

Structural Insights on the *Mycobacterium tuberculosis* Proteasomal ATPase Mpa

Tao Wang,¹ Hua Li,¹ Gang Lin,² Chunyan Tang,¹ Dongyang Li,¹ Carl Nathan,² K. Heran Darwin,^{3,*} and Huilin Li^{1,4,*}

¹Biology Department, Brookhaven National Laboratory, Upton, NY 11973-5000, USA

²Department of Microbiology and Immunology, Weill Medical College of Cornell University, New York, NY 10021, USA

³Department of Microbiology, New York University School of Medicine, New York, NY 10016, USA

⁴Department of Biochemistry and Cell Biology, Stony Brook University, Stony Brook, NY 11794, USA

*Correspondence: hli@bnl.gov (H.L.), heran.darwin@med.nyu.edu (K.H.D.)

DOI 10.1016/j.str.2009.08.010

SUMMARY

Proteasome-mediated protein turnover in all domains of life is an energy-dependent process that requires ATPase activity. *Mycobacterium tuberculosis* (Mtb) was recently shown to possess a ubiquitin-like proteasome pathway that plays an essential role in Mtb resistance to killing by products of host macrophages. Here we report our structural and biochemical investigation of Mpa, the presumptive Mtb proteasomal ATPase. We demonstrate that Mpa binds to the Mtb proteasome in the presence of ATP γ S, providing the physical evidence that Mpa is the proteasomal ATPase. X-ray crystallographic determination of the conserved interdomain showed a five stranded double β barrel structure containing a Greek key motif. Structure and mutational analysis indicate a major role of the interdomain for Mpa hexamerization. Our mutational and functional studies further suggest that the central channel in the Mpa hexamer is involved in protein substrate translocation and degradation. These studies provide insights into how a bacterial proteasomal ATPase interacts with and facilitates protein degradation by the proteasome.

INTRODUCTION

The proteasome carries out the majority of cytosolic protein degradation inside a eukaryotic cell. In doing so, the proteasome regulates a multitude of cellular functions such as protein homeostasis, signaling, cell cycle progression, and the production of antigenic peptides (Goldberg, 2007; Hershko and Ciechanover, 1998). The proteasome has emerged as an effective drug target, and the proteasome inhibitor bortezomib (Velcade), a dipeptide boronate, has been approved for the treatment of multiple myeloma (Borissenko and Groll, 2007). The proteasome is found ubiquitously in eukaryotes and archaea and some bacteria (Voges et al., 1999). The structure of the 20S core particle (CP) and its proteolytic mechanism are highly conserved across all domains of life: the enzyme is a threonine protease composed of α and β subunits arranged in $\alpha_7\beta_7\beta_7\alpha_7$ fashion into a cylindrical structure (Groll et al., 1997).

The CP alone degrades only small peptides or unstructured proteins. Degradation of normally folded protein substrates is usually an energy-dependent process that requires ATPase activity provided by additional structures (Smith et al., 2006; Voges et al., 1999). The eukaryotic proteasome contains a 19S complex at one or both ends of the CP. The yeast 19S cap contains at least 19 protein subunits, six of which (Rpt1–6) are ATPases that contact the CP (Schmidt et al., 2005). Due to its large size and highly dynamic nature, crystallographic solution of the 19S cap or the entire 26S proteasome has not been achieved, leaving critical questions as to how proteasomal ATPases unfold and translocate protein substrates for degradation by the CP.

Archaeal proteasomes have a simpler ATPase system in which six identical proteins form a homohexamer, loosely capping the seven-fold CP (Smith et al., 2007). These hexameric ATPases hydrolyze ATP to actively unfold and translocate substrates through a gate into the proteolytic chamber (Benaroudj et al., 2003; Medalia et al., 2009). Unlike archaea, most bacteria do not contain proteasomes and instead synthesize other protease complexes, such as ClpAP, ClpXP, HslUV, Lon, and FtsH (Baker and Sauer, 2006; Butler et al., 2006). However, proteasome genes, together with proteasomal ATPase genes, have been identified in bacteria of the order *Actinomycetales* (Nagy et al., 1998; Wolf et al., 1998; Zhang et al., 2004). The functions of these bacterial proteasomes are largely unknown. Recently, the proteasome and its associated ATPase in *Mycobacterium tuberculosis* (Mtb), a species of *Actinomycetales*, were found to play a critical role in bacterial resistance to killing by reactive nitrogen intermediates, which are products of host macrophages (Darwin et al., 2003). Moreover, suppression of expression of the Mtb proteasome led to death of the pathogen in mice (Gandotra et al., 2007). Thus, the Mtb proteasome is currently being targeted for the development of antituberculosis agents (Nathan et al., 2008).

The Mtb proteasomal genes *prcA* and *prcB* encode bona fide proteasome subunits (Lin et al., 2006), and the Mtb 20S proteasome structure is essentially identical to eukaryotic and archaeal proteasomes (Hu et al., 2006). The small post-translational modifier Pup in Mtb covalently tags proteasome substrates for degradation in a manner that is functionally analogous to ubiquitylation in eukaryotes (Pearce et al., 2008). Thus, the Mtb 20S-Mpa pathway may be a simpler model system for the structural investigation of how certain tagged proteins are recognized, unfolded, and degraded by proteasomes.

