognized and degraded, such as the model substrate casein (Raju et al, 2012a).

One way cells ensure that only completely synthesized proteins are circulated is by having incompletely translated proteins removed by the Clp system. In case of stalled translation by the ribosome due to e.g. mRNA truncation or tRNA depletion, a rescue transfer-messenger RNA (tmRNA) resolves the block. It carries in its mRNA part the sequence for the SsrA peptide tag (*E. coli* sequence: AANDENYALAA; *Mtb* sequence: ADSHQRDYALAA) that is added C-terminally to the unfinished proteins. The stalled chain is transferred to the tRNA portion of tmRNA, after which the mRNA portion is translated. Then the mRNA-ribosome-protein complex dissociates, the stalled ribosome is released, and the ssrA-tagged protein is degraded (Moore & Sauer, 2007). ClpX, ClpA and ClpC have all been found to recognize ssrA-tagged proteins as their substrates (Gottesman et al, 1998; Laederach et al, 2014; Moore & Sauer, 2007).

The class of N-end rule substrates is another broad substrate class recognized by the Clp system. The N-end rule relates the stability of a protein to its N-terminal residue. In case of a primary destabilizing residue at the N-terminus of the protein (Phe, Trp, Tyr or Leu) the protein is recognized by the adaptor protein ClpS and delivered to the ClpAP complex for protein degradation. This reduces the half-life of these proteins in relation to other cellular proteins (Dougan et al,

2012; Dougan et al, 2010). ClpS is a small molecule (around 12 kDa) that is composed of a cone-shaped globular domain and an unstructured N-terminal extension. One of two conserved motifs on the globular domain and the N-terminal extension are important for ClpA binding and delivery of the N-end rule substrate. A second conserved motif is responsible for recognition and binding of the N-terminal residue of the substrate. A hydrophobic binding pocket accommodates the side chain of the N-terminal residue, while negatively charged residues at the rim of the pocket coordinate the α-amino group (Dougan et al, 2010; Roman-Hernandez et al, 2009; Schuenemann et al, 2009; Wang et al, 2008). The residues important for N-end rule and chaperone binding are highly conserved in ClpS homologues (Lupas & Koretke, 2003).

1.4.2 Substrates of regulative degradation

In addition to the general quality control substrate degradation, the Clp chaperone protease system also degrades substrates that are important players in regulative processes, ranging from stress response and cellular replication to regulation of virulence. These processes are often mediated by adaptor proteins. Below several examples are presented to illustrate regulatory degradation by the Clp system.

In *E. coli*, ClpXP is responsible for the degradation of the general stress response adaptation sigma factor RpoS (Vijayakumar et al, 2004). During normal growth conditions, RpoS is produced, subsequently recognized by its adaptor protein RssB (response regulator of RpoS) and delivered to ClpXP for degradation (Becker et al, 1999; Zhou & Gottesman, 1998). During stress conditions, the adaptor protein RssB is blocked by stress-induced anti-adaptor proteins (Bougdour et al, 2008). Therefore, RpoS is no longer delivered to the ClpXP protease and can act as a transcription factor.

In *Caulobacter crescentus*, a bacterium with a complicated life cycle, where asymmetric division results in two very different cell-types, a motile swarmer cell and a non-motile stalked cell. Replication of DNA only occurs in the stalked cell. In the motile cell, replication is inhibited by CtrA (cell cycle transcriptional regulator), until dephosphorylaction causes dissociation of CtrA from DNA. CtrA is then bound by the adaptor RcdA (regulator of CtrA degradation) and degraded by the ClpXP complex, thereby switching the cell to the DNA replicating stalked cell type (Iniesta et al, 2006; Jenal, 2009; Jenal & Fuchs, 1998).

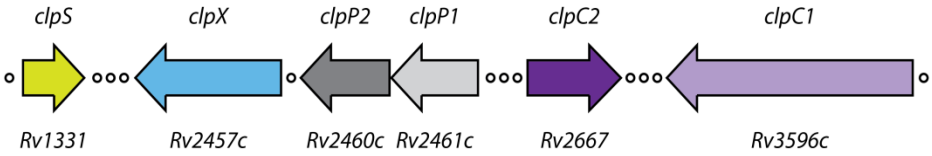


Figure 1.7: Gene loci of the Clp chaperone-protease system. The gene name is listed above the arrows indicating the gene, the respective gene number below.

Amongst the substrates of the ClpAP protease in *E. coli* is also the ClpA protein itself (Gottesman et al, 1990). In the absence of other substrates, ClpA recognizes its C-terminal residues and degrades itself in an autoregulative process (Maglica et al, 2008).

In *Staphylococcus aureus*, the ClpCP complex is responsible for the proteolytic regulation of toxin-antitoxin (TA) complexes (Donegan et al, 2010). TA complexes exist on plasmids or in the genome and are co-transcribed from the same operon. The most abundant TA complexes are type II systems, where both toxin and antitoxin are proteins and the antitoxin bound to the toxin blocks toxin action. Degradation of the antitoxin by cellular proteases frees the toxin, which is usually a ribonuclease, to perform its function (Brzozowska & Zielenkiewicz, 2013; Sala et al, 2014). In *S. aureus*, three TA systems exist, and are e.g. involved in regulation of stress response during heat or antibiotic stress (Donegan & Cheung, 2009), and cause bacteriostatic effects by selective mRNA degradation under stress conditions that are reversible by antitoxin expression (Fu et al, 2009). In all three TA systems of *S. aureus*, the antitoxin partners are degraded by the ClpCP complex (Donegan et al, 2010).

1.5 The Clp system of *Mtb*

In the *Mtb* genome, six genes have been annotated that express proteins belonging to the Clp chaperone-protease system (Figure 1.7). The two protease subunits ClpP1 and ClpP2 (Rv2461c, Rv2460c) form the proteolytic core to which the chaperones ClpX (Rv2457c) or ClpC1 (Rv3596c) associate. All four of these proteins are essential in *Mtb* and are the building elements of the two alternate AAA+ protease complexes ClpXP and ClpCP (Griffin et al, 2011; Sassetti et al, 2003). Furthermore, there are two non-essential proteins annotated as belonging to the Clp system, ClpS (Rv1331) and ClpC2 (Rv2667) (Griffin et al, 2011; Sassetti et al, 2003). ClpS is a putative adaptor for N-end rule substrates and not essential for *in vitro* growth of *Mtb*, but necessary for survival of the bacterium in macrophages (Rengarajan et al, 2005). ClpC2 is a small protein which shows homology to the ClpC1 N-terminal domain. Neither one of them is associated with either the *clpX* or the *clpC1* gene locus.

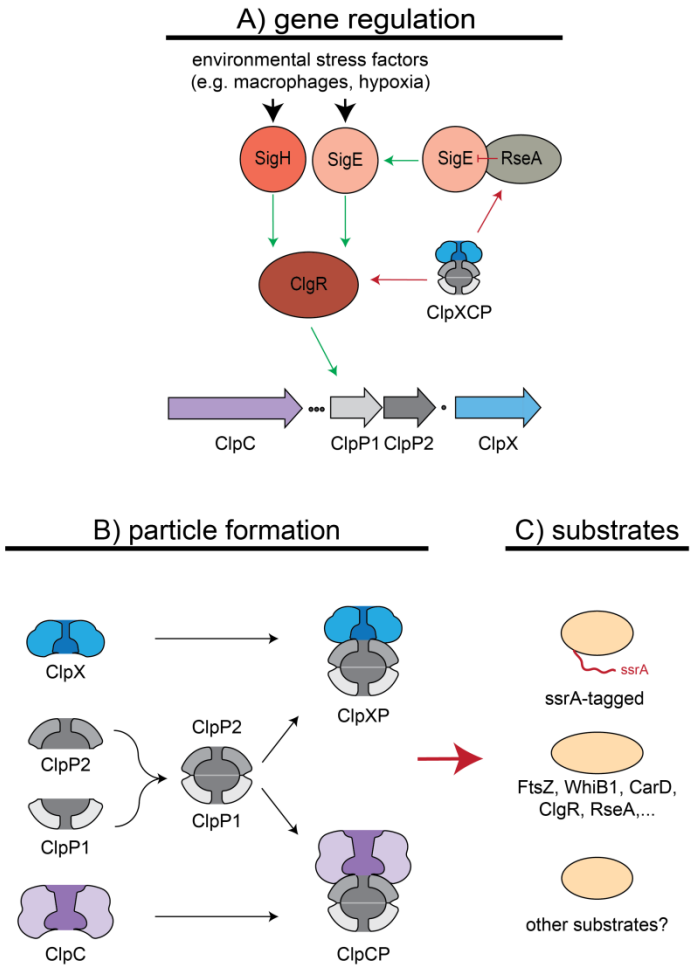


Figure 1.8: The Clp system of M. tuberculosis. A. Gene regulation of the Clp system. The Clp chaperone and protease genes are upregulated by ClgR (Clp gene regulator), which is in turn activated by the SigH and SigE, sigma factors that respond to environmental stress factors. The Clp system is involved in its own regulation. On the one hand, it is responsible for activation of SigE by degrading the anti-sigma factor RseA which in turn activates Clp expression. On the other hand it is involved in a negative feed-back loop by degrading its transcriptional activator ClgR. **B.** Clp chaperone-protease particle formation. The protease subunits ClpP1 and ClpP2 assemble to the active ClpP1P2 double-ring complex to which the chaperones ClpX and ClpC associate to form the degradation-competent complex. **C.** Substrates that are degraded by the Clp chaperone-protease system include ssrA-tagged proteins. Furthermore, several proteins were identified as substrates, like the cell-division protein FtsZ, the transcriptional regulators WhiB1, CarD and ClgR, and the anti-sigma factor RseA. (Figure adapted from (Laederach et al., 2014))

The Clp system in *Mtb* is regulated by the ClgR (Clp gene regulator) transcriptional activator. ClgR has close homologues in other actinomycetes, *Corynebacterium glutamicum* and *Streptomyces coelicolor*, where it activates Clp gene expression, but is negatively regulated by being a substrate of the Clp chaperone-protease itself (Figure 1.8 A) (Bellier et al, 2006; Bellier & Mazodier, 2004; Engels et al, 2005; Engels et al, 2004). In *Mtb,* ClgR up-regulates protease genes in response to reaeration after prolonged hypoxia (Sherrid et al, 2010) and ClgR-dependent regula-

tion of Clp is required for *Mtb* survival inside macrophages, indicating a link of the Clp system to the virulence of *Mtb* (Estorninho et al, 2010). Additionally, in response to environmental stress conditions the Clp system is induced via ClgR by the alternative sigma factor σ^H (Figure 1.8 A) (Mehra & Kaushal, 2009).

In *Mtb,* both chaperone subunits, ClpP1 and ClpP2, are essential and were found to associate into a ClpP1P2 double-ring complex (Figure 1.8 B). When affinity-tagged *Mtb* ClpP1 and ClpP2 were co-expressed in *M. smegmatis*, one subunit co-purified with the respective other (Akopian et al, 2012; Raju et al, 2012b) and cross-linking experiments showed that the rings formed homoheptamers composed of one ClpP subunit each (as opposed to rings formed of mixed ClpP1 and ClpP2 subunits as observed in Cyanobacteria) (Akopian et al, 2012). Both subunits, ClpP1 and ClpP2, are necessary for growth of the bacterium *in vitro* and in a mouse infection model (Raju et al, 2014; Raju et al, 2012b).

Studies on the *in vitro* assembly behavior of the ClpP1P2 complex reported somewhat puzzling results. Assembly of separately expressed ClpP1 and ClpP2 subunits was either not observed at all (Benaroudj et al, 2011), or was shown to only occur in the presence of a so-called "activator peptide". The activator peptide is usually either an N-blocked dipeptide, Z-Leu-Leu (Benzyloxycarbonyl-L-Leucyl-L-Leucine) or a related molecule (Akopian et al, 2012). In the latter study (Akopian et al, 2012), when ClpP1 and ClpP2 were expressed separately, they formed inactive ClpP1P1 or ClpP2P2 homo-tetradecameric double rings. In presence of the activator, these homooligomeric double rings dissociated and reassociated into active ClpP1P2 mixed double-rings (Akopian et al, 2012). So far, *in vitro* assembly of the ClpP1P2 core in absence of synthetic activator peptide was not observed. Interestingly, the activator rendered the *Mtb* ClpP1P2 complex that had exhibited no peptidase activity when co-expressed and co-purified, into an active peptidase (Akopian et al, 2012). How the activator performs this function was visible in a later crystal structure of the ClpP1P2 complex. The activator binds close to the active site of the protease and stabilizes the extended conformation of ClpP1P2 where the protease active site residues are in an ordered arrangement (Schmitz et al, 2014).

Inactivation of either the ClpP1 or the ClpP2 active site in the context of the assembled ClpP1P2 particle showed that they have distinct cleavage specificities for model peptides, and that both ClpP1 and ClpP2 can contribute to protein substrate degradation (Akopian et al, 2012; Schmitz & Sauer, 2014). ClpP1 and ClpP2 subunits are produced as proenzymes carrying N-terminal pro-

peptides that are cleaved off to form the mature complex (Akopian et al, 2012; Benaroudj et al, 2011).

Both chaperone components ClpX and ClpC1 can cooperate with the *Mtb* ClpP1P2 proteolytic core *in vitro* to perform degradation of model substrates (Figure 1.8 B) (Akopian et al, 2012; Schmitz & Sauer, 2014). The mycobacterial ClpC1, even though homologous to *B. subtilis* ClpC (around 60% sequence identity), does not need an adaptor for assembly and function, the presence of ATP suffices (Bajaj & Batra, 2012; Kar et al, 2008; Schmitz & Sauer, 2014).

SsrA-tagged substrates were described *in vivo* and *in vitro* as substrates for the *Mtb* Clp system, and are recognized by both chaperones, ClpX and ClpC1 (Raju et al, 2012b; Schmitz & Sauer, 2014). Furthermore, several non-ssrA substrates for the *Mtb* Clp system were discovered (Figure 1.8 C). ClpX, for example, is involved in regulation of the cell division protein FtsZ (Dziedzic et al, 2010). Additionally, two transcriptional regulators, WhiB1 and CarD, were identified as Clp protease substrates (Raju et al, 2014). However, it was not tested which of the two chaperones, ClpX or ClpC1, recognizes the substrates for degradation. Another substrate, which involves ClpC1P1P2 in feedforward regulation, is the anti-sigma factor RseA (Figure 1.8 A and C). The degradation of RseA activates the sigma factor σ^E that in turn activates Clp expression via ClgR (Barik et al, 2010; Mehra & Kaushal, 2009).

1.6 The *Mtb* Clp system as a drug target

1.6.1 Mycobacterium tuberculosis

Mtb is member of the genus mycobacterium in the actinobacterial phylum. Actinobacteria are a diverse group of gram positive bacteria with a high GC content in their genome. Mycobacteria, even though they have the gram-positive cell wall setup, cannot be stained by the usual gram stain due to the unusual properties of their cell wall. It contains in addition to peptidoglycans a thick mycolic layer that gives them a hydrophobic, waxy cell surface that functions as an extra protective layer.

Mtb is a prominent pathogenic member of mycobacteria and the causative agent of tuberculosis, a disease that infects around one third of the worlds' population. It colonizes lung tissue and is spread via droplet infection. In the lung tissue, *Mtb* is taken up by macrophages and has evolved several ways to survive inside macrophages (Smith, 2003). During macrophage response, invading bacteria are taken up into phagosomes which mature to phagolysosomes in which the engulfed bacterium is usually treated with for example acid stress and reactive oxygen

and nitrogen species. *Mtb* can on the one hand prevent maturation of the phagosomes and has furthermore the ability to enter a dormant state inside macrophages where it cannot be accessed by antibiotic treatments (Flannagan et al, 2009).

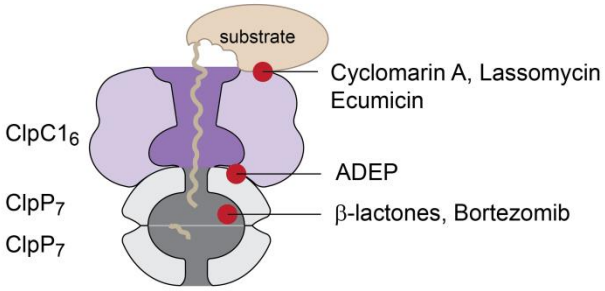


Figure 1.9: Antibacterial compounds act on the Clp chaperone-protease system in different places. Cartoon depiction of the ClpC1P1P2 complex (purple and grey). Antibacterial compounds act on different levels in the complex. Compounds like β-lactones and Bortezomib inhibit the complex by covalently modifying the active site of the protease. Acyldepsipeptides (ADEPs) bind into the hydrophobic patch on the protease and prevent chaperone binding. Additionally they open the axial pore of ClpP and thereby allow for unregulated substrate degradation. Cyclomarin A, Lassomycin and Ecumicin bind to the N-terminal domain of the ClpC1 chaperone and deregulate chaperone action either by overactivating degradation or decoupling protein unfolding from degradation.

There are several antibiotics that are used to treat *Mtb*, however due to the slow doubling time of *Mtb* (around 24 hours), and its ability to enter dormant states, drug treatment has to be performed over a long time frame and more and more drug-resistant strains are evolving, which necessitates the search for new drugs (WHO, 2015).

A group of drug targets that shows increasing promise are mycobacterial chaperone-protease systems (Raju et al, 2012a). Deregulation of essential, controlled proteolysis by overstimulation as well as inhibition of proteolytic complexes can lead to detrimental, bactericidal effects. Several promising compounds have been discovered in the last years targeting the Clp chaperone-protease system, that act either on the chaperone or the protease components, or in the interaction interface between them (Figure 1.9).

1.6.2 Antibacterial compounds acting against components of the *Mtb* Clp system

The three compounds Cyclomarin, Lassomycin and Ecumicin all act on the ClpC1 chaperone, and most likely all bind to its N-terminal domain. Cyclomarin A is a small cyclic peptide originally produced by marine *Streptomyces.* Biochemical and structural assays have shown that it binds to the ClpC1 N-terminal domain and it stimulates the degradation of a tagged GFP model-substrate *in vivo* (Schmitt et al, 2011; Vasudevan et al, 2013). Lassomycin was extracted from *Lentzea ken-*

tuckyensis and acts on ClpC1 by increasing its ATPase and probably substrate unfolding activity, while at the same abolishing protein degradation by uncoupling ATPase activity from cooperation with the ClpP1P2 proteolytic core (Gavrish et al, 2014). Mutational studies and *in silico* docking revealed the N-terminal domain of ClpC1 as the binding partner of Lassomycin (Gavrish et al, 2014). Ecumicin, a compound isolated from *Nonomurea*, probably acts in similar way as Lassomycin based on the increased ATPase activity of ClpC1 and abolished degradation of substrates by the ClpP1P2 protease observed *in vitro* (Gao et al, 2015; Gao et al, 2014)

Antibiotics that target the ClpP protease core on its interaction interface with the chaperone are acyldepsipeptides (ADEPs). ADEPs are natural peptide compounds that were initially found as natural products of *Steptomyces hawaiiensis* (Brotz-Oesterhelt et al, 2005). ADEPs bind to the hydrophobic patch on the ClpP surface, which is the binding site for the LGF-loop of the chaperone, and they thus compete for chaperone interaction. Furthermore, ADEP binding was shown to open the axial pore of the ClpP particle, resulting in increased degradation of unfolded and nascent proteins (Conlon et al, 2013; Kirstein et al, 2009a; Lee et al, 2010; Li et al, 2010). Activating degradation and decoupling it from regulation makes ADEP a powerful antibiotic compound, especially in bacteria where the Clp system is essential. In *Mtb,* ADEP action against the bacterium was found to be enhanced by efflux pump inhibitors (Ollinger et al, 2012). A crystal structure of ClpP1P2 in presence of ADEP showed binding of ADEP only to the hydrophobic patch of the ClpP2 subunit (Schmitz et al, 2014).

Inhibition of the core particle active sites is another way of disrupting Clp chaperone-protease function. For example β-lactones are compounds that inhibit ClpP protease activity (Bottcher & Sieber, 2008). They are suicide inhibitors that covalently bind to the active site serine, thereby inhibiting it, and are toxic to *Mtb* growth (Compton et al, 2013). Recently, the human proteasome inhibitor bortezomib was shown to likewise act on *Mtb* ClpP1P2 by covalently modifying the active site (Moreira et al, 2015).

The examples presented above show that the essential *Mtb* Clp system is a promising drug target, where interference on several levels (activity, assembly and possibly substrate recognition) causes severe dysregulation of the complex. A more detailed understanding of the mechanism and the interactors of the Clp chaperone-protease system could provide the basis for the development of further antibacterial compounds and eventually antitubercular drugs.

1.7 Aim of the thesis

The Clp system of *Mycobacterium tuberculosis* is an essential energy-dependent degradation system in this highly pathogenic organism. The mode of action of several recent drug compounds that affect correct assembly and cooperation of chaperone and protease components (Section 1.6) illustrates the importance of correct complex formation and substrate degradation. This thesis investigates the Clp chaperone-protease system by studying the assembly of its subcomplexes to an active protease machine and by assessing potential new interactors including known putative candidates and by employing an unbiased discovery screen.

The first part of this thesis focuses on the assembly of the ClpP1P2 protease core with its dedicated interaction partners, the chaperone components ClpX and ClpC1. In contrast to the well-studied *E. coli* Clp system, which features only one type of ClpP subunit, the *Mtb* protease core is composed of two different subunits (ClpP1 and ClpP2), resulting in an asymmetric hetero double-ring particle. Reports of the assembly on the ClpP1P2 protease core complex have been contradictory (Section 1.5), and no information is available on how the chaperones ClpX and ClpC1 interact with this asymmetric ClpP1P2 core, for example if ClpX or ClpC1 preferentially bind to the ClpP1 or the ClpP2 face of the double-ring particle. To assess this question, biochemical methods and characterization of single-point mutations in the interaction interface between the chaperone and protease are employed.

The second part of this thesis focuses on interactors of the chaperone ClpC1 and the putative adaptor protein ClpC2. The chaperone ClpC1 is composed of two AAA+ modules and an N-terminal domain, which is likely involved in substrate and adaptor protein interaction. ClpC2 on the other hand is a small protein that does not feature any domains of chaperone or protease function, but is homologous to the N-terminal interactor-binding domain of ClpC1 and therefore classified as a putative adaptor protein. So far, the known non-protease interactor spectrum of ClpC1 encompasses only the model substrate casein, ssrA-tagged substrates, and the anti-sigma factor RseA (section 1.5), while for ClpC2 no interactors have been described.

Here, I aim to find new interactors for the chaperone ClpC1 to expand a) the knowledge of its substrate clientele and b) to potentially identify adaptor proteins associating with the ClpC1 N-domains. As the Clp system is essential in *Mtb*, and as this is likely hinging on specific degradation substrates, widening the known substrate spectrum will give valuable information about the function of the Clp system in *Mtb*. To this aim, I will employ a bacterial two hybrid screening method, where the bait protein ClpC1 is screened against a library composed of around 4000

Mtb genes. The screen is performed in an *E. coli* strain deficient in adenylate cyclase. Bait and prey proteins are each fused to complementary inactive subdomains of *B. pertussis* adenylate cyclase that need to be brought into proximity by bait and prey interaction, before adenylate cyclase activity is restored and can be monitored. The same screen will also be performed using the putative adaptor protein ClpC2 as bait to find interactors of this small potential adaptor and thereby gain insights into its role in regulating Clp-dependent degradation.

Chapter 2 The *Mycobacterium tuberculosis* ClpP1P2 Protease Interacts Asymmetrically with its ATPase Partners ClpX and ClpC1

2.1 Introductory statement

This result chapter is published as and directly taken from:

Leodolter J, Warweg J, Weber-Ban E (2015). The *Mycobacterium tuberculosis* ClpP1P2 Protease Interacts Asymmetrically with Its ATPase Partners ClpX and ClpC1. PLoS ONE 10(5)

The first author (JL) perfomed all the experiments appearing in this publication, partially based on previous experiments of the second author (JW), and was responsible for the writing of the manuscript together with EWB.

2.2 Abstract

Clp chaperone-proteases are cylindrical complexes built from ATP-dependent chaperone rings that stack onto a proteolytic ClpP double-ring core to carry out substrate protein degradation. Interaction of the ClpP particle with the chaperone is mediated by an N-terminal loop and a hydrophobic surface patch on the ClpP ring surface. In contrast to *E. coli*, *Mycobacterium tuberculosis* harbors not only one but two ClpP protease subunits, ClpP1 and ClpP2, and a homo-heptameric ring of each assembles to form the ClpP1P2 double-ring core. Consequently, this hetero double-ring presents two different potential binding surfaces for the interaction with the chaperones ClpX and ClpC1. To investigate whether ClpX or ClpC1 might preferentially interact with one or the other double-ring face, we mutated the hydrophobic chaperone-interaction patch on either ClpP1 or ClpP2, generating ClpP1P2 particles that are defective in one of the two binding patches and thereby in their ability to interact with their chaperone partners. Using chaperone-mediated degradation of ssrA-tagged model substrates, we show that both *Mycobacterium tuberculosis* Clp chaperones require the intact interaction face of ClpP2 to support degradation, resulting in an asymmetric complex where chaperones only bind to the ClpP2 side of the proteolytic core. This sets the Clp proteases of *Mycobacterium tuberculosis,* and probably other Actinobacteria, apart from the well-studied *E. coli* system, where chaperones bind to both sides

of the protease core, and it frees the ClpP1 interaction interface for putative new binding partners.

2.3 Introduction

Mycobacterium tuberculosis (*Mtb*) is a gram-positive bacterium of the phylum Actinobacteria and the causative agent of tuberculosis. More and more strains are evolving resistance to the antibiotics currently in use (Gandhi et al; Jassal & Bishai, 2009), but several new and promising compounds have been discovered in recent years (acyldepsipeptides (Brotz-Oesterhelt et al, 2005; Kirstein et al, 2009a; Ollinger et al, 2012), β-lactones (Bottcher & Sieber, 2008), cyclomarin A (Schmitt et al, 2011), lassomycin (Gavrish et al, 2014)), all targeting the Clp (caseinolytic protease) chaperone-protease system. The Clp chaperone-protease is a bacterial multi-subunit protein complex involved in intracellular protein degradation. It is active in general protein quality control as well as specific degradation of proteins participating in regulatory processes (Baker & Sauer, 2012; Gur et al, 2013; Laederach et al, 2014). The best described substrate class comprises proteins tagged with the ssrA tag, a short peptide sequence C-terminally added to proteins by the tmRNA system to rescue stalled ribosomes (Gottesman et al, 1998; Moore & Sauer, 2007). To mediate substrate degradation, the Clp chaperone-proteases form cylindrical complexes built from rings of protease and chaperone subunits stacked on top of one another. The core of the structure consists of the ClpP proteolytic subunits that assemble into a double-ring stack of two heptameric rings enclosing a sequestered space (Wang et al, 1997). The protease active sites line the inside of this chamber and are made up of the Ser-His-Asp catalytic triad typical for serine proteases. The access to the chamber is controlled by hexameric unfoldases (ClpX, ClpA or ClpC) of the AAA+ type (**A**TPase **a**ssociated with various cellular **a**ctivities) that recognize substrates, unfold them in an ATP-dependent manner and thread them into the proteolytic chamber (Sauer & Baker, 2011; Snider & Houry, 2008). Two conserved interaction elements on the protease particle are involved in the association with the chaperones. One interaction feature is an N-terminal loop positioned at the axial pore (N-loop), where substrates pass from the chaperone to the protease (Bewley et al, 2006; Gribun et al, 2005; Kang et al, 2004). The other interaction feature is a hydrophobic patch located on the face of the protease ring to which binds a loop of the chaperone containing a conserved LGF-motif in case of the *Mtb* chaperones ClpX and ClpC1 (LGF-loop) (Joshi et al, 2004; Kim et al, 2001; Wang et al, 1997). Interestingly, antibiotics of the acyldepsipeptide (ADEP) class bind to this patch in place of the LGF-loop and deregulate the Clp protease by mimicking chaperone binding (Lee et al, 2010; Li et al, 2010; Schmitz et al, 2014).

In contrast to the well-studied Clp system of *E. coli,* Actinobacteria contain not only one ClpP subunit, but carry two or more homologous genes. The *Mtb* system harbors two ClpP genes, *clpP1* and *clpP2*, which are co-expressed from one operon (Personne et al, 2013). Both subunits are essential (Griffin et al, 2011; Sassetti et al, 2003) and both are necessary for the activity of the complex (Raju et al, 2012b). Like other ClpP protease subunits, *Mtb* ClpP1 and ClpP2 carry N-terminal propeptides that are cleaved off in a processing step, as has been demonstrated *in vivo* (Akopian et al, 2012; Benaroudj et al, 2011). To produce the mature, functional ClpP particle, ClpP1 and ClpP2 assemble from one homoheptameric ring of each subunit into a ClpP1P2 hetero double-ring. So far, this assembly has only been observed *in vitro* in the presence of a synthetic non-natural activator peptide (Akopian et al, 2012; Schmitz et al, 2014). This activator peptide is an N-blocked dipeptide, usually Z-Leu-Leu (Benzyloxycarbonyl-L-Leucyl-L-Leucine), Z-Leu-Leu-H (Benzyloxycarbonyl-L-Leucyl-L-Leucinal) or a similar molecule, that binds near the active sites of the proteolytic particle and stabilizes the active conformation of the ClpP1P2 double-ring (Akopian et al, 2012; Schmitz et al, 2014). This functional conformation of the complex is also stabilized by the presence of the *Mtb* chaperone ClpX and protein substrate, acting synergistically with the activator peptide (Schmitz & Sauer, 2014). ClpC1 also appears to have a stabilizing effect, although here ClpC1 of *M. smegmatis* was used in combination with *Mtb* ClpP1P2 (Schmitz & Sauer, 2014). Interestingly, *E. coli* ClpX rings (EcClpX) an interact with the *Mtb* ClpP1P2 complex and even promote substrate degradation more than ten-fold faster compared to the *Mtb* chaperones (Schmitz & Sauer, 2014).

In the *E. coli* ClpAP and ClpXP complexes, chaperone binding to both sides of the ClpP cylinder was shown to be the preferred state (Kessel et al, 1995; Maglica et al, 2009). However, the *E. coli* ClpP double-ring is a symmetric particle made of 14 identical subunits resulting in two equal binding surfaces, a situation fundamentally different from that in the *Mtb* ClpP1P2 double-ring, which presents a different binding surface on each face of the cylinder. The *Mtb* ClpP1P2 double-ring structure shows considerable differences between its two potential chaperone-interaction faces (Schmitz et al, 2014) which could indicate that specific chaperone binding to one face or the other occurs.

Here, we investigate the effects of the asymmetry of the *Mtb* ClpP1P2 hetero double-ring particle on its assembly, propeptide processing and interaction with the ATP-dependent unfoldases ClpX and ClpC1 of *Mtb*. We show that assembly of ClpP1P2 is independent of the presence of the propeptide and we provide *in vitro* evidence that an asymmetric behavior is already apparent during the propeptide processing reaction, with ClpP1 as the main carrier of the processing ac-

tivity. To investigate whether ClpP1 and ClpP2 exhibit differences in interaction with the chaperones ClpX and ClpC1, we designed variants of both proteins impaired in chaperone interaction by mutating the hydrophobic patch necessary for LGF-loop binding. Our results show that both chaperones need the intact interaction face of ClpP2 to support chaperone-mediated substrate degradation, suggesting that the ClpP2 ring face is the sole interaction platform for these chaperones.

2.4 Materials and Methods

Alignment

Protein sequences of M. tuberculosis (*Mtb*) ClpP1 (P9WPC5), *Mtb* ClpP2 (P9WPC3) and their actinobacterial homologues, as well as *E. coli* ClpP (P0A6G7), were extracted from the Uniprot database (http://www.uniprot.org/) and aligned using the Clustal Omega algorithm (Goujon et al, 2010; Sievers et al, 2011). Visualisation was performed using Jalview (Waterhouse et al, 2009).

Cloning, expression and protein purification

All genes were amplified by PCR with Phusion DNA polymerase (New England Biolabs) from *M. tuberculosis* H37Rv genomic DNA. N-terminal deletions and active-site mutations of *clpP1* and *clpP2* were introduced by site-directed mutagenesis.

ClpP1 and clpP2, containing a C-terminal His_4-tag, were separately ligated into the pET-Duet-1 coexpression plasmid. ClpX was fused to a C-terminal Tobacco Etch Virus (TEV) endopeptidase cleavage site followed by GFP-His6 in a pET24 vector. ClpC1 was ligated into a p7XC3H FX vector (Geertsma & Dutzler, 2011), including a stop codon to produce the untagged protein. The malate dehydrogenase (*mdh*) gene with the *Mtb* ssrA-tag (AADSHQRDYALAA) added C-terminally was ligated into a pET20 vector. GFP-ssrA was recloned from the construct used in (Kress et al, 2009b), with the *E. coli* ssrA-tag changed to the *Mtb* ssrA-tag sequence. Correct insertion and sequences of the genes were verified by sequencing.

All constructs were transformed into *E. coli* Rosetta (DE3) (Invitrogen) and grown in LB broth supplemented with the respective antibiotic. Expression was induced at an OD_{600} of 0.8 by addition of 0.1 mM IPTG and expression was carried out overnight at 20°C. Cells were cracked by sonication or with a microfluidizer. The ClpP1 and ClpP2 proteins and variants were purified by Ni-NTA affinity chromatography and subsequent gel filtration on a Superose 6 column (GE Healthcare) in Buffer A (50 mM Hepes-NaOH pH 7.5, 300 mM NaCl, 10% glycerol). ClpX and

MDH-ssrA were purified by Ni-NTA affinity chromatography, TEV cleavage and a reverse Ni-NTA step, followed by dialysis into Buffer A. ClpC1 was purified on an anion exchange Fast Flow Q column, followed by ammonium sulfate (AS) precipitation and a Superdex 200 (GE Healthcare) gel filtration step. The final fractions were again precipitated by AS, then resuspended and dialysed into Buffer A. GFP-ssrA, with either the *Mtb* or *E. coli* ssrA-tag was purified as described (Kress et al, 2009b). EcClpX was expressed and purified as described (Maglica et al, 2009). ClpX and ClpC1 were further dialysed into Buffer J (50 mM Hepes-KOH pH 7.5, 150 mM KCl, 15% glycerol, 20 mM MgCl2). Correct molecular mass of the final proteins was verified by mass spectrometry.

Analytical gel filtration

Analytical gel filtrations at room temperature were performed on a Superdex 200 10/300 GL column (GE Healthcare, 24 ml) in Buffer A at a flow rate of 1 ml/min on an ÄKTA Purifier System. Prior to loading, the sample (25 µM ClpP1 and/or ClpP2 protomer) was centrifuged in an Eppendorf table-top centrifuge for 10 minutes at 13'000 rpm. 100 µl of the sample were injected onto the column. Proteins were detected by absorption at 280 nm. For analytical gel filtration runs at 4°C the same Superdex 200 column was used in Buffer A with a flowrate of 0.6 ml/min on a different ÄKTA Purifier. The column was calibrated at room temperature and at 4°C with the same set of standard proteins (GE Healthcare Gel filtration Calibration Kit), and a calibration curve was calculated for each.

Processing of the N-terminal ClpP propeptides

For processing assays of the N-terminal propeptides in the presence of the activator Z-Leu-Leu-H (Benzyloxycarbonyl-L-Leucyl-L-Leucinal) (PeptaNova), 25 µM full-length ClpP1 and ClpP2 protomer each (referred to as proClpP1 and proClpP2) were first incubated for 1 hour at room temperature in Buffer A without the activator. Then the processing reaction was started by the addition of 1 mM activator dissolved in DMSO, resulting in a final DMSO concentration of 2%. Alternatively, for processing in presence of 1 µM ClpC1 hexamer and 10 µM GFP-ssrA or 1 µM ClpX hexamer and 10 µM MDH-ssrA, the reaction was performed in Buffer J, pH 7 or 7.5, respectively, supplemented with 5 mM ATP, 1 mM DTT, 40 mM phosphocreatine and 1 U/ml creatine phosphokinase, with 0.5 µM preassembled proClpP1P2 double-ring. The reaction was stopped at the indicated time points by the addition of Laemmli buffer and the samples were heated for 10 minutes at 95°C. The samples were analyzed on a 15% SDS-PA gel. For the production of mature ClpP1P2 for use in further biochemical assays, proClpP1 and proClpP2 were incubated at a concentration of 50 µM (for wild-type) or 70 µM (for hydrophobic patch variants) protomer each in

Buffer A overnight at room temperature in the presence of 1 mM activator. Processing and assembly were verified by SDS-PAGE and analytical gel filtration. The mature ClpP1P2 complex was separated from the activator using a PD-10 desalting column (GE Healthcare).

