Published in final edited form as:

Subcell Biochem. 2013 ; 66: 297–327. doi:10.1007/978-94-007-5940-4_11.

Archaeal Proteasomes and Sampylation

Julie A. Maupin-Furlow

Department of Microbiology and Cell Science, University of Florida, Gainesville, FL 32611-0700, USA

Julie A. Maupin-Furlow: jmaupin@ufl.edu

Abstract

Archaea contain, both a functional proteasome and an ubiquitin-like protein conjugation system (termed sampylation) that is related to the ubiquitin proteasome system (UPS) of eukaryotes. Archaeal proteasomes have served as excellent models for understanding how proteins are degraded by the central energy-dependent proteolytic machine of eukaryotes, the 26S proteasome. While sampylation has only recently been discovered, it is thought to be linked to proteasome-mediated degradation in archaea. Unlike eukaryotes, sampylation only requires an E1 enzyme homolog of the E1-E2-E3 ubiquitylation cascade to mediate protein conjugation. Furthermore, recent evidence suggests that archaeal and eukaryotic E1 enzyme homologs can serve dual roles in mediating protein conjugation and activating sulfur for incorporation into biomolecules. The focus of this book chapter is the energy-dependent proteasome and sampylation systems of Archaea.

Introduction

Archaea are one of three major lineages of life which share deep evolutionary roots with eukaryotes. Although typically single-celled, many Archaea thrive in conditions once thought to be uninhabitable to life such as temperatures above 100°C, saturating salt and extreme pH. Thus, Archaea provide an exciting opportunity to examine how proteolytic systems can function in maintaining the quality of proteomes when cells are growing at physical conditions considered suboptimal for protein folding and stability.

Although the hydrolysis of peptide bonds is exergonic (releasing energy), many proteolytic machines (e.g. the proteasome) require energy for protein degradation. Of the various types of proteases, it is often the energy-dependent proteases that are important in regulating central processes of the cell such as protein quality, cell division and survival after exposure to stress [1, 2]. The energy-dependence of the proteolytic system adds assurance that the appropriate protein substrate has been selected prior to its destruction, which would otherwise come at a high energy cost.

Energy-dependent proteases (e.g. Lon, Clp, FtsH, HslUV and the proteasome), while not necessarily conserved in primary amino acid sequence, share a common protein architecture of self-compartmentalization [3]. In general, the proteolytic active-sites of the protease are sequestered within a chamber that has narrow openings for substrate entry. These openings are often gated to reduce the non-specific entry of proteins into the chamber [4–7]. Typically, the chamber alone does not degrade folded proteins but requires association with regulatory ATPases (either a protein domain or separate protein complex). The ATPases

couple the hydrolysis of ATP to the unfolding and the translocation of the substrate protein, so that the substrate protein can access the proteolytic active sites lining the central chamber of the protease.

The ATPase regulators of energy-dependent proteolysis belong to the AAA + (ATPases associated with a variety of cellular activities) superfamily [8, 9]. These AAA + proteins form hexameric rings that associate with a compartmentalized cylindrical peptidase. In some cases, the peptidase forms a hexameric ring (e.g. HslV), while in other cases the peptidase exhibits a seven-fold symmetry. Hence, symmetry mismatch can exist between the ATPase and protease components. For example, the hexameric ATPases (ClpA, ClpC and ClpX) associate with ClpP (which is composed of two heptameric rings). Similarly, the proteasomal ATPases (ARC/Mpa of actinobacteria, PAN of Archaea, and Rpt1-6 of eukaryotes) all form hexamers and interact with a 20S core particle (CP), which has seven-fold symmetry. Energy-dependent proteases such as Lon and FtsH differ from most AAA + proteases, in that both the ATPase and proteolytic domains are located on the same polypeptide chain [10] (see also accompanying reviews [11, 12]).

Archaeal Proteasomes

Similarly to eukaryotes, all Archaea are predicted to use proteasomes as one of their major systems for energy-dependent proteolysis [13]. All species of Archaea with sequenced genomes are predicted to encode a 20S CP of α - and β -type subunit composition as well as an AAA + regulator (PAN and/or Cdc48) thought to associate with the CP and mediate protein degradation [13, 14]. Archaea also contain the membrane bound B-type Lon protease [15, 16]. However in contrast to bacteria, Archaea lack homologs of the membrane-spanning FtsH protease and, in most cases also lack HslUV, Clp and the A-type Lon proteases [15, 16]. Like other bacteria, the actinobacteria (including species from *Mycobacterium*, *Rhodococcus* and *Frankia*) encode many energy-dependent proteases including the proteasomal CPs and its associated AAA + components termed ARC or Mpa, in addition to Lon, Clp and FtsH proteases, but do not encode the HslUV protease [17–20]. Overall, the distribution of energy-dependent proteases in Archaea is more like eukaryotes than either bacteria or bacterial-like organelles of eukaryotes.

Proteasomal Core Particles (CPs)

Proteasomes are composed of a 20S CP that is required for the hydrolysis of unfolded proteins and small peptides [21]. The CP is formed from structurally related α - and β -type subunits that are arranged into four-stacked heptameric rings (Fig. 11.1). The two outer rings are composed of α -subunits, and the two inner rings are of β -subunits. These subunits are arranged to form a cylindrical particle with openings on both ends of the particle, gated by the N-termini of the α -subunits. The gates open into a central channel that connects three interior chambers (two antechambers and one central chamber). The central chamber is lined with the proteolytic active sites that are formed by the N-terminal threonine residues of the β -subunits and are exposed after cleavage of the β propeptide during CP assembly. In contrast to eukaryotic CPs, which are formed from seven different α -subunits ($\alpha 1$ to $\alpha 7$) and seven different β -subunits ($\beta 1$ to $\beta 7$) that assemble into a 28-subunit complex of dyad symmetry. Core particles from Archaea and actinobacteria are typically assembled from one to two different α -subunits and one to two different β -subunits.

Proteolytic Active Sites of Proteasomes

The proteolytic active sites of proteasomes, sequestered within the CP, are of the N-terminal nucleophile (Ntn) hydrolase superfamily [24–26]. This superfamily includes penicillin acylases, aspartylglucosaminidases and other hydrolytic enzymes. While Ntn hydrolases are

not necessarily conserved in primary sequence, these enzymes hydrolyze amide bonds and undergo autocatalytic post-translational processing to expose an N-terminal residue (Thr, Ser or Cys) that is used as nucleophile (the side chain OH or SH group) and proton donor (the N α -amino group) in a structurally related active site configuration. The Ntn hydrolase reaction is initiated when the Ntn residue donates the proton of its side chain to its own α -amino group for attack of the carbonyl carbon of the substrate (Fig. 11.2). A negatively charged tetrahedral intermediate is formed that is stabilized by hydrogen bonding to the oxyanion hole. The Ntn α -amino group completes the acylation step by donating a proton to the nitrogen of the substrate bond undergoing hydrolysis. While part of the cleaved substrate (the portion with the R or P1 group) is released, a covalent bond is formed between the nucleophile of the Ntn and the remaining substrate (with the R' or P1' group) that must be deacylated for product release. Deacylation is initiated when the hydroxyl group of water attacks the carbonyl carbon of the covalent acyl-enzyme intermediate. The Ntn α -amino group accepts a proton from water and stabilizes the negatively charged enzyme intermediate. Reprotonation of the nucleophile by the Ntn α -amino group completes this reaction.

While the mechanism of peptide hydrolysis by the CP active site resembles serine proteases, only one amino acid (Thr1) serves as the catalytic center, rather than the three amino acids (Asp-His-Ser) that constitute the catalytic triad of serine proteases [27]. Both the nucleophile (hydroxyl side chain) and the base (α -amino group) of the CP proteolytic reaction exist in the same amino acid (Thr). Like β -lactamase and some esterases, there is no acid component in the CP active site [28]. Serine can substitute for threonine as the Ntn active site in CPs (T1S variants) for the hydrolysis of peptide amide substrates such as N-succinyl-Leu-Leu-Val-Tyr-7-amido-4-methyl coumarin (Suc-LLVY-amc). However, CP T1S variants are severely impaired in their ability to cleave longer peptides, revealing the importance of the threonine side chain methyl group in active site conformation [29, 30].

Cleavage of N-Terminal Propeptide to Expose Proteolytic Active Site

Proteasomal active sites are formed through intramolecular autolysis of β -type precursor proteins during CP assembly [31]. The CP β -subunits are typically translated as precursor proteins with N-terminal propeptides that are removed during formation of the CP. The β -subunits remain monomeric, unprocessed and inactive until assembled with α -subunits into the CP. In most cases, the propeptide of the β -subunit is autocatalytically removed at a conserved G $\downarrow$ T motif (where $\downarrow$ represents the site of peptide bond cleavage) to generate a mature β -subunit that contains an active N-terminal threonine. However, some β -subunits that assemble into CPs are not cleaved at conserved G $\downarrow$ T motifs [32, 33]. For example, the eukaryotic β 7 and β 6 precursors are cleaved at N $\downarrow$ T and H $\downarrow$ Q sites, respectively, resulting in the removal of N-terminal propeptides and formation of inactive β subunits in CPs. Likewise, the eukaryotic β 3 and β 4 subunits are not cleaved but are assembled into CPs as inactive subunits.

The position and residues surrounding the cleavage site of the propeptide of proteasomal β -subunits is predicted to be highly diverse among archaea. Like eukaryotes and actinobacteria, Archaea are known to synthesize β -type precursor proteins that are cleaved at G $\downarrow$ T motifs resulting in the removal of N-terminal propeptides of variable length (6–49 residues) and the formation of mature β -subunits with active CPs. This knowledge is based on N-terminal sequencing of CP subunits purified from Archaea of the phylum Euryarchaeota (i.e., *Thermoplasma acidophilum*, *Methanosarcina thermophila*, *Haloferax volcanii* and *Pyrococcus furiosus*) [34–37]. Similarly, most archaea are predicted to encode at least one β -type precursor protein that is cleaved at a G $\downarrow$ T motif to release an N-terminal propeptide and generate an active β -type subunit. Interestingly, select Crenarchaeota (i.e.,

Thermoproteus and *Pyrobaculum* species) and Candidatus *Caldiarchaeum subterraneum* of the novel division Aigarchaeota have β -subunit homologs devoid of propeptides (only an initiator methionine residue, which may be removed by a methionine aminopeptidase, is predicted to precede the conserved active site threonine residue). Furthermore, most Crenarchaeota are predicted to encode two different β -type proteins with at least one of these proteins harbouring an N-terminal extension that lacks the classical G↓T cleavage site. Whether or not these latter β -subunits associate in CPs and form active sites remains to be determined.

The propeptides of β -type precursors (' β -propeptides') can modulate the function and assembly of proteasomal CPs. The β -propeptides of yeast protect the N-terminal threonine active sites of the β -subunits from N^α-acetylation and promote CP assembly [38]. In actinobacteria, the β -propeptide of *Rhodococcus erythropolis* is important for the formation of CPs, while the β -propeptide of *Mycobacterium tuberculosis* mediates temperature-dependent inhibition of CP assembly [41, 42]. By contrast to the actinobacteria and yeast, the β -propeptides of archaea appear to be dispensable for CP assembly. Archaeal β -subunits can be produced in recombinant *E. coli* without an N-terminal propeptide ($\beta\Delta$ pro) and then reconstituted with α -type subunits *in vitro* to form active CPs (i.e., *Methanocaldococcus jannaschii* CPs) [43]. Likewise, archaeal CPs can form active complexes when $\beta\Delta$ pro variants are co-expressed with α -type subunits in recombinant *E. coli* (i.e., *T. acidophilum* and *M. thermophila* CPs) [30, 31]. By modulating the ionic strength of the buffer, haloarchaeal CPs (of *H. volcanii*) can also be disassembled into monomers and reconstituted into active CPs (based on gel filtration and activity assays) [37].