Mpa was shown to have ATPase activity (Darwin et al., 2005), and degradation of proteasome substrates in Mtb required the

presence of Mpa, together with an additional proteasome accessory factor (PafA), a Pup ligase (Festa et al., 2007; Pearce et al., 2006, 2008; Striebel et al., 2009). However, there were no data that showed any bacterial ATPase interacted with its cognate proteasome. In this paper, we demonstrate for the first time that Mpa could indeed directly interact with the Mtb 20S CP. Mpa is a multidomain structure, with an N-terminal predicted coiled coil domain, a 150 amino acid interdomain (Mpa-ID) that is unique to the proteasome-associated ATPases, a canonical AAA (ATPase associated with various activities) domain, and a small C-terminal domain. We determined the crystal structure of the Mpa-ID, which formed a tightly packed ring-shaped hexamer in the crystal structure as well as in solution. Finally, we demonstrate that specific point mutations in Mpa selected based on our structural analysis could either abolish hexamerization in vitro or proteasome substrate degradation in Mtb.

RESULTS

Mpa Directly Interacts with the Mtb 20S Proteasome Core Particle

The requirement for Mpa in Mtb 20S proteolysis has been strongly implicated (Darwin et al., 2003; Pearce et al., 2006, 2008). However, a direct interaction between the 20S CP and Mpa has not been demonstrated. We investigated by electron microscopy (EM) the binding between the full-length Mpa or an N-terminal 97 residue deletion mutant Mpa (Mpa_{N97Δ}) with the wild-type or the open gate mutant 20S (20SOG) CP in the presence of various nucleotides (see Electron Microscopy of Mpa-20SOG Complex for sample preparation) (Figure 1). Mpa_{N97Δ} has the N-terminal predicted coiled coil domain removed. The 20SOG CP is an N-terminal octapeptide deletion mutant at the α subunit of the 20S CP. The proteolytic activity of this mutant increased by an order of magnitude, and the mutant was thus designated open gate (Lin et al., 2006). No significant binding was observed between Mpa or Mpa_{N97Δ} and the wild-type 20S CP (data not shown). However, we found that both forms of Mpa were able to bind the 20SOG CP in the presence of 1 mM ATP γ S (Figure 1 and see Figure S1 available online). Under these conditions, Mpa_{N97Δ} had higher affinity for the 20SOG than the full-length Mpa, based on the frequency with which the 20SOG was found to be in complex, which was \sim 15% and \sim 2% for Mpa_{N97Δ} and Mpa, respectively. The orientation of the bound Mpa on 20SOG appears to be highly flexible (Figure 1D and Figure S1C). From particles in which Mpa binds approximately along the axial position of the 20SOG CP, it appears that the C-terminal end of Mpa contacts the α ring of the CP (Figure 1E).

The Conserved Mpa-ID Forms an Oligomer in Solution

Bioinformatic analysis of the Mpa indicated a separately folded domain, \sim 150 amino acids long, from Pro-97 to Glu-245, that we termed Mpa-ID, in a region between the N-terminal coiled coil domain and the C-terminal AAA domain. This domain appears to be unique to proteasomal ATPases (Figure 2A) (Zhang et al., 2004). This stably-folded domain was identified by limited trypsin digestion and mass spectroscopy (data not shown). Mpa-ID eluted from the gel filtration column at a volume corresponding to \sim 100 kD (Figure 2B), suggesting that Mpa-ID

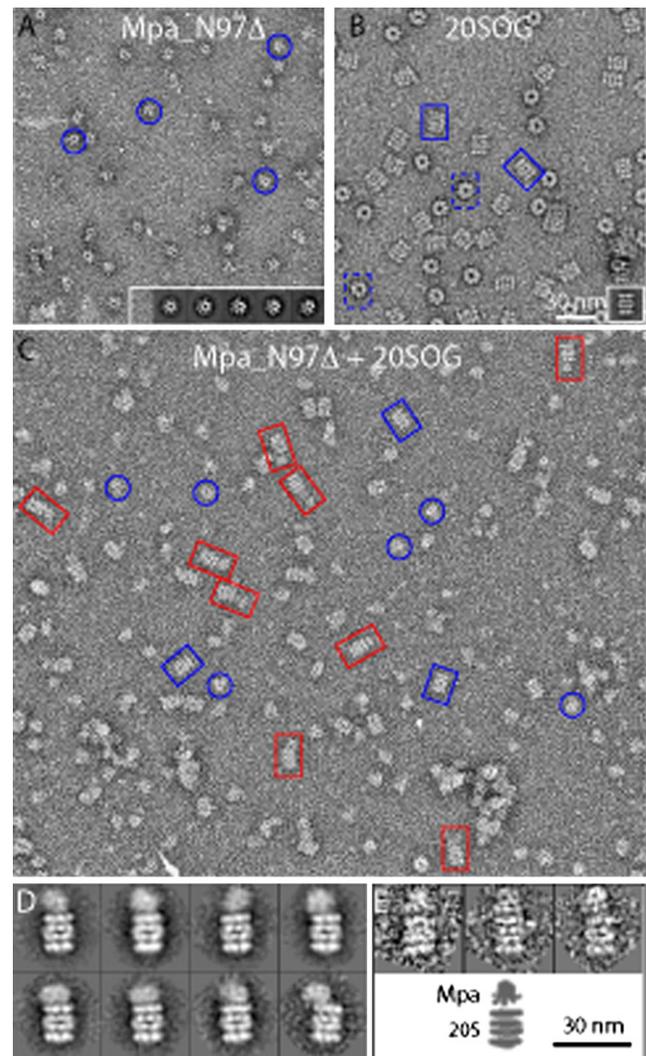


Figure 1. Mpa Directly Interacts with the Open Gate Form (20SOG) of the Mtb Proteasome