Degradation of MDH-ssrA by ClpXP1P2

Protein degradation assays using MDH-ssrA as a model substrate were performed at room temperature in Buffer J with 1 µM ClpX hexamer, 0.5 µM mature ClpP1P2 double-ring complex preassembled for 1 hour at room temperature, 2 µM MDH-ssrA, 5 mM ATP, 1 mM DTT, 20 mM phosphocreatine and 1 U/ml creatine phosphokinase. The reaction was stopped at the indicated time points by the addition of Laemmli buffer and the samples were heated for 10 minutes at 95°C. The samples were analyzed on a 15% SDS-PA gel.

Degradation of GFP-ssrA by ClpC1P1P2 and EcClpXP1P2

Protein degradation assays using GFP-ssrA as a model substrate were carried out by following the loss of the GFP fluorescence signal. The experiments were performed in a BioTek Synergy 2 plate reader with a Tungsten light source, with an excitation wavelength of 360/40 nm and emission wavelength of 528/20 nm (50% optics position, sensitivity: 90) in Corning non-binding 96-well half area assay plates in 50 µl reaction volume. Reaction conditions were the same as given for the degradation of MDH-ssrA, but using ClpC1 and GFP-ssrA (*Mtb* ssrA sequence) instead of ClpX and MDH-ssrA, respectively. For the reaction with ClpC1, the pH of the reaction buffer was adjusted to 7. Before and after the reaction, samples were drawn for SDS-PAGE analysis. Reaction buffer conditions for the degradation of GFP-ssrA (*E. coli* ssrA sequence) by EcClpX and ClpP1P2 were the same as for *Mtb* ClpXP1P2.

Negative stain Electron Microscopy

1.4 µM proClpP1 protomer (200 nM heptamer) sample was applied to a carbon-coated copper mesh grid for 20 seconds and subsequently stained with 2% aqueous uranyl acetate for 2 minutes. Imaging was perfomed in a FEI Morgagni 268 transmission electron microscope operating at 100 kV.

2.5 Results

2.5.1 The ClpP1P2 double-ring complex can be assembled without activator peptide and is processed in a chaperone-dependent manner

The mycobacterial ClpP protease subunits ClpP1 and ClpP2 were previously shown to each form a homo-heptameric ring that assembles with the other into the proteolytically active ClpP1P2

double-ring complex in the presence of synthetic activator peptides (Akopian et al, 2012; Schmitz et al, 2014). To investigate the assembly behavior of ClpP1 and ClpP2 into the double-ring particle in absence of any activator peptide, we carried out analytical gel filtration analysis, both at room temperature and at 4°C. Protease subunits with the propeptides still present were used (proClpP1 and proClpP2), which corresponds most closely to the situation encountered after the subunits are first translated in the cell and are coming together in the initial assembly of the ClpP1P2 complex. At room temperature, recombinant proClpP1 and proClpP2 assemble into the double ring complex in absence of the activator, as shown in analytical gel filtration by an elution shift of the peaks of the single proClpP1 and proClpP2 rings (Figure 2.1 A, dashed light and dark grey traces) to the peak of the double-ring complex at ~11 ml (Figure 2.1 A, dark brown trace). Based on a calibration curve using molecular weight standards this elution volume corresponds to ~300 kDa, which is equivalent to the size of the assembled double-ring particle. The *E. coli* ClpP double-ring elutes at this position (EcClpP, indicated by a grey arrow tip), further supporting a double-ring assembly state. In addition, SDS-PAGE analysis confirms that fractions collected from this elution peak contain equimolar amounts of ClpP1 and ClpP2 (Figure 2.1 A, gel slice). When the two ClpP subunits are run separately, they elute at positions corresponding to lower molecular weight than the double-ring complex. However, while ClpP2 elutes roughly at the expected position of a single ring, ClpP1 elutes even later despite the similar subunit size. To ensure that ClpP1 nevertheless forms rings on its own, we analyzed the ClpP1 sample by negative stain electron microscopy. Top views of ring-shaped complexes with a stain-filled center are clearly visible, indicating that ClpP1 is assembled (Figure S 1). The lower elution volume could be due to unspecific interactions with the column material, potentially involving the propeptide, since the mature ClpP1 (mClpP1) runs at the expected elution volume close to ClpP2 (Figure 2.1 B). It could also indicate a tendency of a portion of the single rings to dissociate into smaller oligomeric states (Benaroudj et al, 2011).

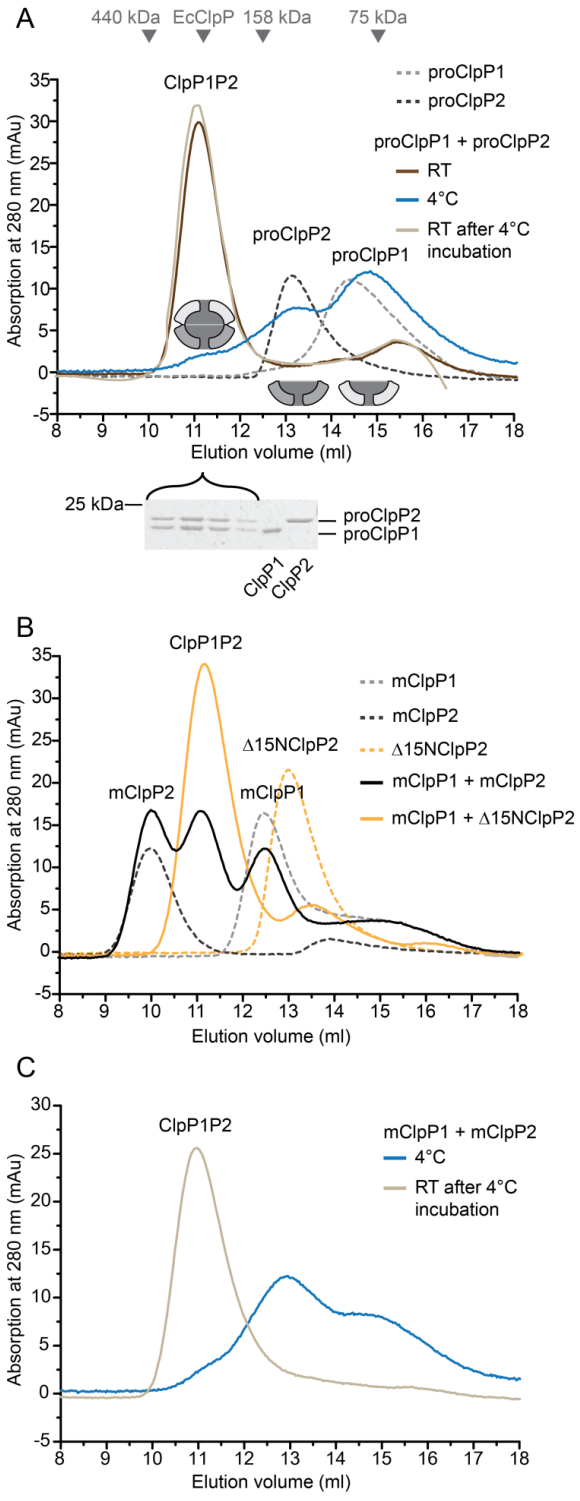


Figure 2.1: Assembly of ClpP1 and ClpP2 to the ClpP1P2 double-ring particle. Analytical gel filtration runs were performed on a Superdex 200 gel filtration column (24 ml). Unless noted otherwise, all runs were performed at room temperature. The assembly state of the proteins is indicated by a cartoon depiction in the topmost graph of either single rings in light grey for ClpP1 and dark grey for ClpP2, or the assembled double-ring particle. Molecular size markers are indicated as arrow tips above the elution profiles in the first graph. A marker for E. coli ClpP (EcClpP) is included as a reference for the double-ring assembly. The concentration of ClpP1 and ClpP2 is always 25 µM protomer. *Continued next page*

Figure 2.1 (continued): A. ClpP1 and ClpP2 containing the N-terminal propeptide (proClpP1, proClpP2) assemble to double-ring complexes in the absence of the activator. When proClpP1 (dashed light grey) and proClpP2 (dashed dark grey) are incubated together at room temperature overnight they assemble into the double-ring complex (dark brown). SDS-PAGE shows a 1:1 ratio of ClpP1 to ClpP2 in the peak fraction. Incubation of the assembled complex at 4°C for 3 hours leads to disassembly (blue). Reincubation of the complex at room temperature for 3 hours leads to reassembly into the double-ring complex (light brown). **B**. Analytical gel filtration of individually loaded mature ClpP1 (mClpP1, dashed light grey), mature ClpP2 (mClpP2, dashed dark grey) and Δ15NClpP2 (dashed orange). mClpP2 shows its main elution peak at 10 ml corresponding to a size of around 450 kDa. Elution profiles were also recorded after overnight incubation of mClpP1 and mClpP2 (black), and of mClpP1 and Δ15NClpP2 (orange). **C.** Elution profiles of *in vitro* processed mClpP1P2. To produce the mature complex proClpP1 and proClpP2 were incubated in the presence of 1 mM activator overnight at room temperature. The activator was then removed by buffer exchange and the mClpP1P2 complex was incubated at 4°C for 3 hours leading to disassembly of the complex (blue). Reincubation of the complex at room temperature for 3 hours leads to reassembly into the double-ring complex (light brown).

After incubation of the assembled complex at 4°C, ClpP1P2 double-rings disassemble into single rings as shown by subsequent gel filtration at 4°C (Figure 2.1 A, blue trace). The disassembly is reversible, since reincubation at room temperature and subsequent analysis by gel filtration shows a single main elution peak at thse double-ring particle position (Figure 2.1 A, light brown trace). This shows that the assembly of the ClpP1 and ClpP2 single rings to the ClpP1P2 double-ring particle occurs also in absence of the activator. It is, however, temperature-dependent and does not occur quantitatively at low temperature.

In the functional ClpP1P2 complex the propeptides are processed to form the mature particle (Akopian et al, 2012; Benaroudj et al, 2011). Previous studies had identified the processing site to be located between Ala12 and Arg13 for the *Mtb* ClpP2 subunit (Akopian et al, 2012; Benaroudj et al, 2011). For the propeptide cleavage in ClpP1, two different cleavage sites were reported by two different studies both expressing ClpP1 in absence of ClpP2, namely between Arg8 and Ser9 (Benaroudj et al, 2011) or between Asp6 and Met7 (Akopian et al, 2012). Using *in vitro* processing of the propeptide-containing ClpP1P2 double-ring complex in presence of activator peptide, we predominantly obtained cleavage between Met7 and Arg8 (Figure S 2). Therefore, to test whether the mature particle can be generated *in vitro* by mixing recombinantly produced ClpP subunits lacking the propeptide and to test if the propeptides might be important for assembly, we produced mature ClpP1 and ClpP2 (mClpP1, mClpP2), where the residues corresponding to the propeptides (the first 6 and 12 residues, respectively) were removed. Methionine 7 was kept in the mature ClpP1 construct, because it forms the translation start site for expression. Interestingly, while mClpP1 eluted at a position corresponding to a size of ~150 kDa and thus one heptameric ring (Figure 2.1 B, dashed light grey trace), mClpP2 eluted much earlier, in fact even earlier than the double-ring particle (Figure 2.1 B, dashed dark grey trace), translat-

ing into an apparent molecular mass of ~450 kDa. This could correspond to a non-native assembly state of three stacked heptameric rings, a behavior observed previously by Benaroudj et al (Benaroudj et al, 2011). Likely due to this non-native stacking, only a fraction of mixed mClpP1 and mClpP2 assembles into ClpP1P2 double-ring particles (Figure 2.1 B, black trace).

To circumvent the aggregation tendencies of mClpP2 and still be able to test the role of the propeptides in assembly, we generated a variant of mClpP2 that was shortened by another three residues beyond the processing site (Δ15NClpP2). This variant is soluble and runs at the position expected for the ClpP2 single ring (Figure 2.1 B, dashed orange trace). Upon incubation with mClpP1, the double-ring particle is formed with the expected elution properties (Figure 2.1 B, orange trace), indicating that the propeptides do not play an active role in ClpP1P2 particle assembly.

The temperature dependence of the mature double-ring complex resembles that of the propeptide-containing complex. When mature ClpP1P2 double-ring particle is generated by addition of activator peptide (see next paragraph for processing) and is then incubated at 4°C, it also disassembles into single rings and reassembles at room temperature (Figure 2.1 C, blue and light brown trace).

To generate the mature ClpP1P2 complex from proClpP1 and proClpP2, the propeptides at their N-termini have to be cleaved off. Formation of the double-ring is a requirement for propeptide processing to occur, since the individual subunits alone, even in presence of activator peptide, show no peptidase activity (Akopian et al, 2012). Processing of the propeptides is the first activity performed by the newly formed ClpP1P2 complex. To investigate this activity and to assess a potential asymmetry across the ClpP1:ClpP2 ring-ring interface in the propeptide processing reaction, we analyzed the propeptide cleavage activity in the assembled double-ring particle. To follow the processing of individual subunits in the wild-type particle, we first incubated proClpP1 and proClpP2 together for one hour at room temperature resulting in the assembly of the unprocessed ClpP1P2 complex. Processing was allowed to occur during overnight incubation either in presence or absence of the activator peptide Z-Leu-Leu-H, from here on referred to as activator. For the reaction in absence of the activator, only an end-point sample was drawn after overnight incubation (Figure 2.2 A, last lane). For the sample in presence of the activator, the progress of the cleavage reaction was followed at specific time intervals by drawing samples that were then analyzed by SDS-PAGE (Figure 2.2 A). Unfortunately, the generated mClpP2 overlays with the band of proClpP1 and is then not visible as a separate band. Therefore, processing of

ClpP2 is followed by assessing the disappearance of the proClpP2 band and processing of ClpP1 is followed by assessing the appearance of the mClpP1 band. After the first 20 minutes of processing, the band for proClpP2 is considerably reduced while on the other hand hardly any mature ClpP1 has yet been produced. This suggests that the ClpP2 propeptide is processed first. While we showed that assembly to the double-ring particle occurs in absence of the activator, this experiment demonstrates that maturation to the processed complex requires the presence of the activator (Figure 2.2 A). As the synthetic activator is not present *in vivo*, this raises the question which endogenous molecules could serve this function. Inside the cell, ClpP particles form assemblies with their chaperone partners. To test whether processing could occur in the absence of the activator as long as a chaperone partner and a protein substrate are present, we incubated proClpP1 and proClpP2 in the presence of the chaperone ClpC1 and the model substrate GFP-ssrA, or the chaperone ClpX and the model substrate MDH-ssrA. In Figure 2.2 B we show that, indeed, processing of proClpP1 and proClpP2 to mClpP1 and mClpP2 occurs when the chaperone ClpX or ClpC1 and a degradation substrate is present. The processing in this case is dependent on ATP hydrolysis, suggesting that substrate must enter the proteolytic particle. Therefore, *in vivo*, the natural interaction partners of the ClpP particle likely serve to activate propeptide processing. As processing in presence of the activator provided a cleaner experimental setup (Figure 2.2 B, +Act), we made use of the activator as a tool for further *in vitro* processing assays.

Given that propeptide cleavage occurs only in the assembled double-ring particle, processing of an excess of one subunit over the other would require dissociation followed by reassociation with the unprocessed subunits. Therefore, the degree of processing of excess ClpP subunits can be used as a measure for the kinetic stability of the complex. When the two subunits are provided in equimolar amounts, overnight incubation with activator results in complete processing of both subunits (Figure 2.2 C, first two lanes). Providing proClpP1 in three-fold excess under the same conditions results in processing of only a fraction of proClpP1 to mClpP1 (Figure 2.2 C, lanes 3 and 4). The unfortunate overlap between the bands of mClpP2 and proClpP1 makes it difficult to assess the exact amount of proClpP1 left after overnight incubation. However, it is clear that the mClpP1 band would have to be three times more intense than the mClpP2/proClpP1 band, if complete processing occurred. This is not the case. Likewise, with proClpP2 in three-fold excess, surplus proClpP2 is not processed and only equimolar bands for mClpP1 and mClpP2 are observed (Figure 2.2 C, lanes 5 and 6). These results demonstrate that the processed ClpP1P2 double-ring particle is kinetically very stable.

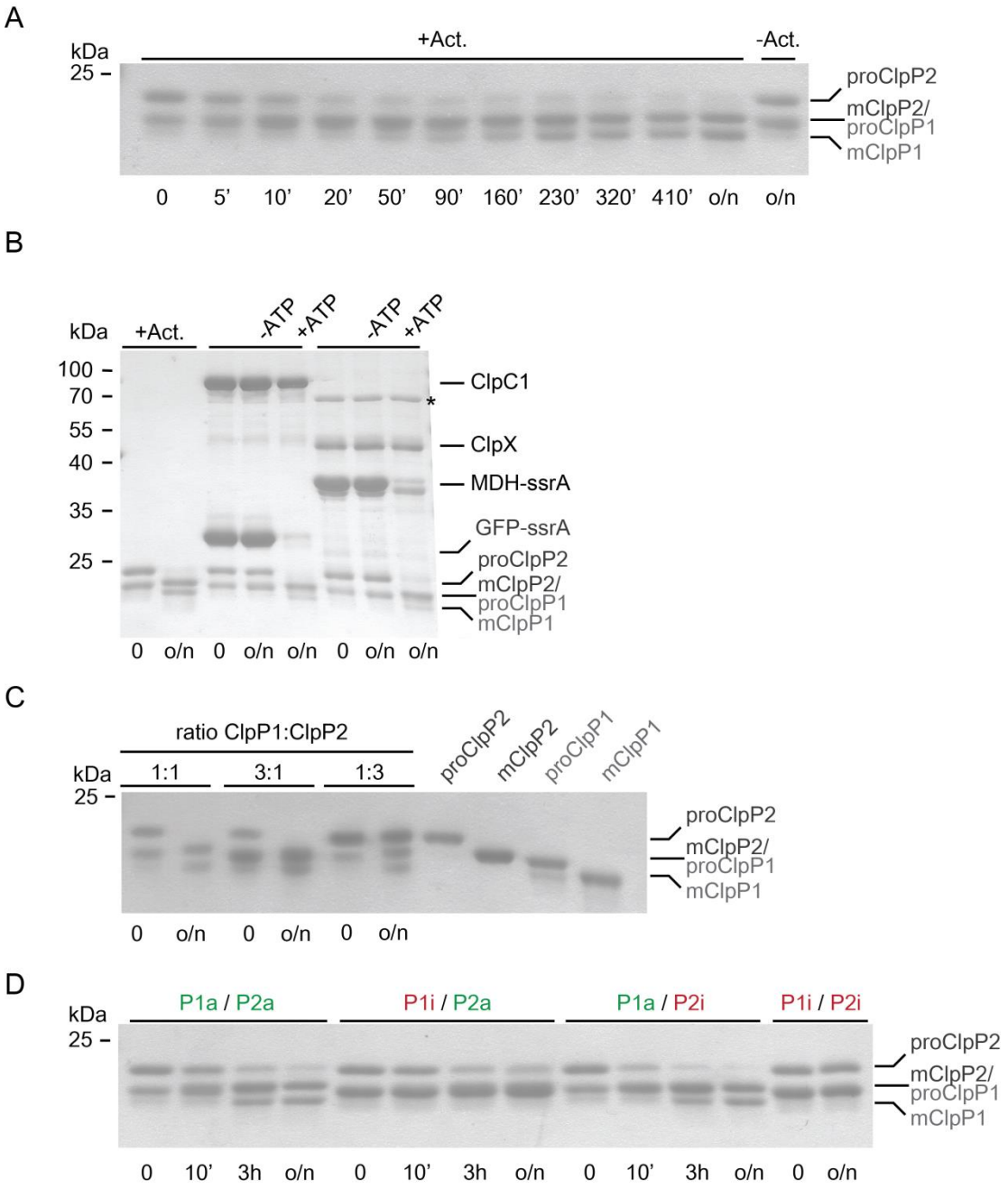


Figure 2.2: Propeptide processing of ClpP1 and ClpP2 in the ClpP1P2 double-ring assembly. Maturation of double-ring assembled ClpP1 and ClpP2 carrying the N-terminal propeptide (proClpP1, proClp2) to the mature proteins (mClpP1, mClpP2) was performed at room temperature and followed by SDS-PAGE analysis. Samples were taken at the indicated time points. A. Propeptide processing of ClpP1 and ClpP2 (25 µM protomer each), in the presence of 1 mM activator (+Act.), or without activator (−Act.). B. Propeptide processing of ClpP1 and ClpP2 (0.5 µM double-ring particle) in the absence of activator and the presence of 1 µM ClpC1 hexamer, 10 µM GFP-ssrA or 1 µM ClpX hexamer and 10 µM MDH-ssrA, and 5mM ATP (+ATP) or without ATP (−ATP). For comparison, overnight processing of ClpP1 and ClpP2 in the presence of 1 mM activator is shown. The asterisk (*) indicates a protein contaminant that was identified by mass spectrometry to contain DnaK from E. coli. C. Propeptide processing of ClpP1 and ClpP2 at different molar ratios of ClpP1:ClpP2 with 25 and 75 µM, respectively, ClpP1 or ClpP2 protomer concentration in the presence of 1 mM activator. D. Propeptide processing of double-ring complexes assembled from wild-type (active) ClpP1 and ClpP2 (P1a, P2a, 25 µM protomer each) and protease-inactive ClpP1S98A, ClpP2S110A (P1i, P2i) in the presence of 1 mM activator.

2.5.2 ClpP1 is the main actor in ClpP1P2 double-ring processing

For the ClpP particle of *Streptomyces lividans*, an organism that is related to *Mtb* and also encodes ClpP1 and ClpP2 subunits, it was suggested based on *in vivo* experiments that processing of the propeptides occurs across the ClpP1P2 double-ring interface (Viala & Mazodier, 2002). For *Mtb* ClpP1 and ClpP2, processing was observed for either subunit expressed on its own in M. smegmatis (Akopian et al, 2012). When both proteins were separately expressed in *E. coli* and purified, only active ClpP1 showed a double band indicative of processing, while ClpP2 only showed a double band when active ClpP1 but not when inactive ClpP1 was coexpressed (Benaroudj et al, 2011). The reports are partly contradictory and in both cases processing activity was assayed during expression in heterologous hosts where interference from endogenous proteins could occur. Furthermore, the latter study stated that ClpP1P2 tetradecamers containing both ClpP1 and ClpP2 are not formed and only partial processing was observed (Benaroudj et al, 2011). Therefore, to investigate this in the background of the functionally relevant hetero oligomer and to exclude possible host interference, we produced mixed active and inactive ClpP1P2 complexes that were then tested for propeptide processing in an *in vitro* setup. Four different double-ring particles were tested, namely particles with both ClpP rings active (P1a/P2a) or both inactive (P1i/P2i) with the active-site serine mutated to alanine (ClpP1S98A, ClpP2S110A), or mixed particles, where either the ClpP1 or the ClpP2 subunit was in the inactive (i) and the other in the active (a) form (P1i/P2a, P1a/P2i). Processing was initiated by adding activator and the reaction was followed by SDS-PAGE analysis (Figure 2.2 D). As expected, when both ClpP1 and ClpP2 subunits are active (P1a/P2a), the fully mature particle with mClpP1 and mClpP2 is produced, while inactivation of both ClpP1 and ClpP2 (P1i/P2i) abolishes processing entirely. When a catalytically active ClpP1 ring was combined with a catalytically inactive ClpP2 ring (P1a/P2i), processing of both rings was complete and even occurred somewhat faster than with wt complex. However, when only ClpP2 was active (P1i/P2a), processing was significantly impaired. Only a fraction of ClpP2 was processed after overnight incubation and no ClpP1 processing was observed. This indicates that while ClpP1 in context of the ClpP1P2 particle can process both itself and ClpP2, ClpP2 is unable to process the trans-ring. Furthermore ClpP2 exhibits little activity even towards its own propeptide, suggesting that ClpP1 is the main propeptide processor of the ClpP1P2 double-ring particle.

2.5.3 ClpP2 is the main interaction platform for the ATPase rings

ClpP1 and ClpP2 show specialization in terms of their processing activity (i.e. ClpP1 is almost solely responsible for processing), and they were previously shown to exhibit different substrate cleavage specificities (Akopian et al, 2012; Compton et al, 2013). However, in the case of substrate degradation, this specialization is not apparent, as protease inactivation of either subunit did not significantly slow down *in vitro* protein degradation (in case of ClpX-dependent degradation), or even enhanced its rate (in case of the heterologous MsmClpC1*Mtb*ClpP1P2 complex) (Schmitz & Sauer, 2014). It is unlikely that the only asymmetric behavior in ClpP1 and ClpP2 would be in their processing activity, as differences between the two molecules are not only apparent in the active-site substrate binding cleft, but also in the N-terminal loop region as well as the hydrophobic patch, both of which were shown in other bacterial ClpP particles to be involved in chaperone interaction (Gribun et al, 2005; Schmitz et al, 2014; Wang et al, 1997). We were interested to find out whether the asymmetry in the ClpP1P2 protease particle also translates into an asymmetry of chaperone binding. Do both chaperones, ClpX and ClpC1, bind to ClpP1 and ClpP2, or is the interaction of one chaperone restricted to one protease partner, e.g. binding of ClpX to ClpP2 or to ClpP1 only?

To answer this question we aimed to create variants of ClpP1 and ClpP2 with impaired chaperone binding. One of the two important interaction features of the Clp proteases is a hydrophobic patch on the ClpP cylinder face, responsible for binding a loop on the chaperone containing a conserved LGF motif (Figure 2.3 A). These hydrophobic patches are located in clefts on the apical surface of the ClpP particle and are formed by residues of two adjacent subunits (Figure 2.3 B). To impair chaperone binding to the protease cylinder, we mutated the hydrophobic patch and termed the resultant proteins hydrophobic patch (hp) variants (hpClpP1and hpClpP2) in contrast to wild-type (wt) ClpP1 and ClpP2 (Figure 2.3 C, impaired ClpP faces are denoted with red crosses). Conserved residues located in this patch have been described for EcClpP, namely Y74, Y76 and F96 (Wang et al, 1997). To identify the corresponding residues in ClpP1 and ClpP2 we aligned the protein sequences of the three proteins (Figure 2.3 D, red arrows) and also ascertained that the selected residues are located on the surface of the *Mtb* ClpP1P2 particle as judged from the published crystal structure (Figure 2.3 E) (Schmitz et al, 2014). To generate hpClpP1, four residues on the ClpP1 ring face were mutated: S61A, Y63V, L83A, Y91V. To generate hpClpP2, two residues were mutated: Y75V, Y95V (Figure 2.3 C). Expression and purification of either variant resulted in a stable protein preparation.

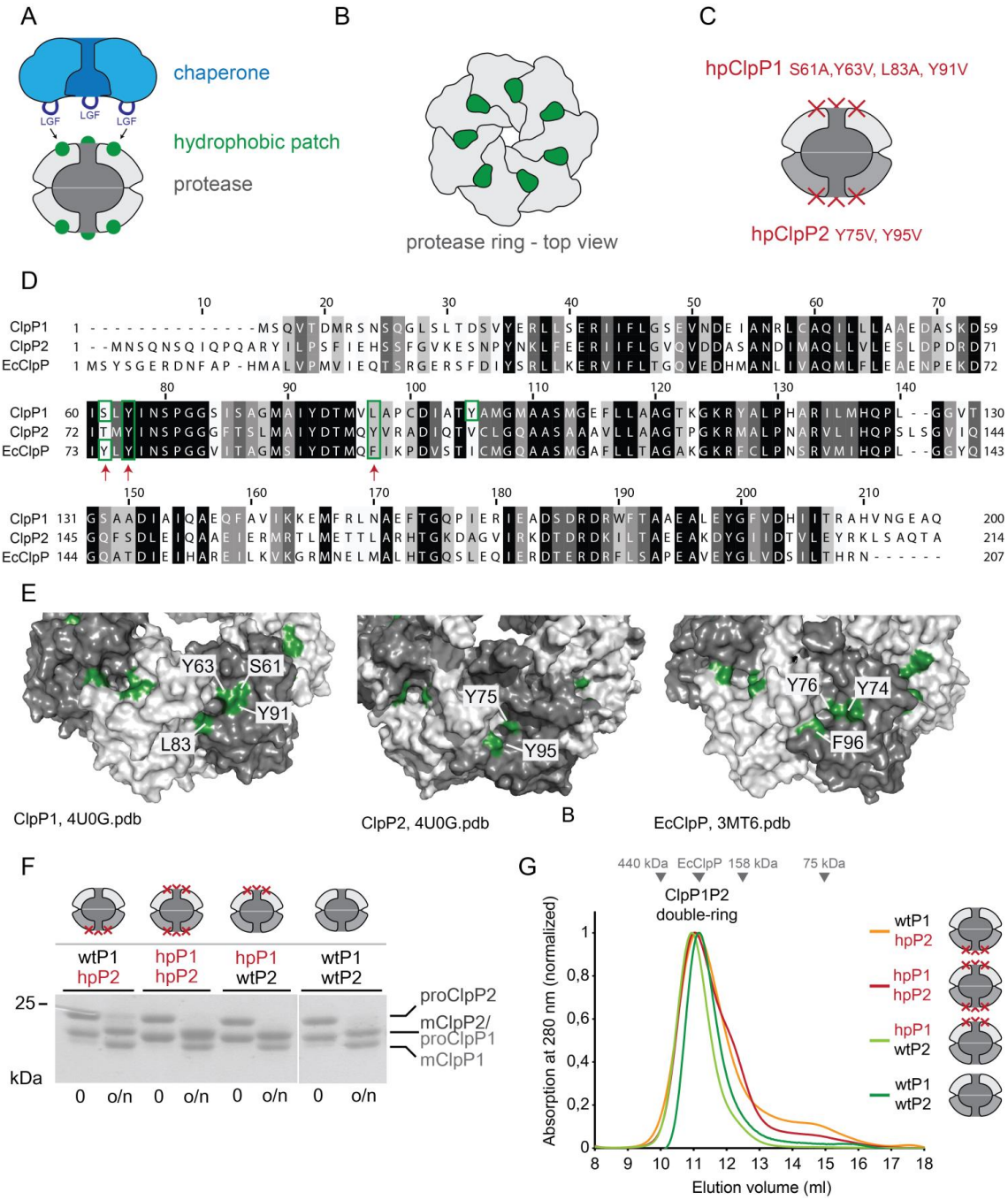


Figure 2.3: Generation of hydrophobic patch variants of ClpP1 and ClpP2. A. Cartoon representation of the LGF-loops (dark blue) of the chaperone binding to hydrophobic surface patches (green) on the protease core (grey). **B.** Top view of a heptameric protease ring. The hydrophobic patches (green) are formed by residues of two adjacent protease subunits (grey). **C**. Mutations introduced in ClpP1 and ClpP2 to create the hydrophobic patch variants hpClpP1 (ClpP1$^{S61A,\ Y63V,\ L83A,\ Y91V}$) and hpClpP2 (ClpP2$^{Y75V,\ Y95V}$). **D.** Alignment of *Mtb* ClpP1 and ClpP2 with EcClpP. Conservation is colored from white (not conserved) to black (identical). The identity between ClpP1/ClpP2 is 39.5%, between EcClpP/ClpP1 46% and EcClpP/ClpP2 44.4%. Red arrows highlight the residue positions of the EcClpP hydrophobic patch residues. Residues depicted in green in panel E are marked with a green box. *Continued next page*

Figure 2.3 (continued): E. Surface representation of ClpP1, ClpP2 (4U0G.pdb) and EcClpP (3MT6.pdb) ring faces. Individual subunits are colored alternatingly in light and dark grey. Hydrophobic patch residues are colored in green and labelled accordingly. For ClpP1 and ClpP2 the hydrophobic patch residues used for mutation are shown, for EcClpP reference residues are shown as described in the literature (Wang et al, 1997). **F**. Creation of a set of mature mixed wild-type ClpP1 and ClpP2 (wtP1, wtP2) and hydrophobic patch variants (hpP1, hpP2). Large-scale processing of ClpP1 and ClpP2 (70 µM protomer each) containing the N-terminal propeptide (proClpP1, proClpP2) to mature ClpP1 and ClpP2 (mClpP1, mClpP2) in the presence of 1 mM activator. The reaction was performed in Buffer A. All samples were run on the same gel. The lane containing the size marker was removed for better visual representation (white line). **G**. Analytical gel filtration was performed with the mature ClpP1P2 complexes created in panel F. The peak at 11 ml shows that all complexes have assembled into double-rings.

From the wild-type subunits and the hpClpP variants, four different proClpP1P2 particles were produced, with either both partners wild-type (wtClpP1/wtClpP2), both partners hp variants (hpClpP1/hpClpP2), or one partner wild-type and the other hp variant in the two mixed particles (wtClpP1/hpClpP2 and hpClpP1/wtClpP2). For the generation of the mature particles, the proClpP1P2 particles were incubated in the presence of the activator and processing was verified by SDS-PAGE analysis. Figure 2.3 F shows that processing to the mature particle occurs for all four particles. The correct size of the resulting mature particles was verified by mass spectrometry. Subsequent analytical gel filtration shows that all particles are assembled into the double-ring complex (Figure 2.3 G). The peak fractions corresponding to the assembled particle were collected and the mature ClpP1P2 particles were used for subsequent experiments. Together, these results show that the hydrophobic patch mutations did not affect the assembly or peptidase function of the ClpP1P2 complexes, as particles containing hpClpP1 and hpClpP2 variants assemble to their double-ring functional state and are active in propeptide cleavage.

The interaction competence of the hydrophobic patch variants with the ATPase partners ClpX and ClpC1, both from *Mtb*, was then tested by chaperone-dependent protein degradation assays with ssrA-tagged model substrates. To assess ClpC1-dependent degradation, GFP carrying the *Mtb* ssrA-tag at its C-terminus (GFP-ssrA) was used, allowing detection of the activity by fluorescence spectroscopy. To measure ClpX-dependent degradation, malate dehydrogenase, extended C-terminally with the *Mtb* ssrA-tag (MDH-ssrA) was employed as a model substrate, because unfolding of the stable GFP-ssrA is not well supported by the weaker ClpX unfoldase activity.

Figure 2.4.A shows ClpX-dependent degradation of MDH-ssrA, measured by the disappearance of the MDH-ssrA protein band in SDS-PAGE. The wild-type complex supports almost complete degradation of MDH-ssrA over the time course of 7 hours. The small band below the MDH-ssrA band (marked with *) was confirmed by mass spectrometry to be MDH, most probably lacking the ssrA tag. As untagged MDH is not recruited to the ATPase, it can consequently not be de-

graded by the complex. As expected, the particle formed from both hydrophobic patch variants (hpP1/hpP2) does not exhibit any degradation of MDH-ssrA within the same time frame. Carrying out the same assay with the hydrophobic patch/wild-type mixed particles produces two opposite outcomes. With the particle in which ClpP2 carries the mutation (wtP1/hpP2), MDH-ssrA degradation is completely abolished, while the particle composed of a wild-type ClpP2 and a mutant ClpP1 ring (hpP1/wtP2) exhibits degradation activity comparable to the wild-type particle (Figure 2.4.A). This demonstrates that ClpX-dependent degradation only occurs in context of the wild-type ClpP2 ring surface, indicating that ClpP2 is the interaction platform for association with ClpX.

ClpC1-dependent degradation of GFP-ssrA was then tested using the four different ClpP double-ring particles. Here, the degradation time course was measured by following the loss of the intrinsic fluorescence signal of GFP, and end point samples were assessed by SDS-PAGE analysis (Figure 2.4 B). Analogous to ClpX-mediated degradation, we again observe degradation with the wild type particle (wtP1/wtP2) and the mixed particle carrying interaction-competent ClpP2 subunits (hpP1/wtP2). No degradation is observed for the hp particle (hpP1/hpP2) or the mixed particle carrying a mutation in the ClpP2 hydrophobic patch (wtP1/hpP2). Hence, ClpC1-dependent degradation also relies on ClpP2 as the interaction platform. Even in a heterologous setup, where EcClpX is used in place of *Mtb* ClpX, the particle with wild-type ClpP1 and a hydrophobic patch mutation in ClpP2 results in loss of activity, while the mixed particle carrying mutated ClpP1 and wild-type ClpP2 is fully active (Figure S 3). These results suggest that ClpP2 generally functions as the interaction platform for chaperone binding partners and that the *Mtb* ClpP1P2 particle in contrast to *E. coli* ClpP forms asymmetric complexes, capped only on one side by a chaperone partner.

All chaperone-dependent degradation assays were carried out in absence of the activator. Addition of the activator leads to a mild increase in the reaction rate, but does not change the overall result (Figure S 4 and S 5). This means that also in the presence of activator, the wild-type ClpP2 interaction surface is required to obtain ClpX- or ClpC1-dependent substrate degradation.

Taken together, our results show that both chaperones of the *Mtb* Clp system, ClpX and ClpC1, need the interaction surface of ClpP2 to support degradation of ssrA-tagged substrates, indicating that ClpP2 is the interaction platform for both ClpX and ClpC1.

