Slicing a protease: Structural features of the ATP-dependent Lon proteases gleaned from investigations of isolated domains

Tatyana V. Rotanova, Istvan Botos, Edward E. Melnikov, Fatima Rasulova, Alla Gustchina, Michael R. Maurizi and Alexander Wlodawer

Protein Sci. 2006 15: 1815-1828
Access the most recent version at doi:10.1110/ps.052069306

REVIEW

Slicing a protease: Structural features of the ATP-dependent Lon proteases gleaned from investigations of isolated domains

TATYANA V. ROTANOVA,[1] ISTVAN BOTOS,[2,3] EDWARD E. MELNIKOV,[1] FATIMA RASULOVA,[4,5] ALLA GUSTCHINA,[2] MICHAEL R. MAURIZI,[4] AND ALEXANDER WLODAWER[2]

[1]Shemyakin–Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, Moscow 117997, Russia
[2]Macromolecular Crystallography Laboratory, National Cancer Institute at Frederick, Frederick, Maryland 21702, USA
[3]Laboratory of Molecular Biology, National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda, Maryland 20892, USA
[4]Laboratory of Cell Biology, National Cancer Institute, Bethesda, Maryland 20892, USA
[5]Enteric Diseases Department, Infectious Diseases Directorate, Naval Medical Research Center, Silver Spring, Maryland 20910, USA

Abstract

ATP-dependent Lon proteases are multi-domain enzymes found in all living organisms. All Lon proteases contain an ATPase domain belonging to the AAA$^+$ superfamily of molecular machines and a proteolytic domain with a serine-lysine catalytic dyad. Lon proteases can be divided into two subfamilies, LonA and LonB, exemplified by the *Escherichia coli* and *Archaeoglobus fulgidus* paralogs, respectively. The LonA subfamily is defined by the presence of a large N-terminal domain, whereas the LonB subfamily has no such domain, but has a membrane-spanning domain that anchors the protein to the cytoplasmic side of the membrane. The two subfamilies also differ in their consensus sequences. Recent crystal structures for several individual domains and sub-fragments of Lon proteases have begun to illuminate similarities and differences in structure–function relationships between the two subfamilies. Differences in orientation of the active site residues in several isolated Lon protease domains point to possible roles for the AAA$^+$ domains and/or substrates in positioning the catalytic residues within the active site. Structures of the proteolytic domains have also indicated a possible hexameric arrangement of subunits in the native state of bacterial Lon proteases. The structure of a large segment of the N-terminal domain has revealed a folding motif present in other protein families of unknown function and should lead to new insights regarding ways in which Lon interacts with substrates or other cellular factors. These first glimpses of the structure of Lon are heralding an exciting new era of research on this ancient family of proteases.

Keywords: AAA$^+$ protein; Lon protease; ATP-dependent protease; Ser-Lys dyad; LonA and LonB subfamilies; domains; crystal structure

Reprint requests to: Alexander Wlodawer, Macromolecular Crystallography Laboratory, NCI at Frederick, P.O. Box B, Frederick, MD 21702, USA; e-mail: wlodawer@ncifcrf.gov; fax: (301) 846-6322; or Tatyana V. Rotanova, Shemyakin–Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, Miklukho-Maklaya st. 16/10, GSP-7, Moscow 117997, Russia; e-mail: rotanova@enzyme.siobc.ras.ru; fax: +7-495-3357103; or Michael R. Maurizi, Laboratory of Cell Biology, National Cancer Institute, Bethesda, MD 20892, USA; e-mail: mmaurizi@helix.nih.gov; fax: (301) 480-2284.

Article and publication are at http://www.proteinscience.org/cgi/doi/10.1110/ps.052069306.

The Lon family of peptidases

The Lon protease family (MEROPS [Rawlings et al. 2004] clan SJ, family S16), which is conserved in the prokaryotes and in eukaryotic organelles such as mitochondria and peroxisomes, is the most widespread family of ATP-dependent proteases. Prokaryotic Lons are key enzymes responsible for intracellular proteolysis, contributing to protein quality and cellular homeostasis by

eliminating mutant and abnormal proteins and participating in rapid turnover of select short-lived regulatory proteins (Goldberg 1992; Gottesman and Maurizi 1992; Gottesman 1996; Wickner et al. 1999). Though less well studied, eukaryotic Lons have been shown to exhibit similarly important regulatory and protein quality control functions in mitochondria (Van Dyck et al. 1994; van Dijl et al. 1998; Bota and Davies 2002). Lon, like all other ATP-dependent proteases (FtsH, ClpAP, ClpXP, HslVU, and the 26 S proteasome), belongs to the AAA^+ protein superfamily (ATPases associated with diverse cellular activities) (Neuwald et al. 1999; Maurizi and Li 2001; Ogura and Wilkinson 2001; Lupas and Martin 2002; Iyer et al. 2004). In addition to protein unfolding and proteolysis, AAA^+ proteins are involved in many cellular processes, including membrane fusion, protein and organelle translocation, DNA and RNA unwinding, assembly and disassembly of multi-protein complexes, and microtubule severing, to name the best known. The cellular activity of AAA^+ proteins is largely defined by the functional partners with which they associate. In the case of Lon, the AAA^+ domain is covalently fused to a protease domain.

Lon proteases function as oligomeric assemblies, which had been variously reported to consist of four to eight identical subunits (Goldberg et al. 1994; Roudiak et al. 1998; Lee et al. 2004). A number of direct and indirect observations indicate that bacterial Lons form rings consisting of six subunits (Lee et al. 2004; Park et al. 2006). The yeast mitochondrial Lon has been shown to form seven-membered rings (Stahlberg et al. 1999), and similar arrangements might also be present in proteins from other sources. The subunits themselves are composed of three or more independently folded domains (see below).

Lons are divided into two subfamilies, LonA and LonB, based on differences in the number of domains and characteristic sequences within the domains (Rotanova et al. 2003, 2004). Both subfamilies contain the ATPase (A) domains that include typical AAA^+ modules, as well as the proteolytic (P) domains, but whereas LonA enzymes contain a large N-terminal (N) domain, the LonB enzymes have no N-domain but have a large transmembrane domain insertion within the AAA^+ module between the Walker motifs A and B (Fig. 1). It had been suggested that Lons are serine proteases (Chung and Goldberg 1981; Waxman and Goldberg 1982; Goldberg et al. 1994), although the amino acid sequences of LonA and LonB P-domains show no homology with serine proteases containing the classical catalytic triad Ser-His-Asp (Amerik et al. 1988, 1990, 1991; Goldberg et al. 1994; Rotanova et al. 2004), or, indeed, to any other proteases. Active-site-directed inhibitors of classical serine proteases, such as sulfonyl fluorides, chloromethyl ketones, or fluorophosphates, are poor inhibitors of Lon, and none has been shown to modify the active site residues. Comparative analysis of the primary structures of the Lon pool and site-directed mutagenesis of the full-length *Escherichia coli* Lon (*Ec*Lon) (Rotanova et al. 2003, 2004), along with the recently determined structure of the *Ec*Lon P-domain (Botos et al. 2004b), established that the active sites of Lon proteases have a Ser-Lys catalytic dyad (Ser679 and Lys722 in the *Ec*Lon numbering).

The domains in *Ec*Lon, a representative member of the LonA subfamily, are assigned by us for the purpose of this review as 1–309 (N-domain), 310–584 (A-domain), and 585–784 (P-domain), although there is no consensus on such a division at this time and the exact position of domain boundaries, especially between the N- and

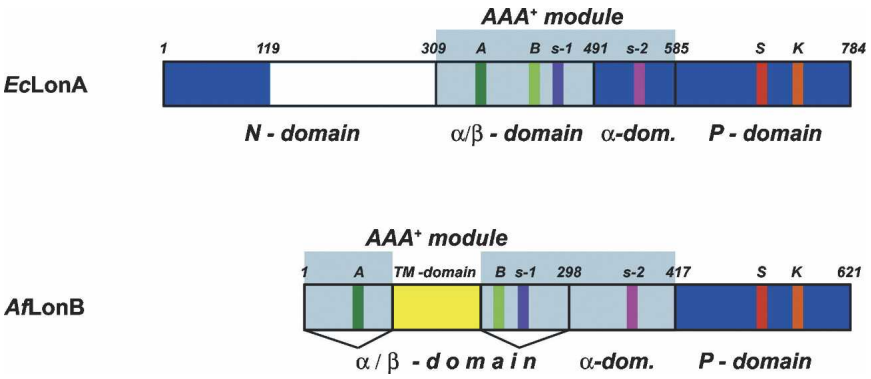

Figure 1. Schematic diagram of the typical members of LonA and LonB subfamilies, represented by *E. coli* and *A. fulgidus* Lon proteases, respectively. Domain definitions are explained in the text. (Dark blue) The segments for which three-dimensional structures have been solved; (green) the locations of the Walker A and B motifs (AAA^+ module), marked A and B; (red and orange) residues forming a catalytic dyad (serine [S] and lysine [K]); (yellow) the transmembrane domain of the LonB proteases; (magenta) the sensor-1 and sensor-2 regions. The positions of putative intein insertions located just after the TM domains in some LonB proteases (but not in *Af*LonB) are not shown.

A-domains, may be subject to future revision. AAA^+ modules generally contribute to target selection and regulation of the activity of the associated functional component (the protease domain in the case of Lon) (Neuwald et al. 1999; Wickner et al. 1999; Ogura and Wilkinson 2001; Lupas and Martin 2002; Iyer et al. 2004). Basic AAA^+ modules consist of two domains: a larger nucleotide-binding domain (α/β-domain) and a smaller helical domain (α-domain). The α/β-domain contains conserved motifs, including Walker A and B and sensor-1, which take part in nucleotide binding and hydrolysis, and an ''Arg finger'' that contributes to activation of ATP hydrolysis upon subunit interaction (Neuwald et al. 1999). The α-domain contains at least one conserved motif, sensor-2, containing an Arg (or rarely Lys) residue, also involved in ATP hydrolysis and in protein substrate remodeling (Neuwald et al. 1999; Iyer et al. 2004).