(A) A small region of a raw electron micrograph of negatively stained Mpa_{N97Δ}. (Inset) Five class averages of the deletion mutant Mpa particles. The blue circles mark several particles.
 (B) A raw EM image of the Mtb 20SOG proteasome. (Inset) 2D class average of the proteasome particles in the side view. The blue rectangles mark individual particles in side (solid line) or top (dashed line) views. (A) and (B) are in the same scale.
 (C) A raw EM micrograph of 20SOG incubated with Mpa_{N97Δ} in the presence of ATP γ S. The red rectangles mark several 20SOG-Mpa complexes.
 (D) Eight selected 2D class averages of the Mpa-20SOG complex particles showing the flexible nature of the interaction between Mpa and proteasome.
 (E) (Top) Three raw particle images of the Mpa-20SOG complex in which the Mpa binds approximately axially on the proteasome. (Bottom) a sketch showing that the Mpa C-terminal portion with weaker density interfaces with the 20SOG. (D) and (E) are in the same scale.

oligomerized in solution as a hexamer. The purified Mpa-ID migrated at \sim 20 kD, near the predicted 16.6 kD in the Coomassie-stained SDS-PAGE (Figure 2C). We inspected the purified and His₆ tag-removed Mpa-ID by negative stain EM and found that Mpa-ID indeed formed a ring-like structure with an apparent six-fold symmetry (Figure 2D). The ring is about 65 Å in diameter

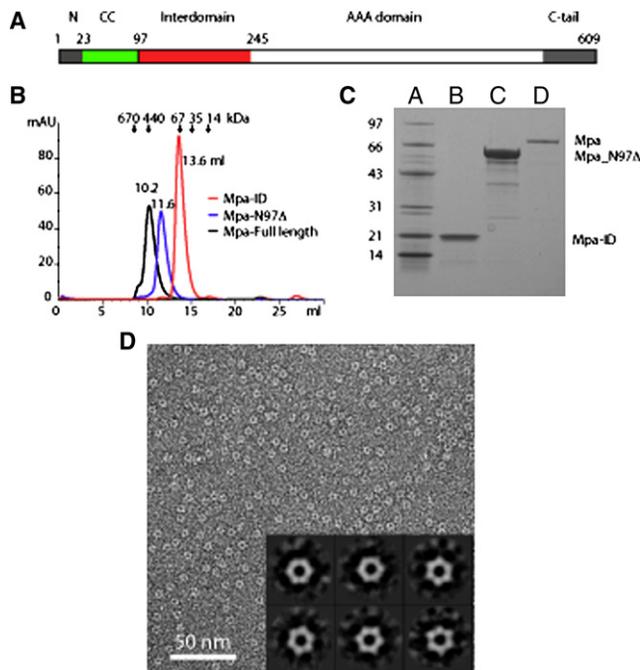


Figure 2. Mpa Has a Multidomain Architecture

(A) Overview of the domain organization of the full-length Mpa. (B) The Superdex-200 gel filtration profiles of the purified full-length Mpa is in black, Mpa_N97 Δ in blue, and Mpa-ID in red. Each of the three proteins is eluted at the volume corresponding to their respective hexameric form. (C) The Coomassie blue-stained SDS-PAGE of the purified Mpa-ID (lane 2), Mpa_N97 Δ (lane 3), and Mpa (lane 4). Lane 1 is the molecular weight marker in kilodaltons. (D) The negative stain EM image of the purified Mpa-ID and the inserted six 2D class averages reveal a hexameric structure about 65 Å in diameter with an \sim 20 Å central low density channel.

with an \sim 20 Å central pore. Based on these observations, we conclude that Mpa-ID on its own is capable of hexamerization and, consequently, this domain may contribute to the stability of the full-length Mpa hexamer that we previously visualized by negative stain EM (Darwin et al., 2005).

The Mpa-ID Has Two Subdomains, Each with a Five-Stranded β Barrel

We determined the atomic structure of Mpa-ID at 2.0 Å resolution by X-ray crystallography (See [Experimental Procedures](#), [Tables 1 and 2](#), and [Figure S2A](#) for the quality of the electron density map). The refined structure has an R_{free} of 26% and R_{work} of 21% and includes all residues except in a long loop region between Leu-188 and Gly-217. This loop can be fully traced only in one monomer ([Figure S2B](#)); in the other monomers, it is partially disordered and has a gap between Ile-193 and Leu-214.