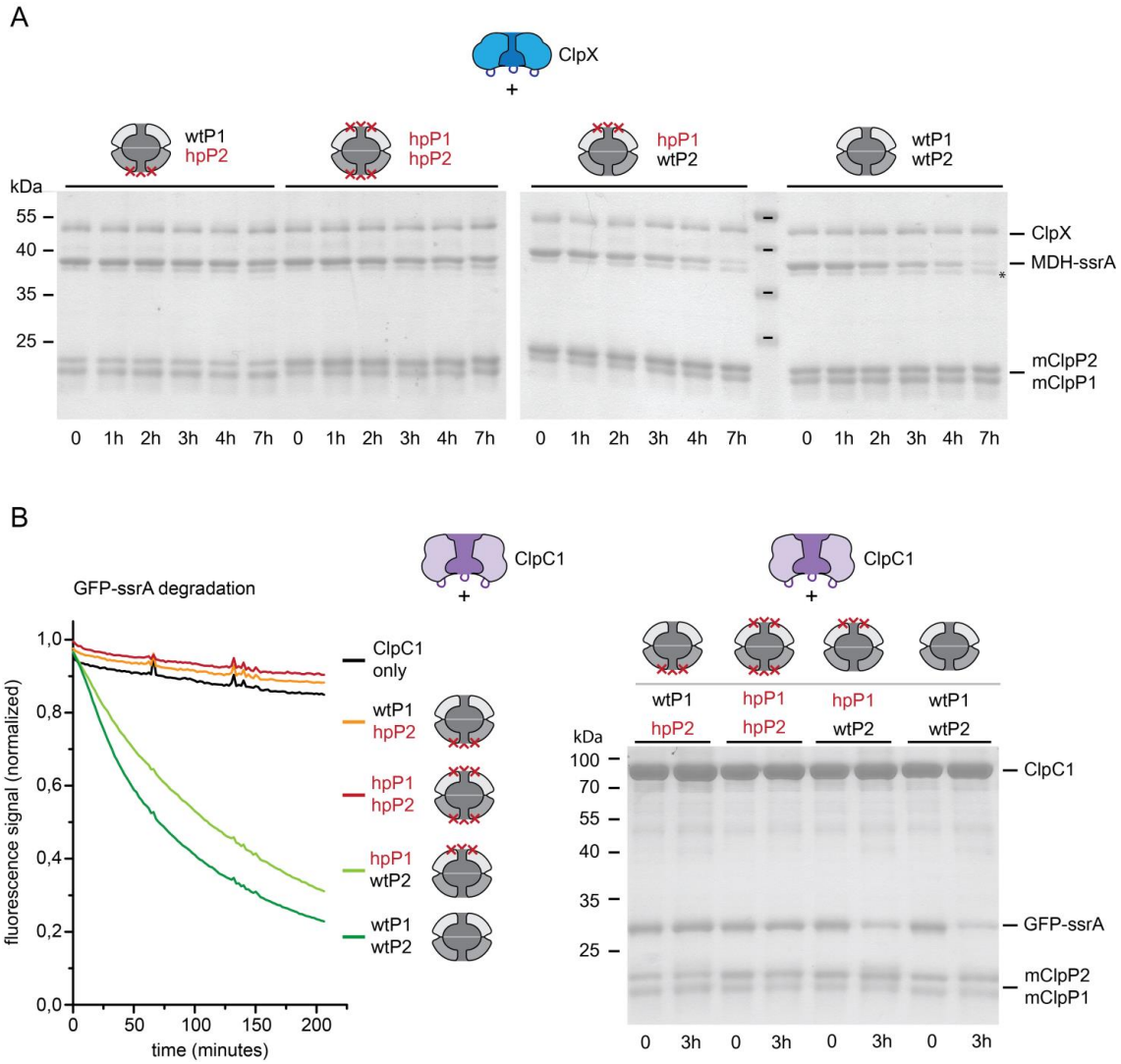


Figure 2.4: **Chaperone-mediated degradation of ssrA-tagged substrates by ClpP1P2 requires the hydrophobic patch on ClpP2.** ClpX and ClpC1-dependent degradation of model substrates was assayed with a set of mature ClpP1P2 particles, created from mixed wild-type (wtP1, wtP2) and hydrophobic patch variants (hpP1, hpP2) of ClpP1 and ClpP2. **A.** Degradation of MDH-ssrA (2 µM) mediated by ClpX (1 µM hexamer) and wt, hp or mixed mature ClpP1P2 particles (0.5 µM double-ring particle), was followed by the disappearance of the MDH-ssrA band in SDS-PAGE. The band just below MDH-ssrA that is not degraded (*) was confirmed by MS/MS to be composed of MDH, most probably lacking the ssrA tag. **B**. Degradation of GFP-ssrA (2 µM) mediated by ClpC1 (1 µM hexamer) and by wt, hp and mixed mature ClpP1P2 particles (0.5 µM double-ring particle) was monitored by the loss of the intrinsic GFP fluorescence signal. The signal was globally normalized. Additionally, time points were taken at the beginning and the end of the reaction and degradation of GFP-ssrA was confirmed by SDS-PAGE.

2.6 Discussion

The Clp chaperone-protease complexes are formed by the coaxial stacking interaction of hexameric ATPase rings on top of the ClpP double-ring particle. In the well-studied *E. coli* Clp system, ATPase partners bind to both sides of the protease core to form symmetric particles (Kessel et al, 1995; Maglica et al, 2009). However, as we show here, in the *Mtb* Clp system this symmetric stacking does not occur. In contrast to *E. coli*, *Mtb* harbors not only one ClpP subunit, but two (ClpP1 and ClpP2), forming a hetero ClpP1P2 double-ring that presents two different chaperone interaction surfaces. While first theories proposed that each of the two chaperones, ClpX and ClpC1, specifically binds to either ClpP1 or ClpP2, our results clearly show that both chaperones only use the ring surface of ClpP2 to build the protein degradation-competent complexes.

There are two motifs on the ClpP protease core involved in chaperone interaction, the N-loop and a hydrophobic surface patch (Bewley et al, 2006; Gribun et al, 2005; Joshi et al, 2004; Kang et al, 2004; Kim et al, 2001; Wang et al, 1997). In order to investigate interaction of both chaperones with the hetero double-ring protease particles, we designed mutants of ClpP1 and ClpP2 where the respective hydrophobic patches were mutated, thereby impairing chaperone binding. Indeed, ClpP1P2 particles where both proteases have the mutated hydrophobic patch are no longer able to support chaperone-mediated degradation. As this particle assembles into the double-ring complex and shows propeptide processing activity, we can be confident that the hydrophobic patch mutations neither impair correct ClpP1P2 complex formation nor do they affect the integrity of the active sites. Our selection of hydrophobic patch residues was based on three aromatic residues described for *E. coli* ClpP Y74, Y76 and F96 (Wang et al, 1997). *Mtb* ClpP2 has two aromatic residues in corresponding positions, Y75 and Y95, that were selected for mutation, as they both were shown to coordinate ADEP, an antibiotic that binds into the hydrophobic patch and mimicks binding of the LGF-loop of the chaperone (Schmitz et al, 2014). In ClpP1, only one of the three corresponding residues is an aromatic residue, Y63, so a nearby aromatic residue that stacked onto Y63, Y91, was also mutated, as well as two less bulky residues. As the hydrophobic patches of ClpP1 and ClpP2 are different in amino acid composition, they also differ slightly in shape. While ADEP binds into the hydrophobic patch of ClpP2, it is too bulky to fit into the hydrophobic patch of ClpP1 (Schmitz et al, 2014), suggesting that a similar reason might cause the LGF-loops of ClpX and ClpC1 (as well as the homologous IGF-loop of Ec-ClpX) to only interact with ClpP2.

The second chaperone-binding element of the Clp proteases could also contribute to the differential interaction of ClpX and ClpC1 with ClpP2. The N-loops of ClpP1 and ClpP2 differ in their

length, about 7 residues in ClpP1 versus 17 residues in ClpP2, and do not show structural similarities. The N-loop of ClpP2 is well resolved in the crystal structure, and contains a β-hairpin necessary for efficient substrate translocation (Alexopoulos et al, 2013; Gribun et al, 2005; Schmitz et al, 2014), while the ClpP1 N-loop residues are resolved neither in the ClpP1P1 nor the ClpP1P2 structure (Ingvarsson et al, 2007; Schmitz et al, 2014), indicating a large degree of flexibility. While the N-loops are not resolved in ClpP1 in either structure, the apparent axial pore size of ClpP1 in the inactive ClpP1P1 structure (12 Å) differs substantially from the one of ClpP1 in the active ADEP-bound ClpP1P2 structure (30 Å), indicating that ADEP-binding to ClpP2 allosterically opens the ClpP1 pore (Schmitz et al, 2014). Therefore, binding of the ClpC1 and ClpX chaperones to ClpP2 might lead to an open ClpP1 pore – a situation that would be detrimental to the cell, as ADEP-activated ClpP was shown to degrade nascent protein chains (Kirstein et al, 2009a). However, it is unlikely that ClpP1 exists in the cell in a deregulated open-pore form. There are indications that pore widening by the chaperone binding partner is less pronounced than pore opening caused by ADEP-binding (Alexopoulos et al, 2012), in which case the ClpP1 pore in a chaperone-bound ClpP1P2 complex might not be as wide as observed in the ADEP-bound crystal structure. Furthermore, the active conformation of the ClpP1P2 complex in absence of activator is stimulated only in the presence of chaperone together with protein substrate (Schmitz & Sauer, 2014). Therefore it is possible that *in vivo* the ClpP1 pore is only open while substrate translocation from the ClpP2 side of the complex takes place, such that the chamber is filled with the translocating substrate chains and peptide products, preventing access of proteins from the ClpP1 side. An open or dynamic ClpP1 pore could thus rather present a pathway for product release, where exiting peptides might even prevent entrance of substrates. Alternatively, the flexible, in the structure unresolved, N-loops of ClpP1 could act as a kind of pore plug and adopt a “down” conformation, restricting access to the chamber from the ClpP1 side (Bewley et al, 2006; Effantin et al, 2010b).

The N-loop region is well conserved throughout ClpP2 homologues from various Actinobacteria (Figure S 7), and also shares residue identity with ClpP from *E. coli*, for example a conserved proline (Figure S 7, yellow box), which was shown to be important for maturation of ClpP and for ClpAP complex formation in *E. coli* (Bewley et al, 2006). This proline is not present in ClpP1 and most of its homologues from Actinobacteria (Figure S 6), and generally, the region that could form an N-loop is less conserved in ClpP1. In addition to the N-loop, the hydrophobic patch residues are also well conserved in ClpP2 (Figure S 7), and less well in ClpP1 (Figure S 6). The different pattern of conservation in ClpP1 and ClpP2 of the chaperone interaction elements suggests

that the specialization of ClpP1 and ClpP2 occurs not only in *Mtb*, but is a general property of actinobacterial ClpP proteases. The asymmetry in interaction with ATPase partners of the ClpP particles might even extend to other gram-positive organisms, as EM reconstructions for the *Listeria monocytogenes* ClpP1P2 complex show a protruding N-terminal density only on ClpP2, which is reminiscent of the well resolved N-loop of ClpP2 in the *Mtb* ClpP1P2 structure (Schmitz et al, 2014; Zeiler et al, 2011).

The conservation of the chaperone interaction motifs in ClpP2 along with our experimental results support the notion that ClpP2 is the canonical subunit involved in chaperone binding, while ClpP1 developed a more varied interaction surface throughout different organisms. If in the assembled chaperone-protease complex both ClpX and ClpC1 bind to ClpP2, this raises the question why both chaperones asymmetrically interact with the ClpP1P2 complex and only bind to one side of the cylinder, when it was shown for the *E. coli* Clp system that symmetric interaction increases the efficiency of the complex (Maglica et al, 2009). A possible competition between ClpX and ClpC1 for protease binding sites most likely poses no problem, since the protease subunits were shown to be amongst the most abundant proteins in the *Mtb* cell (Schubert et al, 2013). Binding of both chaperone partners to ClpP2 would leave the interaction surface of ClpP1 free for as of yet unknown putative interaction partners. A fragment of ADEP was for example shown to activate ClpP1, indicating that in principle binding is still possible (Schmitz et al, 2014).

Apart from the asymmetry in chaperone interaction, the *Mtb* ClpP1P2 complex also shows specialization in its processing activity of the ClpP1 and ClpP2 propeptides, as suggested by coexpression experiments of *Mtb* ClpP1 and ClpP2 in *E. coli* (Benaroudj et al, 2011). We show that the propeptide of ClpP2 is processed first, and that the processing activity is performed mostly by ClpP1. This differential processing could occur due to the length and/or the sequence of the propeptides. With 12 residues the ClpP2 propeptide has almost double the size of the ClpP1 propeptide (~7 residues) (Akopian et al, 2012; Benaroudj et al, 2011), which, depending on the conformation of the putative ClpP1 loop residues, may not reach the active sites of ClpP2. Furthermore, the ClpP1 and ClpP2 active sites have different cleavage specificities and ClpP1 was shown to be especially important for cleavage after hydrohphobic residues, while ClpP2 was not well able to cleave after such a model peptide, which could contribute to a different role in propeptide cleavage (Akopian et al, 2012). We also show that the synthetic activator peptide is not necessary for propeptide processing, but that the presence of the natural interaction partners, chaperone and substrate, is sufficient for this reaction, suggesting that this is how processing is performed *in vivo*. The ATP-dependence of this reaction shows that chaperone assembly and

chaperone-dependent unfolding and translocation of substrates is necessary for ClpP1P2 activation. Substrate binding in or near the active site could stabilize the active conformation of ClpP1P2 in a similar manner as the activator, and act synergistically with chaperone binding to stimulate ClpP1P2 activity (Schmitz & Sauer, 2014). *In vitro*, propeptide processing still seemed more complete in presence of the activator, presumably because the activator provided better long-term stabilization of the active ClpP1P2 conformation as it is not degraded.

During propeptide cleavage the assembled ClpP1P2 complex is stable at room temperature, as an excess of either subunit did not lead to more processing. This shows that the ClpP1P2 complex does not dissociate to reassociate with unprocessed excess subunits, and also that the processing is an intra-particle reaction. The ClpP1P2 complex in presence or absence of the propeptides forms readily at room temperature, without the activator peptide (as opposed to (Akopian et al, 2012)), but dissociates at 4°C, indicating that the interaction between the ClpP1 and ClpP2 rings is mainly mediated by hydrophobic interactions (Baldwin, 1986; Chandler, 2005; Matthews, 2001). Correlating hydrophobic interactions between the rings to the available structural information is not straightforward. The ClpP double-ring can adopt active/extended or inactive/compressed conformations with large differences in the interaction interface as represented by the active ClpP1P2 versus the inactive ClpP1P1 structure (Figure S **8**) (Ingvarsson et al, 2007; Schmitz et al, 2014). Our assembly tests were performed in the absence of activator, and therefore the ClpP1P2 complex is presumably in an inactive conformation, as for activity either activator or chaperone together with substrate were shown to be necessary (Akopian et al, 2012; Schmitz & Sauer, 2014). Analysis of the interface residues between the ClpP1P2 rings and the ClpP1P1 rings shows that irrespective of the difference in the interaction surface area, both interfaces are composed in large portions of hydrophobic interactions that could account for the observed behavior (Figure S 9). However, neither of the complexes exactly represents the interface in our particle and a structure of the inactive conformation of ClpP1P2 would be required for a more quantitative analysis.

Under *in vivo* conditions the ClpP1P2 complex is most likely stable and in an inactive conformation, while the presence of chaperone and substrate dynamically activate the complex at need (Schmitz & Sauer, 2014). It remains to be seen, whether additional factors can regulate the activity and substrate specificity of the fully assembled Clp chaperone-proteases, potentially not only by binding to the ATPase rings as has been observed for various adaptors in other Clp proteases systems, but by binding to the available ClpP1 ring surface in the complex.

2.7 Acknowledgment

We thank Dr. David Ramrath for preparing the electron micrograph image of proClpP1. We thank the Scientific Center for Optical and Electron Microscopy (SCOPEM) of the ETH Zurich for support.

2.8 Supplementary Information

This section contains Supplementary Figures 1-9.

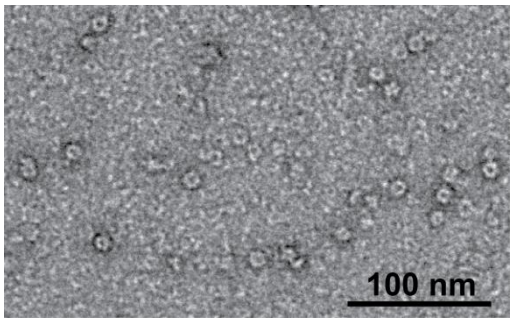


Figure S 1: Negative-stain TEM micrograph of proClpP1. 1.4 µM proClpP1 (protomer) was stained with 2% aqueous uranyl acetate.

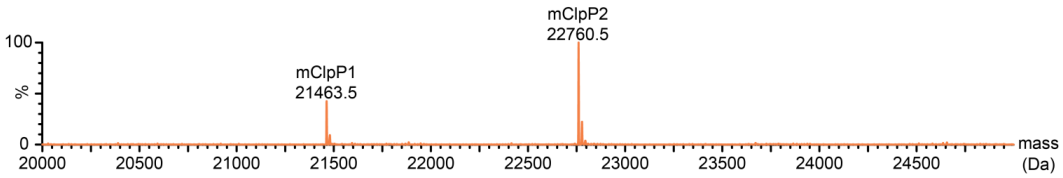


Figure S 2: mClpP1 has a mass weight corresponding to processing after Met7. Electron spray ionisation mass spectrometry of mClpP1P2. proClpP1P2 was processed overnight in the presence of 1 mM activator to produce the mature complex. The expected mass for mClpP1-His4 processed after Met7 is 21463.3 Da. For mClpP2-His4 processed after Ala12 the expected mass is 22760.9 Da.

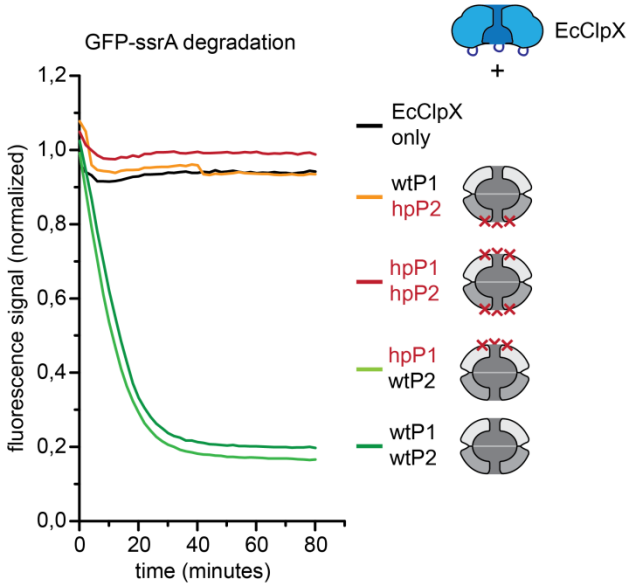


Figure S 3: EcClpX-mediated degradation of GFP-ssrA by ClpP1P2 requires the hydrophobic patch on ClpP2. EcClpX-mediated (1 µM hexamer) degradation of GFP-ssrA (E. coli ssrA tag sequence) (2 µM) by wild-type (wt), hydrophobic patch (hp) and mixed mature ClpP1P2 particles (0.5 µM double-ring particle) was monitored by the loss of the intrinsic GFP fluorescence signal. The signal was globally normalized.

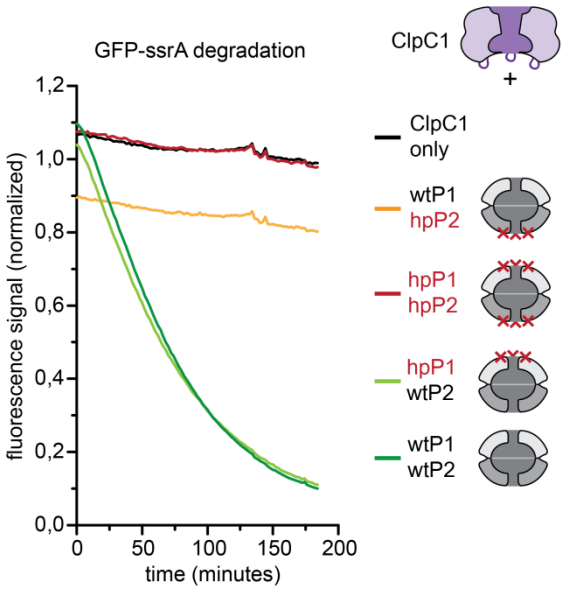


Figure S 4: ClpC1-mediated degradation of GFP-ssrA by ClpP1P2 requires the hydrophobic patch on ClpP2 also in presence of activator. ClpC1-mediated degradation (1 µM hexamer) of GFP-ssrA (2 µM) by wild-type (wt), hydrophobic patch (hp) and mixed mature ClpP1P2 particles (0.5 µM double-ring particle) in the presence of 1 mM activator was monitored by the loss of the intrinsic GFP fluorescence signal. The signal was globally normalized.

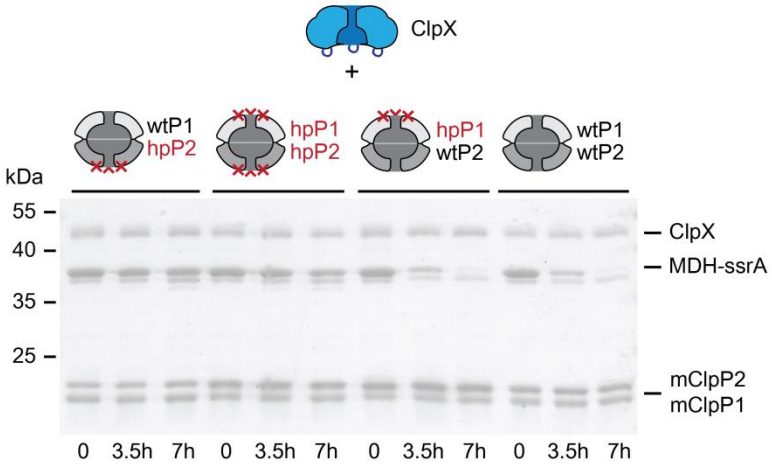


Figure S 5: ClpX-mediated degradation of MDH-ssrA by ClpP1P2 requires the hydrophobic patch on ClpP2 also in presence of activator. ClpX-mediated degradation (1 µM hexamer) of the substrate MDH-ssrA (2 µM) by wild-type (wt), hydrophobic patch (hp) and mixed mature ClpP1P2 particles (0.5 µM double-ring particle) in the presence of 1 mM activator, was followed by the disappearance of the MDH-ssrA band on an SDS-PA gel at the time points indicated below the gel.

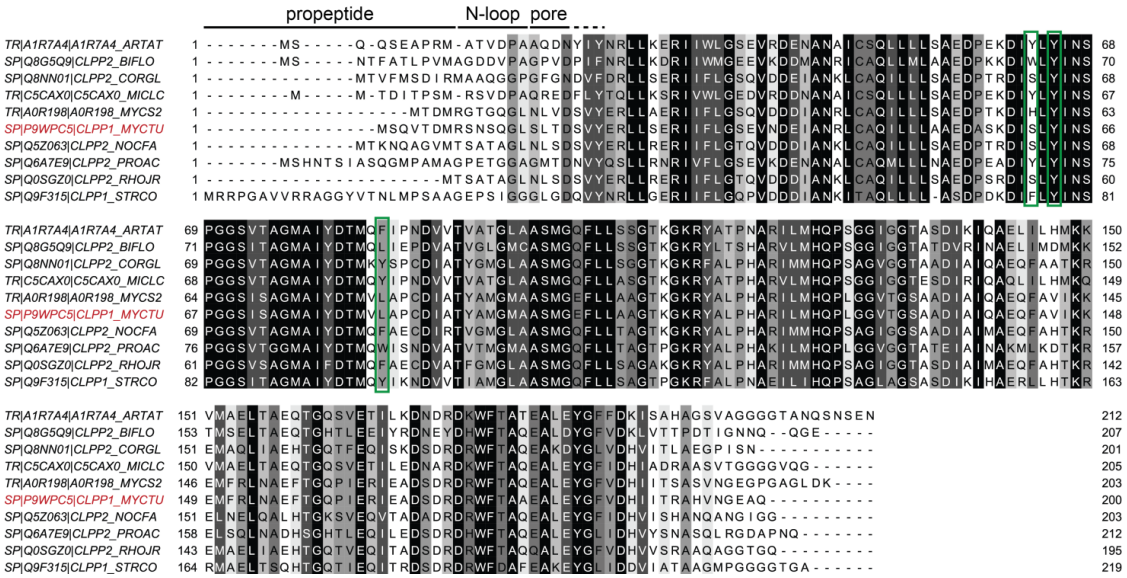


Figure S 6: Sequence conservation amongst *Mtb* ClpP1 and actinobacterial homologues. ClpP1 was aligned with homologous actinobacterial proteins. Conservation is colored from white (not conserved) to black (identical). Sequences and naming were extracted from the Uniprot database. The designation of ClpP subunits as ClpP1 or ClpP2 for different Actinobacteria does not always match the *Mtb* designation. Alignment was based on homology, not on the naming of the subunits. The Uniprot identifiers are given in the sequence labels. Organism abbreviations: ARTAT: *Arthrobacter aurescens*, BIFLO: *Bifidobacterium longum*, CORGL: *Corynebacterium glutamicum*, MICLC: *Micrococcus luteus*, MYCS2: *Mycobacterium smegmatis*, MYCTU*: Mycobacterium tuberculosis*, NOCFA: *Nocardia farcinica*, PROAC: *Propionibacterium acnes*, RHOJR: *Rhodococcus jostii*, STRCO: *Streptomyces coelicolor*. The label for *Mtb* ClpP1 is colored in red, hydrophobic patch residues are marked with green boxes. The annotation of the N-loop and pore residues is based on the *Mtb* ClpP1 structure (2CE3.pdb).

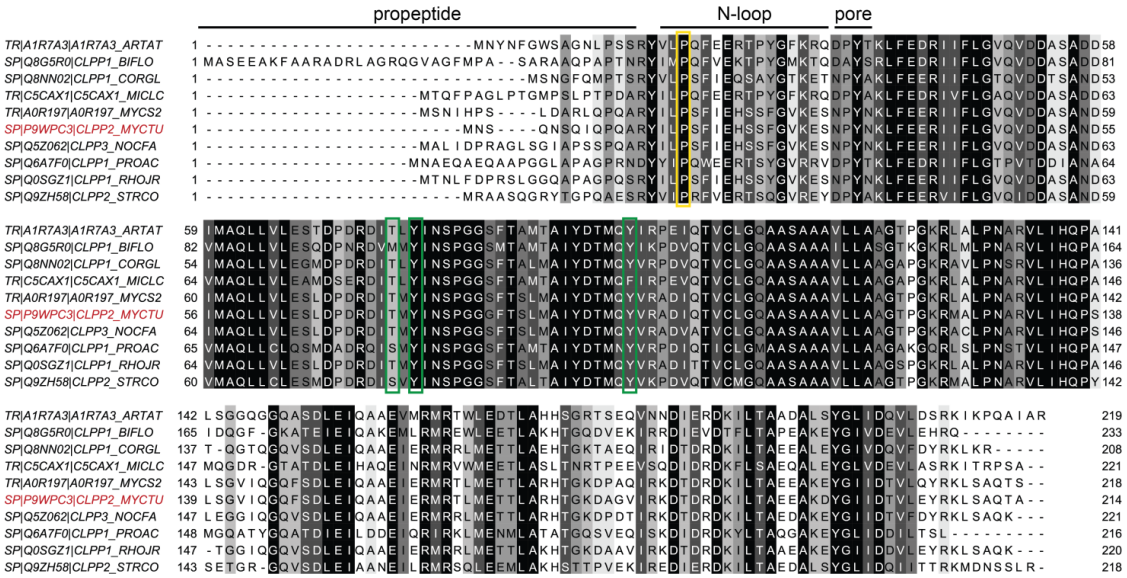


Figure S 7: Sequence conservation amongst *Mtb* ClpP2 and actinobacterial homologues. ClpP2 was aligned with homologous actinobacterial proteins. Conservation is colored from white (not conserved) to black (identical). Sequences and naming were extracted from the Uniprot database. The designation of ClpP subunits as ClpP1 or ClpP2 for different Actinobacteria does not always match the *Mtb* designation. Alignment was based on homology, not on the naming of the subunits. The Uniprot identifiers are given in the sequence labels. Organism abbreviations: ARTAT: *Arthrobacter aurescens*, BIFLO: *Bifidobacterium longum*, CORGL: *Corynebacterium glutamicum*, MICLC: *Micrococcus luteus*, MYCS2: *Mycobacterium smegmatis*, MYCTU*: Mycobacterium tuberculosis*, NOCFA: *Nocardia farcinica*, PROAC*: Propionibacterium acnes,* RHOJR: *Rhodococcus jostii*, STRCO*: Streptomyces coelicolor*. The label for *Mtb* ClpP2 is colored in red, hydrophobic patch residues are marked with green boxes and a conserved proline with a yellow box. The annotation of the N-loop and pore residues is based on the *Mtb* ClpP2 structure (4U0G.pdb).

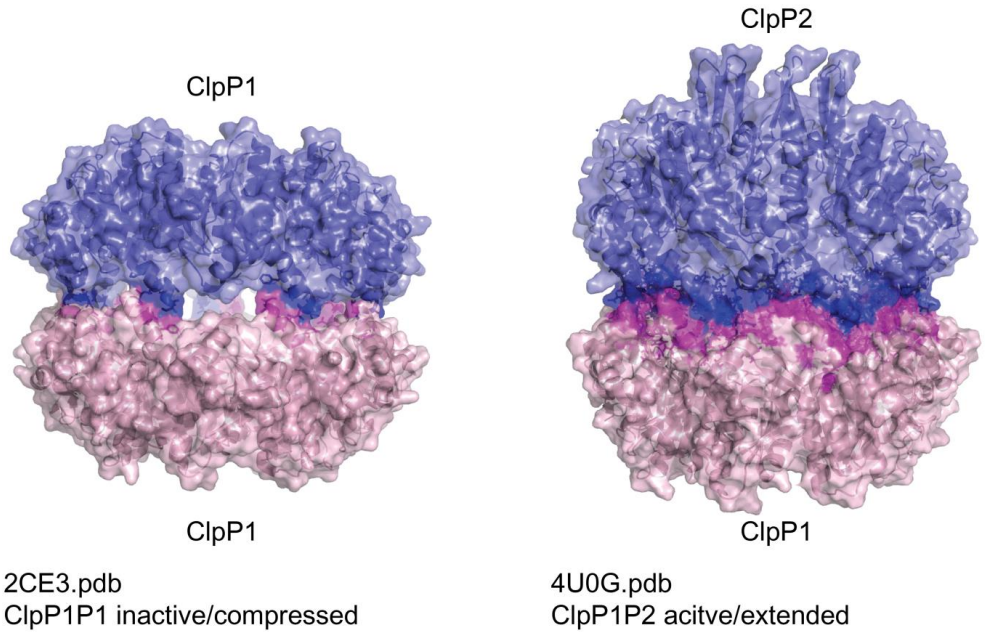


Figure S 8: The *Mtb* ClpP1P1 and ClpP1P2 show differences in their ring-ring interaction surface areas. Interaction residues of ClpP1P1 (2CE3.pdb, left side) and ClpP1P2 (4U0G.pdb, right side) were determined and depicted using the COCOMAPS web application with standard settings (Vangone et al, 2011). The individual rings are colored light violet and light pink in cartoon representation, while the respective interaction residues are colored in dark violet and dark pink and are additionally shown in stick representation.

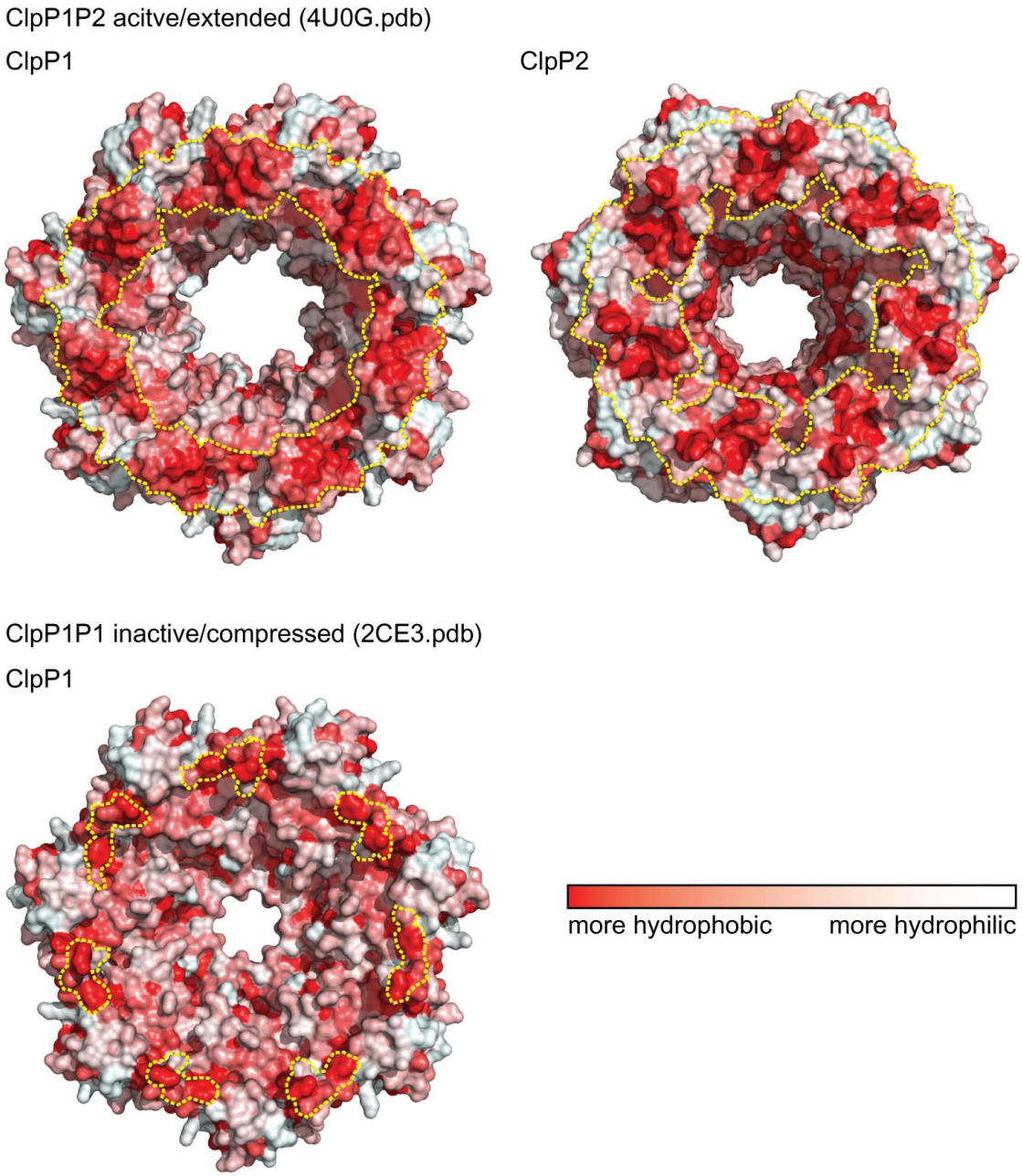


Figure S 9: Hydrophobic interactions in the ClpP1P1 and ClpP1P2 ring-ring interfaces. The interaction surface areas of ClpP1 (upper left) and ClpP2 (upper right) of the active/extended ClpP1P2 structure (4U0G.pdb) and of one ClpP1 ring (lower left) of the inactive/compressed ClpP1P1 structure (2CE3.pdb). The rings are shown from the interface side and the residues involved in the interaction, as determined by the COCOMAPS web application with standard settings (Vangone et al, 2011) are rimmed with a yellow dotted line. Amino acids are colored according to their hydrophobicity using the Eisenberg hydrophobicity scale (http://web.expasy.org/protscale/pscale/Hphob.Eisenberg.html). Red color indicates the most hydrophobic and white color the least hydrophobic residues.

Chapter 3 Interaction partners of the chaperone ClpC1 and its N-terminal domain homologue ClpC2

3.1 Introduction

This part of the thesis focuses on the clientele of protein interactors recruited by the AAA chaperone ClpC1 in addition to the ClpP protease core. These interactors are expected to fall into two categories, degradation substrate proteins and adaptors of substrate specificity. A bacterial two hybrid screen carried out on an *Mtb* ORF library was employed to discover new interaction partners of ClpC1, as well as of ClpC2, a small protein homologous to the ClpC1 N-terminal substrate binding domain. The results from the screen reveal that 10-15 % of the library hits belong to the toxin-antitoxin family. *In vitro* assays examining antitoxin degradation confirm that antitoxins are indeed substrates of the ClpC1P1P2 chaperone-protease.