The binding and hydrolysis of ATP promotes a cycle of conformational changes within the AAA^+ protein, and relative movement between the α/β- and α-domains during the nucleotide binding and hydrolysis cycle generates a mechanical force that acts on the functional partners and associated substrates (Ogura and Wilkinson 2001; Iyer et al. 2004). Thus, many AAA^+ proteins function as chemo-mechanical enzymes to alter the conformation of proteins, disassemble macromolecular complexes, or translocate macromolecules between different compartments within the cell (Baumeister et al. 1998; Langer 2000; Lupas and Martin 2002; Gottesman 2003; Sauer et al. 2004; Wang 2004; Groll et al. 2005; Martin et al. 2005).

The ATPase components of several ATP-dependent proteases have been shown to have protein unfolding activity or at least the ability to locally disrupt the structure of a target protein (Weber-Ban et al. 1999; Kim et al. 2000; Singh et al. 2000). For a majority of ATP-dependent proteases, structural destabilization is a prerequisite for allowing substrate proteins access to the proteolytic active sites, which are sequestered in internal chambers within the holoenzyme complex in such a way that bound substrates must pass through a narrow axial channel to reach them (Hegerl et al. 1991; Lowe et al. 1995; Bochtler et al. 1997; Wang et al. 1997; Groll et al. 2000). Once protein substrates enter the proteolytic sites, peptide bond cleavage occurs rapidly and promiscuously and without any further expenditure of energy, giving rise to a variety of peptide products ranging in size from three to ~15 residues (Kisselev et al. 1999; Choi and Licht 2005). The translocation process itself is also energetically unfavorable, and ATP hydrolysis is needed to drive this process (Hoskins et al. 1998; Singh et al. 2000; Burton et al. 2001; Reid et al. 2001; Kenniston et al. 2003). The distribution of the energy requirement between unfolding and translocation is likely to vary with different substrates, reflecting size, global and local thermodynamic stability, and interactions of unfolded substrates with various binding sites along the translocation pathway in the AAA^+ module and the protease component.

Understanding the mechanism and regulation of the activities of these complex enzymes will require extensive structural information on the functional domains and the interface and interactions between them. Structural details will vary for different families because of variations of the mode of the initial binding interactions between substrates and ATP-dependent proteases, the need for effective force generation without disruption of the enzyme itself, the complexity of translocating a protein through the narrow axial substrate channels, and allosteric effects leading to activation of the proteolytic site or conformational changes needed to release reaction products. So far, significant progress in structure determination has been made with ATP-dependent proteases that are assembled from independently expressed AAA^+ proteins and protease components, such as the Clp proteases (Wang et al. 1997; Guo et al. 2002; Kim and Kim 2003) and proteasome-related proteases (Lowe et al. 1995; Bochtler et al. 1997, 2000; Groll et al. 1997; Sousa et al. 2000; Sousa and McKay 2001; Wang et al. 2001). These structures provide key insights into functionally important features likely to be shared by all such enzymes. The ATP-dependent proteases, Lon and FtsH, in which the AAA^+ protein and the protease are fused in a single polypeptide together with one or more other functional domains, present a unique challenge. Although their global architectural features, such as the fold of the AAA module, are expected to be conserved, as has been already shown for FtsH (Krzywda et al. 2002; Niwa et al. 2002), the single-polypeptide ATP-dependent proteases might differ in important mechanistic and structural details related to changes in the interface between the A- and P-domains and allosteric communication between them. In fact, the structural analysis of Lons has already provided a few surprises and has led to a reconsideration of how this class of ATP-dependent proteases works. For example, the structure of the isolated protease domain (Botos et al. 2004b) does not show the kind of secluded degradation chamber observed in the proteasomes and Clp proteases, suggesting that protection of cellular proteins from degradation might rely on allosteric regulation of the catalytic activity of the protease rather than blocking access to these sites by compartmentalizing them.

Crystallization of intact Lon proteases has been attempted for many years without success. Fortunately, limited proteolysis identified a number of stable domains and their combinations (Ovchinnikova et al. 1998; Roudiak and Shrader 1998; Vasilyeva et al. 2002; Patterson et al.

2004; Rotanova et al. 2004), which has allowed successful purification and crystallization of several Lon fragments, as well as their cloned and expressed equivalents. In some cases, both approaches were used in the study of the same domain, providing complementary information (Botos et al. 2004b, 2005). At this time, detailed structural information is available for a subdomain of the N-domain and for the α-domain of *Ec*LonA, as well as for the P-domains of wild-type and mutated forms of *Ec*LonA, *Archaeoglobus fulgidus* LonB (*Af*LonB), and *Methanococcus jannaschii* LonB (*Mj*LonB). The available structures provide an incomplete picture of Lon molecules and their oligomeric complexes, but they have been valuable in defining structural constraints on the domain organization and assembly of the native protein and in providing information explaining the mode of catalytic activity.

The N-terminal domain of LonA

The N-domain is found only in LonA subfamily members. Intriguingly, the Lon N-domain is related to a widespread and diverse family of proteins of unknown biological function that are present as 200–300 amino acid–long open reading frames in the genomes of a variety of organisms (Li et al. 2005). The Lon N-domain is composed of two or more smaller domains identified by a combination of multiple sequence alignment and limited proteolysis. Limited proteolysis of *Ec*LonA produced several transiently stable N-terminal fragments terminating between Lys223 and Lys239 (cleavage with trypsin or lysyl endopeptidase C), after Glu240 (Staphylococcal V8 protease), and between Tyr228 and Met234 and after Trp303 (chymotrypsin) (Ovchinnikova et al. 1998; Vasilyeva et al. 2002; Patterson et al. 2004; T.V. Rotanova, I. Botos, E.E. Melnikov, G.G. Leffers, F. Rasulova, A. Gustchina, M.R. Maurizi, and A. Wlodawer, unpubl.). Upon extended incubation with chymotrypsin, the N-terminal domain was reduced to one very stable fragment, Lon-N209, which exists as a monomer in solution and is structurally homogeneous, judging from its ability to yield large crystals. The occurrence of insertions within divergent Lon sequences suggested another possible boundary in the region near *Ec*LonA residue 119; a construct of *Ec*Lon-N119 was expressed and purified, and its structure solved (Li et al. 2005). *Ec*Lon-N119 has a novel fold made up of three twisted β-sheets folded into a shallow U shape with a single α-helix nestled in the depression (Fig. 2A). Most of the surface, including the region around the α-helix, is hydrophilic, with the exception of a broad swath of exposed hydrophobic residues cutting across the β-sheets on the surface opposite the α-helix.

A structure similar to that of Lon-N119 was also observed for the β-rich N-terminal domain of a small protein, BPP1347, purified from *Bordetella parapertussis*.

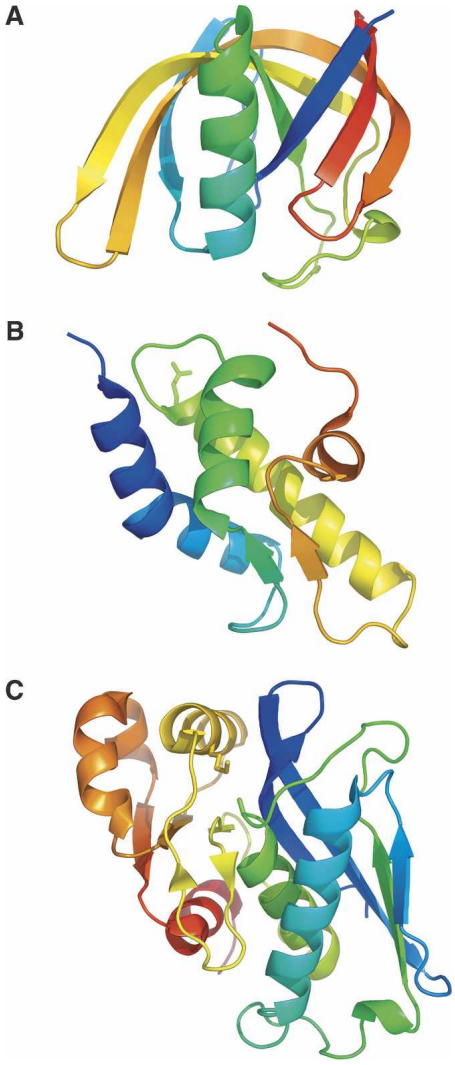

Figure 2. Crystal structures of the fragments of *E. coli* Lon. (*A*) Structure of the N-terminal subdomain of the N-domain (designated *Ec*Lon-N119, PDB code 2ane), shown as a ribbon diagram in rainbow colors (changing from blue at the N terminus to red at the C terminus). (*B*) Structure of the α-domain from the AAA^+ module, shown as a ribbon diagram in rainbow colors (PDB code 1qzm). The side chain of sensor-2 Arg542 is marked in stick representation. (*C*) Structure of the P-domain, shown as a ribbon diagram in rainbow colors (PDB code 1rr9). The side chains of the catalytic dyad are marked in stick representation.

BPP1347, whose function and level of expression (if any) are not known, is the product of a 202-amino-acid open reading frame randomly isolated and crystallized as part of a structural genomics effort (PDB code 1ZBO). The BPP1347 molecule has two domains connected in tandem by a single linker polypeptide segment. Its N-terminal domain has 19% sequence identity to Lon-N119, and, remarkably, a similar degree of identity is also seen throughout the C-terminal part of BPP1347 and residues 120–209 of *Ec*LonA. The C-terminal domain of BPP1347

contains four prominent α-helices with unique topology, and this structure correlates very well with the secondary structure predicted for residues 120–209 of Lon (*Ec*LonA numbering) based on the sequences of >100 LonA proteins, including the number and length of the helices and the positions of turns and loops. We postulated that the ~210 amino acid–long N-terminal fragments of *Ec*LonA and other prokaryotic A-type Lon proteins have the same overall structure as BBP1347 (Li et al. 2005). In BPP1347, there is a large interface between the two subdomains formed mostly by hydrophobic and hydrogen bonding interactions. The structure of Lon-N209 modeled on BPP1347 also shows an extensive interface involving >20 side chain interactions, suggesting that the two subdomains of Lon N-domains are fixed relative to each other and function as a single bimodal structural unit.