While archaeal β -propeptides are not required for the generation of active CPs *in vitro* or when produced in *E. coli*, β -propeptides are likely to be important in the proper formation of proteasomes within archaeal cells. In support of a function for β -propeptides in Archaea, genetic deletion of the 49-residue propeptide of the *H. volcanii* β -subunit results in undetectable levels of this protein ($\beta\Delta$ pro) (either by immunoblot or His-tag purification) when expressed in its native host *H. volcanii* (Kaczowka and Maupin-Furlow, unpublished). In contrast, *H. volcanii* $\beta\Delta$ pro is readily detected and purified (using a His-tag) when expressed in a heterologous host, *E. coli* [44]. Likewise, the β -subunit is synthesized and purified at high levels when the β -propeptide is added to the *H. volcanii* expression system [44]. Whether or not these findings are unique to *H. volcanii* remains to be determined.

Notably, the N-terminal propeptides (typically over 30 residues) of the β -subunits of haloarchaea are predicted to be the longest amongst all Archaea. Haloarchaea are unusual in that they maintain osmotic balance with hypersaline environments by accumulating molar concentrations of KCl in their cytoplasm [45]. In order to cope with these harsh conditions, haloarchaea synthesize proteins with a high negative surface charge, which makes their proteins more soluble and flexible at high concentrations of salt compared to non-halophilic proteins [46]. Thus, the extended N-terminal propeptide of β -subunits may be needed to facilitate folding and assembly of the CP in the 'harsh' cytosolic environment found in haloarchaeal cells. Alternatively, given that N^α-acetylation is quite common in haloarchaea [47–49], the N-terminal propeptide of β -subunits may prevent N^α-acetylation of the active site threonine residue as has been observed in yeast.

CP Active Site Number and Subunit Complexity

While all proteasomes to date have 14 α -type and 14 β -type subunits assembled into each CP, the ratio of β -type subunits with N-terminal threonine active site residues per CP varies among the three domains of life. These differences in active site number do not appear to

influence the overall size of peptide products generated nor the rate of peptide bond hydrolysis, based on comparison of archaeal and eukaryotic proteasomes [50, 51].

In general, proteasomal CPs from eukaryotes have less proteolytic active sites per particle than CPs from either archaea or actinobacteria. Eukaryotic CP subtypes have six active sites per particle formed by three of the seven different β -type subunits per heptameric β -ring [52]. For example, within the housekeeping CP of eukaryotes only the $\beta 1$, $\beta 2$ and $\beta 5$ subunits are proteolytically-active (the $\beta 3$, $\beta 4$, $\beta 6$ and $\beta 7$ subunits are inactive). In response to cytokines and microbial infections, $\beta 1$, $\beta 2$ and $\beta 5$ are replaced by $\beta 1i$, $\beta 2i$ and $\beta 5i$, which are also active and alter CPs to generate peptide products optimized for MHC class I loading in hematopoietic cells [53].

In contrast to eukaryotes, all archaeal (and actinobacterial) CPs purified to date have 14 active sites per particle and are typically purified as a cylindrical ~600–700 kDa complex of a single type of β subunit that associates with a single type of α subunit (Fig. 11.1). Thus, most archaeal CPs have identical proteolytic active sites with some determined by N-terminal sequencing of the β -subunits [34, 36, 37, 54]. Examples of CPs purified from Archaea (or recombinant *E. coli* expressing these archaeal genes) composed of only a single type of α - and β -subunits, include CPs from *Thermoplasma acidophilum* [55], *Thermoplasma volcanium* [56], *Methanosarcina thermophila* TM-1 [30, 36], *Methanosarcina mazei* [57], *Methanocaldococcus jannaschii* [43], and *Archaeoglobus fulgidus* [58].

Although Archaea with CPs composed of only a single type of α - and β -subunit have received most attention, as simplified models for biochemistry, the majority of Archaea are predicted to encode multiple α - and/or β -subunit homologs based on genome sequence. Most Crenarchaeota encode single α -type and two different β -type subunit homologs, while Thaumarchaeota, Korarchaeota and select families of Euryarchaeota (*Halobacteriaceae* and *Thermococcaceae*) commonly encode multiple α -type and/or β -type subunit homologs. Interestingly, the recently described Candidatus *Caldiararchaeum subterraneum* is predicted to encode three different β -subunits and a single α -subunit. Many of the Archaea with two or more coding sequences for β -subunits harbour one gene that encodes a β -subunit devoid of the classical G↓T cleavage motif, suggesting that the protein is synthesized either as an inactive β -subunit or has a propeptide cleavage site that differs from the CPs characterized from Euryarchaeota and eukaryotes.

Archaeal CPs composed of a single type of α subunit and two different β -type subunits (both containing a G↓TTT cleavage motif) have been characterized at the biochemical level. Initially, CPs of only a single α/β composition were purified from the archaeon *Pyrococcus furiosus* [34]. However, DNA sequencing of the complete genome of this hyperthermophile predicted that two β -type subunits ($\beta 1$ and $\beta 2$) and a single α -type subunit [59] existed in this species. Further analysis by microarray revealed that the level of the $\beta 1$ transcripts were upregulated after cells were exposed to heat shock, which might explain why only one β -type protein was detected in CPs purified from *P. furiosus* grown under normal (non-stressed) conditions [60]. To examine the function of CPs which contain multiple β -type proteins, the α -type and both β -type proteins were produced in *E. coli* and reconstituted at different temperatures *in vitro* [60]. Using this approach, an increased ratio of $\beta 1$ to $\beta 2$ was observed in CPs assembled at higher temperatures, which correlated with more thermostable properties. Surprisingly, the α and $\beta 1$ proteins could not be reconstituted into CPs or form particles with peptide hydrolyzing activity independent of $\beta 2$. Based on these results, the *P. furiosus* $\beta 1$ is thought to stabilize $\alpha 2/\beta$ CPs at high temperature and to influence the biocatalytic properties of the proteasome.

CP subtypes, of different α - and β -type subunit composition, have also been directly purified from an archaeon. The haloarchaeon *H. volcanii* encodes a single β -type and two α -type ($\alpha 1$ and $\alpha 2$) CP subunits [37]. Not surprisingly, the β -type gene is essential for growth and the presence of one of the two α -type genes is required for growth (based on conditional mutation) [61]. All three of the different CPs ($\alpha 1 \beta$, $\alpha 2 \beta$ and $\alpha 1 \alpha 2 \beta$) have been purified from *H. volcanii* [37, 44, 62]. Each of these different CP subtypes are active in the hydrolysis of peptides and unfolded proteins with the $\alpha 1 \alpha 2 \beta$ -proteasome apparently asymmetric based on the detection of homooligomeric rings of $\alpha 1$ and $\alpha 2$ [44, 62]. The population of CP subtypes is likely altered during growth based on the finding that levels of $\alpha 2$ increase several-fold as cells transition to stationary phase, while $\alpha 1$ levels remain relatively constant [63]. Although it remains to be established, differences in α -type subunit composition are thought to influence the type of regulator that associates with the CPs in *H. volcanii*. Interestingly, in *H. volcanii* the CPs (and PAN ATPases) are also modified post-/co-translationally (phosphorylation, N^{α} -acetylation and methylation) suggesting an additional layer of regulation in modulating the populations of proteasomes in archaeal cells [64–66].

Proteasomal Regulators

Eukaryotic proteasomal CPs co-purify with a variety of accessory proteins including ATPase regulators, such as the 19S regulatory particle (RP), and non-ATPase regulators such as Blm10/PA200 and the 11S regulators PA28, PA26 and REG [67]. Of these regulators, the 19S RP can associate with either end of the CP cylinder to form the 26S proteasome that catalyzes the ATP-dependent degradation of proteins (Fig. 11.3) [22]. Subunits of the RP are grouped into two major categories including the RP non-ATPase subunits (Rpn) and the RP AAA + (triple-A+) subunits (Rpt) of the ATPases associated with various cellular activities (AAA) family within the AAA + superfamily [68]. In yeast, RPs are dissociated into base and lid subcomplexes by deletion of the gene encoding the Rpn10 subunit [69] (Fig. 11.3). The base subcomplex is composed of Rpn1-2, Rpn13 and six different Rpt subunits (Rpt1-6) that form a hexameric ring in direct contact with the outer rings of the CP complex [69].

While archaeal CPs have yet to be purified in complex with regulatory proteins from its native host, archaeal CPs can associate with ATPase regulators *in vitro* (Fig. 11.3). Studies of this association were initiated after release of the first archaeal genome sequence (*M. jannaschii*), which was predicted to encode a close homolog of the Rpt1-6 subunits of eukaryotic 26S proteasomes [70]. The archaeal Rpt homolog (termed PAN for its role as a proteasome-activating nucleotidase) was synthesized and purified from recombinant *E. coli* and shown to be required for the degradation of substrate proteins (e.g., casein) by CPs including those of *T. acidophilum* [71] and *M. jannaschii* [43]. *M. jannaschii* (Mj) PAN forms a relatively simple homohexameric ATPase complex and, thus, has served as an ideal model to understand how Rpt-like proteins function in the degradation of proteins by proteasomes (see later). While all Archaea encode proteasomal CPs based on genome sequence, many do not encode Rpt/PAN homologs including archaea of the Thaumarchaeota, Korarchaeota, *Thermoplasmata* class of Euryarchaeota, Aigarchaeota (*Candidatus Caldiarchaeum subterraneum*), and Thermoproteales order of Crenarchaeota. Whether or not ATPases beyond PAN can regulate these archaeal CPs remains to be determined. Members of the Cdc48/VCP/p97 subfamily of AAA + proteins are universally distributed among archaea, and, in eukaryotes, these Cdc48-type proteins are associated with the ubiquitin-proteasome system in endoplasmic reticulum-associated protein degradation (ERAD) [72, 73]. The *T. acidophilum* VCP is an archaeal Cdc48-type ATPase that has been purified and shown to function as an ATP-dependent unfoldase and stimulate protein degradation using mutant CPs with open gates [74] (see below). More recently, VCP has

been shown to associate with CPs and stimulate peptide and protein degradation [75]. Thus, archaeal CPs appear to be regulated by multiple types of AAA + proteins.

Gated Openings of Proteasomal Core Particles (CPs)

The proteolytic active sites of proteasomes are sequestered within the interior of the chambered CP, which has gated openings to control substrate entry. The first 3-D structure of the proteasome, which provided insight into its compartmentalization, came from the CP of the archaeon *T. acidophilum* (TaCP) that was produced in *E. coli* [75, 76]. The TaCP structure revealed narrow (13 Å) openings on each end of the cylindrical particle (113 Å diameter by 148 Å length) that connected a central channel leading to an interior chamber lined with proteolytic active sites. Subsequent, X-ray structures of yeast and bovine CPs revealed particles with no apparent openings on the ends for substrates to access the proteolytic active sites or for products to exit the chambered particle [77, 78]. The N-terminal tails of α -type subunits within the eukaryotic CP structures appeared to block the channel openings and, thus, were thought to gate substrate entry (Fig. 11.4).

With the yeast CP as a model, genetic modifications to residues associated with the gate and structural analysis of CPs associated with regulators thought to open the gate provided evidence for a gating mechanism in eukaryotes. In particular, genetic deletion of the N-terminal tail of the yeast $\alpha 3$ ($\alpha 3\Delta N$) was found to open the channel and derepress the peptide hydrolysis mediated by CPs [79]. The $\alpha 3\Delta N$ mutation also increased the size of CP-generated peptide products suggesting CP gating controls product release in addition to substrate entry [80]. Binding of regulators to CPs *in vivo*, was thought to relieve the inhibition mediated by the α -subunit tails and open the CP gates [79]. Indeed, deletion of the gate (N-terminal tail of $\alpha 3$) in mutant proteasomes containing an active site mutation in the ATPase subunit (Rpt2) was sufficient to overcome defects in peptide hydrolysis [80]. Likewise, crystal structures of yeast CPs in complex with the non-ATPase activator PA26 revealed, that loops in PA26 could induce conformational changes in the α -subunit tails to open the gate, separating the CP interior from the intracellular environment [81].

