PROTEASOMES: PERSPECTIVES FROM THE ARCHAEA

Julie A. Maupin-Furlow, Małgorzata A. Gil, Ivanka M. Karadzic, Phillip A. Kirkland and Christopher J. Reuter

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

TABLE OF CONTENTS

1. Abstract
2. Introduction
3. 20S proteasome structure and function
   3.1. Subunit and isoform complexity
   3.2. Mechanism of peptide bond hydrolysis
   3.3. Mechanism of polypeptide chain hydrolysis
4. Proteasome-associated regulatory particles
   4.1. 19S cap and COP9 signalsome
   4.2. Archaeal Rpt homologs
   4.3. Cdc48 homologs
   4.4. Non-ATPase modulators
5. Energy-dependent proteolysis
   5.1. Substrate recognition and binding
   5.2. Substrate unfolding
   5.3. Channel gating
   5.4. Substrate translocation
6. Regulation and modification of proteasome subunits
7. Proteasome assembly
8. Perspectives
9. Acknowledgments
10. References

1. ABSTRACT

The development of whole systems approaches to microbiology (e.g., genomics and proteomics) has facilitated a global view of archaeal physiology. Surprisingly, as archaea respond to environmental signals, the majority of protein concentration changes that occur are not reflected at the mRNA level. This incongruity highlights the importance of post-transcription control mechanisms in these organisms. One of the central players in proteolysis is the proteasome, a multicatalytic energy-dependent protease. Proteasomes serve both proteolytic and non-proteolytic roles in protein quality control and in the regulation of cell function. The proteolytic active sites of these enzymes are housed within a central chamber of an elaborate nanocompartment termed the 20S proteasome or core particle. Axial gates, positioned at each end of this particle, restrict the type of substrate that can access the proteolytic active sites. Assortments of regulatory AAA complexes are predicted to recognize/bind and unfold substrate proteins, open the axial gates, and translocate substrate into the 20S core particle.

2. INTRODUCTION

Proteasomes are energy-dependent proteases found in all three domains of life: Bacteria, Archaea and Eucarya (1). These enzymes maintain quality control by degrading misfolded and denatured proteins in response to cell stress and general protein turnover (2). Proteasomes also play central roles in the regulation of many cellular processes such as cell division, metabolism, and DNA repair (3-5). A growing body of evidence reveals that proteasomes are also intimately involved in controlling the distribution, abundance, and activity of components of the transcription machinery (6, 7). In addition, a functional link between proteasomes and components of translation initiation (eIF3, eucaryotic translation initiation factor 3) have been identified (8). Non-proteolytic roles have also been demonstrated for proteasomes in nucleotide excision repair (9, 10), transcription elongation (11, 12), and cell cycle control (13).

The development of whole systems approaches (e.g., genomics, proteomics) to microbiology has provided insight into the central role proteasomes are likely to play in the physiology of archaea. This is highlighted by the universal distribution of proteasome homologs in archaeal genomes, including that of the recently discovered archaeal parasite Nanoarchaeum equitans, one of the smallest genomes to date (table 1). Based on the apparent absence of other cytosolic energy-dependent proteases, proteasomes are predicted to be the central energy-dependent proteases within the archaeal cell (figure 1). The catalytic core of the proteasome (20S core particle) in combination with various AAA regulatory proteins (e.g., Pan and VCP) is expected to mediate the quality control and regulated turnover of most cytosolic proteins. The 20S proteasome may also associate with AAA proteins located in the cell membrane to aid the archaeal-type Lon protease in the retrograde translocation and degradation of membrane-associated proteins. The central role proteasomes seem to play in the archaea
Archaeal proteasomes