The structural similarity between BPP1347 and fragments of the N-domain of Lon suggests that the two might have some activity or property in common. It is not known if, and under what conditions, BBP1347 or other hypothetical proteins consisting only of this conserved domain and encoded in the genomes of many other organisms are expressed in vivo. The N-domain of *Ec*LonA appears to have some protein-binding ability and might contribute to substrate recognition. *E. coli* Lon lacking 107 N-terminal residues had drastically reduced protein-degrading activity in vitro (Rasulova et al. 1998a). *Mycobacterium smegmatis* Lon lacking 90, 225, or 277 N-terminal residues lost practically all proteolytic activity while exhibiting reduced protein binding activity, as well as alteration of the oligomeric state (Roudiak and Shrader 1998). It is possible, however, that such deletions can cause structural perturbations that affect activity in other parts of the protein and influence its oligomerization (Lee et al. 2004). A more specific indication of substrate interaction by the N-domain is the identification of an *E. coli* Lon mutant altered in substrate specificity (Ebel et al. 1999). A mutation in Lon that converts Glu240 to Lys results in stabilization of one Lon substrate, RcsA, in vivo but does not affect the degradation of another substrate, SulA. Whether Lon has specific protein recognition sites in the N-domain remains to be proven. It is possible that the N-domain contributes to specific substrate interaction by binding disordered regions in substrates that also carry specific motifs recognized elsewhere on Lon. There is evidence of such an auxiliary role for the N-domain of ClpA, the chaperone component of ATP-dependent protease, ClpAP (Xia et al. 2004). We speculate that the more distantly related domains and stand-alone proteins that are structurally similar to the Lon N-domain might also bind unfolded proteins or disordered polypeptides. Such activity should provide some clue as to their still unknown functions.

The α-domain of the AAA$^+$ module

The AAA$^+$ module of Lon did not yield crystals suitable for structure determination; however, digestion of *Ec*LonA by α-chymotrypsin yielded a stable fragment consisting of residues 491–584, which was purified and crystallized (Botos et al. 2004a). This largely α-helical fragment represents the complete small α-domain of the AAA$^+$ module and displays a conserved topology (helix-strand-helix-helix-strand-helix) reported for many similar α-domains (Lupas and Martin 2002) (Fig. 2B). Helix 1 is slightly bent in the middle. The significance of the bend is unclear at present, although it alters the space and the angle between the α-domain and the larger α/β-domain. AAA$^+$ proteins differ depending on whether helix 1 is straight, is bent in the middle, or contains a two-amino-acid bulge in the middle (Ogura and Wilkinson 2001). Following β-strand 1 and helix 2 is a long helix 3, at the beginning of which is located the well-conserved sensor-2 residue, Arg542. The following β-strand loops form a parallel β-sheet with the β-strand 1. The C-terminal helix 4 is unwound at the end where it should be connected to the beginning of the P-domain.

Current models of the function of AAA$^+$ modules of typical AAA$^+$ proteins suggest that the α-domain acts as a rigid body connected to the α/β-domain by a loop that is sensitive to the nucleotide state of the module (Rouiller et al. 2002; Wang 2004; DeLaBarre and Brunger 2005). Stability of subunit interactions in the assembled rings is conferred by the interaction of the α-domain of one subunit with the α/β-domain of an adjacent subunit (Lenzen et al. 1998; Bochtler et al. 2000; Zhang et al. 2000). Binding and release of a nucleotide cause a relative rotation and separation between the α-domain and the α/β-domain (Wang 2004). Depending on the state of assembly and the presence of a bound load, this relative movement can impose a force on the load, impose a force on a polypeptide extension connected to one of the domains, or cause partial separation of the subunits. In the fixed-ring configuration, movement between the α- and α/β-domains would force the α/β-domain to be displaced along the axis and exert a force on a substrate polypeptide within the axial channel. If motion of the α/β-domain is restricted, movement between the α- and α/β-domains will cause the α-domain to rotate out and up from the ring (Wang 2004), and, in the case of Lon, this effect will impinge on the secondary structural elements by which the P-domain is connected to the AAA$^+$ module. In this way, nucleotide binding or release can allosterically affect interactions between the structural elements of the P-domain or even the configuration of the catalytic residues at the active site. Taking into account the oligomeric structure of Lon and assuming the presence of multiple active subunits in the oligomer, it is likely that

these divergent effects occur on different subunits in an ordered or sequential manner, similarly to ATP synthase subunits that are sequentially active, depending on rotary interactions with the γ subunit (Stock et al. 2000).

The structure of the P-domain of Lon

In vitro, purified *Ec*Lon P-domain exhibits no detectable activity against protein substrates degraded by full-length Lon (Botos et al. 2004b), but retains a significant fraction of peptidase activity (Rasulova et al. 1998b). Interestingly, a construct containing residues 793–1133 of yeast Lon, which comprises the P-domain along with most of the α-domain, exhibited low but significant proteolytic activity in vivo (van Dijl et al. 1998). Moreover, activity of this yeast Lon αP-fragment was enhanced when it was coexpressed with a construct containing the N- and A-domains (residues 1–917), indicating that the P-domain is somewhat malleable and may require interactions with other parts of Lon to maintain a fully active configuration (van Dijl et al. 1998).

The crystal structure of the P-domain of *Ec*LonA (residues 585–784, Fig. 2C) (Botos et al. 2004b) elucidated a unique fold that is shared only with the subsequently determined P-domains of *Mj*LonB (residues 456–649) (Im et al. 2004) and *Af*LonB (residues 417–621) (Botos et al. 2005), as well as with viral protease VP4 (Feldman et al. 2006); this structure will be detailed below, with only the differences outlined for the other two structures (Fig. 3).

The first crystals of the P-domain of *Ec*Lon belonged to a $P3_1$ space group. The asymmetric unit contained six subunits, and the putative oligomeric molecule appeared as a ring with pseudo-sixfold molecular symmetry. The P-domain subunit itself has six α-helices and ten β-strands (Fig. 2C) and is composed of two compact subdomains, residues 585–697 and 698–784. The first nine residues of the P-domain (585–593) are disordered. The β-strand 1 and an antiparallel β-strand 2 form a long β-hairpin loop. This loop and parallel β-strands 3 and 4, which are separated by helix 1, form the first large β-sheet, which lies in a plane aligned with the sixfold axis. The distal surface of this sheet forms the interface with the adjacent subunits in the ring. A disulfide bridge between Cys617 and Cys691 connects the end of helix 2 to the end of β strand 2, stabilizing the subdomain. This unusual surface-exposed disulfide bond is unique only to Lon proteases from closely related enteric bacteria. At the base of the subdomain, strand 5 forms a second, small β-sheet producing a shallow concavity toward the center of the ring. Strand 5 is connected to helix 2 by a loop that contains the catalytic Ser679.

Following helix 2, a random coil forms a bridge to the second subdomain. A short β-strand 6 leads into another

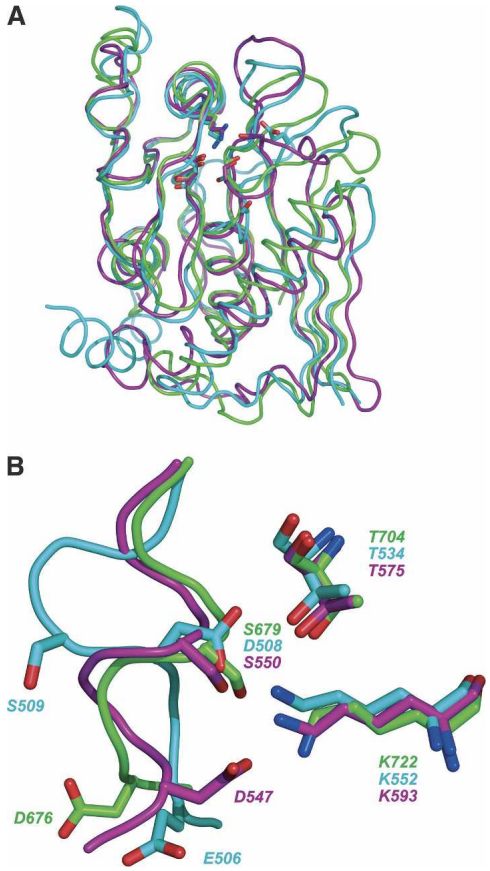

Figure 3. Comparison of the structures of the P-domains of *E. coli* LonA (green), *A. fulgidus* LonB (cyan), and *M. jannaschii* LonB (magenta). (*A*) Ribbon diagram of the superposition of the main chains, with the side chains in the active site marked in sticks. (*B*) Detailed superposition in the vicinity of the active site, with the trace of the chain that contains the catalytic serine shown as a ribbon. The position of the hydroxyl of Ser679 in EcLonA is modeled, whereas all other atoms are from experimental structures.

β-loop formed by antiparallel strands 7 and 8, followed by helix 3. Helix 3, which lies near the base of the subunit and runs nearly parallel to the edge of the hexamer, carries the second catalytic residue, Lys722. Strand 9 returns along helix 3, followed by a short 3_{10} helix 4, α-helix 5, and then parallel strand 10. Strands 6, 9, and 10 form a third small β-sheet, sandwiched by helix 3 and C-terminal helix 6. The last nine residues adopt an extended conformation in only one molecule, while they are disordered in the other ones.

Only a few deviations from the secondary structure of LonA have been reported for LonB. An isolated P-domain of *Af*LonB (residues 417–621, Fig. 3A) (Botos et al. 2005) contains an additional strand S0 on its N terminus, as well as a long C-terminal helix 7. An equivalent helix is not present in the otherwise related *Mj*LonB (residues 456–649, Fig. 3A) (Im et al. 2004). In the latter structure,

helix 2, immediately adjacent to the part of the active site that contains the catalytic Ser550, is elongated (Fig. 3B), leading to the modifications of the active site that will be further discussed below.

