Structural and Functional Studies of the ATP-dependent Clp Proteases in Cyanobacteria

Frida M Ståhlberg

GÖTEBORGS UNIVERSITET

FACULTY OF SCIENCE
DEPARTMENT OF BIOLOGICAL AND ENVIRONMENTAL SCIENCES

Akademisk avhandling för filosofie doktorsexamen i Naturvetenskap med inriktning Biologi, som med tillstånd från Naturvetenskapliga fakulteten kommer att offentligt försvaras fredagen den 26 september 2014 kl. 10.00 i Hörsalen, Institutionen för biologi och miljövetenskap, Carl Skottsbergs gata 22B, Göteborg.

Examinator: Professor Cornelia Spetea Wiklund, Institutionen för biologi och miljövetenskap, Göteborgs Universitet
Fakultetsopponent: Professor Hanne Ingmer, Department of veterinary disease biology, food safety and Zoonoses, University of Copenhagen

ISBN 978-91-85529-68-1

ISBN: 978-91-85529-68-1
Tryck: Ineko AB, Göteborg
E-publicerad: http://hdl.handle.net/2077/36524

Till Erene, Tommy och Iris

Structural and Functional Studies of the ATP-dependent Clp Proteases in Cyanobacteria

Frida M Ståhlberg
Gothenburg University, Department of Biological and Environmental Sciences
Box 461, SE-405 30 Gothenburg, Sweden

ABSTRACT

Proteins are essential in all living organisms and they are involved in a myriad of biological functions. It is vital for cells to have efficient surveillance and quality control systems that ensure damaged proteins are either repaired to their functional state or quickly removed by degradation. Two crucial components of these protein quality systems are molecular chaperones and proteases, of which one major contributor is the AAA+ (ATPases Associated with diverse cellular Activities) family that includes the Clp proteases.

The Clp protease exists in many forms of life, ranging from eubacteria to mammals. In the bacterium *E. coli*, the hexameric ATPases ClpX and ClpA recognize the substrate and once unfolded translocate it into the proteolytic core, which is formed by two heptameric rings of ClpP. The complexity of Clp proteases in terms of both composition and functionality is far greater in photosynthetic organisms compared with their bacterial counterparts. This is highlighted in the cyanobacterium *Synechococcus elongatus* (*Synechococcus*), which has at least two Clp proteases, the essential ClpCP3/R and the non-essential ClpXP1/P2. Of these various Clp proteins, the ClpR subunit is unique to photosynthetic organisms and is proteolytically inactive, while the existence of two ClpS adaptors (ClpS1 and ClpS2) is also unique for cyanobacteria. This thesis focuses on the continued characterization of these Clp proteins in *Synechococcus*.

In paper I, two conserved motifs in the ClpR and one motif in the ClpP3 N-terminus were identified as being crucial for association to ClpC. It was also shown that these motifs were important for the stability of the ClpP3/R complex. A C-terminal motif in ClpC (the R-domain) was also demonstrated as conferring the specificity for ClpP3/R association. In paper II, the subunit stoichiometry of the ClpP1/P2 proteolytic core was determined by non-denaturing mass spectrometry. The proteolytic core was composed of an equal amount of ClpP1 and ClpP2 subunits arranged in an alternating pattern within each heptameric ring. The two double heptameric rings had dual stoichiometry, where two different proteolytic cores could be formed, (4ClpP1+3ClpP2) + (3ClpP1+4ClpP2) and 2×(3ClpP1+4ClpP2). In paper III, the functionality of the ClpP1/P2 protease was further characterized *in vitro*. ClpP1/P2 displayed the expected proteolytic activity with ClpX, but no activity was observed with ClpC. Both types of ClpP subunit appear to contribute to the proteolytic activity of the ClpP1/P2 core, but the arrangement of these two ClpP paralogs somehow limits the overall degradation rate. It was also revealed that ClpP2 alone could not assemble into higher molecular mass complexes, whereas ClpP1 readily formed a homo-tetradecamer that was proteolytically active with both ClpC and ClpX but whose activity was dependent on increased Mg^{2+} concentrations. In paper IV, the cyanobacterial-specific ClpS2 adaptor was shown to be a relatively low-abundant, soluble protein that is essential for phototrophic growth. Like ClpS1, ClpS2 redirects the general substrate specificity of ClpC to N-end rule substrates, but the two adaptors differ in exactly which N-end rule substrates they target. ClpS1 recognizes Phe and Tyr as destabilizing amino acids, while ClpS2 recognizes Leu. In the final paper (paper V), the Δ*clpS*1 and Δ*clpP2* mutants are shown to have greater resistance to exogenously added H_2O_2, while Δ*clpP1* was extremely sensitive. The different phenotypes of these mutants were dependent on the level of the catalase peroxidase KatG, where elevated basal expression of the *katG* gene was responsible for the resistance observed in Δ*clpS1* and Δ*clpP2*. In contrast, increased proteolysis of the KatG protein in Δ*clpP1* caused the extreme sensitivity of this mutant to the oxidative stress

ISBN: 978-91-85529-68-1

LIST OF PUBLICATIONS

This thesis is based on the following papers which are referred to by their Roman numerals in the text:

I. Tryggvesson A.[1], **Ståhlberg F.M.**[1], Mogk A., Zeth K. and Clarke A.K. (2012). Interaction specificity between the chaperone and proteolytic components of the cyanobacterial Clp protease. Biochem. J. 446(2):311-20.[*]

II. Mikhailov A.[1], **Ståhlberg F.M.**[1], Clarke A.K, Robinson C.V. Mass spectrometry reveals dual stoichiometry of the ClpP1/P2 protease from the cyanobacterium *Synechococcus elongatus*. *Manuscript*

III. **Ståhlberg F.M.**, Tanabe N., Lymperopoulos P., Mogk A., Zeth K. and Clarke A.K. Functional characterization of the ClpP1 and ClpP2 proteins from *Synechococcus elongatus*. *Manuscript*

IV. Tryggvesson A., **Ståhlberg F.M**., Tanabe N., Mogk A. and Clarke A.K. Characterization of ClpS2, an essential adaptor protein for the cyanobacterium *Synechococcus elongatus. Manuscript*

V. **Ståhlberg F.M.,** Javahari S. and Clarke A. K. Functional studies of the ClpS1 adaptor protein in the cyanobacterium *Synechococcus* and its importance during oxidative stress. *Manuscript*

[1] Both authors contributed equally to this work

[*] Reprinted with permission from Biochemical Journal©

TABLE OF CONTENTS

TABLE OF CONTENTS

1. INTRODUCTION

Proteins are essential macromolecules in all living organisms. They are involved in a diverse array of functions, being integral components of membrane structures and active participants in many different cellular processes. It is crucial to cell integrity that proteins remain active during their lifetime and that non-functional variants, resulting from misfolding or other forms of structural damage are quickly removed. If such abnormal proteins are not efficiently eliminated, they can accumulate and jeopardize cell viability by interfering with different cellular activities. This means that surveillance and quality control systems are needed to ensure that damaged proteins are either repaired to a functional state or removed by degradation. As a consequence, all proteins have a certain lifespan and cell homeostasis relies on their constant turnover. This balance between cellular protein synthesis and degradation is termed proteostasis.

Two key components underlying the protein quality control systems in all organisms are molecular chaperones and proteases. Chaperones affect protein structures in many different ways and often require ATP to fuel their activities. They help proteins to correctly fold into their active form and facilitate processes such as organellar protein import. Chaperones also monitor protein structures throughout the cell and can rescue those that inadvertently denature or misfold. This function is particularly important during periods of stress when the occurrence of such damaged proteins is more prevalent. At certain times, at the end of a protein's lifespan or if it is irreversibly damaged, chaperones can also facilitate the degradation of these proteins by specific proteases. Many different families of proteases also exist in the cell and they perform a multitude of roles. They are not only important for removing unwanted proteins, proteases are also necessary for processing certain enzymes and regulatory proteins to their active form. The degradation products from proteolysis can also act as regulatory signals that ultimately affect gene expression (Wickner et al., 1999; Gottesman et al., 1997).

Proteases degrade proteins by breaking the peptide bond between adjoining amino acids, the building blocks of all proteins. Proteases can be designated as either endo- or exopeptidases, depending on the position of the peptide bond cleaved within the polypeptide chain. Endopeptidases cleave the peptide bond of interior amino acids within the polypeptide, typically generating peptide fragments of varying length. Exopeptidases conversely break the terminal peptide bond at either end of the protein and generate single amino acids that can be recycled to support nascent protein synthesis. Proteases can also be mechanistically classified by the type of catalytic site used to cleave the peptide bonds of substrate proteins. There are six such groups: aspartate-, cysteine-, glutamic acid-, metallo-, serine- and threonine-proteases (Beynon and Bond, 2001). Proteases can also be divided into two larger types depending on if they require energy in the form of ATP to perform their function. The best characterized of the energy-independent proteases include Deg and OmpT, whereas the main group

of energy-dependent proteases is the AAA+ (ATPases Associated with diverse cellular Activities) proteases (Wickner et al., 1999; Sauer et al., 2004).

1.1. AAA

AAA+ proteases are a diverse group of ATP-dependent proteases that includes the 20S and 26S proteasomes, FtsH, Lon, HslUV and Clp proteases (Neuwald et al., 1999). They are key components of the major protein surveillance systems in all cells and are important in the regulation of several major cellular events. AAA+ proteases often function in the essential process of cell maintenance or "housekeeping", such as the 26S proteasome in eukaryotes or Clp proteases in oxygenic photosynthetic organisms. They can also be stress inducible, such as the Lon and Clp proteases in *Escherichia coli* (*E. coli*), providing the extra proteolytic activity needed to deal with the accumulation of irreversibly damaged proteins (Sauer and Baker, 2006, 2011).

AAA+ proteases consist of two distinct parts: an ATPase belonging to the AAA+ super-family and a proteolytic core. The two parts can either be separate domains within the same polypeptide as for FtsH and Lon, or they can be divided into two or more different subunits as for HslUV, the 20S and 26S proteasomes, and the Clp proteases (Figure 1). In either case, the ATPase component has at least one AAA+ domain that contains the Walker A and B domains where nucleotide binding and hydrolysis occurs (Neuwald et al., 1999). The ATPase components are responsible for substrate recognition. They typically form hexameric ring structures with a central axial pore, in which the bound protein substrates are unfolded and then translocated into the proteolytic core. The proteolytic core has a barrel-like structure formed by rings of six (HslV, Lon, FtsH) or seven subunits (ClpP, 20S proteasome, 26S proteasome), where the active sites are enclosed inside the barrel. The type of active sites and thereby the mechanism of degradation differs between the various AAA+ proteases. The cylindrical shape of the proteolytic chamber has very narrow entrances through which only unfolded proteins can pass, which is why the substrate requires unfolding by the ATPase component before translocation (Gottesman 2003; Baker and Sauer 2006) .

The different AAA+ proteases vary in their substrate specificity but how are these targeted proteins recognized? For certain AAA+ proteases, substrates are tagged with additional peptide sequences at the C-terminus, such as SsrA, or protein at the N-terminus as in the case of ubiquitin (Ub). Structural changes to the protein substrate such as partial unfolding/misfolding can also act as recognition signals as well as modifications like oxidation. Once the substrate is identified and bound, each AAA+ protease has the same basic mechanism for unfolding the protein and threading it into the proteolytic core. Powered by ATP binding and hydrolysis, the ATPase component undergoes conformational changes that lead to the unfolding of the protein substrate. The unfolded polypeptide is then translocated through the central axial pore into the proteolytic core, where it is progressively degraded at several sites to produce small peptide products. How these peptides fragments are released is not clear, although it

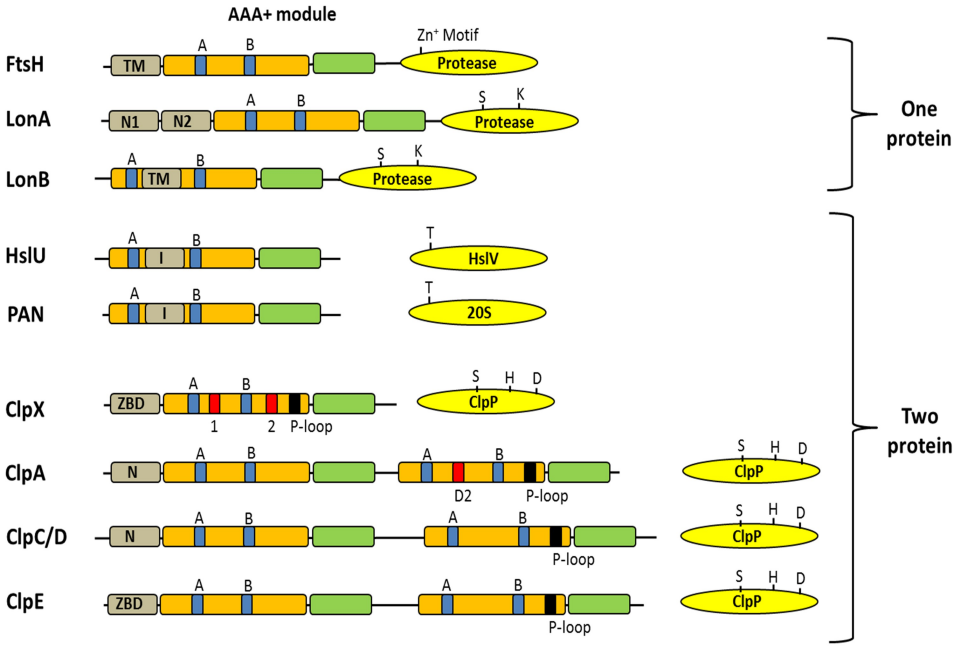

Figure 1. Pictorial representation of the different AAA+ proteases. The AAA+ proteases can be divided into two groups; one-component (FtsH and Lon) and two-component proteases (HslUV, 26S proteasome, 20S proteasome and Clp). Shown are the large (orange box) and small (green box) AAA+ domains within each protein, as well as the unique family domain (gray box). The protease part/protein is depicted in yellow, with the catalytic amino acids indicated. The blue strips indicate the Walker A and Walker B domains, while the black strip shows the location of the P-loop. The red strips are important regions involved in substrate and ATPase association (adapted from Sauer and Baker, 2011; Gur et al., 2013).

might occur by diffusion through the axial pores or via gaps between the flexible rings in the proteolytic core (Sauer and Baker, 2006, 2011).

1.1 1. 26S proteasome

1.1.1.1 Structure

The 26S proteasome is the best studied of all the AAA+ proteases. It is located in the cytoplasm and nucleus of all eukaryotes, and carries out an essential housekeeping role in both compartments (Voges et al., 1999; Tanaka, 2009). Many different protein substrates have now been identified for the 26S proteasome, with most being recognized via ubiquitination. The 26S proteasome shares the same basic architecture as the other AAA+ proteases, with a hexameric ATPase component and a large cylindrical proteolytic core, but its overall structure is by far the most complex (Peters et al., 1993). The ATPase component is termed 19S and consists of two associated sub-complexes, the lid and base. The base contains ten subunits, six distinct ATPase subunits

that form a hetero-hexameric ATPase ring, and four other peripherally-bound subunits that include an Ub receptor. The lid is also composed of ten different subunits that assemble into linear structures that overlay the base and include one deubiquitination protein and a second Ub receptor subunit. It is the 19S complex that recognizes and binds ubiquitinated proteins, and then sequentially removes the Ub tag, unfolds the protein substrate and translocates it into the proteolytic core complex (Glickman et al., 1998; Marques et al., 2009; Tomko et al., 2010).

The proteolytic core is known as 20S and consists of 28 different subunits arranged in the barrel-like shape characteristic of AAA+ proteases. The overall structure is formed by four heptameric rings stacked on top of each other. The two outer rings consist of proteolytically inactive α-subunits while the two inner rings are composed by proteolytically active β-subunits (Groll et al 1997, Baumeister et al 1998; Unno et al., 2002). Access to the inner β-rings is restricted by the N-termini of the α-subunits in the adjacent α-rings. The N-termini extend into the central pore of the α-ring and form an interfering network that blocks protein translocation (Groll et al., 2000). Entry of substrates occurs once the 19S complex associates to the outer α-ring, which causes conformation changes to the α-subunits that allows passage of the substrate from the 19S complex into the β-rings for degradation (Smith et al., 2007; Kim et al., 2011). Each β-ring consists of seven distinct subunits named for their position within the ring (i.e., β1-7). Of these seven subunits, only three are proteolytically active (β1, 2 and 5) and each has distinct endopeptidase activity - β1 cleavages after acidic residues (peptidylglutamyl-peptide hydrolytic), β2 after basic residues (trypsin-like) and β5 after hydrophobic residues (chymotrypsin-like) (Myung et al., 2001; Gallastegui and Groll, 2010).

1.1.1.2 Ubiquitin-mediated degradation pathway

Most substrates of the 26S proteasome are targeted through addition of Ub, which involves three different enzymes in a specific pathway. The first enzyme (E1) is the Ub–activating enzyme that as its name implies activates Ub via the ATP-dependent formation of a bond between the active-site cysteine on E1 and the C-terminus of Ub. This is then followed by the transfer of the activated Ub on E1 to the active-site cysteine on the second enzyme (E2), an Ub-conjugating enzyme. Proteins destined for degradation by the 26S proteasome are first recognized by the third enzyme (E3, Ub protein ligase) that then transfers the Ub from E2 to a lysine on the protein substrate. Several Ub can be added to the same protein substrate through the action of E3, either to build a poly-Ub chain or addition of Ub to several different lysine residues (Myung et al., 2001). Addition of at least four Ub in a chain appears necessary for the substrate to be recognized by the 26S proteasome (Figure 2). However, for a protein substrate to be degraded by the 26S proteasome it also needs an unstructured region. Control of the Ub system and its broad substrate specificity occurs through the regulation of E3, of which there are numerous in most eukaryotes; there are more than 600 different E3 enzymes in humans and over 1000 in the model plant species Arabidopsis thaliana

(Mazzucotelli et al., 2006). Protein degradation by the 26S proteasome begins when the Ub chain on the substrate binds to one of the Ub receptors in the 19S complex. Once the unstructured recognition sites are situated close to the ATPase ring pore, translocation starts and the poly-Ub chain is removed. As the protein substrate passes through the base, it unfolds and is then threaded into the proteolytic core for proteolysis. The overall importance of the Ub-mediated degradation pathway was recognized in 2004 with the awarding of the Nobel Prize to those who discovered and characterized much of this vital process (Myung et al., 2001).

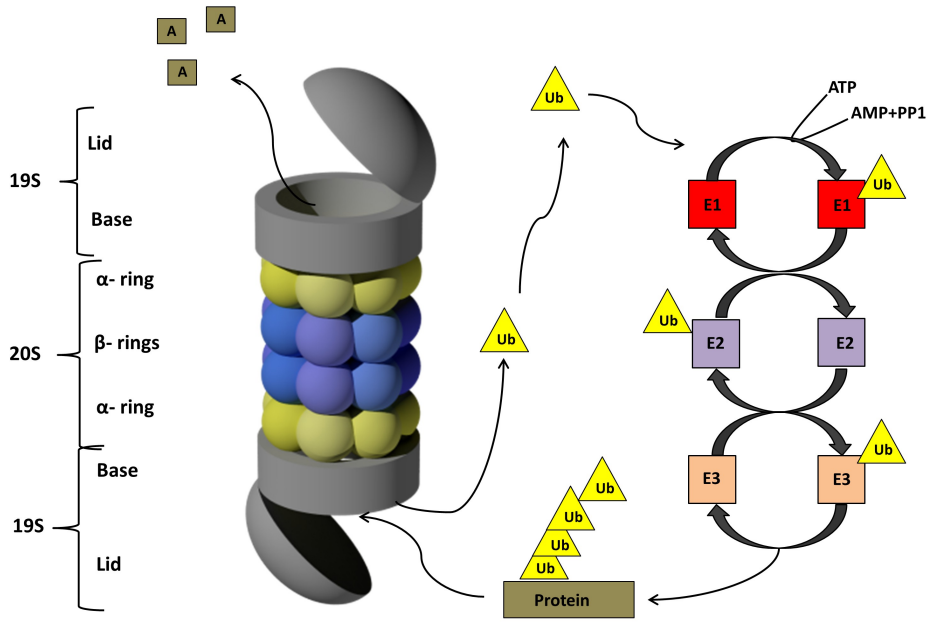

Figure 2. Ubiquitin-mediated degradation pathway. Protein substrates are targeted for degradation with a polyubiquitin chain placed by the enzymatic cascade of E1-E2-E3. A ubiquitin-activating enzyme (E1) binds to ubiquitin (Ub) in an ATP-dependent reaction and then transfers the activated Ub to a Ub-conjugating enzyme (E2). E2 then interacts with the ubiquitin ligase (E3) that transfers the Ub to the protein substrate. The 26S proteasome recognizes the Ub-chain and degrades the protein, with the Ub recycled for tagging additional substrates (adapted from Myung et al., 2001).

1.1.1.3 Substrate recognition by the N-end rule

The main identifying feature within proteins recognized by the ubiquitin system is based upon the N-end rule. The N-end rule refers to the type of amino acids at the N-terminus of a protein and how these affect its stability, with so-called “destabilizing” residues reducing the half-life of a protein *in vivo* (Varshavsky et al., 1996). Different amino acids

are destabilizing in different organisms. In eukaryotic cells, it is phenylalanine (Phe), leucine (Leu), isoleucine (Ile), tryptophan (Trp), lysine (Lys), arginine (Arg) and histidine (His) that are ubiquitinated by E3 (Varshavsky et al., 1996, 2003). There are two classes of regions in E3 that recognize N-end rule substrates: the UBR box class (Lys, Arg and His) or the ClpS-like class (hydrophobic side chains). This differs somewhat in Gram-negative bacteria in that Leu, Phe, tyrosine (Tyr), Trp, Lys and Arg that are main destabilizing amino acids (Tobias et al., 1991; Shrader et al., 1993). These bacterial N-end rule substrates are recognized by the adaptor ClpS that delivers them to the ClpAP protease for degradation (discussed later) (Erbse et al., 2006). There are three levels at which N-end rule substrates can be recognized. Primary and secondary destabilizing amino acids are found in both prokaryotes and eukaryotes, while tertiary destabilizing residues have only so far been found in eukaryotes. Primary destabilizing amino acids are directly identified by the ClpS or the E3 ligase, while the secondary and tertiary destabilization amino acids require modification to be recognized. This modification is done be specific enzymes, like the amino acyltransferase (Aat) in *E. coli* that positions the primary amino acid Leu and Phe to the secondary destabilization amino acids (Shrader et al., 1993).

1.1.2. PAN/20S proteasome

Apart from eukaryotes, archaea also possess a proteasome but one in which the 20S proteolytic core associates to an ATPase component known as PAN (Proteasome-activating-nucleotidase) (Zwickl et al., 1998). The chaperone part of PAN is formed by six identical multi-domain subunits, while the proteolytic 20S component is a threonine-type protease composed of two different subunits, α and β (Löwe et al., 1995; Rubin et al., 1995; Zwickl et al., 1998; Smith et al., 2005). Like its eukaryotic counterpart, the archaeal 20S component consists of four heptameric rings arranged as $\alpha_7\beta_7\beta_7\alpha_7$ but differs in that there is only one type of α- and β-subunit. Since all the β-subunits are identical, all are therefore proteolytically active (Löwe et al., 1995; Rubin et al., 1995). The mechanism by which protein substrates are targeted for degradation by the PAN/20S proteasome remains unclear but it does appear to involve the addition of SAMPs (small archaeal modifier proteins)(Maupin-Furlow et al., 2006; Humbard et al., 2010).

1.1.3. FtsH

FtsH (Filamentous temperature sensitive H) proteases are found in all eubacteria and the mitochondria and plastids of eukaryotes, but not in the archaea. In bacteria like *E. coli*, it is the only protease essential for cell viability, as well as the only AAA+ protease that is anchored to the membrane through two transmembrane domains (Begg et al., 1992; Akiyama et al., 1995; Jayasekera et al., 2000). Within the soluble region, the FtsH protein has both an ATPase domain in the N-terminal half and a proteolytic domain in the C-terminal half (Figure 1) (Tomoyasu et al., 1993). Crystallography and EM studies of

FtsH from different organisms has shown that the protein forms a single oligomeric structure resembling two stacked heptameric rings, one formed by the AAA+ domains and the second by the protease domains (Bieniossek et al., 2006; Suno et al., 2006; Beniossel et al., 2009; Lee et al., 2011). The N-terminal region of FtsH is also important for its oligomerization, while the transmembrane region is needed for the degradation of membrane proteins (Makino et al., 1999).

FtsH is a metalloprotease, with a conserved Zn (II) binding motif that makes the protease dependent on Zn^{2+} (Tomoyasu et al., 1995). The FtsH protease can extract integral protein substrates within the lipid bilayer and degrade them in an ATP-dependent manner. It degrades membrane proteins that are misfolded or otherwise damaged, and subunits of large multimeric complexes that have misassembled. Several membrane proteins have been identified as FtsH substrates, such as the F_0a subunit of ATP synthase (Akiyama et al., 1996a, 1996b). There are also cytosolic substrates for FtsH including σ32 (a heat shock sigma factor) and LpxC (a metallo-deacetylase involved in endotoxin biosynthesis) (Herman et al., 1995; Tomoyasu et al., 1995; Kanemori et al., 1997; Langklotz et al., 2011) . Despite this, the exact substrate specificity of the FtsH protease remains unknown, although it does appear to preferentially cut at amino acids with positively charged or hydrophobic side groups. FtsH can also degrade mistranslated polypeptides that are tagged with the C-terminal SsrA motif, as well as proteins with free unstructured N- and C-terminal ends around 10-20 amino acids in length (Herman et al., 1998; Chiba et al., 2000, 2002). Compared with the Lon and Clp proteases, FtsH has a relatively low unfolding activity and thus preferentially degrades proteins with low thermo-stability; a preference that has been proposed to influence the protein substrate selectivity for the FtsH protease (Herman et al., 2003).

Human, yeast and plant mitochondria have at least two FtsH proteases anchored to the inner mitochondrial membrane. They are named according to which soluble compartment the catalytic domains are in contact with; i-AAA (intermembrane space) and m-AAA (matrix) (Leonhard et al., 1996). The m-AAA protease has two distinct subunits, paraplegin and AFG3L2 (ATPase), that form hexamers either with AFG3L2 only or with both AFG3L2 and paraplegia (Atorina et al., 2003; Koppen et al., 2007). Two human diseases are connected to the m-AAA protease, hereditary spastic paraplegia (mutations in paraplegin) and hereditary spinocerebellar ataxia (mutation in AFG3L2) (Casari et al., 1998; Atorina et al., 2003).

The number and complexity of FtsH proteases increases dramatically in oxygenic photosynthetic organisms, for example the cyanobacterium *Synechocystis sp.* PCC 6803 (*Synechocystis*) has four and *Arabidopsis* 17 FtsH paralogs, respectively (García-Lorenzo et al., 2006). Of the latter, five appear to be inactive due to the absence of the Zn motif (Sokolenko et al., 2010). All five inactive FtsH proteins plus eight active ones are localized in plastids, whereas three are exclusively located in mitochondria (Ferro et al., 2010; Janska et al., 2005). The remaining active paralog (FtsH11) appears to be dual localized in both mitochondria and plastids (Urantowka et al., 2005). The different plastid FtsH proteins can form either homo- or hetero-oligomeric complexes, attached to either the thylakoid or envelope membranes (Yu et al., 2004, 2005; Zaltsman et al.,

2005). Few protein substrates for the plastid FtsH protease have so far been identified but one that has is the photosystem II reaction center protein D1, a crucial component in the photosynthetic electron transport chain (Lindahl et al., 2000; Bailey et al., 2002; Kato, 2009).

1.1.4. Lon

The Lon protease is a serine-type protease with a catalytic dyad of Ser and Lys. It exists in almost all bacteria and eukaryotes (Amerik, et al 1991; Rotanova et al., 2004; Tsilibaris et al., 2006). Based on structures, the Lon proteases can be divided into two groups, LonA and LonB. Both LonA and B have the ATPase (located centrally within the protein) and protease domains (C-terminal location, Figure 1), however LonA also has an N-terminal domain while LonB is often membrane anchored. Most eukaryotes have both LonA and LonB, and while certain bacteria can also possess both Lon types (e.g., *Bacillus subtillis*) LonA is more common in eubacteria and LonB in Archaea (Rotanova et al., 2004). Depending on the organism, the Lon proteases exist as either a single hexameric (bacteria) or heptameric (yeast) ring (Ståhlberg et al., 1999; Park et al., 2006). Its expression patterns can also differ between organisms, being heat stress inducible in bacteria and yeast mitochondria but down-regulated during heat stress in plant mitochondria (Riga et al., 2009).