Table 1. Distribution of 20S proteasome, Pan, and Cdc48 homologs in archaea

<table>
<thead>
<tr>
<th>Organism</th>
<th>20S α-type (COG0638)</th>
<th>20S β-type (COG0638)</th>
<th>Pan (COG1222)</th>
<th>VCP/Cdc48 (COG0464)</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Archaeoglobus fulgidus DSM 4304</td>
<td>AF0490*</td>
<td>AF0481</td>
<td>AF1976</td>
<td>AF0477</td>
<td>161</td>
</tr>
<tr>
<td>Archaeoglobus fulgidus</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ferroplasma acidarmanus</td>
<td>u.d.</td>
<td>u.d.</td>
<td>Fac1393</td>
<td>GenBank</td>
<td></td>
</tr>
<tr>
<td>Halobacterium sp. NRC-1</td>
<td>Contig88_15652_15780</td>
<td>Contig137_16344_17066</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Haloferax volcanii DS2</td>
<td>AAD35404</td>
<td>AAD35405</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Methanothermobacter thermoautotrophicus str. Delta H</td>
<td>MTH686</td>
<td>MTH1202</td>
<td>MTH1639</td>
<td></td>
<td>145, 172, 73</td>
</tr>
<tr>
<td>Pyrococcus furiosus DSM 3638</td>
<td>PAB0417</td>
<td>PAB2199</td>
<td>PAB1867</td>
<td></td>
<td>168</td>
</tr>
<tr>
<td>Pyrococcus furiosus</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thermoplasma acidophilum</td>
<td>Tai0612</td>
<td>—</td>
<td>Tai175</td>
<td></td>
<td>171</td>
</tr>
<tr>
<td>Thermoplasma volcanii</td>
<td>TVN0304</td>
<td>TVN0636</td>
<td></td>
<td></td>
<td>174</td>
</tr>
<tr>
<td>Crenarchaeota</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aeropyrum pernix K1</td>
<td>APE1449</td>
<td>APE0521</td>
<td>APE2012</td>
<td></td>
<td>175</td>
</tr>
<tr>
<td>Pyrococcus furiosus</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sulfolobus acidocaldarius* DSM 639</td>
<td>SS02673</td>
<td>ST0946</td>
<td>ST0330</td>
<td>ST0356</td>
<td>177</td>
</tr>
<tr>
<td>Nanoarchaeota</td>
<td>u.d.</td>
<td>u.d.</td>
<td></td>
<td></td>
<td>178</td>
</tr>
<tr>
<td>Nanoarchaeum equitans KIND-M</td>
<td>NEK586</td>
<td>NEK2171</td>
<td>NEK252</td>
<td>NEK516</td>
<td>179</td>
</tr>
</tbody>
</table>

Asterisks (*) indicates complete genome sequence not available, u.d., undeetermined, d sequence submitted directly to GenBank, Open reading frames determined from H. marismortui and H. volcanii unfinished genomes available at http://zdna2.umbi.umd.edu, PanA and PanB sequences unpublished, — no homolog identified based on complete genome sequence.

contrasts with that of bacteria, which encode multiple energy-dependent proteases in the cytosol (e.g. Lon, Clp) that provide redundant functions (14). However, even with this redundancy, 20S proteasomes are critical to the survival of some bacteria after exposure to stress (15).

Recent evidence highlights the importance of post-transcriptional regulation in the archaea. Examination of Halobacterium sp. NRC-1 at the global level reveals that the majority of protein concentration changes that occur are not reflected at the mRNA level (16). Hence, there appears to be a significant degree of post-transcriptional control in this haloarchaeon (and likely other archaea), which may be mediated by proteases such as the proteasome. Elucidating the role proteasomes play in this regulation as well as in general protein turnover is expected to have far reaching...
Archaeal proteasomes

Figure 1. The centrality of archaeal proteasomes in protein quality control and regulated protein turnover. The 20S core particle and AAA+ (e.g., Pan and VCP) homologs are predicted to be central players in the quality control and post-transcriptional regulation of cytosolic proteins. Proteasomes, in association with their regulatory particles, are expected to serve not only proteolytic roles but also to mediate non-proteolytic processes such as protein remodeling. Proteasomes may also assist in the hydrolysis of membrane-associated proteins. GenBank protein accession numbers are indicated for a select number of protease, chaperone, and regulatory protein homologs based on genome sequence of Halobacterium sp. NRC-1 (162). Protease homologs include: 20S proteasome (this review); TET, tetrahedral aminopeptidase (180); PfpI, P. furiosus protease I homolog (181); Lon, archaeal-type Lon protease (182); HtpX, membrane metalloprotease with a cytosolic active site (183); HtrA (DegQ), serine protease with twin-arginine motif (184). Chaperone and protease regulator homologs include: Pan, proteasome-activating nucleotidase and VCP(Cdc48) (this review); prohibitin (HflX, HflC, HflK), (185); thermosome (Hsp60), prefoldin, DnaK (Hsp70), DnaJ (Hsp40), GrpE, and Hsp20 (186).

3. 20S PROTEASOME STRUCTURE AND ACTIVITY

The 20S proteasome or 20S core particle refers to the multicatalytic protease component of proteasomes. This complex is responsible for many facets of proteolysis within the cell and is universal among archaea, eucaryotes, and Gram-positive actinomycetes (1). Much is known about the detailed structure of 20S proteasomes thanks to a number of X-ray diffraction studies (17-20). In general, 20S proteasomes have a highly conserved barrel-like structure formed by four stacked heptameric rings of subunits from a family of related proteins, α and β (21). The outer two rings are composed of α-type subunits and the inner two rings are of β-type subunits (22). The proteolytic active sites are located at the N-termini of subunits and line an inner chamber, flanked by two antechambers which are accessed through a central channel (17, 23). Axial gates, positioned at each end of the barrel, limit the ability of globular substrates to enter the central channel (24-27).