Although gated openings were not evident in the early archaeal CP crystal structures (most likely due to the disordered structure of the α subunit N-terminal tail) [58, 76], several lines of evidence now support a gating mechanism for archaeal CPs. For example, deletion of the N-terminal tail of the *T. acidophilum* α -subunit ($\alpha\Delta N$) reduces the central mass of heptameric rings formed by the α -subunit in electron micrographs, supporting the localization of these tails near the central channel of mature CPs [82]. Increased rates of unfolded protein hydrolysis are also observed for TaCPs containing $\alpha\Delta N$ -subunits compared to wild-type suggesting that deletion of the N-terminal residues of the α -subunit generates a “gateless” CP [82, 83]. The recent crystal structure of a CP assembled from *M. jannaschii* α - and β -subunits produced in *E. coli* also provides evidence that an archaeal CP can be stabilized in closed gate conformation [84]. Furthermore, addition of specific peptides (based on the C-terminal tail of MjPAN) causes a conformational change in TaCPs that is consistent with gate opening (Fig. 11.5) (see later) [85, 86]. Interestingly, the N-terminal tails of the α -subunits that form the gates can be highly dynamic and extend inside and outside the CP cylinder, based on methyl-transverse relaxation optimized nuclear magnetic resonance spectroscopy experiments of TaCPs [87].

Functional Association of Proteasomal ATPases and CPs

While proteasomal CPs have yet to be purified in association with defined regulatory proteins from archaeal cells, MjPAN (an archaeal Rpt homolog) can be reconstituted with CPs *in vitro*. Formation of MjPAN:TaCP complexes has been detected by immunoprecipitation, electron microscopy and surface plasmon resonance. These complexes

form when recombinant MjPAN and TaCP are incubated in the presence of ATP or a non-hydrolyzable ATP analogue (AMPPNP or ATP γ S), but not when incubated with ADP or in the absence of nucleotide [83]. Thus, ATP binding is required for MjPAN:TaCP association. In electron micrographs of negatively stained complexes, MjPAN is associated as a two-ringed structure that caps either one or both ends of the TaCP [83]. The outer ring of MjPAN (distal to the TaCP) is thought to correspond to the N-terminal domain of the protein based on analogy to a related AAA + complex from bacteria, HslU [88–90]. The C-terminal domain of MjPAN is important for docking with CPs and opening the CP gates (*see below*).

Docking of the Proteasomal ATPase (PAN) to the CP

Proteasomal ATPases such as MjPAN have a C-terminal tail that is essential for docking with CPs and opening the CP gates. A series of elegant studies focused on the association of the archaeal MjPAN with CPs facilitated this discovery [85, 86, 91]. These studies on MjPAN were guided by similarities to the bacterial HslUV protease and eukaryotic non-ATPase 11S regulators (PA28 and PA26) in association with proteasomal CPs. While somewhat controversial at the time, the C-terminal domain of HslU (the AAA + component of the HslUV protease) was thought to be required for association with HslV [90, 92]. Likewise, the extreme C-terminus of the 11S regulator PA28 was known to be required for association with CP [93], and structural analysis revealed that the C-terminal tail of PA26 (a related 11S regulator) docked into pockets of the CP, formed by adjacent α -subunits [81, 94].

Biochemical studies using MjPAN, demonstrated that a tripeptide motif (HbYX, where Hb=a hydrophobic amino acid and x=any amino acid) located at the extreme C-terminus of the proteasomal ATPase was important for proteasome function. Removal of the MjPAN C-terminus dramatically reduced its ability to stimulate the TaCP-mediated degradation of a nine-residue long peptide substrate [83]. Importantly, this MjPAN variant retained wild type ATPase activity, unfoldase activity and ability to stimulate hydrolysis of small peptides by TaCPs [83]. Systematic mutation of the C-terminal residues of MjPAN demonstrated that both the length of this tail and penultimate tyrosine residue of the HbYX motif was absolutely essential for MjPAN to stimulate the TaCP-mediated degradation of LFP and, thus, CP gate opening (where LFP represents a nine residue peptide substrate) [83]. To further investigate this activation step, a tryptophan residue was introduced into either the Hb or X position of the HbYX motif in MjPAN, and analyzed by tryptophan fluorescence and polarization [83]. This approach provided further evidence that MjPAN could associate with TaCP and suggested that the C-terminal tail of MjPAN moves from an aqueous to a hydrophobic environment upon association. Interestingly, peptides based on the C-terminus of MjPAN could compete with MjPAN for TaCP binding and serve as “gate openers” based on their stimulation of TaCP-mediated hydrolysis of LFP [83]. In contrast, peptides based on the C-terminus of the non-ATPase 11S regulator PA26 were unable to open the TaCP gate but could compete with MjPAN for TaCP binding [83]. These findings were consistent with studies, which revealed docking of the C-terminal tail of PA26 to the CP was associated with binding of a distant activation domain of PA26 to a region of the α -subunit N-terminal tails that form the CP gate [81, 94]. Thus, MjPAN and PA26 appear to use a different mechanism to open the CP gate.

Structural studies have been performed to further understand how proteasomal ATPases with C-terminal HbYX motifs open CP gates. In particular, cryoelectron microscopy (cryoEM) has been used to identify sites in the TaCP that are bound by peptides mimicking the C-terminal tail of MjPAN and to determine the structures of the TaCP gate in the “closed” and “open” forms induced by these peptides (Fig. 11.5) (gated and ungated

openings of 9 and 20 Å diameter, respectively) [85]. An artificial hybrid activator has also been constructed using the heptameric PA26 structure as a scaffold for attachment of eight residues corresponding to the C-terminal tail of MjPAN [86]. This hybrid activator is heptameric and forms a stable complex with the TaCP in which the activator caps both ends of the TaCP and opens the gate of the CP through an HbYX-dependent mechanism. Using single particle cryoEM and X-ray crystallography, the structure of the hybrid activator was determined in association with TaCP. In both studies [85, 86], residues corresponding to the MjPAN C-terminal tail were found to bind to the α -rings of TaCP in the ‘inter-subunit’ pocket (between the adjacent α -subunits) and induce and stabilize the open-gate conformation of TaCP. Although PA26 and MjPAN bind to a similar pocket on the CP, PA26 also requires binding to the CP α -subunit N-terminal tails, by a distant activation domain, to open and stabilize the CP gates [94, 95].

Interestingly, most haloarchaea and many methanogens are predicted to encode two distinct homologs of PAN, often one of which is lacking the conserved C-terminal HbYX motif. Thus, the C-terminal motif needed for CP interaction and gate opening, by PAN, may vary among archaeal species or alternatively archaea may synthesize PAN subtypes of different functions (e.g. PANs that unfold proteins but do not bind the CP or PANs that bind the CP but do not open the gates). Among the C-terminal sequences predicted for haloarchaeal PANs, ~43% include HbYX motifs, ~33% possess TFA motifs and ~23% harbour C-terminal tails that lack tyrosine or phenylalanine in the penultimate position. In *H. volcanii*, two PAN proteins have been identified, PAN-A with a C-terminal sequence, AFA and PAN-B with a C-terminal sequence, YQY [63]. As the penultimate tyrosine or phenylalanine of C-terminal tails appears essential for opening CP gates, the PAN-A and PAN-B proteins may have distinct roles in regulating proteasome-mediated degradation. Consistent with this possibility, a number of biological differences have been detected for the *H. volcanii* PAN-A and PAN-B proteins that support distinct functional roles. First, PAN-A is synthesized at high levels throughout growth, while the levels of PAN-B are low and induced several fold as cells transition to stationary phase [63]. In addition, while a phenotype for *panB* mutant strains has yet to be identified, deletion of *panA* renders cells more resistant to thermal stress and more sensitive to nitrogen limitation, hypo-osmotic stress, and exposure to L-canavanine [62]. In further support of functional differences between the two PAN proteins, the *panA* mutation can be complemented *in trans* by providing a wild type copy of *panA* but not *panB* [62].

Multiple Roles of the ATPase (PAN) in CP-Mediated Proteolysis

MjPAN is not only required for CP gate opening, but also for the unfolding and translocation of protein substrates for hydrolysis by the CP. Indeed “gateless” TaCP (containing $\alpha\Delta N$ subunits) still require MjPAN and ATP for the degradation of globular substrates (such as SsrA-tagged GFP) [82]. Thus, while the N-terminal tails of CP α -subunits appear to serve as a gate to prevent entry of large peptides as well as folded and unfolded proteins, removal of this gate is not sufficient to stimulate the degradation of structured proteins. This type of proteolysis still requires hydrolysis of ATP by an AAA + regulatory component, such as MjPAN in association with the CP.

Archaeal PAN X-ray Crystal Structures

In the quest to solve the overall structure of an energy-dependent proteasome, a number of recent breakthroughs have been made examining the structure of various archaeal PAN subdomains. In one study, MjPAN was subjected to limited proteolysis resulting in the generation of two subcomplexes (I and II) that were amenable to X-ray crystallography [96]. Subcomplex I, which forms a stable hexamer, was derived from a fragment that spanned residues 74–150 including a portion of the predicted N-terminal coiled-coil (CC) domain

(Fig. 11.6). In the subcomplex I structure (resolved to 2.1 Å), the CC domains of the six protein monomers formed pairs that protruded from a donut-shaped particle with an axial pore of 13 Å and an unexpected oligonucleotide/oligosaccharide binding (OB)-fold domain (Fig. 11.6) [96]. In contrast, subcomplex II, which dissociated into a mixture of hexamers to monomers, was composed of a fragment spanning residues 155–430 including the conserved AAA + domain and contained the C-terminal HbYX motif (Fig. 11.6). Given the level of conservation between MjPAN and the bacterial HslU, the atomic coordinates of MjPAN were superimposed onto the six subunits of HslU to generate a hexameric model for MjPAN (Fig. 11.6) [88, 90, 96]. Guided by these structures, a mutagenesis study, examined which amino acids and structural motifs were required for MjPAN function [84]. The structures of *Archaeoglobus fulgidus* PAN (AfPAN) N-terminal subdomain containing a eukaryotic GCN4 leucine zipper (in place of its CC-domain) has also helped to elucidate how folded proteins are degraded by archaeal proteasomes [97].

Proteasome-Mediated Proteolysis

From a combination of structural and biochemical information, a complete view of how the proteasomal ATPase component functions in the unfolding of proteins for translocation into the CP for degradation is beginning to emerge. Based on electron micrographs of PAN:CP complexes and mass spectrometry of purified PAN complexes, archaeal PAN is thought to function as a hexameric complex [83, 98], with its C-terminal domain facing the cylindrical ends of four-stacked heptameric rings that form the CP (Fig. 11.6). The central channel of PAN is aligned with the gated central channel of the CP, to form a long tunnel with narrow restrictions for the passage of substrate proteins to the central proteolytic chamber. The N-domain of PAN is positioned distal to the CP with the N-terminal CC domains (one from each subunit) associating in pairs, to form three tentacles that appear to stretch out and search for protein substrates (or possibly protein partners). The OB domain forms a stable ringed-platform for this network of tentacles and a relatively narrow pore (16 Å in diameter) that is thought to function as a molecular sieve restricting access of folded proteins from entering the central channel of the ATPase. While early evidence suggested MjPAN could unfold proteins on its surface in the absence of translocation [99], more recent work supports a model in which proteins are unfolded by energy-dependent translocation through the ATPase ring and that this can be coupled to threading the protein into the CP for destruction [96, 100]. Consistent with a threading model, the Ar-Φ loop (where Ar is any aromatic amino acid and Φ symbol is any hydrophobic amino acid), also known as the pore-1 loop, lines the central passage of the AAA + domain and defines one of the constriction points proximal to the OB domain in the subcomplex II model (Fig. 11.6) [84, 96]. Similar to related AAA + proteins, the Ar-Φ loop of PAN is thought to facilitate protein unfolding by gripping and tugging the substrate protein into the ATPase channel by ATP fuelled motions of the AAA + domain [96, 101–105]. The OB domain is proposed to exert a passive force on protein unfolding by providing a pore with a stable platform that blocks the movement of folded protein structure through the pore as the Ar-Φ loops tug and pull the protein [84]. Ultimately, the unfolded protein is translocated for degradation by the proteolytic active sites sequestered within the central chamber of the CP.