Several protein substrates have been identified for the Lon protease, with one of the first being SulA, a regulatory protein involved in bacterial cell division. Like FtsH, Lon in *E. coli* also degrades mistranslated proteins tagged with the C-terminal SsrA sequence (Tsilibaris et al., 2006). Despite this, not much is known about how Lon recognizes its protein substrates, although it does appear to recognize exposed regions rich in hydrophobic, aromatic amino acids that are usually buried within the native structure. It is also thought that the addition of poly-phosphates to a substrate might target it for degradation by Lon (Gur et al., 2008; Venkatesh et al 2012). Lon has been shown to bind DNA, suggesting it might directly regulate the expression of certain genes (Chung et al., 1987; Fu et al., 1997). Indeed, Lon has been shown to degrade the β-subunit of HU, a nucleoid-binding protein that alters DNA structures and thereby controls which promoters are exposed for transcription (Liao et al., 2009).

1.1.5. HslUV

While most other AAA+ proteases are found in all kingdoms of life, the HslUV protease has so far only been identified in eubacteria although some genomic evidence suggests it might be in archaea as well (Couvreur et al., 2002) The HslUV protease in *E. coli* is involved in resistance to different stresses and both the HslU and HslV subunits are induced during heat stress (Change et al., 1993). HslUV was the first complete AAA+ protease to be crystallized and its structure resolved in detail (Bochtler et al., 2000; Sousa et al., 2000). The HslUV protease consists of a central proteolytic core comprised of two heptameric rings of HslV (ClpQ) flanked at either end by a hexameric ring of the

HslU (ClpY) ATPase components (Bochtler et al., 2000; Sousa et al., 2000; Song et al., 2003). HslV is a threonine-type protease that requires the HslU component to recognize and bind the protein substrates, then unfold and translocate them into the HslV core complex for degradation (Change et al., 1993; Huang and Goldberg, 1997; Kwon et al 2003). Little is known about the actual degradation mechanism of the HslUV protease but it does require both ATP and Mg^{2+} to bind the targeted substrates (Burton et al., 2005) . Natural substrates for HslUV in *E. coli* have been identified and several are shared with other AAA+ proteases, such as the cell division inhibitor SulA that is also degraded by the Lon protease, and the heat shock sigma factor σ^{32} degraded by the FtsH protease (Kanemori et al., 1997; Cordell et al., 2003). Another substrate is the Arc repressor (Burton et al., 2005), a DNA-binding protein that inhibits bacteriophage P22. The Arc repressor is now often used as the model substrate for the HslUV protease during *in vitro* studies, which have revealed that a degradation tag in the N-terminal region of the substrate (Burton et al., 2005).

1.1.6. Clp

The Clp proteases are found in most domains of life, from bacteria to human, as well as in parasites and plants. Clp are serine-type proteases where the catalytic triad consists of active site Ser, His and Asp residues, with all three amino acids being essential for catalytic activity (Maurizi et al., 1990a). The proteolytic core usually consists of a single type of catalytically active subunit (ClpP) but the type and activity of the subunits can vary considerably depending on the organism. As the other AAA+ proteases, Clp has an ATPase part in the form of a hexameric ring and a proteolytic core consisting of twin heptameric rings (Kessler et al., 1995; Wang et al., 1997; Gottesman et al., 1997). The ATPase components of Clp proteases are now recognized as members of the HSP100 family of molecular chaperones and they can be divided into two major groups based on the number of AAA+ domains that they contain. The first group of Clp ATPases contains two AAA+ domains and can be further divided into ClpA-E and ClpL based on conserved amino acid sequences and the length of sequence separating the AAA+ domains. Members of the second group differ from the first by having only one AAA+ domain and include ClpX and ClpY (HslU) (Schirmer et al., 1996, Figure 1). Apart from the AAA+ domains, other types can also be found in various members of the HSP100 family. For example, both ClpX and ClpE have Zn-finger motifs in the N-terminal region that are involved in DNA binding (Donaldsson et al., 2003). All the Clp ATPases except ClpB and ClpL also contain the so-called P-loop (IGF/L-motif) that is necessary for association to the Clp proteolytic core (Kim et al., 2001; Singh et al., 2001), and therefore they have the potential of operating as a chaperone both independently and

	H. sapiens	*E. coli*	*B. subtillis*	*S. aureus*	*M. tuberculosis*	*S. elongatus*	*P. falciparum*	*A. thaliana*
	ClpP	ClpP	ClpP	ClpP	ClpP1	ClpP1	ClpP	ClpP1
					ClpP2	ClpP2	ClpR	ClpP2
						ClpP3		ClpP3
						ClpR		ClpP4
Proteolytic								ClpP5
								ClpP6
								ClpR1
								ClpR2
								ClpR3
								ClpR4
	ClpX	ClpA	ClpC	ClpC	ClpC1	ClpC	ClpC	ClpC1
ATPases		ClpX	ClpX	ClpX	ClpC2	ClpX		ClpC2
with			ClpE		ClpX			ClpD
P-loop								ClpX1
								ClpX2
								ClpX3
ATPases		ClpB		ClpL	ClpB	ClpB1	ClpM	ClpB1
without				ClpB		ClpB2	ClpB1	ClpB2
P-loop							ClpB2	ClpB3
		ClpS			ClpS	ClpS1	ClpS	ClpS
Adaptors						ClpS2		ClpT1
								ClpT2

Table 1. The diversity and function of Clp proteins in different organisms. The Clp protein composition in *Homo sapiens*, *Escherichia coli*, *Bacillus subtilis*, *Streptococcus aureus*, *Mycobacterium tuberculosis*, *Synechococcus elongatus*, *Plasmodium falciparum* and *Arabidopsis thaliana* is shown. The far left column indicates the different functional groups of the Clp proteins. The different colored text indicates the location of the protein: cytosol (black), mitochondrial (blue), chloroplastic (green) and apicoplastic (red).

as the ATPase component of Clp proteases. The role for these ATPases within the Clp protease is similar to that in other AAA+ proteases, i.e., to recognize and bind protein substrates, and then translocate the unfolded protein into the proteolytic core for degradation.

The complexity of Clp proteases in terms of composition and types differs tremendously between different organisms (Table 1). Among the eubacteria, the Gram-negative species typically have ClpA and ClpX ATPases and a single ClpP, whereas Gram-positive bacteria possess ClpC, ClpE and ClpX along with one-to-five ClpP paralogs (Ingmer et al., 1999; Frees et al., 2007). The diversity of Clp proteolytic core subunits increases further in oxygenic photosynthetic organisms, with cyanobacteria usually containing three ClpP paralogs and vascular plants having up to six, along with one or more of a unique variant termed ClpR (Clarke et al., 1999). The functional importance of the Clp protease also varies significantly from organism to organism. In *E. coli*, for example, loss of Clp proteolytic activity has no obvious effect on cell viability or exponential growth, but does affect certain growth transitions and stress responses (Chuang et al., 1993; Dougan et al., 2002; Thomsen et al., 2002; Erbese et al., 2006). In contrast, the Clp proteases in cyanobacteria and plants are essential for normal growth and appear to have little or no role during stresses (Schelin et al., 2002; Zheng et al., 2002, Peltier et al., 2004). Clp proteases are also important for virulence in several

different organisms, including pathogenic Gram-positive bacteria and certain protozoan parasites (Mei et al., 1997; Frees et al., 2003; Raju et al., 2012, 2014). In general, Clp proteases degrade a wide range of enzymes and regulatory proteins within the different organisms and as such influence many different cellular pathways.

1.2. Clp Proteases in Different Organisms

1.2.1. *E. coli*

Of all the different Clp proteases, the one in *E. coli* has been the most extensively studied and therefore most of what we know today about the mechanism of Clp proteases comes from this model system. The *E. coli* Clp protease consists of four Clp proteins: ClpA, ClpX, ClpP and ClpS (Katayama et al., 1988; Hwang et al., 1988; Gottesman et al., 1993; Wojtkowiak et al., 1993; Dougan et al., 2002). The *clpX* gene is in an operon with *clpP* and both are co-expressed constitutively (under the control of σ70) and during heat stress (σ32) (Maurizi et al., 1990; Gottesman et al., 1993). The *clpA* and *clpS* genes are situated in a second operon and expressed constitutively under the control of σ70 (Dougan et al., 2002). The overall amount of these Clp proteins is relatively low during normal growth but they can rise during stresses such as high temperatures (Chuang et al., 1993). Mutational studies have shown that the different Clp proteins in *E. coli* are not essential for normal growth, but they are crucial for stress survival and certain growth transitions (Dougan et al., 2002; Thomsen et al., 2002; Erbese et al., 2006).

ClpP in *E. coli* is synthesized as a precursor with a 14 amino acid extension at the N-terminus that is later autolytically processed to generate the mature protein of 193 amino acids (Maruizi et al., 1990). The Clp proteolytic core consists of a barrel-shaped tetradecamer characteristic of AAA+ proteases, in which the two heptameric rings of ClpP subunits are stacked on top of each other (Kessler et al., 1995; Wang et al., 1997). The two heptameric rings associate to each other via the handle region of opposing ClpP subunits (Wang et al., 1998), while the subunits within each ring bind through hydrogen bonding between certain amino acids (Bewley et al., 2006). The entrance pore into the degradative chamber is very narrow and restricts entry to all proteins apart from short, unfolded peptides (Thompson and Maurizi, 1994; Wang et al., 1997) It is only when the ClpA or ClpX chaperones associate to the ClpP core complex that longer unfolded polypeptides can be translocated inside for degradation (Gottesman et al., 1997; Joshi et al., 2004; Kim et al., 2008; Kolygo et al., 2009). Not only do ClpA and ClpX confer substrate specificity for the Clp protease but this specificity varies between the two types of ATPases (Gottesman et al., 1993; Flynn et al., 2003; Mogk et al., 2004). It appears that most of the Clp protease in *E. coli* consists of a hexameric ring of ClpA or ClpX at either end of the proteolytic core, although only a single protein substrate is translocated inside the core complex at any given time. It is also possible that a ClpA hexamer can bind to one end of the proteolytic core with a ClpX hexamer bound to the other (Grimaud et al., 1998).

1.2.1.1. Mechanism

Structural studies on the ClpA hexamer has shown that the two AAA+ domains form two stacked rings, with the ring formed by the second AAA+ domain closest to the proteolytic core within the Clp protease (Kessler et al., 1995; Guo et al 2002; Hinnerwisch et al. 2005). With only one AAA+ domain, ClpX forms a single hexameric ring structure but one in which there are two distinct subunits. The first is termed "loadable" (L) where the small and large part of the AAA+ domain are oriented in such a way that a clef is formed in which the nucleotide can bind. The other is the unloadable (U) subunit where the clef site is destroyed by a rotation in the hinge region. Within the known atomic structure of *E. coli* ClpX, these two forms of subunits are arranged in the following configuration L/U/L/L/U/L (Stinson et al., 2013).

Protein substrates specific for either ATPase component are bound to the N-terminal region of ClpA/X (Singh et al., 2001; Wojtyra et al., 2003). Substrates are then pulled into the central cavity of the hexamer and are unfolded through conversion of the energy from ATP hydrolysis to mechanical motion (Weber-Ban et al., 1999; Reid et al., 2001). When nucleotide binds to ClpX it leads to a stepwise alteration of the neighboring subunit, eventually causing the loadable subunit to be converted to an unloadable one. It is this conversion of subunits stimulated by ATP hydrolysis that results in the mechanical force that unfolds the protein substrate (Stinson et al 2013). The mechanical pulling is linked to conformation changes in ClpX close to the pore-1 loop, a region that lines the central cavity of the hexamer (Martin et al., 2008; Glynn et al., 2009; Wang et al., 2001). A single pulling, or so called "power stroke" can fail several times *in vitro* to unfold a region within the protein substrate, but it is not until a power stroke coincides with destabilization of that region that the unfolding process of the substrate can continue. This would mean that in theory the complete unfolding of a stable protein substrate would require hydrolysis of only one ATP molecule per power stroke, but the high cost of failure could increase this cost dramatically to several hundred ATP molecules. However, it remains unclear if the rate of power-stroke failure *in vivo* is as high as that shown *in vitro* (Martin et al., 2005). In contrast, it is the D2 loop in ClpA, situated in the axial channel of the ClpA hexamer that is important for the substrate unfolding and translocation into the ClpP (Hinnerwisch et al., 2005; Bohon et al. 2008; Farbman et al. 2008).

Association between the ClpA/X hexamers and the ClpP proteolytic core occurs at more than one region. The first involves the P-loop motif in the C-terminal region of ClpA and ClpX that extends down and binds to a hydrophobic clef in the surface of the heptameric ring of ClpP (Figure 3). This association probably leads to conformation changes that open up the narrow entrance in ClpP to enhance passage of unfolded substrates inside (Kim et al., 2001; Singh et al., 2001; Joshi et al., 2004). A second interaction occurs between the N-terminal region of the ClpP subunits and the pore-2-

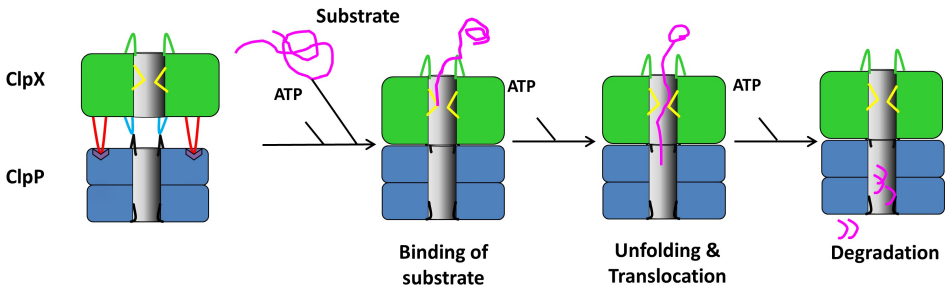

Figure 3. Mechanism of protein degradation by ClpXP. Shown is a schematic view of the regions important for association between ClpX and ClpP. The essential P-loop in ClpX (red) associates to the hydrophobic clef in ClpP (purple arrows). The N-terminus of ClpP (black loops) interacts with the 2-pore loop from ClpX (light blue). The protein substrate is recognized by ClpX, where it first associates with the N-terminus (pink) and then the substrate is pulled down by internal loops in ClpX (yellow) (adapted from Gur et al., 2013).

loop in ClpX (Gribun et al., 2005; Bewley et al., 2006; Martin et al., 2007, 2008; Jennings et al., 2008; Figure 3). Structural studies have also shown that the N-terminal region of ClpP is highly flexible and can form different conformations called “up” and “down”. In the “up” conformation, part of the N-terminus protrudes out from the access pore while in the “down” conformation most of the N-terminus resides within the access pore. It has been suggested that these N-terminal structures could also provide a symmetrical match between the hexameric ClpA/X and heptameric ClpP rings if six of the seven ClpP N-termini have the same conformation simultaneously (Bewley et al., 2006). It is also thought that the N-terminal region of ClpP, presumably in the down configuration closes the entrance channel and stabilizes the acyl-enzyme intermediate during proteolysis (Jennings et al., 2008). Later it was suggested that charged amino acids in the N-terminal region of ClpP that line the channel are involved in determining the maximal rate of degradation (Lee et al., 2010). The degradative efficiency of the ClpXP/ClpAP proteases is ensured by the high concentration of active sites inside the barrel chamber and that the unfolded substrate can bind to more than one active site simultaneously and be cleaved at multiple sites. How resulting peptide fragments are released from the proteolytic core remains unknown but they are considered to freely diffuse out via the axial entrance pores or through side gaps between the two rings (Kang et al., 2005). The released peptide fragments are then degraded by exopeptidases to single amino acids.

1.2.1.2. ClpA

Of the two ATPase components, ClpA has a higher affinity for the ClpP proteolytic core than ClpX, and that during normal growth there are more ClpAP proteases than ClpXP (Grimaud et al., 1998). To date, the only well-defined substrate for the ClpAP protease is RepA, a P1 plasmid initiator protein (Wickner et al., 1994; Pat et al., 1997). Most of the

RepA protein in *E. coli* exists as inactive dimers, but they are converted to active monomers by ClpA in an ATP-dependent manner, enabling the active RepA to associate to *ori*P1 DNA (Wickner et al., 1994). ClpA can also deliver RepA to the ClpP proteolytic core for degradation (Sharman et al., 2005). Proteins with the C-terminal SsrA-tag are also degraded by ClpAP *in vitro*, although their degradation *in vivo* appears to be done primarily by ClpXP (Gottesman et al., 1998; Farell et al., 2005). The ClpA protein itself is autoregulated, with any excess ClpA protein relative to that of ClpP being degraded by the ClpAP protease (Gottesman et al., 1990).

1.2.1.3. ClpS adaptor

ClpS is a small protein (12 kDa) that when bound changes the substrate specificity of ClpA to N-end rule substrates, while simultaneously blocking substrates normally recognized by ClpA alone such as SsrA-tagged proteins (Dougan et al., 2002; Erbse et al., 2006). ClpS has a cone-shaped structure comprised of two parts, an N-terminal region that extends out from the core region (Zeth et al., 2002; Roman-Hernandez et al., 2011). In the core structure, there are two conserved domains, one of which is involved in the interaction to ClpA and the other a hydrophobic pocket that binds *via* hydrogen bonding to the primary destabilizing amino acid of N-end rule substrates (Guo et al., 2002; Zeth et al., 2002; Erbse et al., 2006; Wang et al., 2008a; Scuenemann et al., 2009). The hydrophobic pocket in ClpS is small but it can accommodate the side-chains of primary destabilizing amino acids Leu, Phe, Tyr and possibly Trp (Wang et al., 2008a; Roman-Hernandez et al., 2009; Schuenemann et al., 2009).

The N-terminal region of the ClpS adaptor is necessary for delivery of the N-end rule substrates to ClpA, but it is not needed for the actual substrate binding (Hou et al., 2008; Roman-Hernandez et al., 2011). This was clearly shown using a truncated version of ClpS lacking the N-terminal region, which was still capable of associating to the substrate but not initiating its degradation. It was also shown that this truncated version could still inhibit the degradation of SsrA-tag substrates by the ClpAP protease. It appears that it is not the actual amino acid sequence of the N-terminal region in ClpS that is important but its length, suggesting it is the peptide backbone of the amino acids that are important for the role of the N-terminal region (Hou et al. 2008). The N-terminal region and the junction between this and the core structure enhance, but are not essential for the association to ClpA (De Donatis et al., 2010; Roman–Hernandes et al., 2011). One model suggests that ClpS first binds to the N-end rule substrate and then associates to the ClpA hexamer *via* the flexible N-domain of ClpA and the core structure of ClpS. The N-terminal domain of ClpS then also binds to ClpA, probably near the access pore so that the N-end rule substrate is in close proximity. This is followed by ClpA pulling in the N-terminal domain of ClpS, thereby causing a conformational change to the ClpS core that releases the N-end rule substrate. The substrate is then transported into the ClpA pore and protein unfolding begins, while ClpS is released. It has been implied that this association between the N-terminal region of ClpS and ClpA

ensures that only one substrate is delivered for eventual degradation at any given time (Figure 4; Roman-Hernandes et al., 2011).

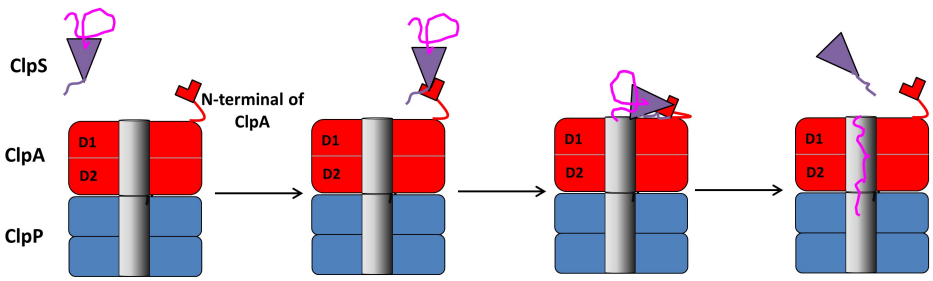

Figure 4. Substrate delivery by ClpS to the ClpAP protease. Shown is a schematic view of the suggested model for substrate degradation by ClpAPS. ClpS recognizes and binds the N-end rule substrate (pink), followed by the association to ClpA N-terminus via the region between the N-terminus and the core domain of ClpS. Next, the N-terminus of ClpS binds to an unidentified site near the pore entry, which positions the substrate at the entry. ClpA then finally pulls on the substrate, which probably triggers conformational changes in ClpS that releases the substrate (adapted from Roman-Hernandes et al., 2011).

In regards to the ClpS-ClpA interaction, there remains debate over exactly how many ClpS proteins need to associate to a ClpA hexamer in order to maximize degradation of targeted substrates. It is clear that a single ClpS monomer is sufficient to change the substrate specificity of a ClpA hexamer and to inhibit the degradation of SsrA-tagged proteins (Hou et al., 2008; De Donatis et al., 2010). There is however conflicting evidence on the extent of this stimulation and inhibition. One study showed that a single ClpS monomer associated to ClpA degrades N-end rule substrates at maximum efficiency as well as blocking degradation of other substrates. They claim that only one ClpS monomer *via* its N-terminal region is likely to bind to the ClpA hexamer with high affinity, whereas additional ClpS monomers would associate only weakly (De Donatis et al., 2010). However, another study has shown that when three or less ClpS monomers associate to the ClpA hexamer that the degradation of SsrA-tagged substrates is inhibited by only 30%. They demonstrated that ClpA had higher affinity for N-end rule substrates with only two ClpS monomers attached but that four were needed to maximize degradation of N-end rule substrates while also completely blocking degradation of SsrA-tagged proteins (Hou et al. 2008).

In regards to the proteins recognized by ClpS in *E. coli*, there are two types of destabilizing amino acids - primary and secondary. The primary destabilizing amino acids are Leu, Phe, Tyr and Trp, and they interact directly with the ClpS adaptor (Tobias et al., 1991). In contrast, the secondary destabilizing amino acids (Lys and Arg) must first be modified before being recognized as N-end rule substrates by the ClpS. The modification is the addition of the primary destabilizing amino acid Leu or Phe to the secondary destabilizing amino acid by the enzyme amino acyltransferase (Aat) (Shrader et al., 1993). Also required for the degradation of N-end rule substrates is an unstructured region between the primary destabilizing amino acid and the folded area

of the protein (Erbese et al., 2006). This unstructured region must be at least four amino acids long and include a hydrophobic element for delivery of the substrate to ClpA (Wang et al., 2008b; Ninnis et al., 2009).

Two native substrates targeted by ClpS for the ClpAP protease were identified many years ago (Ninnis et al., 2009; Schmidt et al., 2009); the Dps protein that helps protect DNA during several different stress conditions including starvation and oxidation (Almiron et al., 1992; Lomovskaya et al., 1994; Altuvia et al., 1994; Martinez et al., 1997), as well as the putrecine aminotransferase (PATase) (Ninnis et al., 2009). Studies of the Dps protein revealed the primary destabilizing Leu in position 6, although how this truncated Dps is generated is unknown (Ninnis et al., 2009; Schmidt et al., 2009). It is important to note that the full-length Dps protein is a substrate for the ClpXP *in vivo* (Flynn et al., 2003; Stephani et al., 2003). In a recent study, an additional 100 substrates for ClpS were identified, of which some has been identified earlier. Most of these new substrates (90%) contained the primary destabilizing amino acids Leu or Phe at the N-terminus. However, in the native protein this sequence was internal indicating that the proteins had undergone an earlier proteolytic event. It was also shown that this proteolytic event seemed to occur mostly in accessible regions such as the N-terminal and flexible surface exposed regions (Humbard et al., 2013).

1.2.1.4. ClpX

ClpXP is the protease *in vivo* primarily responsible for degrading SsrA-tag substrates in *E. col* (Farell et al., 2005; Lies et al., 2008). The SsrA-tag is added to the C-terminus of nascent polypeptides by the tmRNA when ribosomes stall during translation, ensuring that unfinished proteins are degraded by the ClpXP protease before they interfere with cellular functions (Keiler et al., 1996; Gottesman et al., 1998). The SsrA-tag in *E. coli* is 11 amino acids, of which only two are essential for recognition by ClpX (Kim et al., 2000; Flynn et al., 2001). Proteins marked with SsrA substrates are recognized either directly by ClpX or delivered by the adaptor SspB (Levchenko et al., 2000). ClpX recognizes the C-terminal dipeptide and the α-carboxylate within the SsrA sequence, while SspB recognizes the N-terminal sequence of the SsrA-tag and presents it to ClpX (Levchenko et al., 2000; Flynn et al., 2001). The action of the SspB adaptor enables the ClpXP protease to degrade SsrA-tag substrates even at very low concentrations. Association between SspB and ClpX occurs at several multivalent contacts, although the adaptor readily releases the bound substrate once ClpX begins to pull it at the start of the unfolding process (Bolon et al., 2004).

Another ClpX adaptor is RssB that recognizes the global stress sigma factor RpoS (σ38) and presents it as a substrate for the ClpXP protease (Zhou et al., 2001; Zhou and Gottesman, 1998). Other substrates for the ClpXP protease are the LexA and RecN proteins, both are involved in the regulation of DNA repair (Neher et al., 2003; Nagashima et al., 2006; Neher et al., 2006). ClpX can also function as a separate ATPase, disassembling for example the MuA-transposase tetramer that associates to DNA, an important step in phage replication (Mhammedi-Alaoui et al., 1994; Levchenko et al.,

1995; Burton et al., 2001; Burton and Baker, 2003, 2005). Proteomic and bioinformatics studies have revealed many potential substrates for the ClpXP protease and within them five general recognition tags could be detected. Two of them are C-terminal sequences that resemble the SsrA- and MuA-tags. Another at the N-terminus has sequence similarity to the λO-protein, a known substrate for ClpXP, while a second N-terminal sequence is associated secretion of proteins. Of all the potential substrates for the ClpXP protease, around 25% contain two recognition tags (Gonciarz-Swiatek et al., 1999; Flynn et al., 2003; Neher et al., 2006). Some of these tags associate to the axial pore of ClpX, whereas others like SsrA cross-link to the pore-1 and pore-2 loops in the axial chamber of ClpX (Martin et al., 2007, 2008b). For other substrates, the N-terminal domain of ClpX is also important for their recognition (Wojtyra et al., 2003), as is a loop in the entrance pore of ClpX containing a RKH sequence (Figure 4) (Farrell et al., 2007).

1.2.2. *Gram-positive bacteria*

Gram-positive bacteria are a diverse group that includes several human pathogens such as *Listeria monocytogenes* (*L. monocytogenes*), *Staphylococcus aureus* (*S. aureus*) and *Streptococci*. There are also many non-pathogenic species like *B. subtilis*, which decompose organic matter in the soil. Most of the Gram-positive bacteria studied so far have three Clp ATPases (ClpX, ClpC and ClpE) and one Clp proteolytic core consisting of a single type of ClpP (review by Frees et al., 2007). In many of these bacteria, the *clpP* gene is inducible by heat stress and when mutated growth at high temperatures is restricted (similar to that observed in *E. coli*) (Msadek et al., 1998; Frees and Ingmer, 1999; Gaillot et al., 2000; Lemos and Burne, 2002; Robertson et al., 2002; Frees et al., 2003; Nair et al., 2003). In contrast, the function of ClpX appears to vary significantly between the different Gram-positive bacteria and to *E. coli* (Nair et al., 1999; Chastanet et al., 2001; Miethke et al., 2006). Deletion of the *clpX* gene in *B. subtilis* leads to restricted growth at high temperatures (Gerth et al., 1998), whereas the same mutation in *S. aureus* confers increased tolerance to heat stress (Frees et al., 2003; 2004). In *L. lactis*, the ClpX is also inducible during heat stress but also at low temperatures, while *Streptococci* ClpX is essential for normal growth (Skinner et al., 2001; Robertson et al., 2003).