3.1. 20S proteasome subunit and isoform complexity

Although the basic structure of 20S proteasomes is conserved, modest differences in the complexity of subunits exist among organisms. Most prokaryotic 20S proteasomes are composed of one α-type and one β-type...
Archael proteasomes

subunit; however, some contain 3 to 4 different subunits (28, 29). In contrast, eukaryal 20S proteasomes are characterized by 7 different α-type and 7 different β-type subunits (30, 31). Comparative genomics predicts these β subtypes differentiated earlier than did the α subtypes (32).

The number of 20S proteasome isoforms also varies among organisms. Primitive eucaryotes (e.g. yeast, Caenohabditis elegans) synthesize only one 20S proteasome. In contrast, higher eucaryotes have constitutive housekeeping and inducible ancillary 20S proteasomes. For example, the immunoproteasome of vertebrates is induced by IFN-γ (33), and a spermatogenesis-specific proteasome has been identified in insects (34). Multiple proteasomes are also common in plants where up to 23 different α- and β-type genes have been identified (35), and mixtures of proteasome isoforms have been purified (36). Recent dissection of 20S proteasomes from human erythrocytes reveals at least 32 different subunit types with many subunits modified post-translational (37). Surprisingly, the halorarchaeon Haloferax volcanii synthesizes at least two different 20S proteasomes including: a constitutive complex of only α1 and β subunits (38) and an ancillary asymmetric complex of homooligomeric rings arranged in an αββα2 configuration (29).

3.2. Mechanism of peptide bond hydrolysis

20S proteasomes belong to the amino-terminal (Ntn) hydrolyase family (39). A mechanism similar to serine proteases is envisioned in which the N-terminal threonine hydroxyl group of β subunits initiates hydrolysis by attacking the carbonyl carbon of a peptide bond (40, 41). This results in the formation of a tetrahedral intermediate that collapses into an acyl-enzyme and releases the peptide product generated downstream of the cleavage site. Nucleophilic attack of this acyl-enzyme intermediate by water yields free enzyme and release of the second peptide product upstream of the cleavage site. In contrast to serine proteases, however, 20S proteasomes require the additional methyl group of threonine to support rapid rates of protein breakdown (41, 42).

3.3. Mechanism of polypeptide chain hydrolysis

The mechanism of how 20S proteasomes degrade polypeptide chains into short peptides is not fully understood. Products range from 3 to 30 amino acids in length and fit a log-normal distribution (43, 44). The number of catalytic sites does not influence the average length of product (45); however, regulatory components that associate with and modify the axial gates do (44, 46). Thus, the dimensions of the 20S proteasome axial gates are likely to play a role in determining the size of products released. The rate-limiting step is entry of substrate protein into the 20S proteasome and/or translocation of this substrate to the proteolytic active sites. This is based on the observation that the rate of bond cleavage decreases with increasing chain length of unfolded polypeptide (41).

An intrinsic feature of 20S proteasomes is the processive degradation of proteins from free N- or C-termini and may be due to a trapping of the substrate protein inside the 20S cylinder (47, 48). Regulatory components are not required for processive degradation but appear to be necessary for the rates of degradation required in the cell (45). Interestingly, 20S proteasomes degrade some proteins by non-processive hydrolysis (49) and do not require substrates to have free N- or C-termini (50). In fact, some substrates (e.g., NF-xB p105, NF-xB p100) are predicted to have disordered, internal loops that enter the axial channel of 20S proteasomes resulting in substrate processing and activation (50). This model is supported by the follow studies of 20S proteasomes: endoproteolytic activity has been detected using green fluorescent protein (GFP) fusions (50), the open gate conformation is predicted to accommodate β-hairpin structures (24, 25), and three extended polypeptide chains can be modeled to fit within the central proteolytic chamber (51).

There is growing evidence that 20S proteasomes do not cleave proteins at random; instead, there are preferred amino acid motifs that are recognized and hydrolyzed (proline at P4, leucine at P1 and amino acids that promote β-turns at P1') (45, 46, 52). In addition, allosteric binding of effector molecules to non-catalytic sites influences protein degradation (53-56). Hydrophobic peptides act as positive effectors and promote an open gate conformation of the axial channel of 20S proteasomes that stimulates peptidase activity (56). This open gate transition is consistent with the two distinct, inter-converting forms of 20S proteasomes observed by atomic force microscopy (57). These two allosteric states include R (closed-gate barrel-like) and T (open-gate cylinder-like) states in which the T state is stabilized by hydrophobic substrates (57).