Archaeal PAN – An Ordered Reaction Cycle

Interestingly, subunits of MjPAN appear to bind ATP in pairs, which results in distinct effects on proteasome function that imply an ordered reaction cycle. Subunits of MjPAN are thought to partner with subunits opposite each other, in the hexameric ring, team up and cycle around the ring like a clock between one of three different states including ATP bound, ADP bound and nucleotide free [106]. Thus, MjPAN subunit pairs mediate an ordered reaction cycle of ATP binding, hydrolysis and release that coordinates and drives conformational changes around the MjPAN ring. These ATP-driven conformational changes

are proposed to be critical for MjPAN-mediated protein unfolding. In particular, the Ar- Φ loop within the channel of the MjPAN ring is thought to grip proteins, tugging them up and down with repeated cycles of ATP hydrolysis. Tugging the proteins through the pore formed by the OB-fold domain is likely to serve as a rigid platform for resisting the entry of protein structure and, thus, facilitating an unfolding process through the central channel of the PAN ATPase.

Non-ATPases Associated with Archaeal CPs

In addition to AAA + regulatory proteins such as PAN, archaeal CPs are also proposed to associate with non-ATPase components. Early evidence for this possibility, arose when a protein inhibitor was co-purified with *T. acidophilum* CP [107], however the identity of the 20 kDa subunit of the inhibitor remains to be determined. More recently, immature archaeal CPs (composed of β -subunits that retain the N-terminal propeptide due to a Thr to Ala active site mutation) have been shown *in vitro*, to associate with proteins of the DUF75 superfamily [108]. Archaea and eukaryotes typically encode at least two members of the DUF75 superfamily that form heterodimeric complexes [Pba1-Pba2 (yeast), PAC1-PAC2 (mammals) and PbaA-PbaB in the archaeon *Methanococcus maripaludis*] [108–111]. Interestingly, the archaeal PbaA and eukaryotic Pba1/PAC1 proteins all have a conserved C-terminal HbYX motif that is required for binding the immature CPs [108]. Based on site-directed mutagenesis, this C-terminal tail docks to the same surface pocket on the α -subunit of the CP that is used by ATPase activators such as MjPAN [108]. Unlike MjPAN which opens the CP gate, the DUF75 proteins do not appear to alter CP activity and instead are thought to shield proteasomal subunit intermediates from nonproductive associations until assembly is complete or protect the cell from misassembled complexes [109–111]. While DUF75 protein homologs are also found in actinobacteria, their function is thought to be distinct from the DUF75 proteins of Archaea and eukaryotes due to the addition of a large C-terminal domain and absence of a C-terminal HbYX motif [108].

Protein Modifications

Ubiquitylation and Its Role in Marking Proteins for Proteolysis by Proteasomes in Eukaryotes

Ubiquitin is a highly conserved small protein modifier that is covalently attached to substrate proteins by a process termed ubiquitylation (Fig. 11.7a) [112]. Ubiquitylation involves the covalent attachment of the C-terminal carboxylate of Gly76 of ubiquitin to a protein substrate. Ubiquitin can be attached to proteins by an isopeptide bond to the ϵ -amino group of substrate lysine residues, a thioester bond to substrate cysteine residues, an oxy-ester bond to substrate serine or threonine residues, or peptide bond to the N ^{α} -amino group of a protein substrate [113, 114]. The attachment of a single moiety of ubiquitin to a protein (mono-ubiquitylation) can alter the activity and/or location of the protein and can even signal the protein for destruction in lysosomes [115–117]. Chains of ubiquitin can also form through isopeptide bonds between the C-terminal carboxyl group of an incoming ubiquitin to one of the seven lysine residues of a ubiquitin (K6, K11, K27, K29, K33, K48 and K63) moiety covalently attached to a protein substrate [118]. Of these chains, the K63- and K48-linked chains of ubiquitin are best characterized. The K63-linked ubiquitin chains have non-proteolytic roles such as signalling for sorting into the multivesicular body pathway [119], while the K48-linked ubiquitin chains typically target a protein for destruction by 26S proteasomes [120].

Ubiquitylation involves a series of enzyme catalyzed reactions with the formation of adenylated ubiquitin and thioester intermediates that result in the covalent attachment of ubiquitin to proteins (Fig. 11.7a) [121–123]. In the first step, a ubiquitin-activating enzyme

(E1) adenylates the C-terminal carboxylate of Gly76 of ubiquitin (part of the a conserved diglycine motif, Gly75Gly76) in an ATP-dependent reaction that releases PPi. The adenylated form of ubiquitin is attacked by a conserved cysteine residue of E1 to form a thioester intermediate between the C-terminal Gly76 of ubiquitin and the conserved catalytic cysteine of the E1. In the next step, a conserved cysteine residue of an ubiquitin-conjugating enzyme (E2) attacks the E1-ubiquitin thioester intermediate. This E2-mediated reaction results in the transfer of ubiquitin to the E2 enzyme and formation of a thioester bond between the E2 and the C-terminal Gly76 of ubiquitin. Finally, ubiquitin is transferred to the substrate protein with assistance from an ubiquitin ligase (E3). The E3 enzyme either directly transfers ubiquitin from E2 to the protein substrate or forms a thioester intermediate between a conserved cysteine of E3 and the C-terminal Gly76 of ubiquitin prior to transfer. In the formation of ubiquitin chains, the concerted action of the E1-E2-E3 enzymes is repeated with the transfer of ubiquitin to one of the seven lysine residues of ubiquitin that is attached to the modified protein [124]. In some cases, the ubiquitylation process is more elaborate with the use of a ubiquitin chain elongation factor termed an E4 [122], while other ubiquitylation events occur independent of an E3-ubiquitin ligase [125–127].

Multifunctional Roles of Proteins Related to Ubiquitin and E1 in Bacteria

While ubiquitin is not conserved, at the amino acid level, among either Archaea or bacteria, proteins with a predicted β -grasp ubiquitin-like fold are common to all organisms [128]. For example, the bacterial proteins ThiS and MoaD both exhibit a ubiquitin-like structure and have been extensively studied for their role in the incorporation of sulfur into thiamine and molybdenum cofactors, respectively [129, 130]. Additional biochemical roles for bacterial proteins with a ubiquitin-like structure have been identified that are independent of protein modification including QbsE-mediated thioquinolobactin siderophore biosynthesis in *Pseudomonas fluorescens* [131], CysO-mediated cysteine biosynthesis in *Mycobacterium tuberculosis* [132, 133] and others [134].

Similar to eukaryotic ubiquitin, bacterial proteins with a ubiquitin-like fold can be activated at their C-terminus by enzymes of the E1-like superfamily [135]. For example, the bacterial E1 homologs, MoeB and ThiF, associate with and adenylate the C-terminal carboxylate of their cognate ubiquitin-like partner protein, MoaD and ThiS, respectively [136, 137]. However, in contrast to ubiquitin, the adenylated form of the bacterial ubiquitin-like proteins is typically modified by an E1-like enzyme to accept sulfur from either a cysteine desulfurase or rhodanese [138, 139]. This transfer of sulfur results in the formation of a C-terminal thiocarboxylated form of the ubiquitin-like protein that can be used as a source of activated-sulfur for biosynthetic reactions [140].

E1 and Ubiquitin-Like Protein Homologs of Archaea

While all Archaea encode homologs of E1 and proteins with a ubiquitin-like protein structure based on genome sequence, the role of these proteins in archaeal cell function has only recently been examined. Using the halophilic archaeon *H. volcanii* as a model system, two ubiquitin-like proteins, termed SAMP1 and SAMP2 (small archaeal modifier protein 1 and 2) and an E1 homolog (UbaA, ubiquitin-like activating protein of Archaea) were examined for their role in protein modification and/or sulfur incorporation into biomolecules such as molybdenum cofactor and tRNA (2-thiouracil) [141, 142]. Similar to most archaeal species, halophilic Archaea are predicted to synthesize biomolecules that require sulfur such as molybdenum/tungsten cofactors [143]. Furthermore, like most Archaea, *H. volcanii* is predicted to encode multiple ubiquitin-like proteins and a single E1 homolog, but not E2 or E3 homologs [144–146]. Thus, Archaea were anticipated to use ubiquitin-like proteins for sulfur incorporation into biomolecules and not for protein conjugation [128].

Sampylation – An Archaeal Form of Ubiquitylation

Recently it was shown that ubiquitin-like proteins can be conjugated to proteins in Archaea, by a process termed sampylation that is dependent upon the synthesis of an E1 homolog (Fig. 11.7b). These surprising findings are based on the following experimental evidence. Using *H. volcanii* as a model system, the ubiquitin-like proteins SAMP1 (87 a.a.) and SAMP2 (66 a.a.) were expressed with an N-terminal FLAG-tag and the formation of SAMP1/2 protein conjugates was analyzed by α -FLAG immunoblot [141]. Under certain growth conditions, an array of large α -FLAG specific protein bands was detected for both FLAG-SAMP1 and FLAG-SAMP2 expression strains [141]. Deletion of the conserved C-terminal diglycine motif of the ubiquitin-like protein or deletion of the *ubaA* gene (encoding the *H. volcanii* E1 homolog UbaA) abolished the ability to detect SAMP protein conjugates in *H. volcanii* cells [141, 142]. This effect could be complemented *in trans* by the wild type *ubaA* (but not by *ubaA* with a mutation in the conserved active site cysteine) suggesting that SAMP1 and SAMP2 are activated by an E1-like mechanism [142].

To determine the type of peptide bond formed between archaeal ubiquitin-like proteins and their protein substrates, SAMP2 was selected for further analysis [141], as it contains a lysine residue immediately preceding its C-terminal diglycine motif (i.e. KGG). The location of this lysine residue enabled the use of tryptic digestion and tandem mass spectrometry (MS/MS) to analyze the SAMP2 protein conjugates, in a similar to approach to that used to analyze ubiquitylated proteomes of eukaryotes [147]. If an isopeptide bond was present, between the C-terminal carboxylate of Gly66 in SAMP2 and an internal lysine residue on the protein substrate, the modified lysine residue of the protein substrate would be resistant to cleavage by trypsin and hence retain the GG footprint derived from the SAMP2 C-terminus. With this approach, SAMP2-protein conjugates were isolated by α -FLAG affinity chromatography from an *H. volcanii* FLAG-SAMP2 expression strain (compared to the control strain) and isopeptide bonds were detected between Gly66 (of SAMP2) and the lysine residues of eight different proteins predicted to mediate a variety of functions, from metabolism to transcription [141]. Similar to ubiquitylation [148], SAMP2 modification (samp2ylation) was also detected on multiple sites within a single target protein (i.e. a TATA-box binding protein and tandem rhodanese domain protein) [141]. Furthermore, Gly66 of SAMP2 also formed an isopeptide bond with one of the two SAMP2 lysine residues (i.e. Lys58) revealing the presence of K58-linked SAMP2 chains [141]. Whether the SAMP2 chains are associated with a protein substrate remains to be determined. In eukaryotes, ubiquitin genes are translated into polypeptide chains of repeating units of ubiquitin fused head-to-tail [149–151] or ubiquitin fused to unrelated amino acid sequences [152]. These ubiquitin-protein fusions are cleaved by deubiquitylases to generate free pools of mature ubiquitin [153]. Unlike ubiquitin, SAMP2 is translated as a single protein. While speculative, it is possible that archaeal cells might regulate the pools of free SAMP2 by the synthesis and cleavage of unanchored SAMP2 chains. Alternatively, if SAMP2 chains are anchored, this might enhance the diversity of SAMP modifications available for protein targeting.