Of the ClpC proteins, the one in *B. subtilis* is by far the most studied and whose function relies on the adaptor MecA. *B. subtilis* ClpC can only form a hexamer when MecA is bound to the protein (Kirstein et al., 2006). MecA associates to the N-terminal region of ClpC, and following ClpC oligomerization it can present substrates to ClpC (Persuh et al., 1999; Kirstein., 2006; Mei et al., 2009). ClpCP is the protease mainly responsible in *B. subtilis* and *S. aureus* for degrading denatured or otherwise damaged polypeptides during heat stress, although the ClpEP protease can also be involved (Kock et al., 2004). ClpE in *B. subtilis* is relatively low abundant during normal growth, but it is induced several fold in the early stages of heat stress (Miethke et al., 2006). The level of ClpE is also regulated by the related ClpCP and ClpXP proteases (Derre et al. 1999; Gerth et al. 2004). Together, the various Clp proteases in *B. subtilis* are not only important for

protein degradation during stress conditions, but they contribute significantly to protein turnover during normal growth. This is highlighted by the fact that 20-30% of nascent polypeptides in *B. subtilis* aggregate when ClpP is inactivated (Kock et al., 2004).

Expression of most of Clp proteins in Gram-positive bacteria, apart from ClpX is controlled through a fine-tuning system that involves the repressor CtsR (Derre et al. 1999). CtsR is itself regulated by all three Clp proteases in *B. subtilis* depending on the growth conditions (Derre et al. 2000; Miethke et al., 2006; Kirsten et al. 2005). Other repressor proteins also degraded by the Clp protease in Gram-positive bacteria include Spx, a regulator of protein during oxidative stress (Nakano et al., 2002, 2003). Of the different Gram-positive bacteria studied so far, *Streptomyces lividans* stands out as unusual by having five *clpP* genes. These genes are organized in two operons with *clpP1clpP2* in one and *clpP3clpP4* in the other, while *clpP5* is a monocistronic gene. Little is yet known about the regulation and function of these multiple *clpP* genes, although *clpP3clpP4* appear not to be expressed during normal growth but are induced when the *clpP1* gene is inactivated (Viala et al., 2002). Also the *L. monocytogenes* has two ClpP homologues (LmClpP1 and LmClpP2) that together form a proteolytic core. The proteolytic core consists of one heptameric ring of only LmClpP1 and one with only LmClpP2. The LmClpP1 is also different because the catalytic triad is Ser, His and Asn (Zeiler et al., 2011, 2013)

1.2.2.1. Virulence

When pathogens invade another organism they must survive many different adverse environments including high temperatures, oxidative stress and antimicrobial peptides, most of which will lead to extensive protein unfolding. As mentioned above, many of the Clp proteins in Gram-positive bacteria are induced during different stresses and they have now been shown to be crucial for virulence of several pathogenic Gram-positive bacteria. One such example is *S. aureus*, a human pathogen that gives rise to life-threatening conditions such as bacteremia as well as other less harmful infections (Lowy, 1998). Clp proteins are essential for virulence in *S. aureus*. When either *clpX* or *clpP* are inactivated in *S. aureus*, the pathogen has a lower infection rate of different host cells (Mei et al., 1997; Frees et al., 2003), due possibly to the stress condition inflicted on the mutant cells during the infection. *S. aureus* mutants of *clpX*, *clpP* or *clpC* are also more sensitive to oxidative stress but they respond differently when exposed to higher temperatures. The *clpX* mutant survives better at high temperature than the wild type, whereas the *clpP* and *clpC* mutants are more sensitive (Frees et al., 2003; Chatterjee et al., 2005). However, Clp proteins are not only important for virulence indirectly by their importance in stress tolerance, they are also more directly involved. In both the *S. aureus clpX* and *clpP* mutants, expression of the major hemolysin proteins (α- and β-hemolysin) is greatly reduced (Frees et al., 2003). Several regulatory proteins of virulence factors and stress adaption are now known to be substrates of the Clp proteases in *S. aureus* (Frees et al., 2003; Michel et al., 2006; Feng et al., 2013). Clp proteases are also involved in virulence in another Gram-positive bacterium *L.*

monocytogenes, a human pathogen that can be found in food and can cause severe infections. When the *clpP* gene is mutated, *L. monocytogenes* cells have a drastically lower survival rate in macrophages, due to the mutant secreting a much lower concentration of functional listerolysin O than the wild type (Gallito et al., 2000). ClpC is also involved in the survival of the pathogen in macrophages (Rouquette et al., 1996, 1998), as is ClpE in virulence of *L. monocytogenes* more generally (Nair et al., 1999).

A major medical dilemma facing society today is the growing spread of multi-drug resistant bacteria worldwide (Otto 2012), for which the discovery of new drug targets is of vital and immediate importance. As described above, Clp proteases are involved in the virulence of several pathogenic bacteria and as such they are an interesting target for antibiotics. The acyl depsipeptides (ADEPs) were the first antibiotic shown to function directly upon the Clp protease (Brötz-Oesterhelt et al., 2005, Kirstein et al., 2009). ADEPs bind directly to the ClpP subunit, causing the Clp proteolytic core to become proteolytically active without the ATPase partner and no longer ATP dependent (Brötz-Oesterhelt et al., 2005; Kirstein et al., 2009; Li et al., 2010). This leads to uncontrolled degradation of proteins that eventually becomes cytotoxic (Sass et al., 2011). Unlike most other type of antibiotics, ADEPs are effective on non-dividing or dormant cells, and in these cells the unregulated ClpP core complex degrades over 400 different proteins (Conlon et al., 2013). Another example of new antibiotics that targets the ClpP subunit is the β-Lactones (Böttcher and Sieber, 2008). It should also be noted that Clp proteases can also help confer antibiotic resistance, as for example in vancomycin resistance (Shoji et al., 2011).

1.2.3. *Mycobacterium tuberculosis*

Tuberculosis is a major global health issue, with nearly 1.3 million people killed by the disease annually (WHO 2013). The causal agent of tuberculosis is the bacterium *Mycobacterium tuberculosis*. The Clp protease is essential for normal growth of *M. tuberculosis* as well as for its virulence (Sassetti et al., 2003; Ollinger et al,. 2011; Raju et al. 2012, 2014). *M. tuberculosis* possesses two different *clpP* genes (*clpP1* and *clpP2*) that are co-expressed in a single operon controlled by ClgR (Engles et al., 2004; Estorninho et al., 2010; Personne et al., 2013). The *clpP1* and *clpP2* genes are constitutively expressed but also stress inducible under conditions such as aerobic and hypoxic growth (Muttucumaru et al., 2004) as well as during infection of macrophages (Estorninho et al., 2010).

There are three potential ATPase partners for the ClpP1 and ClpP2 proteins: ClpC1, ClpC2 and ClpX. Earlier studies showed that ClpC1 associated to ClpP2 and not ClpP1 (Singh et al., 2006) and that ClpC1 was important for the degradation of RseA, an anti-sigma factor. In this latter study, they showed that RseA was degraded by ClpC1P2 *in vitro* and not by ClpXP1, ClpXP2 or ClpC1P1 (Barik et al., 2010). More recent work has revealed that recombinant ClpP1 and ClpP2 form a mixed proteolytic core that consists of one heptameric ring of ClpP1 and one of ClpP2 (Akopian et al., 2012; Raju et al., 2012). Each of the recombinant ClpP1 and ClpP2 proteins can form homo-

tetradecamers, but the resulting core complexes are proteolytically inactive against model peptides or casein despite the ClpP subunits having catalytic triads in the active configuration. Formation of the mixed ClpP1P2 proteolytic core and its continued proteolytic activity only occurs in the presence of a certain activator peptide (N-blocked peptide aldehydes). Degradation of casein was considerably faster when ClpC1 was added along with the activator, suggesting ClpC1 functions as the chaperone partner for the mixed proteolytic core (Akopian et al., 2012).

ClpP1/P2 is involved in recognizing SsrA-tagged substrates (Raju et al., 2012; Personne et al., 2013). However, when ClpP1 and/or ClpP2 was overexpressed *in vivo* along with the LacZ protein tagged with the SsrA-tag (AANDENYA**LAA**) or an altered tag (AANDENYA**GGG**), both ClpP1 and ClpP2 recognized the SsrA-tagged LacZ but ClpP2 could also recognize the altered tagged LacZ (Personne et al., 2013). Site-directed mutagenesis also revealed that if either the ClpP1 or ClpP2 subunits were inactivated the entire ClpP1/P2 complex had reduced degradation activity and the inhibition was greater when ClpP1 was inactive than ClpP2 (Akopian et al., 2012).

One substrate that has been identified for the ClpP1/P2 core is WhiB1, an essential transcriptional repressor of several genes. It is the ClpP1/P2 core that controls the level of the repressor *via* degradation, explaining at least in part why ClpP1/P2 is essential for *M. tuberculosis* cell viability (Raju et al., 2014). ClpP1/P2 is also important for the degradation of misfolded proteins in *M. tuberculosis* and is responsible for the degradation of SsrA-tagged proteins (Personne et al., 2013).

1.2.4. Cyanobacteria

Cyanobacteria are the only prokaryotes that perform oxygenic photosynthesis and they can be found in almost all habitats globally. Ancestors of modern-day cyanobacteria are responsible for the oxygenation of our atmosphere and are the progenitors for the plastid in algae and plants. Cyanobacterial genomes typically code for multiple Clp proteins, with the model species *Synechococcus elongatus* (*S. elongatus*) having ten Clp proteins: ClpB1-2, ClpC, ClpX, ClpP1-3, ClpR and ClpS1-2 (Figure 3)(Clarke et al., 2005).

The ClpC in cyanobacteria and other photosynthetic organisms has only low sequence similarity to the type of ClpC in Gram-positive bacteria. It also differs from ClpC in Gram-positive bacteria by not requiring adaptors for its assembly or chaperone activity (Andersson et al., 2006). A phylogenetic analysis of the three cyanobacterial ClpP proteins suggests that ClpP1 is specific to cyanobacteria, ClpP2 the ortholog to ClpP in *E. coli* and ClpP3 the ortholog to the plastid-encoded ClpP1 in algae and plants (Peltier et al., 2001). ClpR is a variant of ClpP that has only been found to date in cyanobacteria and plastids (Clarke et al., 1999). It has a similar amino acid sequence to ClpP but crucially lacks some or all of the active site amino acids within the catalytic triad. All ClpR proteins also have an insertion in the sequence that when modeled upon known ClpP structures is situated within the head domain and blocks the substrate groove. As a consequence, ClpR was presumed to be proteolytically inactive and this was later shown for the *S. elongatus* protein *in vitro* (Andersson et al., 2009).

In most cyanobacteria, the *clpP2* and *clpX* genes as well as *clpP3* and *clpR* are arranged in bicistronic operons (Kaneko et al., 1996; Clarke et al., 1998), although in *S. elongatus* the four genes can also be expressed as monocistronic messages (Schelin et al., 2002). ClpC, ClpX, ClpP3 and ClpR are all essential in *S. elongatus* for cell viability and photosynthetic growth (Clarke et al., 1996; Schelin et al., 2002). ClpP1 is inducible by low temperatures, high light and UV-B radiation (Clarke et al., 1998; Porankiewicz et al., 1998), while the amount of ClpX, ClpP2, ClpP3 and ClpR also increase at high light intensities (Schelin et al., 2002). *S. elongatus* has at least two Clp proteases, ClpCP3/R and ClpXP1/P2. The ClpCP3/R protease is essential while the ClpXP1/P2 protease is not, but the latter is important during certain stress conditions. The two adaptor proteins ClpS1 and ClpS2 appear to associate to only the ClpCP3/R protease (Stanne et al. 2007). ClpS1 is phylogenetically the ortholog to *E. coli* ClpS, while ClpS2 is so far unique to cyanobacteria (Lupas and Koretke, 2003, Nishimura et al., 2013). Although all the Clp proteins in *S. elongatus* are soluble, a proportion of ClpC, ClpP1, ClpR and ClpS2 appear to associate to membranes. As a result, it was suggested that a third Clp protease might exist attached to membranes, a ClpCP1/R protease with and without the ClpS2 adaptor (Stanne et al., 2007). Indeed, it has been proposed that a membrane-bound ClpCP1/R protease is responsible for degrading phycobilisomes (the main light-harvesting array in cyanobacteria) in *Synechocystis* together with NblA as the targeting adaptor (Karradt et al., 2008; Baier et al., 2014).

Our group has shown that the *S. elongatus* ClpP3/R proteolytic core consists of two heptameric rings, each containing four ClpR and three ClpP3 monomers (Andersson et al. 2009). The subunits in each ring are in alternating positions, except for two ClpR monomers that are side by side (i.e., P3-R-P3-R-P3-R-R). When modelled against the known structure of the *E. coli* ClpP core, the ClpP3/R complex appears to have entrance pores narrower in diameter, so narrow that the entrance is essentially sealed. It is this restricted entrance that could explain why the ClpP3/R core does not exhibit the peptidase activity characteristic of ClpP proteolytic cores (Anderssons et al. 2009). Recombinant ClpC and ClpP3/R can degrade model substrates such as α-casein *in vitro*, as well as N-end rule substrates when bound to the ClpS1 adaptor (Andersson et al., 2006, 2009*)*.

1.2.5. Apicomplexa

Apicomplexa is a family of protozoan parasites that possess a unique organelle called the apicoplast. The apicoplast appears to originate from ancestral red algae and is a non-photosynthetic plastid that contains several metabolic pathways, the most important of which is fatty acid biosynthesis. Apicomplexan parasites are the causal agents of several severe medical and veterinary diseases, the best known of which is malaria caused by the genus *Plasmodium*. The most severe form of human malaria is caused by *P. falciparum*, where the apicoplast is essential for parasite virulence (Ralph et al., 2004). *P. falciparum* has several Clp proteins: ClpB1-2, ClpC, ClpM, ClpP, ClpR and ClpS. All are nuclear-encoded except ClpM which originates from an apicoplast gene

(Ralph et al., 2004; El Bakkouri et al., 2010). ClpM, ClpB1, ClpC, ClpP and ClpR are all located in the apicoplast, with ClpS also predicted to be apicoplast-localized (Ralph et al. 2004; El Bakkouri et al., 2010; Rathore et al., 2010; El Bakkouri et al., 2012). All the Clp ATPases in *P. falciparum* have an insertion in the first AAA+ domain not present in their counterparts from other organisms. ClpM has two AAA+ domains like ClpC, but the first domain is missing conserved sequences in both the Walker A and B domains, indicating that this domain in ClpM is inactive (El Bakkouri et al., 2010). This and the fact that only ClpC has a recognizable P-loop suggest ClpC is the likely chaperone partner for any Clp proteolytic core in the apicoplast (El Bakkouri et al., 2010).

Both recombinant *P. falciparum* ClpP and ClpR proteins appear to form separate heptameric rings, but no interaction between the two different rings has yet been observed (El Bakkouri et al., 2010, 2012). As expected, the ClpR protein has no peptidase activity but ClpP exhibits only low activity. Crystal structures of *P. falciparum* ClpP revealed the characteristic tetradecamer consisting of two heptameric rings. The tetradecamer, however, has a more compressed structure compared to that of *E. coli* ClpP, with the catalytic triad of each ClpP subunit in an inactive conformation (El Bakkouri et al., 2010). Similar structures for ClpR also revealed heptameric rings that were in a more flattened conformation than is usual for ClpP proteases. The ClpR structure also indicated that the handle region of each ClpR monomer could not associate to the corresponding handle region within ClpP monomers, thus indicating that the formation of a ClpPR tetradecamer is unlikely in *P. falciparum* (El Bakkouri et al., 2012).

1.2.6. Photosynthetic eukaryotes

The main Clp protease in photosynthetic eukaryotes such as green algae and vascular plants is far more complex than those in bacteria. The model plant *Arabidopsis* has 23 Clp proteins, 16 of which are located inside plastids (Figure 3). These Clp proteins are much more abundant in leaf chloroplasts than in other types of plastids and almost all are constitutively expressed and essential for plant viability (Zheng et al., 2002; Peltier et al., 2004).

Arabidopsis has six ClpP (ClpP1-6) and four ClpR (ClpR1-4) paralogs, of which all but ClpP2 are located in chloroplasts. All are nuclear-encoded and post-translationally imported into chloroplasts, except for ClpP1 that is encoded within the plastome (Adam et al. 2001; Peltier et al. 2001). All chloroplast ClpP and ClpR proteins form a single heterogeneous proteolytic core comprised of two distinct rings, one consisting of ClpP3-6 (P-ring) and the other ClpP1 and ClpR1-4 (R-ring) (Sjögren et al. 2006). The R-ring appears to contain three ClpP1 and one of each ClpR subunit (Olinares et al., 2011), matching the ClpP:R stoichiometry of the cyanobacterial ClpP3R core complex and possibly its alternating subunit arrangement. In contrast, the P-ring consists of four ClpP paralogs, all of which are predicted to catalytically active, and to contain one ClpP3 monomer, three ClpP4, two ClpP5 and one ClpP6 (Olinares et al. 2011; Sjögren et al. 2014). The only other chloroplast Clp proteolytic core studied to date is that in the

green alga *Chlamydomonas reinhardtii*, which consists of three ClpP and five ClpR subunits (Majeran et al. 2005; Derrien et al. 2012). Assembly of the chloroplast Clp proteolytic core in *Arabidopsis* involves the two accessory ClpT1-2 proteins unique to plants. It appears that ClpT1 first associates to the P-ring followed by ClpT2, after which only then can the P-ring associate to the R-ring to form the intact proteolytic core and associate with the ATPase component (Sjögren and Clarke 2011). It has also been suggested that ClpT1 and ClpT2 might also been involved in substrate recognition (Olinares et al., 2011b).

Chloroplasts also contain four Clp ATPases (ClpC1-2, ClpB3 and ClpD), of which only ClpC1-2 and ClpD contain the P-loop domain and are likely to function as regulatory partners for the Clp proteolytic core (Adam et al., 2001; Peltier et al., 2004; Clarke et al., 2005). ClpC1 is by far the most abundant of the three ATPase components throughout vegetative growth, contributing to the bulk of Clp proteolytic activity observed in *Arabidopsis* chloroplasts. ClpC2 in comparison is relatively low abundant and constitutes no more than 10% of the total chloroplast ClpC protein (Sjögren et al. 2014). ClpD, which is a variant of ClpC only found in vascular plants, exists in similarly low levels during vegetative growth as that for ClpC2. However, ClpD is induced several fold in leaves during certain stresses and senescence (Nakabayashi et al. 1999; Zheng et al. 2002). Within chloroplasts, ClpD is only found in the stroma whereas a significant proportion of both ClpC1 and ClpC2 (ca. 30%) associate to the envelope membranes in addition to their stromal localization (Akita et al 1997; Nielsen et al 1997; Sjögren et al. 2014). Recent quantifications have revealed that more of the Clp proteolytic core exists in the stroma than the combined amount of ClpC1-2 and ClpD hexamers, suggesting that these HSP100 proteins function only as ATPase components of the Clp protease and not as independent chaperones (Sjögren et al., 2014).

A small proportion of the Clp proteolytic core has also recently been discovered attached to ClpC on the envelope membrane and in sufficient amounts to bind to all the available ClpC (Sjögren et al., 2014). Membrane ClpC associates to Tic40 and Tic110 (Akita et al., 1997; Nielsen et al., 1997), key components of the Tic-Toc multiprotein complexes responsible for translocating the majority of chloroplast-localized proteins across the outer and inner envelope membrane following their synthesis in the cytosol (review by Flores-Pérez and Jarvis, 2013). It has been proposed that ClpC via its ATPase activity helps drive protein transport through the Tic complex and aid in the subsequent refolding of the newly imported polypeptide (Akita et al., 1997; Nielsen et al., 1997; Jackson-Constan et al., 2001). The discovery of the Clp proteolytic core on the envelope membranes now suggests that ClpC might have an additional role during import, such as conferring a quality control mechanism on imported polypeptides as recently proposed (Sjögren et al., 2014).

Chloroplasts also contain an ortholog to bacterial ClpS that associates to ClpC1 and ClpC2 in *Arabidopsis*. Inactivation of *Arabidopsis* ClpS however, produces no obvious phenotypic changes, indicating that it plays a minor role in targeting protein substrates for the chloroplast Clp protease (Nishimura et al. 2013). Up to 25 putative substrates have now been identified for the stromal Clp protease, most of which perform various

regulatory roles in maintaining chloroplast functions (Sjögren et al. 2006; Stanne et al. 2009). Several additional substrates specifically recognized by ClpS have also recently been identified (Nishimura et al. 2013).

2. Results and Discussion

2.1. ClpC + ClpP3/R

2.1.1. Proteolytic core

The essential Clp protease in *Synechococcus* consists of the ATPase ClpC and a proteolytic core composed of two types of subunits, ClpP3 and ClpR (Clarke et al., 1996; Schelin et al., 2002; Andersson et al., 2006; Stanne et al., 2007). While ClpP3 is proteolytically active, ClpR is not and is found in only photosynthetic organisms and Apicomplexan parasites. The proteolytic core consists of two heptameric rings with four ClpR and three ClpP3 subunits in each arranged in an alternating configuration, except for two ClpR that are situated next to each other (Andersson et al., 2009). It is clear that the ClpP3/R proteolytic core associates to ClpC both *in vivo* and *in vitro* (Stanne et al., 2007; Andersson et al., 2006, 2009). There is certain specificity in the interaction between ClpC and ClpP3/R, in that ClpC does not associate to EcClpP while EcClpA cannot associate to ClpP3/R (Andersson et al., 2009, **paper I**). Several of the interactive determinants that have been identified for the *E. coli* Clp proteins are conserved in their cyanobacterial counterparts, such as the P-loop in ClpC. An essential interaction region is the N-terminus of EcClpP that associates directly to EcClpA and EcClpX, where the interaction site on EcClpX is on the pore-2 loop (Fig. 4) (Gribun et al., 2005; Martin et al., 2007, 2008; Jennings et al., 2008). There is some conservation between the N-terminus of ClpP3 to that of EcClpP, while there is no similarity in this region between ClpR and EcClpP. This suggests that the N-terminus of ClpR is likely to be important for the specific interaction with ClpC. That the N-termini of both ClpR and ClpP3 are involved in the interaction with ClpC is supported by 3D-modeling of the two subunits, with this region in ClpP3 and ClpR extending further out from the main proteolytic core compared to that of EcClpP (Andersson et al., 2009, **paper I**). To further investigate the importance of the N-terminus, we prepared alignments of many cyanobacterial ClpP3 and ClpR orthologs, of which several representative sequences are shown in Figure 6 to identify conserved regions in the N-terminal domain of both types of subunit.

Two conserved motifs were identified in the ClpR N-terminus, where the first is situated in position 12-15 (YYGD, Tyr-motif) and the second at 19-23 (RTPPP, Pro-motif). The ClpP3 N-terminus has one highly conserved region, that of the first six amino acids (MPIGVP, MPIG-motif) (Figure 5, **paper I**). To investigate these conserved regions, three different recombinant chimeric ClpR proteins were constructed, where different lengths of the N-terminus were replaced with the corresponding sequence from ClpP3. Two additional chimeric proteins were constructed in which the N-terminus of ClpP3 was

changed to the corresponding region in ClpR (Figure 5, **paper I**). The functionality of the constructed chimeric proteins was examined *in vitro* by assaying their proteolytic activity as well as assaying the ATPase activity of the ClpC chaperone partner; CN-PAGE was also performed to examine possible changes to the oligomeric structure of the chimeric proteins. Both the proteolytic and ATPase assays showed that all three of the conserved motifs are important for association to ClpC. All but one of the chimerics had a lower stimulation of the ClpC ATPase activity, indicating that they have a less stable association to ClpC. The exception was ClpR-N3, the chimeric with both the ClpR motifs intact but with the introduction of the MPIG-motif to the ClpR N-terminus. The ClpR-N3 core stimulated the ATPase activity of ClpC to a greater extent than the wild type ClpP3/R. This indicates that when all the subunits in the heptameric ring have the MPIG-motif the association between them and ClpC is stronger than normal.

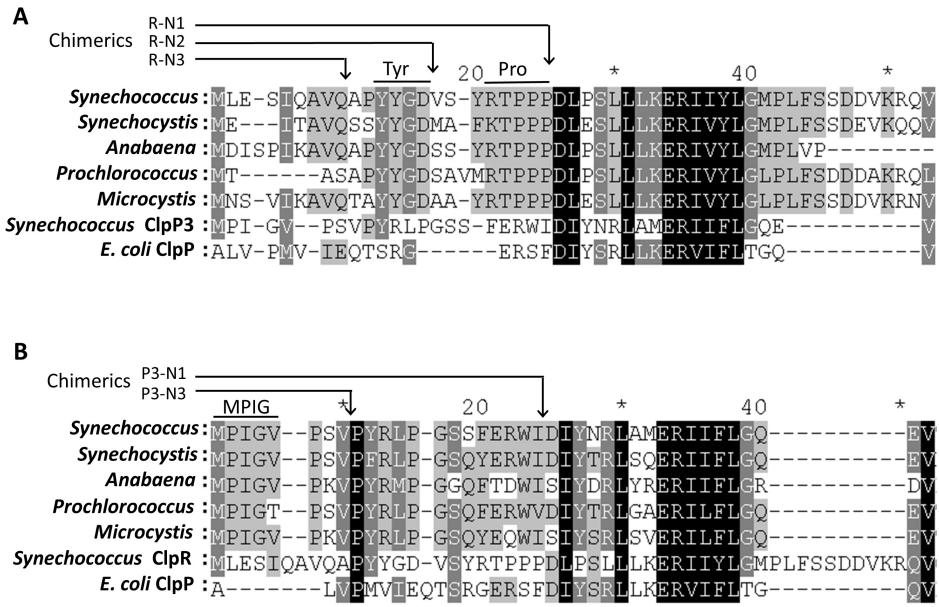

Figure 5. Identification of conserved motifs in the N-termini of ClpR and ClpP3. A representative sequence alignment of the N-terminal domain of five cyanobacterial ClpR (**A**) and ClpP3 (**B**) proteins. ClpP3 and ClpR sequence were from *Synechococcus elongatus* PCC7942 (*Synechococcus*), *Synechocystis* PCC6803 (*Synechocystis*), *Anabaena* sp 7120 (*Anabaena*), *Prochlorococcus* ma 1375 (*Prochlorococcus*) and *Microcystis* ae NIES.843 (*Microcystis*). Included for comparison was EcClpP plus the other subunit within the ClpP3/R core from *Synechococcus*. Functionally conserved amino acids are shaded in grey or black depending on the degree of conservation. The three conserved motif Tyr, Pro and MPIG are indicated, with arrows showing which parts were altered in the different chimeric (adapted from Tryggvesson et al., 2012).

However, as seen with the other chimerics if any of the conserved motifs are removed the interaction is negatively affected (**paper I**).

The N-terminal regions of ClpP3 and ClpR are not only involved in association to ClpC but they also affect proteolytic activity and oligomerization. This is clear when observing the degradation rate of the model substrate α-casein and the oligomeric structure of the different chimerics. When the Pro-motif was removed from the ClpR subunits, the ClpP3 and ClpR-N1 proteins formed two distinct core formations. One of the core complexes was similar in size to the wild type ClpP3/R while the other core was larger, corresponding in size to a double nonomeric core (**paper I**). In the *E. coli* ClpP core structure, a hydrogen bond formed between the Arg12 in one subunit to a Ser21 in the neighboring subunit is crucial for ring formation (Bewley et al., 2006). In comparison, all cyanobacterial ClpP3 proteins have a matching Arg at position 11 but no Ser, whereas the ClpR proteins have no Arg in that region but there is a Ser at position 27, which could explain why the ClpP3 and ClpR subunits alternate within the heptameric rings. Removing the Pro-motif (amino acids 19-23) as in ClpR-N1 is therefore likely to disrupt the association between the ClpR and ClpP3 subunit and lead to the formation of an aberrant core. This also explains why the ClpP3-N1 core is less stable than the others since the Arg11 is missing in all subunits (**paper I**). This indicates that the N-terminal region of both native ClpP3 and ClpR subunits is important for the correct formation of the proteolytic core, and that both subunits have been modified over time to strictly adhere to only one conformation of mixed heptameric ring.