4. PROTEASOME-ASSOCIATED REGULATORY PARTICLES

A variety of regulatory components associate with 20S proteasomes including: both type I and type II AAA proteins (ATPases associated with various cellular activities) (58-60) as well as non-ATPase modulators.

4.1. 19S cap and COP9 signalsome

The 19S cap (PA700) is a proteasome regulatory complex of eukaryotes that is composed of at least 17 subunits. In yeast, the 19S cap can be separated into two multisubunit substructures including: a “lid” composed of nine Regulatory particle non-ATPases (Rpn) subunits and a “base” composed of six Regulatory particle triple-A type I proteins (Rpt) and two Rpn proteins (61, 62). The 20S core and 19S cap together form 26S proteasomes, which recognize and degrade substrates tagged with ubiquitin (Ub) (63). “26S proteasome” commonly refers to either a 30S complex consisting of a 20S particle capped at both ends by 19S complexes or a 26S particle capped only at one end (64). The 20S core and base alone can degrade globular proteins; however, the presence of the lid, in addition to the 20S core and base, is essential for the degradation of ubiquitin-tagged proteins (62).

Duplicated genes encoding 19S cap homologs are present in a variety of organisms including Trypanosoma, Arabidopsis, and rice (65-67). The differential expression of these homologs appears to have diversified the
Archaal proteasomes

functional capacity of proteasomes. Mixtures of 19S caps with different Rpt isoforms have been identified, and the relative amounts and expression patterns of these isoforms vary in a tissue-specific manner (68). A more distantly related homolog of the 19S cap, the COP9 signalosome or CSN, also interacts with the ubiquitin-proteasome system to regulate protein turnover (69, 70).

4.2. Archaeal Rpt homologs

Archaea encode Rpt homologs that resemble the base of the 19S cap in both structure and function (71, 72). These proteins have been designated Pan for Proteasome-activating nucleotidases. Most archaea encode a single Pan protein while some encode two highly related paralogs (table 1). Interestingly, *Halofex volca*nti*es* synthesizes at least three Pan isoforms including both homo and heterologimeric complexes of PanA and PanB (Reuter and Maupin-Furlow, unpublished results). Since Pan proteins alone are able to catalyze the unfolding of substrate proteins (73), they appear to be directly involved in substrate recognition. It will be interesting to see whether the diversification of Pan and other AAA family members (e.g. Cdc48 homologs discussed below) enhances the number of different motifs recognized as substrates for degradation by archaeal 20S proteasomes. If so, the haloarchaea encode a tremendous number of AAA proteins (table 1) that may be used in different combinations with 20S proteasome isoforms for the regulated turnover of proteins. This would be a new paradigm with analogy to the model recently proposed for haloarchaeal gene regulation in which a diversity of transcription factors interact in up to 42 different combinations to recognize a large set of promoters (74).

The Pan protein (MJ1176) from *Methanocaldococcus jannaschii* is the most thoroughly characterized archaeal Rpt homolog. It forms an irregular ring-shaped dodecameric ATPase of 600 kDa (71, 72). In the presence of ATP or CTP this Pan stimulates 20S proteasome-dependent hydrolysis of proteins including casein and GFP-SsrA (GFP with an 11-residue C-terminal peptide tag) (71, 72). Substrate binding to this triple-A ATPase *in vitro*, activates ATP hydrolysis, which successively promotes substrate unfolding, opening of the axial gate, and possibly substrate translocation into the 20S core (27, 75) (see below).

In addition to Pan, small proteins with Jab1/MPN* motifs common to the eucaryal 19S cap are predicted for archaea and bacteria (76). Although the Jab1/MPN* motif has been implicated in the de-Ub activity of eucaryal 26S proteasomes (77), an archaeal protein (AF2198) with this motif does not appear to hydrolyze peptide bonds. Instead this archaeal protein is proposed to catalyze the removal of lysine side chain modifications (78).

4.3. Cdc48 homologs

Not all archaea encode Rpt homologs (*i.e. Thermoplasma* sp. and *Pyrobaculum aerophilum*) (table 1), which has lead to the suggestion that archaeal Cdc48 (VCP, VAT, p97) homologs may also facilitate proteasome-mediated degradation. Cdc48 homologs are type II AAA proteins found in all three domains and purify as barrel-like structures of two stacked hexameric rings (79, 80) with chaperone-like activity (81, 82). In eucaryotes, the p97 protein in complex with Ufd1 and Npl4 has been implicated in the ATP-dependent movement of polyUb substrates into the cytosol via retrotranslocation from the ER for proteasome-mediated degradation (2, 83, 84).