Using various strains of *H. volcanii* expressing either FLAG-SAMP1 or FLAG-SAMP2, together with α -FLAG immunoprecipitation, several interacting proteins were identified by tandem MS, following “in-gel” digested with trypsin [141]. Cells used for this analysis were grown in different culture conditions to enhance coverage of proteins modified by SAMP1 and/or SAMP2 (termed the “sampylome”). While SAMP modification sites were not mapped using this second approach, 32 proteins specific for the FLAG-SAMP expression strains were identified by MS/MS. These included the 8 proteins described above (with their SAMP2 modification sites mapped) as well as 24 additional homologs of proteins involved

in ubiquitylation, sulfur incorporation, stress responses, metabolism and information processing such as transcription and translation [141].

Regulation of Sanylation

Environmental conditions can signal changes in the level of SAMP protein conjugates that are formed in *H. volcanii* cells. For example, growth in the presence of dimethyl sulfoxide and other culture conditions can increase the levels of SAMP protein conjugates [141, 142]. Thus, sanylation is regulated, most likely at the posttranscriptional level, based on the use of a constitutive promoter for expression of the genes encoding the SAMP proteins throughout these experiments [141, 142].

Functional Role of Sanylation

Sanylation (the formation of SAMP protein conjugates) appears to mark some proteins for degradation by proteasomes in Archaea. However, the evidence supporting a biological connection between sanylation and proteasomes remains indirect. The first piece of evidence in support of this association is that SAMP1-modified proteins accumulate in *H. volcanii* cells that contain chromosomal deletions of the genes encoding the proteasomal CP α 1-subunit (*psmA*) or the Rpt-like ATPase PAN-A (*panA*) [141]. Thus, SAMP1-modified proteins may not be efficiently degraded by proteasomes in these mutant cells. In addition to this finding, many of the proteins that are modified by SAMP1 and/or SAMP2 or associated with these SAMP proteins [141] were also found to accumulate in *H. volcanii* cells that are disrupted in proteasome function either by deletion of *panA* [154] or treatment with the proteasome-specific inhibitor clasto-Lactacystin β -lactone [155]. However, currently there is no direct evidence that sanylation of a protein leads to its degradation by the proteasome.

To further understand the role of sanylation in archaeal cell function, the genes encoding the ubiquitin-like protein modifiers (SAMP1 and SAMP2) and the E1 homolog (UbaA) were deleted from the genome of *H. volcanii* [142]. The resulting mutant strains were examined for biochemical and phenotypic differences compared to wild type cells. With this approach, UbaA was found to be required for the formation (or stabilization) of SAMP1 and SAMP2 protein conjugates in *H. volcanii* [142]. Thus, an E1-type mechanism of SAMP protein activation has been proposed (Fig. 11.7b).

In addition to protein conjugation, UbaA and SAMP2 appear essential for the thiolation of tRNA (most likely 2-thiouridine) based on comparison of tRNA from wild type and mutant cells, using [(*N*-acryloylamino)phenyl]mercuric chloride (APM) gel electrophoresis followed by hybridization to a specific probe for lysine tRNAs with anticodon UUU ($\text{tRNA}^{\text{Lys}}_{\text{UUU}}$) [142]. Because E1- and ubiquitin-like proteins can catalyze sulfur incorporation into tRNA in bacteria and eukaryotes [156, 157], the biochemical differences in the level of thiolated tRNA in *H. volcanii* *ubaA* and *samp2* mutants (devoid of slow migrating “thiolated” $\text{tRNA}^{\text{Lys}}_{\text{UUU}}$) compared to wild type are thought to be independent of protein conjugation (Fig. 11.7b).

UbaA and SAMP1 also seem important for sulfur incorporation into the pterinbased molybdenum cofactor (MoCo) (Fig. 11.7b). This is based on the finding that, in comparison to wild type cells, *ubaA* and *samp1* mutant cells are unable to grow under anaerobic conditions with dimethyl sulfoxide (DMSO) as a terminal electron acceptor or produce DMSO reductase activity when switched to anaerobic conditions with DMSO [142]. DMSO reductase requires incorporation of MoCo into the catalytic subunit DmsA for activity. Similarly to wild type, the *ubaA* and *samp1* mutant strains were still able to produce transcript specific for *dmsA* (encoding the catalytic subunit of DMSO reductase).

Together, these results suggest UbaA and SAMP proteins mediate not only sampylation but also sulfur incorporation into tRNA and MoCo (Fig. 11.7b). The E1 homolog, UbaA, appears at the crossroads of activating the SAMP proteins for formation of an isopeptide bond between the C-terminal carboxylate of SAMP and the target protein and for the formation of a putative C-terminal thiocarboxylate group on the SAMP1 and SAMP2 proteins for sulfur transfer to MoCo and tRNA, respectively. Thus, any phenotypic differences that are detected for archaeal mutants of E1 and ubiquitin-like homologs (in comparison to wild type cells) will need to be further examined for whether these changes are due to perturbations in sampylation (protein conjugation) or sulfur transfer to biomolecules.

Summary

While Archaea are often thought of as prokaryotes, due to the absence of a membrane-bound nucleus, these incredible microbes have self-compartmentalized proteases and protein conjugation systems that are closely related to the ubiquitin-proteasome systems of eukaryotes. Unlike bacteria, which have multiple energy-dependent proteases in their cytosol [i.e. Lon, Clp and HslUV (or proteasomes in actinobacteria)], Archaea appear highly dependent on the proteasome system (with an unknown contribution by the membrane-bound Lon protease) as their major energy-dependent proteases. Based on numerous studies that have investigated how archaeal PAN and CP complexes function at the atomic level, an elaborate mechanism of how proteasomes degrade unfolded and/or folded proteins is rapidly emerging. Surprisingly, Archaea also contain a system in which ubiquitin-like proteins are conjugated to substrate proteins, through an apparent E1-like mechanism termed sampylation. Sampylation is predicted to exist in all Archaea and occurs independent of E2 or E3 homologs. Whether or not sampylation marks proteins for degradation by the proteasome remains to be determined; however, indirect evidence suggests these two pathways are functionally connected. Improving our understanding of how proteins are targeted for degradation in archaea promises to be an exciting area of research. Sampylation is likely a simplified system for the covalent marking proteins, which has close evolutionary roots to ubiquitylation and other related pathways that play an important role in eukaryotic cell function. Interestingly, more elaborate protein conjugation mechanisms than sampylation may exist in a select group of bacteria and archaea that harbour E1, E2 and E3 homologs, based on recent metagenomic sequencing and comparative genomics [144, 158].

Acknowledgments

This work was funded in part by grants from the National Institutes of Health (GM57498) and the Department of Energy Office of Basic Energy Sciences (DE-FG02-05ER15650).

References

1. Gottesman S. Regulation by proteolysis: developmental switches. *Curr Opin Microbiol.* 1999; 2(2): 142–147. [PubMed: 10322171]
2. Wolf DH, Hilt W. The proteasome: a proteolytic nanomachine of cell regulation and waste disposal. *Biochim Biophys Acta.* 2004; 1695(1–3):19–31. [PubMed: 15571806]
3. Lupas A, Flanagan JM, Tamura T, Baumeister W. Self-compartmentalizing proteases. *Trends Biochem Sci.* 1997; 22(10):399–404. [PubMed: 9357316]
4. Cha SS, An YJ, Lee CR, Lee HS, et al. Crystal structure of Lon protease: molecular architecture of gated entry to a sequestered degradation chamber. *EMBO J.* 2010; 29(20):3520–3530. [PubMed: 20834233]
5. Effantin G, Maurizi MR, Steven AC. Binding of the ClpA unfoldase opens the axial gate of ClpP peptidase. *J Biol Chem.* 2010; 285(19):14834–14840. [PubMed: 20236930]

6. Lee ME, Baker TA, Sauer RT. Control of substrate gating and translocation into ClpP by channel residues and ClpX binding. *J Mol Biol.* 2010; 399(5):707–718. [PubMed: 20416323]
7. Smith DM, Benaroudj N, Goldberg A. Proteasomes and their associated ATPases: a destructive combination. *J Struct Biol.* 2006; 156(1):72–83. [PubMed: 16919475]
8. Ammelburg M, Frickey T, Lupas AN. Classification of AAA + proteins. *J Struct Biol.* 2006; 156(1): 2–11. [PubMed: 16828312]
9. Neuwald AF, Aravind L, Spouge JL, Koonin EV. AAA+: a class of chaperone-like ATPases associated with the assembly, operation, and disassembly of protein complexes. *Genome Res.* 1999; 9(1):27–43. [PubMed: 9927482]
10. Sauer RT, Baker TA. AAA+ proteases: ATP-fueled machines of protein destruction. *Annu Rev Biochem.* 2011; 80:587–612. [PubMed: 21469952]
11. Gur, E.; Ottofuelling, R.; Dougan, DA. Machines of destruction – AAA + proteases and the adaptors that control them. In: Dougan, DA., editor. *Regulated proteolysis in microorganisms.* Vol. 66. Springer: Subcell Biochem; 2013. p. 3-33.
12. Samanovic, M.; Li, H.; Darwin, KH. The Pup-Proteasome system of *Mycobacterium tuberculosis*. In: Dougan, DA., editor. *Regulated proteolysis in microorganisms.* Vol. 66. Springer: Subcell Biochem; 2013. p. 267-295.
13. Maupin-Furlow JA, Humbard MA, Kirkland PA, Li W, et al. Proteasomes from structure to function: perspectives from Archaea. *Curr Top Dev Biol.* 2006; 75:125–169. [PubMed: 16984812]
14. Bar-Nun S, Glickman MH. Proteasomal AAA-ATPases: structure and function. *Biochim Biophys Acta.* 2012; 1823(1):67–82. [PubMed: 21820014]
15. Besche H, Tamura N, Tamura T, Zwickl P. Mutational analysis of conserved AAA+ residues in the archaeal Lon protease from *Thermoplasma acidophilum*. *FEBS Lett.* 2004; 574(1–3):161–166. [PubMed: 15358558]
16. Fukui T, Eguchi T, Atomi H, Imanaka T. A membrane-bound archaeal Lon protease displays ATP-independent proteolytic activity towards unfolded proteins and ATP-dependent activity for folded proteins. *J Bacteriol.* 2002; 184(13):3689–3698. [PubMed: 12057965]
17. De Mot R, Nagy I, Walz J, Baumeister W. Proteasomes and other self-compartmentalizing proteases in prokaryotes. *Trends Microbiol.* 1999; 7(2):88–92. [PubMed: 10081087]
18. Lupas A, Zuhl F, Tamura T, Wolf S, et al. Eubacterial proteasomes. *Mol Biol Rep.* 1997; 24(1–2): 125–131. [PubMed: 9228293]
19. Lupas A, Zwickl P, Baumeister W. Proteasome sequences in eubacteria. *Trends Biochem Sci.* 1994; 19(12):533–534. [PubMed: 7846763]
20. Tamura T, Nagy I, Lupas A, Lottspeich F, et al. The first characterization of a eubacterial proteasome: the 20S complex of *Rhodococcus*. *Curr Biol.* 1995; 5(7):766–774. [PubMed: 7583123]
21. Bochtler M, Ditzel L, Groll M, Hartmann C, et al. The proteasome. *Annu Rev Biophys Biomol Struct.* 1999; 28:295–317. [PubMed: 10410804]
22. Voges D, Zwickl P, Baumeister W. The 26S proteasome: a molecular machine designed for controlled proteolysis. *Annu Rev Biochem.* 1999; 68:1015–1068. [PubMed: 10872471]
23. Larsen CN, Finley D. Protein translocation channels in the proteasome and other proteases. *Cell.* 1997; 91(4):431–434. [PubMed: 9390550]
24. Brannigan JA, Dodson G, Duggleby HJ, Moody PC, et al. A protein catalytic framework with an N-terminal nucleophile is capable of self-activation. *Nature.* 1995; 378(6555):416–419. [PubMed: 7477383]
25. Ekici OD, Paetzel M, Dalbey RE. Unconventional serine proteases: variations on the catalytic Ser/His/Asp triad configuration. *Protein Sci.* 2008; 17(12):2023–2037. [PubMed: 18824507]
26. Oinonen C, Rouvinen J. Structural comparison of Ntn-hydrolases. *Protein Sci.* 2000; 9(12):2329–2337. [PubMed: 11206054]
27. Seemuller E, Lupas A, Stock D, Lowe J, et al. Proteasome from *Thermoplasma acidophilum*: a threonine protease. *Science.* 1995; 268(5210):579–582. [PubMed: 7725107]
28. Dodson G, Wlodawer A. Catalytic triads and their relatives. *Trends Biochem Sci.* 1998; 23(9): 347–352. [PubMed: 9787641]