The MPIG motif at the very N-terminus of ClpP3 might also be important for the actual degradation of substrates. The proteolytic core formed by ClpR-N2 and ClpP3 stimulated less the ATPase activity of ClpC, but it still degraded α-casein at the wild type rate. This indicates that removing the Tyr-motif in the ClpR subunit negatively affects the chaperone association, but that the incorporation of the MPIG-motif enhances the degradation rate. This is also observed in the ClpR-N3 chimeric that has both a better association to ClpC and a faster degradation rate than the wild type. In *E. coli*, the N-terminal region of ClpP appears to extend into the entry pore and interact with the acyl-enzyme intermediate, thereby affecting the maximum degradation rate (Bewley et al., 2006; Jennings et al., 2008; Lee et al., 2010). It is possible, therefore, that the ClpP3 N-terminus might stimulate the proteolytic activity of the ClpP3/R core in a similar way. Another explanation for the relatively fast degradation rate of the ClpR-N3/ClpP3 core could be that the ClpR N-terminus blocks the entry pore in the wild type structure. However, this does not fit with the new model of the ClpP3/R core where both the N-termini are unstructured and protrude from the core body (**paper I**). According to these results, it is clear that both the ClpR and ClpP3 subunits contribute to the interaction with ClpC.

To identify potential regions in ClpC that might also contribute to the specific interaction with ClpP3/R, sequence alignments of ClpC orthologs from various cyanobacteria and plants were prepared and included EcClpA for comparison. From these alignments, a short region (now termed the R-domain) was identified just downstream of the conserved P-loop of all ClpC proteins but was absent in EcClpA

(**paper I**). The importance of this region was investigated by constructing a recombinant chimeric of *Synechococcus* ClpC (ClpCA), in which the R-domain was replaced by the shorter corresponding region in EcClpA. The resulting chimeric ClpCA differed in its proteolytic core specificity from the wild type ClpC, in that it no longer associated to the ClpP3/R core but it did to EcClpP. This was observed both in the stimulation of ClpCA ATPase activity and in the degradation of α-casein. These results suggest that the R-domain is a decisive factor in determining the proteolytic core specificity of the ClpC chaperone partner. Also involved in this specificity could be differences in the hydrophobic clef on the surface of the proteolytic core of EcClpP and ClpP3/R. In EcClpP, the hydrophobic clef is formed by Tyr60, Tyr62, Phe82, Ile90, Phe112 and Leu189, and it is known that if the Phe112 is changed to Ala the association between ClpA and ClpP is compromised (Bewley et al., 2006). Interestingly, all but one of these amino acids is conserved in ClpR and ClpP3, with the Phe112 changed in both proteins to either an Ala as in ClpR or Leu/Val in ClpP3. This again suggests that few amino acid differences between the *E. coli* and cyanobacterial Clp subunits could well be responsible for their differences in interactive specificity.

Quantifications of the amount of ClpP3 and ClpR protein in *Synechococcus* show there is around 0.20 pmol of complex / μg Chl. In comparison, there is relatively less of the ClpC ATPase partner, with only ca. 0.13 pmol of ClpC hexamer / μg Chl. This suggests that ClpC is unlikely to function as a separate chaperone, and that its availability is the limiting factor in the formation of the ClpCP3/R protease (paper III). This situation corresponds to that in *Arabidopsis*, where the level of ClpC ATPases is also the limiting factor in the assembly of the chloroplast Clp protease (Sjögren et al., 2014); this similarity could well be due to the endosymbiotic origin of the chloroplast Clp protease from its ancestral cyanobacterial counterpart. Earlier it had been shown that *Synechococcus* ClpC is situated in both the soluble and membrane protein fractions (Stanne et al., 2007), and we later confirmed that 20% of the total ClpC content was attached to the membranes (**paper III**). However both ClpP3 and ClpR are found exclusively in the soluble fraction, inferring that the membrane-bound ClpC functions as an independent chaperone, although it cannot yet be excluded that a proportion of ClpP3/R detached from the membrane-bound ClpC during fractionation (**paper III**). The chloroplast ClpC orthologs in vascular plants are similarly distributed between the stroma and envelope membranes (Akita et al., 1997; Nielsen et al., 1997; Sjögren et al., 2014), but it is unlikely that the cyanobacterial and chloroplast ClpC proteins share the same exact function on the membranes. The envelope membrane-bound ClpC associates to the Tic110 protein, which is a major component of the translocation complex in the inner membrane (Tic) that facilitates preprotein import into chloroplasts (Akita et al., 1997; Nielsen et al., 1997). It has been proposed that ClpC acts as a motor protein driving translocation of preproteins through the Tic complex (Flores-Pérez and Jarvis, 2013), although the recent discovery of the Clp proteolytic core attached to ClpC on the envelope membrane raises the possibility of an additional role as part of a quality control system on chloroplast protein import (Sjögren et al., 2014). Given that there is no protein translocation system in cyanobacteria similar to the Tic complex, it is

almost certain that the specific function of membrane-bound ClpC in cyanobacteria differs from that of the chloroplast orthologs. It is tempting to speculate that following the endosymbiotic event the function of ClpC bound to the membranes changed in response to the transfer of most genes to the nucleus and the subsequent requirement for an efficient chloroplast protein import mechanism.

In paper I, we have continued the characterization of the ClpP3/R proteolytic core and identified three motifs in the N-terminal regions of ClpP3 and ClpR that are important not only for the association to ClpC but also for the correct assembly of the tetradecamer. The N-terminal regions of both ClpP3 and ClpR also influence the actual proteolytic activity of the ClpCP3/R protease. Furthermore, we have identified the R-domain in ClpC that is essential for the specific interaction between ClpC and ClpP3/R.

2.1.2. Adaptors

The ClpS adaptor in *E. coli* changes the substrate specificity of the ClpAP protease towards N-end rule substrates (Dougan et al., 2002; Erbse et al., 2006). The N-end rule pathway is dependent on destabilization amino acids at the N-terminus of the proteins targeted for degradation and has been identified in both eukaryotes and prokaryotes (Varshavsky et al., 1996). In eukaryotes, there are a multitude of distinct E3 ligases that recognize the destabilizing amino acids in the N-end rule substrates and deliver them to the 26S proteasome for degradation (Varshavsky et al., 1996, 2003)., It is ClpS in prokaryotes that recognizes the primary destabilization amino acids (Erbse et al., 2006), which in *E. coli* are Leu, Phe, Tyr and Trp (Tobias et al., 1991). Cyanobacteria possess two ClpS adaptors, of which the closest ortholog to *E. coli* ClpS is ClpS1 (Nishismura et al., 2013; **paper IV**). Our group had earlier shown that both *Synechococcus* ClpS1 and ClpS2 associate to ClpC *in vivo* (Stanne et al., 2007). Binding of ClpS1 to ClpC changes the substrate specificity of the chaperone *in vitro* towards N-end rule substrates (Andersson et al., 2006, 2009). To further investigate the role of ClpS1 and ClpS2 in *Synechococcus*, we constructed inactivation constructs for each *clpS* gene. Inactivation of the *clpS1* gene produced no obvious phenotypic changes during phototrophic growth, but it did alter the susceptibility of *Synechococcus* to oxidative stress (**paper V**). In contrast, all attempts to inactivate the *clpS2* gene proved lethal, suggesting that ClpS2 function is essential for cell viability (paper IV). During phototrophic growth, the relative amounts of both ClpS adaptors are low compared to that of ClpC (0.13 pmol), with ClpS1 almost an order of magnitude more abundant than ClpS2 (0.093 pmol ClpS1 *vs* 0.012 pmol ClpS2) (**paper V**). It is important to note that these values are based on ClpS functioning as monomers and ClpC as hexamers. A single EcClpS monomer can change the substrate specificity of EcClpA, although there is conflicting evidence on the extent of the stimulation of N-end rule substrate degradation. One study claims that only a single EcClpS monomer is required to bind to a ClpA hexamer to maximize the degradation of N-end rule substrates (De Donaties et al., 2010), while another claims four EcClpS monomers are needed for the same maximal activity (Hou et al., 2008).

Another study has also suggested that the proportion of ClpS (as monomers) to ClpA hexamer is ca. 2:1 during exponential growth of *E. coli* (Farell et al., 2005). If there is a need for more than one ClpS monomer to change the ClpC substrate specificity in *Synechococcus*, then even fewer ClpC hexamers would be occupied by a ClpS adaptor during phototrophic growth. It is clear that the relative ratio between ClpS:ClpA (2:1) in *E. coli* is higher than *Synechococcus* ClpS vs ClpC (0.8/1), which could indicate that relatively fewer N-end rule substrates exist in *Synechococcus*. The fact that ClpS2 function is vital for phototrophic growth, however, would suggest that the continued turnover of these few N-end rule substrates is essential for cell viability.

The *Synechococcus clpS2* ORF codes for a polypeptide with an unusually long N-terminal extension compared to that of other cyanobacterial ClpS2 proteins as well as all other bacterial ClpS orthologs including ClpS1 (**paper IV**). An antibody made to a 15 amino acid synthetic peptide matching the C-terminal sequence of *Synechococcus* ClpS2 detected in cell extracts a 17 kDa protein that matched the predicted size of the full-length ClpS2 and so it was concluded that this long N-terminal was translated (Stanne et al., 2007). Subsequent work, however, cast doubt on the true identity of the protein detected by immunoblotting, and so further investigations were initiated. First, a new antibody was made against a full-length recombinant *Synechococcus* ClpS2 protein, which eventually detected a protein of ca. 12 kDa in wild type *Synechococcus* but not the original 17 kDa protein (**paper IV**). To confirm the identity of the 12 kDa protein, the native *clpS2* gene was replaced in *Synechococcus* with a modified version coding for a ClpS2 protein with a C-terminal His_6-tag. Immunoblotting with the new antibody of cell extracts detected a single protein in the transformant that was slightly larger than the one in wild type *Synechococcus*, but the size difference corresponded to that of the His_6-tag. Later examination of the N-terminal sequence of ClpS2 revealed an internal Val (position 64) that is coded for by GTG, a codon that is well known to function as a translational start for many bacterial genes. Translation from the Val64 codon would produce a protein of similar size to ClpS1, which was later shown by phylogenetic analysis to have an N-terminus that more closely matched those from other cyanobacterial ClpS2 orthologs (**paper IV**). Over-expression in *E. coli* of *Synechococcus* ClpS2 starting from the Val64 position (which was changed to a Met) produced a recombinant protein that again matched the size of the native protein with the addition of the His_6-tag (**paper IV**), further supporting that Val64 is the true start of the native ClpS2 protein. Given the misidentification of *Synechococcus* ClpS2 in the earlier study (Stanne et al., 2007), we re-examined the cellular localization of ClpS2. Previously it was shown that ClpS1 was exclusively a soluble protein whereas ClpS2 was equally distributed between the membrane and soluble protein fractions (Stanne et al., 2007). In the new fractionations, both ClpS1 and ClpS2 were exclusively soluble proteins (paper IV), where also 80% of the ClpC is situated (**paper III**).

In **paper IV**, we show that *Synechococcus* ClpS2 associates to ClpC and inhibits the degradation of α-casein while promoting the degradation of certain N-end rule substrates. The ClpS1 and ClpS2 adaptors differ in the types of N-end rule substrates they recognize. ClpS1 targets proteins with a Tyr and Phe as the primary destabilization

amino acid but not Leu, while ClpS2 recognizes Leu but not Tyr or Phe. This difference in substrate specificity between ClpS1 and ClpS2 varies from that of EcClpS, which can recognize all three destabilization amino acids. This variation can be explained at least in part by conserved changes in important amino acids in the ClpS1 and ClpS2 proteins. Two amino acids in EcClpS define the size of the cavity and thereby its specificity - Met40 and Met62 (Wang et al., 2008a; Schuenemann et al., 2009). Both these amino acids are conserved in ClpS1 but different in ClpS2, where Met40 is changed to a Phe and Met62 to a Thr. In EcClpS, when Met40 was changed to an Ala the substrate specificity broadened to include Ile, but at the expense of weakening the association to the usual N-terminal destabilization amino acids (Wang et al., 2008a; Schuenemann et al., 2009). That ClpS2 lacks Met in these positions probably indicates that it has evolved to recognize another set of de-stabilizing N-terminal residue(s). It is also possible that this change in ClpS2 reduces its affinity for other destabilization amino acids, which would explain why it fails to recognize the Phe or Tyr. The aromatic side group of the Phe residue in position 42 would presumably make the substrate cavity of ClpS2 smaller and thus more restrictive in the types of substrates that could be accommodated, which could in part explain why ClpS2 recognizes Leu but not the larger Phe or Tyr residues. This might furthermore indicate that the destabilization amino acids recognized by ClpS2 are ones with less bulky side groups, although it is clearly not size alone that determines the amino acid specificity since ClpS2 does not recognize Ser, Ala or Lys (**paper IV**).

Another interesting observation in **paper IV** was that ClpS1 did not recognize N-end rule substrates with Leu despite having the conserved Met40 and Met 62 residues as well as other amino acids important for substrate association in EcClpS. One such important amino acid is Leu99 that is also conserved in all cyanobacterial ClpS orthologs, but in many ClpS1 proteins the three following amino acids are unusually polar in nature, which could influence its substrate specificity. The different substrate specificities between *Synechococcus* ClpS1 and ClpS2 *in vitro* also appear to occur *in vivo*, given that the more abundant ClpS1 cannot compensate for the inactivation of ClpS2, or that ClpS2 levels do not change when ClpS1 is mutated is indicated by the result observed in paper V, that the ClpS2 is not induced when the ClpS1 is missing (**paper IV**). It is almost certain that the second ClpS adaptor in cyanobacteria has evolved to recognize a distinct set of substrates, while at the same time not losing the specificity of ClpS1. One selective advantage for having a second ClpS protein is that if the additional substrate recognition was instead incorporated into the already existing single ClpS adaptor, then the broadened specificity could compromise its association to the other previously-recognized substrates. Whatever the reason underlying the appearance of ClpS2, it is clear from the phylogenetic analysis that ClpS2 is present in all cyanobacterial species examined and is targeting proteins whose turnover is crucial for cell viability. It is also evident that whatever these substrates of ClpS2 are, their regulation in this way is specific to cyanobacteria and was not retained in chloroplasts after the endosymbiotic event.

The only substrates that ClpS2 has so far degraded *in vitro* are LVK-Dps and LVK-GFP, both of which were not degraded by ClpS1. In *E. coli*, it is only the truncated form of the native Dps protein that is a substrate for EcClpS, while the full-length Dps is instead degraded by ClpXP (Flynn et al., 2003; Stephani et al., 2003; Schmidt et al., 2009). The fact that DspA levels do not change in the *Synechococcus clpS1* mutant further supports it is ClpS2 and not ClpS1 that recognizes the DpsA protein *in vivo* (**paper V**). Of course, DpsA cannot be the only native substrate for ClpS2. Simply searching the *Synechococcus* proteome for proteins with destabilization amino acids at the N-terminus is limited, since most of the possible substrates for EcClpS are now known to have internal destabilization amino acids that are only recognizable after the upstream N-terminal sequence is removed (Humbard et al., 2013). Despite this, screening the *Synechococcus* proteome for the sequence Met- Leu-Val at or near the N-terminus revealed ten potential substrates for ClpS2, four of which were hypothetical proteins. The other six proteins perform diverse roles in cyanobacteria and include Arabinose efflux permease, Anthranilate phophoribysltransferase, HrcA and proton-translocating NADH-quinone oxidoreductase (NDH-1), chain M.

Another interesting observation in both **paper IV** and **V** is the possibility that ClpS2 in some cyanobacteria, including *Synechococcus,* might be a substrate for the ClpXP1/P2 protease. Many ClpS2 sequences include a recognition tag for ClpX of non-polar amino acids at the C-terminus, ending most frequently with two Ala residues. Although this twin Ala combination is essential for the recognition of the substrate CtrA by ClpX (Domian et al., 1997; Flynn et al 2003), it is more commonly found at the end of the SsrA tag that is typically added to the C-terminus of mistranslated polypeptides and targets them for degradation by the ClpXP protease (Tu et al., 1995; Keiler et al., 1996). The two Ala residues in the SsrA sequence are crucial for ClpX recognition and their presence at the C-terminus of many ClpS2 orthologs suggests they might be regulated by the ClpXP1/P2 protease in cyanobacteria. This is supported by the fact that ClpS2 accumulates in the *Synechococcus clpP1* knockout mutant that lacks a functional ClpXP1/P2 protease (**paper V**). To investigate the possibility that ClpS2 is a substrate for ClpX we performed an *in vitro* degradation assay with recombinant ClpS2, ClpP1/P2 and EcClpX. Although no significant degradation of ClpS2 was observed, we cannot exclude the possibility that ClpS2 recognition is specific to cyanobacterial ClpX and therefore lacking in EcClpX, especially since ClpS2 is unique to cyanobacteria.

In paper IV, we have shown that both ClpS adaptors in Synechococcus are relative low abundant compared to ClpC, with the level of ClpS1 almost an order of magnitude higher than ClpS2. Despite its low level, however, ClpS2 activity is essential for phototropic growth of Synechococcus while that of ClpS1 is not. Both ClpS adaptors recognize N-end rule substrates but with different specificity - ClpS1 targets proteins with destabilization Tyr and Phe residues at the N-terminus while ClpS2 recognizes those with Leu.

2.2. ClpX + ClpP1/P2

2.2.1. Structure

There are at least two Clp proteases in *Synechococcus*, ClpCP3/R and ClpXP1/P2 (Stanne et al., 2007). Despite the ClpX chaperone being essential for phototrophic growth, the ClpXP1/P2 protease is not (Schelin et al., 2002; Clarke et al., 1998; Porankiewicz et al., 1998). It has earlier been shown that the ClpP1 is induced during three different stress conditions: low temperatures, UV-B irradiation and high light intensities (Clarke et al., 1998; Porankiewicz et al., 1998). The fact that ClpX and ClpP2 were also induced by high light intensities (Schelin et al., 2002) suggested that the activity of ClpXP1/P2 protease in *Synechococcus* was more important under stress conditions. The work done in this thesis has now continued the characterization of the ClpXP1/P2 protease both structurally and functionally (**papers II**, **III** and **V**). We could show that the ClpP2 in *Synechococcus* is autolytically processed, removing the first 30 amino acids. EcClpP also post-translationally processes its N-terminus to produce the mature protein, but only 14 amino acids are removed (Maruizi et al., 1990). Using a similar strategy to that adopted for ClpP3/R (Andersson et al., 2009), recombinant ClpP1 and ClpP2 were co-expressed in *E. coli* with only ClpP2 containing the His_6 affinity at the C-terminus. That both ClpP1 and ClpP2 were purified together after affinity and gel filtration chromatography confirmed the two recombinant ClpP proteins interacted. Separation of the recombinant ClpP1/P2 proteins by CN-PAGE revealed an oligomer matching a single mixed heptamer (**paper II**), similar to that observed for the native ClpP1/P2 *in vivo* (Stanne et al., 2007), suggesting the intact tetradecamer is unstable under electrophoretic conditions.

In order to determine how ClpP1 and ClpP2 are organized within the proteolytic core, the recombinant proteins were studied by non-denaturing mass spectrometry. These experiments showed that the proteolytic core is composed of the same amount of ClpP1 and ClpP2 in the form of two different mixed heptameric rings - 4ClpP1+3ClpP2 and 3ClpP1+4ClpP2. The two different ClpP subunits alternate in the assembly of each ClpP1/P2 heptamer, similar to how the ClpP3 and ClpR subunits are arranged within the ClpP3/R proteolytic core (Andersson et al., 2009). These two different heptamers then appear to form two distinct tetradecamers - (4ClpP1+3ClpP2) + (3ClpP1+4ClpP2) and 2×(3ClpP1+4ClpP2), with little evidence for the possible third combination of 2(4ClpP1+3ClpP2) (**paper II**). This means that both the ClpP1/P2 and ClpP3/R proteolytic cores in *Synechococcus* are composed of heterogeneous heptameric rings (Andersson et al., 2009), the existence of which has only been identified to date in Clp proteases from photosynthetic organisms. Certain Gram-positive bacteria also contain a ClpP1/P2 proteolytic core, but these are composed of one homogeneous heptameric ring of ClpP1 and another of ClpP2 (Akopian et al., 2012; Raju et al., 2012; Zeiler et al., 2011, 2013).

In paper II, we show that the Synechococcus ClpP1/P2 proteolytic core is composed of mixed heptameric rings, either 4ClpP1+3ClpP2 or 3ClpP1+4ClpP2. Within each of these rings, the ClpP1 and ClpP2 subunits are arranged in an alternating pattern. Two distinct tetradecamers appear to assemble from these heptamers consisting of either (4ClpP1+3ClpP2) + (3ClpP1+4ClpP2) or 2×(3ClpP1+4ClpP2).

2.2.2. Function

We also further investigated the functionality of the ClpP1/P2 proteolytic core. We made numerous attempts to purify recombinant *Synechococcus* ClpX by over-expression in *E. coli*, but failed each time due to the majority of the protein forming inclusion bodies. Attempts to resolubilize the inclusion bodies using different denaturants along with different protein refolding protocols also proved unsuccessful, with the ClpX protein either precipitating during the procedure or the final purified protein being inactive. As a consequence, we instead used EcClpX to study the proteolytic activity of ClpP1/P2 *in vitro*. ClpP1/P2 together with EcClpX degraded the model GFP-SsrA substrate under the same standard conditions used for the EcClpXP protease (**papers II** and **IV**). We could also show that the recombinant ClpP1/P2 has a specificity in which chaperon partner it associate to. The ClpP1/P2 has only proteolytic activity with EcClpX and not with ClpC, EcClpA or the chimeric ClpCA (**paper IV** and result not shown).

Although it had been assumed that the non-essential ClpXP1/P2 protease would be less abundant in *Synechococcus* than the essential ClpCP3/R protease during phototrophic growth, this now appears not to be the case. Quantifications of ClpP1 and ClpP2 levels in wild type *Synechococcus* revealed equimolar amounts of both ClpP paralogs, consistent with the near 1:1 stoichiometry of both subunits in the ClpP1/P2 core as determined by native mass spectrometry (**paper II**). Based on this stoichiometry, there is ca. 0.15 pmol of the ClpP1/P2 proteolytic core, which is considerably less than that of the chaperone partner ClpX (0.40 pmol hexamers / µg Chl). Approximately one-third of this ClpX, however, is bound to membranes leaving 0.28 pmol of ClpX hexamer available for the soluble ClpP1/ClpP2 proteolytic core. This suggests therefore that it is the amount of ClpP1/P2 proteolytic core that dictates how much ClpXP1/P2 exists in *Synechococcus* (i.e., 0.15 pmol complex), which now appears to be similar to that of ClpCP3/R (0.13 pmol) (**paper III**). This also indicates that up to two-thirds of the ClpX content in *Synechococcus* could be functioning as an independent chaperone (**paper III**), which would be consistent with the known chaperone activity of EcClpX (Mhammedi-Alaoui et al., 1994; Burton et al., 2001; Levchenko et al., 1995; Burton and Baker, 2003, 2005).

It has been observed in *L. monocytogenes* that in addition to the LmClpP1/P2 proteolytic core LmClpP2 can also form an active proteolytic core by itself, whereas LmClpP1 alone assembles only into inactive heptameric rings (Zeiler et al., 2011, 2013). In *Synechococcus*, it has long been considered that ClpP1 might form a proteolytic core separate from the ClpP1/P2. This assumption was based on the observation that loss of

ClpP1 in the Δ*clpP1* strain led to an almost complete loss of the ClpP2 protein, whereas ClpP1 levels in the Δ*clpP2* strain remained unchanged (Schelin et al., 2002; Stanne et al., 2007; **paper V**). To further investigate the potential for separate ClpP1 and ClpP2 proteolytic cores, we purified both individual ClpP paralog by over-expression in *E. coli*. We show that recombinant ClpP2 is unable to form higher molecular mass structures and as such displays no proteolytic activity with any of the chaperone partners tested. ClpP1, on the other hand readily formed a stable tetradecameric proteolytic core that was proteolytically active with all the Clp ATPases tested (i.e., ClpC, EcClpX, EcClpA and the ClpCA chimeric) (**paper III**). However, the proteolytic activity of ClpP1 was dependent on the $MgCl_2$ concentration, with ClpP1 being inactive at 4 mM $MgCl_2$ but steadily increasing in activity from 7 mM until reaching the maximum degradation rate at 16 mM $MgCl_2$. The reason for this effect remains unclear, but it does not appear to affect the formation of the ClpP1 tetradecamer (**paper III**). In *L. monocytogenes*, the assembly of the LmClpP1/P2 proteolytic core activates the LmClpP1 subunit, presumably by the association between the LmClpP1 and LmClpP2 rings modifying the structure of the catalytic triad in LmClpP1 from a disordered state to one that is active (Zeiler et al., 2011, 2013). In the absence of crystal structures for the ClpP1/P2 or ClpP1 proteolytic cores, it remains unclear if the catalytic triad of ClpP1 is disordered at low $MgCl_2$ concentration or if another as yet unknown factor is involved.

Both ClpP1 and ClpP2 contribute to the catalytic activity of the ClpP1/P2 proteolytic core, but this rate is several times slower than that of the ClpP1 only core when assayed at the higher $MgCl_2$ concentrations. Interestingly, when either ClpP1 or ClpP2 was inactivated the catalytic activity of the proteolytic core was faster than that of the native ClpP1/P2 complex. This increase in proteolytic activity was observed at both 4 and 20 mM $MgCl_2$ when it was the ClpP2 subunit that was mutated, but only at 20 mM $MgCl_2$ when it was the ClpP1 subunit that was inactivated. Instead, at 4mM $MgCl_2$ the proteolytic core with inactive ClpP1 subunits was catalytically slower than the native ClpP1/P2 complex. These results indicate that the incorporation of both ClpP1 and ClpP2 into the one tetradecamer actually generates a proteolytic core whose catalytic activity is tightly regulated.

In paper III, we show that both the ClpP1 and ClpP2 subunits contribute to the proteolytic active of the ClpP1/P2 core, but that its catalytic activity is relatively slow. ClpP1 can also form an active tetradecamer by itself that has relatively fast catalytic activity but only at higher $MgCl_2$ concentrations, whereas ClpP2 alone is unable to form higher molecular mass complexes. Although the exact effect of $MgCl_2$ on the different Clp proteolytic cores remains unclear, it appears to affect more the actual catalytic sites of ClpP1 and ClpP2 and not their oligomerization.

2.3. A third Clp proteolytic core?

For many years our group has considered the possibility of a third Clp protease in *Synechococcus*. Previous fractionations had shown that a small proportion of ClpC,

ClpP1 and ClpR were associated to the membranes, raising the possibility that they could form an additional Clp protease (Stanne et al., 2007). That there could be a third protease including ClpP1 was suggested even earlier by the fact that ClpP1 is induced in wild type *Synechococcus* during cold and UV-B irradiation stress but ClpP2 is not (Clarke et al. 1998; Porankiewicz et al. 1998), and that loss of ClpP2 has no effect of ClpP1 levels (Schelin et al., 2002). To further investigate this, we repeated the fractionation studies to quantify the amount of each Clp protein associated to the membrane relative to that in the soluble fraction. However, these new fractionations revealed that only ClpC and ClpX were bound to the membrane, while ClpP1 and ClpR was now exclusively soluble proteins along with ClpP2 and ClpP3 (**paper III**). We also tested the possibility of a ClpP1/R proteolytic core *in vitro* by co-expressing both proteins in *E. coli* using the same system that successfully purified the ClpP3/R and ClpP1/P2 core complexes. However, it was only the ClpP1 protein with the His_6-tag that was eventually purified with no trace of ClpR. Quantification of the relative amounts of each Clp protein also did not support the possibility of a third Clp proteolytic core under standard growth conditions. The amounts of ClpP3 and ClpR (1.31 and 1.46 pmol/μg Chl, respectively) were consistent with their known 3:4 stoichiometry within the ClpP3/R proteolytic core (Andersson et al., 2009), while to equal amounts of ClpP1 and ClpP2 also corresponded to the subunit stoichiometry of the ClpP1/P2 proteolytic core as recently documented (**paper II**). Taking all these results together suggests that the possible existence of a third Clp protease consisting of ClpCP1/R is now highly unlikely.