4.4. Non-ATPase modulators

There are several non-ATPases that modulate 20S proteasome activity. Most of these have been isolated from eucaryotic cells including the IFN-γ inducible 11S (PA28, PA26, REG) activator (85) as well as the CF-2, β-amyloid, PI31, and Hsp90 inhibitors (86-89). An inhibitor of the Ca2+-dependent proteinase activity of an archaeal 20S proteasome has also been described (90). Of these, the mechanism of activation of the 11S regulator is best understood and is mediated by a loop which opens the axial gates of 20S proteasomes (91, 92). Whether the inhibitors cap the axial pores, plug the channel and or promote conformational changes in the substrate binding sites of 20S proteasomes is not clear.

5. ENERGY-DEPENDENT PROTEOLYSIS

Regulatory AAA particles may serve multiple roles in stimulating the energy-dependent degradation of proteins by 20S proteasomes. These include substrate recognition and binding, substrate unfolding, opening the axial gates of 20S proteasomes, and translocation of unfolded substrates into the proteolytic chamber of 20S proteasomes. A growing list of AAA* protein structures [i.e. HsIU (93-95), ClpA (96, 97), ClpX (98), FtsH (99)] has enhanced our understanding of how these molecular machines couple energy to the unfolding and/or remodeling of proteins for proteolysis.

5.1. Substrate recognition and binding

Self-compartmentalized proteases such as 20S proteasomes rely upon upstream energy-dependent enzymes for substrate discrimination. In eucaryotes, 26S proteasomes recognize substrates covalently linked to polyUb chains (100). Ubiquitination is an energy-dependent process mediated by a series of enzymes including Ub-activating (E1), Ub conjugating/carrier (E2) and Ub protein ligases (E3). The specificity of a Ub-proteolytic pathway is conferred by the E3 ligase (101). Once a protein is modified by poly-Ub, the Rpt5 (102) and Rpn10 (103) subunits of the 26S proteasome can bind. In addition, the Ub-like and Ub-associated domains of proteins can interact with E3 Ub ligases and 26S proteasomes to provide a link between the ubiquitination and degradation of substrates (104). For example, Rpn1 and Rpn2 subunits of 26S proteasomes bind the Ub-like domains of the poly-Ub binding proteins Rad23 and Dsk2 (105-107). In addition, the poly-Ub binding activity of the N-terminus of p97-VCP, a Cdc48 homolog, is necessary for targeting a subset of proteins for degradation by 26S proteasomes (108, 109). It should be noted, however, that not all proteins degraded by 26S proteasomes are
Archaeal proteasomes

conjugated to Ub (e.g. ornithine decarboxylase, CDK inhibitor p21\textsuperscript{WAF1/CIP1}) (110).

In archaea, as well as bacteria, homologs of the 19S lid subunits (Rpn) and other enzymes essential to the ubiquitin pathway have not been identified. Therefore, it is not surprising that archaea do not use the ubiquitin-labelling pathway to tag substrates for degradation. In fact, a pathway for proteolytic targeting has yet to be determined. Based on analogy to structurally related proteases such as Clp, it is anticipated that proteasomal ATPases such as Pan directly recognize and bind non-ub substrates. The N-terminal coiled-coil domain of Pan is proposed to mediate substrate binding as well as subunit interaction (111). However, it is not yet known the mechanism by which Pan recognizes substrates for degradation. Interestingly, the distantly related Clp ATPase (ClpX) has both a processing site which recognizes degradation signals at or near the C- or N- terminus of proteins as well as tethering sites which interact with substrate delivery/adaptor proteins (e.g. UmuD, SspB, RssB) (112-114). These adaptor proteins appear to improve the efficiency of degradation at low substrate concentration via tethering to the proteolytic complex. Similarly, adaptor proteins may also be needed for archael proteasome function.

5.2. Substrate unfolding

The hexameric ring-like structures formed by many AAA proteins appear to be physiologically advantageous in the catalysis of protein unfolding, a process required for entry of substrate into the 20S proteasome. The central pore and internal cavity/chamber of the ring structure may enable cells to sequester substrate proteins from the cytosol during the unfolding process. Consistent with this, unfoldase and/or chaperone activity has been detected for several proteasome-associated AAA proteins (115). Pan catalyzes the ATP-dependent unfolding of GFP-SsrA, a step required for degradation of this protein (27). However, the proteasomes require Pan and hydrolysable ATP for the degradation of unfolded GFP-SsrA (27). However, the translocation step does not appear to increase the overall amount of ATP hydrolyzed per molecule of protein degraded (27).