The active site of the P-domain

The first structure determined for any P-domain of Lon was of the inactive S679A mutant of *Ec*LonA (Botos et al. 2004b); thus, the mutual disposition of the side chains of the catalytic dyad residues Ser679 and Lys722 could only be modeled. In the experimentally determined structure, Lys722 makes a hydrogen bond to the carbonyl oxygen of Gly717 (Fig. 3B), which is strictly conserved in LonA and LonB proteins, in addition to also interacting with either a bound sulfate or the C terminus of another monomer. The position of the hydroxyl group of Ser679 was modeled by making minor adjustments to bring the Nζ atom of Lys722 to within ~3 Å of Oγ1 of Thr704 and O of Gly717. This model suggested an important role for Thr704, which belongs to a strictly conserved Tyr-Gly pair located 25 residues downstream of the catalytic serine (Fig. 3B). The resulting distance between Ser679 Oγ and Lys722 Nζ was predicted to be ~2.8 Å.

Important questions about the relevant disposition of the residues in the catalytic center of Lon proteases have been raised on the basis of the arrangement of the active site of the isolated P-domain of *Mj*LonB (Im et al. 2004). In this structure, the catalytic dyad formed by Ser550 and Lys593 (Fig. 3B) was observed (distances of 3.0 Å between Oγ of the former and Nζ of the latter in the two independent molecules in the asymmetric unit). In addition, a third residue, Asp547, was found within hydrogen bonding distance of Lys593 (distances 2.7 and 3.3 Å). The distances between Lys593 and Thr575 were ~3.4 Å. These observations led to a postulate that the details of the catalytic mechanism might be different between the two Lon subfamilies (Im et al. 2004).

Further departure from the predicted arrangement of the active site residues was observed in the atomic-resolution (1.2 Å) structure of *Af*LonB (Botos et al. 2005), which differed from both *Ec*LonA and *Mj*LonB. The side chain of Ser509 points out into the solvent and its Oγ group occupies two positions, one within 2.5 Å of a water molecule, and another 3.27 Å away from Oε1 of Glu472, a residue that is strictly conserved among all known LonB proteases other than *Mj*Lon (Fig. 3B). *Af*Lon Lys552 is very well ordered and its Nζ group makes three hydrogen bonds, being 2.77 Å away from the carbonyl oxygen of Gly547, 2.88 Å away from Oγ1 of Thr534, and 2.69 Å from Oδ1 of Asp508. The Oδ2 group of Asp508 makes another hydrogen bond to Oγ1 of Thr534 (2.64 Å) and accepts a bond from the main chain amide nitrogen of Gly535 (2.99 Å). Removal of the side chain of Asp508 in the D508A mutant does not change significantly the positions of either Lys552 (which is still well ordered and makes hydrogen bonds to Thr534 and Gly547) or Ser509, for which only a single orientation, hydrogen bonded to a water molecule, is observed.

Enzymatic mechanism at the proteolytic active site

Biochemical, mutational, and structural data have all established that the Ser-Lys dyad is responsible for the catalytic activity of proteases that belong to the Lon family. However, in the absence of structural data describing a substrate bound in the active site of Lon, it is necessary to examine the structures of other proteins that use such a dyad in order to model the mechanism of action of Lon. Related enzymes include *E. coli* type 1 signal peptidase (SPase) (MEROPS, clan SK; PDB codes 1kn9, 1b12) (Paetzel et al. 2002); the autoproteolytic enzymes LexA (1jhc, 1jhe) (Luo et al. 2001) and UmuD (1umu) (Peat et al. 1996); and λ cI protein (1f39) (the latter all belonging to the MEROPS clan SF) (Bell et al. 2000). A comparative analysis of consensus sequences surrounding the catalytic residues of these peptide hydrolases, as well as LonA and LonB proteases, revealed only remote similarity between Lons and LexA (Rotanova 2002). At the same time, structural comparisons reveal the overall similarity among the folds of these proteins in the vicinity of their catalytic centers (Fig. 4). Also, in those structures, the catalytic serine and lysine residues are thought to be in an ''active'' configuration, because they overlap very well with the catalytic residues of TEM1 β-lactamase, a related hydrolase with a Ser-Lys catalytic dyad. Although the P-domain of Lon has a completely different overall fold, the stretch of residues that includes strand 5 with the catalytic Ser679 (*Ec*Lon)

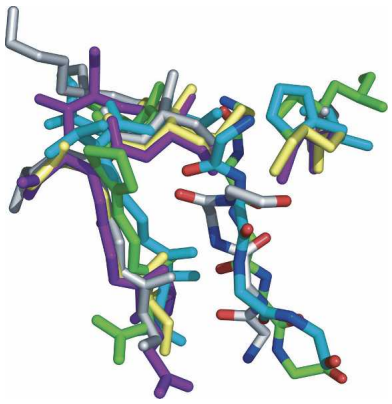

Figure 4. Active site superposition of Lon with other serine proteases: *E. coli* Lon protease (green; PDB code 1rr9); subtilisin complexed with eglin (cyan; PDB code 1cse); λ cI (magenta; PDB code 1f39); umuD (yellow; PDB code 1umu); LexA (gray; PDB code 1jhe).

and the beginning of helix 2 can be superimposed quite well onto the corresponding segments in the four structures listed above. This alignment also brings the catalytic lysines into excellent superposition.

Comparative analysis of the active sites of these enzymes identifies a third residue that might assist the Ser-Lys dyad during catalysis. This residue is either a serine or a threonine, with the side chain Oγ located within hydrogen-bonding distance of the catalytic lysine. The role of such a residue (Ser278 in SPase, Thr154 in LexA, Thr704 in *Ec*LonA, Thr534 in *Af*LonB, and Thr575 in *Mj*LonB) might be similar to that of the aspartate present in the classic catalytic triad of serine proteases (Dodson and Wlodawer 1998). However, it is noteworthy that mutant forms *Ec*LonT704A and *Af*LonT534A retain significant proteolytic activity. Thus, a full role of these threonine residues for the function of Lon proteases needs to be clarified in further studies.

Do the LonB subfamily enzymes have a different mechanism of action as postulated based on the structural studies on *Mj*LonB (Im et al. 2004)? In that structure, Asp547 formed a salt bridge with Lys593 (Fig. 3B), suggesting that it directly impinges on a catalytic residue and could be required for activity. However, Asp547 is not universally conserved in all B-type Lons; for example, Glu506 is present at the equivalent position in *Af*LonB, and mutation of the latter residue to alanine did not significantly influence the enzymatic activity (Botos et al. 2005). It is possible that the structure observed for *Mj*LonB represents an alternative state of LonB that can undergo a change to the active configuration, but further data are needed to address the question of the role of this aspartate residue. Given the significant degree to which the details of the active site differ substantially in the structures of the P-domains solved to date, it may be that none fully represents the active enzyme. In addition, the active site of the viral protein VP4 in which the active site serine was present (Feldman et al. 2006) is in excellent agreement with the structure of *Ec*LonA.

An additional structural feature, the ''oxyanion hole,'' is critical for enzymatic activity in serine proteases, because it helps to stabilize the formation of a tetrahedral intermediate during catalysis (Kraut 1977). In the LexA mutant 1jhe, the oxyanion hole was identified on the basis of the hydrogen-bonded interactions of the carbonyl oxygen of Ala84, which forms the scissile bond targeted during autoproteolysis (Luo et al. 2001). Two main chain amide nitrogens, one from the catalytic serine and the other from the preceding residue, form the oxyanion hole. In three other structures of LexA in which the loop containing the Ala84-Gly85 bond is not in a conformation allowing self-cleavage, a water molecule occupies the position of the Ala84 carbonyl. An equivalent water is found in the structures of the *Ec*LonA P-domain, the λ cI protein, and UmuD′; we note that the latter structure is lacking the polypeptide segment with the scissile bond. The water makes an additional hydrogen bond to the carbonyl oxygen of the residue at the third position N-terminal to the catalytic serine (Asp676 in the case of *Ec*LonA). Considering the similarity of the active site configurations of the P-domain and the other enzymes, it was possible that the amide nitrogens of Ser679 and the preceding residue, Pro678, together formed the oxyanion hole in Lon. However, as proline has not been shown to participate in the formation of oxyanion holes in other enzymes, its ability to serve this function in Lon is rather questionable. A second candidate would be the amide nitrogen of the nearby Trp603, which contributes to a highly conserved motif, GLAW/Y in LonAs or GLAV/I in LonBs, for which no function has yet been identified.

Oligomeric structure of Lon proteins

Isolated forms of both A and B type Lon proteases are oligomeric, but the stoichiometry of the subunits, and whether the stoichiometry is the same in all Lon species, have not been established. The stoichiometry was reported to be seven subunits in yeast mitochondrial Lon based on a number of independent criteria, including electron microscopic image analysis (Stahlberg et al. 1999). For *Ec*Lon, sedimentation data pointed to two states putatively composed of four and eight subunits (Goldberg et al. 1994), whereas for *M. smegmatis* Lon, an equilibrium between hexamer/tetramer/dimer or hexamer/trimer was reported (Rudyak et al. 2001). STEM analysis of intact *Ec*Lon and an N-domain-deleted form of Lon (residues 309–784) showed predominantly hexameric assemblies or, for intact Lon only, complexes of two associated hexamers (FSR, MRM; M. Kessel, R. Leapman, unpubl.). Very recently, negative stain electron microscopy of *Ec*Lon gave particles that produced averaged images with sixfold symmetry, also suggesting that *Ec*Lon forms rings with six subunits (Park et al. 2006). *Bacillus thermoruber* Lon was reported to be a hexamer based on a combination of gel filtration and sedimentation velocity experiments and on cross-linking of intact and truncated species (Lee et al. 2004). Finally, it was suggested that the oligomeric state of *Ec*Lon could be modulated by reaction conditions in vitro (Vineyard et al. 2005), as well as by various cellular events in vivo (Nishii et al. 2005).

