2. Introduction

Bacillus subtilis is a well-studied model organism for the Gram-positive bacteria. In its natural habitat, *B. subtilis* faces a wide array of environmental challenges including starvation for nutrients and various stress conditions. Thus, *B. subtilis* developed adaptation strategies involving developmental and differentiation processes such as (i) motility and synthesis of extracellular proteases upon entry into stationary phase, (ii) biofilm formation, (iii) competence development and (iv) sporulation. Accordingly, on the molecular level, *B. subtilis* employs a wide range of regulatory circuits including two-component systems, alternative σ factors, regulatory small RNAs, regulated proteolysis and a protein quality control network.

2.1. General and regulated proteolysis

2.1.1. Protein quality control

A network of chaperones and proteases ensures the functionality of the cellular proteome and can act at any stage of the lifetime of a particular protein: (i) upon *de novo* folding at its synthesis at the ribosome (ii) when targeted to its cellular destination, (iii) maintaining its fold and function upon stress condition which might lead to misfolding and / or aggregation and (iv) when it is finally degraded either in a regulated way as part of a signal transduction pathway or as removal strategy when previous disaggregation and refolding trials failed.

The cytosol is a highly crowded molecular environment. Severe stress conditions such as exposure to oxidative agents or heat treatment lead to protein misfolding and aggregation. The accumulation of aggregated proteins does not only correlate with a loss of function of these particular proteins. The more so as the exposure of hydrophobic residues by these malfolded proteins could lead to a further aggregation of native proteins in the cytosolic environment (Fig. 2).

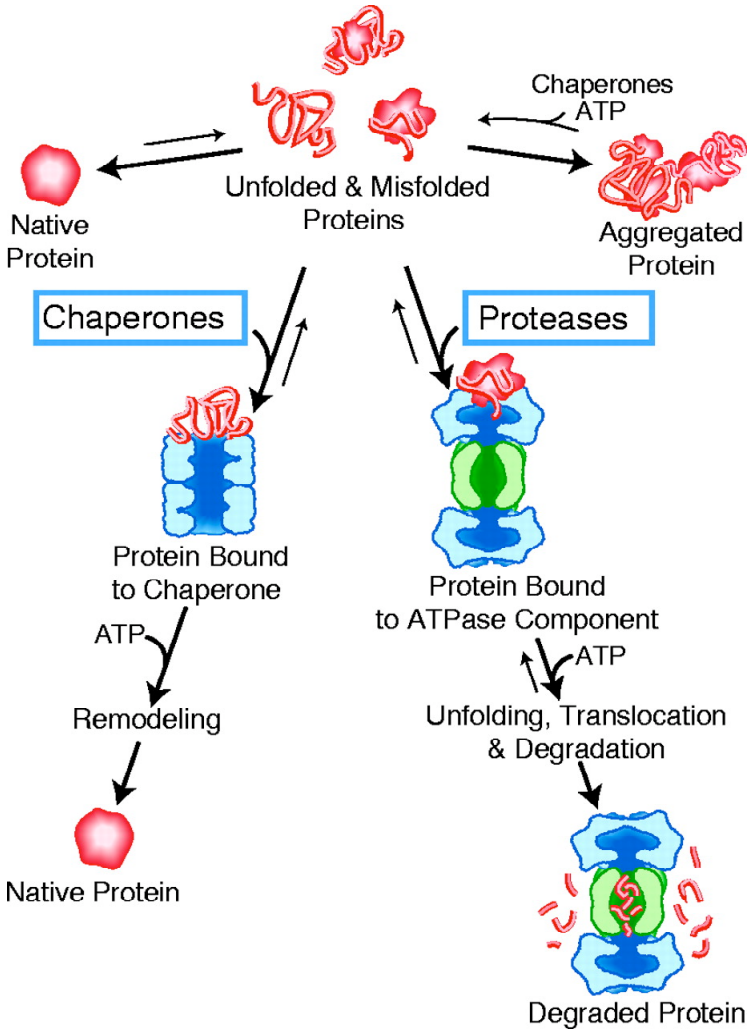

Fig. 1 Overview of the protein quality control system with an emphasis on the decision between the two alternative strategies to cope with unfolded and aggregated proteins: remodeling by chaperones (left path) vs. degradation by proteases (right path). (from Wickner et al., 1999)

Therefore, the cell had to evolve a system to cope with these situations (i) to prevent misfolding or aggregation of already unfolded proteins and (ii) to rescue these proteins by disaggregation and subsequent refolding or degradation (Fig. 1 + 2). Both processes are carried out by chaperones, which guard the folding status of the client protein by (i) a ´holder´ function such as the DnaK system, which bind to unfolded stretches of the particular protein to prevent its aggregation and (ii) by a disaggregation / refolding activity by ´folders´ such as ClpB or the ubiquitous GroEL/GroES. The degradation of misfolded / aggregated proteins is an alternative

strategy for those proteins, which are irreversibly damaged. This proteolytic clearance is mainly carried out by the Lon and Clp proteases (see 2.2.). The decision making process whether a misfolded protein will be subject to refolding or proteolytic turnover (Fig. 1) is not very well understood, but might be determined by kinetic parameters. An illustration summarizing the specific activity of the different members of the protein quality control network is given in figure 2.