The Mpa-ID contains two subdomains ([Figure 3A](#)). Each subdomain is a five stranded β barrel. Unexpectedly, the two β barrels are arranged in tandem and connected by a 6 amino acid loop between Thr154 and Gly159. In each subdomain, the five β strands form a g^-b type fold containing a left-handed Greek key motif, where g refers to Greek key motif and b is $\beta 5$ ([Figure 3B](#)) (Zhang and Kim, 2000). The interaction between the two subdomains is mainly of hydrophobic nature, involving

Table 1. Data Reduction and Structure Determination Statistics

	Native I (1°–360°)	Native II (1°–270°)	Hg ²⁺ derivative (1°–360°)
Data Set			
Oscillation	1°	1°	1°
Space group	P2 ₁	P2 ₁	P2 ₁
Unit cell parameters (Å)	a = 74.86	a = 75.39	a = 75.03
	b = 74.99	b = 75.92	b = 75.59
	c = 200.85	c = 192.74	c = 192.37
	α = 90	α = 90	α = 90
	β = 90.319	β = 90.015	β = 90.179
	γ = 90°	γ = 90°	γ = 90°
Resolution (Å)	20– \sim 2.0	20–2.6	20– \sim 2.8
Wavelength (Å)	1.000	1.000	1.0066
$R_{\text{merge}}^{\text{a,b}}$	0.21 (0.55) ^c	0.06 (0.60)	0.06 (0.49)
Completeness ^b	0.98 (0.89)	0.97 (0.93)	0.91 (0.82)
Mean of redundancy	7.2	4.0	3.2
$I / \sigma(I)^{\text{b}}$	13.0 (1.9)	22.1 (1.9)	14.9 (1.8)
f' (Hg)			–12.98
f'' (Hg)			9.74
Measured reflections	1,064,345	287,826	328,382
Unique reflections	147,655	66,845	98,072
Heavy atom sites			17

^a $R_{\text{merge}} = \sum (|I - \langle I \rangle|) / \sum (I^2)$.

^b Values in parentheses refer to the highest resolution shell.

^c High R_{merge} value maybe due to the presence of twinning or a small crack.

four large hydrophobic residues (Phe-116 and Phe-155 in subdomain 1 and Tyr-228 and Phe-230 in subdomain 2) and three smaller ones (Val-103, Leu-105, and Ala-226) ([Figure 3C](#)). In addition, there are three hydrogen bonds that further stabilize the interface (two H bonds between Lys-137 and Asp-190 and one H bond between Try-228 and Leu-105). Although no sequence repeat can be observed within the Mpa-ID, the two β barrels can be superimposed with a root mean square deviation (rmsd) of 1.9 Å ([Figure 3D](#)). The only significant difference is the connecting loop between the third and the fourth β strands in each β barrel, as highlighted in blue and orange, respectively ([Figure 3D](#)). The Mpa-ID tandem β barrels are followed by a C-terminal α helix ([Figure 3A](#)). This helix is a linker between the ID and the following AAA domain, because the predicted α/β subdomain of AAA domain starts immediately after the α helix.

The five-stranded β barrel with the Greek key motif as observed in Mpa-ID is also called the oligonucleotide and oligosaccharide binding fold (OB fold), due to the nature of its binding substrates (Murzin, 1993). The first β barrel of Mpa-ID can be superimposed with an rmsd of 2.5 Å with the OB fold structure found in the vaccinia virus protein K3L that is involved in substrate recognition and subversion of host immune defense ([Figure 3E](#)) (Dar and Sicheri, 2002) and with an rmsd of 1.5 Å with the OB fold in the transcription factor 5A (TF5A) ([Figure 3E](#)) (Kloks et al., 2002). Interestingly, TF5A has 21% sequence identity with Mpa-ID. In a canonical OB fold, an α helix links the $\beta 3$ and $\beta 4$ strands and consequently caps the β barrel ([Figure 3E](#)). However, in the first β barrel of Mpa-ID, a 14 amino acid loop

Table 2. Refinement Statistics of the Mpa-ID Crystal Structure

Resolution	20–2.0 (Å)
Number of reflections used	140,155
Number of nonhydrogen atoms	13,735
Number of water	866
R_{work}^a	0.213
R_{free}^a	0.266
Mean B value (Å ²)	33.4
Rmsd of bond angle (°)	1.682
Rmsd of bond length (Å)	0.018
Mean figure of merit	0.83

^a $R_{\text{work}} = \sum (|F_o - F_c|) / \sum |F_o|$; R_{free} is the same as R_{work} calculated with a test set of 5% random selected subset of the reflections that was not included in refinement calculations.

(Thr-125 to Gly-138), instead of an α helix, connects the β 3 and β 4 strands and in the second β barrel, a long (30 amino acids) and partially disordered loop (Leu 188 to Gly 217) connects the two corresponding β strands (Figure 3D). In the OB fold proteins, the nucleotide or sugar binding sites are clustered in a similar region: on the side surface of the β barrel where three loops, the loop between β 1 and β 2, the beginning section of the long loop between β 3 and β 4, and the loop between β 4 and β 5, come together (Murzin, 1993). In Figure 3A, these loops in the first β barrel are highlighted in blue and the region is labeled by an orange arrow. The corresponding region in the first β barrel roughly translates to the top surface in the Mpa-ID hexamer (Figure 3A and Figure 4). This region in the Mpa-ID second β barrel is shielded by intramolecular interaction. Given the role of Mpa in protein degradation, this analysis suggests that the top surface of the Mpa-ID hexamer is involved in protein-substrate recognition and binding.