The crystal structures of the *Ec*LonA and *Af*LonB P-domains have provided data favoring hexameric ring structures for these enzymes. In the *Ec*LonA P-domain crystals (Botos et al. 2004b), as well as one of the crystal forms of *Af*LonB P-domain (Botos et al. 2005), the asymmetric unit contained six molecules related by an

approximate sixfold non-crystallographic symmetry (NCS) axis, revealing a hexamer in which individual monomers form a ring (Fig. 5). Viewed from the side, the ring is dome-shaped, with a diameter of ~100 Å at the base and ~50 Å at the top. A solvent-accessible central pore ~32 Å long runs through the hexamer. It has a diameter of ~18 Å at the entrance from the proximal side and widens slightly toward the distal end. The pore entrance from the proximal surface (16–18 Å) is significantly larger in size than the entry channel in *E. coli* ClpP (~10 Å) (Wang et al. 1997) and could be expected to accommodate two folded α-helices or β-strands or unstructured loops with two polypeptide chains. A number of negatively charged and polar residues are located toward the distal end of the pore, making the distal end rather negatively charged, while the proximal half of the pore has several positively charged residues, which might serve a gating function for substrate entry.

The presence of hexameric symmetry in several crystal forms of *Ec*LonA P-domains suggests that the hexameric ring assembly is not an artifact of crystallization but rather a biologically significant unit. Interactions between isolated P-domains are weak; *Ec*LonA P-domains are monomeric in solution at concentrations as high as 100 μM subunit equivalents. No higher order oligomers were reported in the structure of *Mj*LonB P-domain (Im et al. 2004), in which the two molecules in the asymmetric unit formed a loose isologously bonded dimer. In addition, a more complicated picture was observed for the *Af*LonB P-domain (Botos et al. 2005), which is present in solution as a monomer, but its oligomeric state in crystals differs among the four reported forms. In two of them, orthorhombic and monoclinic, the asymmetric units contain hexamers very similar to the one described above for *Ec*LonA. Charge distribution on the surface of the hexamer of *Af*LonB P-domain has very similar characteristics to the distribution described for *Ec*LonA. However, two related hexagonal crystal forms found for the wild-type enzyme and for the D508A mutant have only a single molecule in the asymmetric unit, with the molecules following a helical packing pattern along the 6_5 axis. Translation of the molecules in steps of ~6.5 Å along the hexagonal axis yields a hexameric ring that superimposes perfectly on the hexameric rings present in the orthorhombic and monoclinic crystal forms. These observations suggest that the hexameric ring is the favored configuration at high protein concentrations, with the caveat that the contacts favored under crystallization conditions may differ slightly from those in the intact oligomer under more physiological conditions. Since, in most AAA^+ proteins, the major contacts stabilizing the oligomeric states are made through the AAA^+ modules themselves, it is likely that P-domain interactions are strongest when they are brought into close proximity by assembly of the AAA^+ domains. Quite possibly, contacts between P-domains may be somewhat variable and responsive to the conformational state of the AAA^+ domain.

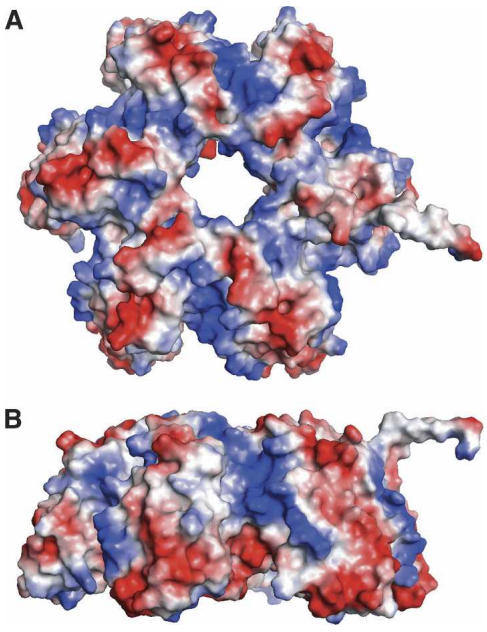

Figure 5. Oligomeric structure of the P-domain of LonA. Top (*A*) and side (*B*) views of the *Ec*LonA hexamers as seen in the crystals are colored according to charge distribution: (blue) positive areas, (red) negative areas.

P-domain interactions may also influence the conformation and activity of the A-domain, which could explain the properties of some previously isolated mutants. In earlier attempts to identify catalytic Lon residues, the very highly conserved His665 and His667 were mutated, and both mutant enzymes lack protein-degrading activity (Starkova et al. 1998). These mutants also caused a 90% reduction in ATPase activity, suggesting that they perturbed communication between the P- and A-domains or indirectly altered the structure of the A-domain. In the P-domain crystal, His665 and His667 both lie at the oligomeric interface. Their mutation should alter the subunit interactions, which in turn could perturb the interactions between the A-domains or could alter the timing of conformational changes in the A-domain. One of the earliest described mutants of *E. coli* Lon, called CapR9 because it caused a change in production of the colanic acid capsular polysaccharide (Hua and Markovitz 1972), was shown to have a single point mutation in Glu614 (Oh et al. 1998). Lon-E614K is a dominant-negative mutant, because it can form mixed oligomers with wild-type Lon and interfere with its activity. In the P-domain crystal, the side chain of Glu614 lies at the interface between subunits and makes a salt bridge with Arg710 from the adjacent subunit. Glu614 is present in virtually all bacterial Lon proteases, and Arg710 is highly conserved or is replaced by other

positively charged or hydrophilic residues. In the crystal, Glu614 also makes a hydrogen bond with His667 of the same subunit and thus affects subunit bonding by interacting with this residue as well. The region around Glu614, His667, and His665 contains five structural water molecules, indicating that perturbation of any of these residues could perturb a significant part of the subunit interface. The correlation between the mutational data and the model placing these residues at the subunit interface suggests that the model of the P-domain in the crystal is quite likely to be biologically significant.

Functional insights from models of full-length Lon

The recently acquired structural data can now be combined with the structures of other ATP-dependent proteases to assemble a working model of how Lon might look and work. Most AAA^+ proteins form hexameric rings in which the small α-domain interacts with the α/β-domain of the adjacent subunit. These assemblies are stabilized by ATP binding, and in many cases ATP hydrolysis requires the assembled hexamer state. The nucleotide binds in a crevice formed by the α- and α/β-domains from one subunit, but at least one residue, usually an arginine (referred to as an ''Arg finger''), from the adjacent subunit protrudes into the nucleotide binding site and can activate nucleotide hydrolysis. In the case of assemblies in which the functional partner is expressed independently, interaction between the AAA^+ hexamer and the functional partner is mediated by an extension from the AAA^+ domain that fits into a specific docking groove on the functional component. In the HslUV complex, the C-terminal heptapeptide of HslU penetrates a deep depression in the surface of HslV, anchoring the two hexamers together (Seong et al. 2002). In isolated HslU, the C-terminal peptide is folded back and docks into a hydrophobic groove on HslU itself (Bochtler et al. 2000), suggesting that a nucleotide-regulated conformational change is required to flip out the docking peptide to allow it to interact with HslV. In Clp proteases, a loop with a highly conserved motif at its apex protrudes from the AAA^+ module and docks into a groove on the surface of ClpP. Despite a symmetry mismatch between hexameric ClpA or ClpX and heptameric ClpP, which would prevent all the ''ClpP loops'' from being utilized simultaneously, the nucleotide-promoted interaction between the functional partners is quite tight (Kd < 10 nM). One proposal regarding the mismatch between the components in Clp proteases is that it allows flexibility and relative movement between the components during the catalytic cycles (Beuron et al. 1998).

Lon, in which the protease domain is a polypeptide extension joined to the AAA^+ module, presents yet another arrangement, but it might resemble more closely the interactions seen in the two-component HslUV. As in HslUV, the interaction between the AAA^+ module and the protease in Lon is symmetrical and is mediated through a linear polypeptide linkage, although the covalent linkage between the domains in Lon perhaps allows more direct mechanical transduction of conformational changes in the two domains in the subunit. Moreover, intersubunit domain–domain interactions between ATPase and proteolytic sites were revealed in *Ec*Lon by complementation of two mutant forms, LonK362Q and LonS679A (Melnikov et al. 2001; Tsirulnikov et al. 2003). When the α-domain structure of Lon is modeled onto the hexamer of HslU, the C-terminal end of the AAA^+ module is predicted to lie on the ring surface proximal to the P-domain. The C terminus of the α-domain is linked to the P-domain by a short length of polypeptide, which defines the orientation of the protease domain with respect to the AAA^+ module. In the proposed model of the combined Lon A- and P-domains (Lon-AP), we assume that the P-domain structure is correct with respect to the position of its N-terminal peptide. This assumption is justified because that portion of the structure is well determined in all the models constructed so far and because the peptide is stabilized by numerous interactions with the folded domain. These restraints place the P-domain hexamer with the domed surface abutting the AAA^+ module and the concave surface of the P-domain distal to the AAA^+ module. Assuming Lon domains are disposed as in other AAA^+ proteins, the N-domain of Lon would be positioned on or near the opposite face of the AAA^+ module, where they would assist in the recruitment of substrates to binding sites on the proximal ring surface or within the axial channel. Bound substrates would then be translocated through the AAA^+ module to the protease.

The model has a few expected, as well as some unexpected, features. The axial pore described above would align with the exit pore from the AAA^+ hexamer, allowing an unfolded polypeptide or unstructured loop to pass directly from the AAA^+ module through the pore, which would then bring it into contact with the proteolytic active sites. One surprise is that the active sites are not sequestered in this model, but lie in a solvent-exposed shallow crevice on the distal surface of the protease ring. In order to assure processive protein degradation with this arrangement, movement of the substrate protein from the AAA^+ module through the protease would need to be rate-limiting and timed to allow a sufficient number of cleavage events to generate the small peptides that are the usual product of degradation by Lon.