29. Kisselev AF, Songyang Z, Goldberg AL. Why does threonine, and not serine, function as the active site nucleophile in proteasomes? *J Biol Chem.* 2000; 275(20):14831–14837. [PubMed: 10809725]
30. Maupin-Furrow JA, Aldrich HC, Ferry JG. Biochemical characterization of the 20S proteasome from the methanarchaeon *Methanosarcina thermophila*. *J Bacteriol.* 1998; 180(6):1480–1487. [PubMed: 9515917]
31. Zwickl P, Kleinz J, Baumeister W. Critical elements in proteasome assembly. *Nat Struct Biol.* 1994; 1(11):765–770. [PubMed: 7634086]
32. Groll M, Heinemeyer W, Jager S, Ullrich T, et al. The catalytic sites of 20S proteasomes and their role in subunit maturation: a mutational and crystallographic study. *Proc Natl Acad Sci U S A.* 1999; 96(20):10976–10983. [PubMed: 10500111]
33. Jager S, Groll M, Huber R, Wolf DH, et al. Proteasome beta-type subunits: unequal roles of propeptides in core particle maturation and a hierarchy of active site function. *J Mol Biol.* 1999; 291(4):997–1013. [PubMed: 10452902]
34. Bauer MW, Bauer SH, Kelly RM. Purification and characterization of a proteasome from the hyperthermophilic archaeon *Pyrococcus furiosus*. *Appl Environ Microbiol.* 1997; 63(3):1160–1164. [PubMed: 16535545]
35. Dahlmann B, Kopp F, Kuehn L, Nidel B, et al. The multicatalytic proteinase (prosome) is ubiquitous from eukaryotes to archaeobacteria. *FEBS Lett.* 1989; 251(1–2):125–131. [PubMed: 2502434]
36. Maupin-Furrow JA, Ferry JG. A proteasome from the methanogenic archaeon *Methanosarcina thermophila*. *J Biol Chem.* 1995; 270(48):28617–28622. [PubMed: 7499378]
37. Wilson HL, Aldrich HC, Maupin-Furrow J. Halophilic 20S proteasomes of the archaeon *Haloferax volcanii*: purification, characterization, and gene sequence analysis. *J Bacteriol.* 1999; 181(18):5814–5824. [PubMed: 10482525]
38. Arendt CS, Hochstrasser M. Eukaryotic 20S proteasome catalytic subunit propeptides prevent active site inactivation by N-terminal acetylation and promote particle assembly. *EMBO J.* 1999; 18(13):3575–3585. [PubMed: 10393174]
39. Kwon YD, Nagy I, Adams PD, Baumeister W, et al. Crystal structures of the *Rhodococcus* proteasome with and without its pro-peptides: implications for the role of the pro-peptide in proteasome assembly. *J Mol Biol.* 2004; 335(1):233–245. [PubMed: 14659753]
40. Zuhl F, Seemuller E, Golbik R, Baumeister W. Dissecting the assembly pathway of the 20S proteasome. *FEBS Lett.* 1997; 418(1–2):189–194. [PubMed: 9414124]
41. Li D, Li H, Wang T, Pan H, et al. Structural basis for the assembly and gate closure mechanisms of the *Mycobacterium tuberculosis* 20S proteasome. *EMBO J.* 2010; 29(12):2037–2047. [PubMed: 20461058]
42. Lin G, Hu G, Tsu C, Kunes YZ, et al. *Mycobacterium tuberculosis* *prcBA* genes encode a gated proteasome with broad oligopeptide specificity. *Mol Microbiol.* 2006; 59(5):1405–1416. [PubMed: 16468985]
43. Wilson HL, Ou MS, Aldrich HC, Maupin-Furrow J. Biochemical and physical properties of the *Methanococcus jannaschii* 20S proteasome and PAN, a homolog of the ATPase (Rpt) subunits of the eucaryal 26S proteasome. *J Bacteriol.* 2000; 182(6):1680–1692. [PubMed: 10692374]
44. Kaczowka SJ, Maupin-Furrow JA. Subunit topology of two 20S proteasomes from *Haloferax volcanii*. *J Bacteriol.* 2003; 185(1):165–174. [PubMed: 12486053]
45. Ginzburg M, Sachs L, Ginzburg BZ. Ion metabolism in a *Halobacterium*. I. Influence of age of culture on intracellular concentrations. *J Gen Physiol.* 1970; 55(2):187–207. [PubMed: 5413077]
46. Mevarech M, Frolow F, Gloss LM. Halophilic enzymes: proteins with a grain of salt. *Biophys Chem.* 2000; 86(2–3):155–164. [PubMed: 11026680]
47. Falb M, Aivaliotis M, Garcia-Rizo C, Bisle B, et al. Archaeal N-terminal protein maturation commonly involves N-terminal acetylation: a large-scale proteomics survey. *J Mol Biol.* 2006; 362(5):915–924. [PubMed: 16950390]
48. Kirkland PA, Humbard MA, Daniels CJ, Maupin-Furrow JA. Shotgun proteomics of the haloarchaeon *Haloferax volcanii*. *J Proteome Res.* 2008; 7(11):5033–5039. [PubMed: 18816081]

49. Soppa J. Protein acetylation in archaea, bacteria, and eukaryotes. *Archaea*. 2010; 2010:9. (Article ID 820681).
50. Kisselev AF, Akopian TN, Goldberg AL. Range of sizes of peptide products generated during degradation of different proteins by archaeal proteasomes. *J Biol Chem*. 1998; 273(4):1982–1989. [PubMed: 9442034]
51. Kisselev AF, Akopian TN, Woo KM, Goldberg AL. The sizes of peptides generated from protein by mammalian 26 and 20S proteasomes. Implications for understanding the degradative mechanism and antigen presentation. *J Biol Chem*. 1999; 274(6):3363–3371. [PubMed: 9920878]
52. Gallastegui N, Groll M. The 26S proteasome: assembly and function of a destructive machine. *Trends Biochem Sci*. 2010; 35(11):634–642. [PubMed: 20541423]
53. Vigneron N, Van den Eynde BJ. Proteasome subtypes and the processing of tumor antigens: increasing antigenic diversity. *Curr Opin Immunol*. 2012; 24(1):84–91. [PubMed: 22206698]
54. Zwickl P, Grziwa A, Puhler G, Dahlmann B, et al. Primary structure of the *Thermoplasma* proteasome and its implications for the structure, function, and evolution of the multicatalytic proteinase. *Biochemistry*. 1992; 31(4):964–972. [PubMed: 1734972]
55. Zwickl P, Lottspeich F, Baumeister W. Expression of functional *Thermoplasma acidophilum* proteasomes in *Escherichia coli*. *FEBS Lett*. 1992; 312(2–3):157–160. [PubMed: 1426246]
56. Kocabiyyik S, Ozdemir I, Zwickl P, Ozdogan S. Molecular cloning and co-expression of *Thermoplasma volcanium* proteasome subunit genes. *Protein Expr Purif*. 2010; 73(2):223–230. [PubMed: 20460155]
57. Medalia N, Sharon M, Martinez-Arias R, Mihalache O, et al. Functional and structural characterization of the *Methanosarcina mazei* proteasome and PAN complexes. *J Struct Biol*. 2006; 156(1):84–92. [PubMed: 16690322]
58. Groll M, Brandstetter H, Bartunik H, Bourenkow G, et al. Investigations on the maturation and regulation of archaeobacterial proteasomes. *J Mol Biol*. 2003; 327(1):75–83. [PubMed: 12614609]
59. Robb FT, Maeder DL, Brown JR, DiRuggiero J, et al. Genomic sequence of hyperthermophile, *Pyrococcus furiosus*: implications for physiology and enzymology. *Methods Enzymol*. 2001; 330:134–157. [PubMed: 11210495]
60. Madding LS, Michel JK, Shockley KR, Connors SB, et al. Role of the beta1 subunit in the function and stability of the 20S proteasome in the hyperthermophilic archaeon *Pyrococcus furiosus*. *J Bacteriol*. 2007; 189(2):583–590. [PubMed: 17114253]
61. Zhou G, Kowalczyk D, Humbard MA, Rohatgi S, et al. Proteasomal components required for cell growth and stress responses in the haloarchaeon *Haloferax volcanii*. *J Bacteriol*. 2008; 190(24):8096–8105. [PubMed: 18931121]
62. Karadzic I, Maupin-Furlow J, Humbard M, Prunetti L, Singh P, Goodlett DR. Chemical cross-linking, mass spectrometry, and in silico modeling of proteasomal 20S core particles of the haloarchaeon *Haloferax volcanii*. *Proteomics*. 2012; 12(11):1806–1814. [PubMed: 22623373]
63. Reuter CJ, Kaczowka SJ, Maupin-Furlow JA. Differential regulation of the PanA and PanB proteasome-activating nucleotidase and 20S proteasomal proteins of the haloarchaeon *Haloferax volcanii*. *J Bacteriol*. 2004; 186(22):7763–7772. [PubMed: 15516591]
64. Humbard MA, Reuter CJ, Zuobi-Hasona K, Zhou G, et al. Phosphorylation and methylation of proteasomal proteins of the haloarchaeon *Haloferax volcanii*. *Archaea*. 2010; 2010:481725. [PubMed: 20671954]
65. Humbard MA, Stevens SM Jr, Maupin-Furlow JA. Posttranslational modification of the 20S proteasomal proteins of the archaeon *Haloferax volcanii*. *J Bacteriol*. 2006; 188(21):7521–7530. [PubMed: 16950923]
66. Humbard MA, Zhou G, Maupin-Furlow JA. The N-terminal penultimate residue of 20S proteasome alpha1 influences its N(alpha) acetylation and protein levels as well as growth rate and stress responses of *Haloferax volcanii*. *J Bacteriol*. 2009; 191(12):3794–3803. [PubMed: 19376868]
67. Stadtmueller BM, Hill CP. Proteasome activators. *Mol Cell*. 2011; 41(1):8–19. [PubMed: 21211719]