The Mpa-ID Hexamer Structure Has a Negatively Charged Outside Surface and a Positively Charged Central Channel

Mpa-ID formed hexamers in the crystal structure (Figure 4A). In fact, two hexamers stacked end to end, forming a dodecamer, were the packing unit in the crystal structure. The interface between the two hexamers in the crystals is maintained by three hydrophobic residues in the C-terminal helix (Figure S3). The Mpa-ID hexamer is ~ 70 Å in diameter and ~ 60 Å high, comparable to the dimensions observed by EM (Figure 2D). Intriguingly, the Mpa-ID hexamer has a highly negatively charged side surface, but several positively charged patches at the top, which we have described above as contributing to a potential substrate-binding site (Figures 4A and 4C). The hexamer has a central channel. At the opening of the channel on the top surface, the six Glu-145 of the six β hairpin loops (β 4– β 5) define a large and negatively charged pore (28 Å) (Figures 4B–4D). The interior of the channel is, however, lined by highly positively charged residues in three layers formed by β hairpins from top to bottom: six Arg-120 of β 2– β 3 at the first layer define a 20 Å pore, six Lys-225 of β 9– β 10 at the middle layer define a 15 Å pore, and six His-179 of β 7– β 8 at the bottom layer also form a 15 Å pore (Figures 4B and 4D). The central channel of Mpa-

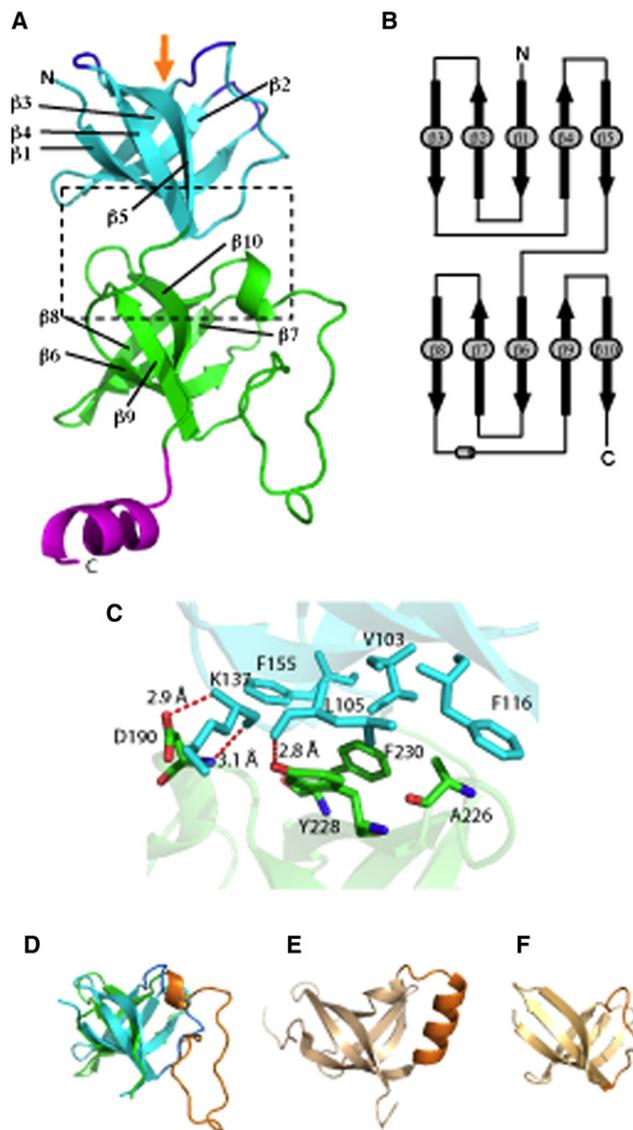


Figure 3. The Crystal Structure of the Mpa-ID

(A) Mpa-ID contains two five stranded β barrel subdomains connected by a loop. The C-terminal helix might be a linker between the interdomain and the following AAA domain. The orange arrow indicates the substrate recognition/binding region at the top as defined by the three β hairpin loops highlighted in blue.

(B) The topology of Mpa-ID contains two Greek key motifs formed by the first four β strands in each subdomain.

(C) An enlarged view of the interface between the two tandem β barrels as boxed in (A), showing the hydrophobic and H bond interactions.

(D) The two β barrel subdomains have a similar fold; they can be overlapped with an rmsd of 1.9 Å.

(E) and (F) are the OB fold structures of the vaccinia virus K3L (E; PDB ID 1LUZ; rmsd = 2.5 Å) and the TF5A (F; PDB ID 1EIF; rmsd = 1.5 Å; sequence identity = 21%), respectively, aligned with the first β barrel subdomain in Mpa-ID.

ID is funnel-shaped, constricting gradually from the top to the bottom by a factor of 2, from 28 Å to 15 Å (Figure 4B).