The bacterial two hybrid experiments performed on MacConkey agar, the size exclusion experiments and library screen with ClpC2, and degradation experiments with GFP-ssrA were performed by the master student Anne Kerschenmeyer under the author's supervision. The bacterial two hybrid system and the *Mtb* ORF library was established and provided by Michał Ziemski, another PhD student in the Weber-Ban group.

3.2 Results

3.2.1 Sequence relationship between ClpC1 and ClpC2

ClpC1 is a chaperone of the *Mtb* Clp system whose function is the recognition of target proteins and their subsequent unfolding and threading into the ClpP1P2 protease core for protein degradation (Laederach et al, 2014). It is composed of an N-terminal domain (residues 1-164) and two AAA+ modules, D1 (residues 165-420) and D2 (residues 477-810). Both AAA+ modules harbor the nucleotide-binding motifs Walker A (D1: residues 216-233; D2: residues 553-560) and Walker B (D1: residues 282-290; D2: 622-628) (Figure 3.1 A and B). In its active, assembled hexameric form, ClpC forms a ring-shaped structure with a two-tiered appearance due to the two AAA+ modules (Liu et al, 2013; Sauer & Baker, 2011). A loop containing a conserved GYVG-motif (residues 600-604) is important for substrate translocation and a loop carrying the

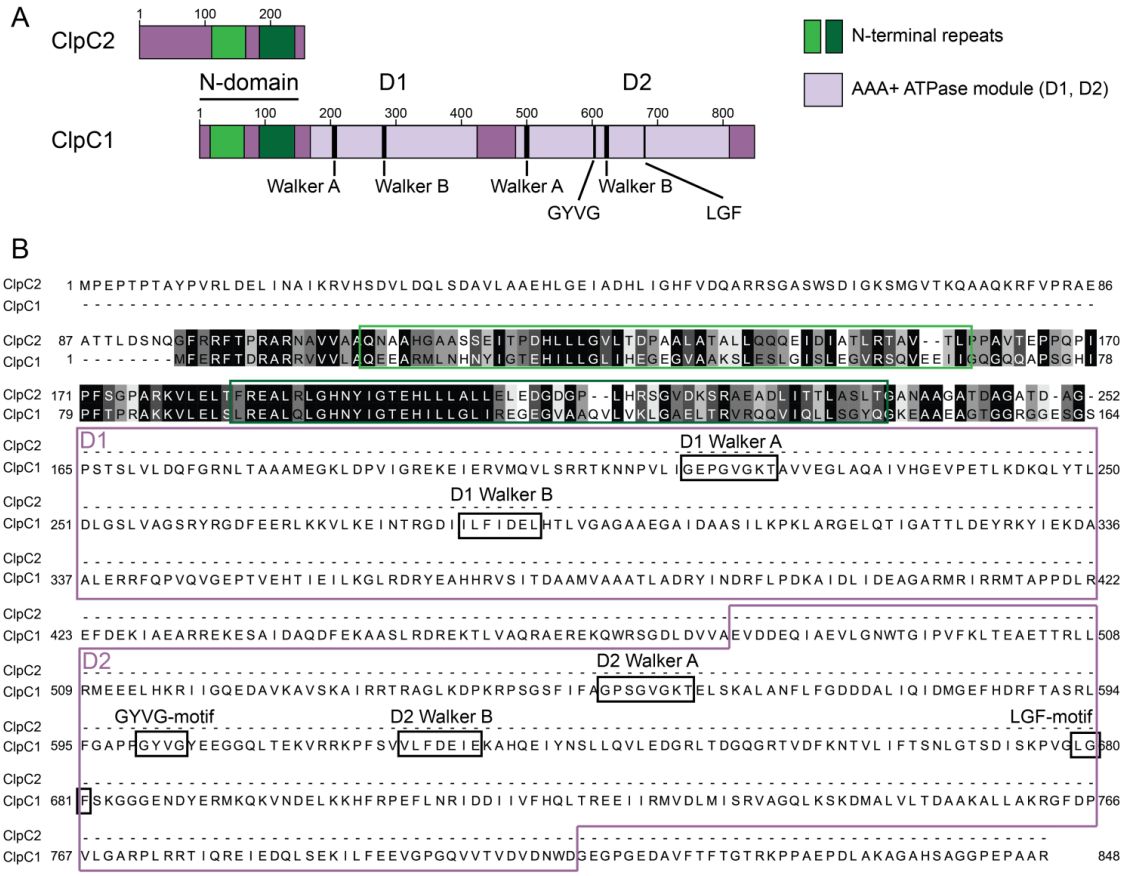


Figure 3.1: Sequence conservation between the ClpC1 N-terminal domain and ClpC2. A. Cartoon representation of the ClpC1 and ClpC2 domain arrangement. The N-terminal repeat domains are colored light and dark green, the AAA+ ATPase modules D1 and D2 in light violet, and further ClpC1 sequence motifs in black (see B for details). Numbers above the cartoon indicate amino acid residue number. **B**. Sequence alignment of *Mtb* ClpC2 and ClpC1. Sequence conservation is indicated by color from black (identical) to white (not conserved). The green boxes indicate the N-terminal repeats as annotated by the NCBI database for *Mtb* ClpC1 (light green: N-terminal repeat 1: residues 16-68; dark green: N-terminal repeat 2: residues 91-143). The ClpC1 N-terminal domain is composed of residues 1-164. The AAA+ modules D1 (residues 165-420) and D2 (residues 477-810) are outlined by violet boxes. The sequence motifs Walker A (D1: residues 216-233; D2: residues 553-560) and Walker B (D1: residues 282-290; D2: 622-628), GYVG (residues 600-604) and LGF (residues 679-681) are marked in black boxes.

LGF-motif (residues 679-681) is important for association with the ClpP2 side of the ClpP1P2 protease complex (Figure 3.1 A and B) (Kim et al, 2001; Leodolter et al, 2015). The ClpC1 N-terminal domain is contains two conserved N-terminal repeat domains (Figure 3.1 A and B, light and dark green) and forms a tightly packed domain composed of eight α-helices (Vasudevan et al, 2013). The N-terminal domains of AAA chaperones are often directly responsible for substrate recognition, as well as adaptor protein recruitment, thereby switching substrate specificity (Guo et al, 2002; Mogk et al, 2004; Zeth et al, 2002). They are the least conserved part between the different AAA chaperones, conferring different

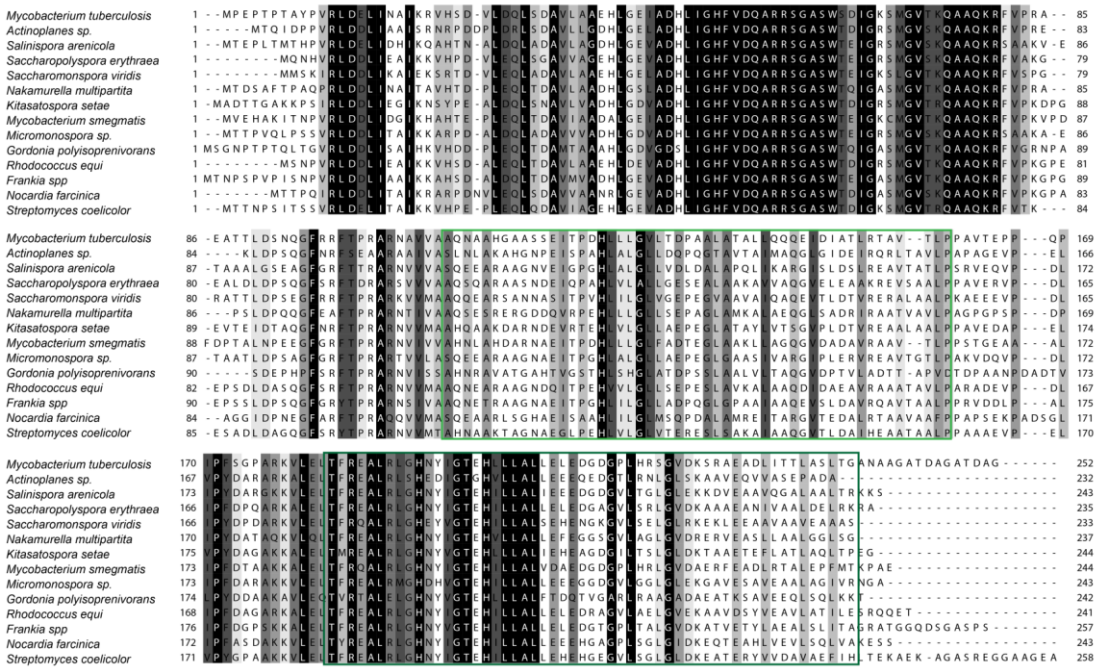


Figure 3.2: Sequence conservation of ClpC2 and related homologues. ClpC2 from *Mtb* was aligned with homologues from related actinobacterial species using the ClustalO algorithm. The names and sequences are as used in the UniProt database. Sequence conservation is indicated by color from black (identical) to white (not conserved). The UniProt identifiers are indicated in the sequence labels. The green boxes indicate the N-terminal repeats as annotated by NCBI for *Mtb* ClpC1 (light green: N-terminal repeat 1; dark green: N-terminal repeat 2). UniProt identifiers for the ClpC2 homologues of the respective organisms: *Mycobacterium tuberculosis*: P9WPC7; *Actinoplanes sp.*: G8SA70; *Salinispora arenicola*: A8M715; *Saccharopolyspora erythraea*: A4FEV4; *Saccharomonospora viridis*: C7MWY2; *Nakamurella mulitpartita*: C8XC43; *Kitasatospora setae*: E4N688; *Mycobacterium smegmatis:* A0QW35; *Micromonospora sp.*: C4RR89; *Gordonia polyisoprenivorans*: H6MXU3; *Rhodococcus equi*: E4WAI4; *Frankia sp.*: A8L6Y7; *Nocardia farcinica*: Q5YNW2; S*treptomyces coelicolor:* Q9X8L2.

specificity to the alternate protease complexes, for example the ClpXP versus the ClpCP proteases (Mogk et al, 2004).

In *Mtb,* in addition to ClpC1, a second, much smaller gene is annotated as belonging to the ClpC subfamily, and was termed *clpC2* (*rv2667),* encoding a protein of 26 kDa molecular weight. Sequence analysis of ClpC2 shows that it does not contain any domains indicating ATPase function, but features a conserved sequence stretch that is homologous to the N-terminal domain of ClpC1 (Figure 3.1 A and B). The conserved region is preceded by an extended N-terminal sequence stretch that is not present in ClpC1 (or any other protein found), but is conserved in ClpC2 homologues from other Actinobacteria (Figure 3.2).

The homology of ClpC2 to the ClpC1 N-terminal domain, the chaperones' substrate and adaptor interface, supports the hypothesis that ClpC2 could be a binder for similar set of proteins as ClpC1. However, the absence of conserved enzymatic domains (e.g. ATPase or protease), ex-

cludes a direct role in substrate unfolding/degradation, but rather suggests putative roles for ClpC2 either as an adaptor protein, perhaps docking to other AAA chaperones via its conserved N-terminus, or as a competitive substrate binder titrating substrates away from ClpC1.

3.2.2 Screening for interactors of ClpC1 in an *Mtb* ORF library using the BACTH system.

Substrates described for the *Mtb* ClpC1P1P2 chaperone-protease so far are limited to proteins carrying an ssrA-tag (Leodolter et al, 2015; Raju et al, 2012b), and the anti-sigma factor RseA (Barik et al, 2010). Other substrates, for example the ClgR gene regulator, were shown to be degraded by ClpP1P2, but recognition was not attributed to either chaperone ClpX or ClpC1 (Sherrid et al, 2010). Additionally, so far no adaptor proteins have been described for the *Mtb* ClpC1 protein.

To identify potential specificity factors interacting with ClpC1 or to discover new substrate classes, we screened an ORF library containing around 4800 genes of the *Mtb* strains H37Rv and CDC1551. The method used to screen for interaction was the Bacterial Two-Hybrid approach based on adenylate cyclase activity (BACTH: bacterial adenylate cyclase two hybrid)(Karimova et al, 1998), reviewed in detail in (Battesti & Bouveret, 2012)). This screening method reads out the reconstitution of *B. pertussis* adenylate cyclase activity in an *E. coli cya-* strain. The adenylate cyclase domain is split in two inactive subdomains, T18 and T25 that are fused to the bait and prey proteins. Upon interaction of a bait and prey pair, the two subdomains are brought together and adenylate cyclase activity is restored. The resulting cAMP production activates (amongst others) the lactose and maltose operons by binding to CAP (catabolite activator protein). As β-galactosidase is encoded in the lactose operon (lacZ gene), this leads to β-galactosidase activity that can be monitored (Figure 3.3 A). In our experiments we used selective reporter agar plates to detect interaction between the tested proteins. Plating of cells transformed with plasmids containing both bait and prey vectors on MacConkey agar supplemented by 1% maltose reports on the ability of the cells to ferment the sugar in the medium. Interaction and the resulting reconstitution of the adenylate cyclase activity enables this fermentation, which in turn results in a pH drop in the medium that is detected by a change in color from white/yellowish to red (Karimova et al, 1998). Another medium that serves both as selection as well as reporter is minimal M63 agar, supplemented with lactose and X-gal as reporter substances. On these plates, only cells able to use lactose as a sugar source grow (i.e. cells where protein interaction produces cAMP), while additionally

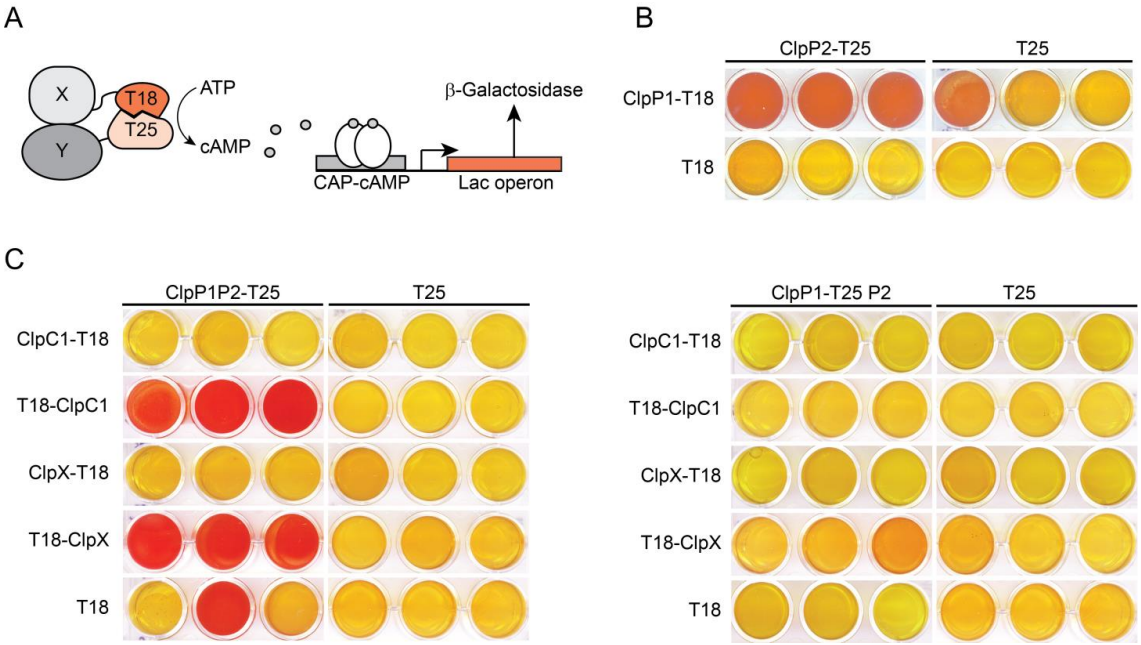


Figure 3.3: Proof of principle experiment for Clp protease and chaperone interaction in the bacterial two hybrid experimental setup. **A**. Cartoon of the BACTH (bacterial adenylate cyclase two hybrid) principle. The proteins of interest (X and Y in light and dark grey) are fused to the T18 (dark orange) and the T25 (light orange) adenylate cyclase subdomains. Upon interaction of target proteins, the adenylate cyclase activity is restored and ATP is converted to cAMP (grey circles). cAMP binds (amonst others) to the catabolite activator protein (CAP), which activates the Lac operon, enhancing the expression of the reporter ß-Galactosidase. **B**. Interaction of the Clp protease subunits ClpP1 and ClpP2. ClpP1 and ClpP2, C-terminally fused to the T18 and T25 domains, repectively, were co-transformed with each other and with the respective empty vector and plated on MacConkey reporter agar. Red coloration of the agar indicates interaction of the proteins. **C.** Interaction of the chaperones ClpX and ClpC1 with the protease core ClpP1P2. The indicated constructs were co-transformed and plated on MacConkey reporter agar. Red coloration of the agar indicates interaction of the co-transformed proteins. Left side: Interaction of ClpX and ClpC1 with the T18 domain fused to either N- or C-terminus with the ClpP1P2 complex, with the T25 domain fused to ClpP2 (ClpP1 coexpressed with ClpP2-T25). Right side: Interaction of ClpX and ClpC1 with the T18 domain fused to either N- or C-terminus with the ClpP1P2 complex, with the T25 domain fused to ClpP1 (ClpP1-T25 coexpressed with ClpP2).

the presence of β-galactosidase is evidenced by the cleavage of X-gal, which leads to blue coloration of colonies.

3.2.3 Proof of principle tests for the BACTH system show expected interactions between the Clp chaperone and protease subunits

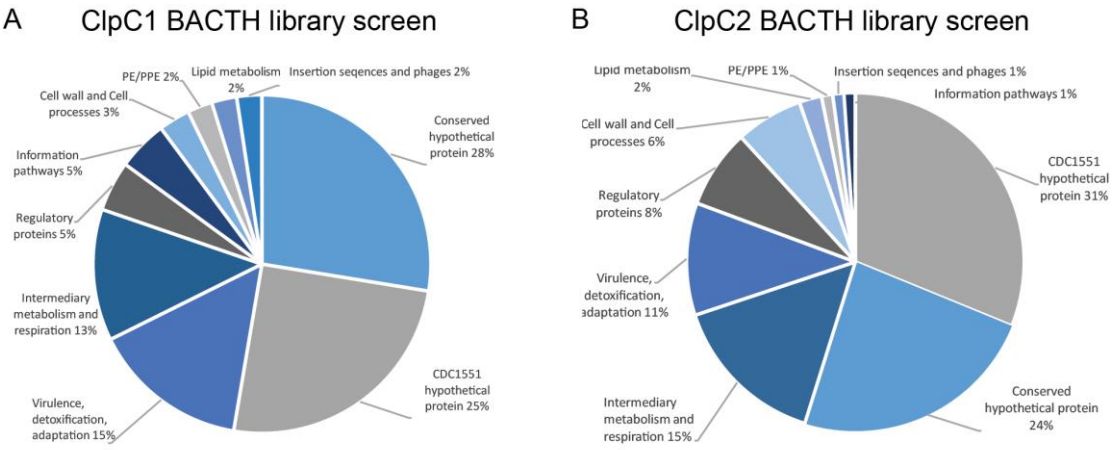
Prior to screening the library, a proof-of-principle experiment was performed to determine the suitability of the Clp system in the BACTH setup. Known interactions between components of the Clp system were probed, namely the interaction between the two protease core subunits ClpP1 and ClpP2, as well as the interaction between the chaperones ClpC1 and ClpX with the protease core ClpP1P2.

In the cell, the protease core subunits ClpP1 and ClpP2 form heptameric rings that assemble a double-ring complex composed of one ring ClpP1 and one ring ClpP2 (Akopian et al, 2012; Raju et al, 2012b). For the purpose of the BACTH screen, the domains T18 and T25 were fused to the C-terminus of ClpP1 and ClpP2, respectively (Figure 3.3 B). The respective plasmids were co-transformed and cells plated on MacConkey reporter agar. The color change of the reporter agar from yellowish to red upon co-transformation of constructs containing ClpP1 and ClpP2 indicates interaction between the two proteins. In the negative controls where each ClpP-containing vector was co-transformed with the corresponding vector without a target gene no color change was observed, except for one well in the triplicate. False positive signals in one of the three wells of the negative control are occasionally observed in experiments using the MacConkey reporter medium (Personal communication from A. Kerschenmeyer and M. Ziemski).

To test the interaction of the chaperones ClpC1 and ClpX with the protease core ClpP1P2, the ClpP1 and ClpP2 subunits were co-expressed on one plasmid and either ClpP1 or ClpP2 was fused on their C-termini to the T25 domain. The plasmids carrying ClpP1-T25/ClpP2 and ClpP1/ClpP2-T25 were co-transformed with the chaperones ClpC1 or ClpX, with the T18 domain fused to N- or C-termini. Figure 3.3 C shows red coloration of the reporter medium only if the T25 domain is fused to the ClpP2 subunit and not the ClpP1 subunit, corroborating my previous results described in chapter 2 of this thesis, where ClpP2 was shown to be the main interaction platform for both chaperones (Leodolter et al, 2015). Additionally, chaperone interaction was only observed if the T18 domain was placed at the N-terminus of the chaperones, indicating that domain placement at the C-terminus sterically hinders interaction with the protease.

3.2.4 Screening of ClpC1 and ClpC2 against an *Mtb* ORF library

Correct ClpC1 expression in BACTH vectors was confirmed by the previous proof of principle experiment, and ClpC1 was subsequently used as a bait protein to screen for unknown interactors. As prey, an *Mtb* ORF library was used, composed of about 3800 genes (3294 ORFs from *M. tuberculosis* H37Rv (>80%) and 430 unique ORFs from *M. tuberculosis* CDC1551) N- or C-terminally fused to the T18 domain. As the library prey proteins feature the T18 domain, the bait proteins were fused to the T25 domain. In the case of ClpC1, the domain was attached to the C-terminus of the protein, as N-terminal fusion might block the interaction with its N-terminal domain, the location where we expect substrates and adaptors to dock.



Functional Category	Unique hits
Conserved hypothetical protein	46
CDC1551 hypothetical protein	42
Virulence, detoxification, adaptation	25
Intermediary metabolism and respiration	21
Regulatory proteins	8
Information pathways	8
Cell wall and Cell processes	5
PE/PPE	4
Lipid metabolism	4
Insertion seqences and phages	4
Grand Total	**167**

Functional Category	Unique hits
CDC1551 hypothetical protein	29
Conserved hypothetical protein	22
Intermediary metabolism and respiration	14
Virulence, detoxification, adaptation	10
Regulatory proteins	7
Cell wall and Cell processes	6
Lipid metabolism	2
PE/PPE	1
Insertion seqences and phages	1
Information pathways	1
Grand Total	**93**

Figure 3.4: Result overview of the ClpC1 and ClpC2 BACTH screen against an *Mtb* ORF library. Unique protein hits of the respective library screen were analyzed in respect to the functional category according to the Tuberculist database. The pie charts show the percentage of hits in each category while the table gives the number of unique hits. **A.** Results of library screen using ClpC1 as bait. **B.** Results of library screen using ClpC2 as bait.

We also performed the library screen using ClpC2, but here the T25 domain was attached to either the N- or the C-terminus of the bait protein. The rationale for this was that unlike for ClpC1, where we know that the ClpC1 ring face with the C-termini interacts with the ClpP particle, for ClpC2 we do not know its assembly state, which type of interaction partners to expect and on which side of the protein they might bind.

For the library screen, competent cells harboring the respective bait plasmids (expressing ClpC1 or ClpC2 fusions to the T25 domain) were transformed with the library plasmids and plated on selective M63/maltose minimal agar supplemented with X-gal for selection. Blue clones that appeared after 5 days of incubation at 30°C were picked and streaked out again on the same medium to remove eventual false positives (~5%). On those clones that appeared after the second round of selection, colony-PCR was performed to amplify plasmid inserts containing the library hits and identify them by sequencing.

The results of the library screens with ClpC1 and ClpC2 are summarized in Figure 3.4 A and B. In the ClpC1 screen 167 unique protein hits were discovered, in the ClpC2 screen 93. The hits were classified into functional categories based on the classification used in the TubercuList database

(http://tuberculist.epfl.ch/). Hits corresponding to genes from *Mtb* strain CDC1551 for which no homologous H37Rv genes could be found and no protein function was annotated, were classified into an additional category termed "CDC1551 hypothetical protein". The number of hits for each category is shown in a table in Figure 3.4 A and B, while the percentage of hits of each group with respect to the total is represented in a pie chart.

For both the ClpC1 and the ClpC2 screen approximately half of the hits consist of hypothetical proteins of unknown function, either from the H37Rv or the CDC1551 strain, which can therefore not be included in the functional category analysis. Of the remaining identified genes, the ones falling into the functional context "Virulence, detoxification and adaptation" present the largest group (15% and 11% of all hits for ClpC1 and ClpC2, respectively). Another significant portion is formed by genes belonging to "Intermediary metabolism and respiration" (13% and 15%). "Regulatory proteins" show with 5% and 8% occurrence also a significant portion.

Appendix Table 2 and 3 list, by functional category, all unique hits for the ClpC1 and the ClpC2 screen. The count of how often a respective hit occurred in the screen gives an indication as to the strength of the interaction. In this context it should be mentioned that 113 out of 167 hits in the ClpC1 screen occurred only once, 49 hits between 2 and 9 times, and 5 hits occurred more than 10 times (Rv0106, Rv3586, MT3135, Rv2012, Rv3916; Table 3.1). For ClpC2, 64 out of 93 ClpC2 hits occurred only once, 25 hits between 2 and 9 times, and 4 hits occurred more than 10 times (MT3135, Rv1693, Rv2012, Rv3916c; Table 3.2).

Table 3.1: Hits from the ClpC1 screen with the most counts. Confirmed false positives are colored in red.

Locus tag	Functional category	Function	Protein name	Counts	also found with bait
Rv0106	Conserved hypothetical protein			50	ClpC2
Rv3586	Information pathways	DNA integrity scanning protein DisA	DisA	43	ClpC2
MT3135	CDC1551 hypothetical protein			38	ClpC2
Rv2012	Conserved hypothetical protein			20	ClpC2
Rv3916c	Conserved hypothetical protein			14	ClpC2

However, hits occurring in multiple screens at high count could also indicate false positives that might occur for reasons like aggregation. Analysis of overlapping hits between the ClpC1 and ClpC2 screen is given in Appendix Table 1. Two of these, which also show a high count, have

been confirmed as false positives: Rv2012 and Rv3916c. Both of these proteins appeared in a negative control library screen performed with the bait vector containing only the T25 domain, indicating that these proteins interacted directly with the T25 domain. Rv2012 furthermore showed a signal when the plasmid was co-transformed with the T25 plasmid without insert (personal communication, M. Ziemski).

Table 3.2: Hits from the ClpC2 screen with the most counts. Confirmed false positives are colored red.

Locus Tag	Functional Category	Protein name	Counts	also found with bait
MT3135	CDC1551 hypothetical protein	MT3135	115	ClpC1
Rv1693	Conserved hypothetical protein	Rv1693	37	ClpC1
Rv2012	Conserved hypothetical protein	Rv2012	19	ClpC1
Rv3916c	Conserved hypothetical protein	Rv3916c	16	ClpC1

The two functional groups that have the highest percentage of unique hits (apart from hypothetical proteins, which span likely a range of functional categories) are “Intermediate metabolism and respiration” and “Virulence, detoxification, adaptation”. Strikingly, while hits found in the first group are composed of diverse enzymes from different pathways, hits found in the second group are with one exception (Metallothionein MymT, a copper binding protein) composed of proteins comprising Toxin-Antitoxin systems. Toxin-antitoxin (TA) systems are small bipartite systems, which are transcribed from one operon and always contain a ribonuclease toxin and an antitoxin that blocks toxin action. The active toxins often act in response to stresses and, in *Mtb,* have been implicated in the pathogenicity and persistence of the bacterium (Sala et al, 2014). Intrestingly, *Mtb* features an unusally high number of TA systems in its genome.

One of the hits identified in the ClpC2 screen was the ClpC2 protein itself, indicating that ClpC2 can form homodimers or higher homooligomers. The protease subunits ClpP1 and ClpP2 were not identified as ClpC1 interactors, although we had shown previously that ClpC1 interacts with the ClpP2 side of the ClpP1P2 particle (Leodolter et al, 2015). The lack of this interaction in the library screen is due to the distance and steric constraints that arose through domain placement on the bait and prey pair in the BACTH assay. As Figure 3.3 C shows, a pairwise complementation assay using ClpC1 with the T25 domain attached to its C-terminus, as is the case in the library bait vector, and ClpP1P2 did not result in a positive signal.

3.2.5 ClpS as a binder of ClpC1

Another gene annotated as belonging to the Clp system, the putative N-end rule adaptor ClpS, is unfortunately not present in the *Mtb* ORF library and could therefore not be picked up in the screen. It was thus analyzed separately to test for interaction with ClpC1 or ClpC2.

In *E. coli*, the adaptor protein ClpS switches the specificity of the chaperone ClpA from ssrA-tagged substrates to N-end rule substrates (Dougan et al, 2002; Erbse et al, 2006). Recent evidence for the cyanobacterium *Synechococcus elongatus* shows homologous adaptor proteins, ClpS1 and ClpS2, interacting with the ClpA orthologue ClpC (Andersson et al, 2009; Stanne et al, 2007). In *Mtb*, the gene Rv1331 is annotated as the adaptor protein ClpS. Based on the high sequence conservation of residues involved in substrate and chaperone binding, *Mtb* ClpS is likely also an adaptor for N-end rule degradation (Lupas & Koretke, 2003; Schuenemann et al, 2009). However, Rv1331 has not been characterized *in vitro* or tested as an interactor of ClpC1 and ClpC2.

Interaction between ClpC1/ClpC2 and ClpS was therefore first tested in a BACTH co-transformation experiment, where a plasmid expressing ClpC1 (or ClpC2) fused to one of the adenylate cyclase subdomains was co-transformed with a plasmid carrying ClpS fused to the other adenylate cyclase subdomain. Figure 3.5 A shows that ClpC1 and ClpS do not interact in this setup (first two rows). However, when ClpS is co-transformed with ClpC2, interaction is evident by the red coloration of the MacConkey reporter agar (Figure 3.5 A, row 3 and 4; in row 4 “T*-ClpC2” means that in wells 1-3 T18-ClpC2 was used and in wells 4-6 T25-ClpC2). These results indicate that ClpS either can only bind to ClpC2 and not ClpC1, or that ClpC1 binding could not be detected in the BACTH test due to steric hindrances of reporter domain placement in case of the much larger ClpC1.

The potential interaction of ClpS with ClpC1 was therefore tested with the recombinantly produced and purified proteins *in vitro*. As mentioned above, in *E. coli*, the adaptor protein ClpS switches the specificity of the chaperone ClpA away from ssrA-tagged substrates to N-end rule substrates (Dougan et al, 2002; Erbse et al, 2006). If ClpS provided a similar specificity switch for ClpC1, then it should inhibit the ClpC1-dependent degradation of ssrA-tagged substrates as well as promote the degradation of an N-end rule model substrate.

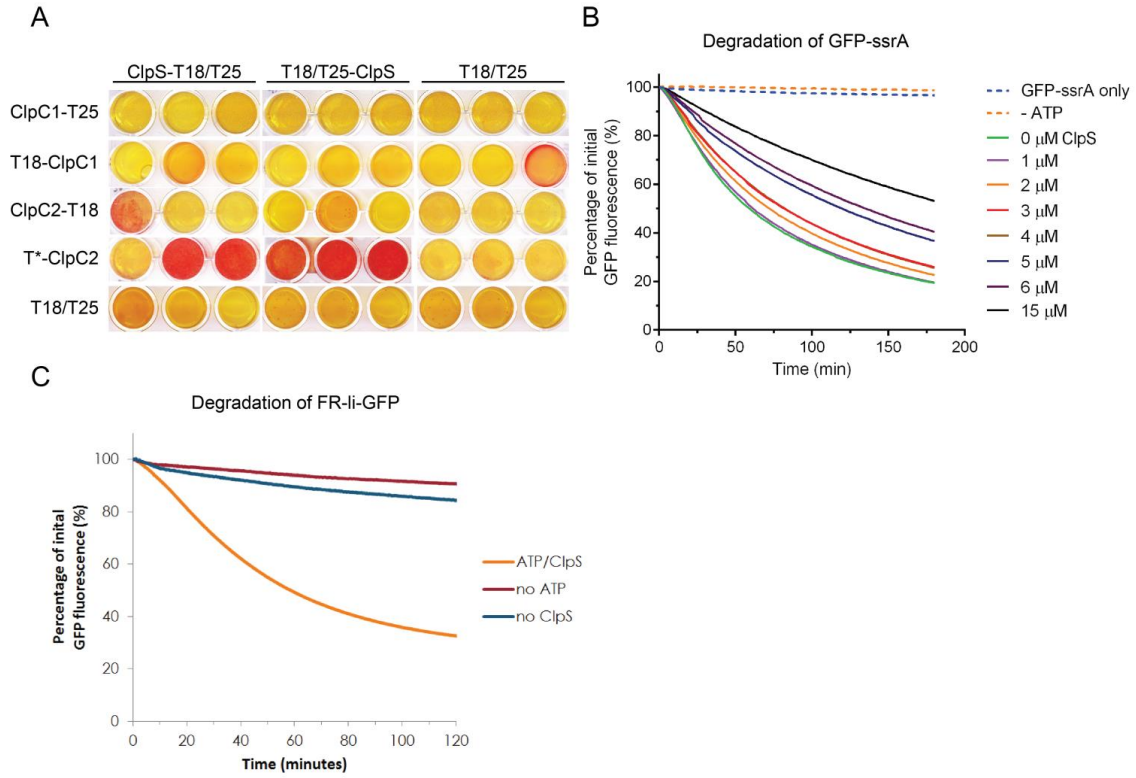


Figure 3.5: ClpS is a binder for ClpC1 and ClpC2. A. Bacterial Two-Hybrid Assay testing for ClpS binding to ClpC1 and ClpC2 on a MacConkey agar matrix of co-transformed vectors carrying the indicated target gene fused to either the T25 or T18 domain. Red coloration of the agar medium indicates successful interaction of the target proteins. Experiments were performed in triplicates. Controls were performed by co-transformation of a vector containing insert with the corresponding vector without insert. In the figure, the indicated ClpC1 and ClpC2 constructs were co-transformed with the corresponding ClpS construct (T25 + T18). For the T*-ClpC2 row, in wells 1-3 T18-ClpC2 was co-transformed with ClpS-T25, and in wells 4-6 T25-ClpC2 with T18-ClpS. **B.** Inhibition of ClpC1-mediated degradation of GFP-ssrA by ClpS. ClpC1-dependent (1 µM) degradation of GFP-ssrA (2 µM) by mature ClpP1P2 (0.5 µM) in the presence of 0-15 µM ClpS (full lines). Negative controls are depicted in dashed lines, and represent the reaction mix lacking ClpC1P1P2 (GFP-ssrA only) and a control lacking ATP (-ATP). Loss of fluorescent signal indicates GFP-ssrA degradation. **C.** ClpS promotes degradation of FR-li-GFP as seen by the loss of GFP signal (orange trace). FR-li-GFP (6 µM) was degraded by the ClpC1P1P2 (1.5 µM ClpC1-hexamer, 0.5 µM mature ClpP1P2 14-mer) in presence of ClpS (10 µM) and 10 mM ATP. Controls are performed with the complete reaction mix in either the absence of ClpS or ATP (black and red traces, respectively).

To test inhibition of ssrA-tagged substrate degradation, we used the intrinsic fluorescence of the model substrate GFP-ssrA, and followed substrate degradation by ClpC1P1P2 over time, monitoring the loss of GFP fluorescence (Figure 3.5 B). Upon titrating in *Mtb* ClpS, the degradation of GFP-ssrA is increasingly inhibited, suggesting that similar to *E. coli* ClpS, *Mtb* ClpS switches the specificity of ClpC1 away from ssrA-tagged substrates. To then test the degradation of N-end rule substrates, we used the N-end rule model substrate FR-li-GFP (which is GFP carrying the N-end rule residue phenylalanine at its N-terminus followed by an arginine and a linker sequence (Erbse et al, 2006)) in a fluorescence-based *in vitro* degradation assay. Figure 3.5 C shows that

ClpC1P1P2-dependent degradation of FR-li-GFP occurs as expected in presence of both ClpS and ATP. The reaction only proceeds to around 30% of the initial GFP concentration, because 20-30% of the FR-li-GFP sample actually lack the first 2 residues which are important for recognition, as was confirmed by mass spectrometry (not shown). Together, these results show that *Mtb* ClpS is indeed a binding partner of the chaperone ClpC1 and changes the specificity of the chaperone from ssrA-tagged to N-end rule substrates.