5.4. Substrate translocation

Proteasome-associated AAA regulators are likely to assist in translocation of unfolded protein through the axial pore of 20S proteasomes for hydrolysis in the central chamber. The pore of the 19S cap forms a continuous passage with the axial channel of the 20S core and is presumed to assist in the transfer of unfolded substrate proteins (120). It is currently unknown whether substrate translocation is an energy-dependent step. ‘Open-gate’ 20S proteasomes require Pan and hydrolysable ATP for the degradation of unfolded GFP-SsrA (27). However, the translocation step does not appear to increase the overall amount of ATP hydrolyzed per molecule of protein degraded (27).

6. REGULATION AND MODIFICATION OF PROTEASOME SUBUNITS

Alterations in the levels of 20S proteasome and proteasome associated AAA regulators play a role in regulating proteasome activity. In eucaryotes, these changes occur after proteasome inhibition (121-123), IFN-γ induction (124), during rapid growth (125), during differentiation and development (126-128), after heat shock and canavanine treatment (129), and after transition from log to stationary phase (130-132). Rpn4 appears to be a major player in the transcriptional control of balanced levels of proteasome subunits in yeast. Rpn4 is not only a subunit of 26S proteasomes, but also a transcriptional activator that binds to a common cis-element (proteasome-associated control element or PACE) upstream of almost all of 26S proteasome genes (133). Once Rpn4 induces proteasome formation, it is destroyed by mature proteasomes in an autoregulatory feedback mechanism (134). One notable exception to Rpn4-mediated control is Rpn10, which is also the only 26S proteasome subunit found at significant levels free in the cytosol (135, 136).
Archael proteasomes

<table>
<thead>
<tr>
<th>Protein ID</th>
<th>Deduced N-terminal sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>SSO0278</td>
<td>MEELP-ATAVGL</td>
</tr>
<tr>
<td>BAB65301</td>
<td>MEELP-ATAVGL</td>
</tr>
<tr>
<td>PAF3595</td>
<td>MEELP-ATAVGL</td>
</tr>
<tr>
<td>APE0507</td>
<td>MDSYTHOCWAKTVIQTPGRSEQYIYGPASDLYASGCCCLESFAG-ATAVGL</td>
</tr>
<tr>
<td>PAF0807</td>
<td>MTTTIGI</td>
</tr>
<tr>
<td>NEK221</td>
<td>MTTTIGI</td>
</tr>
<tr>
<td>PB02199</td>
<td>MNNKRTG-TTTVG</td>
</tr>
<tr>
<td>DB0245</td>
<td>MNNKRTG-TTTVG</td>
</tr>
<tr>
<td>HJ1237</td>
<td>MNDTVKIG-TTTVG</td>
</tr>
<tr>
<td>Fac10876</td>
<td>MEVLTIG-TTTVG</td>
</tr>
<tr>
<td>TA0612</td>
<td>MnQTLTIG-TTTVG</td>
</tr>
<tr>
<td>TYN0663</td>
<td>MnQTLTIG-TTTVG</td>
</tr>
<tr>
<td>M83173</td>
<td>MMOQTVKIG-TTTVG</td>
</tr>
<tr>
<td>mPO694</td>
<td>MMDTVKIG-TTTVG</td>
</tr>
<tr>
<td>MTH1202</td>
<td>MKNKTVKIG-TTTVG</td>
</tr>
<tr>
<td>PSMB_MetaTE</td>
<td>MMDTVKIG-TTTVG</td>
</tr>
<tr>
<td>PH1402</td>
<td>MLQLTKEFKTG-TTTVG</td>
</tr>
<tr>
<td>RAB1667</td>
<td>MLQLTKEFKTG-TTTVG</td>
</tr>
<tr>
<td>MK1226</td>
<td>MKQEOPLTG-TTTVG</td>
</tr>
<tr>
<td>AF0401</td>
<td>MMSEEEKTVG-TTTVG</td>
</tr>
<tr>
<td>Meth0136</td>
<td>MFTMONRKTVG-TTTVG</td>
</tr>
<tr>
<td>BB55470</td>
<td>MEVMTKEIKLG-TTTVG</td>
</tr>
<tr>
<td>APE0521</td>
<td>MGAVCKVKYTEFLGGLGEPAGRLDEPYTVBG-TTTVG</td>
</tr>
<tr>
<td>SSO0766</td>
<td>MNQKLVTETLLLMVIMGNEQKLEIKLG-PPVGL</td>
</tr>
<tr>
<td>Hma_βeta2</td>
<td>MHVERKLLTVDEYPGNLNPDEHSIGHDEENKTVG-PPVGL</td>
</tr>
<tr>
<td>Hma_βeta1</td>
<td>MHDEEKDLLTVDEYPGNLNPDEHSIGHDEENKTVG-PPVGL</td>
</tr>
<tr>
<td>AAD53406</td>
<td>MTTTIDFSGRLDSlingenfVFGELGEFSNADDRADGELKTVKTVG</td>
</tr>
<tr>
<td>VNG0080G</td>
<td>MTTTVGKLTVGLTVG</td>
</tr>
</tbody>
</table>