We tested the possible role of these positively charged amino acids on the recognition of a known proteasomal degradation substrate, malonyl coA-acyl carrier protein transacylase

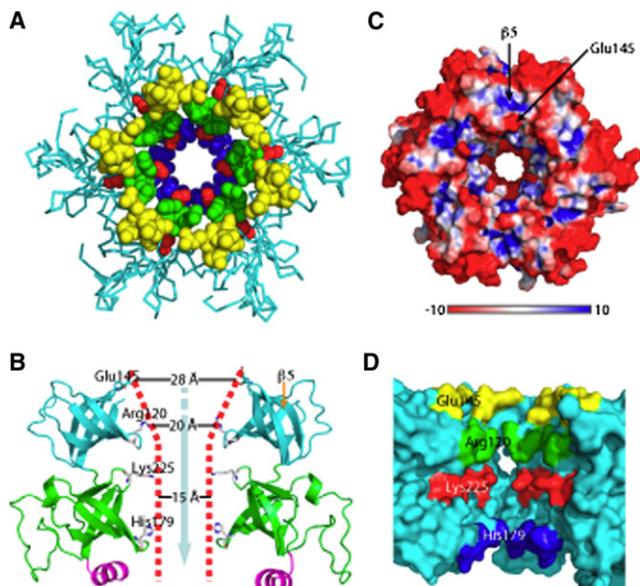


Figure 4. Crystal Structure of the Mpa-ID Hexamer

(A) The top view of the hexamer, with residues of the four β hairpins lining the central channel shown as spheres in different colors. (B) Side view showing only two diagonal subunits. The four residues lining the funnel-like central channel are highlighted in the ball-and-stick presentation. The potential peptide translocation path is illustrated by a cyan arrow. (C) Surface-charge presentation of the Mpa-ID hexamer in the top view. Arrows indicate the substrate-binding site in the OB fold above the $\beta 5$ strand and the negatively charged Glu-145 at the opening of the channel, respectively. (D) Surface rendered side view of the Mpa-ID hexamer with the front two subunits removed to show the interior channel. Charged residues lining the channels are labeled and shown in different colors.

(FabD). FabD is targeted for degradation by Pup, which is highly negatively charged (Pearce et al., 2008). Point mutations in Mpa-ID were made at either Arg-120 or Lys-225 and introduced into an *mpa* null mutant to determine if they could restore degradation of FabD. The predicted α/β subdomain, but not the α -helical subdomain, of the Mpa AAA domain can be aligned with mammalian p97 (43% identity and 61% similarity) and shares 38% identity with AAA domain of the bacterial membrane-associated protease FtsH (Figure 5A). The D2 domain of p97 was recently implicated in protein substrate binding and possibly unfoldase activity (DeLaBarre et al., 2006). ATPases involved in protein substrate unfolding and translocation contain a conserved Ar- Φ pore loop that was proposed to pull the unfolded peptide through the central channel (Martin et al., 2008a, 2008b). The Ar- Φ pore loop motif, containing single aromatic residue and a hydrophobic amino acid, is also conserved in Mpa (Figure 5A). Therefore, we also mutated the Φ residue Val-342. In addition, we analyzed several previously reported mutants that disrupt the in vitro ATPase activity or the penultimate tyrosine (Tyr) that is highly conserved in most proteasome-associated ATPases (Darwin et al., 2005).

Under steady-state conditions, mutations in the Walker boxes (required for ATPase activity), the penultimate Tyr (Tyr-608), or the pore loop dramatically abrogated FabD degradation (Figure 5B, lanes 1–7 and 9). Importantly, Mpa is itself a target

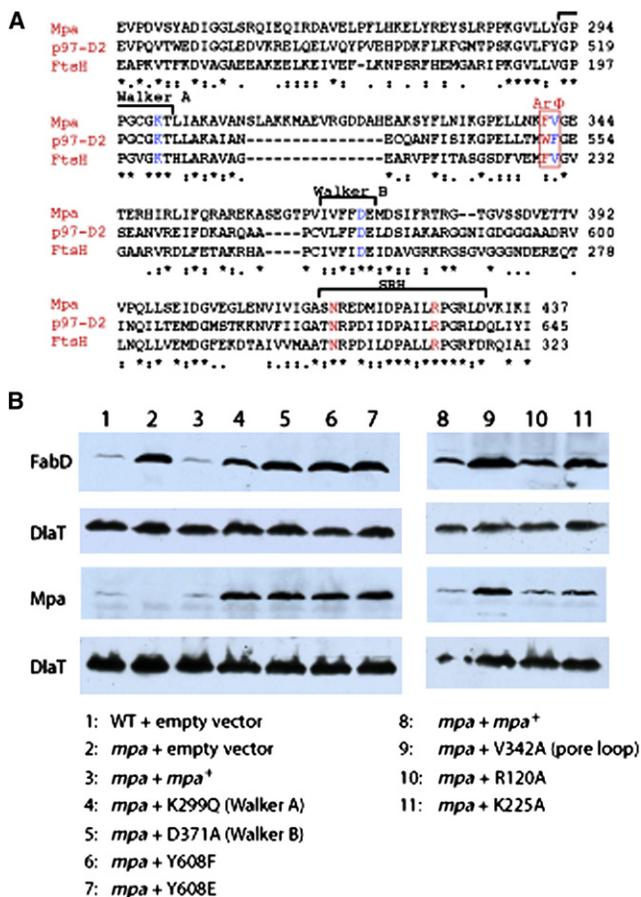


Figure 5. Site-Directed Mutations in Mpa Disrupt Protein Degradation

(A) Sequence alignment of Mpa with p97/VCP and FtsH reveals unexpected high identity at the α/β subdomain of the AAA domain with p97 D2 domain at 46% and FtsH at 38.5%. The conserved aromatic and hydrophobic residues in the pore-loop region of AAA domain are highlighted by a red box (Ar Φ). The Walker A/B motifs and the conserved sensor-1 residue (N) and the arginine finger (R) in the second region of homology (SRH) are also marked.