This configuration also raises a puzzling question regarding how proteins in the surrounding milieu avoid being degraded, since the Lon active sites are directly accessible from the distal surface. One possibility is that, although the active sites are accessible, the catalytic

residues are not configured properly for proteolysis without some signal from the AAA^+ module or from a protein within the axial channel. Such a condition exists in the HslUV system, in which it was shown that the isolated C-terminal decapeptide of HslU can allosterically activate peptide cleavage by HslV (Seong et al. 2002), implying that the proteolytic activity of HslV is dependent on HslU not only for delivery of substrates but also for regulating that catalytic activity of the active site. Making the functionality of the active sites dependent on allosteric effects produced by substrates bound on the AAA^+ module or within the channel would provide protection from unwanted cleavage of non-substrate proteins. The properties of the mutant, Lon-D676N, can be interpreted in terms of allosteric communication between the axial channel of the P-domain and the ATPase active sites of Lon. Lon-D676N is completely inactive for protein degradation; it retains some basal ATPase activity, but no activation of ATPase activity occurs upon binding of protein substrates. Asp676 lies at the distal end of the axial channel and stabilizes an α/β-turn made by helix 1 and strand 3, which together make up the walls of the axial channel. We propose that, in Lon-D676N, destabilization of the channel blocks substrate passage, which in turn prevents activation of the ATPase activity, which is needed to couple movement of the substrate through the channel with additional rounds of ATP hydrolysis.

Recent data obtained with mitochondrial Lon suggest that some substrates bind Lon and can be acted on by Lon while they are still folded (Ondrovicova et al. 2005). The proteins, the α-subunit of the mitochondrial processing protease and steroidogenic acute regulatory protein, are degraded by Lon in vitro under conditions in which they are resistant to trypsin digestion and folded in functionally competent states. The initial peptide bonds cleaved are at sites well away from the ends of the polypeptide, and are present in the three-dimensional structure in regions exposed on the surface in the folded proteins. Subsequent degradative steps occur processively from those sites. These data clearly indicate that polypeptide loops can gain access to the proteolytic sites of Lon and that Lon does not need to unravel a protein from the N or C termini. In fact, studies with fusion protein substrates suggest that yeast Lon has a relatively poor ability to unravel proteins and is only able to degrade proteins that have unstable tertiary structure (von Janowsky et al. 2005).

How do these data affect models of ATP-dependent degradation by Lon? It was suggested by Ondrovicova et al. (2005) that Lon rings might open to allow the folded protein into the interior of the AAA^+ module and into the proteolytic sites. Studies of yeast Lon by electron microscopy indicate large conformational differences between unliganded Lon and Lon with ATP bound, which could reflect a loosening or opening of contacts between subunits in a ring (Stahlberg et al. 1999). Nicking of the protein would destabilize its tertiary structure, allowing the substrate to be encapsulated when the rings come back together, leading to processive unfolding and translocation into the proteolytic sites. An alternative possibility is that Lon can unfold a protein from the middle and translocate a polypeptide loop or multiple polypeptide segments. The size of the axial channel is not necessarily fixed. Gating of the axial channel has been confirmed for the proteasome, in which the channels can undergo a major rearrangement in response to binding of an activator protein (Groll et al. 2000; Whitby et al. 2000), and recent data on ClpP indicate that the N-terminal peptide can adopt different configurations that alter the size and character of the axial channel (Kang et al. 2004; Bewley et al. 2006). A question raised by the openness of the proteolytic sites in the P-domain crystals is whether folded substrates can directly access the proteolytic sites. The initial cleavage could result in structural destabilization, allowing the A-domain to engage the protein, further unfold it, and translocate it to the proteolytic domain to complete the degradation. It is not clear in this mechanism how release of the nicked protein would be avoided to ensure processive degradation, but the proximity to the A-domain or N-domains might be sufficient to permit efficient retention of the nicked protein. Further insight into Lon's mechanism of action must await additional structural and functional studies, but it is likely that the diversity of ATP-dependent proteases reflects the wide variation in substrates targeted for degradation and that the mechanisms by which they operate are optimized for the types of substrates they select and the cellular conditions under which they must function. The ability of Lon active sites to switch between latent and active configurations might even provide a means by which metabolic effectors could activate more promiscuous proteolytic activity under exigent circumstances. Recently, ligand-induced stimulation of indiscriminate proteolytic activity was proposed for ClpP (Brotz-Oesterhelt et al. 2005), and it is possible that such activity is expressed by Lon and other proteolytic components of ATP-dependent proteases under stress conditions or when damaged or unfolded proteins accumulate.

Conclusions

Although Lons were the first ATP-dependent proteases to be studied, our knowledge of their structural and biochemical properties has lagged behind those of other similarly regulated proteolytic enzymes. In the absence of a crystal structure of a full-length enzyme, some important questions must remain unanswered. Nevertheless, the

recently solved structures of individual Lon domains, coupled with biochemical, biophysical, and electron microscopy data, are providing our first insights into how the structural features of Lon relate to its functions and activities. Already there are intriguing hints of differences in mechanisms of action and modes of regulation between Lon and other ATP-dependent proteases. Given the important biological roles played by Lon proteases, it is our hope that the still fragmentary information presented in this review will provide a basis for more detailed studies on the structure and function of this ubiquitous family of enzymes.

Acknowledgments

This work was supported in part by a grant from the Russian Foundation for Basic Research (project no. 05-04-48383) to T.V.R, by the U.S. Civilian Research and Development Foundation grant to T.V.R. and A.W. (project no. RB1-2505-MO-03), and by the Intramural Research Program of the NIH, National Cancer Institute, Center for Cancer Research.

References

Amerik, A.Y., Chistyakova, L.G., Ostroumova, N.I., Gurevich, A.I., and Antonov, V.K. 1988. Cloning, expression and structure of the functionally active shortened *lon* gene in *Escherichia coli*. *Bioorg. Khim.* **14:** 408–411.

Amerik, A.Y., Antonov, V.K., Ostroumova, N.I., Rotanova, T.V., and Chistyakova, L.G. 1990. Cloning, structure and expression of the full *lon* gene of *Escherichia coli* coding for ATP-dependent protease La. *Bioorg. Khim.* **16:** 869–880.

Amerik, A.Y., Antonov, V.K., Gorbalenya, A.E., Kotova, S.A., Rotanova, T.V., and Shimbarevich, E.V. 1991. Site-directed mutagenesis of La protease. A catalytically active serine residue. *FEBS Lett.* **287:** 211–214.

Baumeister, W., Walz, J., Zuhl, F., and Seemuller, E. 1998. The proteasome: Paradigm of a self-compartmentalizing protease. *Cell* **92:** 367–380.

Bell, C.E., Frescura, P., Hochschild, A., and Lewis, M. 2000. Crystal structure of the λ repressor C-terminal domain provides a model for cooperative operator binding. *Cell* **101:** 801–811.

Beuron, F., Maurizi, M.R., Belnap, D.M., Kocsis, E., Booy, F.P., Kessel, M., and Steven, A.C. 1998. At sixes and sevens: Characterization of the symmetry mismatch of the ClpAP chaperone-assisted protease. *J. Struct. Biol.* **123:** 248–259.

Bewley, M.C., Graziano, V., Griffin, K., and Flanagan, J.M. 2006. The asymmetry in the mature amino-terminus of ClpP facilitates a local symmetry match in ClpAP and ClpXP complexes. *J. Struct. Biol.* **153:** 113–128.

Bochtler, M., Ditzel, L., Groll, M., and Huber, R. 1997. Crystal structure of heat shock locus V (HslV) from *Escherichia coli*. *Proc. Natl. Acad. Sci.* **94:** 6070–6074.

Bochtler, M., Hartmann, C., Song, H.K., Bourenkov, G.P., Bartunik, H.D., and Huber, R. 2000. The structures of HslU and the ATP-dependent protease HslU-HslV. *Nature* **403:** 800–805.

Bota, D.A. and Davies, K.J. 2002. Lon protease preferentially degrades oxidized mitochondrial aconitase by an ATP-stimulated mechanism. *Nat. Cell Biol.* **4:** 674–680.

Botos, I., Melnikov, E.E., Cherry, S., Khalatova, A.G., Rasulova, F.S., Tropea, J.E., Maurizi, M.R., Rotanova, T.V., Gustchina, A., and Wlodawer, A. 2004a. Crystal structure of the $AAA^+\alpha$ domain of *E. coli* Lon protease at 1.9 Å resolution. *J. Struct. Biol.* **146:** 113–122.

Botos, I., Melnikov, E.E., Cherry, S., Tropea, J.E., Khalatova, A.G., Rasulova, F., Dauter, Z., Maurizi, M.R., Rotanova, T.V., Wlodawer, A., et al. 2004b. The catalytic domain of *Escherichia coli* Lon protease has a unique fold and a Ser-Lys dyad in the active site. *J. Biol. Chem.* **279:** 8140–8148.

Botos, I., Melnikov, E.E., Cherry, S., Kozlov, S., Makhovskaya, O.V., Tropea, J.E., Gustchina, A., Rotanova, T.V., and Wlodawer, A. 2005. Atomic-resolution crystal structure of the proteolytic domain of *Archaeoglobus fulgidus* Lon reveals the conformational variability in the active sites of Lon proteases. *J. Mol. Biol.* **351:** 144–157.

Brotz-Oesterhelt, H., Beyer, D., Kroll, H.P., Endermann, R., Ladel, C., Schroeder, W., Hinzen, B., Raddatz, S., Paulsen, H., Henninger, K., et al. 2005. Dysregulation of bacterial proteolytic machinery by a new class of antibiotics. *Nat. Med.* **11:** 1082–1087.

Burton, R.E., Siddiqui, S.M., Kim, Y.I., Baker, T.A., and Sauer, R.T. 2001. Effects of protein stability and structure on substrate processing by the ClpXP unfolding and degradation machine. *EMBO J.* **20:** 3092–3100.

Choi, K.H. and Licht, S. 2005. Control of peptide product sizes by the energy-dependent protease ClpAP. *Biochemistry* **44:** 13921–13931.