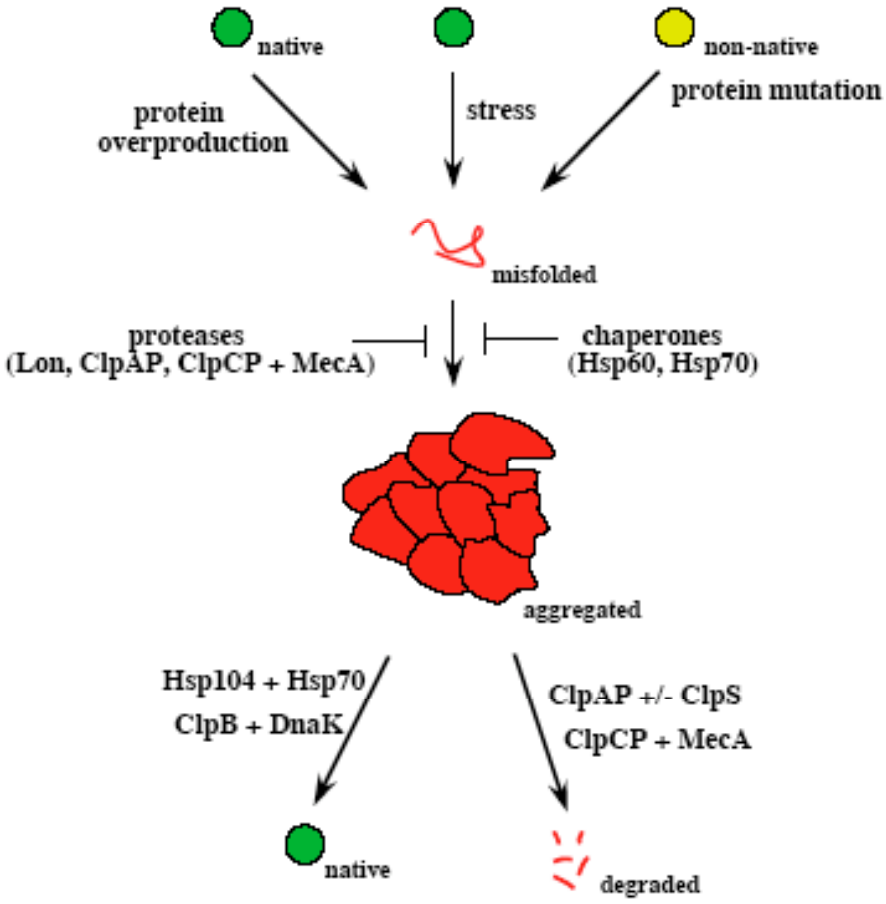

Fig. 2 Intervention at the different stages of protein unfolding and aggregation by specific members of the protein quality network of *E. coli* and *B. subtilis*. from (Weibezahn et al., 2004)

Although the members of the protein quality network are highly conserved, the composition and interplay of the chaperone and proteolytic system differs remarkably

between the two well-studied model organisms, *E. coli* and *B. subtilis*. For instance, in *E. coli* the DnaK/J chaperone system shows an overlapping substrate spectrum with the ribosome associated trigger factor (TF). A double knockout exhibits a synthetic lethality being only viable in a very narrow temperature range (Deuerling et al., 2003). In contrast, a *B. subtilis* Δ*dnaK*/Δ*tig* double deletion is viable even under heat shock conditions (T. Schlothauer, personal communication) suggesting that other chaperones are involved in the maintenance of the cellular integrity and the *de novo* folding of proteins. Moreover, it was recently reported that the secretion specific chaperone SecB acts as suppressor for DnaK / TF in *E. coli* (Ullers et al., 2004). *B. subtilis* however, lacks a clear homolog of SecB. In addition, the *B. subtilis* genome does also not encode a ClpB homolog, which is the major disaggregation and refolding machine cooperating with DnaK in *E. coli*. Possibly, this chaperone activity is accomplished by the Clp ATPases, ClpC, ClpE and ClpX (see 2.2.2.) in *B. subtilis*. Electron microscopy and immunogold-labelling analysis´ already showed that ClpP and all three ATPases localize to heat-aggregated proteins (Krüger et al., 2000; Miethke et al., 2006) as well as to inclusion bodies generated by overexpression of heterologous proteins (Jürgen et al., 2001). Additionally, *in vitro* studies demonstrated that the adaptor proteins MecA and YpbH (see 2.3.1.) together with ClpC form a chaperone system, exhibiting disaggregation and refolding activity (Schlothauer et al., 2003).

2.1.2. Regulated proteolysis

Regulated proteolysis accounts for the selective removal of signaling proteins, which is important for the control of gene expression especially when cells undergo differentiation or developmental processes or face rapid environmental changes. Examples of regulated proteolysis are competence development (Turgay et al., 1998), sporulation (Pan et al., 2001), cell wall biogenesis (Kock et al., 2004b) and adaptation to heat shock in *B. subtilis* (Krüger et al., 2001), the general stress response in *E. coli* (Zhou and Gottesmann, 1998) and cell cycle control in *Caulobacter crescentus* (McGrath et al., 2006). The precisely controlled degradation of a conditionally unstable regulatory protein allows a fast adaptation of its cellular level to a specific stimulus. Several key regulators could already be identified as substrates for the Clp proteases in *B. subtilis* e.g. SpoIIAB, MurAA, ComK, ComS, Spx and Sda (Kock et

al., 2004b; Pan et al., 2001; Ruvolo et al., 2006; Turgay et al., 1998), thus emphasizing their importance in the regulated proteolysis.

Additionally and in contrast to *E. coli,* where Lon seems to be the dominant protease, the Clp proteases are the major determinants of the protein turnover in *B. subtilis* (Kock et al., 2004a). Consequently, a knockout of *clpP* has a severe phenotype e.g. a filamentous and poor growth even under normal, but especially under stress conditions and starvation, an impairment in competence, sporulation, motility, cell division and thermo-tolerance (Gerth et al., 1998; Msadek et al., 1998). Again, supporting the vital role of the Clp proteins in all major cellular processes.

2.1.2.1. Regulation of proteins involved in the protein quality control by regulated proteolysis – the heat shock response in *B. subtilis*

As noted above, the repressor of the class III heat shock genes, CtsR, is subject to regulated proteolysis. The CtsR regulon comprises the *clpC* operon and the monocistronic genes *clpE* and *clpP*. Thus, the cellular protein level of ClpP and the two Clp ATPases, ClpC and ClpE, which are members of the protein quality control network are controlled through regulated proteolysis.