(B) Immunoblot analysis of total Mtb lysates of equivalent cell numbers of wild-type, *mpa* mutant, or *mpa* mutant Mtb transformed with plasmids encoding wild-type or mutated *mpa*. Samples were analyzed with either polyclonal antibodies to FabD or Mpa. The same blots were stripped and reanalyzed with antibodies to dihydroliipoamide dehydrogenase (DlaT), the loading control.

of proteasomal degradation (Pearce et al., 2006), thus these mutations also increased the steady-state levels of Mpa, but not of a control protein, DlaT. In contrast, Arg-120 or Lys-225, which we hypothesized may interact with pupylated substrate, did not appear to dramatically affect the degradation of FabD, (Figure 5B, lanes 10 and 11). This indicates that the central channel in the OB ring might not be directly involved in Pup recognition. However it is still possible that these amino acids are important for efficient degradation under different growth conditions or of other proteins.

The Second β Barrel Ring in Mpa-ID Contributes to the Stability of the Mpa Hexamer

The interactions between the monomers within the hexamer are quite different for the two subdomains: in the first or top β barrel

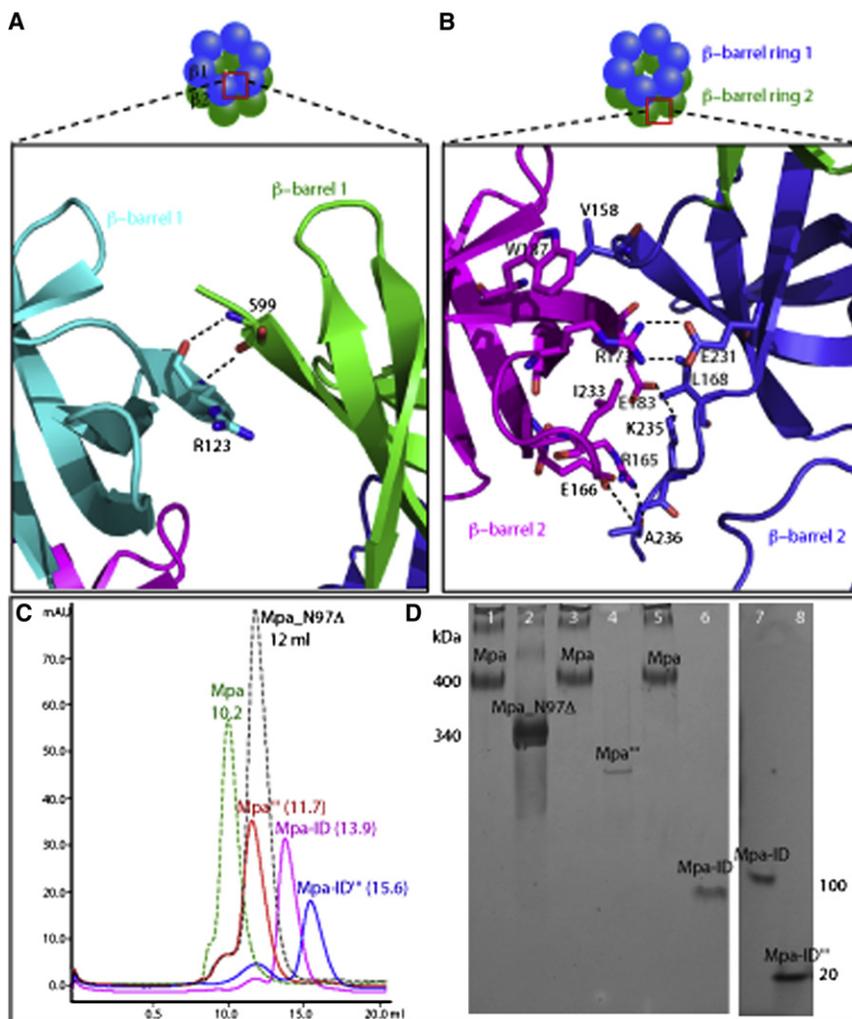


Figure 6. Subunit-Subunit Interactions in the Mpa-ID Hexamer

(A) The interface between two neighboring β barrels; one is limited, involving only two main chain H bonds between the Arg-123 and Ser-99. (B) The interactions between the two neighboring β barrels; two are extensive. At the top of the interface, Trp-187 and Val-158 form a hydrophobic interaction; at the middle section, the interface involves two salt bridges/H bonds (between Arg-173 and Glu-231 and between Glu-183 and Lys-235) and a pair of hydrophobic residues (between Ile-233 and Leu-168); and at the bottom section, Arg-165 and Glu-166 form two H bonds with the Ala-236. (C) Gel filtration profiles of the native full-length Mpa, Mpa_{N97} Δ , the full-length Mpa triple mutant (Mpa^{'''}), Mpa-ID, and Mpa-ID with the triple mutations (Mpa-ID^{'''}). (D) The blue native gel of the corresponding Mpa constructs purified as shown in (C). The right panel (lanes 7 and 8) was run separately for a shorter period of time than the left panel. This was to prevent the small Mpa-ID^{'''} (~20 kD) from running out of the gel.