68. Finley D, Tanaka K, Mann C, Feldmann H, et al. Unified nomenclature for subunits of the *Saccharomyces cerevisiae* proteasome regulatory particle. *Trends Biochem Sci.* 1998; 23(7):244–245. [PubMed: 9697412]
69. Glickman MH, Rubin DM, Coux O, Wefes I, et al. A subcomplex of the proteasome regulatory particle required for ubiquitin-conjugate degradation and related to the COP9-signalosome and eIF3. *Cell.* 1998; 94(5):615–623. [PubMed: 9741626]
70. Bult CJ, White O, Olsen GJ, Zhou L, et al. Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*. *Science.* 1996; 273(5278):1058–1073. [PubMed: 8688087]
71. Zwickl P, Ng D, Woo KM, Klenk HP, et al. An archaeobacterial ATPase, homologous to ATPases in the eukaryotic 26S proteasome, activates protein breakdown by 20S proteasomes. *J Biol Chem.* 1999; 274(37):26008–26014. [PubMed: 10473546]
72. Stolz A, Hilt W, Buchberger A, Wolf DH. Cdc48: a power machine in protein degradation. *Trends Biochem Sci.* 2011; 36(10):515–523. [PubMed: 21741246]
73. Buchberger, A. Roles of Cdc48 in regulated protein degradation in yeast. In: Dougan, DA., editor. *Regulated proteolysis in microorganisms*. Vol. 66. Springer: Subcell Biochem; 2013. p. 3p. 195-222.
74. Gerega A, Rockel B, Peters J, Tamura T, et al. VAT, the thermoplasma homolog of mammalian p97/VCP, is an N domain-regulated protein unfoldase. *J Biol Chem.* 2005; 280(52):42856–42862. [PubMed: 16236712]
75. Barthelme D, Sauer RT. Identification of the Cdc48•20S proteasome as an ancient AAA+ proteolytic machine. *Science.* 2012; 337(6096):843–846. [PubMed: 22837385]
76. Lowe J, Stock D, Jap B, Zwickl P, et al. Crystal structure of the 20S proteasome from the archaeon *T. acidophilum* at 3.4 Å resolution. *Science.* 1995; 268(5210):533–539. [PubMed: 7725097]
77. Groll M, Ditzel L, Lowe J, Stock D, et al. Structure of 20S proteasome from yeast at 2.4 Å resolution. *Nature.* 1997; 386(6624):463–471. [PubMed: 9087403]
78. Unno M, Mizushima T, Morimoto Y, Tomisugi Y, et al. Structure determination of the constitutive 20S proteasome from bovine liver at 2.75 Å resolution. *J Biochem.* 2002; 131(2):171–173. [PubMed: 11820928]
79. Groll M, Bajorek M, Kohler A, Moroder L, et al. A gated channel into the proteasome core particle. *Nat Struct Biol.* 2000; 7(11):1062–1067. [PubMed: 11062564]
80. Kohler A, Cascio P, Leggett DS, Wool KM, et al. The axial channel of the proteasome core particle is gated by the Rpt2 ATPase and controls both substrate entry and product release. *Mol Cell.* 2001; 7(6):1143–1152. [PubMed: 11430818]
81. Whitby FG, Masters EI, Kramer L, Knowlton JR, et al. Structural basis for the activation of 20S proteasomes by 11S regulators. *Nature.* 2000; 408(6808):115–120. [PubMed: 11081519]
82. Benaroudj N, Zwickl P, Seemuller E, Baumeister W, et al. ATP hydrolysis by the proteasome regulatory complex PAN serves multiple functions in protein degradation. *Mol Cell.* 2003; 11(1):69–78. [PubMed: 12535522]
83. Smith DM, Kafri G, Cheng Y, Ng D, et al. ATP binding to PAN or the 26S ATPases causes association with the 20S proteasome, gate opening, and translocation of unfolded proteins. *Mol Cell.* 2005; 20(5):687–698. [PubMed: 16337593]
84. Zhang F, Hu M, Tian G, Zhang P, et al. Structural insights into the regulatory particle of the proteasome from *Methanocaldococcus jannaschii*. *Mol Cell.* 2009; 34(4):473–484. [PubMed: 19481527]
85. Rabl J, Smith DM, Yu Y, Chang SC, et al. Mechanism of gate opening in the 20S proteasome by the proteasomal ATPases. *Mol Cell.* 2008; 30(3):360–368. [PubMed: 18471981]
86. Yu Y, Smith DM, Kim HM, Rodriguez V, et al. Interactions of PAN's C-termini with archaeal 20S proteasome and implications for the eukaryotic proteasome-ATPase interactions. *EMBO J.* 2010; 29(3):692–702. [PubMed: 20019667]
87. Religa TL, Sprangers R, Kay LE. Dynamic regulation of archaeal proteasome gate opening as studied by TROSY NMR. *Science.* 2010; 328(5974):98–102. [PubMed: 20360109]
88. Bochtler M, Hartmann C, Song HK, Bourenkov GP, et al. The structures of HsIU and the ATP-dependent protease HsIU-HsIV. *Nature.* 2000; 403(6771):800–805. [PubMed: 10693812]

89. Rohrwild M, Pfeifer G, Santarius U, Muller SA, et al. The ATP-dependent HslVU protease from *Escherichia coli* is a four-ring structure resembling the proteasome. *Nat Struct Biol.* 1997; 4(2): 133–139. [PubMed: 9033594]
90. Sousa MC, Trame CB, Tsuruta H, Wilbanks SM, et al. Crystal and solution structures of an HslUV protease-chaperone complex. *Cell.* 2000; 103(4):633–643. [PubMed: 11106733]
91. Smith DM, Chang SC, Park S, Finley D, et al. Docking of the proteasomal ATPases' carboxyl termini in the 20S proteasome's alpha ring opens the gate for substrate entry. *Mol Cell.* 2007; 27(5):731–744. [PubMed: 17803938]
92. Ramachandran R, Hartmann C, Song HK, Huber R, et al. Functional interactions of HslV (ClpQ) with the ATPase HslU (ClpY). *Proc Natl Acad Sci U S A.* 2002; 99(11):7396–7401. [PubMed: 12032294]
93. Ma CP, Willy PJ, Slaughter CA, DeMartino GN. PA28, an activator of the 20S proteasome, is inactivated by proteolytic modification at its carboxyl terminus. *J Biol Chem.* 1993; 268(30): 22514–22519. [PubMed: 8226760]
94. Forster A, Masters EI, Whitby FG, Robinson H, et al. The 1.9 Å structure of a proteasome-11S activator complex and implications for proteasome-PAN/PA700 interactions. *Mol Cell.* 2005; 18(5):589–599. [PubMed: 15916965]
95. Forster A, Whitby FG, Hill CP. The pore of activated 20S proteasomes has an ordered 7-fold symmetric conformation. *EMBO J.* 2003; 22(17):4356–4364. [PubMed: 12941688]
96. Zhang F, Wu Z, Zhang P, Tian G, et al. Mechanism of substrate unfolding and translocation by the regulatory particle of the proteasome from *Methanocaldococcus jannaschii*. *Mol Cell.* 2009; 34(4): 485–496. [PubMed: 19481528]
97. Djuranovic S, Hartmann MD, Habeck M, Ursinus A, et al. Structure and activity of the N-terminal substrate recognition domains in proteasomal ATPases. *Mol Cell.* 2009; 34(5):580–590. [PubMed: 19481487]
98. Medalia N, Beer A, Zwickl P, Mihalache O, et al. Architecture and molecular mechanism of PAN, the archaeal proteasome regulatory ATPase. *J Biol Chem.* 2009; 284(34):22952–22960. [PubMed: 19363223]
99. Navon A, Goldberg AL. Proteins are unfolded on the surface of the ATPase ring before transport into the proteasome. *Mol Cell.* 2001; 8(6):1339–1349. [PubMed: 11779508]
100. Matouschek A. Protein unfolding – an important process in vivo? *Curr Opin Struct Biol.* 2003; 13(1):98–109. [PubMed: 12581666]
101. Hinnerwisch J, Fenton WA, Furtak KJ, Farr GW, et al. Loops in the central channel of ClpA chaperone mediate protein binding, unfolding, and translocation. *Cell.* 2005; 121(7):1029–1041. [PubMed: 15989953]
102. Inobe T, Kraut DA, Matouschek A. How to pick a protein and pull at it. *Nat Struct Mol Biol.* 2008; 15(11):1135–1136. [PubMed: 18985068]
103. Martin A, Baker TA, Sauer RT. Pore loops of the AAA + ClpX machine grip substrates to drive translocation and unfolding. *Nat Struct Mol Biol.* 2008; 15(11):1147–1151. [PubMed: 18931677]
104. Martin A, Baker TA, Sauer RT. Diverse pore loops of the AAA + ClpX machine mediate unassisted and adaptor-dependent recognition of ssrA-tagged substrates. *Mol Cell.* 2008; 29(4): 441–450. [PubMed: 18313382]
105. Wang J, Song JJ, Seong IS, Franklin MC, et al. Nucleotide-dependent conformational changes in a protease-associated ATPase HslU. *Structure.* 2001; 9(11):1107–1116. [PubMed: 11709174]
106. Smith DM, Fraga H, Reis C, Kafri G, et al. ATP binds to proteasomal ATPases in pairs with distinct functional effects, implying an ordered reaction cycle. *Cell.* 2011; 144(4):526–538. [PubMed: 21335235]
107. Ehlers C, Kopp F, Dahlmann B. Screening for molecules interacting with proteasomes in *Thermoplasma acidophilum*. *Biol Chem.* 1997; 378(3–4):249–253. [PubMed: 9165078]
108. Kusmierczyk AR, Kunjappu MJ, Kim RY, Hochstrasser M. A conserved 20S proteasome assembly factor requires a C-terminal HbYX motif for proteasomal precursor binding. *Nat Struct Mol Biol.* 2011; 18(5):622–629. [PubMed: 21499243]

109. Hirano Y, Hendil KB, Yashiroda H, Iemura S, et al. A heterodimeric complex that promotes the assembly of mammalian 20S proteasomes. *Nature*. 2005; 437(7063):1381–1385. [PubMed: 16251969]
110. Le Tallec B, Barrault MB, Courbeyrette R, Guerois R, et al. 20S proteasome assembly is orchestrated by two distinct pairs of chaperones in yeast and in mammals. *Mol Cell*. 2007; 27(4): 660–674. [PubMed: 17707236]
111. Li X, Kusmierczyk AR, Wong P, Emili A, et al. beta-Subunit appendages promote 20S proteasome assembly by overcoming an Ump1-dependent checkpoint. *EMBO J*. 2007; 26(9): 2339–2349. [PubMed: 17431397]
112. Hochstrasser M. Origin and function of ubiquitin-like proteins. *Nature*. 2009; 458(7237):422–429. [PubMed: 19325621]
113. Iwai K, Tokunaga F. Linear polyubiquitination: a new regulator of NF-kappaB activation. *EMBO Rep*. 2009; 10(7):706–713. [PubMed: 19543231]
114. Wang X, Herr RA, Hansen TH. Ubiquitination of substrates by esterification. *Traffic*. 2012; 13(1):19–24. [PubMed: 21883762]
115. Frappier L, Verrijzer CP. Gene expression control by protein deubiquitinases. *Curr Opin Genet Dev*. 2011; 21(2):207–213. [PubMed: 21411309]
116. Hicke L, Dunn R. Regulation of membrane protein transport by ubiquitin and ubiquitin-binding proteins. *Annu Rev Cell Dev Biol*. 2003; 19:141–172. [PubMed: 14570567]
117. Ramaekers CH, Wouters BG. Regulatory functions of ubiquitin in diverse DNA damage responses. *Curr Mol Med*. 2011; 11(2):152–169. [PubMed: 21342128]
118. Xu P, Duong DM, Seyfried NT, Cheng D, et al. Quantitative proteomics reveals the function of unconventional ubiquitin chains in proteasomal degradation. *Cell*. 2009; 137(1):133–145. [PubMed: 19345192]
119. Lauwers E, Erpapazoglou Z, Haguenauer-Tsapis R, Andre B. The ubiquitin code of yeast permease trafficking. *Trends Cell Biol*. 2010; 20(4):196–204. [PubMed: 20138522]
120. Pickart CM. Targeting of substrates to the 26S proteasome. *FASEB J*. 1997; 11(13):1055–1066. [PubMed: 9367341]
121. Metzger MB, Weissman AM. Working on a chain: E3s ganging up for ubiquitylation. *Nat Cell Biol*. 2010; 12(12):1124–1126. [PubMed: 21124306]
122. Pastore A. Further insights into the ubiquitin pathway: understanding the scarlet letter code. *Structure*. 2010; 18(8):891–892. [PubMed: 20696387]
123. Scheffner M, Nuber U, Huibregtse JM. Protein ubiquitination involving an E1-E2-E3 enzyme ubiquitin thioester cascade. *Nature*. 1995; 373(6509):81–83. [PubMed: 7800044]
124. Ye Y, Rape M. Building ubiquitin chains: E2 enzymes at work. *Nat Rev Mol Cell Biol*. 2009; 10(11):755–764. [PubMed: 19851334]
125. Hochstrasser M. Ubiquitin ligation without a ligase. *Dev Cell*. 2007; 13(1):4–6. [PubMed: 17609103]
126. Hoeller D, Hecker CM, Wagner S, Rogov V, et al. E3-independent monoubiquitination of ubiquitin-binding proteins. *Mol Cell*. 2007; 26(6):891–898. [PubMed: 17588522]
127. Sorkin A. Ubiquitination without E3. *Mol Cell*. 2007; 26(6):771–773. [PubMed: 17588512]
128. Iyer LM, Burroughs AM, Aravind L. The prokaryotic antecedents of the ubiquitin-signaling system and the early evolution of ubiquitin-like beta-grasp domains. *Genome Biol*. 2006; 7(7):R60. [PubMed: 16859499]
129. Jurgenson CT, Begley TP, Ealick SE. The structural and biochemical foundations of thiamin biosynthesis. *Annu Rev Biochem*. 2009; 78:569–603. [PubMed: 19348578]
130. Schwarz G, Mendel RR, Ribbe MW. Molybdenum cofactors, enzymes and pathways. *Nature*. 2009; 460(7257):839–847. [PubMed: 19675644]
131. Godert AM, Jin M, McLafferty FW, Begley TP. Biosynthesis of the thioquinolobactin siderophore: an interesting variation on sulfur transfer. *J Bacteriol*. 2007; 189(7):2941–2944. [PubMed: 17209031]