Figure 2. Deduced N-terminal sequence of archaeal β-type proteasomal proteins. The amino acid residues and number of residues predicted to be removed from the mature β subunit are highlighted in yellow and indicated on the right, respectively. Those β-type proteins which do not have a conserved N-terminal threonine residue are indicated by an asterisk. The N-terminal threonine residues of PSMB_HVO, PSMB_METTE, and TA0162 were determined by sequencing the β subunit of purified 20S proteasomes (38, 146, 172). Protein ID as in Table 1.

Post-transcriptional modification controls proteasome assembly, activity and subcellular location. One of the most common controls is the autocatalytic removal of the N-terminal β propeptide to generate an active site N-terminal threonine. In yeast, the β-propeptide has been shown to protect the Thr1 active site from acetylation and inactivation (137). Acetylation and N-myristoylation have also been observed for other 26S proteasome proteins (138, 139); however, the rationale for these modifications remains to be determined. In eucaryotes, subunits of the 19S cap (i.e., Rpt2, Rpt3, Rpt4, Rpt6, Rpn8), 20S proteasome (i.e., α2 to α7) and Cdc48 homologs (140-143) are phosphorylated in a tissue and organism specific manner. Phosphorylation of Rpt6 has been shown to be linked to the assembly of 26S proteasomes (144).

In archaea, very little is known about the control of 20S proteasome, Pan, or VCP (a Cdc48 homolog) activity. Whether the β-propeptide serves to protect the active site of some archaeal proteasomes is not known. However, it is not required for the biological function of 20S proteasomes from either Pyrococcus aerophilum or Nanoarchaeum equitans, based on its absence from the deduced protein sequence (figure 2). It is possible that the α-subunits of archaeal 20S proteasomes are modified by post-transcriptional mechanisms. One can imagine that this would influence a variety of 20S proteasome functions including axial pore gating and interaction with regulatory proteins. Primary sequence analysis reveals most archaea α subunits have conserved phosphorylation sites (145, 146); however, this has not been confirmed at the protein level. Some archaea α subunits appear to be modified at their N-termini (based on the inability to obtain an N-terminal protein sequence vs. internal sequence) (38, 145). Furthermore, 20S proteasomes purified from Methanosarcina thermophila contain a mixture of α-subunits encoded by the same gene with one of the α subunits four amino acids shorter than the other (146). Recently, the transcription of 20S proteasome genes has been shown to be induced by heat shock in Pyrococcus furiosus (147). Whether this increase in mRNA translates to an increase in proteasome proteins remains to be determined; however, this finding is consistent with the requirement for
Archaeal proteasomes

archaeal 20S proteasomes to be active in order to survive heat shock (148).

7. PROTEASOME ASSEMBLY

The most recent advances in understanding eucaryotic proteasome assembly have been in the identification of the maturation factors Nob1 and Pno1, in addition to the previously identified Ump1 (149). Final assembly of eucaryal proteasomes has been shown to occur in the nucleus (150). Nob1p facilitates the maturation of 20S proteasomes prior to nuclear import (151) and remains associated at the interface of 20S proteasomes and 19S cap (or pre19S) complexes (151, 152). Thus, Nob1 is also predicted to assist in assembly of 26S proteasomes where it is degraded after a tight association has been made between 20S proteasomes and 19S cap complexes. Pno1 associates with Nob1p and assists in transport of proteasome intermediates into the nucleus. Interestingly, although archaea do not have a nucleus, Nob1 and Pno1 homologs are present in this domain. Whether these homologs serve to stabilize 20S proteasome and AAA regulatory particle associations or play other roles in proteasome function is not known.

In yeast, the transition to stationary phase has been shown to induce assembly of doubly capped 26S proteasomes from 20S and 19S complexes (131). As cells reach late stationary phase, there is a down-regulation of proteolytic activity which appears to be mediated by disassembly of 26S proteasomes into 19S cap and 20S core particles (153). The reason for this has yet to be determined but may serve to inhibit proteolysis and/or enable the proteasome particles to play independent roles. For example, the 20S proteasome may hydrolyze certain proteins independent of ubiquitin and ATP (50) while the 19S caps refolds proteins (116, 154). Similar control of proteasome assembly/disassembly in the archaea remains to be determined.