3.2.6 BACTH co-transformation confirms library screen hits of toxins and antitoxins as binders of ClpC1 and ClpC2

The library screens for ClpC1 and ClpC2 revealed the components of toxin-antitoxin systems, in some cases the toxins and in other cases the antitoxins, as putative interactions partners. Figure 3.4 shows that library genes grouped into “Virulence, detoxification and adaptation” represent 15% and 11% of all hits for ClpC1 and ClpC2 respectively, making them the highest and second-highest, respectively, represented functional group. Strikingly, all the hits in this functional category (with the exception of one gene, MT0196) classify as toxin-antitoxin modules, with toxins as well as antitoxins identified in the screen. The percentages given correspond to 24 unique toxin-antitoxin hits in the ClpC1 screen and 9 unique hits in the ClpC2 screen (Table 3.3 and 3.4).

Mtb features an exceptionally high number of toxin-antitoxin modules (79 annotated), most of which are of the type II, where the toxin as well as the antitoxin components are proteins (Sala et al, 2014). So far, six different families of type II toxin-antitoxin modules were found in the *Mtb* genome: VapBC (50), MazEF (10), RelBE (2), YefM/YoeB (1), HigBA (2), ParDE (2) (Sala et al, 2014). While most toxins are ribonucleases, the families differ in their substrate range (mRNA, tRNA, rRNA,…), their recognition sequences and binding partners (some antitoxins associate to the ribosomal subunits while others do not) (Sala et al, 2014).

In the ClpC1 screen, eight different antitoxins were found, two belonging to the ParD family, and 6 to the VapB family. In the ClpC2 screen seven different antitoxin hits occurred, one of the Rel, five of the VapB family, and YefM. While in both screens a similar number of antitoxin interactors were found, in the ClpC1 screen a markedly higher number of toxins occurred. There, 16 unique hits were found in the toxin class: two of the MazF family and 14 of the VapC family. In the ClpC2 screen, only three toxins were identified, one each of the MazF, the VapC family, and Rv1546, which is not attributed to any family.

Table 3.3: Toxin-antitoxin hits from the ClpC1 screen. Hits overlapping with the ClpC2 screen are colored blue, and toxins with the corresponding antitoxins found in the ClpC2 screen are colored orange. "Confirmed" refers to results of the co-transformation validation experiment described in Figure 3.6.

Locus tag	Function	Protein name	confirmed
Rv1960	Antitoxin	ParD1	yes
MT2201/Rv2142A	Antitoxin	ParD2	yes
Rv0300	Antitoxin	VapB2	no
Rv2550	Antitoxin	VapB20	yes
MT0290/Rv0277A	Antitoxin	VapB25	yes
MT2035/Rv1982A	Antitoxin	VapB36	yes
MT2676/Rv2601A	Antitoxin	VapB41	no
MT3800/Rv3697A	Antitoxin	VapB48	yes
Rv1102	Toxin	MazF3	no
MT2123/Rv2063A	Toxin	MazF7	maybe
Rv1561	Toxin	VapC11	yes
Rv1720	Toxin	VapC12	yes
Rv1838	Toxin	VapC13	yes
Rv2010	Toxin	VapC15	yes
Rv2548	Toxin	VapC19	yes
Rv2863	Toxin	VapC23	maybe
Rv0598	Toxin	VapC27	no
Rv1242	Toxin	VapC33	no
Rv1741	Toxin	VapC34	yes
Rv1962c	Toxin	VapC35	yes
Rv2596	Toxin	VapC40	yes
Rv2872	Toxin	VapC43	yes
Rv3320	Toxin	VapC44	yes
Rv3180	Toxin	VapC49	yes

Although in both screens many members of the Vap family were identified, for only two, VapBC12 and VapBC34, the antitoxin as well as the corresponding toxin were identified (marked orange in Table 3.3 and 3.4), with ClpC1 binding to the toxins VapC12 and VapC34 and ClpC2 binding to the antitoxins VapB12 and B34. The only overlapping toxin-antitoxin hits between the ClpC1 and the ClpC2 screen are the toxins MazF3 (Rv1102c) and VapC49 (Rv3190c) (marked blue in Table 3.3 and 3.4).

For the validation of hits direct BACTH co-transformation experiments were used, where the individual toxin and antitoxin genes were fused to the respective T18 domain as found in the screen, and co-transformed with either ClpC1 or ClpC2, carrying the T25 domain to verify the hit, as well as with the T25 domain without insert to control for false positives. All toxin-antitoxin

hits were tested against ClpC1 as well as ClpC2 to find out whether both proteins have the ability to interact with the same substrates. Co-transformation of toxins/antitoxins with ClpC1/ClpC2 might allow for detection of interactions that were too weak to be picked up in the library screen.

Table 3.4: Toxin-antitoxin hits from the ClpC2 screen. Hits overlapping with the ClpC1 screen are colored blue, and antitoxins with the corresponding toxins found in the ClpC1 screen are colored orange. "Confirmed" refers to results of the co-transformation validation experiment described in Figure 3.6.

Locus Tag	Function	Protein name	confirmed
Rv1247c	Antitoxin	RelB1	no
Rv3357	Antitoxin	YefM	yes
Rv1721c	Antitoxin	VapB12	yes
Rv1952	Antitoxin	VapB14	maybe
Rv1740	Antitoxin	VapB34	no
Rv2104c	Antitoxin	VapB37	yes
MT0987/Rv0959A	Antitoxin	VapB9	yes
Rv1102c	Toxin	MazF3	no
Rv1546	Toxin	Rv1546	yes
Rv3180c	Toxin	VapC49	no

Figure 3.6 shows the results of the co-transformations of the toxin and antitoxin hits with ClpC1 and ClpC2. The appearance of blue colonies on the M63/lactose/X-gal minimal selective reporter medium confirms interactions of proteins, while the absence of colonies indicates no interaction (the co-transformations were grown and plated in triplicates). In the negative control where the toxins/antitoxins in the pUT18/18C vectors were co-transformed with the empty T25 vector, no colony growth was observed, showing that the appearance in the library was not due to interaction with toxins/antitoxins with the T25 domain.

Figure 3.6 A shows co-transformation of hits from the ClpC2 screen (Table 3.4). 4 out of 7 antitoxin hits (blue label) were confirmed for ClpC2, while 6 out of 7 show interaction with ClpC1 as well. The one toxin hit shown in Figure 3.6 A (Rv1546, red label) also shows interaction with ClpC2 and ClpC1. The interaction of antitoxins/toxins with ClpC1 and ClpC2 both indicates that antitoxins likely bind to the region conserved between both proteins.

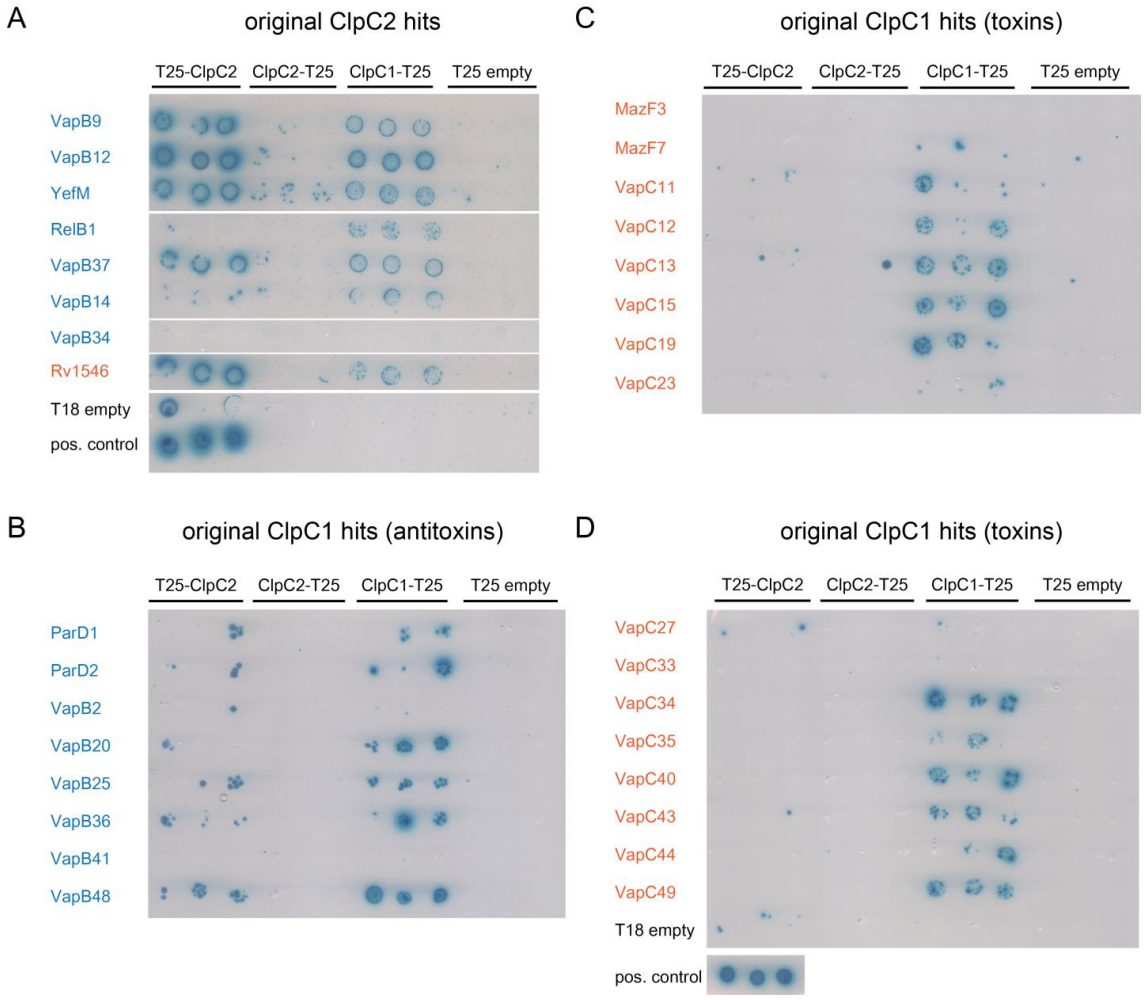


Figure 3.6: Validation of BACTH library hits for toxin-antitoxin complexes. Toxin (red writing) and anti-toxin (blue writing) hits from the ClpC1 and ClpC2 screen were co-transformed with ClpC2 and ClpC1 fused to the T25 domain (T25-ClpC2, ClpC2-T25, ClpC1-T25), or an empty vector expressing only the T25 (T25 empty) as a control. Successful interaction between two proteins results in growth and blue color of colonies on M63/agar supplemented with lactose and X-gal. Co-transformations were grown and plated in triplicates.

Figure 3.6 B-D shows the results of co-transformation of hits from the ClpC1 screen (Table 3.3). Antitoxin hits are labelled in blue writing (Figure 3.6 B), and toxin hits in red writing (Figure 3.6 C and D). 6 out of 8 antitoxin hits were confirmed for ClpC1 interaction. Most of them seem to interact with T25-ClpC2 as well, but only to a small degree as there are only few blue colonies observed. Co-transformation of the toxins (Figure 3.6 C and D) confirms 11 hits of 16 and 2 hits are uncertain, since they show very low blue colony number (MazF7, VapC23). For ClpC2, no interactions with any toxins from the ClpC1 screen are observed.

Interestingly, while all confirmed hits from the ClpC2 screen also interact with ClpC1 (Figure 3.6 A), it seems that for the ClpC1 screen hits only the antitoxins show a tendency of interacting with

ClpC2 in addition to ClpC1 (Figure 3.6 B-D). This indicates that ClpC2 might be more involved with antitoxins than toxins.

In the ClpC2 screen 3 toxins were identified, Rv1546, MazF3 and VapC49, but only the interaction with Rv1546 was confirmed (Table 3.4, Figure 3.6). The toxins MazF3 and VapC49 are the only two hits overlapping between both the ClpC1 and the ClpC2 screens. VapC49 seems to interact only with ClpC1, while for MazF3 interaction was observed with neither ClpC1 nor ClpC2, suggesting that this hit is a false positive.

3.2.7 Antitoxins of the Vap and Rel class are degraded by the ClpC1P1P2 chaperone-protease

In toxin-antitoxin complexes, the antitoxins bind to their respective toxins and thereby inhibit toxin action. Antitoxins are, however, less stable than the toxins. Under certain conditions, antitoxins are rapidly degraded by proteases, allowing toxins to perform their function (Brzozowska & Zielenkiewicz, 2013; Sala et al, 2014). Most of the reports on antitoxin degradation stem from *in vivo* studies in *E. coli*, describing mostly Lon, but also the ClpAP and ClpXP system as the degradation systems responsible (Brzozowska & Zielenkiewicz, 2013). The only gram positive bacterium studied so far was *S. aureus,* which, like *Mtb*, lacks the Lon protease and where the ClpCP complex was shown to degrade antitoxins *in vivo* (Donegan et al, 2010).

In our screen for binders of ClpC1 and ClpC2 we detected 15 antitoxins. Therefore, to test whether antitoxin degradation in *Mtb* is indeed performed by the Clp chaperone protease system, we selected two antitoxin hits for an *in vitro* experiment, VapB37 of the Vap and RelB1 of the RelBE family. For the VapBC37 system a link to *Mtb* latent infection was proposed as the toxin concentration was found elevated in the extracellular space in a nutrient starvation model (Albrethsen et al, 2013). The RelBE1 system was suggested to be involved in *Mtb* persistence, as it was found upregulated in the lung tissues of infected mice, and toxin overexpression enhanced the amount of drug-tolerant persister cells (Singh et al, 2010).

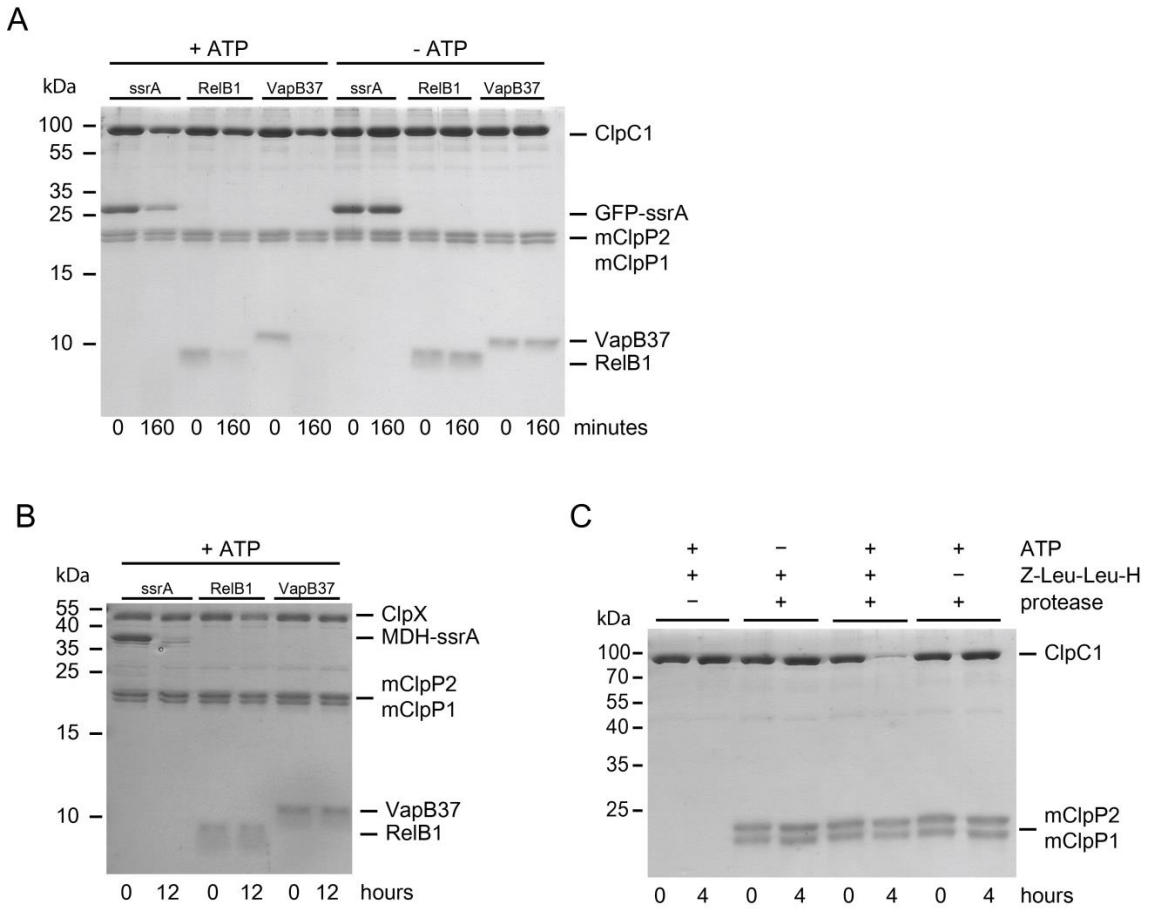


Figure 3.7: *In vitro* degradation of the Antitoxins VapB37 and RelB1. A. ClpC1-dependent degradation of VapB37 and RelB1 by ClpP1P2. ClpC1 (0.5 µM hexamer), mature ClpP1P2 (0.25 µM 14-mer) and substrate (GFP-ssrA: 2 µM; VapB27/RelB1: 15 µM) were incubated at 37°C for 160 minutes in the presence of 1 mM activator peptide Z-Leu-Leu-H, and ATP regeneration system. +ATP/-ATP indicates the presence or absence of ATP (5 mM). Degradation was followed by the disappearance of the substrate band on an SDS-PA gel (18%). **B.** ClpX-dependent degradation of VapB37 and RelB1 by ClpP1P2. ClpX (0.5 µM hexamer), mature ClpP1P2 (0.25 µM 14-mer) and substrate (MDH-ssrA: 2 µM; VapB27/RelB1: 15 µM) were incubated at 37°C for 12 hours in the presence of 1 mM activator peptide Z-Leu-Leu-H and 5 mM ATP and ATP regeneration system. Degradation was followed by the disappearance of the substrate band on SDS-PA gel (18%). **C.** Autodegradation of ClpC1 in presence of activator. Autodegradation of ClpC1 (0.5 µM hexamer) was assessed in the presence or absence of mature ClpP1P2 (0.25 µM 14-mer), 1 mM activator peptide Z-Leu-Leu-H and 5 mM ATP. Degradation was followed by the disappearance of the substrate band on an SDS-PA gel (18%).

VapB37 and RelB1 were recombinantly expressed in *E. coli*, purified and subjected to *in vitro* degradation by the ClpC1P1P2 and ClpXP1P2 chaperone-protease complexes. Figure 3.7 A shows degradation of VapB37 and RelB1 by the ClpC1P1P2 complex after 160 minutes of incubation at 37°C. The bands of the two antitoxins disappear almost completely after the incubation time. The first two lanes show degradation of the model substrate GFP-ssrA as a control. The degradation of the substrates is ATP-dependent, showing that antitoxin degradation is dependent on ClpC1 recognition and unfolding action.

As the *Mtb* Clp system features a second chaperone component, ClpX, we furthermore tested whether antitoxins are degraded when presented to the ClpXP1P2 complex. However, for ClpX-dependent degradation, no antitoxin degradation is observed under conditions where the model substrate MDH-ssrA is degraded (Figure 3.7 B).

A time-trace degradation experiment with both antitoxins (Figure 3.8 A and B) shows that VapB37 and RelB1 are not degraded at the same speed by ClpC1P1P2. While Vap37 is completely degraded after 120 minutes, for RelB1 a slight band is still visible at 160 minutes, indicating differences in either recognition or unfolding efficiency by ClpC1.

Figure 3.7 A and Figure 3.8 A and B confirm the results of the BACTH co-transformation experiment, showing that ClpC1 recognizes two model antitoxins, VapB37 and RelB1 and mediates their degradation by the ClpP1P2 protease.

During degradation assays with the antitioxins (Figure 3.7 A and Figure 3.8), we observed a slight ATP-dependent diminishing of the ClpC1 chaperone band itself, which was reminiscent of the autodegradation observed for *E. coli* ClpA. *E. coli* ClpA performs autodegradation as a regulative mechanism in the absence of substrate (Gottesman et al, 1990; Maglica et al, 2008). To test for a similar behavior of ClpC1 in the absence of substrate, we tested the degradation of ClpC1 in the absence of substrate. Figure 3.7 C shows that ClpC1 indeed autodegrades, as the presence of ATP as well as the protease is necessary. Additionally, it requires the presence of the activator peptide, a compound that stabilizes the interaction of the protease with the chaperone *in vitro* (Schmitz et al, 2014). As the experiments in Chapter 2 of this thesis were performed in absence of activator peptide, this effect was not observed before.

Taken together, these results indicate that in *Mtb* antitoxins are degraded by the Clp chaperone protease system, and that ClpC1P1P2 rather than ClpXP1P2 is the chaperone-protease responsible for antitoxin degradation.

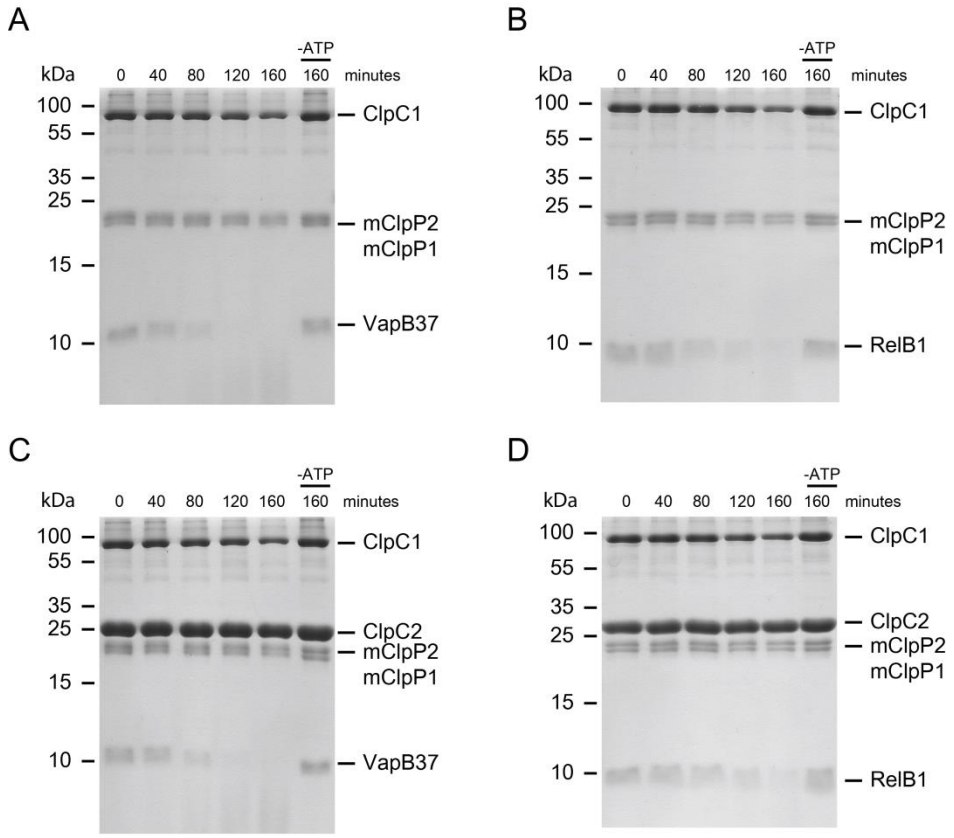


Figure 3.8: *In vitro* degradation time traces of VapB37 and RelB1 by ClpC1P1P2 in presence and absence of ClpC2. ClpC1-dependent (0.5 µM hexamer) degradation of VapB37/RelB1 (15 µM) by mature ClpP1P2 (0.25 µM 14-mer) in the presence of 1 mM activator peptide Z-Leu-Leu-H and 5 mM ATP (except in the control lacking ATP (-ATP)). Degradation was followed over time (0-160 minutes; timepoints indicated on top of the gel) by the disappearance of the substrate band on SDS-PA gel (18%). **A.** Degradation of VapB37 in the absence of ClpC2. **B.** Degradation of RelB1 in the absence of ClpC2. **C.** Degradation of VapB37 in the presence of ClpC2 (15 µM). **D.** Degradation of RelB1 in the presence of ClpC2 (15 µM)

3.2.8 The putative adaptor ClpC2 does not significantly influence antitoxin degradation.

We established the model antitoxins VapB37 and RelB1 as degradation substrates of the ClpC1P1P2 chaperone-protease. However, as most of the antitoxins show not only interaction with ClpC1, but also ClpC2 in the BACTH system, ClpC2 might equally interact with antitoxins. ClpC2 might, for example, modulate degradation of antitoxins by ClpC1P1P2 in a similar manner as an adaptor protein. Binding of antitoxins to ClpC2 could for example compete with ClpC1-dependent degradation in an anti-adaptor like function.

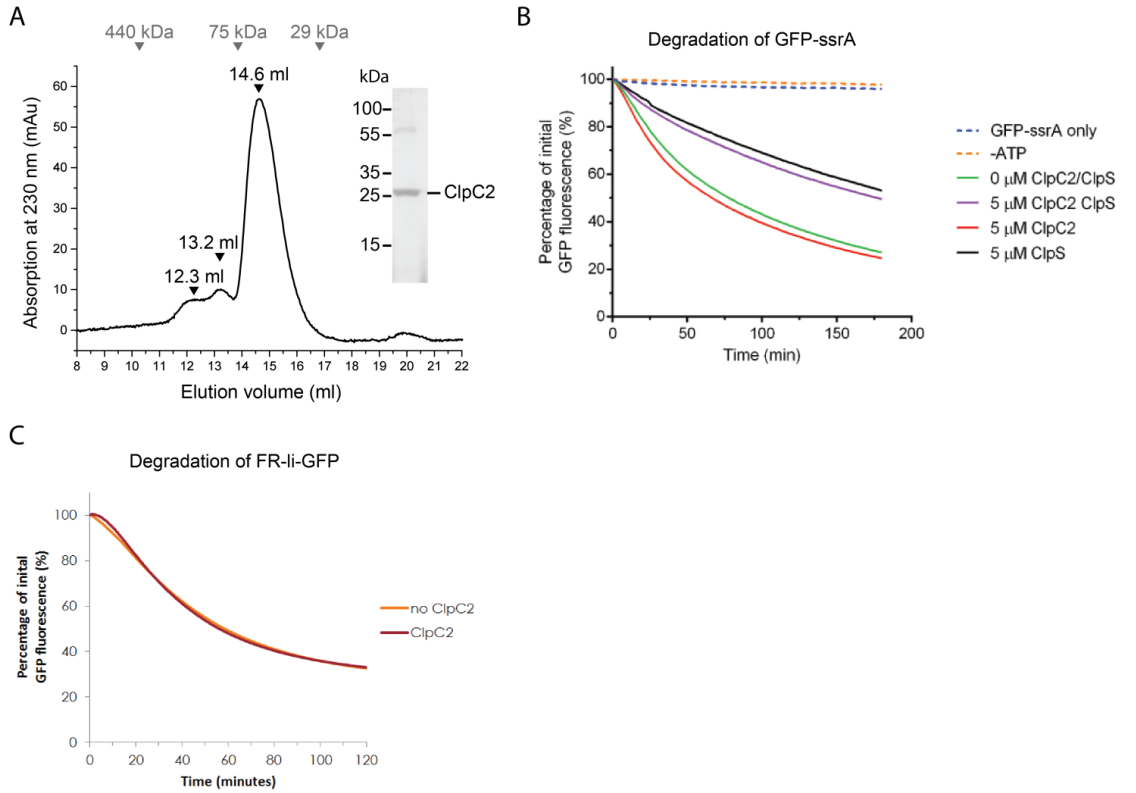


Figure 3.9: **Analysis of purified ClpC2. A**. Analytical gel filtration of ClpC2 (100 µM) on a Superdex 200, 24 ml column. The main peak of ClpC2 elution is marked at 14.6 ml (corresponding to an approximate mass weight of 60 kDa), with minor peaks eluting at 12.3 ml and 13.2 ml. Molecular weight standards (440 kDa, 75 kDa, 29 kDa) are indicated above the graph. An SDS-PA gel showing the final purification step is depicted. **B.** ClpC2 (5 µM) has no effect of GFP-ssrA (2 µM) degradation by ClpC1P1P2 (ClpC1: 1 µM hexamer, ClpP1P2: 0.5 µM hexamer) (green line: - ClpC2; red line: + 5 µM ClpC2). Furthermore, ClpC2 does not influence the partial inhibition of GFP-ssrA degradation by ClpS (black line; + 5 µM ClpS; violet line: + 5 µM ClpS/+ 5 µM ClpC2). A control was performed with the complete reaction mix in absence of ATP (-ATP) and with GFP-ssrA only in the absence of other protein components (GFP-ssrA) only. **C.** ClpC2 (10 µM) does not influence ClpS-dependent degradation of FR-li-GFP (red trace). FR-li-GFP (6 µM) was degraded by the ClpC1P1P2 (1.5 µM ClpC1-hexamer, 0.5 µM mature ClpP1P2 14-mer) in presence of 10mM ATP and ClpS (10 µM) (orange trace).

To test the influence of ClpC2 on the degradation of antitoxins by the ClpC1P1P2 complex, ClpC2 was recombinantly expressed in *E. coli* and purified. Analytical size exclusion chromatography of the purified ClpC2 protein on a 24 ml Superdex 200 column (Figure 3.9 A) shows a main peak with an elution volume of 14.6 ml, which corresponds to an approximate mass weight of 60 kDa under the experimental conditions. Given the 26 kDa molecular mass of ClpC2, this elution volume suggests a dimeric ClpC2 assembly state. Furthermore, there are two peaks of smaller size at lower elution volumes, namely at 13.2 and 12.3 ml, yielding approximate mass weights of 115 and 170 kDa, that could correspond to higher assembly states of the dimer, namely a tetrameric species or hexameric species, respectively. This indicates that ClpC2, even if dimeric under the experimental conditions, can assemble to higher order oligomers. The results of the analytical

gel filtration furthermore agree with results from the BACTH library screen of ClpC2, where we identified ClpC2 as a binder of ClpC2 (Appendix Table 3; Section 3.2.4), indicating that ClpC2 oligomerizes under *in vivo* conditions in *E. coli*, as well as *in vitro.*

As we observed interaction of ClpC2 and ClpS in a BACTH setup (Figure 3.5 A), we additionally tested whether ClpC2 could influence ClpS action by titrating ClpS away from ClpC1. This should then reactivate degradation of ssrA-tagged substrates on the one hand and inhibit degradation of N-end rule substrates on the other. However, the inhibition of GFP-ssrA degradation is not reversed by adding ClpS (compare black and purple time traces in Figure 3.9 B). Neither is the degradation of FR-li-GFP that is observed in presence of ClpS inhibited by addition of ClpS (compare red and orange time traces in Figure 3.9 C). These results suggest that on the level of ClpS binding, ClpC2 cannot compete with ClpC1.

It is nevertheless possible that the direct binding of certain substrates to the ClpC1 N-domain is competitively inhibited by ClpC2. For this reason, to test for eventual competitive behavior between ClpC1 and ClpC2 for antitoxins, we performed the same time-trace experiments as depicted in Figure 3.8 A and B, in the presence of an excess of ClpC2 (Figure 3.8 C and D). However, under the conditions tested, no significant influence of ClpC2 on the dynamics of ClpC1-dependent antitoxin degradation is observed. For VapB37 degradation, no change in band intensity is observed, while RelB1 degradation a slight slowing of degradation can be seen.

3.3 Materials and Methods

Alignments

Protein alignments were performed by first extracting the protein sequences from the UniProt database (http://www.uniprot.org/) and subsequent alignment using the the Clustal Omega algorithm (Goujon et al, 2010; Sievers et al, 2011). Visualisation was performed using Jalview (Waterhouse et al, 2009). The UniProt identifiers for *Mtb* ClpC1 and Clp2 are P9WPC9 and P9WPC7 respectively, and the identifiers for ClpC2 homologous are given in the respective figure legend.

Cloning, expression and protein purification

ClpC1, ClpX, ClpP1 and ClpP2, MDH-ssrA and GFP-ssrA were cloned and purified as described in Chapter 2. FR-linker-GFP (FR-li-GFP) was available in the lab and prepared as published (Cranz-Mileva et al, 2008).The mature ClpP1P2 complex was assembled and processed as described in Chapter 2.

The *clpC2, clpS, vapB37* and *relB1* genes were amplified by PCR from *M. tuberculosis* H37Rv genomic DNA with Phusion DNA polymerase (New England Biolabs). The *clpC2* gene was then ligated into a p7XC3H FX vector (Geertsma & Dutzler, 2011), including a stop codon to produce an untagged ClpC2 protein. *ClpS, vapB37* and *relB1* were ligated into a modified pET20 vector, resulting in the addition of a His_{10}-tag followed by a thioredoxin (Trx) sequence and a TEV (tobacco etch virus) protease cleavage site to the N-terminus of the proteins (His_{10}-Trx-TEV-protein of interest).

The plasmids were transformed into *E. coli* Rosetta (DE3) (Invitrogen) cells and grown in LB medium (ClpC2, ClpS) or 2xYT medium (VapB37, RelB1) supplemented with the respective antibiotic. Protein expression was induced at an OD_{600} of 0.8 by addition of 0.1 mM IPTG. The expression was performed overnight at a temperature 20°C for ClpC2 and 25°C for ClpS, VapB37 and RelB1. The cells were harvested and resuspended in respective buffers supplemented with Complete protease inhibitor cocktail (Roche), 1 mM PMSF, 0.03 U/µL DNase and 5 mM MgCl2 (ClpS: Buffer L (50 mM Hepes-KOH, pH 7.5, 300 mM KCl, 10 % glycerol); Antitoxins Buffer N: 50 mM Hepes-KOH, pH 7.5, 150 mM KCl). The cells were cracked using a mircofluidizer device (M-110L; Microfluidics) at 50 PSI and the supernatant containing the soluble protein of interest separated from the insoluble proteins and cell debris by ultracentrifugation (Ti-70 rotor, 45000 rpm). ClpS, VapB37 and RelB1 were purified by standard Ni-NTA affinity chromatography using the respective breaking buffers supplemented with imidazole, subsequent TEV cleavage and a

reverse Ni-NTA step, followed by size exclusion chromatography on a Superdex 75 (16/60) column (GE healthcare) equilibrated in Buffer L (50 mM Hepes-KOH, pH 7.5, 300 mM KCl, 10 % glycerol) for ClpS and in Buffer Q (50 mM Hepes-KOH, pH 7.5, 150 mM KCl) for VapB37 and RelB1. For storage, 10% glycerol were added to the VapB37 and RelB1 samples. For the purification of ClpC2, the cells were first resuspended in Buffer LS (50 mM Hepes-NaOH, pH7.5, 100 mM NaCl, 2 mM EDTA, 10% glycerol) and loaded on an anion exchange Fast Flow Q column (60ml, GE Healthcare). ClpC2 was eluted with gradient to buffer HS (50 mM Hepes pH 7.5, 1 M KCl, 2 mM EDTA, 10 % glycerol), ClpC2 containing fractions were pooled and precipitated with 50% ammonium sulfate (AS) for 2.5 hours at 4°C. The precipitate was pelleted (SS-34; 15k rpm, 20 mins), resuspended and dialysed against buffer LS. Next, the sample was loaded on a 23 mL Source 30Q anion exchange column (GE Healthcare), and eluted with a gradient to buffer HS. The fractions containing ClpC2 were pooled and concentrated using Vivaspin centrifugal centrifugation devices (cutoff 10 kDa) and injected on a Superdex 200 gel filtration column (320 ml, GE Healthcare), equilibrated in Buffer LS. ClpC2 containing fractions were pooled and precipitated with 30% ammonium sulfate to remove an ATPase impurity that co-precipitated when using a higher concentration of ammonium sulfate. The resulting pellet was then finally dialysed into Buffer (50 mM Hepes-KOH, pH 7.5, 150 mM KCl, 10% glycerol). The correct size for all purified proteins was verified by mass spectrometry and proteins were stored at -20°C.

Analytical gel filtration

Analytical gel filtration of 100 µM ClpC2 was performed at room temperature on a Superdex 200 10/300 GL column (GE Healthcare, 24 ml) in Buffer L (50 mM Hepes-KOH, pH 7.5, 300 mM KCl, 10 % glycerol) at a flow rate of 0.5 ml/min on an ÄKTA Purifier System. 30 µl of the sample were injected onto the column and proteins were detected by absorption at 230 nm. Prior to loading, the ClpC2 sample was centrifuged for 5 minutes at 14'000 rpm in an Eppendorf table-top centrifuge. The column was calibrated with a set of standard proteins (GE Healthcare Gel filtration Calibration Kit), from which a calibration curve was calculated.