Although it is almost certain that a third Clp protease does not exist in *Synechococcus* during phototrophic growth, there remains the possibility during certain stress conditions. Given that ClpP1, and not ClpP2 is inducible during stresses such as cold and UV-B irradiation (Clarke et al., 1998; Porankiewicz et al., 1998) and that ClpP1 can form its own proteolytic core *in vitro*, together suggests that it might form a third Clp protease in *Synechococcus*. According to the *in vitro* assays, the degradation rate of the ClpP1 proteolytic core is considerably faster than that of both the ClpP1/P2 and ClpP3/R complexes, which would be an advantage for a stress-inducible protease given the increased propensity for protein damage during periods of stress. It is also plausible that a stress-inducible ClpP1 proteolytic core could function together with the ClpX chaperone, since according to the quantifications there is significantly more ClpX hexamer in *Synechococcus* than the ClpP1/P2 proteolytic core (**paper III**). However, the apparent lack of chaperone specificity of the ClpP1 proteolytic core suggests it could also function with ClpC as the chaperone partner, thereby increasing the range of potential protein substrates that could be degraded during the stress. We have attempted to test the affinity of ClpC for either ClpP1 or ClpP3/R by performing *in vitro* competition experiments using either inactive ClpP1 with active ClpP3/R or inactive ClpP3/R with active ClpP1 and testing the resulting rates of α-casein degradation. However, we could not detect any reduction in the degradation of α-casein with either combination of active and inactive proteolytic cores, suggesting that ClpC has equal affinity for both ClpP1 and ClpP3/R. These experiments, however, need to be repeated and fine-tuned before we can make this conclusion.

Another observation that suggests a ClpP1 only proteolytic core can function with ClpX *in vivo* comes from the oxidative stress experiments discussed below. The level of both ClpP1 and ClpP2, along with ClpX increased in wild type *Synechococcus* during the recovery phase after H_2O_2 treatment, suggesting a role for the ClpXP1/P2 protease. The possibility that ClpP1 can form two distinct proteolytic cores, however, helps explain why the *clpP1* mutant was so sensitive to the oxidative stress while the *clpP2* mutant was not. If the ClpXP1/P2 protease is normally involved in the recovery from oxidative protein damage, then ClpXP1 could substitute for this activity when ClpP2 is absent as in Δ*clpP2*. When ClpP1 is missing, however, not only is ClpXP1/P2 activity lost but so is any from a ClpXP1 protease, resulting in the extremely slow recovery of the Δ*clpP1* strain after addition of H_2O_2 (**paper V**).

It appears likely that a third Clp protease consisting of a homogeneous ClpP1 proteolytic core is induced in Synechococcus during certain stress conditions. The potential association of the ClpP1 proteolytic core to ClpC and/or ClpX could provide the extra degradative capacity needed in response to the accumulation of stress-damaged polypeptides.

2.4. Involvement of the Clp protease in phycobilisome degradation

When cyanobacteria are starved of certain critical nutrients such as nitrogen or sulfur, the cells begin to bleach due to the rapid degradation of phycobilisomes (PBS), the main light-absorbing antenna complex. This degradation of PBS is thought to reduce light absorption during the stress period and thus minimize the risk of photo-damage, as well as provide a source of recycled amino acids for continued protein synthesis. The PBS is a large complex that is constructed from many protein subunits and pigments. In *Synechococcus*, the core structure is formed by the allophycocyanins from which several rods extend out from. The rods consist of hexamers of phycobiliproteins that bind both α and β phycocyanin, with each hexamer in the rod separated by specific linker proteins (Fig. 7A). A small polypeptide called NblA is induced during such nutrient stresses and is known to promote PBS degradation by presumably destabilizing the PBS complex (Collier and Grossman, 1994; Dines et al., 2008). A study from another group showed that NblA from *Nostoc* could associate to ClpC *in vitro* (Karradt et al., 2008), raising the intriguing possibility that NblA functions as an adaptor for ClpC in targeting the PBS for degradation by the Clp protease. We therefore investigated if NblA could also associate to the Clp proteins in *Synechococcus.* We first obtained a construct to overexpress *Synechococcus* NblA as a His-tagged protein from Professor Noah Adir (Hebrew University of Jerusalem, Israel), but the purification of the recombinant NblA proved difficult. After trying many combinations of buffers and purification procedures, the recombinant NblA could only be kept soluble in the presence of high concentrations of imidazole (500 mM). However, this high concentration of imidazole also had an inhibitory effect on the degradation activity of ClpCP3/R, which could only be partly overcome by the addition of extra ATP and $MgCl_2$ to the assay buffers. Using these

modified conditions, we first tested if *Synechococcus* NblA functioned as an adaptor for ClpC in the same way as ClpS by blocking the degradation of α-casein. However, when NblA was added to ClpC and ClpP3/R no inhibition of α-casein degradation could be observed. Interestingly, the one protein that was degraded during this assay was NblA itself (Fig. 7B). This unexpected finding was later confirmed by a study showing that *Synechocystis* NblA was degraded *in vitro* by a supposed ClpCP1/R protease (Baier et al., 2014). They also showed that NblA could associate to ClpC, the PBS and a phycocyanin (PC)-GFP construct, concluding that this indicated NblA was an adaptor for ClpC in PBS degradation although no actual PBS degradation was observed (Baier et al., 2014). To understand why *Synechococcus* NblA did not block α-casein degradation by ClpCP3/R, we tested the solubility of NblA when combined with ClpC and the PBS (purified from wild type *Synechococcus*). The NblA was indeed soluble when added to the purified PBS, but was insoluble when combined with ClpC, indicating that the recombinant NblA could bind to the PBS but not to ClpC.

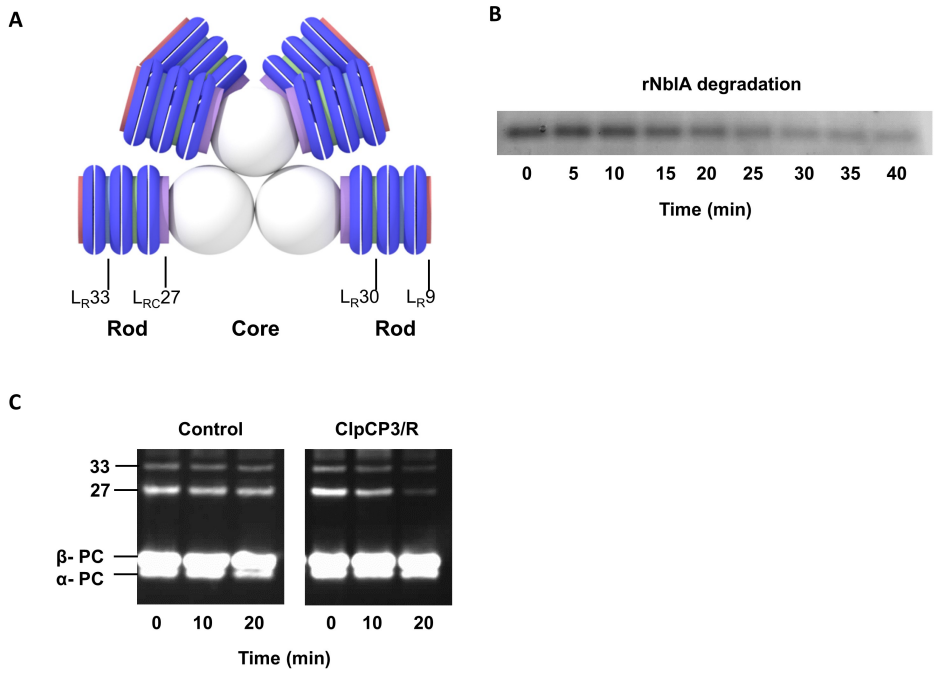

Figure 6. Degradation of *Synechococcus* NblA and phycobilisome rod linkers by the Clp protease *in vitro*. (**A**) A schematic overview of the basic structure of the phycobilisomes (PBS) in S*ynechococcus*, where the position of the different linkers are indicated. (**B**) Degradation of recombinant *Synechococcus* NblA (rNblA) by the ClpCP3/R protease as visualized by denaturing-PAGE and colloidal coomassie blue staining. (**C**) PBS degradation assay with the ClpCP3/R protease performed at 37°C with an ATP-regeneration system. The control reaction contained all the components of the test sample except no Clp proteins were added. The result was visualized by immunoblotting with the PBS antibody. The protein bands that correspond to the 27 and 33 kDa linkers, as well as α- and β-phycocyanin (PC) are indicated.

The lack of ClpC association was unlikely due to the His_6-tag on the NblA protein since the *Synechocystis* NblA contained the much larger GFP-tag and was still capable of binding ClpC (Baier et al., 2014). Instead, it is possible that high concentration of imidazole in the *Synechococcus* NblA sample could have adversely affected the association to ClpC.

To further investigate the role of the Clp proteases during PBS degradation, we used the purified PBS in a proteolytic assay. We first examined the stability of the PBS under the assay conditions and observed some instability of the PBS linkers over time (Fig. 7C). However, when ClpC and ClpP3/R was included in the assay there were a clear degradation of the 27 and 33 kDa PBS linkers, but no degradation of the phycocyanins. There might also have been degradation of the 30 kDa linker, but the level of this protein was too low to quantify accurately. It should be noted that prolonging the assay also resulted in no degradation of the phycocyanins. These results indicate that the ClpCP3/R protease might be involved in the degradation of the PBS linkers. Unfortunately, we could not test the effect of NblA on this degradation because of the inhibitory effect of the imidazole. One suggested function of NblA during PBS degradation is to destabilize the entire complex by binding in the gap between PBS-subunits (Dines et al., 2008). Since a slight instability of the PBS was observed under the conditions that the degradation assay was performed, we repeated the degradation assays with added phosphate to maintain the PBS in a more stable state. Under these conditions, no instability of the PBS was observed, but also no degradation of any of the PBS proteins when ClpCP3/R was added with or without NblA. It is plausible that the degradation of the PBS linkers by the ClpCP3/R protease is dependent on the PBS complex being first destabilized. It has been suggested that the degradation of the *Synechocystis* PBS linkers requires them to first be dephosphorylated by a phosphatase (Baier et al., 2014). It has also been suggested in *Nostoc* that dephosphorylation of the PBS linkers is a signal for PBS degradation (Ke and Hassletorn, 2013). It is therefore possible that in the absence of added phosphate the *Synechococcus* PBS linkers are dephosphorylated and thus susceptible to degradation by the ClpCP3/R protease. The role for NblA in this scenario *in vivo* would therefore be to destabilize the PBS complex in such a way that the linkers become accessible. NblA might then recognize the linkers as substrates (which could include a dephosphorylation step) and initiate their degradation by associating to the Clp protease. Once the linkers are degraded, the large amount of phycocyanins could then be freely degraded by other proteases.

We have observed degradation of the PBS linkers by the ClpCP3/R protease in vitro, suggesting that the Clp protease might be involved in PBS degradation in vivo. It is likely that NblA functions to destabilize the PBS so that the linkers become accessible for degradation by the ClpCP3/R protease. We could not confirm however if NblA functions as an adaptor for the Clp protease.

2.5. Involvement of Clp proteins during oxidative stress

Inactivation of the *clpS1* gene in *Synechococcus* resulted in no obvious phenotypic changes under standard growth conditions. The Δ*clpS1* strain, however, was more resistant to exogenously added H_2O_2 than the wild type, a phenomenon that was also observed for the Δ*clpP2* strain. In contrast, the Δ*clpP1* strain was extremely sensitive to the H_2O_2 treatment. The behavior of all three mutants was due at least in part to changes in the level of the peroxidase-catalase KatG (**paper V**). The KatG enzyme dismutates the H_2O_2 to water and molecular oxygen by a mechanism in which the two oxygen atoms that form the O_2 originate from the same H_2O_2 molecule (reviewed by Vlasits et al., 2010). It was earlier shown that the inactivation of *katG* in *Synechococcus* causes increased sensitivity to exogenously added H_2O_2 (Perelman et al., 2003). In the two less sensitive mutants, Δ*clpS1* and Δ*clpP2*, the basal level of KatG was considerably higher than that in the wild type due to elevated expression of the *katG* gene. This would explain why the mutants are more resistant to the oxidative stress, in that the increased amount of basal KatG could more quickly dismutase the added H_2O_2. The sensitivity of the Δ*clpP1* strain could also be explained by its relatively low basal level of KatG. However, the reduced level of KatG in Δ*clpP1* was not due to a down-regulation of *katG* gene expression but instead to increased instability of the enzyme (paper V). It is therefore the elevated rate of KatG degradation before and during the oxidative stress that contributes to the sensitivity of Δ*clpP1* to the added H_2O_2. The fact that KatG is not induced in cyanobacteria during oxidative stress as shown in paper V and by others (Li et al., 2004; Kanesaki et al., 2007), only intensifies the loss in KatG content as observed in the Δ*clpP1* strain. It might be that the ClpCP3/R protease in association with ClpS2 is responsible for this accelerated degradation of KatG, since all the components of this protease are induced in Δ*clpP1* (**paper V**). To investigate this possibility, we performed an *in vitro* degradation assay with recombinant KatG and ClpCP3/R with or without ClpS1/S2, but could only observe a slow rate of KatG degradation by ClpCP3/R alone and in the presence of ClpS2. Although KatG has not been identified as a substrate for the Clp protease in other organisms, several other proteins involved in oxidative stress have, including TrxB and the peroxidase AhpCF in both *Staphylococcus aureus* and *E. coli* (Weichart et al., 2003; Michel et al., 2006).

While KatG has been shown to be important during the removal of exogenous added H_2O_2, the peroxidase 2-Cys prx is more important at relatively low levels of H_2O_2 (Perelman et al., 2003). The 2-Cys prx functions as homodimer and catalyzes the reduction of H_2O_2 by using the thioredoxin system as reducing agents. The basic mechanism of 2-Cys prx is that a thiol group in the catalytic cysteine is oxidized by the H_2O_2 to form a sulphenic acid. The sulphenic acid is then reduced to a thiol once more and H_2O is released (Wood et al., 2003). Our results are consistent with those from the earlier study (Perelman et al., 2003) since in the highly sensitive Δ*clpP1* strain the doubling of both 2-Cys prx transcript and protein cannot compensate for the loss of KatG (**paper V**). Despite this, the induction of 2-Cys prx in Δ*clpP1* might be directly connected to the endogenous levels of H_2O_2. It is known that 2-Cys prx is induced in

response to elevated cellular levels of H_2O_2 (Stork et al., 2002). Given that the decrease in basal KatG levels in Δ*clpP1* would likely elevate the endogenous H_2O_2 concentration in this strain, this could explain the observed induction of 2-Cys prx (**paper V**). This would also explain the reduced transcriptional level of *2-cys prx* in Δ*clpS1* and Δ*clpP2*, in which the higher level of KatG would almost certainly lower the endogenous H_2O_2 concentration.

Dps is another protein that has been linked to oxidative stress in *E. coli*, where it is the truncated version of the protein that is a substrate for EcClpS (Ninnis et al., 2009; Schmidt et al., 2009). As discussed earlier, it is likely that the DpsA ortholog in *Synechococcus* is a substrate for ClpS2 rather than ClpS1. However, it is possible that the less abundant ClpS2 targets DpsA during normal growth, while the more abundant ClpS1 recognizes it during oxidative stress, which would explain why DpsA accumulates in Δ*clpS1* during oxidative stress but not during normal growth (**paper V**). It has previously been shown that the Δ*clpP* and Δ*clpX* strains in *Staphylococcus aureus* (*Staphylococcus*) are sensitive towards oxidative stress (Frees et al., 2003). That *Synechococcus* Δ*clpP1* is sensitive to H_2O_2 is not surprising given its susceptibility to other stresses such as high light intensity, UV-B irradiation and low temperature that all likely lead to the accumulation of reactive oxygen species (ROS) (Clarke et al., 1998; Porankiewicz et al., 1998). As discussed in a previous section, the reason for the different phenotypes between Δ*clpP1* and Δ*clpP2* is probably due to the ability of ClpP1 to form its own proteolytic core that can associate with ClpX and thus replace the activity of ClpP1/2. The fact that ClpX, ClpP1 and ClpP2 levels increase during the recovery phase after the H_2O_2 treatment suggests that the ClpXP1/P2 protease plays a role during this stage. It is also possible that KatG during both normal growth and oxidative stress is degraded by the ClpCP3/R protease in association to ClpS1 or ClpS2. The effect on *katG* expression in the Δ*clpS1* and Δ*clpP2* strain also suggests that the Clp protease influences the regulation of certain genes during normal growth.

Little is known about the protein regulation that underlies the oxidative stress responses in cyanobacteria. In *Bacillus subtilis*, the ClpXP protease has been shown to degrade the important oxidative regulator, Spx (Nanako et al., 2001, 2002, 2003). Spx controls the expression of several genes that are important for the response to oxidative stress, including the induction of thioredoxin (*trxA*) and thioredoxin reductase (*trxB*). The thioredoxin in turn regulates several other enzymes by reduction cycles, including peroxidases (e.g., 2-Cys prx) (reviewed by Apel et al., 2004). Although there is no obvious Spx ortholog in cyanobacteria, there is one for PerR that controls the expression of the *dps*, *katG* and *ahpCF* genes in several Gram-positive bacteria (Lee and Helmann, 2007). However, the PerR ortholog in *Synechocystis* represses the expression of *dpsA* and *ahpC*, but not that of *katG* (Li et al., 2004; Kobayashi et al., 2004). When the *perR* gene is inactivated, the resulting mutant is more resistant to H_2O_2 (Houot et al., 2007). It is possible therefore that the Clp protease degrades similar inducer/repressor proteins during both standard growth and recovery after oxidative stress. It might also be that the expression of *katG* is indirectly affected by the accumulation of reactive oxygen species (ROS). This is the case in both the Δ*psbU* strain in *Synechococcus* as well

as the *Synechocystis* Δ*psaE* strain. PsbU is a subunit of photosystem II and is situated in the cluster of Mn, Ca and Cl ions within the oxygen evolving complex (Shen et al., 1997, Otha et al., 1999, Inoue-Kashino et al 2005). The *Synechococcus* Δ*psbU* strain is less sensitive to H_2O_2 treatment because of increased catalase and peroxidase activity. It is thought that the loss of PsbU leads to increased generation of ROS, which in turn induces *katG* expression (Balint et al., 2006). A similar induction of *katG* is supposedly behind the increased resistance to H_2O_2 observed in the Δ*psaE* strain (Jeanjean et al., 2008). PsaE is a subunit of photosystem I and is involved in the binding of ferredoxin/flavodoxin (Weber and Strotmant, 1993; Rousseau et al., 1993; Xu et al., 1994; Barth et al., 1998; Meimberg et al., 1998). It is possible therefore that in Δ*clpS1* and Δ*clpP2* the level of ROS increases and leads to the induction of *katG* gene expression.

Another indirect effect that could influence *katG* expression is through the thioredoxin system. TrxB, as mention above, is a substrate for the Clp protease in *E. coli* and when *trxB* is inactivated the resulting mutant is more resistant to H_2O_2 during exponential growth as a result of increased *katG* expression (Takemoto et al., 1998). It has been shown that *Synechocystis* KatG interacts with both thioredoxin (Trx) (Pererz-Perez et al., 2006) and glutaredoxin (Grx) (Li et al., 2007), and thus any change to either the Trx or Grx system in *Synechococcus* could explain the phenotypes of all three *clp* mutants. Interestingly, when TrxA in *E. coli* is inactivated, the mutant is instead more sensitive to H_2O_2 (Takemoto et al., 1998). A preliminary experiment with *Synechococcus* Δ*clpP1* indicated that the levels of TrxA are lower than in the wild type under standard growth conditons. A recent study on the chloroplast ClpS in *Arabidopsis* has identified several putative substrates, including proteins involved in redox signaling that affect the levels of iron superoxide dismutase and thioredoxins, as well as enzymes involved in the tetrapyrrole biosynthetic pathway (Nishimura et al., 2013). The fact that KatG is a heme-catalase suggests its turnover might be affected if the tetrapyrrole pathway is compromised, which could be a contributing factor to the elevated KatG levels observed in the *Synechococcus* Δ*clpS1* strain.

The results in paper V suggest that Clp proteins are involved in the response of Synechococcus to oxidative stress in several ways. Firstly, the degradation of KatG during both standard growth and oxidative stress might be performed by the ClpCP3/R protease in association with ClpS1 and ClpS2. In addition, the ClpXP1/P2 protease appears to participate in the recovery phase following the initial oxidative damage. Clp proteins also appear to influence the regulation of katG gene expression, although their exact role remains unclear.

3. Future perspectives

In this thesis, many questions concerning the structure and function of different Clp proteins in *Synechococcus* have been answered. However, there remain many more questions left unanswered and much additional experimental work is still required to finally resolve the more important aspects of this important protease family in cyanobacteria. Below are the most crucial issues I believe need to be addressed in the future, and I have detailed how these could possibly be investigated.

What is the evolutionary advantage for the existence of ClpR in photosynthetic organisms?

In **paper I**, we have identified motifs important for the association between ClpC and ClpP3/R. In this interaction, both the ClpP3 and ClpR subunits are necessary for the association to ClpC, which indicates ClpR did not evolve solely to specify the binding to this chaperone partner, which is further confirmed by the fact ClpP1 can also associate to ClpC (**paper IV**). It was also clear from **paper I** that ClpR is necessary for the stable formation of the ClpP3/R core, but this again does not answer why the mixed proteolytic core has evolved. If we compare this situation with that of the 26S proteasome, we also see the development of inactive subunits incorporated into the proteolytic core of the eukaryotic enzyme, which differs from that in the proteasome from Archaebacteria. In fact, the β-rings of the eukaryotic 20S core have the same ratio of three proteolytically active subunits and four inactive ones (Myung et al., 2001; Gallastegui and Groll, 2010) to that in *Synechococcus* ClpP3/R. The persistence of ClpR in photosynthetic organisms suggests that any loss in Clp proteolytic activity that could result from incorporating an inactive subunit does not compromise the required rates of substrate degradation. In fact, it could be that ClpR confers a slower but more tightly regulated degradation activity that is crucial for certain key substrates. Evidence that ClpR regulates the activity of the Clp proteolytic core was shown in paper I, in which the addition of the MPIG motif from ClpP3 to the N-terminus of ClpR produced a core complex with considerably faster catalytic activity. This indicates that the corresponding region in the N-terminus ClpR somehow restricts the activity of the ClpP3/R core, possibly by regulating translocation of the substrate into the degradative chamber. Another indication that ClpR might play an essential regulatory role comes from when a re-activated ClpR along with ClpP3 was transformed into *Synechococcus* and proved lethal (unpublished result).

It is not only ClpP3/R that appears to have evolved a more tightly regulated proteolytic activity, but also the other main Clp proteolytic core in *Synechococcus*, ClpP1/P2. It is clear that the combination of ClpP1 and ClpP2 in the one proteolytic core slows down the degradation rate compared to that of the ClpP1 only core, or even that of the partly inactivated ClpP1/P2 cores (**paper IV**). The fact that the ClpP1 and ClpP2 subunits in ClpP1/P2 are arranged within each heptameric ring in an alternating

pattern, the same as the different subunits in the ClpP3/R core suggests that this configuration is important in conferring the regulated catalytic activity. It is interesting that in Gram-positive bacteria which have proteolytic cores of more than one type of ClpP, that each ClpP paralog forms a separate heptameric ring and not mixed rings as observed for the ClpP/R orthologs in photosynthetic organisms (Zelier et al., 2011, 2013, Raju et al., 2012, Akopian et al., 2012). As a consequence, it is likely that the degradation activity of these ClpP1/P2 proteolytic cores in the Gram-positive bacteria is not restricted by the incorporation of different subunits within the one heptameric ring. It is also possible that the inactive ClpR actually enhances the catalytic activity of ClpP3, as was observed for the partially inactivate ClpP1/P2 variants (**paper III**), while at the same time eliminating any potential limitation on ClpP3 activity, such as the Mg^{2+}-dependency of ClpP1.

In the case of ClpP3/R, the alternating arrangement of different subunits could also affect the structure of the substrate-binding groove within the proteolytic core, since ClpR has an insertion that protrudes into the groove according to 3D models (**paper I** and Andersson et al., 2009). This apparent obstruction could affect the kinds of substrates the ClpP3/R core can degrade and thus limits its specificity. However, this does not seem the case given that ClpP3/R can readily degrade α-casein, SsrA-tagged GFP as well as N-end rule substrates with the ClpS adaptors.

Ultimately, the questions involving the exact function of ClpR will probably not be resolved until a proteolytic core like *Synechococcus* ClpP3/R is crystallized and the atomic structure determined. It will only be then that an explanation might be found to how the alternating arrangement of ClpR and ClpP3 subunits within each heptameric ring affects the substrate-binding groove of the proteolytic core and its overall catalytic activity. Such information could also be helpful in explaining the obvious regulation of the ClpP1/P2 proteolytic core that occurs as a result of a similar alternating subunit configuration. Several attempts over the years have been made to crystallize the *Synechococcus* ClpP3/R complex but with no success to date. Efforts still continue to find the correct conditions for ClpP3/R crystallization and we have now included ClpP1/P2 into this project. In the future, attempts should also be made to crystallize the ClpP1 proteolytic core in the presence of both low and high concentrations of $MgCl_2$ to investigate how the catalytic site changes under these conditions.

What are the native substrates for the different Clp proteases?

One outstanding issue that must be addressed in the near future is the identity of the native protein substrates for both the ClpCP3/R and ClpXP1/P2 proteases. Several studies on Clp proteases from different organisms have successfully used different pull-down and trap techniques to identify putative native substrates (Flynn et al., 2003; Feng et al., 2013; Humbard et al., 2013; Nishimura et al., 2013). To identify substrates for ClpP3/R, we have pursued the trap approach by preparing a variant of the proteolytic core in which the ClpP3 subunit has been inactivated by site-directed mutagenesis (Andersson et al., 2009). Since ClpP3/R is essential for cell viability, we instead

transformed this inactive construct into a neutral site locus in the *Synechococcus* genome under the control of the *tac* promoter. Although the transformations were successful, problems occurred with the expression levels of the construct. When the *lacIq* repressor gene was included in the transformation, over-expression of the construct in *Synechococcus* by addition of IPTG was too low for sufficient amounts of the inactive ClpP3/R trap to be purified. Later when the *lacIq* gene was removed, no variable transformants were obtained presumably due to too high constitutive expression of the inactive trap interfering with normal ClpP3/R activity. Alternative promoters are now being considered to replace *tac*, which will hopefully provide the necessary controlled expression of the inactive ClpP3/R trap in *Synechococcus*.

An alternative approach to the *in situ* expression of the inactive ClpP3/R core could be to simply incubate soluble protein extracts from wild type *Synechococcus* with large amounts of recombinant ClpC and inactive ClpP3/R. Afterwards, the recombinant Clp proteins could be purified on Ni^{2+} affinity columns since they all contain His_6-tags, thereby removing the bulk of nonspecific *Synechococcus* soluble proteins. The eluted Clp proteins would then be denatured by urea treatment, disrupting the inactive ClpP3/R core and liberating the native substrates trapped inside. Passing the sample once more through the Ni^{2+} affinity columns would then separate the His-tagged Clp proteins from the native substrates, which would simply flow through. The identity of the native substrates could then be determined by mass spectrometry against the known *Synechococcus* proteome. Possible controls for this approach could be to perform the same steps with the active ClpP3/R core or to denature a proportion of the inactive ClpP3/R trap before the first Ni^{2+} affinity column. In either case, only those proteins found in the first trap isolations and not in the controls would be identified as native substrates, thereby minimizing the risk of non-specific proteins being incorrectly identified. If successful for ClpCP3/R, the same approach could be adapted for the ClpP1/P2 proteolytic core as long as the solubility issue with ClpX could be solved, although in this case overexpressing the inactive ClpP1/P2 trap in *Synechococcus* would likely be more successful.

For ClpS1 and ClpS2, it would also be feasible to use a pull-down approach similar to that done for the chloroplast ClpS from *Arabidopsis* (Nishimura et al., 2013). For the chloroplast ClpS, a fusion with GST was used since a simple His-tag did not bind the ClpS protein tightly enough to the affinity column. Although similar GST fusions could be made for the *Synechococcus* ClpS proteins, I would suggest using instead the maltose-binding protein (MBP). Our group has past experience over-expressing and purifying MBP fusion proteins, which are remain soluble in *E. coli* at very high concentrations. The purified MBP-ClpS1 or MBP-ClpS2 fusion could then be bound to an amylose column and soluble cell extracts from wild type *Synechococcus* washed over the column. Proteins bound to the MBP-ClpS fusion would then be eluted from the column using a buffer containing maltose and later identified by mass spectrometry. The control for this experiment would be to perform the same experiment but with just the MBP bound to the amylose column.