The heat shock response in *B. subtilis* is mainly regulated at the level of transcription either (i) by an activator such as the alternative σ factor σ^B (class II) and the response regulator CssR (class V) or (ii) by a repressor protein such as HrcA (class I) and CtsR (class III) or (iii) by an unknown mechanism (class V). CtsR is encoded by the first gene of the tetracistronic *clpC* operon and binds as a dimeric repressor to a highly conserved heptanucleotide direct repeat (Derre et al., 1999; Krüger and Hecker, 1998). The activity of CtsR is thought to be controlled on the level of its concentration and thus its availability as repressor. It is assumed that under normal growth conditions the CtsR concentration is kept at a basal level ensuring an efficient binding to its target sites, whereas degraded by ClpCP upon heat shock (Krüger et al., 2001). It is assumed that the two adjacent downstream genes of *ctsR*, *mcsA* and *mcsB*, are involved in the regulation of the CtsR activity (Krüger et al., 2001). However, the precise role of both modulators especially with regard to ClpC is not very well understood.

2.2. Clp/HSP100 proteins belong to the AAA+ superfamily of ATPases

The AAA+ (ATPases associated with various cellular activities) superfamily comprises ATPases with a structurally conserved ATP binding module as defining feature. They are found in all kingdoms of life and participate in diverse cellular processes including membrane fusion, DNA replication, protein remodeling and proteolysis. This functional variety is reflected in their extensive number of accessory domains and interaction partners.

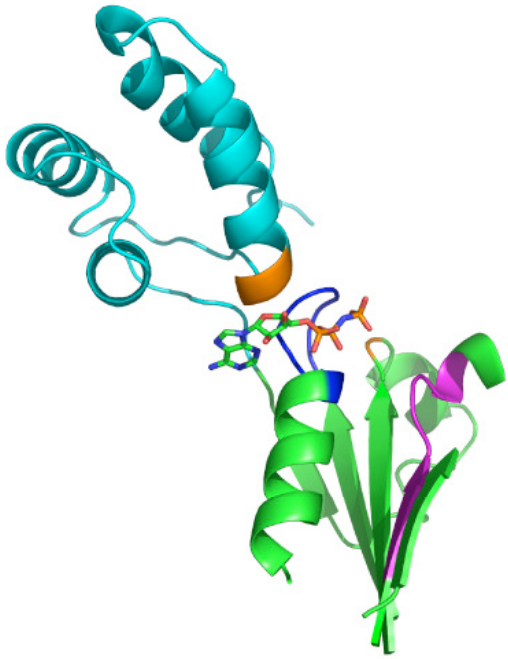

Fig. 3 Core of an AAA+ domain. Depicted are the secondary structural elements of an ATP bound pair of adjacent subunits of the NSF-AAA-2 hexamer (pdb: 1d2n). The Walker A and B motif are highlighted in blue and magenta, respectively. (from Ogura & Wilkinson, 2001)

The AAA module, typically spanning 200-250 aa, contains Walker A (GX_4GKT) and B (h_4DE) signature sequences enabling ATP binding and hydrolysis (Patel and Latterich, 1998) and is activated upon formation of the oligomer. The nucleotide binding pocket is located (i) between a large αβ-domain and the proximate smaller α-helical domain within an AAA+ module and (ii) at the interface of two adjacent subunits in the hexamer (Bochtler et al., 2000). An intermolecular catalysis mechanism, which involves a conserved arginine protruding as ´arginine finger´ into

the neighbouring subunit, explains the greatly enhanced ATPase activity in the oligomeric state (Karata et al., 1999). AAA+ proteins use the chemical energy of a series of ATP hydrolysis events to drive the remodeling of the substrate. The ATP mediated conformational changes within the assembled oligomer are thus transferred to the target molecule, protein or DNA, thereby resulting in the remodeling. The AAA+ superfamily embraces the classical AAA protein family, which contains in addition to the Walker motifs a specific signature, referred to as the SRH (second region of homology). Sequence alignment analysis´ led to an extension of the AAA family forming the wider AAA+ superfamily, including now the Clp/HSP100 proteins (caseinolytic proteins; heat shock proteins; (Neuwald et al., 1999). Orthologs of the Clp proteins are found in eubacteria, but also in mitochondria and chloroplasts and function as large molecular machines. Based on their signature motifs, the AAA+ family is subdivided into seven clades. The ClpABC family harbors two ATPase modules (referred to as AAA-1 and AAA-2) and is termed class I. Whereas the N-terminal AAA-1 domain shares the signature of the classic clade, the C-terminal AAA-2 module contains a β-hairpin insertion, which is characteristic for the AAA+ domains of clade five. Thus, the Clp ATPases of the ClpABC family originated from two distantly related AAA+ modules. The class II ATPases, ClpX and HslU (ClpY), contain only one AAA+ domain, which is homologous to the AAA-2 module of the class I proteins.

The association with other domains confers the AAA+ proteins their functional diversity. Two different domains are found to be linked to the AAA+ module (i) a protease domain, either of the type of a serine or a metalloprotease and (ii) domains usually at the N-terminus, which mediate either the targeting to a specific subcellular location or the interaction with adaptor proteins. The peptidase and ATPase activity can reside on one polypeptide (Lon and FtsH) or can be separated onto two different proteins (Clp proteases). The separation of the AAA+ ATPase from the peptidase allows a proteolytic independent function of the Clp ATPase as chaperone (Levchenko et al., 1995). The interaction of the hexameric ATPase module with the heptameric ClpP protease proceeds via a conserved surface exposed IGF/L loop of the ATPase subunit (Kim et al., 2001). The association with the ATPase component is required for the recognition, unfolding and subsequent translocation into the proteolytic ClpP chamber. Figure 4 shows a schematic representation of the Clp/HSP100 proteins of *E. coli* and *B. subtilis*. The highest variation can be observed