Degradation of substrate proteins by ClpC1P1P2 and ClpXP1P2

Degradation of substrate proteins in gel based assays was performed in Buffer J (50 mM Hepes-KOH, pH 7.5, 150 mM KCl, 15% glycerol, 20 mM MgCl2) for ClpX -dependent degradation and in Buffer K (50 mM Hepes-KOH, pH 7, 150 mM KCl, 15% glycerol, 20 mM MgCl2) for ClpC1-dependent degradation. All degradations were performed at 37°C, in the presence of 1 mM activator (Benzyloxycarbonyl-L-Leucyl-L-Leucinal (Z-Leu-Leu-H); from PeptaNova). ClpC1 or ClpX (0.5 µM hexamer), mature ClpP1P2 (0.25 µM 14-mer; produced as described in chapter 2.4; preas-

sembled at room temperature for one hour) and substrate (GFP-ssrA/MDH-ssrA: 2 µM; VapB27/RelB1: 15 µM) were incubated at 37°C for the indicated time span in the presence of 1mM DTT and ATP regeneration system (20 mM phosphocreatine and 1 U/ml creatine phosphokinase). Unless otherwise indicated in control reactions, ATP (5 mM) was used in the reaction. In case of reactions containing ClpC2 (15 µM), ClpC2 was preincubated with the substrate and the reaction mix for 15 minutes. The reaction was stopped at the indicated time points by the addition of Laemmli buffer. Before loading on a 18% SDS-PA gel, the samples were heated for 10 minutes at 95°C. SDS-PA gels containing antitoxins were first run for 30 minutes at 90V, and then approximately one hour at 200V.

For fluorescence based degradation assays, the degradation of GFP-ssrA (*Mtb* ssrA tag sequence: AADSHQRDYALAA) and FR-li-GFP (Phe-Arg-linker-GFP)) was followed by loss of fluorescence signal in a BioTek Synergy 2 plate reader in Corning non-binding 96-well half area assay plates in 50 µl reaction volume. The excitation wavelength was 360/40 nm and the emission wavelength of 528/20 nm. For experiments measuring GFP-ssrA degradation a tungsten light source was used (50% optics position, sensitivity: 90) at 23 °C and for experiments measuring FR-li-GFP degradation a Xenon light souce was used (50% optics position, sensitivity: 45) at 30°C. The mixed samples were incubated for 10 minutes before the reaction was started by addition of ATP. GFP-ssrA (2 µM) degradation was performed by ClpC1ClpP1P2 (1 µM ClpC1 hexamer; 0.5 µM preassembled mature ClpP1P2 14-mer) in the presence of 5 mM ATP, 1 mM DTT and ATP regeneration system (20 mM phosphocreatine and 1 U/ml creatine phosphokinase) in Buffer K. ClpC2 and ClpS were added in concentrations as indicated. Degradation of FR-li-GFP was performed with 1.5 µM ClpC1-hexamer, 0.5 µM mature, assembled ClpP1P2 14-mer, 10 µM ClpS, and 6 µM FR-li-GFP, in the presence of 10 mM ATP, 1mM DTT, and ATP regeneration system (20 mM phosphocreatine and 1 U/ml creatine phosphokinase) and 1 mM activator (Z-Leu-Leu-H) in Buffer K for the indicated time frame.

BACTH (bacterial adenylate cyclase two hybrid) methods

Cloning

For the BACTH assays, the genes of interest (GoI) were cloned into the respective bacterial two hybrid vectors provided by Euromedex, that contained either the *B. pertussis* adenylate cyclase subdomain T18 or T25 for fusion to the GoI (Battesti & Bouveret, 2012; Karimova et al, 1998). The respective vectors were pUT18 (GoI-T18), pUT18C (T18-GoI), pKT25 (T25-GoI) and pKNT25 (GoI-T25). Vectors harboring the T18 domain carry ampicillin resistance, those harboring the T25 domain a kanamycin resistance. For the construction of plasmids not being part of the library,

the respective genes were amplified by PCR from *M. tuberculosis* H37Rv genomic DNA with Phusion DNA polymerase (New England Biolabs) and combined with the respective vectors using Gibson Assembly method (Gibson et al, 2009). In the constructs used in this study a linker was introduced by PCR that corresponded to 8 amino acids (GGSGSGSG) in the translated protein between the protein of interest and the fused adenylate cyclase domain.

The *Mtb* ORF library was provided from BEI resources, as a collection of *E. coli* glycerol stocks, each one containing a pDONR221 Gateway® vector containing one library gene. The library was prepared by M. Ziemski, by growing the respective cells in LB, extracting the plasmids and subsequently transferring the inserts from the pDONR221 vectors into Gateway compatible BACTH (created by M. Ziemski) vectors using a Clonase Master Mix for LR recombination (Invitrogen). The library was cloned into both the pUT18 and pUT18C vectors. The genes of the library were pooled into 4 pools of roughly ~ 1000 genes each. Constructs used for the validation of library screen hits were prepared in the same way with the exception that they were not pooled.

BACTH co-transformation tests performed with proteins of the Clp system.

For the interaction assays between components of the Clp system (Figure 3.3 and 3.5), the experiments were performed as follows. From each of the two indicated plasmid 50 ng were co-transformed into a 50 µl aliquot of chemo-comptent *E. coli* BTH101 cells (Euromedex) (F', cya-99, araD139, galE15, galK16, rpsL1 (Str^R), hsdR2, mcrA1, mcrB1, relA1). The cells were incubated 30 mins on ice, heat-shocked at 42°C for 30 seconds, and incubated for 2 minutes on ice, before 950 µL of LB medium containing 1 % glucose was added and cells were recovered for 1 hour at 37°C. After recovery, 150 µL of the cell suspesion was plated on LB-agar plates supplementes with ampicillin (100 µg/ml) and kanamycin (50 µg/ml). After incubation of the plates for 48h at 30°C, five to six clones of each plate were picked an inoculated in triplicates in 1ml LB-Amp/Kan and 0.5 mM IPTG each inside a 96-deep-well block sealed with Breathe-Easy membrane. Incubation occurred over-night in a shaker (160 rpm) with the deep-well-blocks tilted to an approximate 45° position. Subsequently, 30 µL of the liquid cultures were transferred on MacConkey medium (Sigma Aldrich) containing Amp, Kan and 1% maltose, and incubated at 30°C for 24 hours. Plates were scanned and the picture was enhanced with digital software (Adobe Photoshop) by setting brightness to +75, contrast to +50, hue to +25 - +30, and saturation to +50. As a positive control, pKT25-zip and pUT18-zip plasmids were co-transformed (Battesti & Bouveret, 2012).

For the co-transformation assays used for verification of library hits, the respective hits were cloned out of the library as described above. About 30 ng of toxin/antitoxin and ClpC1/ClpC2 containing plasmid were co-transformed into 10 µl chemo-competent *E. coli* BTH101 *ΔclpXP* cells (provided by M. Ziemski), in 96-well PCR plates. The cells were incubated 30 mins on ice, heat-shocked at 42°C for 30 seconds, and incubated for 2 minutes on ice, before 180 µL SOC medium was added and cells allowed to recover for 1 hour at 37°C. 20 µL of the recovered mix was added in triplicates to 500 µL LB-Amp/Kan in a 96-deep well block, which was incubated for 16 hours at 30°C in a tilted position. Following, 500 µL of expression inducing LB-Amp/Kan/IPTG medium were added to a final IPTG concentration of 0.5 mM and incubation was continued for 8 hours. 2.5 µL of every well were then plated on M63/lactose agar supplementd with Amp (50 µg/ml) and Kan (30 µg/ml) and plates incubated for 2-3 days at 30°C.

BACTH library screen

For the BACTH library screen, the respective bait proteins were first transformed into *E. coli* BTH101 *ΔclpXP* cells, and electro-competent cells containing the respective bait proteins were prepared (Sambrook & Russel, 2001), and the resulting cell suspensions were diluted to a final concentration of 2.5 x 10^{10} cells/ml. 1 µl from each of the 4 library plasmid pools (100 ng/µl) were added to 50 µL electro-competent *ΔclpXP* BHT101 cells containing the bait plasmid and incubated on ice for 5 minutes, before they were transformed by electroporation. Cells were recovered in 950 µL LB-medium for 1.5 hours at 30°C. The OD_{600} of the cell suspension was measured and cells were diluted to a final concentration of 1.33 x 10^7 cells/ml in M63 medium. Of these, 75 µl each were plated on two M63/maltose plates (square petri dishes, 120 mm x 120 mm) supplemented with Amp (50 µg/ml) and Kan (30 µg/ml), resulting in 2x10^6 total cells plated per transformed pool. After 5 days of incubation at 30°C, blue colonies were picked and re-streaked on fresh M63/maltose/Amp/Kan plates and grown for 1-2 days at 30° C to remove eventual false positive hits (about 10% of hits did not grow after restreaking). The positive hits were then picked and amplified by colony PCR using primers specific for the prey plasmids and OneTaq GC Polymerase Master Mix (NEB). The PCR products were checked on 1% agarose gels and subsequently sent for sequencing to GATC.

The ClpC1 screen was performed with ClpC1-T25 as bait against prey libraries harbouring the T18 domain either at the N-terminus of the target protein (prey vector pUT18C) or at the C-terminus of the target protein (pUT18). The ClpC2 screen was performed with either ClpC2-T25 (bait vector pKNT25) or T25-ClpC2 (bait vector pKT25) as bait against a prey library where the the T18 domain is located at the C-terminus of the target protein (pUT18).

Table 3.5: Recipes

MacConkey agar plates	1.7% (w/v) peptone, 0.3% (w/v) Proteose peptone, 1% (w/v) lactose, 0.15% (w/v) Bile Salts No. 3, 0.5% (w/v) Sodium Chloride, 1.35% (w/v) Agar, 0.03 g/L Neutral Red, 0.001 g/L Crystal Violet
M63 medium	100 mM KH_2PO_4, 15 mM $(NH_4)_2SO_4$, 500 ug/L $FeSO_4$ x 7 H_2O, 1 mM MgSO4 x 7 H2O, pH 7
M63/maltose plates	M63 medium supplemented with 0.2% maltose, 30 µg/mL kanamycin, 50 µg/mL ampicillin, 0.5 µg/mL Vitmain B1, 0.5 mM IPTG, 40 µg/mL X-Gal and 1.5% agar
Luria-Bertani (LB) medium	1% (w/v) NaCl, 1% (w/v) tryptone, 0.5% (w/v) yeast extract in H_2O
2xYT medium/litre	16g tryptone, 10g yeast extract, 5g NaCl

3.4 Discussion

3.4.1 BACTH library screens for ClpC1 and ClpC2 reveal toxin-antitoxin systems as binders

To gain a better understanding of the functional importance of the *Mtb* ClpC1P1P2 chaperone-protease complex, expansion of the knowledge of its interaction partners, the classes of recruited degradation substrates as well as potential regulatory factors, is important. Therefore, to identify new interaction partners, a bacterial two-hybrid screen on the *Mtb* proteome was carried out with ClpC1 and its partial homolog ClpC2 as bait proteins. The homology of ClpC2 to the N-terminal domain of ClpC1, which is involved in substrate and interactor binding, suggested that these two proteins share a part of their interactor profile and indeed, in addition to 18 protein hits that overlap between ClpC1 and ClpC2 (out of 167 and 93 unique protein hits, respectively), 15% and 11% of hits of the ClpC1 and the ClpC2 screen belong to the same functional group of toxin-antitoxin (TA) systems. While more than half of the 18 directly overlapping hits are hypothetical proteins, TA systems present a coherent group of hits belonging to a distinct functional class. TA systems are small two-component systems, composed of a ribonuclease toxin whose action is blocked by antitoxin binding. *Mtb* features an unusually high number of TA systems (79) in its genome and toxin action has been linked to the persistence and pathogenicity of *Mtb* (Sala et al, 2014). The function of TA systems together with their comparatively high occurrence in the library screen for ClpC1 and ClpC2 makes them a prime target for further investigation.

Interestingly, TA system hits occurring in the ClpC2 screen were largely antitoxins (7 out of 10), while for the chaperone ClpC1 antitoxins comprised about one third of the hits (8 of 24), with toxin hits making up the rest. When the respective TA hits of both screens were co-transformed with ClpC1 and ClpC2 each in a validation experiment, ClpC1 showed interaction with toxins as well as antitoxins from both screens, but ClpC2 almost exclusively interacts with antitoxins. This suggests that antitoxin interaction is mediated by the homologous region between the ClpC1 N-domain and ClpC2. Toxin interaction with only ClpC1 might occur due to specific, not conserved residues on the N-domain, or alternatively toxin interaction requires additional interfaces of the D1 or D2 AAA modules of ClpC1. On the other hand, toxins might likewise interact with ClpC2 *in vivo,* but in the context of the BACTH the presence of the reporter domain inhibits binding.

Interaction of antitoxins with ClpC1 suggests antitoxins as substrates of the ClpC1P1P2 complex, as antitoxin removal by chaperone-protease degradation is the mechanism that permits toxin

action (Brzozowska & Zielenkiewicz, 2013). *In vitro* degradation experiments showed that both of the antitoxins chosen, VapB37 and RelB1, are indeed substrates for the ClpC1P1P2 complex, while degradation by ClpXP1P2, the alternative Clp protease assembly present in *Mtb*, does not occur. The main difference between the ClpX and ClpC1 chaperones, apart from the number of AAA modules, lies in their N-terminal domains. This, together with the observation that antitoxins were found with the N-terminal domain homologue ClpC2, strongly suggests that antitoxin recognition is mediated by the ClpC1 N-domain. In addition, substrate classes that are generally recognized by both chaperones, like ssrA-tagged substrates, are usually directly recognized by the pore loop regions, without the mediation of the N-terminal domain (Farrell et al, 2007; Hinnerwisch et al, 2005a; Martin et al, 2008a). In fact, for *E. coli* ClpAP it is has been shown that deletion of the ClpA N-domains leads to even more efficient recruitment of an ssrA-tagged model substrate (Cranz-Mileva et al, 2008). To unambiguously test whether antitoxins are indeed recruited by the ClpC1 N-terminal domain, a similar approach could be taken, where *in vitro* degradation assays are performed with a ClpC1 variant lacking the N-terminal domain.

The degradation of the antitoxins tested, VapB37 and RelB1, occurred on slightly different time scales. While VapB37 was completely degraded after 120 minutes, RelB1 still showed a faint residual band on SDS-PA gel after 160 minutes of degradation. These differences in antitoxin degradation efficiency are more likely due to the recognition by ClpC1 rather than unfolding capability, as antitoxins are usually at least partially unstructured, especially in the absence of the toxin, and therefore generally good degradation substrates (Brzozowska & Zielenkiewicz, 2013). Preliminary sequence analysis of antitoxin hits from the ClpC2 screen did not reveal obvious conserved sequence motifs on the N- or C-termini of the proteins, suggesting that an unstructured terminal region is responsible for recognition rather than a conserved sequence motif. However, as we did not assess the secondary structure of the purified antitoxin, performing CD spectroscopy could provide information as to a possible different secondary structure of VapB37 and RelB1.

Antitoxins were also identified as interactors of ClpC2, however this protein lacks an ATPase domain as well as a known interaction motifs for protease interaction. One possible role for ClpC2 that requires antitoxin binding could be as a modulator of ClpC1-mediated antitoxin degradation. Potential hypotheses were that ClpC2 binds to antitoxins via its region conserved with ClpC1, and then either delivers the substrate more efficiently to ClpC1 employing its N-terminal stretch which would result in increased degradation speed, or it could sequester antitoxins and thereby competitively inhibit degradation, resulting in decreased degradation speed. However,

addition of ClpC2 to ClpC1-mediated antitoxin degradations did not show a significant effect on the time courses, neither for VapB37 nor for RelB1. This observation indicates, at least under the chosen *in vitro* conditions, that the binding of antitoxins to ClpC1 occurs much more efficiently than to ClpC2. It is possible that under different reaction conditions or in presence of additional factors the relative affinities might be different. Optimization of buffer conditions in respect to ionic strength, pH or glycerol content should be systematically tested. For a more efficient read-out of degradation than SDS-PA gel bands, a fluorophore could be coupled to the substrates for utilization of a continuous spectrophotometric readout.

Summarizing, we showed that two antitoxins of different functional families are degraded by the ClpC1P1P2 complex. The implication of the VapBC37 system in the establishment of latent infection and of the RelBE1 system in persistence of *Mtb* (Albrethsen et al, 2013; Singh et al, 2010) stresses the importance of degradation of the respective antitoxins by the Clp chaperone-protease and its possible influence on the virulence of *Mtb.* Interestingly, in addition to the antitoxin substrates, the BACTH screen also identified the toxin components as interactors of ClpC1, indicating that the ClpC1P1P2 particle is likewise responsible for degradation of the toxin part of the TA systems. In TA systems, the toxin is far more stable than the labile antitoxin and less prone to degradation (Brzozowska & Zielenkiewicz, 2013; Sala et al, 2014), however even toxins have to be removed from the cellular environment eventually. To follow up this line of thought, recombinant production of the toxin in complex with its antitoxin and subsequent investigation of degradation should answer that question. Should the hypothesis prove correct, antitoxins will be degraded preferentially by the ClpC1P1P2 complex, and the stably folded toxin will only be degraded later.

Other functional categories contain potentially interesting hits, namely “Information Pathways” and “Regulatory Proteins”, as in these categories DNA binding proteins and regulators are found and the *Mtb* Clp system was already implicated in the regulatory removal of several transcriptional regulators. It degrades for example its own transcriptional activator ClgR (Laederach et al, 2014), the anti-sigmaE factor RseA (Barik et al, 2010) and the essential transcriptional repressor WhiB1, whose controlled degradation could be one of the reasons why the Clp system is essential in *Mtb* (Raju et al, 2014). Given the importance of regulatory degradation, it would be worthwhile to identify further substrates involved in regulative processes, for example in stress response, which is important for survival of *Mtb* under hostile conditions. One hit, RsbW (Rv3287c/MT3386) is the anti-sigma sigma factor for SigmaF, an alternative sigma factor involved in the stress response of *Mtb* to e.g. antibiotic or nutrient starvation (Sachdeva et al,

2010), while another hit, the transcriptional repressor LexA regulates a part of the *Mtb* DNA stress response (Smollett et al, 2012). As Clp-dependent degradation of the anti-sigma factor RseA was already described (Barik et al, 2010), the anti-sigma factor RsbW might also be a true degradation target. Likewise, LexA was described in *E. coli* as a substrate of the ClpXP system under stress conditions (Neher et al, 2003), indicating that it might be similarly regulated in *Mtb* by the ClpC1P1P2 complex.

Interestingly, none of the degradation substrates described so far for the Clp system appeared in the BACTH library screen with ClpC1. For some of these substrates, e.g. WhiB1 and CarD, the identity of the chaperone involved in the degradation, ClpX or ClpC1, was not investigated (Raju et al, 2014). These substrates could therefore be exclusively recognized by ClpX, or reporter domain placement in the context of the BACTH system could obstruct the respective ClpC1/substrate interaction motifs. The latter could for example be the reason why the ClpC1-dependent substrate RseA was not found in the BACTH library screen (Barik et al, 2010). Alternatively, degradation of certain substrates might require an *Mtb* adaptor protein or additional modification like phosphorylation that does not occur in the *E. coli* cells in which the screen was performed.

In the BACTH screen, about half of the hits of the ClpC1 and ClpC2 screen were hypothetical proteins, either originating from the *Mtb* strain H37Rv or CDC1551, for which mostly no putative function was annotated. In addition, the CDC1551 hypothetical proteins were often very small ORFs encoding for 40-90 amino acid long products found in the intergenic region between genes annotated with a function. Unfortunately, the 4-5 hits that were most abundantly observed in the ClpC1 and ClpC2 screens (Table 3.1 and 3.2), i.e. where each hit occurred in more than 10 sequenced clones, are almost exclusively hypothetical proteins. The number of clones a certain hit was observed in gives a rough estimation of the strength of the interaction, but is of course influenced strongly by how well a certain protein is expressed in a cell and if and how attachment of the reporter domain affects its binding behavior. Two of the proteins observed in high abundance are for example confirmed false positive hits. The only annotated protein in Table 3.1, the DNA integrity scanning protein DisA, was also found in screens performed with members of the Pup-Proteasome system and could be a common binder or a false positive hit (personal communication, M. Ziemski).

Some general caution needs to be taken due to the fact that the BACTH screen is performed in *E. coli*, which also has a Clp system. This means that proteins expressed by *E. coli* might influence

the outcome of the screen, and interaction between bait and prey proteins might be mediated by a third party. Another issue is that ClpC1 is an active unfoldase that might actually unfold potential substrate interactors in the cell. However, this issue is maybe not dramatic, as we observed numerous hits with antitoxin substrates, indicating that the interaction of the reporter domains prevents efficient unfolding by ClpC1. However, one might increase hits of ClpC1 substrates in a future BACTH library screen by mutating the ATPase active site of ClpC1 to prevent ATP hydrolysis and thereby unfolding.

3.4.2 The putative adaptor protein ClpC2

The protein ClpC2 is annotated as part of the Clp system, but is so far not described in the literature. Alignment of ClpC2 with ClpC1 show that it has sequence homology to the ClpC1 N-domain that is responsible for substrate and adaptor binding, indicating that these two proteins have a similar substrate clientele. As mentioned above, the BACTH library screen and subsequent validation experiments confirmed this hypothesis as we found about 20 overlapping hits in the library, and could show in co-transformation experiments that ClpC1 and ClpC2 interact with the same set of antitoxins. Additionally, we observed interaction of both proteins with the protein ClpS. In a BACTH co-transformation experiment ClpS interacted with ClpC2, while *in vitro* assays showed that ClpS is the canonical N-end rule adaptor for the chaperone ClpC1.

The occurrence of overlapping substrates or binders for ClpC1 and ClpC2 suggests that recognition of these interactors occurs via the conserved N-terminal region and therefore they compete for the same set of interactors. However, in experiments performed with ClpS, we observe no decrease in the degradation of the N-end rule substrate FR-li-GFP upon addition of ClpC2. Neither does ClpC2 prevent ClpS from inhibiting GFP-ssrA degradation by ClpC1P1P2. Also, the results from the competition of ClpC2 on the *in vitro* degradation of antitoxins indicate that even if ClpC2 binds a similar substrate set as ClpC1, it has a much lower affinity. It could potentially be a kind of “overflow” mechanism that sequesters ClpC1 substrates only when the chaperone is occupied.

So far, the function of ClpC2 could be not clearly determined with our experiments. Analysis of the ClpC2 sequence conservation reveals no ClpC2 homologs outside of Actinobacteria, while it was observed in actinobacterial orders such as Mycobacteria, Streptomycetales, Nocardia, Frankia, Rhodococcus, Micromonosporales and others while being absent in e.g. Corynebacterium, Bifidobacterium and Arthrobacter. An alignment of several actinobacterial ClpC2 homologues shows a moderate overall percentage of identity (~20 %), but reveals two highly con

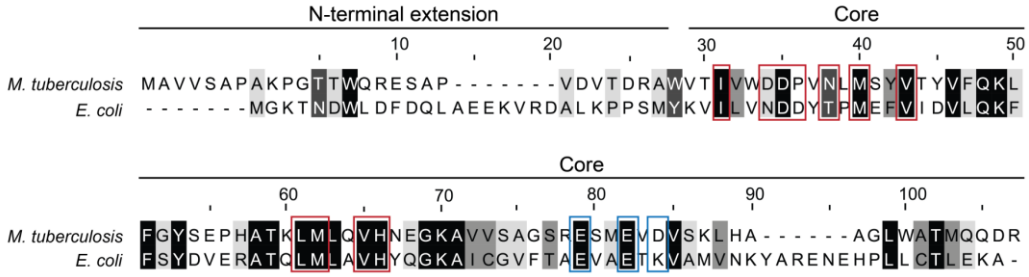


Figure 3.10: Alignment of *Mtb* ClpS with E. coli ClpS. Sequence conservation is colored from white (not conserved) to black (conserved). Stretches belonging to the N-terminal extension and to the globular core domain are indicated above. Residues important for N-end rule substrate binding are boxed in red, and the three residues on the core domain involved in chaperone binding are boxed in blue.

served stretches (Figure 3.2). One of them is in the N-terminal extension of ClpC2 (residue 11-83, *Mtb* numbering), the second partly coincides with the second N-terminal repeat region (residue 169-204, *Mtb* numbering). The middle part of the protein containing the first N-terminal repeat is less conserved.

The conservation of the N-terminal stretch of ClpC2, which is not homologous to the ClpC1 N-terminal domain, and not related to any known protein sequence, indicates that this stretch is important for a unique ClpC2 function. Further investigations of ClpC2, focusing on this N-terminal stretch, might help to elucidate the function of this actinobacterial Clp protein.

3.4.3 Rv1331 is the N-end rule adaptor ClpS in *Mtb*

In this chapter, we investigated the protein Rv1331, which was annotated as the putative adaptor ClpS. Our results show that Rv1331 inhibits degradation of GFP-ssrA and promotes the degradation of the N-end rule model substrate Fr-li-GFP, which is the characteristic behavior of ClpS proteins (Dougan et al, 2002; Erbse et al, 2006) and we conclude that Rv1331 is indeed ClpS that acts as an adaptor to the ClpC1 chaperone and switches its substrate specificity from ssrA-tagged substrates to N-end rule substrates.

Interestingly, while *Mtb* ClpS has homologous function to *E. coli* ClpS, it shows a reduced inhibition of ssrA-tagged substrates compared to *E. coli* ClpS. While *E. coli* ClpS inhibits substrate-ssrA degradation completely at a 1:1 ratio of ClpS to ClpA (monomer) (Hou et al, 2008), the same effect was not observed in our experiments. Even a more than 2-fold excess of ClpS over ClpC1 only markedly slowed degradation, but did not inhibit it completely. One reason could be that in this experiment the reaction conditions were not optimal, or that the affinity of *Mtb* ClpS to ClpC1 is not as strong as that of *E. coli* ClpS to ClpA. A reduced affinity could be due to differ-

ences in the binding interfaces between the respective ClpS' and chaperones. The binding of *E. coli* ClpS to ClpA is mediated by three conserved residues on the ClpS globular core domain, as well as its N-terminal extension (Roman-Hernandez et al, 2011; Zeth et al, 2002). The extension is however not conserved between *E. coli* and *Mtb* ClpS and could account for the decreased inhibition of *Mtb* ClpS (Figure 3.10). In general, the overall sequence similarity of ClpS homologues is not extremely high (e.g. around 22% sequence identity between *E. coli* and *Mtb* ClpS), and only residues critical for N-end rule substrate binding and the three residues for chaperone binding are highly conserved (Lupas & Koretke, 2003). Further differences between *Mtb* ClpS as a ClpC-associating adaptor and *E. coli* ClpS as a ClpA-associated adaptor include their genomic arrangement: ClpA and ClpS usually occur in one operon and ClpC-associated ClpS adaptors generally exist on distant genomic locations, affording them the opportunity to be differently regulated (Lupas & Koretke, 2003). In this context it is interesting to note that while ClpC1 is essential in *Mtb,* ClpS is not under *in vitro* growth conditions (Griffin et al, 2011; Sassetti et al, 2003). However, for survival of *Mtb* in macrophages, ClpS seems to become also a required gene (Rengarajan et al, 2005), indicating that ClpS might be specifically up- or downregulated under these conditions and regulation of the N-end rule pathway of *Mtb* is primarily necessary under infectious conditions.

3.4.4 The chaperone ClpC1 autodegrades in absence of substrate

In this work, we observe autodegradation of the ClpC1 chaperone (Figure 3.7 C). A similar behavior was described for *E. coli* ClpA, a ClpC ortholog, which regulates its cellular levels by autodegradation (Gottesman et al, 1990). So far, autodegradation of ClpC was not observed in cyanobacteria (Andersson et al, 2009), while on the other hand *S. aureus* ClpC it is suspected to be its own substrate (Feng et al, 2013).

ClpA autodegradation is a regulative process and in the presence of alternative substrates or binders, e.g. ssrA-tagged substrates or the adaptor ClpS, it does not occur (Maglica et al, 2008). In gel based experiments performed with ClpA and the model substrate GFP-ssrA, ClpA autodegradation only started after most of the substrate was already degraded (Maglica et al, 2008). In the degradation time traces performed in this thesis for the degradation of antitoxins (Figure 3.8 A and B), slow ClpC1 degradation was observed, and after 160 minutes the ClpC1 band has diminished to about half its size. However, due to the diffuse nature of the substrate band, it is hard to judge from these gels if ClpC1 autodegradation starts only after degradation of the substrate or if it is degraded in parallel. Further investigation of ClpC1 autodegradation requires experiments where the time points resolve not only substrate degradation, but also ClpC1 auto-

degradation. ClpC1 autodegradation time traces should also be performed in the absence of substrate to compare if substrate degradation competes for autodegradation. The use of a substrate that shows sharper bands in SDS-PAGE (e.g. GFP-ssrA) would be advantageous.

The recognition sequence on ClpA for autodegradation are the C-terminal 9 amino acids that are unstructured and show a similar recognition motif as the ssrA-tag (Maglica et al, 2008). Comparison of the ClpA and ClpC1 C-termini with the ssrA tag shows some similarity between the putative recognition motifs (Table 3.6). As the loops implicated in substrate binding near the ClpA and ClpC1 axial pore are almost identical (Flynn et al, 2001; Kress et al, 2009a), the recognition of the ClpC1 C-terminal stretch could occur by a similar mechanism as in ClpA. Furthermore, it is unlikely that degradation of ClpC1 is initiated at its N-terminus, as the crystal structure shows a tightly packed domain with no flexible residues for engagement by the unfolding loops (Vasudevan et al, 2013). To verify this hypothesis, the C-terminal residues of ClpC1 could be deleted and if they are indeed the recognition signal, autodegradation should be mostly abolished, as observed for ClpA. Furthermore, transplantation of this sequence to a model protein should render it a degradation substrate (Maglica et al, 2008).

Table 3.6: Proposed interaction motifs for ClpC1 autodegradation. In the *E. coli* ssrA sequence, the residues important for interaction with ClpA are marked with grey background (Flynn et al, 2001). In the *Mtb* ssrA tag, the second stretch of residues seems to be conserved ("ALA"). In row 3 and 4, the C-terminal 11 amino acids of ClpA and ClpC1, respectively are shown. In grey are similar regions that we suspect could be recognized by the chaperone. Both ClpC1 and ClpA contain a positively charged residue at their C-terminus that could be responsible for the reduced affinity compared to the ssrA tag (Maglica et al, 2008).

E. coli ssrA sequence	AANDENYALAA
Mtb ssrA sequence	ADSHQRDYALAA
ClpA C-terminal residues	SAQKHKAEAAH
ClpC1 C-terminal residues	ASAGGPEPAAR

3.4.5 Notes on the activator peptide

Interestingly, in our experiments with ClpC1 the autodegradation (and also antitoxin and FR-li-GFP degradation, data not shown) was only observed in the presence of the activator peptide Z-Leu-Leu, in contrast to ssrA-tagged protein degradation that also occurs without (Chapter 2). As shown by Schmitz et al., the activator stabilizes the active conformation of the ClpP1P2 complex, and the same holds true for substrate translocation by the chaperone and degradation by ClpP1P2 (Schmitz & Sauer, 2014). One hypothesis to the requirement of the activator for ClpC1 autodegradation is that the ClpC1 recognition motif is not as efficiently recognized as the ssrA-tag. In fact, *E. coli* ClpA for example binds an ssrA-tagged model substrate with twice the affinity than the same substrates fused to the ClpA autodegradation sequence (Maglica et al, 2008). A

weaker recognition motif might not allow for efficient gripping and unfolding of the substrate and thereby not stabilize the active ClpC1P1P2 assembly long or strongly enough for degradation to occur, at least under the given *in vitro* conditions in absence of activator. *In vivo*, peptides present in the cell might stabilize the ClpP1P2 complex with the ClpX and ClpC1 chaperones, or so far unknown additional factors might play a role in the assembly of the Clp chaperone-protease system.

Chapter 4 Conclusion

The Clp chaperone-protease system is essential in the highly pathogenic *Mycobacterium tuberculosis (Mtb)* that infects about one third of the world population (Griffin et al, 2011; Sassetti et al, 2003; WHO, 2015). It is one of three chaperone-protease systems harbored by this organism; the others are FtsH, also essential, and the Pup-proteasome system. The Pup-Proteasome system is not essential under *in vitro* growth conditions, but is required for persistence of *Mtb* in macrophages (Darwin et al, 2003; Griffin et al, 2011; Lamichhane et al, 2006; Sassetti et al, 2003). As FtsH is membrane-associated and largely responsible for the degradation of membrane proteins, and the Pup-proteasome system degrades proteins specifically modified with the Pup tag, the Clp system is the main cytosolic house-keeping protease that takes care of e.g. unfolded proteins or incompletely synthesized proteins modified with the ssrA tag (Laederach et al, 2014). In addition to general proteolysis, the Clp system of *Mtb* is involved in the degradation of regulatory proteins such as transcriptional regulators (Barik et al, 2010; Dziedzic et al, 2010; Raju et al, 2014). Its essential nature and multi-component assembly make it a promising drug target. Indeed, several compounds have been found acting on the Clp system, illustrating that its deregulation is detrimental to the *Mtb* cell. Known drugs act either through prevention of correct complex assembly and subsequent decoupling of unfolding from degradation, through inhibition of protease action or through overstimulation of chaperone activity (Gavrish et al, 2014; Raju et al, 2012a). As drug action on so many faces of this complex has already been observed, more detailed investigations of the Clp system are important to fully understand how this complex works and to pave the way for further drug development. This thesis therefore investigates this complex, from the assembly of its protease, chaperones and adaptor components to identifying new substrates for the functional complex.

The ClpP1P2 protease core is an unusual ClpP core particle as it differs from well-studied Clp systems like the one of *E. coli* by being composed of two subunits that form an asymmetric particle, where the ClpP1 and ClpP2 sides of the complex present different binding interfaces. This thesis shows that, surprisingly, the chaperone partners ClpX and ClpC1 interact only with the ClpP2 side of the ClpP1P2 proteolytic particle, which is a somewhat unexpected behavior as for *E. coli,* chaperone association to both sides of the proteolytic particle was shown to be the more efficient assembly (Maglica et al, 2009). However, this might prove an advantage as it frees the

ClpP1 face of the ClpP1P2 complex for so far undescribed binders. For example, for the Pup-Proteasome system recently a non-ATPase interactor was described (Delley et al, 2014) and having the ClpP1 interaction face free could afford the Clp system the possibility to bind such an additional regulator. In fact, using the bacterial two hybrid library screen that was already employed in this thesis to find interactors of the Clp components ClpC1 and ClpC2, new interactors could be identified when using ClpP1 as a bait protein (Battesti & Bouveret, 2012). To corroborate our results, *in vivo* studies could be performed with conditional mutants of ClpP1 and ClpP2 in *Mtb* (Raju et al, 2014; Raju et al, 2012b), where hydrophobic patch mutations of ClpP2, but not ClpP1 should render *Mtb* non-viable.

Assembly of the chaperone and protease components is indispensable for the correct function for the complex, the controlled degradation of substrate proteins. The Clp system is involved in general protein control and degrades broad substrate classes such as unfolded proteins, incompletely synthesized proteins tagged with the ssrA-peptide and N-end rule substrates (Sauer & Baker, 2011). Of these, the *Mtb* Clp system has been described to degrade unfolded as well as ssrA-tagged model substrates (Akopian et al, 2012; Schmitz & Sauer, 2014). This thesis completes the picture by showing that *Mtb* also harbors an N-end-rule degradation pathway, as the protein *Rv1331* is the *Mtb* ClpS adaptor that delivers an N-end-rule model protein to the ClpC1P1P2 complex. In addition to these general protein quality control substrates, the *Mtb* Clp system has been implicated in the regulatory degradation of several transcriptional regulators and their binders that probably contribute to it being an essential degradation system in the cell (Laederach et al, 2014; Raju et al, 2014). In this thesis, we add the substrate class of toxin-antitoxin systems to the substrate spectrum of the ClpC1 chaperone, and show that two antitoxins of the Rel and Vap class are degraded by the ClpC1P1P2 complex. In the cell, antitoxin degradation frees the respective toxin to perform its function, which in the case of the selected examples connects the Clp system to regulatory degradation relevant for *Mtb* persistence (Sala et al, 2014). Future studies could include *in vitro* degradation of more functional antitoxin classes, as well as degradation of the corresponding toxins. Furthermore, bioinformatical analysis of the library screen data could reveal conserved recognition motifs of the ClpC1 and ClpC2 proteins.