Chung, C.H. and Goldberg, A.L. 1981. The product of the *lon (capR)* gene in *Escherichia coli* is the ATP-dependent protease, protease La. *Proc. Natl. Acad. Sci.* **78:** 4931–4935.

DeLaBarre, B. and Brunger, A.T. 2005. Nucleotide dependent motion and mechanism of action of p97/VCP. *J. Mol. Biol.* **347:** 437–452.

Dodson, G. and Wlodawer, A. 1998. Catalytic triads and their relatives. *Trends Biochem. Sci.* **23:** 347–352.

Ebel, W., Skinner, M.M., Dierksen, K.P., Scott, J.M., and Trempy, J.E. 1999. A conserved domain in *Escherichia coli* Lon protease is involved in substrate discriminator activity. *J. Bacteriol.* **181:** 2236–2243.

Feldman, A.R., Lee, J., Delmas, B., and Paetzel, M. 2006. Crystal structure of a novel viral protease with a serine/lysine catalytic dyad mechanism. *J. Mol. Biol.* **358:** 1378–1389.

Goldberg, A.L. 1992. The mechanism and functions of ATP-dependent proteases in bacterial and animal cells. *Eur. J. Biochem.* **203:** 9–23.

Goldberg, A.L., Moerschell, R.P., Chung, C.H., and Maurizi, M.R. 1994. ATP-dependent protease La (lon) from *Escherichia coli*. *Methods Enzymol.* **244:** 350–375.

Gottesman, S. 1996. Proteases and their targets in *Escherichia coli*. *Annu. Rev. Genet.* **30:** 465–506.

———. 2003. Proteolysis in bacterial regulatory circuits. *Annu. Rev. Cell Dev. Biol.* **19:** 565–587.

Gottesman, S. and Maurizi, M.R. 1992. Regulation by proteolysis: Energy-dependent proteases and their targets. *Microbiol. Rev.* **56:** 592–621.

Groll, M., Ditzel, L., Lowe, J., Stock, D., Bochtler, M., Bartunik, H.D., and Huber, R. 1997. Structure of 20S proteasome from yeast at 2.4 Å resolution. *Nature* **386:** 463–471.

Groll, M., Bajorek, M., Kohler, A., Moroder, L., Rubin, D.M., Huber, R., Glickman, M.H., and Finley, D. 2000. A gated channel into the proteasome core particle. *Nat. Struct. Biol.* **7:** 1062–1067.

Groll, M., Bochtler, M., Brandstetter, H., Clausen, T., and Huber, R. 2005. Molecular machines for protein degradation. *ChemBioChem* **6:** 222–256.

Guo, F., Maurizi, M.R., Esser, L., and Xia, D. 2002. Crystal structure of ClpA, an Hsp100 chaperone and regulator of ClpAP protease. *J. Biol. Chem.* **277:** 46743–46752.

Hegerl, R., Pfeifer, G., Puhler, G., Dahlmann, B., and Baumeister, W. 1991. The three-dimensional structure of proteasomes from *Thermoplasma acidophilum* as determined by electron microscopy using random conical tilting. *FEBS Lett.* **283:** 117–121.

Hoskins, J.R., Pak, M., Maurizi, M.R., and Wickner, S. 1998. The role of the ClpA chaperone in proteolysis by ClpAP. *Proc. Natl. Acad. Sci.* **95:** 12135–12140.

Hua, S.S. and Markovitz, A. 1972. Multiple regulator gene control of the galactose operon in *Escherichia coli* K-12. *J. Bacteriol.* **110:** 1089–1099.

Im, Y.J., Na, Y., Kang, G.B., Rho, S.H., Kim, M.K., Lee, J.H., Chung, C.H., and Eom, S.H. 2004. The active site of a Lon protease from *Methanococcus jannaschii* distinctly differs from the canonical catalytic dyad of Lon proteases. *J. Biol. Chem.* **279:** 53451–53457.

Iyer, L.M., Leipe, D.D., Koonin, E.V., and Aravind, L. 2004. Evolutionary history and higher order classification of AAA^+ ATPases. *J. Struct. Biol.* **146:** 11–31.

Kang, S.G., Maurizi, M.R., Thompson, M., Mueser, T., and Ahvazi, B. 2004. Crystallography and mutagenesis point to an essential role for the N-terminus of human mitochondrial ClpP. *J. Struct. Biol.* **148:** 338–352.

Kenniston, J.A., Baker, T.A., Fernandez, J.M., and Sauer, R.T. 2003. Linkage between ATP consumption and mechanical unfolding during the protein processing reactions of an AAA^+ degradation machine. *Cell* **114:** 511–520.

Kim, D.Y. and Kim, K.K. 2003. Crystal structure of ClpX molecular chaperone from *Helicobacter pylori*. *J. Biol. Chem.* **278:** 50664–50670.

Kim, Y.I., Burton, R.E., Burton, B.M., Sauer, R.T., and Baker, T.A. 2000. Dynamics of substrate denaturation and translocation by the ClpXP degradation machine. *Mol. Cell* **5:** 639–648.

Kisselev, A.F., Akopian, T.N., Woo, K.M., and Goldberg, A.L. 1999. The sizes of peptides generated from protein by mammalian 26 and 20 S

proteasomes. Implications for understanding the degradative mechanism and antigen presentation. *J. Biol. Chem.* **274:** 3363–3371.

Kraut, J. 1977. Serine proteases: Structure and mechanism of catalysis. *Annu. Rev. Biochem.* **46:** 331–358.

Krzywda, S., Brzozowski, A.M., Verma, C., Karata, K., Ogura, T., and Wilkinson, A.J. 2002. The crystal structure of the AAA domain of the ATP-dependent protease FtsH of *Escherichia coli* at 1.5 Å resolution. *Structure* **10:** 1073–1083.

Langer, T. 2000. AAA proteases: Cellular machines for degrading membrane proteins. *Trends Biochem. Sci.* **25:** 247–251.

Lee, A.Y., Hsu, C.H., and Wu, S.H. 2004. Functional domains of *Brevibacillus thermoruber* lon protease for oligomerization and DNA binding: Role of N-terminal and sensor and substrate discrimination domains. *J. Biol. Chem.* **279:** 34903–34912.

Lenzen, C.U., Steinmann, D., Whiteheart, S.W., and Weis, W.I. 1998. Crystal structure of the hexamerization domain of N-ethylmaleimide-sensitive fusion protein. *Cell* **94:** 525–536.

Li, M., Rasulova, F., Melnikov, E.E., Rotanova, T.V., Gustchina, A., Maurizi, M.R., and Wlodawer, A. 2005. Crystal structure of the N-terminal domain of *E. coli* Lon protease. *Protein Sci.* **14:** 2895–2900.

Lowe, J., Stock, D., Jap, B., Zwickl, P., Baumeister, W., and Huber, R. 1995. Crystal structure of the 20S proteasome from the archaeon *T. acidophilum* at 3.4 Å resolution. *Science* **268:** 533–539.

Luo, Y., Pfuetzner, R.A., Mosimann, S., Paetzel, M., Frey, E.A., Cherney, M., Kim, B., Little, J.W., and Strynadka, N.C. 2001. Crystal structure of LexA: A conformational switch for regulation of self-cleavage. *Cell* **106:** 585–594.

Lupas, A.N. and Martin, J. 2002. AAA proteins. Curr. Opin. Struct. *Biol.* **12:** 746–753.

Martin, A., Baker, T.A., and Sauer, R.T. 2005. Rebuilt AAA^+ motors reveal operating principles for ATP-fuelled machines. *Nature* **437:** 1115–1120.

Maurizi, M.R. and Li, C.C. 2001. AAA proteins: In search of a common molecular basis. International Meeting on Cellular Functions of AAA Proteins. *EMBO Rep.* **2:** 980–985.

Melnikov, E.E., Tsirulnikov, K.B., and Rotanova, T.V. 2001. Coupling of proteolysis to ATP hydrolysis upon *Escherichia coli* Lon protease functioning II. Hydrolysis of ATP and activity of the enzyme peptide hydrolase sites. *Bioorg. Khim.* **27:** 120–129.

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

Nishii, W., Suzuki, T., Nakada, M., Kim, Y.T., Muramatsu, T., and Takahashi, K. 2005. Cleavage mechanism of ATP-dependent Lon protease toward ribosomal S2 protein. *FEBS Lett.* **579:** 6846–6850.

Niwa, H., Tsuchiya, D., Makyio, H., Yoshida, M., and Morikawa, K. 2002. Hexameric ring structure of the ATPase domain of the membrane-integrated metalloprotease FtsH from *Thermus thermophilus* HB8. *Structure* **10:** 1415–1423.

Ogura, T. and Wilkinson, A.J. 2001. AAA^+ superfamily ATPases: Common structure–diverse function. *Genes Cells* **6:** 575–597.

Oh, J.Y., Eun, Y.M., Yoo, S.J., Seol, J.H., Seong, I.S., Lee, C.S., and Chung, C.H. 1998. LonR9 carrying a single Glu614 to Lys mutation inhibits the ATP- dependent protease La (Lon) by forming mixed oligomeric complexes. *Biochem. Biophys. Res. Commun.* **250:** 32–35.

Ondrovicova, G., Liu, T., Singh, K., Tian, B., Li, H., Gakh, O., Perecko, D., Janata, J., Granot, Z., Orly, J., et al. 2005. Cleavage site selection within a folded substrate by the ATP-dependent lon protease. *J. Biol. Chem.* **280:** 25103–25110.

Ovchinnikova, T.V., Martynova, N.Yu., Potapenko, N.A., Vasilyeva, O.V., Golikova, N.Yu., Ginodman, L.M., and Rotanova, T.V. 1998. Study of heat shock bacterial proteins: Limited proteolysis of ATP-dependent protease Lon from. *E. coli. Biomed. Tekhnol.* **8:** 62–65.

Paetzel, M., Dalbey, R.E., and Strynadka, N.C. 2002. Crystal structure of a bacterial signal peptidase apoenzyme: Implications for signal peptide binding and the Ser-Lys dyad mechanism. *J. Biol. Chem.* **277:** 9512–9519.