132. Jurgenson CT, Burns KE, Begley TP, Ealick SE. Crystal structure of a sulfur carrier protein complex found in the cysteine biosynthetic pathway of *Mycobacterium tuberculosis*. *Biochemistry*. 2008; 47(39):10354–10364. [PubMed: 18771296]
133. O'earry SE, Jurgenson CT, Ealick SE, Begley TP. O-phospho-L-serine and the thiocarboxylated sulfur carrier protein CysO-COSH are substrates for CysM, a cysteine synthase from *Mycobacterium tuberculosis*. *Biochemistry*. 2008; 47(44):11606–11615. [PubMed: 18842002]
134. Burroughs AM, Balaji S, Iyer LM, Aravind L. Small but versatile: the extraordinary functional and structural diversity of the beta-grasp fold. *Biol Direct*. 2007; 2:18. [PubMed: 17605815]
135. Burroughs AM, Iyer LM, Aravind L. Natural history of the E1-like superfamily: implication for adenylation, sulfur transfer, and ubiquitin conjugation. *Proteins*. 2009; 75(4):895–910. [PubMed: 19089947]
136. Lehmann C, Begley TP, Ealick SE. Structure of the *Escherichia coli* ThiS-ThiF complex, a key component of the sulfur transfer system in thiamin biosynthesis. *Biochemistry*. 2006; 45(1):11–19. [PubMed: 16388576]
137. Schmitz J, Wuebbens MM, Rajagopalan KV, Leimkuhler S. Role of the C-terminal Gly-Gly motif of *Escherichia coli* MoaD, a molybdenum cofactor biosynthesis protein with a ubiquitin fold. *Biochemistry*. 2007; 46(3):909–916. [PubMed: 17223713]
138. Dahl JU, Urban A, Bolte A, Sriyabhaya P, et al. The identification of a novel protein involved in molybdenum cofactor biosynthesis in *Escherichia coli*. *J Biol Chem*. 2011; 286(41):35801–35812. [PubMed: 21856748]
139. Zhang W, Urban A, Mihara H, Leimkuhler S, et al. IscS functions as a primary sulfur-donating enzyme by interacting specifically with MoeB and MoaD in the biosynthesis of molybdopterin in *Escherichia coli*. *J Biol Chem*. 2010; 285(4):2302–2308. [PubMed: 19946146]
140. Mueller EG. Trafficking in persulfides: delivering sulfur in biosynthetic pathways. *Nat Chem Biol*. 2006; 2(4):185–194. [PubMed: 16547481]
141. Humbard MA, Miranda HV, Lim JM, Krause DJ, et al. Ubiquitin-like small archaeal modifier proteins (SAMPs) in *Haloferax volcanii*. *Nature*. 2010; 463(7277):54–60. [PubMed: 20054389]
142. Miranda HV, Nembhard N, Su D, Hepowit N, et al. E1- and ubiquitin-like proteins provide a direct link between protein conjugation and sulfur transfer in archaea. *Proc Natl Acad Sci U S A*. 2011; 108(11):4417–4422. [PubMed: 21368171]
143. Zhang Y, Gladyshev VN. Molybdoproteomes and evolution of molybdenum utilization. *J Mol Biol*. 2008; 379(4):881–899. [PubMed: 18485362]
144. Burroughs AM, Iyer LM, Aravind L. Functional diversification of the RING finger and other binuclear treble clef domains in prokaryotes and the early evolution of the ubiquitin system. *Mol Biosyst*. 2011; 7(7):2261–2277. [PubMed: 21547297]
145. Hartman AL, Norais C, Badger JH, Delmas S, et al. The complete genome sequence of *Haloferax volcanii* DS2, a model archaeon. *PLoS One*. 2010; 5(3):e9605. [PubMed: 20333302]
146. Makarova KS, Koonin EV. Archaeal ubiquitin-like proteins: functional versatility and putative ancestral involvement in tRNA modification revealed by comparative genomic analysis. *Archaea*. 2010; 2010
147. Wagner SA, Beli P, Weinert BT, Nielsen ML, et al. A proteome-wide, quantitative survey of in vivo ubiquitylation sites reveals widespread regulatory roles. *Mol Cell Proteomics*. 2011; 10(10):M111 013284.
148. Yang XJ. Multisite protein modification and intramolecular signaling. *Oncogene*. 2005; 24(10):1653–1662. [PubMed: 15744326]
149. Baker RT, Board PG. The human ubiquitin gene family: structure of a gene and pseudogenes from the Ub B subfamily. *Nucleic Acids Res*. 1987; 15(2):443–463. [PubMed: 3029682]
150. Lund PK, Moats-Staats BM, Simmons JG, Hoyt E, et al. Nucleotide sequence analysis of a cDNA encoding human ubiquitin reveals that ubiquitin is synthesized as a precursor. *J Biol Chem*. 1985; 260(12):7609–7613. [PubMed: 2581967]
151. Wiborg O, Pedersen MS, Wind A, Berglund LE, et al. The human ubiquitin multigene family: some genes contain multiple directly repeated ubiquitin coding sequences. *EMBO J*. 1985; 4(3):755–759. [PubMed: 2988935]

152. Ozkaynak E, Finley D, Solomon MJ, Varshavsky A. The yeast ubiquitin genes: a family of natural gene fusions. *EMBO J.* 1987; 6(5):1429–1439. [PubMed: 3038523]
153. Katz EJ, Isasa M, Crosas B. A new map to understand deubiquitination. *Biochem Soc Trans.* 2010; 38(Pt 1):21–28. [PubMed: 20074029]
154. Kirkland PA, Gil MA, Karadzic IM, Maupin-Furrow JA. Genetic and proteomic analyses of a proteasome-activating nucleotidase A mutant of the haloarchaeon *Haloferax volcanii*. *J Bacteriol.* 2008; 190(1):193–205. [PubMed: 17965165]
155. Kirkland PA, Reuter CJ, Maupin-Furrow JA. Effect of proteasome inhibitor clasto-lactacystin-beta-lactone on the proteome of the haloarchaeon *Haloferax volcanii*. *Microbiology.* 2007; 153(Pt 7):2271–2280. [PubMed: 17600071]
156. Leidel S, Pedrioli PG, Bucher T, Brost R, et al. Ubiquitin-related modifier Urm1 acts as a sulphur carrier in thiolation of eukaryotic transfer RNA. *Nature.* 2009; 458(7235):228–232. [PubMed: 19145231]
157. Shigi N, Sakaguchi Y, Asai S, Suzuki T, et al. Common thiolation mechanism in the biosynthesis of tRNA thiouridine and sulphur-containing cofactors. *EMBO J.* 2008; 27(24):3267–3278. [PubMed: 19037260]
158. Nunoura T, Takaki Y, Kakuta J, Nishi S, et al. Insights into the evolution of Archaea and eukaryotic protein modifier systems revealed by the genome of a novel archaeal group. *Nucleic Acids Res.* 2011; 39(8):3204–3223. [PubMed: 21169198]

Fig. 11.1. 20S proteasome core particles (CPs)

CPs are composed of four stacked heptameric rings of α - and β -type subunits (indicated by α_7 and β_7 , respectively) that form a cylindrical structure. The central channel of the CP is accessed by gated pores on each end of the cylinder and connects three interior chambers. Proteolytic active sites formed by β -type subunits (indicated in red) line the central chamber that is flanked by two antechambers. CPs of (a) the eukaryote *Saccharomyces cerevisiae* and (b) the archaeon *Thermoplasma acidophilum* are presented as examples. (Figure modified from [22, 23] with permission)

Fig. 11.2. Mechanism of proteasomes and other related Ntn-hydrolyases in the hydrolysis of an amide bond

Substrate, *shaded in blue*; Ntn active site residue, *shaded in red*; oxyanion hole, *shaded in green*; charged state, indicated by + or – in *red font*; Y, oxygen or sulfur; X, nitrogen or oxygen; R-R', substrate bond cleaved (comparable to the P1–P1' peptide bond cleaved by proteases as depicted in the *inset*) (Figure modified from [26] with permission)

Fig. 11.3. ATP-dependent proteasomes of eukaryotes compared to archaea

Proteasomal CPs can associate on each end of their cylindrical structure with heptameric rings of AAA + proteins including eukaryotic Rpt1-6 and archaeal PAN. Rpt1-6 are AAA + subunits of the 19S regulatory particle (RP) that associates with CPs to form 26S proteasomes. In yeast, the 19S RP can be dissociated into lid and base subcomplexes by deletion of the *RPN10* gene. PAN is an AAA ATPase that forms a hexameric ring and associates with CPs *in vitro*. PAN is common to many but not all archaea

Fig. 11.4. Eukaryal proteasomal CPs are gated

In X-ray crystal structures of eukaryotic (e.g., yeast and bovine) CPs, the central channel used for protein substrate entry is gated by the N-terminal tails of α -type subunits (Figure modified from [67] with permission)

Fig. 11.5. Archaeal proteasomal CPs are gated

Cryo-electron microscopy reveals: (a) sites within the intersubunit pockets of α subunits in the TaCP that are bound by peptides mimicking the C-terminal tail of MjPAN and (b) structures of the TaCP gate in the closed and open forms induced by these peptides (Figure modified from [85] with permission)

Fig. 11.6. Archaeal PAN subdomain structures provide a foundation for modelling the overall structure of an energy-dependent proteasome

(a) and (b) Subcomplex I and II generated by limited proteolysis of MjPAN are indicated with the N-terminal coiled-coil (CC), OB fold and AAA domains highlighted. (c) Evidence supports the docking of MjPAN C-terminal tails with pockets formed at the α - α intersubunit interface of TaCPs. The N-terminal CC-domains are assumed to be distal to the CP and act as tentacles that grab protein substrates, thus, enabling the Ar- Φ loop within subcomplex II to grip regions of the substrate that extend through the pore formed by the OB fold domain. The Ar- Φ loop is also thought to undergo ATP-fuelled conformational changes resulting in pulling and tugging at the protein substrate while the OB fold domain provides a passive force that blocks the movement of folded protein structure through the pore. As the protein is unfolded by this mechanism, it is translocated through the ATPase into the central channel of the CP for degradation (Figure modified from [84] with permission)

Fig. 11.7. Ubiquitylation and sampylation

(a) Eukaryotes use an elaborate E1-E2-E3 mediated mechanism for the attachment of ubiquitin to protein substrates. (b) Similarly to eukaryotes, small archaeal ubiquitin-like modifier proteins (termed *SAMPs*) can form protein conjugates in the archaeon *H. volcanii* by a pathway that is dependent upon the synthesis of an E1 homolog (termed *UbaA*). Based on genome sequence, E2 and E3 homologs are not predicted for this pathway. *UbaA* and *SAMP* proteins appear to also be linked to sulfur incorporation pathways (such as the biosynthesis of MoCo and thiolated tRNA) that require an E1-type adenylation reaction for the formation of a thiocarboxylated sulfur carrier protein with a ubiquitin-type β -grasp fold