For each set of potential substrates for the different Clp proteases and adaptors, several could be confirmed *in vitro* by preparing recombinant versions and testing their degradation using our proteolytic assays. If suitable antibodies are available, certain substrates identified for the ClpXP1/P2 protease and the ClpS1 adaptor could also be confirmed by examining their level in the Δ*clpP1/2* and Δ*clpS1* strains, respectively. If the protein accumulated in the mutant relative to the wild type, then this would be consistent with its identification as a substrate.

What are the substrate specificities of ClpS1 and ClpS2?

Although we have begun to define the substrate specificities of *Synechococcus* ClpS1 and ClpS2 (**paper V**), additional work is needed to provide a comprehensive analysis. To identify more combinations of N-terminal sequences recognized by either ClpS1 or ClpS2, a library of synthetic peptides previously used to test the substrate specificity of EcClpS will soon be screened (Erbes et al., 2006). However, the number of peptide sequences in this library is limited and so other approaches should also be considered. If money was no object, custom-made peptide libraries could be commercially synthesized that include at least 10 000 sequences to cover all the possible amino acid combinations in the first three positions at the N-terminus. An alternative peptide library could be one covering the N-terminal sequence of all *Synechococcus* proteins (ca. 3 000), although this would not identify those substrates of ClpS1 or ClpS2 that first requires internal processing.

It would be interesting to explore the importance of the conserved amino acid differences between the cyanobacterial ClpS1 and ClpS2 proteins. Site-directed mutagenesis could be used to make changes to key amino acids in the *Synechococcus* ClpS proteins. For example, what would happen if the Phe40 and Thr62 in ClpS2 were changed to Met, which is conserved in ClpS1 orthologs as well as EcClpS? Would the modified ClpS2 then recognize N-end rule substrates normally recognized by only ClpS1 or EcClpS? Such modifications would begin to unravel the structural differences between ClpS1 and ClpS2 that underlies their difference in substrate specificity. In this regard, crystal structure of both *Synechococcus* ClpS1 and ClpS2 would provide valuable insights into the important amino acid variations between these two cyanobacterial ClpS adaptors, and there are plans to soon begin crystallization trial of the recombinant proteins.

Does Synechococcus ClpX function as EcClpX?

We have made numerous attempts to purify recombinant *Synechococcus* ClpX by over-expression in *E. coli* but failed each time to produce an active protein. Although using EcClpX has allowed us to study the activity of the ClpP1/P2 and ClpP1 proteolytic cores *in vitro*, there remains the distinct possibility that the characteristics of the *E. coli* ortholog do not fully represent those of the cyanobacterial ClpX. One possible solution would be to construct chimeric proteins in which key regions of the *Synechococcus* ClpX

are engineered into the more soluble EcClpX. It has been shown that the N-terminal domain in EcClpX is important for substrate association, while the C-terminal region including the P-loop is crucial for association to the proteolytic core (Singh et al., 2001; Wojtyra et al., 2003). Two chimerics could therefore be made, in which either the N- or C-terminal domain of EcClpX is replaced with the corresponding region from *Synechococcus* ClpX. The chimeric ClpX with the C-terminal change could then be used to confirm the core specificity for ClpP1/P2 and if the cyanobacterial ClpX actually associates to the ClpP1 only proteolytic core. The chimeric ClpX with the N-terminal change could be used test if ClpS2 is a substrate for ClpX in *Synechococcus*, as well as to identify other native substrates for the cyanobacterial ortholog. Although it is hoped that the solubility of these chimeric ClpX proteins is better than the native *Synechococcus* ClpX, this is certainly not guaranteed. As such, an alternative could be to over-express the ClpX protein from another cyanobacterium such as *Anabaena*, although the solubility of this ClpX protein is not guaranteed to be better than the *Synechococcus* ortholog. If we were sitting on a goldmine, we could also attempt to commercially synthesize the *Synechococcus* ClpX protein, although again this would not necessarily guarantee improved solubility.

Does the Clp protease degrade the phycobilisome?

We have preliminary evidence that the ClpCP3/R protease degrades the rod-linker proteins of the *Synechococcus* PBS *in vitro*, but it would be interesting to investigate this phenomenon in more detail. Firstly, a recombinant NblA would need to be purified without requiring high concentrations of imidazole to remain soluble. One possibility is to purify the NblA from another cyanobacterium such as *Anabaena* that is more soluble (Bienert et al., 2006). The new NblA could then be used to investigate its possible interaction with ClpC and how this affects PBS degradation *in vitro*. We also have evidence that *Synechococcus* ClpP1 could be involved in PBS degradation. When the Δ*clpP1* strain was starved for sulfur, the normally rapid degradation of the PBS was significantly slower, while crosslinking studies showed that both ClpC and ClpP1 are bound to wild type PBS during its rapid degradation (Barker-Åström et al., unpublished). It would therefore be interesting to further test the degradation of PBS *in vitro* using the recombinant ClpC and ClpP1 and in combination with ClpS1 and ClpS2.

Do Clp proteases play a role during oxidative stress?

In **paper V**, we demonstrated that certain *Synechococcus* Clp proteins are somehow involved in the response to oxidative stress, although some of the characteristics of this involvement could not be satisfactorily explained. More specifically, how basal KatG levels are influenced by different Clp proteins needs to be further explored. Additional assays should be performed to determine if the degradation of recombinant KatG by the ClpCP3/R protease can be optimized using different buffer conditions and pre-treatments. Why *katG* expression is induced in the Δ*clpP2* and Δ*clpS1* strains should

also be investigated further. The possibility that changes in the amount of enzymes such as thioredoxin and glutaredoxin are involved could be examined in the different mutants by immunoblotting. Protein profile comparisons between the wild type and *clp* mutants by techniques such as iTraq could also reveal other proteins affected by the various Clp proteins and how these might influence the response to oxidative stress. Similarly, the identification of native substrates for the different Clp proteases using the pull-down or trap approaches described above could also provide insights into the different phenotypes observed during oxidative stress.

4. References

Adam Z., Adamska I., Nakabayashi K., Ostersetzer O., Haussuhl K., Manuell A., Zheng B., Vallon O., Rodermel S.R., Shinozaki K., Clarke A.K.. (2001). Chloroplast and mitochondrial proteases in *Arabidopsis*. A proposed nomenclature. *Plant. Physiol*. **125**, 1912-8.

Akita M., Nielsen E., Keegstra K. (1997).Identification of protein transport complexes in the chloroplastic envelope membranes via chemical cross-linking. *J. Cell. Biol.* **36**, 983-94.

Akiyama Y., Yoshihisa T., Ito K. (1995). FtsH, a membrane-bound ATPase, forms a complex in the cytoplasmic membrane of *Escherichia coli*. *J. Biol. Chem*. **270**, 23485-90.

Akiyama Y., Kihara A., Tokuda H., Ito K. (1996). FtsH (HflB) is an ATP-dependent protease selectively acting on SecY and some other membrane proteins. *J. Biol. Chem.* **271**,31196-201.

Akiyama Y., Kihara A., Ito K. (1996). Subunit a of proton ATPase F0 sector is a substrate of the FtsH protease in *Escherichia coli*. *FEBS. Lett.* **399**, 26-8.

Akopian T., Kandror O., Raju R.M., Unnikrishnan M., Rubin E.J., Goldberg A.L. (2012). The active ClpP protease from M. tuberculosis is a complex composed of a heptameric ClpP1 and a ClpP2 ring. *EMBO. J.* **31**, 1529-41.

Almirón M., Link A.J., Furlong D., Kolter R. (1992). A novel DNA-binding protein with regulatory and protective roles in starved Escherichia coli. *Genes. Dev.* **6**, 2646-54

Altuvia S., Almirón M., Huisman G., Kolter R., Storz G. (1994). The dps promoter is activated by OxyR during growth and by IHF and sigma S in stationary phase. *Mol. Microbiol.* **13**, 265-72.

Amerik A.Y., Antonov V.K., Gorbalenya A.E., Kotova S.A., Rotanova T.V., Shimbarevich E.V.(1991). Site-directed mutagenesis of La protease. A catalytically active serine residue. FEBS. Lett. 287, 211-4.

Andersson F.I., Blakytny R., Kirstein J., Turgay K., Bukau B., Mogk A., Clarke A.K. (2006). Cyanobacterial ClpC/HSP100 protein displays intrinsic chaperone activity. *J Biol. Chem*. **281**, 5468-75.

Andersson F.I., Tryggvesson A., Sharon M., Diemand A.V., Classen M., Best C., Schmidt R., Schelin J., Stanne T.M., Bukau B., Robinson C.V., Witt S., Mogk A., Clarke A.K. (2009). Structure and function of a novel type of ATP-dependent Clp protease. *J. Biol. Chem*. **284**, 13519-32.

Atorino L., Silvestri L., Koppen M., Cassina L., Ballabio A., Marconi R., Langer T., Casari G. (2003). Loss of m-AAA protease in mitochondria causes complex I deficiency and increased sensitivity to oxidative stress in hereditary spastic paraplegia. *J. Cell. Biol.* **163**, 777-87.

Baker T.A., Sauer R.T. (2006). ATP-dependent proteases of bacteria: recognition logic and operating principles. *Trends. Biochem. Sci.* **31**, 647-53.

Baier A., Winkler W., Korte T., Lockau W., Karradt A. (2014). Degradation of Phycobilisomes in Synechocystis sp. PCC6803: evidence for essential formation of an NblA1/NblA2 hetetodimer and its coderadation bb a Clp protease complex. *J. Biol. Chem*. **289**, 11755-66.

Bailey S., Thompson E., Nixon P.J., Horton P., Mullineaux C.W., Robinson C., Mann N.H. (2002). A critical role for the Var2 FtsH homologue of *Arabidopsis thaliana* in the photosystem II repair cycle *in vivo*. *J. Biol. Chem.* **277**, 2006-11.

Balint I., Bhattacharya J., Perelman A., Schatz D., Moskovitz Y., Keren N., Schwarz R. (2006). Inactivation of the extrinsic subunit of photosystem II, PsbU, in *Synechococcus PCC 7942* results in elevated resistance to oxidative stress. FEBS. Lett. **580**, 2117-22.

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.

Barth P., Lagoutte B., Sétif P. (1998). Ferredoxin reduction by photosystem I from S*ynechocystis sp. PCC 6803*: toward an understanding of the respective roles of subunits PsaD and PsaE in ferredoxin binding. *Biochemistry.* **37**, 16233-41.

Barth P., Lagoutte B., Sétif P. (1998). Ferredoxin reduction by photosystem I from S*ynechocystis sp. PCC 6803*: toward an understanding of the respective roles of subunits PsaD and PsaE in ferredoxin binding. *Biochemistry.* **37**, 6233-41.

Baumeister W., Walz J., Zühl F., Seemüller E. (1998). The proteasome: paradigm of a self-compartmentalizing protease. *Cell*. **92**, 367-80.

Begg K.J, Tomoyasu t., Donachie W.D., Khatfar M., Niki H. Yamanaka K., Hiraga S., Ogura T. *Escherichia coli* Mutant Y16 Is a Double Mutant Carrying Thermosensitive ftsH and ftsI Mutations *J. Bacteriol.* **174**, 2416.

Bienert R., Baier K., Volkmer R., Lockau W., Heinemann U. (2006). Crystal structure of NblA from *Anabaena sp. PCC 7120*, a small protein playing a key role in phycobilisome degradation. *J. Biol. Chem.* **281**, 5216-23.

Benaroudj N., Goldberg A.L. (2000) PAN, the proteasome-activating nucleotidase from archaebacteria, is a protein-unfolding molecular chaperone. *Nat. Cell. Biol.* **2**, 833-9.

Bewley M.C., Graziano V., Griffin K., Flanagan J.M. (2006). The asymmetry in the mature amino-terminus of ClpP facilitates a local symmetry match in ClpAP and ClpXP complexes. *J. Struct. Biol*. **153**, 113-28.

Beynon, R. and Bond, J.S (2001). Proteolytic enzymes, second edition. *Oxford university press*.

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**, 3066-71.

Bieniossek C., Niederhauser B., Baumann U.M. (2009). The crystal structure of apo-FtsH reveals domain movements necessary for substrate unfolding and translocation Proc. *Natl. Acad. Sci. U. S. A*., **106**, 21579–21584

Bohon J., Jennings L.D., Phillips C.M., Licht S., Chance M.R. (2008). Synchrotron protein footprinting supports substrate translocation by ClpA via ATP-induced movements of the D2 loop. *Structure*. 16, 1157-65.

Bolon D.N., Wah D.A., Hersch G.L., Baker T.A., Sauer R.T. (2004). Bivalent tethering of SspB to ClpXP is required for efficient substrate delivery: a protein-design study. *Mol. Cell.* **13**, 443-9.

Bochtler M., Hartmann C., Song H.K., Bourenkov G.P., Bartunik H.D., Huber R. (2000). The structures of HsIU and the ATP-dependent protease HsIU-HsIV. *Nature.* **403**, 800-5.

Brötz-Oesterhelt H., Beyer D., Kroll H.P., Endermann R., Ladel C., Schroeder W., Hinzen B., Raddatz S., Paulsen H., Henninger K., Bandow J.E., Sahl H.G., Labischinski H. (2005). Dysregulation of bacterial proteolytic machinery by a new class of antibiotics. *Nat. Med*. **11**, 1082-7.

Burton B.M., Baker T.A. (2003). Mu transpososome architecture ensures that unfolding by ClpX or proteolysis by ClpXP remodels but does not destroy the complex. *Chem Biol*. **10**, 463-72.

Burton R.E., Baker T.A., Sauer R.T.(2005). Nucleotide-dependent substrate recognition by the AAA+ HslUV protease. *Nat. Struct. Mol. Biol.* **12**, 245-51.

Burton B.M., Baker T.A. (2005). Remodeling protein complexes: insights from the AAA+ unfoldase ClpX and Mu transposase. *Protein. Sci*. **14**, 1945-54.

Bukau B., Weissman J., Horwich A. (2006). Molecular chaperones and protein quality control. *Cell* **125**, 443-51.

Casari G., De Fusco M., Ciarmatori S., Zeviani M., Mora M., Fernandez P., De Michele G., Filla A., Cocozza S., Marconi R., Dürr A., Fontaine B., Ballabio A. (1998). Spastic paraplegia and OXPHOS impairment caused by mutations in paraplegin, a nuclear-encoded mitochondrial metalloprotease. *Cell.* **93,** 973-83.

Chatterjee I., Becker P., Grundmeier M., Bischoff M., Somerville G.A., Peters G., Sinha B., Harraghy N., Proctor R.A., Herrmann M. (2005). *Staphylococcus aureus* ClpC is required for stress resistance, aconitase activity, growth recovery, and death. *J. Bacteriol.* **187**, 4488-96.

Chastanet A., Prudhomme M., Claverys J.P., Msadek T. (2001). Regulation of *Streptococcus pneumoniae* clp genes and their role in competence development and stress survival. *J Bacteriol.* **183,** 7295-307.

Chiba S., Akiyama Y., Mori H., Matsuo E., Ito K. (2000) Length recognition at the N-terminal tail for the initiation of FtsH-mediated proteolysis. *EMBO*. **1**, 47-52.

Chiba S., Akiyama Y., Ito K. (2002). Membrane protein degradation by FtsH can be initiated from either end. *J Bacteriol.* **184**, 4775-82.

Chung C.H., Goldberg A.L. (1982). DNA stimulates ATP-dependent proteolysis and protein-dependent ATPase activity of protease La from *Escherichia coli*. *Proc. Natl. Acad .Sci. U. S. A.* **79**, 795-9.

Chuang S.E., Blattner F.R. (1993). Characterization of twenty-six new heat shock genes of *Escherichia coli. J. Bacteriol.* **175**, 5242-52.

Chuang S.E., Burland V., Plunkett G. 3rd., Daniels D.L., Blattner F.R. (1993). Sequence analysis of four new heat-shock genes constituting the hslTS/ibpAB and hslVU operons in *Escherichia coli*. *Gene*. **134**, 1-6.

Clarke AK., Eriksson M.J. (1996). The cyanobacterium *Synechococcus sp. PCC 7942* possesses a close homologue to the chloroplast ClpC protein of higher plants. *Plant. Mol. Biol.* **31**, 721-30.

Clarke A.K., Schelin J., Porankiewicz J. (1998). Inactivation of the *clpP1* gene for the proteolytic subunit of the ATP-dependent Clp protease in the cyanobacterium *Synechococcus* limits growth and light acclimation. *Plant. Mol. Biol.* **37**, 791-801.

Clarke A.K . (1999). ATP-dependent Clp Proteases in Photosynthetic Organisms— A Cut Above the Rest! Ann Bot 83, 593-599.

Clarke A K., MacDonald T.M., Sjögren L.L.E. (2005). The ATP-dependent Clp protease in chloroplasts of higher plants. *Physiologia Plantarum* **123**, 406–412.

Collier J.L., Grossman A.R. (1994). A small polypeptide triggers complete degradation of light-harvesting phycobiliproteins in nutrient-deprived cyanobacteria. *EMBO. J.* **13**, 1039-47.

Cordell S.C., Robinson E.J., Lowe J. (2013). Activated ClpP kills persisters and eradicates a chronic biofilm infection. *Nature.* **503**, 365-70.

Cordell S.C., Robinson E.J., Lowe J. (2003). Crystal structure of the SOS cell division inhibitor SulA and in complex with FtsZ. Proc. *Natl. Acad. Sci. U. S. A.* **100,** 7889-94.

Couvreur B., Wattiez R., Bollen A., Falmagne P., Le Ray D., Dujardin J.C.. Eubacterial HslV and HslU subunits homologs in primordial eukaryotes. *Mol Biol Evol.* **19**, 2110-7.

De Donatis G.M., Singh S.K., Viswanathan S., Maurizi M.R. (2010). A single ClpS monomer is sufficient to direct the activity of the ClpA hexamer. *J. Biol. Chem*. **285**, 8771-81.

Derré I., Rapoport G., Devine K., Rose M., Msadek T. (1999). ClpE, a novel type of HSP100 ATPase, is part of the CtsR heat shock regulon of Bacillus subtilis. *Mol. Microbiol.* **32**, 581-93.

Derré I., Rapoport G., Msadek T. (2000). The CtsR regulator of stress response is active as a dimer and specifically degraded in vivo at 37 degrees C. *Mol. Microbiol.* **38**, 335-47.

Derrien B., Majeran W., Effantin G., Ebenezer J., Friso G., van Wijk K.J., Steven A.C., Maurizi M.R., Vallon O. (2012). The purification of the *Chlamydomonas reinhardtii* chloroplast ClpP complex: additional subunits and structural features. Plant. Mol. Biol. 80, 189-202.

Dines M., Sendersky E., David L., Schwarz R., Adir N. (2008). Structural, functional, and mutational analysis of the NblA protein provides insight into possible modes of interaction with the phycobilisome. *J. Biol. Chem.* **283**, 30330-40.

Domian I.J., Quon K.C., Shapiro L. (1997). Cell type-specific phosphorylation and proteolysis of a transcriptional regulator controls the G1-to-S transition in a bacterial cell cycle. *Cell*. **90**, 415-24.

Donaldson L.W., Wojtyra U., Houry W.A. (2003). Solution structure of the dimeric zinc binding domain of the chaperone ClpX. J. Biol. Chem. 278, 48991-6.

Dougan D.A., Reid B.G., Horwich A.L., Bukau B. (2002). ClpS, a substrate modulator of the ClpAP machine. *Mol. Cell.* **9**, 673-83.

El Bakkouri M., Rathore S., Calmettes C., Wernimont A.K., Liu K., Sinha D., Asad M., Jung P., Hui R., Mohmmed A., Houry W.A.. Structural insights into the inactive subunit of the apicoplast-localized caseinolytic protease complex of *Plasmodium falciparum*. *J. Biol. Chem.* **288,** 1022-31.

El Bakkouri M., Pow A., Mulichak A., Cheung K.L., Artz J.D., Amani M., Fell S., de Koning-Ward T.F., Goodman C.D., McFadden G.I., Ortega J., Hui R., Houry W.A. (2010). The Clp chaperones and proteases of the human malaria parasite *Plasmodium falciparum*. *J. Mol. Biol*. **404**, 456-77.

Engels S., Schweitzer J.E., 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. *Mol. Microbiol.* **52,** 285-302.

Erbse A., Schmidt R., Bornemann T., Schneider-Mergener J., Mogk A., Zahn R., Dougan D.A., Bukau B. (2006). ClpS is an essential component of the N-end rule pathway in *Escherichia coli. Nature*. **439**, 753-6.

Estorninho M., Smith H., Thole J., Harders-Westerveen J., Kierzek A., Butler R.E., Neyrolles O., Stewart G.R. (2010). ClgR regulation of chaperone and protease systems is essential for *Mycobacterium tuberculosis* parasitism of the macrophage. *Microbiology.* **156**, 3445-55.

Farbman M.E., Gershenson A., Licht S. (2008). Role of a conserved pore residue in the formation of a prehydrolytic high substrate affinity state in the AAA+ chaperone ClpA. *Biochemistry.* **47,** 13497-505.

Farrell CM1, Baker TA, Sauer RT. Altered specificity of a AAA+ protease. *Mol .Cell.* **25**, 161-6.

Feng J., Michalik S., Varming A.N., Andersen J.H., 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**, 547-58.

Ferro M., Brugière S., Salvi D., Seigneurin-Berny D., Court M., Moyet L., Ramus C., Miras S., Mellal M., Le Gall S., Kieffer-Jaquinod S., Bruley C., Garin J., Joyard J., Masselon C.., Rolland N. (2010). AT_CHLORO, a comprehensive chloroplast proteome database with subplastidial localization and curated information on envelope proteins. *Mol. Cell. Proteomics.* **9**, 1063-84.

Flores-Pérez Ú., Jarvis P. (2013). Molecular chaperone involvement in chloroplast protein import. *Biochim. Biophys. Acta.* **1833**, 332-40.

Flynn J.M., Levchenko I., Seidel M., Wickner S.H., Sauer R.T., Baker T.A. (2001). Overlapping recognition determinants within the ssrA degradation tag allow modulation of proteolysis. *Proc. Natl. Acad. Sci. U S A.* **98**, 10584-9.

Flynn J.M., Neher S.B., Kim Y.I., Sauer R.T., Baker T.A. (2003). Proteomic discovery of cellular substrates of the ClpXP protease reveals five classes of ClpX-recognition signals. *Mol. Cell.* **11**, 671-83.

Frees D., Qazi S.N., Hill P.J., Ingmer H. (2003). Alternative roles of ClpX and ClpP in Staphylococcus aureus stress tolerance and virulence. *Mol. Microbiol.* **48**, 1565-78.

Frees D., Ingmer H. (1999). ClpP participates in the degradation of misfolded protein in *Lactococcus lactis. Mol. Microbiol.* **31**, 79-87.

Frees D., Chastanet A., Qazi S., Sørensen K., Hill P., Msadek T., Ingmer H. (2004). Clp ATPases are required for stress tolerance, intracellular replication and biofilm formation in *Staphylococcus aureus*. *Mol. Microbiol*. **54**, 1445-62.

Frees D., Savijoki K., Varmanen P., Ingmer H. (2007). Clp ATPases and ClpP proteolytic complexes regulate vital biological processes in low GC, Gram-positive bacteria. *Mol. Microbiol.* **63** 1285-95.

Fu G.K., Smith M.J., Markovitz D.M. Bacterial protease Lon is a site-specific DNA-binding protein. *J. Biol. Chem.* **272**, 534-8.

Gallastegui N., Groll M. (2010). The 26S proteasome: assembly and function of a destructive machine. *Trends. Biochem. Sci.* 35, 634-42.

Gaillot O., Pellegrini E., Bregenholt S., Nair S., Berche P. (2000). The ClpP serine protease is essential for the intracellular parasitism and virulence of *Listeria monocytogenes*. *Mol. Microbiol*. 35, 1286-94.

García-Lorenzo M., Sjödin A., Jansson S., Funk C. (2006). Protease gene families in *Populus* and *Arabidopsis*. *BMC. Plant.* **6**:30.

Gerth U., Krüger E., Derré I., Msadek T., Hecker M. (1998). Stress induction of the *Bacillus subtilis clpP* gene encoding a homologue of the proteolytic component of the Clp protease and the involvement of ClpP and ClpX in stress tolerance. *Mol. Microbiol.* **28**, 787-802.

Gerth U., Kirstein J., Mostertz J., Waldminghaus T., Miethke M., Kock H., Hecker M. (2004). Fine-tuning in regulation of Clp protein content in *Bacillus subtilis*. *J. Bacteriol.* **186,** 179-91.

Glickman M.H., Rubin D.M., Fried V.A., Finley D. (1998). The regulatory particle of the *Saccharomyces cerevisiae* proteasome. *Mol. Cell. Biol.* **18**, 149-62.

Glynn S.E., Martin A., Nager A.R., Baker T.A., Sauer R.T. (2009). Structures of asymmetric ClpX hexamers reveal nucleotide-dependent motions in a AAA+ protein-unfolding machine. *Cell.* **139**, 744-56.

Gonciarz-Swiatek M., Wawrzynow A., Um S.J., Learn B.A., McMacken R., Kelley W.L., Georgopoulos C., Sliekers O., Zylicz M. (1999). Recognition, targeting, and hydrolysis of the lambda O replication protein by the ClpP/ClpX protease. *J. Biol. Chem.* **274**, 13999-4005.

Gottesman S., Clark W.P., Maurizi M.R. (1990). The ATP-dependent Clp protease of *Escherichia coli*. Sequence of *clpA* and identification of a Clp-specific substrate. *J. Biol. Chem.* **265**, 7886-93.

Gottesman S., Clark W.P., de Crecy-Lagard V., Maurizi M.R. (1993). ClpX, an alternative subunit for the ATP-dependent Clp protease of Escherichia coli. Sequence and in vivo activities. *J. Biol. Chem.* 268, 22618-26.

Gottesman S., Wickner S., Maurizi M.R. (1997). Protein quality control: triage by chaperones and proteases. *Genes .Dev.* **11**,815-23.

Gottesman S., Roche E., Zhou Y., Sauer R.T. (1998). The ClpXP and ClpAP proteases degrade proteins with carboxy-terminal peptide tails added by the SsrA-tagging system. *Genes. Dev.* **12**, 13, 38-47.

Gottesman S. (2003). Proteolysis in bacterial regulatory circuits. *Annu. Rev. Cell. Dev. Biol.* **19**, 565-87.

Gribun A., Kimber M.S., Ching R., Sprangers R., Fiebig K.M., Houry W.A. (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. *J. Biol. Chem.* **280**, 16185-96.

Grimaud, R., Kessel, M., Beuron, F., Steven, A.C., and Maurizi M.R. (1998). Enzymatic and structural similarities between the Escherichia coli ATP-dependent proteases, ClpXP and ClpAP. *J. Biol. Chem.* **273**, 12476-12481.

Groll M., Ditzel L., Löwe J., Stock D., Bochtler M., Bartunik H.D., Huber R. (1997). Structure of 20S proteasome from yeast at 2.4 A resolution. *Nature.* **386**, 463-71.

Groll M., Bajorek M., Köhler A., Moroder L., Rubin D.M., Huber R., Glickman M.H., Finley D. (2000). A gated channel into the proteasome core particle. *Nat. Struct. Biol*. **7,** 1062-7.

Guo F., Esser L., Singh S.K., Maurizi M.R., Xia D. (2002). Crystal structure of the heterodimeric complex of the adaptor, ClpS, with the N-domain of the AAA+ chaperone, ClpA. *J. Biol. Chem*. **277**, 46753-62.

Gur E., Sauer R.T. (2008). Recognition of misfolded proteins by Lon, a AAA(+) protease. *Genes. Dev*. **22**, 2267-77.

Herman C., Thevenet D., D'Ari, R., Bouloc P. (1995). Degradation of s32, the heat shock regulator in Escherichia coli, is governed by HflB. *Proc. Natl. Acad. Sci. USA*. **92**,3516–3520.

Herman C., Thévenet D., Bouloc P., Walker G.C., D'Ari R. (1998). Degradation of carboxy-terminal-tagged cytoplasmic proteins by the *Escherichia coli* protease HflB (FtsH). *Genes. Dev.* **12**, 1348-55.