in terms of the accessory domains. The NTD of ClpC for instance is homologous to the one of ClpA and ClpB. ClpX on the other hand harbors a C4-type Zn binding domain (ZBD) in its NTD, which can be found in the NTD of ClpE as well. It could be demonstrated for *E. coli* ClpX that the ZBD extends out of the AAA+ base and contributes to substrate recognition and translocation (Thibault et al., 2006). The second additional functional element, the Linker domain, is present in ClpB, ClpC and ClpE. The Linker has a coiled coil structure, varies in size and is inserted into the smaller C-terminal subdomain of AAA-1. Except for ClpB, all HSP100/Clp proteins possess in their AAA-2 the conserved P-loop motif enabling the interaction with ClpP (Kim et al., 2001). Thus, with respect to its domain properties, ClpC is positioned between ClpA and ClpB. Nevertheless, ClpC differs from the homologous AAA+ proteins in that it requires the interaction with an adaptor protein (see below 2.3.) for virtually all its activities.

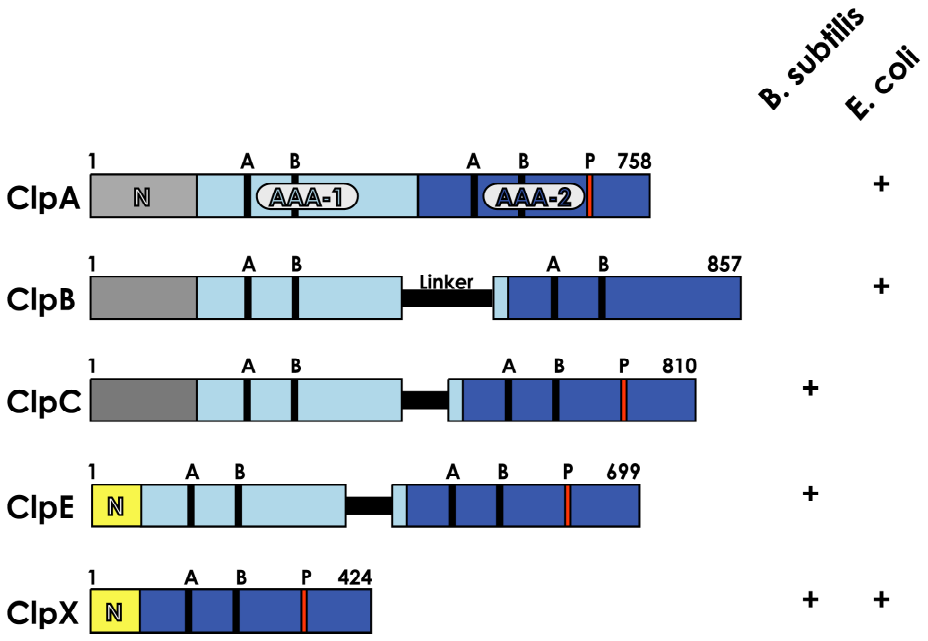

Fig. 4 Domain organization of the Clp/HSP100 proteins in *B. subtilis* and *E. coli*. The arrangement and position of the AAA+ modules (AAA-1 and AAA-2), the Walker A and B sites (A, B), the P-loop, enabling the interaction with ClpP, (P) and the accessory domains: NTD (N) and Linker are indicated.

2.2.1. Architecture of Clp proteases and the degradation process

The eubacterial ClpP proteases resemble in their structure the eukaryotic 26S proteasome (Fig. 5). ClpP is a homo-tetradecamer composed of two heptameric rings, which are stacked face to face to build a barrel shaped degradation chamber (Wang et al., 1998). The 14 serine proteolytic active sites that catalyze the peptide bond cleavage are sequestered in the interior of the proteolytic core and thereby shielded from the cytosol. Substrates enter the aqueous cavity through axial pores. However, the channel is only 50 Å in diameter and therefore too narrow to allow the entry of a folded protein. The unfolding and subsequent translocation of a native protein requires the activity of the Clp ATPases. The hexameric ATPases bind in an asymmetric manner either on one or on both axial sites of the double heptameric ClpP. It was shown, that the interaction of the ClpX ATPase with ClpP relies on two kinds of contacts. The interaction of the six ClpX IGF loops with the hydrophobic pockets in ClpP sets the general position of the two asymmetric rings in the proteolytic complex. This interaction is independent of the nucleotide state of different subunits of ClpX and stabilizes an open-pore conformation of ClpP. A different kind of interaction is mediated between the pore-2 loops of ClpX and N-terminal loops of ClpP. These rather weak interactions vary with the nucleotide state of the particular subunit of ClpX and allow a fine-tuning of the ATP hydrolysis rates (Martin et al., 2007). Repetitive conformational changes in the AAA+ ring upon ATP hydrolysis are required to translocate the polypeptide into the degradation chamber of ClpP (Fig. 5). The cleavage reaction itself occurs without sequence specificity (Thompson et al., 1994). By means of cryo and negatively stained EM, the process of substrate unfolding and translocation could be directly visualized for ClpAP and ClpXP. The substrate was bound at the distal end of the ATPase component and using an inactive ClpP variant, the substrates were trapped either within the ATPase or the degradation chamber (Ortega et al., 2000). Interestingly, both ATPase components displayed similar affinities for ClpP. Thus, the association of ClpP with the ATPases occurred randomly and resulted in the formation of bifunctional hybrid ClpAPX complexes (Ortega et al., 2004). The translocation of substrates appeared to alternate between the two oppositely bound ATPases. Whether the *in vitro* observation of the formation of heterologous Clp complexes holds true for the living cell remains elusive. Nevertheless, these data suggest a communication between the poles of the proteolytic complex.