Park, S.C., Jia, B., Yang, J.K., Van, D.L., Shao, Y.G., Han, S.W., Jeon, Y.J., Chung, C.H., and Cheong, G.W. 2006. Oligomeric structure of the ATP-dependent protease La (Lon) of *Escherichia coli*. *Mol. Cells* **21:** 129–134.

Patterson, J., Vineyard, D., Thomas-Wohlever, J., Behshad, R., Burke, M., and Lee, I. 2004. Correlation of an adenine-specific conformational change with the ATP-dependent peptidase activity of *Escherichia coli* Lon. *Biochemistry* **43:** 7432–7442.

Peat, T.S., Frank, E.G., McDonald, J.P., Levine, A.S., Woodgate, R., and Hendrickson, W.A. 1996. Structure of the UmuD′ protein and its regulation in response to DNA damage. *Nature* **380:** 727–730.

Rasulova, F.S., Dergousova, N.I., Melnikov, E.E., Ginodman, L.M., and Rotanova, T.V. 1998a. Synthesis and characterisation of ATP-dependent forms of Lon-proteinase with modified N-terminal domain from *Escherichia coli*. *Bioorg. Khim.* **24:** 370–375.

Rasulova, F.S., Dergousova, N.I., Starkova, N.N., Melnikov, E.E., Rumsh, L.D., Ginodman, L.M., and Rotanova, T.V. 1998b. The isolated proteolytic domain of *Escherichia coli* ATP-dependent protease Lon exhibits the peptidase activity. *FEBS Lett.* **432:** 179–181.

Rawlings, N.D., Tolle, D.P., and Barrett, A.J. 2004. MEROPS: The peptidase database. *Nucleic Acids Res.* **32:** D160–D164.

Reid, B.G., Fenton, W.A., Horwich, A.L., and Weber-Ban, E.U. 2001. ClpA mediates directional translocation of substrate proteins into the ClpP protease. *Proc. Natl. Acad. Sci.* **98:** 3768–3772.

Rotanova, T.V. 2002. Peptide hydrolases with catalytic dyad Ser-Lys. Similarity and distinctions of the active centers of ATP-dependent Lon proteases, LexA repressors, signal peptidases and C-terminal processing proteases. *Vopr. Med. Khim.* **48:** 541–552.

Rotanova, T.V., Melnikov, E.E., and Tsirulnikov, K.B. 2003. Catalytic dyad Ser-Lys at the active site of *Escherichia coli* ATP-dependent Lon-proteinase. *Bioorg. Khim.* **29:** 97–99.

Rotanova, T.V., Melnikov, E.E., Khalatova, A.G., Makhovskaya, O.V., Botos, I., Wlodawer, A., and Gustchina, A. 2004. Classification of ATP-dependent proteases Lon and comparison of the active sites of their proteolytic domains. *Eur. J. Biochem.* **271:** 4865–4871.

Roudiak, S.G. and Shrader, T.E. 1998. Functional role of the N-terminal region of the Lon protease from *Mycobacterium smegmatis*. *Biochemistry* **37:** 11255–11263.

Roudiak, S.G., Seth, A., Knipfer, N., and Shrader, T.E. 1998. The lon protease from *Mycobacterium smegmatis*: Molecular cloning, sequence analysis, functional expression, and enzymatic characterization. *Biochemistry* **37:** 377–386.

Rouiller, I., DeLaBarre, B., May, A.P., Weis, W.I., Brunger, A.T., Milligan, R.A., and Wilson-Kubalek, E.M. 2002. Conformational changes of the multifunction p97 AAA ATPase during its ATPase cycle. *Nat. Struct. Biol.* **9:** 950–957.

Rudyak, S.G., Brenowitz, M., and Shrader, T.E. 2001. Mg^{2+}-linked oligomerization modulates the catalytic activity of the Lon (La) protease from *Mycobacterium smegmatis*. *Biochemistry* **40:** 9317–9323.

Sauer, R.T., Bolon, D.N., Burton, B.M., Burton, R.E., Flynn, J.M., Grant, R.A., Hersch, G.L., Joshi, S.A., Kenniston, J.A., Levchenko, I., et al. 2004. Sculpting the proteome with AAA+ proteases and disassembly machines. *Cell* **119:** 9–18.

Seong, I.S., Kang, M.S., Choi, M.K., Lee, J.W., Koh, O.J., Wang, J., Eom, S.H., and Chung, C.H. 2002. The C-terminal tails of HslU ATPase act as a molecular switch for activation of HslV peptidase. *J. Biol. Chem.* **277:** 25976–25982.

Singh, S.K., Grimaud, R., Hoskins, J.R., Wickner, S., and Maurizi, M.R. 2000. Unfolding and internalization of proteins by the ATP-dependent proteases ClpXP and ClpAP. *Proc. Natl. Acad. Sci.* **97:** 8898–8903.

Sousa, M.C. and McKay, D.B. 2001. Structure of *Haemophilus influenzae* HslV protein at 1.9 Å resolution, revealing a cation-binding site near the catalytic site. *Acta Crystallogr.* **D57:** 1950–1954.

Sousa, M.C., Trame, C.B., Tsuruta, H., Wilbanks, S.M., Reddy, V.S., and McKay, D.B. 2000. Crystal and solution structures of an HslUV protease–chaperone complex. *Cell* **103:** 633–643.

Stahlberg, H., Kutejova, E., Suda, K., Wolpensinger, B., Lustig, A., Schatz, G., Engel, A., and Suzuki, C.K. 1999. Mitochondrial Lon of *Saccharomyces cerevisiae* is a ring-shaped protease with seven flexible subunits. *Proc. Natl. Acad. Sci.* **96:** 6787–6790.

Starkova, N.N., Koroleva, E.P., Rumsh, L.D., Ginodman, L.M., and Rotanova, T.V. 1998. Mutations in the proteolytic domain of *Escherichia coli* protease Lon impair the ATPase activity of the enzyme. *FEBS Lett.* **422:** 218–220.

Stock, D., Gibbons, C., Arechaga, I., Leslie, A.G., and Walker, J.E. 2000. The rotary mechanism of ATP synthase. *Curr. Opin. Struct. Biol.* **10:** 672–679.

Tsirulnikov, K.B., Melnikov, E.E., and Rotanova, T.V. 2003. Proteolysis coupled with ATP. Regulation of activity of proteolytic centers of *Escherichia coli* lon protease. *Bioorg. Khim.* **29:** 486–494.

van Dijl, J.M., Kutejova, E., Suda, K., Perecko, D., Schatz, G., and Suzuki, C.K. 1998. The ATPase and protease domains of yeast mitochondrial Lon: Roles in proteolysis and respiration-dependent growth. *Proc. Natl. Acad. Sci.* **95:** 10584–10589.

Van Dyck, L., Pearce, D.A., and Sherman, F. 1994. PIM1 encodes a mitochondrial ATP-dependent protease that is required for mitochondrial function in the yeast *Saccharomyces cerevisiae*. *J. Biol. Chem.* **269:** 238–242.

Vasilyeva, O.V., Kolygo, K.B., Leonova, Y.F., Potapenko, N.A., and Ovchinnikova, T.V. 2002. Domain structure and ATP-induced conformational changes in *Escherichia coli* protease Lon revealed by limited proteolysis and autolysis. *FEBS Lett.* **526:** 66–70.

Vineyard, D., Patterson-Ward, J., Berdis, A.J., and Lee, I. 2005. Monitoring the timing of ATP hydrolysis with activation of peptide cleavage in *Escherichia coli* Lon by transient kinetics. *Biochemistry* **44:** 1671–1682.

von Janowsky, B., Knapp, K., Major, T., Krayl, M., Guiard, B., and Voos, W. 2005. Structural properties of substrate proteins determine their proteolysis by the mitochondrial AAA^+ protease Pim1. *Biol. Chem.* **386:** 1307–1317.

Wang, J. 2004. Nucleotide-dependent domain motions within rings of the RecA/AAA(+) superfamily. *J. Struct. Biol.* **148:** 259–267.

Wang, J., Hartling, J.A., and Flanagan, J.M. 1997. The structure of ClpP at 2.3 Å resolution suggests a model for ATP-dependent proteolysis. *Cell* **91:** 447–456.

Wang, J., Song, J.J., Franklin, M.C., Kamtekar, S., Im, Y.J., Rho, S.H., Seong, I.S., Lee, C.S., Chung, C.H., and Eom, S.H. 2001. Crystal structures of the HslVU peptidase-ATPase complex reveal an ATP-dependent proteolysis mechanism. *Structure* **9:** 177–184.

Waxman, L. and Goldberg, A.L. 1982. Protease La from *Escherichia coli* hydrolyzes ATP and proteins in a linked fashion. *Proc. Natl. Acad. Sci.* **79:** 4883–4887.

Weber-Ban, E.U., Reid, B.G., Miranker, A.D., and Horwich, A.L. 1999. Global unfolding of a substrate protein by the Hsp100 chaperone ClpA. *Nature* **401:** 90–93.

Whitby, F.G., Masters, E.I., Kramer, L., Knowlton, J.R., Yao, Y., Wang, C.C., and Hill, C.P. 2000. Structural basis for the activation of 20S proteasomes by 11S regulators. *Nature* **408:** 115–120.

Wickner, S., Maurizi, M.R., and Gottesman, S. 1999. Posttranslational quality control: Folding, refolding, and degrading proteins. *Science* **286:** 1888–1893.

Xia, D., Esser, L., Singh, S.K., Guo, F., and Maurizi, M.R. 2004. Crystallographic investigation of peptide binding sites in the N-domain of the ClpA chaperone. *J. Struct. Biol.* **146:** 166–179.

Zhang, X., Shaw, A., Bates, P.A., Newman, R.H., Gowen, B., Orlova, E., Gorman, M.A., Kondo, H., Dokurno, P., Lally, J., et al. 2000. Structure of the AAA ATPase p97. *Mol. Cell* **6:** 1473–1484.