Herman C., Prakash S., Lu C.Z., Matouschek A., Gross C.A. (2003) Lack of a robust unfoldase activity confers a unique level of substrate specificity to the universal AAA protease FtsH. *Mol. Cell.* **11**, 659-69.

Hinnerwisch J., Reid B.G., Fenton W.A., Horwich A.L. (2005). Roles of the N-domains of the ClpA unfoldase in binding substrate proteins and in stable complex formation with the ClpP protease. *J. Biol. Chem*. **280**, 40838-44.

Hou J.Y., Sauer R.T., Baker T.A. (2008). Distinct structural elements of the adaptor ClpS are required for regulating degradation by ClpAP. *Nat. Struct. Mol. Biol*. **15**, 288-94.

Houot L., Floutier M., Marteyn B., Michaut M., Picciocchi A., Legrain P., Aude J.C., Cassier-Chauvat C., Chauvat F. (2007). Cadmium triggers an integrated reprogramming of the metabolism of *Synechocystis PCC6803*, under the control of the Slr1738 regulator. *BMC. Genomics*. **8**, 350.

Huang H., Goldberg A.L. (1997). Proteolytic activity of the ATP-dependent protease HslVU can be uncoupled from ATP hydrolysis. *J. Biol. Chem.* **272**, 21364-72.

Humbard M.A., Miranda H.V., Lim J.M., Krause D.J., Pritz J.R., Zhou G., Chen S., Wells L., Maupin-Furlow J.A. (2010). Ubiquitin-like small archaeal modifier proteins (SAMPs) in *Haloferax volcanii*. Nature. 463, 54-60..

Humbard M.A., Surkov S., De Donatis G.M., Jenkins L.M., Maurizi M.R. (2013). The N-degradome of Escherichia coli: limited proteolysis in vivo generates a large pool of proteins bearing N-degrons. *J. Biol. Chem*. **288**, 28913-24.

Hwang B.J., Woo K.M., Goldberg A.L., Chung C.H. (1988). Protease Ti, a new ATP-dependent protease in *Escherichia coli*, contains protein-activated ATPase and proteolytic functions in distinct subunits. J. Biol. Chem. 263, 8727-34.

Ingmer H., Vogensen F.K., Hammer K., Kilstrup M. (1999). Disruption and analysis of the *clpB, clpC*, and *clpE* genes in *Lactococcus lactis*: ClpE, a new Clp family in gram-positive bacteria. *J. Bacteriol.* 181, 2075-83.

Inoue-Kashino N., Kashino Y., Satoh K., Terashima I., Pakrasi H.B.(2005). PsbU provides a stable architecture for the oxygen-evolving system in cyanobacterial photosystem II. *Biochemistry.* **44**, 12214-28.

Jackson-Constan D., Akita M., Keegstra K. (2001). Molecular chaperones involved in chloroplast protein import. Biochim. Biophys. Acta. 1541,

Janska H. (2005). ATP-dependent proteases in plant mitochondria: What do we know about them today? *Physiologia. Plantarum.* **123**, 399–405.

Jayasekera M.M., Foltin S.K., Olson E.R., Holler T.P. (2000). *Escherichia coli* requires the protease activity of FtsH for growth. *Arch. Biochem. Biophys*. **380,** 103-7.

Jeanjean R., Latifi A., Matthijs H.C., Havaux M. (2007). The PsaE subunit of photosystem I prevents light-induced formation of reduced oxygen species in the cyanobacterium *Synechocystis sp. PCC 6803. Biochim. Biophys. Acta.* **1777**, 308-16.

Jennings L.D., Bohon J., Chance M.R., Licht S (2008). The ClpP N-terminus coordinates substrate access with protease active site reactivity. *Biochemistry.* **47**, 11031-40.

Joshi S.A., Hersch G.L., Baker T.A., Sauer R.T. (2004) Communication between ClpX and ClpP during substrate processing and degradation. *Nat. Struct. Mol. Biol.* **11**, 404-11.

Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y., Miyajima N., Hirosawa M., Sugiura M., Sasamoto S., Kimura T., Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S., Shimpo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M., Tabata S. (1996). Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis sp. strain PCC6803*.II. Sequence determination of the entire genome and assignment of potential protein-coding regions. *DNA Res*. **3**, 109-36.

Kanemori, M., Nishihara, K., Yanagi, H., Yura, T. (1997). Synergistic roles of HslVU and other ATP-dependent proteases in the controlling in vivo proteins in Escherichia coli turnover of s32 and abnormal *J. Bacteriol.* **179**, 7219–7225.

Kanesaki Y., Yamamoto H., Paithoonrangsarid K., Shoumskaya M., Suzuki I., Hayashi H., Murata N. Histidine kinases play important roles in the perception and signal transduction of hydrogen peroxide in the cyanobacterium, *Synechocystis sp. PCC 6803*. *Plant. J.***49**, 313-24. .

Kang S.G., Maurizi M.R., Thompson M., Mueser T., Ahvazi B. (2004). Crystallography and mutagenesis point to an essential role for the N-terminus of human mitochondrial ClpP. *J. Struct. Biol.* **148**, 338-52

Katayama Y., Gottesman S., Pumphrey J., Rudikoff S., Clark W.P., Maurizi M.R. (1988). The two-component, ATP-dependent Clp protease of *Escherichia coli*. Purification, cloning, and mutational analysis of the ATP-binding component. *J. Biol. Chem.* **263**, 15226-36.

Kato Y., Miura E., Ido K., Ifuku K., Sakamoto W. (2009). The variegated mutants lacking chloroplastic FtsHs are defective in D1 degradation and accumulate reactive oxygen species. *Plant. Physiol*. **151**, 1790-801.

Karradt A., Sobanski J., Mattow J., Lockau W., Baier K. (2008). NblA, a key protein of phycobilisome degradation, interacts with ClpC, a HSP100 chaperone partner of a cyanobacterial Clp protease. *J. Biol. Chem.* **283**, 32394-403.

Ke S., Haselkorn R. (2012) Fluorescence spectroscopy study of heterocyst differentiation in *Anabaena PCC 7120* filaments. *Microbiology.* **159**, 253-8.

Keiler K.C., Waller P.R., Sauer R.T. (1996). Role of a peptide tagging system in degradation of proteins synthesized from damaged messenger RNA. *Science*. **271**, 990-3.

Kessel M., Maurizi M.R., Kim B., Kocsis E., Trus B.L., Singh S.K., Steven A.C. (1995) Homology in structural organization between E. coli ClpAP protease and the eukaryotic 26 S proteasome. *J. Mol. Biol.* **250**, 587-94.

Kim Y.I., Burton R.E., Burton B.M., Sauer R.T., Baker T.A. (2000). Dynamics of substrate denaturation and translocation by the ClpXP degradation machine. *Mol. Cell.* **5**, 639-48.

Kirstein J., Zühlke D., Gerth U., Turgay K., Hecker M. (2005). A tyrosine kinase and its activator control the activity of the CtsR heat shock repressor in *B. subtilis*. *EMBO. J*. **24**, 3435-45.

Kirstein J., Schlothauer T., Dougan D.A., Lilie H., Tischendorf G., Mogk A., Bukau B., Turgay K. (2006) Adaptor protein controlled oligomerization activates the AAA+ protein ClpC. *EMBO J.* **25**, 1481-91.

Kirstein J., Hoffmann A., Lilie H., Schmidt R., Rübsamen-Waigmann H., Brötz-Oesterhelt H., Mogk A., Turgay K. The antibiotic ADEP reprogrammes ClpP, switching it from a regulated to an uncontrolled protease. *EMBO Mol. Med.* **1**, 37-49.

Kim Y.I., Levchenko I., Fraczkowska K., Woodruff R.V., Sauer R.T., Baker T.A. (2001) Molecular determinants of complex formation between Clp/Hsp100 ATPases and the ClpP peptidase. *Nat.Struct. Biol*. **8**, 230-3.

Kim D.Y., Kim K.K. (2008). The structural basis for the activation and peptide recognition of bacterial ClpP. *J. Mol. Biol.* **379**, 760-71.

Kim Y.C., DeMartino G.N. (2011). C termini of proteasomal ATPases play nonequivalent roles in cellular assembly of mammalian 26 S proteasome. *J. Biol. Chem.* **286**, 26652-66.

Kock H., Gerth U., Hecker M. (2004) The ClpP peptidase is the major determinant of bulk protein turnover in *Bacillus subtilis*. *J. Bacteriol.* **186,** 5856-64.

Kolygo K., Ranjan N., Kress W., Striebel F., Hollenstein K., Neelsen K., Steiner M., Summer H., Weber-Ban E. (2009). Studying chaperone-proteases using a real-time approach based on FRET. *J. Struct. Biol.* **168**, 267-77.

Koppen M., Metodiev M.D., Casari G., Rugarli E.I., Langer T. (2006). Variable and tissue-specific subunit composition of mitochondrial m-AAA protease complexes linked to hereditary spastic paraplegia. *Mol. Cell. Biol.* **27**, 758-67.

Kwon A.R., Kessler B.M., Overkleeft H.S., McKay D.B. (2003). Structure and reactivity of an asymmetric complex between HslV and I-domain deleted HslU, a prokaryotic homolog of the eukaryotic proteasome. *J. Mol. Biol* **330**, 185-95.

Langklotz S., Schäkermann M., Narberhaus F. (2010) Control of lipopolysaccharide biosynthesis by FtsH-mediated proteolysis of LpxC is conserved in enterobacteria but not in all gram-negative bacteria. *J. Bacteriol.* **193**, 1090-7.

Lemos J.A., Burne R.A. (2002). Regulation and Physiological Significance of ClpC and ClpP in *Streptococcus mutans*. *J. Bacteriol.* **184**, 6357-66.

Leonhard K., Herrmann J.M., Stuart R.A., Mannhaupt G., Neupert W., Langer T. (1996) AAA proteases with catalytic sites on opposite membrane surfaces comprise a proteolytic system for the ATP-dependent degradation of inner membrane proteins in mitochondria. *EMBO. J.* **15**, 4218-29.

Lee M.E., Baker T.A., Sauer R.T. (2010) Control of substrate gating and translocation into ClpP by channel residues and ClpX binding. *J. Mol. Biol.* **399**, 707-18.

Lee S., Augustin S., Tatsuta T., Gerdes F., Langer T., Tsai F.T. (2011). Electron cryomicroscopy structure of a membrane-anchored mitochondrial AAA protease. *J. Biol. Chem*. **286**, 4404-11.

Levchenko I., Luo L., Baker T.A. (1995). Disassembly of the Mu transposase tetramer by the ClpX chaperone. *Genes. Dev.* **9,** 2399-408.

Levchenko I., Seidel M., Sauer R.T., Baker T.A. (2000). A specificity-enhancing factor for the ClpXP degradation machine. *Science*. **289**, 2354-6.

Li H., Singh A.K., McIntyre L.M., Sherman L.A. (2004). Differential gene expression in response to hydrogen peroxide and the putative PerR regulon of *Synechocystis sp. strain PCC 6803*. *J. Bacteriol*. **186,** 3331-45.

Li D.H., Chung Y.S., Gloyd M., Joseph E., Ghirlando R., Wright G.D., Cheng Y.Q., Maurizi M.R., Guarné 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. *Chem. Biol.* **17**, 959-69.

Liao J.H., Lin Y.C., Hsu J., Lee A.Y., Chen T.A., Hsu C.H., Chir J.L., Hua K.F., Wu T.H., Hong L.J., Yen P.W., Chiou A., Wu S.H. (2010). Binding and cleavage of *E. coli* HUbeta by the E. coli Lon protease. *Biophys. J.* **98**, 129-37.

Lies M., Maurizi M.R. (2008). Turnover of endogenous SsrA-tagged proteins mediated by ATP-dependent proteases in Escherichia coli. *J. Biol. Chem.* **283,** 22918-29.

Lindahl M., Spetea C., Hundal T., Oppenheim A.B., Adam Z., Andersson B (2000). The thylakoid FtsH protease plays a role in the light-induced turnover of the photosystem II D1 protein. *Plant. Cell*. **12,** 419-31.

Lomovskaya O.L., Kidwell J.P., Matin A. (1994). Characterization of the sigma 38-dependent expression of a core *Escherichia coli* starvation gene, *pexB*. *J. Bacteriol.* **176**, 3928-35.

Lowy F.D. (1998). *Staphylococcus aureus* infections. *N. Engl. J. Med*. **339**, 520-32.

Lupas A.N., Koretke K.K. (2003). Bioinformatic analysis of ClpS, a protein module involved in prokaryotic and eukaryotic protein degradation. *J. Struct. Biol.* **141**, 77-83.

Löwe J., Stock D., Jap B., Zwickl P., Baumeister W., Huber R. (1995). Crystal structure of the 20S proteasome from the archaeon *T. acidophilum* at 3.4 A resolution. *Science*. **268**, 533-9.

Makino S., Makino T., Abe K., Hashimoto J., Tatsuta T., Kitagawa M., Mori H., Ogura T., Fujii T., Fushinobu S., Wakagi T., Matsuzawa H. (1999). Second transmembrane segment of FtsH plays a role in its proteolytic activity and homo-oligomerization. *FEBS. Lett.* **460**, 554-8.

Martin A., Baker T.A., Sauer R.T. (2005). Rebuilt AAA + motors reveal operating principles for ATP-fuelled machines. *Nature.* **437**, 1115-20.

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

Martin A., Baker T.A., Sauer R.T. (2008). Pore loops of the AAA+ ClpX machine grip substrates to drive translocation and unfolding. *Nat. Struct. Mol. Biol.* **15**, 1147-51.

Martin A., Baker T.A., Sauer R.T. (2008). Diverse pore loops of the AAA+ ClpX machine mediate unassisted and adaptor-dependent recognition of ssrA-tagged substrates. *Mol. Cell.* **29**,441-50.

Martinez A., Kolter R. (1997). Protection of DNA during oxidative stress by the nonspecific DNA-binding protein Dps. *J. Bacteriol*. **179**, 5188-94.

Marques A.J., Palanimurugan R., Matias A.C., Ramos P.C., Dohmen R.J. (2009). Catalytic mechanism and assembly of the proteasome. *Chem. Rev.* **109**, 1509-36.

Maurizi M.R., Clark W.P., Kim S.H., Gottesman S. (1990). Clp P represents a unique family of serine proteases. *J. Biol. Chem*. ***265*,** 12546-52.

Maurizi M.R., Clark W.P., 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*. *J. Biol. Chem*. **265**,12536-45.

Mazzucotelli E., Belloni S., Marone D., De Leonardis A., Guerra D., Di Fonzo N., Cattivelli L., Mastrangelo A. (2006). The e3 ubiquitin ligase gene family in plants: regulation by degradation. *Curr. Genomics*. **7**,509-22.

Maupin-Furlow J.A., Humbard M.A., Kirkland P.A., Li W., Reuter C.J., Wright A.J., Zhou G. (2006). Proteasomes from structure to function: perspectives from Archaea. *Curr. Top. Dev Biol.* **75**, 125-69.

Majeran W., Friso G., van Wijk K.J., Vallon O. (2005). The chloroplast ClpP complex in *Chlamydomonas reinhardtii* contains an unusual high molecular mass subunit with a large apical domain. *FEBS. J*. **272**, 5558-71.

Mei J.M., Nourbakhsh F., Ford C.W., Holden D.W. (1997). Identification of *Staphylococcus aureus* virulence genes in a murine model of bacteraemia using signature-tagged mutagenesis. *Mol. Microbiol*. **26**, 399-407.

Mei Z., Wang F., Qi Y., Zhou Z., Hu Q., Li H., Wu J., Shi Y. (2009). Molecular determinants of MecA as a degradation tag for the ClpCP protease. J. Biol. Chem. **284**, 34366-75.

Meimberg K., Lagoutte B., Bottin H., Mühlenhoff U. (1998). The PsaE subunit is required for complex formation between photosystem I and flavodoxin from the cyanobacterium *Synechocystis sp. PCC 680*3. *Biochemistry.* **37**, 9759-67.

Mhammedi-Alaoui A., Pato M., Gama M.J., Toussaint A. (1994). A new component of bacteriophage Mu replicative transposition machinery: the *Escherichia coli* ClpX protein. *Mol. Microbiol* **11**, 1109-16.

Michel A., Agerer F., Hauck C.R., Herrmann M., Ullrich J., Hacker J., Ohlsen K. (2006). Global regulatory impact of ClpP protease of *Staphylococcus aureus* on regulons involved in virulence, oxidative stress response, autolysis, and DNA repair. *J. Bacteriol.* **188**, 5783-96.

Miethke M., Hecker M., Gerth U. (2006). Involvement of *Bacillus subtilis* ClpE in CtsR degradation and protein quality control. *J. Bacteriol*. **188**, 4610-9.

Mogk A., Dougan D., Weibezahn J., Schlieker C., Turgay K., Bukau B. (2004). Broad yet high substrate specificity: the challenge of AAA+ proteins. *J. Struct. Biol*. **146**, 90-8.

Msadek T., Dartois V., Kunst F., Herbaud M.L., Denizot F., Rapoport G. (1998). ClpP of *Bacillus subtilis* is required for competence development, motility, degradative enzyme synthesis, growth at high temperature and sporulation. *Mol Microbiol.* **27**, 899-914.

Muttucumaru D.G., Roberts G., Hinds J., Stabler R.A, Parish T. (2004).Gene expression profile of *Mycobacterium tuberculosis* in a non-replicating state. *Tuberculosis.* (Edinb) **84**, 239-46.

Myung J., Kim K.B., Crews C.M. (2001). The ubiquitin-proteasome pathway and proteasome inhibitors. *Med. Res. Rev.* **21,** 245-73.

Nakabayashi K., Ito M., Kiyosue T., Shinozaki K., Watanabe A. (1999).Identification of clp genes expressed in senescing *Arabidopsis* leaves. *Plant. Cell. Physiol.* **40**, 504-14.

Nair S., Frehel C., Nguyen L., Escuyer V., Berche P. (1999). ClpE, a novel member of the HSP100 family, is involved in cell division and virulence of *Listeria monocytogenes*. *Mol. Microbiol.* **3**, 185-96.

Nair S., Poyart C., Beretti J.L., Veiga-Fernandes H., Berche P., Trieu-Cuot P. (2003). Role of the *Streptococcus agalactiae* ClpP serine protease in heat-induced stress defence and growth arrest. *Microbiology.* **149**, 407-17.

Nagashima K., Kubota Y., Shibata T., Sakaguchi C., Shinagawa H., Hishida T. (2006). Degradation of *Escherichia coli* RecN aggregates by ClpXP protease and its implications for DNA damage tolerance. *J. Biol. Chem.* **13**, 281, 30941-6.

Nakano M.M., Hajarizadeh F., Zhu Y., Zuber P. (2001). Loss-of-function mutations in yjbD result in ClpX- and ClpP-independent competence development of *Bacillus subtilis*. *Mol. Microbiol*. **42**, 383-94.

Nakano S., Zheng G., Nakano M.M., Zuber P. (2001). Multiple pathways of Spx (YjbD) proteolysis in *Bacillus subtilis*. *J. Bacteriol*. **184**, 3664-70.

Nakano S., Nakano M.M., Zhang Y., Leelakriangsak M., Zuber P. (2003). A regulatory protein that interferes with activator-stimulated transcription in bacteria. *Proc. Natl. Acad. Sci. U. S. A.* **100**,4233-8.

Neher S.B., Flynn J.M., Sauer R.T., Baker T.A. (2003). Latent ClpX-recognition signals ensure LexA destruction after DNA damage. *Genes .Dev.* **17**, 1084-9.

Neher S.B., Villén J., Oakes E.C., Bakalarski C.E., Sauer R.T., Gygi S.P., Baker T.A. (2006). Proteomic profiling of ClpXP substrates after DNA damage reveals extensive instability within SOS regulon. *Mol. Cell.* **22**, 193-204.

Neuwald A.F., Aravind L., Spouge J.L., Koonin E.V. (1999). AAA+: A class of chaperone-like ATPases associated with the assembly, operation, and disassembly of protein complexes. *Genome. Res* **9**, 27-43.

Nielsen E., Akita M., Davila-Aponte J., Keegstra K. (1997).Stable association of chloroplastic precursors with protein translocation complexes that contain proteins from both envelope membranes and a stromal Hsp100 molecular chaperone. *EMBO. J*. **16**, 935-46.

Ninnis R.L., Spall S.K., Talbo G.H., Truscott K.N., Dougan D.A. (2009). Modification of PATase by L/F-transferase generates a ClpS-dependent N-end rule substrate in *Escherichia coli.* *EMBO. J* **28**, 1732-44.

Nishimura K., Asakura Y., Friso G., Kim J., Oh S.H., Rutschow H., Ponnala L., van Wijk K.J. (2013). ClpS1 is a conserved substrate selector for the chloroplast Clp protease system in *Arabidopsis*. *Plant .Cell.* **25**, 2276-301.

Ohta H., Okumura A., Okuyama S., Akiyama A., Iwai M., Yoshihara S., Shen J.R., Kamo M., Enami I. (1999).Cloning, expression of the *psbU* gene, and functional studies of the recombinant 12-kDa protein of photosystem II from a red alga Cyanidium caldarium. *Biochem. Biophys. Res. Commun.* **260**, 245-50.

Olinares P.D., Kim J., Davis J.I., van Wijk K.J. (2011). Subunit stoichiometry, evolution, and functional implications of an asymmetric plant plastid ClpP/R protease complex in *Arabidopsis*. *Plant. Cell*. **23**,2348-61.

Olinares P.D., Kim J., van Wijk K.J. (2011). The Clp protease system; a central component of the chloroplast protease network. *Biochim. Biophys. Acta*. **1807**, 999-1011.

Ollinger J., O'Malley T., Kesicki E.A, Odingo J., Parish T. (2012). Validation of the essential ClpP protease in Mycobacterium tuberculosis as a novel drug target. *J. Bacteriol*. **194**, 663-8.

Otto M. (2012). MRSA virulence and spread. *Cell. Microbiol* **14**, 1513-21.

Pak M., Wickner S. (1997). Mechanism of protein remodeling by ClpA chaperone. *Proc. Natl .Acad .Sci. U. S. A.* **94**, 4901-6.

Park S.C., Jia B., Yang J.K., Van D.L., Shao Y.G., Han S.W., Jeon Y.J., Chung C.H., Cheong G.W (2006). Oligomeric structure of the ATP-dependent protease La (Lon) of *Escherichia coli.* *Mol. Cells.* **21**, 129-34.

Peters J.M., Cejka Z., Harris J.R., Kleinschmidt J.A., Baumeister W. (1993). Structural features of the 26 S proteasome complex. *J. Mol. Biol*. **234**, 932-7.

Peltier J.B., Ytterberg J., Liberles D.A., Roepstorff P., van Wijk K.J. (2001). Identification of a 350-kDa ClpP protease complex with 10 different Clp isoforms in chloroplasts of *Arabidopsis thaliana*. *J. Biol. Chem*. **276**, 16318-2.

Peltier J.B., Ripoll D.R., Friso G., Rudella A., Cai Y., Ytterberg J., Giacomelli L., Pillardy J., van Wijk K.J (2004). Clp protease complexes from photosynthetic and non-photosynthetic plastids and mitochondria of plants, their predicted three-dimensional structures, and functional implications. *J. Biol. Chem*. **279**, 4768-81.

Pérez-Pérez M.E., Florencio F.J., Lindahl M. (2006). Selecting thioredoxins for disulphide proteomics: target proteomes of three thioredoxins from the cyanobacterium *Synechocystis sp. PCC 6803*. *Proteomics*. **1,** S186-95.

Perelman A., Uzan A., Hacohen D., Schwarz R. (2003). Oxidative Stress in *Synechococcus sp. Strain PCC 7942*: Various Mechanisms for H2O2 Detoxification with Different Physiological Roles. *J. Bacteriol*. **185**, 3654–3660.

Personne Y., Brown A.C., Schuessler D.L., Parish T. (2013). Mycobacterium tuberculosis ClpP proteases are co-transcribed but exhibit different substrate specificities. *PLoS. One*. **8,** e60228.

Persuh M., Turgay K., Mandic-Mulec I., Dubnau D. (1999). The N- and C-terminal domains of MecA recognize different partners in the competence molecular switch. *Mol. Microbiol*. **33**, 886-94.

Porankiewicz J., Schelin J., Clarke A.K. (1998). The ATP-dependent Clp protease is essential for acclimation to UV-B and low temperature in the cyanobacterium *Synechococcus. Mol. Microbiol*. **29**, 275-83.

Ralph S.A., van Dooren G.G., Waller R.F., Crawford M.J., Fraunholz M.J., Foth B.J., Tonkin C.J., Roos D.S., McFadden G.I.. (2004). Tropical infectious diseases: metabolic maps and functions of the Plasmodium falciparum apicoplast. *Nat. Rev. Microbiol.* **2**, 203-16.

Raju R.M., Unnikrishnan M., Rubin D.H., Krishnamoorthy V., Kandror O., Akopian T.N., Goldberg A.L., Rubin E.J.. (2012). *Mycobacterium tuberculosis* ClpP1 and ClpP2 function together in protein degradation and are required for viability in vitro and during infection. *PLoS. Pathog*. **8**, e1002511.

Raju R.M., Jedrychowski M.P., Wei J.R., Pinkham J.T., Park A.S., O'Brien K., Rehren G., Schnappinger D., Gygi S.P., Rubin E.J.. (2014). Post-translational regulation via Clp protease is critical for survival of Mycobacterium tuberculosis. *PLoS. Pathog*. **10**, e1003994..

Reid B.G., Fenton W.A., Horwich A.L., Weber-Ban E.U. (2001). ClpA mediates directional translocation of substrate proteins into the ClpP protease. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 3768-72.

Rathore S., Sinha D., Asad M., Böttcher T., Afrin F., Chauhan V.S., Gupta D., Sieber S.A., Mohmmed A. (2010). A cyanobacterial serine protease of *Plasmodium falciparum* is targeted to the apicoplast and plays an important role in its growth and development. *Mol. Microbi.* **77**, 873–890,

Rigas S., Daras G., Laxa M., Marathias N., Fasseas C., Sweetlove L., Hatzopoulos P.. (2009). Role of Lon1 protease in post-germinative growth and maintenance of mitochondrial function in *Arabidopsis thaliana*. *New. Phytol.* **181**, 588-600.

Robertson G.T., Ng W.L., Foley J., Gilmour R., Winkler M.E. (2002). Global transcriptional analysis of clpP mutations of type 2 *Streptococcus pneumoniae* and their effects on physiology and virulence. *J. Bacteriol.* **18**, 3508-20.

Robertson G.T., Ng W.L., Gilmour R., Winkler M.E. (2003). Essentiality of *clpX*, but not *clpP, clpL, clpC*, or *clpE*, in *Streptococcus pneumoniae* R6. *J. Bacteriol*. **185**, 2961-6.

Román-Hernández G., Grant R.A., Sauer R.T., Baker T.A. (2009). Molecular basis of substrate selection by the N-end rule adaptor protein ClpS. *Proc. Natl. Acad. Sci. U. S. A*. **106**, 8888-93.

Román-Hernández G., Hou J.Y., Grant R.A., Sauer R.T., Baker T.A. (2011). The ClpS adaptor mediates staged delivery of N-end rule substrates to the AAA+ ClpAP protease. *Mol. Cell.* **43**, 217-28.

Rousseau F., Sétif P., Lagoutte B. (1993). Evidence for the involvement of PSI-E subunit in the reduction of ferredoxin by photosystem I. *EMBO. J.* **12**, 1755-65.

Rotanova T.V., Melnikov E.E., Khalatova A.G., Makhovskaya O.V., Botos I., Wlodawer A., Gustchina A. (2004) Classification of ATP-dependent proteases Lon and comparison of the active sites of their proteolytic domains. *Eur. J. Biochem*. **271**, 4865-71.

Rouquette C., Ripio M.T., Pellegrini E., Bolla J.M., Tascon R.I., Vázquez-Boland J.A., Berche P (1996). Identification of a ClpC ATPase required for stress tolerance and in vivo survival of *Listeria monocytogenes. Mol. Microbiol.* **21**, 977-87.

Rouquette C., de Chastellier C., Nair S., Berche P. (1998). The ClpC ATPase of *Listeria monocytogenes* is a general stress protein required for virulence and promoting early bacterial escape from the phagosome of macrophages. *Mol. Microbiol*. **27**, 1235-45.