The limited size of the passageway of both, the channel of the ATPase and the pore of the ClpP cylinder, suggests a unidirectional translocation process. Fluorescently labeled probes were used to monitor the time dependent translocation of GFP-ssrA, labeled either at its N- or C-terminus into the ClpP cavity by FRET (Reid et al., 2001). The energy transfer between the fluorophore of the ClpP cavity and the fluorescein labeled substrate was faster for the C- than for the N-terminally labeled GFP-ssrA. Thus indeed, the unfolding and translocation takes place in a unidirectional manner starting from the attachment point of the degradation tag.

Moreover, it could be shown for *E. coli* ClpX, that the NTD and in particular the ZBD undergoes conformational block movements towards ClpP in response to ATP binding. This nucleotide dependent conformational change of ClpX accounts for a switch between a substrate capture state (ADP-bound) and the inward movement of the ZBD (ATP bound) and thus resulting in the threading of the substrate into the ClpP chamber (Thibault et al., 2006).

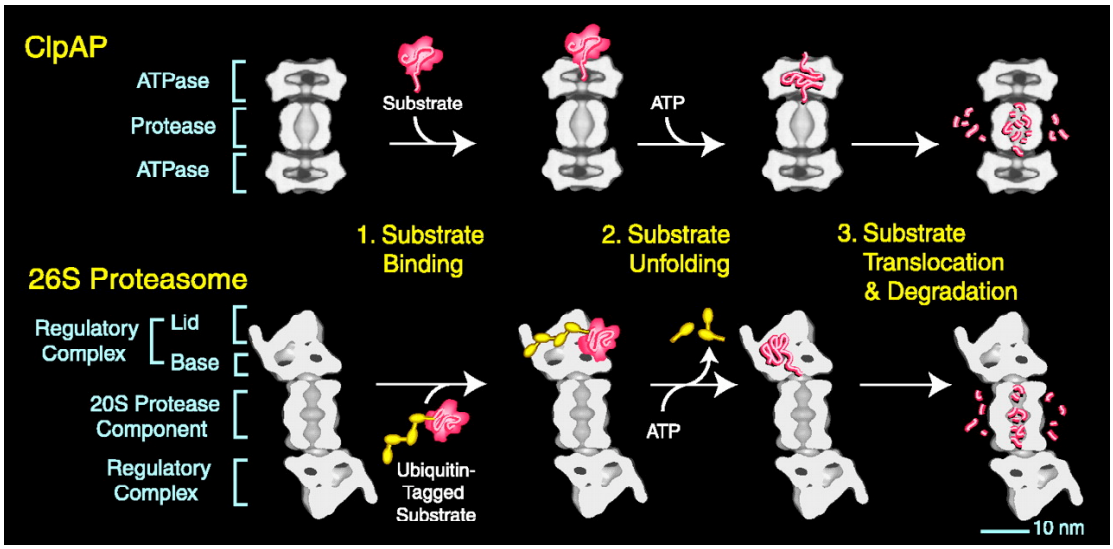

Fig. 5 Schematic model of the different stages of the protein degradation process. Shown is a comparison of the prokaryotic ClpAP (upper panel) with the eukaryotic 26S proteasome mediated proteolysis (lower panel). (Wickner et al., 1999)

As demonstrated for ClpX, the ATPase cycle of the AAA+ protein and the subsequent substrate engagement are stochastic processes. The ATP hydrolysis occurs in a

sequential, but not strictly ordered manner and is independent of the nucleotide status of the neighbouring subunits (Martin et al., 2005). Such a probabilistic way of ATP hydrolysis can be seen as an adaptation to the irregular shape of the substrate protein. Only those subunits, which are in contact with the substrate, are stimulated to hydrolyze ATP. Consequently, the interaction of just one ATPase subunit with the substrate would already allow a successful translocation and thereby avoid stalling or wasted hydrolysis. The ATPase cycle might be more complex for the class I Clp ATPases, which contain two AAA+ modules per polypeptide.

The high energy consumption of the Clp mediated proteolysis is surprising considering that the final hydrolysis of the peptide bond is an exergonic process. However, as outlined above, ATP hydrolysis is required for the denaturing step to unfold the substrate and for the subsequent translocation. It was shown that the enzymatic denaturation requires a minimum of 20 but could increase to as much as 500 ATP molecules (Kenniston et al., 2003). The variation in the ATP consumption is based on the stochastic nature of the unfolding process. Each ATP cycle represents a new unfolding attempt and the probability to successfully unfold the protein depends on the structural stability of the substrate. Finally, the translocation process into the ClpP cavity adds another 0.3-1 ATP per aa. These high costs emphasize the need of a carefully regulated proteolysis of every single substrate (see 2.3.1.).

2.3. Interplay of Clp/HSP100 and adaptor proteins

2.3.1. Adaptor proteins

AAA+ proteins are modulated by a group of otherwise unrelated proteins termed adaptor proteins. They display a huge diversity in sequence and structure and vary in size, but are usually rather small proteins (Dougan et al., 2002a). Adaptor proteins specifically modulate the substrate recognition and / or chaperone activity of their partner AAA+ protein. It was shown for ClpA and ClpX that their adaptor proteins, ClpS, SspB and RssB (see below), bind to the NTD (Dougan et al., 2003; Levchenko et al., 2003; Zeth et al., 2002). Indeed, the NTD is not only positioned distal to the peptidase binding site, it is also the most variable domain of the AAA+ proteins, which might account for the specific interplay of adaptor protein and its cognate ATPase.