The *Mtb* Clp system itself is upregulated in response to various stress conditions, but is also involved in positive and negative feedback regulation of its own expression (Laederach et al, 2014). Based on results in this thesis, we suggest an additional layer of auto-regulation for the ClpC1 chaperone. ClpC1 promotes its own degradation in the absence of substrate, a process reminiscent of its *E. coli* orthologue ClpA, for which autodegradation was proposed to be a self-

regulatory mechanism (Gottesman et al, 1990). Interestingly, the second cytosolic chaperone-protease system, the Pup-proteasome system could also be involved in the regulation of the Clp system. Both protease subunits ClpP1 and ClpP2, as well as the chaperone ClpC1 were found in proteomic studies identifying Pup-tagged substrates of the mycobacterial Pup-proteasome system (termed the "pupylome") (Festa et al, 2010; Poulsen et al, 2010; Watrous et al, 2010). *In vitro* tests of Pup-tagging of Clp components and subsequent degradation by the proteasome could further elucidate the regulation of the Clp chaperone-protease system and establish an interplay between both chaperone-protease systems. There is also indication that the Clp and the Pup-proteasome system share a similar set of substrates, as toxin components of toxin-antitoxin systems were identified in our bacterial two hybrid screen for ClpC1 as well as in the above mentioned pupylome studies. As often bacterial chaperone-proteases show some degree of substrate overlap, comparison of the results of the ClpC1 library screen performed in this thesis with a library screen currently performed in our lab with the proteasomal ATPase Mpa could provide hints whether a substrate overlap occurs between both systems. Degradation of regulatory substrates under stress conditions by two instead of one chaperone-protease system could prove beneficial to the *Mtb* cell.

In the future, to complete the picture of interactors of the Clp system, on the one hand the bacterial two hybrid library screen could be employed to expand the spectrum of substrates, interactors and adaptors for the chaperone ClpX. Furthermore, to reveal the function of the putative adaptor ClpC2, investigations should be continued. The results obtained in this thesis already provide a basis for further biochemical and structural studies that might give insights into the function of this protein. Additionally, creating a knock-out of the protein for example in the model organism *M. smegmatis*, could show a phenotype that indicates its *in vivo* relevance.

Concluding, further knowledge about the Clp system, its assembly, mechanism, adaptors, and substrate clientele will pave the way for the design and development of new antibacterial compounds acting on the *Mtb* Clp system, which is especially important nowadays when more and more multi-drug resistant *Mtb* strains are evolving (WHO, 2015).

References

Akopian T, Kandror O, Raju RM, Unnikrishnan M, Rubin EJ, Goldberg AL (2012) The active ClpP protease from M. tuberculosis is a complex composed of a heptameric ClpP1 and a ClpP2 ring. *The EMBO journal* **31:** 1529-1541

Albrethsen J, Agner J, Piersma SR, Hojrup P, Pham TV, Weldingh K, Jimenez CR, Andersen P, Rosenkrands I (2013) Proteomic profiling of Mycobacterium tuberculosis identifies nutrient-starvation-responsive toxin-antitoxin systems. *Mol Cell Proteomics* **12:** 1180-1191

Alexopoulos J, Ahsan B, Homchaudhuri L, Husain N, Cheng YQ, Ortega J (2013) Structural determinants stabilizing the axial channel of ClpP for substrate translocation. *Mol Microbiol* **90:** 167-180

Alexopoulos JA, Guarne A, Ortega J (2012) ClpP: a structurally dynamic protease regulated by AAA+ proteins. *Journal of structural biology* **179:** 202-210

Andersson FI, Tryggvesson A, Sharon M, Diemand AV, Classen M, Best C, Schmidt R, Schelin J, Stanne TM, Bukau B, Robinson CV, Witt S, Mogk A, Clarke AK (2009) Structure and function of a novel type of ATP-dependent Clp protease. *The Journal of biological chemistry* **284:** 13519-13532

Bajaj D, Batra JK (2012) The C-terminus of ClpC1 of Mycobacterium tuberculosis is crucial for its oligomerization and function. *PLoS One* **7:** e51261

Baker TA, Sauer RT (2012) ClpXP, an ATP-powered unfolding and protein-degradation machine. *Biochimica et biophysica acta* **1823:** 15-28

Baldwin RL (1986) Temperature dependence of the hydrophobic interaction in protein folding. *Proc Natl Acad Sci U S A* **83:** 8069-8072

Barik S, Sureka K, Mukherjee P, Basu J, Kundu M (2010) RseA, the SigE specific anti-sigma factor of Mycobacterium tuberculosis, is inactivated by phosphorylation-dependent ClpC1P2 proteolysis. *Mol Microbiol* **75:** 592-606

Battesti A, Bouveret E (2012) The bacterial two-hybrid system based on adenylate cyclase reconstitution in Escherichia coli. *Methods* **58:** 325-334

Battesti A, Gottesman S (2013) Roles of adaptor proteins in regulation of bacterial proteolysis. *Curr Opin Microbiol* **16:** 140-147

Becker G, Klauck E, Hengge-Aronis R (1999) Regulation of RpoS proteolysis in Escherichia coli: the response regulator RssB is a recognition factor that interacts with the turnover element in RpoS. *Proc Natl Acad Sci U S A* **96:** 6439-6444

Bellier A, Gominet M, Mazodier P (2006) Post-translational control of the Streptomyces lividans ClgR regulon by ClpP. *Microbiology* **152:** 1021-1027

Bellier A, Mazodier P (2004) ClgR, a novel regulator of clp and lon expression in Streptomyces. *Journal of bacteriology* **186:** 3238-3248

Benaroudj N, Raynal B, Miot M, Ortiz-Lombardia M (2011) Assembly and proteolytic processing of mycobacterial ClpP1 and ClpP2. *BMC biochemistry* **12:** 61

Bewley MC, Graziano V, Griffin K, Flanagan JM (2006) The asymmetry in the mature amino-terminus of ClpP facilitates a local symmetry match in ClpAP and ClpXP complexes. *Journal of structural biology* **153:** 113-128

Bewley MC, Graziano V, Griffin K, Flanagan JM (2009) Turned on for degradation: ATPase-independent degradation by ClpP. *Journal of structural biology* **165:** 118-125

Bieniossek C, Schalch T, Bumann M, Meister M, Meier R, Baumann U (2006) The molecular architecture of the metalloprotease FtsH. *Proc Natl Acad Sci U S A* **103:** 3066-3071

Bottcher T, Sieber SA (2008) Beta-lactones as specific inhibitors of ClpP attenuate the production of extracellular virulence factors of Staphylococcus aureus. *Journal of the American Chemical Society* **130:** 14400-14401

Bougdour A, Cunning C, Baptiste PJ, Elliott T, Gottesman S (2008) Multiple pathways for regulation of sigmaS (RpoS) stability in Escherichia coli via the action of multiple anti-adaptors. *Mol Microbiol* **68:** 298-313

Brotz-Oesterhelt H, Beyer D, Kroll HP, Endermann R, Ladel C, Schroeder W, Hinzen B, Raddatz S, Paulsen H, Henninger K, Bandow JE, Sahl HG, Labischinski H (2005) Dysregulation of bacterial proteolytic machinery by a new class of antibiotics. *Nature medicine* **11:** 1082-1087

Brzozowska I, Zielenkiewicz U (2013) Regulation of toxin-antitoxin systems by proteolysis. *Plasmid* **70:** 33-41

Chandler D (2005) Interfaces and the driving force of hydrophobic assembly. *Nature* **437:** 640-647

Choi KH, Licht S (2005) Control of peptide product sizes by the energy-dependent protease ClpAP. *Biochemistry* **44:** 13921-13931

Compton CL, Schmitz KR, Sauer RT, Sello JK (2013) Antibacterial activity of and resistance to small molecule inhibitors of the ClpP peptidase. *ACS chemical biology* **8:** 2669-2677

Conlon BP, Nakayasu ES, Fleck LE, LaFleur MD, Isabella VM, Coleman K, Leonard SN, Smith RD, Adkins JN, Lewis K (2013) Activated ClpP kills persisters and eradicates a chronic biofilm infection. *Nature* **503:** 365-370

Cranz-Mileva S, Imkamp F, Kolygo K, Maglica Z, Kress W, Weber-Ban E (2008) The flexible attachment of the N-domains to the ClpA ring body allows their use on demand. *Journal of molecular biology* **378:** 412-424

Darwin KH, Ehrt S, Gutierrez-Ramos JC, Weich N, Nathan CF (2003) The proteasome of Mycobacterium tuberculosis is required for resistance to nitric oxide. *Science* **302:** 1963-1966

Delley CL, Laederach J, Ziemski M, Bolten M, Boehringer D, Weber-Ban E (2014) Bacterial proteasome activator bpa (rv3780) is a novel ring-shaped interactor of the mycobacterial proteasome. *PloS one* **9:** e114348

Donegan NP, Cheung AL (2009) Regulation of the mazEF toxin-antitoxin module in Staphylococcus aureus and its impact on sigB expression. *Journal of bacteriology* **191:** 2795-2805

Donegan NP, Thompson ET, Fu Z, Cheung AL (2010) Proteolytic regulation of toxin-antitoxin systems by ClpPC in Staphylococcus aureus. *Journal of bacteriology* **192:** 1416-1422

Dougan DA, Micevski D, Truscott KN (2012) The N-end rule pathway: from recognition by N-recognins, to destruction by AAA+proteases. *Biochimica et biophysica acta* **1823:** 83-91

Dougan DA, Reid BG, Horwich AL, Bukau B (2002) ClpS, a substrate modulator of the ClpAP machine. *Mol Cell* **9:** 673-683

Dougan DA, Truscott KN, Zeth K (2010) The bacterial N-end rule pathway: expect the unexpected. *Mol Microbiol* **76:** 545-558

Dziedzic R, Kiran M, Plocinski P, Ziolkiewicz M, Brzostek A, Moomey M, Vadrevu IS, Dziadek J, Madiraju M, Rajagopalan M (2010) Mycobacterium tuberculosis ClpX interacts with FtsZ and interferes with FtsZ assembly. *PloS one* **5:** e11058

Effantin G, Ishikawa T, De Donatis GM, Maurizi MR, Steven AC (2010a) Local and global mobility in the ClpA AAA+ chaperone detected by cryo-electron microscopy: functional connotations. *Structure* **18:** 553-562

Effantin G, Maurizi MR, Steven AC (2010b) Binding of the ClpA unfoldase opens the axial gate of ClpP peptidase. *The Journal of biological chemistry* **285:** 14834-14840

Engels S, Ludwig C, Schweitzer JE, Mack C, Bott M, Schaffer S (2005) The transcriptional activator ClgR controls transcription of genes involved in proteolysis and DNA repair in Corynebacterium glutamicum. *Mol Microbiol* **57:** 576-591

Engels S, Schweitzer JE, Ludwig C, Bott M, Schaffer S (2004) clpC and clpP1P2 gene expression in Corynebacterium glutamicum is controlled by a regulatory network involving the transcriptional regulators ClgR and HspR as well as the ECF sigma factor sigmaH. *Mol Microbiol* **52:** 285-302

Erbse A, Schmidt R, Bornemann T, Schneider-Mergener J, Mogk A, Zahn R, Dougan DA, Bukau B (2006) ClpS is an essential component of the N-end rule pathway in Escherichia coli. *Nature* **439:** 753-756

Estorninho M, Smith H, Thole J, Harders-Westerveen J, Kierzek A, Butler RE, Neyrolles O, Stewart GR (2010) ClgR regulation of chaperone and protease systems is essential for Mycobacterium tuberculosis parasitism of the macrophage. *Microbiology* **156:** 3445-3455

Farrell CM, Baker TA, Sauer RT (2007) Altered specificity of a AAA+ protease. *Mol Cell* **25:** 161-166

Feng J, Michalik S, Varming AN, Andersen JH, Albrecht D, Jelsbak L, Krieger S, Ohlsen K, Hecker M, Gerth U, Ingmer H, Frees D (2013) Trapping and proteomic identification of cellular substrates of the ClpP protease in Staphylococcus aureus. *J Proteome Res* **12:** 547-558

Festa RA, McAllister F, Pearce MJ, Mintseris J, Burns KE, Gygi SP, Darwin KH (2010) Prokaryotic ubiquitin-like protein (Pup) proteome of Mycobacterium tuberculosis [corrected]. *PloS one* **5:** e8589

Flannagan RS, Cosio G, Grinstein S (2009) Antimicrobial mechanisms of phagocytes and bacterial evasion strategies. *Nat Rev Microbiol* **7:** 355-366

Flynn JM, Levchenko I, Seidel M, Wickner SH, Sauer RT, Baker TA (2001) Overlapping recognition determinants within the ssrA degradation tag allow modulation of proteolysis. *Proc Natl Acad Sci U S A* **98:** 10584-10589

Fu Z, Tamber S, Memmi G, Donegan NP, Cheung AL (2009) Overexpression of MazFsa in Staphylococcus aureus induces bacteriostasis by selectively targeting mRNAs for cleavage. *Journal of bacteriology* **191:** 2051-2059

Gandhi NR, Nunn P, Dheda K, Schaaf HS, Zignol M, van Soolingen D, Jensen P, Bayona J (2010) Multidrug-resistant and extensively drug-resistant tuberculosis: a threat to global control of tuberculosis. *The Lancet* **375:** 1830-1843

Gandotra S, Schnappinger D, Monteleone M, Hillen W, Ehrt S (2007) In vivo gene silencing identifies the Mycobacterium tuberculosis proteasome as essential for the bacteria to persist in mice. *Nature medicine* **13:** 1515-1520

Gao W, Kim JY, Anderson JR, Akopian T, Hong S, Jin YY, Kandror O, Kim JW, Lee IA, Lee SY, McAlpine JB, Mulugeta S, Sunoqrot S, Wang Y, Yang SH, Yoon TM, Goldberg AL, Pauli GF, Suh JW, Franzblau SG, Cho S (2015) The cyclic peptide ecumicin targeting ClpC1 is active against Mycobacterium tuberculosis in vivo. *Antimicrob Agents Chemother* **59:** 880-889

Gao W, Kim JY, Chen SN, Cho SH, Choi J, Jaki BU, Jin YY, Lankin DC, Lee JE, Lee SY, McAlpine JB, Napolitano JG, Franzblau SG, Suh JW, Pauli GF (2014) Discovery and characterization of the tuberculosis drug lead ecumicin. *Org Lett* **16:** 6044-6047

Gavrish E, Sit CS, Cao S, Kandror O, Spoering A, Peoples A, Ling L, Fetterman A, Hughes D, Bissell A, Torrey H, Akopian T, Mueller A, Epstein S, Goldberg A, Clardy J, Lewis K (2014) Lassomycin, a ribosomally synthesized cyclic peptide, kills Mycobacterium tuberculosis by targeting the ATP-dependent protease ClpC1P1P2. *Chemistry & biology* **21:** 509-518

Geertsma ER, Dutzler R (2011) A versatile and efficient high-throughput cloning tool for structural biology. *Biochemistry* **50:** 3272-3278

Gibson DG, Young L, Chuang RY, Venter JC, Hutchison CA, 3rd, Smith HO (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods* **6:** 343-345

Glynn SE, Martin A, Nager AR, Baker TA, Sauer RT (2009) Structures of asymmetric ClpX hexamers reveal nucleotide-dependent motions in a AAA+ protein-unfolding machine. *Cell* **139:** 744-756

Gottesman S, Clark WP, Maurizi MR (1990) The ATP-dependent Clp protease of Escherichia coli. Sequence of clpA and identification of a Clp-specific substrate. *The Journal of biological chemistry* **265:** 7886-7893

Gottesman S, Roche E, Zhou Y, Sauer RT (1998) The ClpXP and ClpAP proteases degrade proteins with carboxy-terminal peptide tails added by the SsrA-tagging system. *Genes & development* **12:** 1338-1347

Goujon M, McWilliam H, Li W, Valentin F, Squizzato S, Paern J, Lopez R (2010) A new bioinformatics analysis tools framework at EMBL-EBI. *Nucleic acids research* **38:** W695-699

Gribun A, Kimber MS, Ching R, Sprangers R, Fiebig KM, Houry WA (2005) The ClpP double ring tetradecameric protease exhibits plastic ring-ring interactions, and the N termini of its subunits form flexible loops that are essential for ClpXP and ClpAP complex formation. *The Journal of biological chemistry* **280:** 16185-16196

Griffin JE, Gawronski JD, Dejesus MA, Ioerger TR, Akerley BJ, Sassetti CM (2011) High-resolution phenotypic profiling defines genes essential for mycobacterial growth and cholesterol catabolism. *PLoS pathogens* **7:** e1002251

Grimaud R, Kessel M, Beuron F, Steven AC, Maurizi MR (1998) Enzymatic and structural similarities between the Escherichia coli ATP-dependent proteases, ClpXP and ClpAP. *The Journal of biological chemistry* **273:** 12476-12481

Guo F, Esser L, Singh SK, Maurizi MR, Xia D (2002) Crystal structure of the heterodimeric complex of the adaptor, ClpS, with the N-domain of the AAA+ chaperone, ClpA. *The Journal of biological chemistry* **277:** 46753-46762

Gur E, Ottofueling R, Dougan D (2013) Machines of Destruction – AAA+ Proteases and the Adaptors That Control Them. In *Regulated Proteolysis in Microorganisms*, Dougan DA (ed), Vol. 66, 1, pp 3-33. Springer Netherlands

Hinnerwisch J, Fenton WA, Furtak KJ, Farr GW, Horwich AL (2005a) Loops in the central channel of ClpA chaperone mediate protein binding, unfolding, and translocation. *Cell* **121:** 1029-1041

Hinnerwisch J, Reid BG, Fenton WA, Horwich AL (2005b) Roles of the N-domains of the ClpA unfoldase in binding substrate proteins and in stable complex formation with the ClpP protease. *The Journal of biological chemistry* **280:** 40838-40844

Hou JY, Sauer RT, Baker TA (2008) Distinct structural elements of the adaptor ClpS are required for regulating degradation by ClpAP. *Nature structural & molecular biology* **15:** 288-294

Imkamp F, Ziemski M, Weber-Ban E (2015) Pupylation-dependent and -independent proteasomal degradation in mycobacteria. *Biomol Concepts* **6:** 285-301

Ingvarsson H, Mate MJ, Hogbom M, Portnoi D, Benaroudj N, Alzari PM, Ortiz-Lombardia M, Unge T (2007) Insights into the inter-ring plasticity of caseinolytic proteases from the X-ray structure of Mycobacterium tuberculosis ClpP1. *Acta crystallographica Section D, Biological crystallography* **63:** 249-259

Iniesta AA, McGrath PT, Reisenauer A, McAdams HH, Shapiro L (2006) A phospho-signaling pathway controls the localization and activity of a protease complex critical for bacterial cell cycle progression. *Proc Natl Acad Sci U S A* **103:** 10935-10940

Ishikawa T, Maurizi MR, Steven AC (2004) The N-terminal substrate-binding domain of ClpA unfoldase is highly mobile and extends axially from the distal surface of ClpAP protease. *Journal of structural biology* **146:** 180-188

Jassal M, Bishai WR (2009) Extensively drug-resistant tuberculosis. *The Lancet Infectious Diseases* **9:** 19-30

Jenal U (2009) The role of proteolysis in the Caulobacter crescentus cell cycle and development. *Research in microbiology* **160:** 687-695

Jenal U, Fuchs T (1998) An essential protease involved in bacterial cell-cycle control. *The EMBO journal* **17:** 5658-5669

Jennings LD, Lun DS, Medard M, Licht S (2008) ClpP hydrolyzes a protein substrate processively in the absence of the ClpA ATPase: mechanistic studies of ATP-independent proteolysis. *Biochemistry* **47:** 11536-11546

Joshi SA, Hersch GL, Baker TA, Sauer RT (2004) Communication between ClpX and ClpP during substrate processing and degradation. *Nature structural & molecular biology* **11:** 404-411

Kang SG, Dimitrova MN, Ortega J, Ginsburg A, Maurizi MR (2005) Human mitochondrial ClpP is a stable heptamer that assembles into a tetradecamer in the presence of ClpX. *The Journal of biological chemistry* **280:** 35424-35432

Kang SG, Maurizi MR, Thompson M, Mueser T, Ahvazi B (2004) Crystallography and mutagenesis point to an essential role for the N-terminus of human mitochondrial ClpP. *Journal of structural biology* **148:** 338-352

Kar NP, Sikriwal D, Rath P, Choudhary RK, Batra JK (2008) Mycobacterium tuberculosis ClpC1: characterization and role of the N-terminal domain in its function. *FEBS J* **275:** 6149-6158

Karimova G, Pidoux J, Ullmann A, Ladant D (1998) A bacterial two-hybrid system based on a reconstituted signal transduction pathway. *Proc Natl Acad Sci U S A* **95:** 5752-5756

Kenniston JA, Baker TA, Fernandez JM, Sauer RT (2003) Linkage between ATP consumption and mechanical unfolding during the protein processing reactions of an AAA+ degradation machine. *Cell* **114:** 511-520

Kessel M, Maurizi MR, Kim B, Kocsis E, Trus BL, Singh SK, Steven AC (1995) Homology in structural organization between E. coli ClpAP protease and the eukaryotic 26 S proteasome. *Journal of molecular biology* **250:** 587-594

Kim DY, Kim KK (2003) Crystal structure of ClpX molecular chaperone from Helicobacter pylori. *The Journal of biological chemistry* **278:** 50664-50670

Kim YI, Levchenko I, Fraczkowska K, Woodruff RV, Sauer RT, Baker TA (2001) Molecular determinants of complex formation between Clp/Hsp100 ATPases and the ClpP peptidase. *Nature structural biology* **8:** 230-233

Kimber MS, Yu AY, Borg M, Leung E, Chan HS, Houry WA (2010) Structural and theoretical studies indicate that the cylindrical protease ClpP samples extended and compact conformations. *Structure* **18:** 798-808

Kirstein J, Hoffmann A, Lilie H, Schmidt R, Rubsamen-Waigmann H, Brotz-Oesterhelt H, Mogk A, Turgay K (2009a) The antibiotic ADEP reprogrammes ClpP, switching it from a regulated to an uncontrolled protease. *EMBO molecular medicine* **1:** 37-49

Kirstein J, Moliere N, Dougan DA, Turgay K (2009b) Adapting the machine: adaptor proteins for Hsp100/Clp and AAA+ proteases. *Nat Rev Microbiol* **7:** 589-599

Kress W, Maglica Z, Weber-Ban E (2009a) Clp chaperone-proteases: structure and function. *Research in microbiology* **160:** 618-628

Kress W, Mutschler H, Weber-Ban E (2007) Assembly pathway of an AAA+ protein: tracking ClpA and ClpAP complex formation in real time. *Biochemistry* **46:** 6183-6193

Kress W, Mutschler H, Weber-Ban E (2009b) Both ATPase domains of ClpA are critical for processing of stable protein structures. *The Journal of biological chemistry* **284:** 31441-31452

Laederach J, Leodolter J, Warweg J, Weber-Ban E (2014) Chaperone-Proteases of Mycobacteria. In *The Molecular Chaperones Interaction Networks in Protein Folding and Degradation*, Houry WA (ed), Vol. 1, 16, pp 419-444. Springer New York

Lamichhane G, Raghunand TR, Morrison NE, Woolwine SC, Tyagi S, Kandavelou K, Bishai WR (2006) Deletion of a Mycobacterium tuberculosis proteasomal ATPase homologue gene produces a slow-growing strain that persists in host tissues. *The Journal of infectious diseases* **194:** 1233-1240

Lee BG, Park EY, Lee KE, Jeon H, Sung KH, Paulsen H, Rubsamen-Schaeff H, Brotz-Oesterhelt H, Song HK (2010) Structures of ClpP in complex with acyldepsipeptide antibiotics reveal its activation mechanism. *Nature structural & molecular biology* **17:** 471-478

Lee C, Schwartz MP, Prakash S, Iwakura M, Matouschek A (2001) ATP-dependent proteases degrade their substrates by processively unraveling them from the degradation signal. *Mol Cell* **7:** 627-637

Leodolter J, Warweg J, Weber-Ban E (2015) The Mycobacterium tuberculosis ClpP1P2 Protease Interacts Asymmetrically with Its ATPase Partners ClpX and ClpC1. *PloS one* **10:** e0125345

Li DH, Chung YS, Gloyd M, Joseph E, Ghirlando R, Wright GD, Cheng YQ, Maurizi MR, Guarne A, Ortega J (2010) Acyldepsipeptide antibiotics induce the formation of a structured axial channel in ClpP: A model for the ClpX/ClpA-bound state of ClpP. *Chemistry & biology* **17:** 959-969

Liu J, Mei Z, Li N, Qi Y, Xu Y, Shi Y, Wang F, Lei J, Gao N (2013) Structural dynamics of the MecA-ClpC complex: a type II AAA+ protein unfolding machine. *The Journal of biological chemistry* **288:** 17597-17608

Liu K, Ologbenla A, Houry WA (2014) Dynamics of the ClpP serine protease: a model for self-compartmentalized proteases. *Critical reviews in biochemistry and molecular biology* **49:** 400-412

Lupas AN, Koretke KK (2003) Bioinformatic analysis of ClpS, a protein module involved in prokaryotic and eukaryotic protein degradation. *Journal of structural biology* **141:** 77-83

Maglica Z, Kolygo K, Weber-Ban E (2009) Optimal efficiency of ClpAP and ClpXP chaperone-proteases is achieved by architectural symmetry. *Structure* **17:** 508-516

Maglica Z, Striebel F, Weber-Ban E (2008) An intrinsic degradation tag on the ClpA C-terminus regulates the balance of ClpAP complexes with different substrate specificity. *Journal of molecular biology* **384:** 503-511

Martin A, Baker TA, Sauer RT (2007) Distinct static and dynamic interactions control ATPase-peptidase communication in a AAA+ protease. *Mol Cell* **27:** 41-52

Martin A, Baker TA, Sauer RT (2008a) Diverse pore loops of the AAA+ ClpX machine mediate unassisted and adaptor-dependent recognition of ssrA-tagged substrates. *Mol Cell* **29:** 441-450

Martin A, Baker TA, Sauer RT (2008b) Pore loops of the AAA+ ClpX machine grip substrates to drive translocation and unfolding. *Nature structural & molecular biology* **15:** 1147-1151

Matthews BW (2001) Hydrophobic Interactions in Proteins. In *eLS*. John Wiley & Sons, Ltd

Maurizi MR, Clark WP, Katayama Y, Rudikoff S, Pumphrey J, Bowers B, Gottesman S (1990) Sequence and structure of Clp P, the proteolytic component of the ATP-dependent Clp protease of Escherichia coli. *The Journal of biological chemistry* **265:** 12536-12545

Maurizi MR, Singh SK, Thompson MW, Kessel M, Ginsburg A (1998) Molecular properties of ClpAP protease of Escherichia coli: ATP-dependent association of ClpA and clpP. *Biochemistry* **37:** 7778-7786

Mehra S, Kaushal D (2009) Functional genomics reveals extended roles of the Mycobacterium tuberculosis stress response factor sigmaH. *Journal of bacteriology* **191:** 3965-3980

Mogk A, Dougan D, Weibezahn J, Schlieker C, Turgay K, Bukau B (2004) Broad yet high substrate specificity: the challenge of AAA+ proteins. *Journal of structural biology* **146:** 90-98

Moore SD, Sauer RT (2007) The tmRNA system for translational surveillance and ribosome rescue. *Annual review of biochemistry* **76:** 101-124

Moreira W, Ngan GJ, Low JL, Poulsen A, Chia BC, Ang MJ, Yap A, Fulwood J, Lakshmanan U, Lim J, Khoo AY, Flotow H, Hill J, Raju RM, Rubin EJ, Dick T (2015) Target mechanism-based whole-cell screening identifies bortezomib as an inhibitor of caseinolytic protease in mycobacteria. *MBio* **6:** e00253-00215

Neher SB, Flynn JM, Sauer RT, Baker TA (2003) Latent ClpX-recognition signals ensure LexA destruction after DNA damage. *Genes & development* **17:** 1084-1089

Nishimura K, van Wijk KJ (2015) Organization, function and substrates of the essential Clp protease system in plastids. *Biochimica et biophysica acta* **1847:** 915-930

Ogura T, Okuno T, Suno R, Akiyama Y (2013) Chapter 144 - FtsH Protease. In *Handbook of Proteolytic Enzymes*, pp 685-692. Academic Press

Olivares AO, Baker TA, Sauer RT (2016) Mechanistic insights into bacterial AAA+ proteases and protein-remodelling machines. *Nat Rev Microbiol* **14:** 33-44

Ollinger J, O'Malley T, Kesicki EA, Odingo J, Parish T (2012) Validation of the essential ClpP protease in Mycobacterium tuberculosis as a novel drug target. *Journal of bacteriology* **194:** 663-668

Park E, Rho YM, Koh OJ, Ahn SW, Seong IS, Song JJ, Bang O, Seol JH, Wang J, Eom SH, Chung CH (2005) Role of the GYVG pore motif of HslU ATPase in protein unfolding and translocation for degradation by HslV peptidase. *The Journal of biological chemistry* **280:** 22892-22898

Park SC, Jia BL, Yang JK, Le Van D, Shao YG, Han SW, Jeon YJ, Chung CH, Cheong GW (2006) Oligomeric structure of the ATP-dependent protease La(Lon) of Escherichia coli. *Mol Cells* **21:** 129-134

Personne Y, Brown AC, Schuessler DL, Parish T (2013) Mycobacterium tuberculosis ClpP proteases are co-transcribed but exhibit different substrate specificities. *PloS one* **8:** e60228

Poulsen C, Akhter Y, Jeon AH, Schmitt-Ulms G, Meyer HE, Stefanski A, Stuhler K, Wilmanns M, Song YH (2010) Proteome-wide identification of mycobacterial pupylation targets. *Molecular systems biology* **6:** 386

Raju RM, Goldberg AL, Rubin EJ (2012a) Bacterial proteolytic complexes as therapeutic targets. *Nat Rev Drug Discov* **11:** 777-789

Raju RM, Jedrychowski MP, Wei JR, Pinkham JT, Park AS, O'Brien K, Rehren G, Schnappinger D, Gygi SP, Rubin EJ (2014) Post-translational regulation via Clp protease is critical for survival of Mycobacterium tuberculosis. *PLoS pathogens* **10:** e1003994

Raju RM, Unnikrishnan M, Rubin DH, Krishnamoorthy V, Kandror O, Akopian TN, Goldberg AL, Rubin EJ (2012b) Mycobacterium tuberculosis ClpP1 and ClpP2 function together in protein degradation and are required for viability in vitro and during infection. *PLoS pathogens* **8:** e1002511

Reid BG, Fenton WA, Horwich AL, Weber-Ban EU (2001) ClpA mediates directional translocation of substrate proteins into the ClpP protease. *Proc Natl Acad Sci U S A* **98:** 3768-3772

Rengarajan J, Bloom BR, Rubin EJ (2005) Genome-wide requirements for Mycobacterium tuberculosis adaptation and survival in macrophages. *Proc Natl Acad Sci U S A* **102:** 8327-8332

Ribeiro-Guimaraes ML, Pessolani MC (2007) Comparative genomics of mycobacterial proteases. *Microb Pathog* **43:** 173-178

Roman-Hernandez G, Grant RA, Sauer RT, Baker TA (2009) Molecular basis of substrate selection by the N-end rule adaptor protein ClpS. *Proc Natl Acad Sci U S A* **106:** 8888-8893

Roman-Hernandez G, Hou JY, Grant RA, Sauer RT, Baker TA (2011) The ClpS adaptor mediates staged delivery of N-end rule substrates to the AAA+ ClpAP protease. *Mol Cell* **43:** 217-228

Sachdeva P, Misra R, Tyagi AK, Singh Y (2010) The sigma factors of Mycobacterium tuberculosis: regulation of the regulators. *FEBS J* **277:** 605-626

Sala A, Bordes P, Genevaux P (2014) Multiple toxin-antitoxin systems in Mycobacterium tuberculosis. *Toxins (Basel)* **6:** 1002-1020

Sambrook J, Russel W (2001) *Molecular Cloning: A laboratory Manual*: Cold Spring Harbor Laboratory.

Sassetti CM, Boyd DH, Rubin EJ (2003) Genes required for mycobacterial growth defined by high density mutagenesis. *Mol Microbiol* **48:** 77-84

Sauer RT, Baker TA (2011) AAA+ proteases: ATP-fueled machines of protein destruction. *Annual review of biochemistry* **80:** 587-612

Schmitt EK, Riwanto M, Sambandamurthy V, Roggo S, Miault C, Zwingelstein C, Krastel P, Noble C, Beer D, Rao SP, Au M, Niyomrattanakit P, Lim V, Zheng J, Jeffery D, Pethe K, Camacho LR (2011) The natural product cyclomarin kills Mycobacterium tuberculosis by targeting the ClpC1 subunit of the caseinolytic protease. *Angewandte Chemie* **50:** 5889-5891

Schmitz KR, Carney DW, Sello JK, Sauer RT (2014) Crystal structure of Mycobacterium tuberculosis ClpP1P2 suggests a model for peptidase activation by AAA+ partner binding and substrate delivery. *Proc Natl Acad Sci U S A* **111:** E4587-4595

Schmitz KR, Sauer RT (2014) Substrate delivery by the AAA+ ClpX and ClpC1 unfoldases activates the mycobacterial ClpP1P2 peptidase. *Mol Microbiol*

Schubert OT, Mouritsen J, Ludwig C, Rost HL, Rosenberger G, Arthur PK, Claassen M, Campbell DS, Sun Z, Farrah T, Gengenbacher M, Maiolica A, Kaufmann SH, Moritz RL, Aebersold R (2013) The Mtb proteome library: a resource of assays to quantify the complete proteome of Mycobacterium tuberculosis. *Cell host & microbe* **13:** 602-612

Schuenemann VJ, Kralik SM, Albrecht R, Spall SK, Truscott KN, Dougan DA, Zeth K (2009) Structural basis of N-end rule substrate recognition in Escherichia coli by the ClpAP adaptor protein ClpS. *EMBO Rep* **10:** 508-514

Seol JH, Baek SH, Kang MS, Ha DB, Chung CH (1995) Distinctive roles of the two ATP-binding sites in ClpA, the ATPase component of protease Ti in Escherichia coli. *The Journal of biological chemistry* **270:** 8087-8092

Sherrid AM, Rustad TR, Cangelosi GA, Sherman DR (2010) Characterization of a Clp protease gene regulator and the reaeration response in Mycobacterium tuberculosis. *PloS one* **5:** e11622

Siddiqui SM, Sauer RT, Baker TA (2004) Role of the processing pore of the ClpX AAA+ ATPase in the recognition and engagement of specific protein substrates. *Genes & development* **18:** 369-374

Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Soding J, Thompson JD, Higgins DG (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular systems biology* **7:** 539

Singh R, Barry CE, 3rd, Boshoff HI (2010) The three RelE homologs of Mycobacterium tuberculosis have individual, drug-specific effects on bacterial antibiotic tolerance. *Journal of bacteriology* **192:** 1279-1291

Singh SK, Maurizi MR (1994) Mutational analysis demonstrates different functional roles for the two ATP-binding sites in ClpAP protease from Escherichia coli. *The Journal of biological chemistry* **269:** 29537-29545

Singh SK, Rozycki J, Ortega J, Ishikawa T, Lo J, Steven AC, Maurizi MR (2001) Functional domains of the ClpA and ClpX molecular chaperones identified by limited proteolysis and deletion analysis. *The Journal of biological chemistry* **276:** 29420-29429

Smith I (2003) Mycobacterium tuberculosis pathogenesis and molecular determinants of virulence. *Clin Microbiol Rev* **16:** 463-496

Smollett KL, Smith KM, Kahramanoglou C, Arnvig KB, Buxton RS, Davis EO (2012) Global analysis of the regulon of the transcriptional repressor LexA, a key component of SOS response in Mycobacterium tuberculosis. *The Journal of biological chemistry* **287:** 22004-22014

Snider J, Houry WA (2008) AAA+ proteins: diversity in function, similarity in structure. *Biochemical Society transactions* **36:** 72-77

Snider J, Thibault G, Houry WA (2008) The AAA+ superfamily of functionally diverse proteins. *Genome Biol* **9:** 216

Sousa MC, Trame CB, Tsuruta H, Wilbanks SM, Reddy VS, McKay DB (2000) Crystal and solution structures of an HslUV protease-chaperone complex. *Cell* **103:** 633-643

Sprangers R, Gribun A, Hwang PM, Houry WA, Kay LE (2005) Quantitative NMR spectroscopy of supramolecular complexes: dynamic side pores in ClpP are important for product release. *Proc Natl Acad Sci U S A* **102:** 16678-16683

Staats CC, Boldo J, Broetto L, Vainstein M, Schrank A (2007) Comparative genome analysis of proteases, oligopeptide uptake and secretion systems in Mycoplasma spp. *Genetics and Molecular Biology* **30:** 225-229

Stanne TM, Pojidaeva E, Andersson FI, Clarke AK (2007) Distinctive types of ATP-dependent Clp proteases in cyanobacteria. *The Journal of biological chemistry* **282:** 14394-14402

Stinson BM, Nager AR, Glynn SE, Schmitz KR, Baker TA, Sauer RT (2013) Nucleotide binding and conformational switching in the hexameric ring of a AAA+ machine. *Cell* **153:** 628-639

Sundar S, Baker TA, Sauer RT (2012) The I domain of the AAA+ HslUV protease coordinates substrate binding, ATP hydrolysis, and protein degradation. *Protein Science* **21:** 188-198

Suno R, Niwa H, Tsuchiya D, Zhang X, Yoshida M, Morikawa K (2006) Structure of the whole cytosolic region of ATP-dependent protease FtsH. *Mol Cell* **22:** 575-585

Vangone A, Spinelli R, Scarano V, Cavallo L, Oliva R (2011) COCOMAPS: a web application to analyze and visualize contacts at the interface of biomolecular complexes. *Bioinformatics* **27:** 2915-2916

Vasudevan D, Rao SP, Noble CG (2013) Structural basis of mycobacterial inhibition by cyclomarin A. *The Journal of biological chemistry* **288:** 30883-30891

Viala J, Mazodier P (2002) ClpP-dependent degradation of PopR allows tightly regulated expression of the clpP3 clpP4 operon in Streptomyces lividans. *Mol Microbiol* **44:** 633-643

Vieux EF, Wohlever ML, Chen JZ, Sauer RT, Baker TA (2013) Distinct quaternary structures of the AAA+ Lon protease control substrate degradation. *Proc Natl Acad Sci U S A* **110:** E2002-2008

Vijayakumar SR, Kirchhof MG, Patten CL, Schellhorn HE (2004) RpoS-regulated genes of Escherichia coli identified by random lacZ fusion mutagenesis. *Journal of bacteriology* **186:** 8499-8507

Wang J, Hartling JA, Flanagan JM (1997) The structure of ClpP at 2.3 A resolution suggests a model for ATP-dependent proteolysis. *Cell* **91:** 447-456

Wang KH, Roman-Hernandez G, Grant RA, Sauer RT, Baker TA (2008) The molecular basis of N-end rule recognition. *Mol Cell* **32:** 406-414

Waterhouse AM, Procter JB, Martin DM, Clamp M, Barton GJ (2009) Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25:** 1189-1191

Watrous J, Burns K, Liu WT, Patel A, Hook V, Bafna V, Barry CE, 3rd, Bark S, Dorrestein PC (2010) Expansion of the mycobacterial "PUPylome". *Mol Biosyst* **6:** 376-385

WHO (2015) Global tuberculosis report. *World Health Organisation, Geneva*

Wu WF, Zhou Y, Gottesman S (1999) Redundant in vivo proteolytic activities of Escherichia coli Lon and the ClpYQ (HslUV) protease. *Journal of bacteriology* **181:** 3681-3687

Ye F, Zhang J, Liu H, Hilgenfeld R, Zhang R, Kong X, Li L, Lu J, Zhang X, Li D, Jiang H, Yang CG, Luo C (2013) Helix unfolding/refolding characterizes the functional dynamics of Staphylococcus aureus Clp protease. *The Journal of biological chemistry* **288:** 17643-17653

Zeiler E, Braun N, Bottcher T, Kastenmuller A, Weinkauf S, Sieber SA (2011) Vibralactone as a tool to study the activity and structure of the ClpP1P2 complex from Listeria monocytogenes. *Angewandte Chemie* **50:** 11001-11004

Zeth K, Ravelli RB, Paal K, Cusack S, Bukau B, Dougan DA (2002) Structural analysis of the adaptor protein ClpS in complex with the N-terminal domain of ClpA. *Nature structural biology* **9:** 906-911

Zhou Y, Gottesman S (1998) Regulation of proteolysis of the stationary-phase sigma factor RpoS. *Journal of bacteriology* **180:** 1154-1158

Appendix

Appendix Table 1: Overlapping hits between the ClpC1 and ClpC2 screen. Confirmed false positives are colored red. The list is sorted by locus tag.