Rubin D.M., Finley D. (1995). Proteolysis. The proteasome: a protein-degrading organelle? *Curr .Biol*. **5**, 854-8.

Sass P., Josten M., Famulla K., Schiffer G., Sahl H.G., Hamoen L., Brötz-Oesterhelt H. (2011). Antibiotic acyldepsipeptides activate ClpP peptidase to degrade the cell division protein FtsZ. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 17474-9..

Sassetti C.M., Boyd D.H., Rubin E.J. (2003). Genes required for mycobacterial growth defined by high density mutagenesis. *Mol. Microbiol.* **48**, 77-84.

Sauer R.T., Bolon D.N., Burton B.M., Burton R.E., Flynn J.M., Grant R.A., Hersch G.L., Joshi S.A., Kenniston J.A., Levchenko I., Neher S.B., Oakes E.S., Siddiqui S.M., Wah D.A., Baker T.A. (2004). Sculpting the proteome with AAA(+) proteases and disassembly machines. *Cell.* **119**, 9-18.

Sauer R.T., Baker T.A. (2011). AAA+ proteases: ATP-fueled machines of protein destruction. *Annu, Rev, Biochem.* **80**, 587-612.

Schelin J., Lindmark F., Clarke A.K. (2002). The clpP multigene family for the ATP-dependent Clp protease in the cyanobacterium *Synechococcus*. *Microbiology.* **148**, 2255-65.

Schmidt R., Zahn R., Bukau B., Mogk A. (2009). ClpS is the recognition component for Escherichia coli substrates of the N-end rule degradation pathway. *Mol. Microbiol. 72*, 506-17.

Shen J.R., Ikeuchi M., Inoue Y. (1997). Analysis of the psbU gene encoding the 12-kDa extrinsic protein of photosystem II and studies on its role by deletion mutagenesis in *Synechocystis sp. PCC 6803. J. Biol. Chem.* **272**, 17821-6.

Schirmer E.C., Glover J.R., Singer M.A., Lindquist S. (1996). HSP100/Clp proteins: a common mechanism explains diverse functions. *Trends. Biochem. Sci*. **21**, 289-96.

Schuenemann V.J., Kralik S.M., Albrecht R., Spall S.K., Truscott K.N., Dougan D.A., Zeth K. (2009). Structural basis of N-end rule substrate recognition in *Escherichia coli* by the ClpAP adaptor protein ClpS. *EMBO. Rep.* **10**, 508-14.

Shrader T.E., Tobias J.W., Varshavsky A. (1993). The N-end rule in Escherichia coli: cloning and analysis of the leucyl, phenylalanyl-tRNA-protein transferase gene aat. *J. Bacteriol.* **175,** 4364-74.

Shoji M., Cui L., Iizuka R., Komoto A., Neoh H.M., Watanabe Y., Hishinuma T., Hiramatsu K. (2011). walK and *clpP* mutations confer reduced vancomycin susceptibility in *Staphylococcus aureus*. *Antimicrob Agents Chemother.* **55**, 3870-81.

Singh S.K., Rozycki J., Ortega J., Ishikawa T., Lo J., Steven A.C., Maurizi M.R. (2001). Functional domains of the ClpA and ClpX molecular chaperones identified by limited proteolysis and deletion analysis. *J. Biol. Chem.* **276**, 29420-9.

Singh A., Mai D., Kumar A., Steyn A.J.. Dissecting virulence pathways of *Mycobacterium tuberculosis* through protein-protein association. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 11346-51.

Skinner M.M., Trempy J.E.. (2001). Expression of *clpX*, an ATPase subunit of the Clp protease, is heat and cold shock inducible in *Lactococcus lactis*. *J. Dairy. Sci*. **84**, 1783-5.

Sjögren L.L., Stanne T.M., Zheng B., Sutinen S., Clarke A.K. (2006). Structural and functional insights into the chloroplast ATP-dependent Clp protease in *Arabidopsis*. *Plant. Cell*. **18,** 2635-49.

Sjögren L.L., Clarke A.K. (2011). Assembly of the chloroplast ATP-dependent Clp protease in *Arabidopsis* is regulated by the ClpT accessory proteins. *Plant. Cell*. **23**, 322-32.

Sjögren L.L., Tanabe N., Lymperopoulos P., Khan N.Z., Rodermel S.R., Aronsson H., Clarke A.K. (2014). Quantitative analysis of the chloroplast molecular chaperone ClpC/Hsp93 in *Arabidopsis* reveals new insights into its localization, interaction with the Clp proteolytic core, and functional importance. *J. Biol. Chem*. **289**, 11318-30.

Smith D.M., Kafri G., Cheng Y., Ng D., Walz T., Goldberg A.L. (2005). ATP binding to PAN or the 26S ATPases causes association with the 20S proteasome, gate opening, and translocation of unfolded proteins. *Mol. Cell*. **20**, 687-98.

Smith D.M., Chang S.C., Park S., Finley D., Cheng Y., Goldberg A.L. (2007). Docking of the proteasomal ATPases' carboxyl termini in the 20S proteasome's alpha ring opens the gate for substrate entry. Mol. *Cell.* **27,** 731-44.

Sokolenko A., Pojidaeva E., Zinchenko V., Panichkin V., Glaser V.M., Herrmann R.G., Shestakov S.V. (2002). The gene complement for proteolysis in the cyanobacterium *Synechocystis sp. PCC 6803* and Arabidopsis thaliana chloroplasts. *Curr. Genet*. **41**, 291-310.

Song H.K., Bochtler M., Azim M.K., Hartmann C., Huber R., Ramachandran R. (2003). Isolation and characterization of the prokaryotic proteasome homolog HslVU (ClpQY) from *Thermotoga maritima* and the crystal structure of HslV. *Biophys. Chem*. **100**, 437-52.

Sousa M.C., Trame C.B., Tsuruta H., Wilbanks S.M., Reddy V.S., McKay D.B. (2000). Crystal and solution structures of an HslUV protease-chaperone complex. *Cell.* **103**, 633-43.

Stanne T.M., Pojidaeva E., Andersson F.I., Clarke A.K. (2007). Distinctive types of ATP-dependent Clp proteases in cyanobacteria. *J. Biol. Chem.* **282**, 14394-402.

Stanne T.M., Sjögren L.L., Koussevitzky S., Clarke A.K. (2009). Identification of new protein substrates for the chloroplast ATP-dependent Clp protease supports its constitutive role in *Arabidopsis*. *Biochem. J.* **417**, 257-68.

Stahlberg H., Kutejová E., Suda K., Wolpensinger B., Lustig A., Schatz G., Engel A., Suzuki C.K.(1999). Mitochondrial Lon of *Saccharomyces cerevisiae* is a ring-shaped protease with seven flexible subunits. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 6787-90.

Stinson B.M., Nager A.R., Glynn S.E., Schmitz K.R., Baker T.A., Sauer R.T. (2013). Nucleotide binding and conformational switching in the hexameric ring of a AAA+ machine. *Cell*. **153**, 628-39.

Stork T., Michel K.P., Pistorius E.K., Dietz K.J. (2005). Bioinformatic analysis of the genomes of the cyanobacteria *Synechocystis sp. PCC 680*3 and *Synechococcus elongatus PCC 7942* for the presence of peroxiredoxins and their transcript regulation under stress. *J. Exp. Bot.* **56**, 3193-206.

Stephani K., Weichart D., Hengge R. (2003). Dynamic control of Dps protein levels by ClpXP and ClpAP proteases in *Escherichia coli*. *Mol. Microbiol*. **49**, 1605-14.

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**, 575-85.

Tanaka K.(2009).The proteasome: overview of structure and functions. *Proc. Jpn. Acad. Ser. B. Phys. Biol. Sci.* **85**, 12-36.

Takemoto T., Zhang Q.M., Yonei S. (1998). Different mechanisms of thioredoxin in its reduced and oxidized forms in defense against hydrogen peroxide in *Escherichia coli*. *Free. Radic. Biol. Med*. **24**, 556-62.

Thomsen L.E., Olsen J.E., Foster J.W., Ingmer H. (2002). ClpP is involved in the stress response and degradation of misfolded proteins in *Salmonella enterica serovar Typhimurium*. *Microbiology.* **148**, 2727-33.

Thompson M.W., Maurizi M.R. (1994). Activity and specificity of Escherichia coli ClpAP protease in cleaving model peptide substrates. *J. Biol. Chem*. **269**, 18201-8.

Tobias J.W., Shrader T.E., Rocap G., Varshavsky A. (1991). The N-end rule in bacteria. *Science*. *254*, 1374-7.

Tomko R.J. Jr., Funakoshi M., Schneider K., Wang J., Hochstrasser M. (2010). Heterohexameric ring arrangement of the eukaryotic proteasomal ATPases: implications for proteasome structure and assembly. *Mol .Cell.* **38**, 393-403.

Tomoyasu T., Yamanaka K., Murata K., Suzaki T., Bouloc P., Kato A., Niki H., T,Hiraga. (1993). Topology and subcellular localization of FtsH protein in *Escherichia coli*. *J. Bacteriol*. **175**, 1352-1357

Tomoyasu T., Gamer J., Bukau B., Kanemori M., Mori H., Rutman A.J., Oppenheim A.B., Yura T., Yamanaka K., Niki H., et al. (1995). *Escherichia coli* FtsH is a membrane-bound, ATP-dependent protease which degrades the heat-shock transcription factor sigma 32. *EMBO. J*. **14**, 2551-60.

Tsilibaris V., Maenhaut-Michel G., Van Melderen L. (2006). Biological roles of the Lon ATP-dependent protease. *Res. Microbiol*. **157**, 701-13.

Voges D., Zwickl P., Baumeister W. (1999).The 26S proteasome: a molecular machine designed for controlled proteolysis. *Annu. Rev. Biochem*. **68**,1015-68.

Unno M., Mizushima T., Morimoto Y., Tomisugi Y., Tanaka K., Yasuoka N., Tsukihara T. (2002).Structure determination of the constitutive 20S proteasome from bovine liver at 2.75 A resolution. *J. Biochem*. **1**3, 171-3.

Urantowka A., Knorpp C., Olczak T., Kolodziejczak M., Janska H. (2005). Plant mitochondria contain at least two i-AAA-like complexes. *Plant. Mol. Biol*. **59**, 239-52.

Varshavsky A. (1996). The N-end rule: functions, mysteries, uses. *Proc. Natl. Acad. Sci. U. S. A*. **93**, 12142-9.

Varshavsky A. (2003). The N-end rule and regulation of apoptosis. *Nat. Cell. Biol*. **5**, 373-6.

Venkatesh S., Lee J., Singh K., Lee I., Suzuki C.K. (2012). Multitasking in the mitochondrion by the ATP-dependent Lon protease. *Biochim. Biophys. Acta*. **1823**, 56-66.

Viala J., Rapoport G., Mazodier P. (2000). The clpP multigenic family in *Streptomyces lividans*: conditional expression of the *clpP3 clpP4* operon is controlled by PopR, a novel transcriptional activator. *Mol. Microbiol.* **38**, 602-12.

WHO (2013) Global tuberculosis control. Available at: http://www.who.int/tb/publications/global_report/en/index.html (Accessed 2014073030).

Wang J., Hartling J.A., Flanagan J.M. The structure of ClpP at 2.3 A resolution suggests a model for ATP-dependent proteolysis. *Cell.* **91**, 447-56.

Wang J., Hartling J.A., Flanagan J.M. (1998). Crystal structure determination of Escherichia coli ClpP starting from an EM-derived mask. *J. Struct. Biol* **124**, 151-63.

Wang J., Song J.J., Seong I.S., Franklin M.C., Kamtekar S., Eom S.H., Chung C.H. (2001). Nucleotide-dependent conformational changes in a protease-associated ATPase HslU. *Structure*. **9**, 1107-16.

Wang K.H., Roman-Hernandez G., Grant R.A., Sauer R.T., Baker T.A. (2008). The molecular basis of N-end rule recognition. *Mol. Cell.* **32**, 406-14.

Wang K.H., Oakes E.S., Sauer R.T., Baker T.A. (2008). Tuning the strength of a bacterial N-end rule degradation signal. *J. Biol. Chem*. **283**, 24600-7.

Weber N., Strotmann H.. (1999). On the function of subunit PsaE in chloroplast Photosystem I. *Biochim. Biophys. Acta*. **1143**, 204-10.

Weber-Ban E.U., Reid B.G., Miranker A.D., Horwich A.L. (1999). Global unfolding of a substrate protein by the Hsp100 chaperone ClpA. *Nature*. **401,** 90-3.

Weichart D., Querfurth N., Dreger M., Hengge-Aronis R. (2003). Global role for ClpP-containing proteases in stationary-phase adaptation of Escherichia coli. *J. Bacteriol*. **185**, 115-25.

Wickner S., Gottesman S., Skowyra D., Hoskins J., McKenney K., Maurizi M.R (1994). A molecular chaperone, ClpA, functions like DnaK and DnaJ. *Proc. Natl. Acad. Sci. U. S. A.* **9**, 12218-22.

Wickner S., Maurizi M.R., Gottesman S. (1999). Posttranslational quality control: folding, refolding, and degrading proteins. *Science*. **286**, 1888-93.

Wojtkowiak D., Georgopoulos C., Zylicz M. (1993). Isolation and characterization of ClpX, a new ATP-dependent specificity component of the Clp protease of *Escherichia coli*. J. Biol. Chem. 268, 22609-17.

Wojtyra U.A., Thibault G., Tuite A., Houry W.A. (2003). The N-terminal zinc binding domain of ClpX is a dimerization domain that modulates the chaperone function. *J. Biol. Chem*. **278**, 48981-90.

Yamanaka G., Glazer A. N. (1980). Dynamic aspects of phycobilisome structure. Phycobilisome turnover during nitrogen starvation in *Synechococcus* sp. *Arch. Microbiol*. **124**:39–47

Tu G.F, Reid G.E., Zhang J.G., Moritz R.L., Simpson R.J. (1995). C-terminal extension of truncated recombinant proteins in Escherichia coli with a 10Sa RNA decapeptide. *J. Biol. Chem.* **270**, 9322-6.

Yu F., Park S., Rodermel S.R. (2004). The Arabidopsis FtsH metalloprotease gene family: interchangeability of subunits in chloroplast oligomeric complexes. *Plant. J.* **37,** 864-76.

Yu F., Park S., Rodermel S.R. (2005). Functional redundancy of AtFtsH metalloproteases in thylakoid membrane complexes.*Plant. Physiol*. **138**, 1957-66.

Xu Q., Hoppe D., Chitnis V.P., Odom W.R., Guikema J.A, Chitnis P.R. (1995). Mutational analysis of photosystem I polypeptides in the cyanobacterium *Synechocystis sp. PCC 6803*. Targeted inactivation of psaI reveals the function of psaI in the structural organization of *psaL*. *J. Biol. Chem*. **270**, 16243-50.

Zaltsman A., Ori N., Adam Z. (2005).Two types of FtsH protease subunits are required for chloroplast biogenesis and Photosystem II repair in *Arabidopsis. Plant. Cell.* **17**, 2782-90.

Zeiler E., Braun N., Böttcher T., Kastenmüller A., Weinkauf S., Sieber S.A.. (2011). Vibralactone as a tool to study the activity and structure of the ClpP1P2 complex from *Listeria monocytogenes*. *Angew. Chem. Int. Ed. Engl***. 50,** 11001-4.

Zeiler E., List A., Alte F., Gersch M., Wachtel R., Poreba M., Drag M., Groll M., Sieber S.A. (2013). Structural and functional insights into caseinolytic proteases reveal an unprecedented regulation principle of their catalytic triad.*Proc. Natl. Acad. Sci. U. S. A.* **110**, 11302-7.

Zeth K., Ravelli R.B., Paal K., Cusack S., Bukau B., Dougan D.A. (2002). Structural analysis of the adaptor protein ClpS in complex with the N-terminal domain of ClpA. *Nat. Struct. Biol*. **9**, 906-11.

Zheng B., Halperin T., Hruskova-Heidingsfeldova O., Adam Z., Clarke A.K. (2002). Characterization of Chloroplast Clp proteins in Arabidopsis: Localization, tissue specificity and stress responses. *Physiol. Plant.* **114**, 92-101.

Zhou Y., Gottesman S. (1998).Regulation of proteolysis of the stationary-phase sigma factor RpoS. *J. Bacteriol*. **180**, 1154-8.

Zhou Y., Gottesman S., Hoskins J.R., Maurizi M.R., Wickner S. (2001). The RssB response regulator directly targets sigma(S) for degradation by ClpXP. *Genes Dev.* **15**, 627-37.

Zwickl P., Ng D., Woo K.M., Klenk H.P., Goldberg A.L. (1999). An archaebacterial ATPase, homologous to ATPases in the eukaryotic 26 S proteasome, activates protein breakdown by 20 S proteasomes. *J. Biol. Chem*. **274,** 26008-14.

5. POPULÄRVETENSKAPLIG SAMMANFATTNING

Alla levande organismer är uppbyggda av celler som innehåller flera olika livsviktiga byggstenar. En av dessa är proteiner och en enda cell kan bestå av flera tusen olika typer av proteiner. Alla proteiner är uppbyggda av långa kedjor av aminosyror som veckar sig till olika tredimensionella strukturer och det är aminosyrorna tillsammans med strukturen av proteinet som bestämmer vilken funktion proteinet har. Proteiner fungerar på flera olika sätt i cellen. De kan bl.a. vara en strukturell del i det skyddande membranet som definierar cellen eller påverkar viktiga kemiska reaktioner. Dock kan proteiner även vara skadliga för cellen, om ett protein veckar sig felaktigt eller på något annat sätt blir skadat kan den utgöra en fara för cellens olika funktioner. För det första kan proteinet inte längre utföra sin funktion i cellen, vilket i sig kan leda till problem men det felaktiga proteinet kan också börja klumpa ihop sig med andra proteiner och bilda aggregat som kan vara livsfarliga för cellen. Det är därför viktigt för cellen att kontrollera funktionen av proteinerna genom att se till att de veckar sig korrekt och ta bort skadade eller felveckade proteiner. Detta kontrolleras på flera nivåer och två viktiga typer av proteiner som är involverade i detta system är chaperoner och proteaser.

Chaperonerna kontrollerar så att proteiner veckar sig till sin korrekta tredimensionella struktur, om proteinet inte kan genomföra detta självständigt så binder chaperonet till proteinet och hjälper till. Chaperonerna ser också till att förhindra att de farliga aggregaten bildas och om det ändå sker så kan chaperonen lösa upp dem. Dessutom kan chaperonen se till att proteiner som är obotligt skadade tas bort från cellen. Detta görs genom att chaperonet känner igen det skadade protein genom olika strukturella signaler som leder till att chaperonet binder till proteinet och presenterar det för ett specifikt proteas. Proteaset har den proteolytiska aktiviteten, vilket betyder att den bryter ner proteinet till aminosyror, genom att bryta de kemiska bindingarna mellan aminosyrorna. Det finns flera olika typer av chaperoner och proteaser i en cell, där en viktigt typ är Clp-proteaser.

Clp-proteaser existerar i nästan allt levande från små bakterier till människor och växter. De har blivit noggrant studerade i modellorganismen *E. coli* (tarm-bakterier), där Clp-proteaset består av ett energiberoende chaperon (ClpA eller ClpX) och en proteolytisk kärna. Den proteolytiska kärnan är uppbyggd av två ringar av sju ClpP proteiner (en heptamer) och har formen som en tunna. Inuti tunnan finns de aminosyror som är viktiga för att bryta sönder proteinet, själva ingången till tunnan är väldigt trång. Detta är ett sätt att kontrollera så att inga normala proteiner blir felaktigt nedbrutna. Det hela fungerar på följande sätt: ClpA eller ClpX känner igen ett felaktigt protein (ett substrat), binder till det och vecklar upp hela strukturen av proteinet så att den blir en enda lång aminosyrakedja igen, som kan transporteras igenom den smala

öppningen i ClpP proteaset och proteinet bryts ner. Även om Clp-proteaserna finns i de flesta levande organismer är det hos de fotosyntetiska organismerna som de är mest komplexa och funktionellt livsviktiga. En viktig modellorganism för fotosyntetiska organismer är cyanobakterier (förr kallade blågröna alger). Dessa bakterier kan fotosyntetisera som växter, det vill säga de kan omvandla solljus och koldioxid till energi i en reaktion som också skapar syre. Förfäderna till cyanobakterien var viktiga för syresättningen av atmosfären för cirka 3,5 miljarder år sedan och de var även en del i evolutionen av växter, genom att forma kloroplasten. Cyanobakterier lever på alla möjliga områden i världen, både i världshaven och i små sjöar. En väl använd modellorganism i laboratorierna är *Synechococcus elongatus (Synechococcus). Synechococcus* har ett mer komplext system med Clp-proteaser jämfört med *E. coli*. Flertalet av Clp-proteiner i *Synechococcus* är livsviktiga för cellen medan inga av Clp-proteinerna i *E. coli* är det. *Synechococcus* har två Clp-proteaser, som båda består av proteolytiska kärnor som är uppbyggda av mer än en typ av protein, ClpCP3/R och ClpXP1/P2. Det essentiella ClpCP3/R är uppbyggt av ett proteolytiskt aktivt protein, ClpP3 och ett proteolytisk inaktivt protein, ClpR. ClpR-proteiner är unika för fotosyntetiska organismer. ClpP3/R består av två ringar av tre ClpP3 och fyra ClpR, som sitter alternerade till varandra. *Synechococcus* har också två ClpS (S1 och S2) adaptorer. ClpS1 känner igen felaktiga proteiner, baserat på vilken aminosyra som proteinet börjar på. ClpS1 binder till dessa proteiner och levererar dem till ClpC, som i sin tur ser till att de blir nedbrytet av ClpP3/R. Detta betyder att en adaptor helt enkelt fungerar genom att ändra vilka proteiner ClpC känner igen. I mina studier har jag och mina kollegor vidare undersökt funktionen av Clp-proteaser in *Synechococcus*, för att försöka besvara vilken funktion dessa proteaser har.

I den först studien har vi undersökt hur det kommer sig att ClpC kan binda till ClpP3/R kärnan. Vi kunde där bevisa att tre olika domäner på den N-terminala änden av både ClpP3 och ClpR binder till ClpC, likväl finns en domän i ClpC som gör att chaperonet endast kan binda till specifika proteolytiska kärnor, så som ClpP3/R men inte EcClpP. I nästa studie studerades själva strukturen av den proteolytiska kärnan i ClpP1/P2. Detta gjordes genom en teknik som studerar olika massor av molekylerna (masspektrometri). Genom detta kunde vi visa att ClpP1/P2 är uppbyggd av två ringar av alternerande ClpP1 och ClpP2. Det var tydligt att det faktiskt kunde bildas två olika typer av kärnor: (4ClpP1+3ClpP2) + (3ClpP1+4ClpP2) och 2×(3ClpP1+4ClpP2). I den tredje studien fortsatte vi att studera ClpP1/P2, men fokuserade på funktionaliteten av proteaset. ClpP1/P2 kunde nedbryta proteiner tillsammans med ClpX, men inte med ClpC. Det var också tydligt att både ClpP1 och ClpP2 är proteolytiskt aktiva i kärnan. När vi sedan studerade ClpP1 och ClpP2 var för sig, visade det sig att ClpP2-protein inte kunde forma de typiska ringstrukturerna av en proteolytisk kärna för sig själv, medan

ClpP1 kunde göra det. Dessutom hade denna självständiga ClpP1-kärna aktivitet med både ClpC och ClpX, men aktiviteten var beroende av en högre koncentration av magnesium. I den fjärde studien, fokuserade vi istället på proteinerna ClpS1 och ClpS2. ClpS2 fungerade som ClpS1, men kände igen andra aminosyror än ClpS1. Dessutom är ClpS2 livsviktig för cell, vilket ClpS1 inte är. I den sista studien använde vi oss av mutanter för att undersöka funktionen av ClpS1 i *Synechococcus*. En mutant i de här fallen är att vi har inaktiverat den gen som uttrycker det protein som vi är intresserade av, alltså kommer det proteinet inte längre att finnas i cellen. Vi inaktiverade *clpS1*-genen och denna mutant (Δ*clpS1*) hade en bättre överlevnad när den utsattes för H_2O_2 (väteperoxid). Det samma gäller Δ*clpP2. D*äremot var Δ*clpP1* mer känslig än den vanliga *Synechococcus* för H_2O_2. Dessa skillnader i mutanterna kunde förklaras med skillnader i ett protein som heter KatG. Detta protein har en funktion där den omvandlar det skadliga H_2O_2 till harmlöst vatten och syre. I de mer resistenta Δ*clpS1* och Δ*clpP2* var nivåerna högre av KatG medan i Δ*clpP1* var den lägre. Detta indikerar att Clp-proteinerna är involverade i det försvar *Synechococcus* har emot H_2O_2.

Studierna i denna avhandling ger en ökad förståelse för strukturerna och vilka funktioner de olika Clp-proteinerna har i *Synechococcus*. Mina förhoppningar är att mina resultat kan användas vidare i forskningen för den fortsatta karakteriseringen av Clp-proteinerna både i *Synechococcus* men även andra organismer. Exempelvis finns det nya antibiotika som påverkar Clp-proteinerna i olika mänskliga patogen. Kanske kan mina resultat vara till nytta i den forskningen.

6. ACKNOWLEDGMENTS

First and foremost I would like to thank my supervisor **Adrian Clarke**. I am extremely grateful for all the help during these years (time just flies by), it feels like it was yesterday I started to work with the Clp proteins. I want to thank you for always finding time to discuss problems and solving them, even when you were on the other side of the world. There have been many and long discussions, but they have always been interesting and usually ended in a new project. =) I would like to especially thank you for the help with thesis and manuscript writing here in the end. This summer did not turn out as planned but a big pack of beer is heading your way…

I would like to thank all my collaborators who have contributed to the work in this thesis, I would specially thank:

Noriaki Tanabe: Without your work, there would probably be no thesis. Thanks for all the help in the lab, especially in magically solving the cloning problem. I hope that you take some time there in Japan and do as we Swedes, just stand still in the sun!

Anders Tryggevesson: Well, I survived more than six years in the chaos of the "boys room". =) Thank you for all the help during these years. It has been great to know that I could always ask you for help in the lab or to discuss problems.

Panagiotis Lymperopoulos: Thanks for all the last min experiments!

Lars Sjögren: You have been away from our lab a long time now, but still I need to give you extra thanks for teaching me to do perfect western blots. It has been extremely helpful over the years.

I would also like to thank the department for funding my phD and also Swedish Research Council (VR) and stiftelsen Olle Engkvist Byggmästare funding the projects.

Jag skulle vilja tacka alla i botanhuset, det är ni som gjort att man överlevt alla dessa år på jobbet. Det hade var så otroligt mycket tråkigare att undervisa om man inte delade jobbet med roliga kollegor. **Sara**: Tänk vad roligt det kan vara att förberedda studentlab och vad man lär sig saker på vägen, så man sedan kan gnälla på studenterna när de gör samma misstag**. Johanna**: Vilka otroligt roliga och intressanta diskussioner vi har hunnit med, och alla så vetenskapliga... **Oskar**: Tack för all hjälpliga råd om labarbete undervisning och spel! **Jenny:** Jag har en känsla av att det är dig jag ska "tacka" för min fina 30års present… Jag tog ju mig runt i alla fall. **Anders N**: För alla intressanta diskussioner och för mallen till avhandlingen, den var guld värd. **Monica**: Tack för att jag har fått låna din "pinne" om och om igen. **Sven**: För alla hjälp med smått som stort. **Emelie**, **Anders N**, **Henrik**, **Per**, **Lisa** och alla ni andra som är och har varit på botan, tack för en trevlig tid!

Även ett tack till mina vänner ute i den riktiga världen, det har varit många roliga nyår, midsommar och semestrar. Ett extra tack till Greta och Jonas för sista-minuten-läsning av min svenska sammanfattning. Ett extra tack till min kusin, som såg till att jag fick hela helger av bara spel och inget jobb!

Tack mamma , pappa och mormor för att ni alltid har funnits där när jag behöver er, älskar er. Tack även till mina kära syskon, min bror som tar ner mig till verkligeten och min syster som håller mig i trim och ser till att jag lär mig laga mat.

Till sist men inte minst, tack till min kära Tobben. Utan ditt tålamod under dessa år, framför allt de två sista skulle jag aldrig blivit färdig. Nästa sommar kan vi ha en riktig semester. =)