2.3.1.1. *E. coli* adaptor proteins

Adaptor proteins can (i) enable the substrate recognition in the first place and (ii) enhance the degradation kinetics. It could be demonstrated that SspB, an adaptor protein targeting SsrA-tagged proteins (see below) to ClpX in *E. coli*, facilitates the degradation by tethering AAA+ and substrate protein (Levchenko et al., 2005). Such a delivery complex becomes important especially at low substrate concentrations in which the substrate alone would not bind strongly enough to ClpX. The tethering increases the local substrate concentration leading to a 10-fold increase in the degradation kinetics (McGinness et al., 2006). Such tethering sequences are also present in the two additional *E. coli* ClpX adaptor proteins, UmuD and RssB and enable the binding to a common site on the NTD of ClpX (Dougan et al., 2003; Neher et al., 2003b; Siddiqui et al., 2004; Wah et al., 2003). This shared interaction mode indicates a competition of these adaptor proteins and thereby allows a change in the priority of substrate turnover.

UmuD resembles SspB in that it uses a related peptide motif to bind to the NTD of ClpX for tethering UmuD´ to ClpX. UmuD is processed to UmuD´ in a post-translational process and is as such a component of the SOS response induced error prone DNA polymerase V. The cellular protein level of UmuD´ is strictly limited to the SOS response and proteolytically controlled by ClpXP. UmuD is required for the degradation of UmuD´ to form the substrate delivery complex UmuD´-UmuD-ClpX and can therefore be regarded as adaptor protein for ClpX (Neher et al., 2003b).

RssB is the third ClpX adaptor protein in *E. coli*, enabling the initial recognition of its substrate RpoS (σ^S), the stationary phase σ factor prior to its partner ATPase, ClpX. RssB belongs to the two component response regulator family of proteins, whose activity is regulated by the phosphorylation at D58 (Bouche et al., 1998). The targeting of σ^S to ClpXP occurs in a two-step-mechanism as only the interaction of RssB-P with the 2.5 region of σ^S exposes subsequently the binding site of ClpX (Becker et al., 1999; Stüdemann et al., 2003).

The only known adaptor protein of ClpA is ClpS, which is encoded by the gene upstream of *clpA*. ClpS exhibits unique features in that it regulates the degradation of three different ClpA substrates. On the one hand, ClpS enhances the degradation of aggregated proteins and enables the recognition of N-end rule substrates (see below) and on the other hand, it inhibits binding and thereby degradation of SsrA-tagged

substrates by ClpAP (Dougan et al., 2002b; Erbse et al., 2006). Interestingly, ClpS mediates the switch from the SsrA-tagged substrate recognition to the aggregated protein species by solely binding to the NTD of ClpA. Binding of ClpS could probably induce a conformational change in ClpA and thus triggering the activity switch.

2.3.1.2. *B. subtilis* adaptor proteins

Up until now, only two adaptor proteins could be identified in *B. subtilis,* both, MecA and YpbH, specifically acting on ClpC (Persuh et al., 2002; Turgay et al., 1998). They do not display similarity to any of the adaptor proteins in *E.coli*. In fact, the interplay of ClpC and its adaptor proteins goes beyond a simple modulation of the substrate selectivity of the partner ATPase. ClpC requires an adaptor protein for all its functions. On its own, ClpC displays only a very low ATPase activity, which is activated up to 300 fold upon interaction with either MecA or YpbH (Schlothauer et al., 2003). Consequently, the interaction with an adaptor is a strict prerequisite for its chaperone and in conjunction with ClpP proteolytic activity. Besides this mechanistic activation, both adaptor proteins enable the recognition of general and specific substrates.

MecA can be dissected into two domains, the N- and C-terminal domains, which are connected by a probably unstructured linker (Persuh et al., 1999). Whereas the CTD is necessary and also already sufficient for the ATPase activation of ClpC, the recognition of substrates is mediated via its N-terminal domain. MecA targets the master regulators for competence development, ComK and ComS as well as unfolded proteins for degradation to ClpCP (Schlothauer et al., 2003; Turgay et al., 1998).

The substrate spectrum of YpbH on the other hand is so far limited to unfolded / aggregated proteins (Schlothauer et al., 2003).

Interestingly, both adaptor proteins become in their cargo-free form a substrate for ClpCP as well. This is probably a part of a regulatory circuit, limiting the presence of an adaptor protein to those times or conditions, which require its targeting abilities. The presence of additional ClpCP substrates such as SpoIIAB and MurAA, for whom the cognate adaptor protein could not be identified yet, supports the assumption that ClpC might employ a number of different and rather specific adaptor proteins (Kock et al., 2004a; Pan et al., 2001).

2.3.1.3. Regulation of adaptor proteins

2.3.1.3.1. By post-translational modification

As stated above (see 2.3.1.1.), RssB binds in its phosphorylated form to σ^S to target it to ClpXP for degradation. Thus, RssB is not only controlled by its cellular concentration, which is near-limiting for σ^S proteolysis (Pruteanu and Hengge-Aronis, 2002), but also by its phosphorylation state. RssB was regarded as orphan response regulator, since no cognate histidine sensor kinase is encoded in an operon together with *rssB*. However, it could recently be shown that the sensor kinase ArcB phosphorylates and thereby modulates RssB (Mika and Hengge, 2005). It was suggested that the redox state of the quinones, which controls the autophosphorylation of ArcB, monitors in addition to oxygen the energy supply as well. Hence, the phosphorylation of RssB by ArcB at low O_2, but high energy levels, reflecting the conditions of the logarithmic growth phase, promotes σ^S proteolysis.

2.3.1.3.2. By anti-adaptor proteins

2.3.1.3.2.1. ComS as anti-adaptor for MecA

The key regulatory proteins for competence development, ComK and ComS, are both substrates for the ClpC adaptor MecA. It could be demonstrated that they display a similar core sequence for the interaction with MecA. The MecA binding motif in ComK (FMLYPK) and in ComS (IILYPR) are located near the C- and N-termini, respectively. Importantly, ComS and ComK compete for binding to MecA and the interaction of ComS with MecA displaces ComK. ComK is thereby protected from degradation, allowing the expression of the competence genes (Prepiak and Dubnau, 2007). The antagonistic function of ComS towards the MecA mediated ComK degradation by ClpCP is a novel feature of the adaptor protein controlled regulated proteolysis. And as such, ComK is regarded as an anti-adaptor protein.