8. PERSPECTIVES

Recent advances in proteomics and genomics have greatly assisted in obtaining a global perspective of the motifs and/or substrates recognized by a variety of energy-dependent proteases. Affinity purification of an inactive ClpP variant enabled the trapping and identification of more than 50 protein substrates with 5 recurring amino acid motifs (155). Comparison of protease mutant (clpP and fisH) and parent strains by 2D-PAGE reveals a multitude of previously unknown substrate proteins and suggests new roles for proteases in cell physiology (156, 157). In addition, mass spectroscopy has enabled the identification of the Ub sites for over 70 proteins (158) and the identification of additional proteins which associate with proteasomes from yeast (159, 160). Similar whole systems approaches, coupled with classical genetics, will most certainly assist in understanding the role proteasomes play in the physiology of archaea.

9. ACKNOWLEDGMENTS

We thank A. Toral for her technical assistance with the manuscript. This work was supported in part by the National Institutes of Health Award R01GM57498 and the Florida Agricultural Experiment Station (Journal Series R-09970).

10. REFERENCES

Archaeal proteasomes

nonproteolytic function of the proteasome is required for the dissociation of cdc2 and cyclin B at the end of M phase. *Genes Dev* 14, 2344-2357 (2000)


Archaeal proteasomes


Archaeal proteasomes

the *Saccharomyces cerevisiae* proteasome regulatory particle. *Trends Biochem Sci* 23, 244-245 (1998)


84. Flierman, D., Y. Ye, M. Dai, V. Chau, & T. A. Rapoport: Polyubiquitin serves as a recognition signal, rather than a ratcheting molecule, during retrotranslocation of proteins across the endoplasmic
Archaal proteasomes


109. Song, C., Q. Wang, & C. C. Li: ATPase activity of p97-VCP: D2 mediates the major enzyme activity and D1
Archaeal proteasomes


133. Mannhaupt, G., R. Schnall, V. Karpov, I. Vetter, & H. Feldmann: Rpn4p acts as a transcription factor by binding to PACE, a nonamer box found upstream of 26S
Archaeal proteasomes


Archaeal proteasomes


Archaeal proteasomes

of the crenarchaeon Sulfolobus solfataricus P2. Proc Natl Acad Sci USA (2001)

178. Kawarabayasi, Y., Y. Hino, H. Horikawa, K. Jin-no, M. Takahashi, M. Sekine, S. Baba, A. Ankai, H. Kosugi, A. Hosoyama: Complete genome sequence of an aerobic thermoacidophilic crenarchaeon, Sulfolobus tokodaii strain7. DNA Res 8, 123-140 (2001)


Abbreviations: AAA and AAA+, subfamily and superfamily of ATPases associated with various cellular activities; Clp, Lon, and FtsH, energy-dependent proteases; COP9, a component of a novel signaling complex of 19S cap homologs which mediates light control of development in Arabidopsis; E1, E2 and E3, Ub-activating, Ub-conjugating/carryer and Ub-ligase proteins; ER, endoplasmic reticulum; GFP-SsrA, green fluorescent protein with C-terminal 11 residue SsrA peptide; GroEL/GroES, group I molecular chaperone; IFN, interferon; MPN+, domain first observed at the N-terminus of the yeast proteins Mpr1p and Pad1p; NF-κB/Rel family of transcription factors in which NF-κB1 p105 is precursor to the mature p50 and NF-κB2 p100 is precursor to the mature p52; Ntn, amino-terminal hydrolase family; Pan, proteasome-activating nucleotidase; Rad23 and Dsk2, proteins with Ub-like domains that bind poly-Ub substrates and 26S proteasomes; Rpt; regulatory particle triple-A type I proteins of the 19S cap; Rpn, regulatory particle non-ATPase proteins of the 19S cap; Ub, ubiquitin; Ump1, Nob1 and Pno1, proteins that facilitate 26S proteasome assembly; VCP, valosin-containing protein with amino acid identity to the cell division cycle protein Cdc48.

Key Words: Archaea, Proteasome, AAA family, ATPases, Chaperone, Protease, Protein quality control, Review

Send correspondence to: Julie A. Maupin-Furlow, Department of Microbiology and Cell Science, University of Florida, Gainesville, Florida, 32611-0700, Tel: 352-392-4095, Fax: 352-392-5922, E-mail: jmaupin@ufl.edu