Locus tag	Functional category	Function	Protein name	Counts (ClpC1)	Counts (ClpC2)
MT0085.1 /Rv0078B	Conserved hypothetical protein			1	2
MT2588	CDC1551 hypothetical protein			38	1
MT3135	CDC1551 hypothetical protein			1	115
Rv0030	Conserved hypothetical protein			50	1
Rv0106	Conserved hypothetical protein			1	2
Rv0150c	Conserved hypothetical protein			1	1
Rv0494	Regulatory proteins	Probable transcriptional regulatory protein (probably GntR-family)		1	1
Rv1102c	Virulence, detoxification, adaptation	Toxin-Antitoxin	MazF3	2	1
Rv1159A	Conserved hypothetical protein	Putative pterin-4-alpha-carbinolamine dehydratase		1	9
Rv1693	Conserved hypothetical protein			20	37
Rv2012	Conserved hypothetical protein			1	19
Rv2720	Regulatory proteins	LexA repressor	LexA	1	2
Rv2922A	Intermediary metabolism and respiration	Acylphosphatase	acyP	1	1
Rv3180	Virulence, detoxification, adaptation	Toxin-Antitoxin	VapC49	1	1
Rv3559c	Intermediary metabolism and respiration	Probable oxidoreductase		43	2
Rv3586	Information pathways	DNA integrity scanning protein DisA	DisA	2	2
Rv3745c	Conserved hypothetical protein			14	1
Rv3916c	Conserved hypothetical protein			1	16

Appendix Table 2: Hit list of the ClpC1 library screen sorted by functional category. The screen was performed with ClpC1-T25 as bait against prey libraries harbouring the T18 domain either at the N-terminus of the target protein (prey vector pUT18C) or at the C-terminus of the target protein (pUT18). The category "Rv of MT" indicated if a hit found from the CDC1551 strain (MTxxxx numbering) has the corresponding gene in the H37Rv strain (Rvxxxx numbering). Counts are the number of clones found for each hit. The column "also found with bait" indicated whether the hit was also found in the ClpC2 library screen. False positive hits are colored in red.

Locus tag	Rv of MT	Functional category	Function	Protein name	Prey Vector	Counts	also found with bait
MT0196	Rv0186A	Virulence, detoxification, adaptation	Metallothionein	mymT	pUT18	2	
MT0290	Rv0277A	Virulence, detoxification, adaptation	Toxin-Antitoxin	VapB25	pUT18	1	
MT2035	Rv1982A	Virulence, detoxification, adaptation	Toxin-Antitoxin	VapB36	pUT18C	4	
MT2035	Rv1982A	Virulence, detoxification, adaptation	Toxin-Antitoxin	VapB36	pUT18	1	
MT2123	Rv2063A	Virulence, detoxification, adaptation	Toxin-Antitoxin	MazF7	pUT18	1	
MT2201	Rv2142A	Virulence, detoxification, adaptation	Toxin-Antitoxin	ParD2	pUT18C	2	
MT2676	Rv2601A	Virulence, detoxification, adaptation	Toxin-Antitoxin	VapB41	pUT18C	2	
MT3800	Rv3697A	Virulence, detoxification, adaptation	Toxin-Antitoxin	vapB48	pUT18	1	
Rv0300		Virulence, detoxification, adaptation	Toxin-Antitoxin	VapB2	pUT18C	2	
Rv0598c		Virulence, detoxification, adaptation	Toxin-Antitoxin	VapC27	pUT18C	1	
Rv1102c		Virulence, detoxification, adaptation	Toxin-Antitoxin	MazF3	pUT18C	1	ClpC2
Rv1242		Virulence, detoxification, adaptation	Toxin-Antitoxin	VapC33	pUT18C	1	
Rv1561		Virulence, detoxification, adaptation	Toxin-Antitoxin	VapC11	pUT18C	1	
Rv1720c		Virulence, detoxification, adaptation	Toxin-Antitoxin	VapC12	pUT18C	1	
Rv1741		Virulence, detoxification, adaptation	Toxin-Antitoxin	VapC34	pUT18C	1	
Rv1838c		Virulence, detoxification, adaptation	Toxin-Antitoxin	VapC13	pUT18C	1	
Rv1960c		Virulence, detoxification, adaptation	Toxin-Antitoxin	ParD1	pUT18C	1	
Rv1962c		Virulence, detoxification, adaptation	Toxin-Antitoxin	VapC35	pUT18C	1	
Rv2010		Virulence, detoxification, adaptation	Toxin-Antitoxin	VapC15	pUT18C	1	
Rv2548		Virulence, detoxification, adaptation	Toxin-Antitoxin	vapC19	pUT18C	1	
Rv2550c		Virulence, detoxification, adaptation	Toxin-Antitoxin	VapB20	pUT18	5	
Rv2596		Virulence, detoxification, adaptation	Toxin-Antitoxin	VapC40	pUT18C	1	
Rv2863		Virulence, detoxification, adaptation	Toxin-Antitoxin	VapC23	pUT18C	1	

Locus tag	Rv of MT	Functional category	Function	Protein name	Prey Vector	Counts	also found with bait
Rv2872		Virulence, detoxification, adaptation	Toxin-Antitoxin	VapC43	pUT18	1	
Rv3180		Virulence, detoxification, adaptation	Toxin-Antitoxin	VapC49	pUT18C	1	ClpC2
Rv3320c		Virulence, detoxification, adaptation	Toxin-Antitoxin	VapC44	pUT18C	1	
Rv0081		Regulatory proteins	Probable transcriptional regulatory protein		pUT18C	1	
Rv0474		Regulatory proteins	XRE family transcriptional regulator		pUT18C	2	
Rv0494		Regulatory proteins	Probable transcriptional regulatory protein (probably GntR-family)		pUT18C	1	ClpC2
Rv0576		Regulatory proteins	Probable transcriptional regulatory protein (possibly ArsR-family)		pUT18	1	
Rv2282c		Regulatory proteins	Probable transcription regulator (LysR family)		pUT18	1	
Rv2720		Regulatory proteins	LexA repressor	LexA	pUT18C	1	ClpC2
Rv3291c		Regulatory proteins	Probable transcriptional regulatory protein LrpA (Lrp/AsnC-family)	LrpA	pUT18C	1	
Rv3334		Regulatory proteins	Probable transcriptional regulatory protein (probably MerR-family)		pUT18C (4), pUT18 (1)	5	
Rv1089		PE/PPE	PE family protein PE10	PE10	pUT18C	2	
Rv1806		PE/PPE	PE family protein PE20	PE20	pUT18C	1	
Rv2107		PE/PPE	PE family protein PE22	PE22	pUT18	1	
Rv3425		PE/PPE	PPE family protein PPE57	PPE57	pUT18C	1	
Rv0230c		Lipid metabolism	Probable phosphotriesterase Php (parathion hydrolase)	PhP	pUT18	1	
Rv0974c		Lipid metabolism	Probable acetyl-/propionyl-CoA carboxylase (beta subunit) AccD2	AccD2	pUT18	1	
Rv1141c		Lipid metabolism	Probable enoyl-CoA hydratase EchA11	EchA11	pUT18	1	
Rv3373		Lipid metabolism	Probable enoyl-CoA hydratase EchA18 (enoyl hydrase) (unsaturated acyl-CoA hydratase) (crotonase)	EchA18	pUT18	1	
MT0543	Rv0521	Intermediary metabolism and respiration	Possible methyltransferase/methylase (fragment)	Rv0521	pUT18	2	
Rv0253		Intermediary metabolism and respiration	Probable nitrite reductase [NAD(P)H] small subunit NirD	NirD	pUT18C	1	
Rv0636		Intermediary metabolism and respiration	(3R)-hydroxyacyl-ACP dehydratase subunit HadB	HadB	pUT18C	1	
Rv0771		Intermediary metabolism and respiration	4-carboxymuconolactone decarboxylase		pUT18C	1	
Rv0793		Intermediary metabolism and respiration	Antibiotic biosynthesis monooxygenase		pUT18C	2	

Locus tag	Rv of MT	Functional category	Function	Protein name	Prey Vector	Counts	also found with bait
Rv0808		Intermediary metabolism and respiration	Amidophosphoribosyltransferase PurF	PurF	pUT18	1	
Rv0864		Intermediary metabolism and respiration	Molybdenum cofactor biosynthesis protein MoaC	MoaC2	pUT18C	1	
Rv0956		Intermediary metabolism and respiration	Probable 5'-phosphoribosylglycinamide formyltransferase PurN (GART)	PurN	pUT18C	1	
Rv1095		Intermediary metabolism and respiration	Probable PHOH-like protein PhoH2	PhoH2	pUT18C	3	
Rv1177		Intermediary metabolism and respiration	Probable ferredoxin FdxC	FdxC	pUT18C	1	
Rv1334		Intermediary metabolism and respiration	CysO-cysteine peptidase	Mec	pUT18C	1	
Rv1336		Intermediary metabolism and respiration	Cysteine synthase B CysM (CSASE B)	CysM	pUT18C	1	
Rv1403c		Intermediary metabolism and respiration	Putative methyltransferase		pUT18C	1	
Rv1897c		Intermediary metabolism and respiration	D-aminoacyl-tRNA deacylase	Dtd	pUT18C	1	
Rv2074		Intermediary metabolism and respiration	Possible pyridoxamine 5'-phosphate oxidase (PNP/PMP oxidase)		pUT18C	1	
Rv2122c		Intermediary metabolism and respiration	Phosphoribosyl-AMP pyrophosphatase HisE	HisE	pUT18C	3	
Rv2445c		Intermediary metabolism and respiration	Probable nucleoside diphosphate kinase NdkA	NdkA	pUT18C	1	
Rv2747		Intermediary metabolism and respiration	Probable L-glutamate alpha-N-acetyltranferase ArgA	ArgA	pUT18C (3), pUT18 (1)	4	
Rv2922A		Intermediary metabolism and respiration	Acylphosphatase	acyP	pUT18	1	ClpC2
Rv3559c		Intermediary metabolism and respiration	Probable oxidoreductase		pUT18	1	ClpC2
Rv3567c		Intermediary metabolism and respiration	Flavin-dependent monooxygenase, reductase subunit HsaB	HsaB	pUT18C	1	
Rv1036c		Insertion seqences and phages	Probable IS1560 transposase (fragment)		pUT18C	2	
Rv1149		Insertion seqences and phages	Possible transposase (Rv1042c,Rv1149)		pUT18C	1	
Rv1579c		Insertion seqences and phages	Probable PhiRv1 phage protein	PhiRv1	pUT18C	1	
Rv2424c		Insertion seqences and phages	Probable transposase		pUT18C	1	
MT3386	Rv3287c	Information pathways	Anti-sigma-F factor RsbW	RsbW	pUT18C	1	
Rv0005	Rv0005	Information pathways	DNA gyrase subunit B	GyrB	pUT18	1	

Locus tag	Rv of MT	Functional category	Function	Protein name	Prey Vector	Counts	also found with bait
Rv0233		Information pathways	Ribonucleoside-diphosphate reductase (beta chain) NrdB	NrdB	pUT18C	2	
Rv0708		Information pathways	50S ribosomal protein L16 RplP	RplP	pUT18	2	
Rv2362c		Information pathways	Possible DNA repair protein RecO	RecO	pUT18C	2	
Rv2816c		Information pathways	CRISPR-associated endoribonuclease Cas2	Cas2	pUT18	2	
Rv2822c		Information pathways	CRISPR type III-associated protein Csm2	Csm2	pUT18C	4	
Rv3586		Information pathways	DNA integrity scanning protein DisA	DisA	pUT18C (10), pUT18 (33)	43	ClpC2
MT0085	Rv0078A	Conserved hypothetical protein			pUT18C	1	
MT0085.1	Rv0078B	Conserved hypothetical protein			pUT18	4	ClpC2
MT1364	Rv1322A	Conserved hypothetical protein	4-hydroxyphenylpyruvate dioxygenase C terminal domain containing protein		pUT18C	2	
MT1547	Rv1498A	Conserved hypothetical protein			pUT18C	2	
Rv0030		Conserved hypothetical protein			pUT18	1	ClpC2
Rv0034		Conserved hypothetical protein			pUT18C	1	
Rv0106		Conserved hypothetical protein			pUT18	50	ClpC2
Rv0141c		Conserved hypothetical protein			pUT18C	1	
Rv0150c		Conserved hypothetical protein			pUT18	1	ClpC2
Rv0164		Conserved hypothetical protein	Conserved protein TB18.5		pUT18C	1	
Rv0201c		Conserved hypothetical protein			pUT18C	2	
Rv0250c		Conserved hypothetical protein			pUT18	1	
Rv0313		Conserved hypothetical protein			pUT18C	9	
Rv0424c		Conserved hypothetical protein			pUT18C	1	
Rv0466		Conserved hypothetical protein	Acyl-ACP thioesterase		pUT18C	1	
Rv0628c		Conserved hypothetical protein			pUT18C	1	
Rv0692		Conserved hypothetical protein	Mycofactocin system protein MftB	MftB	pUT18C	2	
Rv1159.1		Conserved hypothetical protein	Pterin-4-alpha-carbinolamine dehydratase		pUT18C	1	
Rv1159A		Conserved hypothetical protein	Putative pterin-4-alpha-carbinolamine dehydratase		pUT18C	2	ClpC2
Rv1489		Conserved hypothetical protein			pUT18C	1	

Locus tag	Rv of MT	Functional category	Function	Protein name	Prey Vector	Counts	also found with bait
Rv1506c		Conserved hypothetical protein			pUT18C	1	
1693		Conserved hypothetical protein			pUT18C	1	ClpC2
Rv1727		Conserved hypothetical protein			pUT18C	1	
Rv1873		Conserved hypothetical protein	Calpastatin		pUT18C	1	
Rv1904		Conserved hypothetical protein	Anti-sigma factor antagonist		pUT18	1	
Rv1948c		Conserved hypothetical protein			pUT18C	2	
Rv1950c		Conserved hypothetical protein			pUT18C	1	
Rv1975		Conserved hypothetical protein			pUT18C	1	
Rv2012		Conserved hypothetical protein			pUT18	20	ClpC2
Rv2078		Conserved hypothetical protein			pUT18C	1	
Rv2117		Conserved hypothetical protein			pUT18 (2), pUT18C (1)	3	
Rv2137c		Conserved hypothetical protein			pUT18C	1	
Rv2184c		Conserved hypothetical protein			pUT18C	1	
Rv2312		Conserved hypothetical protein			pUT18C	1	
Rv2475c		Conserved hypothetical protein			pUT18C	1	
Rv2663		Conserved hypothetical protein			pUT18C	1	
Rv2708c		Conserved hypothetical protein			pUT18C	4	
Rv2803		Conserved hypothetical protein			pUT18C	3	
Rv2901c		Conserved hypothetical protein			pUT18C	2	
Rv2908c		Conserved hypothetical protein	RNA binding		pUT18	1	
Rv2923c		Conserved hypothetical protein			pUT18	1	
Rv2974c		Conserved hypothetical protein	Conserved hypothetical alanine rich protein		pUT18C	1	
Rv2975c		Conserved hypothetical protein			pUT18C	2	
Rv3231c		Conserved hypothetical protein			pUT18C	1	
Rv3745c		Conserved hypothetical protein			pUT18	1	ClpC2
Rv3745c		Conserved hypothetical protein			pUT18C	1	ClpC2
Rv3916c		Conserved hypothetical protein			pUT18 (11), pUT18C (3)	14	ClpC2

Locus tag	Rv of MT	Functional category	Function	Protein name	Prey Vector	Counts	also found with bait
Rv1496		Cell wall and Cell processes	Possible transport system kinase		pUT18C	1	
Rv1687c		Cell wall and Cell processes	Probable conserved ATP-binding protein ABC transporter		pUT18C	1	
Rv1690		Cell wall and Cell processes	Probable lipoprotein LprJ	LprJ	pUT18C	1	
Rv2341		Cell wall and Cell processes	Probable conserved lipoprotein LppQ	LppQ	pUT18C	1	
Rv2965c		Cell wall and Cell processes	Phosphopantetheine adenylyltransferase	CoaD	pUT18C	1	
I6YCL8_MYCTU	LH57_04325	CDC1551 hypothetical protein	Transposase		pUT18C	1	
MT0009		CDC1551 hypothetical protein			pUT18C	2	
MT0116.1		CDC1551 hypothetical protein			pUT18	2	
MT0294		CDC1551 hypothetical protein			pUT18	1	
MT0416		CDC1551 hypothetical protein			pUT18C	1	
MT0494		CDC1551 hypothetical protein			pUT18	1	
MT0553	Rv0530A	CDC1551 hypothetical protein			pUT18	2	
MT0638.1		CDC1551 hypothetical protein			pUT18	2	
MT0726.1		CDC1551 hypothetical protein			pUT18C	1	
MT0761		CDC1551 hypothetical protein			pUT18	1	
MT0910.4		CDC1551 hypothetical protein			pUT18C	3	
MT1077		CDC1551 hypothetical protein			pUT18 (2), pUT18C (1)	3	
MT1172.1		CDC1551 hypothetical protein			pUT18C	1	
MT1178		CDC1551 hypothetical protein			pUT18 (2), pUT18C (2)	4	
MT1367.1		CDC1551 hypothetical protein			pUT18	1	
MT1560.1		CDC1551 hypothetical protein			pUT18	1	
MT2011		CDC1551 hypothetical protein			pUT18	1	
MT2316		CDC1551 hypothetical protein			pUT18	1	
MT2370.1		CDC1551 hypothetical protein			pUT18	1	
MT2514		CDC1551 hypothetical protein			pUT18C	1	
MT2547.2		CDC1551 hypothetical protein			pUT18 (1), pUT18C (1)	2	
MT2588		CDC1551 hypothetical protein			pUT18	1	

Locus tag	Rv of MT	Functional category	Function	Protein name	Prey Vector	Counts	also found with bait
MT2625		CDC1551 hypothetical protein			pUT18	1	
MT2637.1		CDC1551 hypothetical protein			pUT18C	2	ClpC2
MT2779	LH57_14815	CDC1551 hypothetical protein			pUT18 (2), pUT18C (2)	2	
MT2960		CDC1551 hypothetical protein			pUT18	1	
MT3080.1		CDC1551 hypothetical protein			pUT18	1	
MT3102		CDC1551 hypothetical protein			pUT18	1	
MT3135		CDC1551 hypothetical protein			pUT18	38	
MT3280		CDC1551 hypothetical protein			pUT18C	1	
MT3378		CDC1551 hypothetical protein			pUT18	1	ClpC2
MT3429		CDC1551 hypothetical protein			pUT18C	1	
MT3437.1		CDC1551 hypothetical protein			pUT18	1	
MT3491.1		CDC1551 hypothetical protein			pUT18C	1	
MT3510.1		CDC1551 hypothetical protein			pUT18	2	
MT3535		CDC1551 hypothetical protein			pUT18	2	
MT3671.2		CDC1551 hypothetical protein			pUT18	1	
MT3718.2		CDC1551 hypothetical protein			pUT18	2	
MT3767.3		CDC1551 hypothetical protein			pUT18	3	
MT3952		CDC1551 hypothetical protein			pUT18C	2	
MT3974.1		CDC1551 hypothetical protein			pUT18C	1	
MT4019		CDC1551 hypothetical protein			pUT18	1	

Appendix Table 3: Hit list of the ClpC2 library screen sorted by functional category The screen was performed with either ClpC2-T25 (bait vector pKNT25) or T25-ClpC2 (bait vector pKT25) as bait against a prey librariy where the the T18 domain is located at the C-terminus of the target protein (pUT18). The category "Rv of MT" indicated if a hit found from the CDC1551 strain (MTxxxx numbering) has the corresponding gene in the H37Rv strain (Rvxxxx numbering). Counts are the number of clones found for each hit. The column "also found with bait" indicated whether the hit was also found in the ClpC1 library screen. False positive hits are colored in red.

Locus Tag	Rv of Mt	Functional Category	Function	Protein name	Bait Vector	Counts	also found with bait
MT0196	Rv0186A	Virulence, detoxification, adaptation	Metallothionein mymT	MymT	pKT25	2	
MT0987	Rv0959A	Virulence, detoxification, adaptation	Toxin-Antitoxin	VapB9	pKT25	1	
Rv1102c		Virulence, detoxification, adaptation	Toxin-Antitoxin	MazF3	pKT25	1	ClpC1
Rv1247c		Virulence, detoxification, adaptation	Toxin-Antitoxin	RelB1	pKNT25	1	
Rv1546		Virulence, detoxification, adaptation	Toxin-Antitoxin	Rv1546	pKT25	1	
Rv1721c		Virulence, detoxification, adaptation	Toxin-Antitoxin	VapB12	pKT25	2	
Rv1740		Virulence, detoxification, adaptation	Toxin-Antitoxin	VapB34	pKT25	1	
Rv1952		Virulence, detoxification, adaptation	Toxin-Antitoxin	VapB14	pKT25	2	
Rv2104c		Virulence, detoxification, adaptation	Toxin-Antitoxin	VapB37	pKT25	1	
Rv3180c		Virulence, detoxification, adaptation	Toxin-Antitoxin	VapC49	pKT25	1	ClpC1
Rv3357		Virulence, detoxification, adaptation	Toxin-Antitoxin	YefM	pKT25	1	
Rv0491		Regulatory proteins	Two component sensory transduction protein RegX3	RegX3	pKT25	1	
Rv0494		Regulatory proteins	Probable transcriptional regulatory protein (probably GntR-family)		pKT25	1	ClpC1
Rv1909c		Regulatory proteins	Ferric uptake regulation protein FurA	FurA	pKT25	4	
Rv2234		Regulatory proteins	Phosphotyrosine protein phosphatase PtpA	PtpA	pKT25	1	
Rv2250c		Regulatory proteins	Possible transcriptional regulatory protein		pKT25	1	
Rv2359		Regulatory proteins	Probable zinc uptake regulation protein Zur	FurB	pKT25	1	
Rv2720		Regulatory proteins	LexA repressor	LexA	pKT25	2	ClpC1
Rv1214c		PE/PPE	PE family protein PE14	PE14	pKNT25	1	
Rv0644c		Lipid metabolism	Methoxy mycolic acid synthase 2 MmaA2	MmaA2	pKT25	1	
Rv1528c		Lipid metabolism	Probable conserved polyketide synthase associated protein PapA4		pKT25	1	
Rv0013		Intermediary metabolism and respiration	Possible anthranilate synthase component II TrpG (glu-tamine amidotransferase)	TrpG	pKT25	1	

Locus Tag	Rv of Mt	Functional Category	Function	Protein name	Bait Vector	Counts	also found with bait
Rv0772		Intermediary metabolism and respiration	Probable phosphoribosylamine--glycine ligase PurD (GARS)	PurD	pKT25	1	
Rv2007c		Intermediary metabolism and respiration	Ferredoxin FdxA	FdxA	pKT25	1	
Rv2096		Intermediary metabolism and respiration	Proteasome accessory factor B	PafB	pKT25	1	
Rv2435		Intermediary metabolism and respiration	Probable cyclase (adenylyl- or guanylyl-)		pKNT25	1	
Rv2667		Intermediary metabolism and respiration		ClpC2	pKT25 (1), pKNT25 (7)	8	
Rv2922A		Intermediary metabolism and respiration	Probable acylphosphatase AcyP (acylphosphate phosphohydrolase)	AcyP	pKT25	1	ClpC1
Rv3028c		Intermediary metabolism and respiration	Probable electron transfer flavoprotein (alpha-subunit) FixB	FixB	pKT25	1	
Rv3250c		Intermediary metabolism and respiration	Probable rubredoxin RubB	RubB	pKT25	3	
Rv3535c		Intermediary metabolism and respiration	Probable acetaldehyde dehydrogenase	HsaG	pKT25	1	
Rv3559c		Intermediary metabolism and respiration	Probable oxidoreductase		pKT25	2	ClpC1
Rv3592		Intermediary metabolism and respiration	Possible heme degrading protein MhuD	MhuD	pKNT25	1	
Rv3602c		Intermediary metabolism and respiration	Pantoate--beta-alanine ligase PanC	PanC	pKT25	1	
Rv3914		Intermediary metabolism and respiration	Thioredoxin TrxC (TRX)	TrxC	pKT25	1	
Rv1757c		Insertion seqences and phages	Putative transposase for insertion sequence element IS6110 (fragment)		pKT25	1	
Rv2925c		Information pathways	Probable ribonuclease III Rnc (RNase III)	Rnc	pKT25	5	
Rv0030		Conserved hypothetical protein			pKT25	1	ClpC1
Rv0106		Conserved hypothetical protein			pKT25	2	ClpC1
Rv0150c		Conserved hypothetical protein			pKT25	1	ClpC1
Rv0699		Conserved hypothetical protein			pKNT25	1	
Rv0740		Conserved hypothetical protein			pKT25	1	
Rv1159A		Conserved hypothetical protein	Putative pterin-4-alpha-carbinolamine dehydratase		pKT25	9	ClpC1
Rv1192		Conserved hypothetical protein	Alpha/beta hydrolase		pKT25	1	
Rv1590		Conserved hypothetical protein			pKT25	1	

Locus Tag	Rv of Mt	Functional Category	Function	Protein name	Bait Vector	Counts	also found with bait
Rv1693		Conserved hypothetical protein			pKT25	37	ClpC1
Rv1805c		Conserved hypothetical protein			pKNT25	1	
Rv1847		Conserved hypothetical protein			pKT25	4	
Rv2012		Conserved hypothetical protein			pKT25 (5), pKNT25 (14)	5	ClpC1
Rv2562		Conserved hypothetical protein			pKT25	1	
Rv2774		Conserved hypothetical protein			pKT25	1	
Rv3129		Conserved hypothetical protein			pKT25	1	
Rv3258c		Conserved hypothetical protein			pKT25	1	
Rv3586		Conserved hypothetical protein			pKT25	2	ClpC1
Rv3643		Conserved hypothetical protein			pKT25	1	
Rv3678A		Conserved hypothetical protein			pKT25	1	
Rv3745c		Conserved hypothetical protein			pKT25	1	ClpC1
Rv3916c		Conserved hypothetical protein			pKT25	16	ClpC1
MT2038	Rv1984c	Cell wall and Cell processes	Probable cutinase precursor CFP21	Cfp21	pKNT25	1	
Rv0476		Cell wall and Cell processes	Possible conserved transmembrane protein		pKT25	1	
Rv1018		Cell wall and Cell processes	UDP-N-acetylglucosamine pyrophosphorylase GlmU	GlmU	pKT25	2	
Rv1946c		Cell wall and Cell processes	Possible lipoprotein	LppG	pKT25	1	
Rv2077c		Cell wall and Cell processes	Possible conserved transmembrane protein		pKT25	1	
Rv3619c		Cell wall and Cell processes	Putative ESAT-6 like protein EsxV	EsxV	pKT25	1	
MT0085.1	Rv0078B	CDC1551 hypothetical protein			pKT25	2	ClpC1
MT0294		CDC1551 hypothetical protein			pKT25	1	
MT0325		CDC1551 hypothetical protein			pKT25	3	
MT0600		CDC1551 hypothetical protein			pKT25	1	
MT0603		CDC1551 hypothetical protein			pKT25	1	
MT0717.1		CDC1551 hypothetical protein			pKNT25	1	
MT0768.1		CDC1551 hypothetical protein			pKT25	1	
MT0853		CDC1551 hypothetical protein			pKT25	1	

Locus Tag	Rv of Mt	Functional Category	Function	Protein name	Bait Vector	Counts	also found with bait
MT1025.2		CDC1551 hypothetical protein			pKT25	1	
MT1367.1		CDC1551 hypothetical protein			pKT25	1	
MT1409		CDC1551 hypothetical protein			pKT25 (1), pKNT25 (2)	3	
MT1578.1		CDC1551 hypothetical protein			pKT25	1	
MT1821.1		CDC1551 hypothetical protein			pKT25	2	
MT2045		CDC1551 hypothetical protein			pKT25	1	
MT2165.1		CDC1551 hypothetical protein			pKT25	1	
MT2283		CDC1551 hypothetical protein			pKT25	3	
MT2325		CDC1551 hypothetical protein			pKT25	3	
MT2467	Rv2395B	CDC1551 hypothetical protein	Acid and phagosome regulated protein B AprB	AprB	pKT25	3	
MT2588		CDC1551 hypothetical protein			pKT25	1	ClpC1
MT2625	Rv2548A	CDC1551 hypothetical protein			pKT25	4	
MT2722		CDC1551 hypothetical protein			pKT25	2	
MT2779		CDC1551 hypothetical protein			pKNT25	1	
MT2960		CDC1551 hypothetical protein			pKT25	1	
MT2993		CDC1551 hypothetical protein			pKT25	1	
MT3135		CDC1551 hypothetical protein			pKT25 (109), pKNT25 (6)	115	ClpC1
MT3139.1		CDC1551 hypothetical protein			pKT25	1	
MT3269		CDC1551 hypothetical protein			pKT25	1	
MT3273		CDC1551 hypothetical protein			pKNT25	1	
MT3780	Rv3678A	CDC1551 hypothetical protein			pKT25	3	

Abbreviatons

A	Absorption
Å	Ångström
AAA+	ATPase associated with various cellular activities
ADEP	Acyldepsipeptide
ADP	Adenosine diphosphate
ATP	Adenosine triphosphate
BACTH	Bacterial adenylate cyclase two hybrid
Clp	Caseinolytic protease
cAMP	cyclic adenosine monophosphate
ClgR	Clp gene regulator
CtrA	Cell cycle transcriptional regulator
Cya-	Adenylate cyclase knock-out
DMSO	Dimethyl sulfoxide
DTT	Dithiothreitol
GFP	Green Fluorescent Protein
IPTG	Isopropyl β-D-1-thiogalactopyranoside
kDa	kilo Dalton
mClpP1P2	mature ClpP1P2 where the propeptides are processed
Mtb	*Mycobacterium tuberculosis*
RssB	Response regulator of RpoS
SspB	Stringent starvation protein B
SigE	Sigma factor E
SigH	Sigma factor H
Pup	Prokaryotic ubiquitin-like protein
PMSF	phenylmethylsulfonyl fluoride
TEV	Tobacco Etch Virus
Z-Leu-Leu-H	Benzyloxycarbonyl-L-Leucyl-L-Leucinal
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside

Curriculum Vitae

Personal information

Name:	Julia Leodolter
Email address:	lejulia@mol.biol.ethz.ch
Date and Place of Birth:	23.07.1985, Vienna
Nationality:	Austrian
Address:	Institute of Molecular Biology and Biophysics
	Otto-Stern-Weg 5
	8093 Zürich

Education

- 2011-2016: PhD studies, ETH Zürich, Institute of Molecular Biology and Biophysics
- 2003-2010: Study of Molecular Biology, University of Vienna (Mag. rer. nat/MSc)
- 1995-2003: Grammar/High School, BG/BRG Schwechat, Austria

Selected Research experience

- November 2011-present: ETH Zürich, Institute of Molecular Biology and Biophysics
  PhD thesis: "Architecture, Assembly and Interactors of the *Mycobacterium tuberculosis* Clp chaperone-proteases"
  Advisor: Prof. Eilika Weber-Ban
- July 2009-December 2010: Institute of Molecular Pathology (IMP), Vienna
  Master thesis: "*In vitro* reconstitution of DegP-OMP complexes from *Escherichia coli"*
  Advisor: Dr. Tim Clausen

Selected Presentations

- **Poster** presentation at the 2015 EMBO conference "Molecular chaperones: From molecules to cells and misfolding diseases", Heraklion, Crete. Poster title: *"The Mycobacterium tuberculosis ClpX and ClpC1 chaperones interact asymmetrically with the ClpP1P2 proteolytic core by docking onto the ClpP2 face"*
- **Poster** presentation at the 2014 D-BIOL Symposium, Davos, Switzerland. Poster title: *"Investigation of the Clp chaperone-protease complex of Mycobacterium tuberculosis"*
- **Talk** at the 2015 annual retreat of the Biomolecular Structure and Mechanism (BSM) PhD program: *"Asymmetry in the Clp chaperone-protease complex of Mycobacterium tuberculosis"*

PUBLICATIONS

- **Leodolter J**, Warweg J, Weber-Ban E (2015). The *Mycobacterium tuberculosis* ClpP1P2 Protease Interacts Asymmetrically with Its ATPase Partners ClpX and ClpC1. PLoS ONE 10(5)
- Laederach, J**.**, **Leodolter, J**., Warweg, J., and Weber-Ban, E. (2014). Chaperone-Proteases of Mycobacteria. *In:* The Molecular Chaperones Interaction Networks in Protein Folding and Degradation, W.A. Houry, ed. (Springer New York), pp. 419-444.

TEACHING EXPERIENCE

Supervision of Msc students:

- September 2012-December 2012: Ramon Hochstrasser, Research Project I. Title: "*Investigation of pupylation of the Mycobacterium tuberculosis chaperonin GroES."*
- September 2013-December 2013/October 2014-Mai 2015: Anne Kerschenmeyer, Research Project I and Msc thesis.Title: "*Identification of possible adaptor proteins for the Mycobacterium tuberculosis Clp system and their function."*

Teaching Basic Courses

- February 2012-present: Yearly Practical Couse "Grundlagen der Biologie I"
- Yearly Blockcourse "Protein Folding, Assembly and Degradation"