2.3.1.3.2.2. IraP as anti-adaptor for RssB

IraP (inhibitor of RssB activity during phosphate starvation) is the second member of this new class of the anti-adaptor proteins (Bougdour et al., 2006). It was discovered in a genetic screen for modulators of σ^S proteolysis. IraP binds to RssB upon

phosphate starvation and this sequestration of RssB releases σ^S. As a consequence, σ^S is protected from degradation and thus enabled to activate its transcriptional program. IraP differs from ComS in that it is not a substrate for the partner protease (ClpXP). However, the underlying mechanism of the IraP triggered σ^S release from RssB is not known.

In summary, adaptor proteins and especially its post-translational regulation allow a quick and specific response to changing environmental conditions. The functional diversity conferred by adaptor proteins is not restricted to prokaryotes. Several adaptor proteins could be described for p97 and even the poly-ubiquitin recognition signal for the proteasome was discussed as potential equivalent to the bacterial adaptor proteins (Dreveny et al., 2004; Sauer et al., 2004).

2.3.2. Substrate selection

2.3.2.1. Specific recognition motifs

2.3.2.1.1. SsrA tag

Substrates are recognized by a specific motif within the protein sequence. Such a degradation signal can be either already present in the native sequence or alternatively be subsequently added as in case of the SsrA tag. mRNAs, which lack a stop codon probably as a result of a premature transcription termination or the activity of RNases cause an accumulation of stalled ribosomes. The regulatory RNA SsrA rescues these ribosomes in a process termed trans-translation: SsrA acts as a tmRNA and enables (i) the release of a truncated mRNA lacking a stop codon, which is a prerequisite for the binding of release factors to the ribosome and (ii) the addition of an 11 aa tag (AANDENYALAA), which is recognized by ClpX, the ClpX adaptor protein SspB and ClpA and subsequently lead to the degradation. However, *in vivo* the degradation of SsrA tagged substrates relies solely on ClpXP.

2.3.2.1.2. LCN motif

A unique degradation motif could be identified at the very C-terminus of the *B. subtilis* protein SpoIIAB (Pan et al., 2001). The terminal three aa (LCN) are essential for its degradation by ClpCP. Interestingly, SpoIIAB is the one and only protein

harboring this motif at its C-terminus, suggesting a very specific recognition mode. However, ClpC does not recognize this motif on its own, raising the question of the cognate adaptor protein enabling the targeting to and degradation by ClpCP.

2.3.2.1.3. MecA recognition motif in ComK and ComS

ComK and ComS share a motif, which serves as binding site for MecA: [FI]- [MI]-L-Y-P-[RK] (Prepiak and Dubnau, 2007). This motif is not only required for the targeting of both proteins to ClpCP, but also for the regulation of the competence genes (see 2.3.1.3.2.).

2.3.2.1.4. N-degron

A further recognition mode that is conserved throughout pro- and eukaryotes is the N-end rule pathway. Here, the stability of a certain protein is defined by its N-terminus. The N-terminal residues are classified as stabilizing and destabilizing aa. The recognition determinants (N-degrons) are organized hierarchically in that the destabilizing residues are further dissected into primary, secondary and tertiary destabilizing aa. The generation of destabilizing N-termini can be achieved (i) by removal of the N-terminal methionine via the methionine aminopeptidase, (ii) by an endoproteolytic cleavage reaction to uncover ´cryptic´ N-degrons (see 2.3.2.2.), (iii) by an enzymatic modification e.g. deamination and (iv) by recruiting and subsequent conjugation of a primary to a secondary destabilizing aa e.g. by covalent addition of L/F to the secondary residues R and K in *E. coli* (Tobias et al., 1991). In eukaryotes, the recognition of the N-degron is mediated by N-recognins, a class of E3 ligases, which label the N-end rule substrates with ubiquitin to enable its degradation by the 26S proteasome. The absence of the ubiquitin system suggested a different substrate selection mode in prokaryotes. In addition, bacteria do not encode a N-recognin homolog. A limited homology, which is restricted to the substrate interaction site, could only be observed for the *E.coli* ClpS adaptor protein. The sequence of the recognition site is enriched in acidic and hydrophobic aa and thus complementary to the destabilizing residues of the substrate. It could be demonstrated that ClpS binds directly to the N-end rule substrates and delivers them subsequently to ClpAP for its degradation (Erbse et al., 2006). However, the actual contribution of ClpS is currently

a controversial subject (Wang et al., 2007). And a definite bacterial N-end rule substrate could be still not identified, yet.

2.3.2.1.5. Classification of distinct recognition motifs for ClpX

The relatively high number of already identified endogenous substrate proteins for *E. coli* ClpX by a proteomic approach, allowed a first classification of the recognition motifs: Native degradation signals for ClpX are preferably located at the very N- or C-terminus and could be subdivided into five classes: three N-terminal motifs: N1: T-X-K-[ILV], N2: M-[RK]-h_3-X_5-h and N3: h-X-polar-X-polar-X-[RK]-polar and two C-terminal motifs: C1 is the SsrA motif and C2: [RH]-X-[RK]-K-h-[AS] (Flynn et al., 2003).

2.3.2.2. Cryptic recognition sites

In addition, degradation motifs can also be cryptic e.g. being masked by an interaction partner, buried by its fold or simply located in the middle of the protein sequence instead of at one of the termini. Such cryptic recognition signatures require for instance an initial cleavage event as demonstrated for the anti-σ factor of the ECF σ^W, RsiW, in *B. subtilis*. σ^W becomes activated upon cell envelope stress, induced by alkaline treatment or phage infection and is similarly activated as σ^E in *E.coli* involving a membrane located proteolytic process termed RIP (regulated intramembrane proteolysis) (Alba and Gross, 2004; Schöbel et al., 2004). The intramembrane protein, RsiW is processed by two iCLiPs (intramembrane cleaving protease), YpdC and RasP, prior to the degradation of its C-terminal domain by ClpXP. The second cleavage event by the site-2 protease, RasP, uncovers the cryptic SsrA-like ClpX recognition motif (Heinrich and Wiegert, 2006).

Another example of an initial processing enabling the recognition by ClpX is the damage induced autocleavage reaction of LexA during the SOS response (Flynn et al., 2003; Neher et al., 2003a). In brief, the RecA induced autocleavage reaction of LexA produces two fragments bearing a ClpX recognition motif. The N-terminal fragment ends V82-A83-A84 and is thus highly similar to the last three aa of the SsrA consensus sequence or motif C-1, respectively (see above 2.3.2.1.5.). Although the

binding motif (L112-L113-R114-V115-S116) of the C-terminal fragment is not located at the new terminus, it is recognized upon the cleavage event. Presumably, the cleavage reaction leads to a remodeling of the C-terminal fragment and thus exposing this motif for the recognition by ClpX (Neher et al., 2003a).

Thus, the uncovering of cryptic signals by unfolding, subunit dissociation and preceding cleavage reactions allows even more variety in the control mechanisms for the regulated proteolysis by AAA+ proteases.

2.4. Summary of the different modes of substrate recognition

The recognition of a particular substrate can be achieved in multiple ways. Both, the Clp ATPase and the adaptor protein can contribute to the recognition. Especially the NTD of the AAA+ protein plays a central role. First, the NTD can interact directly with the substrate protein as shown for ClpA and ClpX recognizing the SsrA-tag. Alternatively, the NTD could serve as binding site for adaptor proteins (see 2.3.1.), which then allows two different recognition modes: (i) the adaptor protein enables the recognition of the substrate in the first place e.g. ClpA/ClpS recognizing the N-degron and (ii) both, adaptor protein and Clp ATPase contribute to substrate recognition and binding. In the latter case, the adaptor protein could tether substrate and ATPase and thus enhancing its degradation kinetics as demonstrated for ClpX/SspB and SsrA-tagged proteins. Or on the other hand, the substrate binding by adaptor protein and Clp ATPase could occur in a sequential manner. Thus, the interaction of the Clp ATPase would require the preceding binding of the adaptor protein to the substrate in order to expose the recognition site for the ATPase as shown for ClpX/RssB and σ^S. Related to that, recognition sites, usually located in close proximity of one of the termini (2.3.2.1.5.), can be masked and therefore require an initial cleavage / unfolding to generate or expose the recognition site (2.3.2.2.). Finally, adaptor proteins can be subject to regulation in particular on the post-translational level, thus allowing a quick adaptation to the changing environment. So far, it could be shown that the activity of adaptor proteins could be regulated by phosphorylation and by the antagonizing activity of an anti-adaptor protein (2.3.1.3.).

Taken together, this wide array of substrate recognition modes accounts for the specificity of regulated proteolysis and allows a control and interference at various levels.

2.5. Aims of this work

The variety of substrates for ClpC suggests that the number of yet to be identified adaptor proteins will increase. Strikingly, ClpC adaptor proteins fulfill a dual role: (i) they mediate the substrate recognition and (ii) they activate the otherwise non-functional ClpC.

The study of Schlothauer et al. (2003) showed that the ATPase of ClpC and hence its disaggregation and refolding activity and the proteolytic function of ClpCP relies on the presence of an adaptor protein. These data lead to the question of the principle of the adaptor mediated activation. Using a variety of biochemical and biophysical methods, it could be demonstrated in this work that the general activation of ClpC is based upon the ability of the adaptor to facilitate the assembly of ClpC into an active substrate-recognizing hexamer (Kirstein et al., 2006). Once assembled, hexameric ClpC is enabled to assist the oligomerization and thereby activation of the otherwise monomeric ClpP (see 3.2.).

Interestingly, the observed adaptor dependence for *B. subtilis* ClpC and its adaptor proteins is not conserved among the ClpC homologs. Although ClpC from photobiontic organisms share 90 % similarity with *B. subtilis* ClpC, its basic function such as ATPase and remodeling activities do not depend on the presence of an adaptor protein (Andersson et al., 2006). Thus, adaptor proteins exhibit a wide functional divergence even for the homologous partner ATPase (see 3.3.).

In addition to the MecA paralog YpbH, a third, not MecA related, ClpC adaptor protein, McsB, could be identified (Kirstein et al., 2007). McsA and McsB were previously described as modulators of CtsR, the repressor of the class III heat shock genes (Krüger et al., 2001). McsB could be characterized as tyrosine kinase, which utilizes a domain with homology to guanidino phosphotransferase domains (Kirstein et al., 2005). In this study it could be demonstrated that McsB needs the activation by McsA and that McsA, McsB and CtsR form a ternary complex, resulting in the disruption of the CtsR-DNA-interaction and the subsequent phosphorylation of CtsR. The phosphorylated form of CtsR is then subject for degradation by ClpCP, which requires the adaptor targeting function of McsB (Kirstein et al., 2007; Kirstein et al., 2005). Interestingly, the adaptor function of McsB is coupled to its kinase activity. McsB can therefore be considered as a new adaptor protein regulated by phosphorylation analogous to RssB (see 3.4.).

The characterization of McsB led to the question, whether all three adaptor proteins adopt the same mode of ClpC activation. Indeed, the identified binding sites on ClpC (NTD and Linker domain) are shared, suggesting a general adaptor mediated activation mechanism of ClpC. The interaction of the different adaptor proteins with ClpC might therefore not occur simultaneously. Instead, it was assumed that the adaptor proteins compete for binding to ClpC. Indeed, we could demonstrate that McsB and MecA compete for substrate targeting to ClpCP *in vitro* (Kirstein et al., 2007).

The publications are presented and discussed in their chronological order of publication.