ring, interactions between the two neighboring β barrels are minimal, involving only main chain hydrogen bonds (3.1 Å and 3.2 Å) between Ser-99 and Arg-123 (Figure 6A). In contrast, in the second β barrel ring, interactions between the two neighboring β barrels are extensive, involving two salt bridges (between Glu-231 and Arg-173, 2.9 Å; Lys-235 and Glu-183, 2.7 Å), two hydrophobic interactions (between Val-158 and Trp-187; Ile-233 and Leu-168), and two hydrogen bonds (between carboxyl O of Ala-236 with the NH₂ of Arg-165; 3.3 Å; between the main chain NH₂ of Ala-236 with the side chain carboxyl O of Glu-166, 2.73 Å) (Figure 6B). Therefore, the second β barrel ring of Mpa-ID appears to be in a position to contribute to the stability of the full-length Mpa hexamer. Mutagenesis was used to investigate this possibility. We found that disruption of the H bonds between Ala-236 and Arg-165 by deleting the Ala-236 did not abolish hexamerization of Mpa-ID (data not shown). Subsequently, we designed a triple mutant (R173E; W187A; K235E) that disrupts the salt bridges and the hydrophobic interaction. We found by gel filtration (Figure 6C) and blue native gel electrophoresis (Figure 6D) that the triple mutant Mpa-ID (denoted Mpa-ID^{'''}) was monomeric. Furthermore, no ring-like hexameric particles of Mpa-ID^{'''} were observed by negative stain EM (data not shown). The same triple mutations, when introduced

into the full-length Mpa, significantly reduced the production of the recombinant protein in *E. coli* and markedly altered the migration of Mpa in both gel filtration profile and blue native gel electrophoresis (Figures 6C and 6D). The mutant Mpa (Mpa^{'''}) eluted from the sizing column and migrated in the native gel between the hexameric Mpa (400 kD) and the hexameric Mpa-ID (100 kD) and appeared to form loosely associated trimeric particles (predicted mass 200

DISCUSSION

Weak In Vitro Interaction between Mtb 20S and Mpa

To our knowledge, this is the first demonstration that a bacterial proteasome can directly interact with its cognate ATPase. Previously, Smith et al. (2007) reported that proteasome-activating nucleotidase (PAN) can bind to a noncognate proteasome and stimulate protein degradation. Mpa was previously only circumstantially linked to the Mtb proteasome by genetic studies (Darwin et al., 2003; Pearce et al., 2006). We found that the in vitro interaction between Mpa and 20SOG was not stable to gel filtration (data not shown), and must thus be weak. In electron micrographs, only 15% of Mtb 20S CP were bound to the Mpa_{N97} Δ , although the Mpa ATPases were incubated at high molar ratio (4:1) with 20S CP. It is currently unclear why Mpa_{N97} Δ binds better than the full-length Mpa to 20S CP. We only observed Mpa binding to one end of 20S CP. This is likely due to the low in vitro affinity of “apo” Mpa for 20S CP.

Binding of PAN with the proteasome CP was also weak, undetectable by native PAGE, gel filtration, ultracentrifugation, or immunoprecipitation (Forster et al., 2003), and was only demonstrated by highly sensitive techniques such as surface plasmon resonance and EM (Smith et al., 2005). The off-axis and flexible binding of Mpa on 20S CP was also similar to the binding of PAN on 20S CP. Unlike PAN, we were unable to detect Mpa-enhanced 20S protein degradation in vitro with either peptides or β -casein as substrates (data not shown). These observations underscore the fact that other factors such as Pup are required for protein degradation by the Mpa-20S system in Mtb (Pearce et al., 2008; Striebel et al., 2009).

Functional Implications of Mpa-ID Hexamer with the OB Fold

Most recently, the archaeal PAN from *Methanocaldococcus jannaschii* was shown to contain a single OB fold structure (Zhang et al., 2009), and the AAA ATPase forming ring shaped complex (ARC) from the bacterium *Rhodococcus erythropolis* contained the tandem double OB motifs (Djuranovic et al., 2009). All these OB domains, like Mpa-ID, formed hexamers. Thus the hexameric OB ring appears to be a unique feature of all proteasomal ATPases (Djuranovic et al., 2009). The OB fold has been found in more than 80 other unrelated proteins. The OB fold is in general highly stable. For example, the OB fold structures in three unrelated proteins, the staphylococcal nuclease SN, the *E. coli* anticodon binding domain of Lys-tRNA synthetase, and the cold shock protein CspA, were shown by mutagenesis and hydrogen exchange experiments to be resistant to unfolding (Alexandrescu et al., 1995, 1999; Jaravine et al., 2000). We found that Mpa-ID is also stable and resistant to prolonged protease treatment. Therefore, the hexameric Mpa-ID with a total of 12 OB motifs might be mechanically strong. Given its high structural stability, we propose that the Mpa-ID hexamer might provide a rigid platform upon which the mechanical force exerted from the Mpa AAA domain unfolds the substrate protein. The Mpa AAA domain alone is unstable and does not form a hexamer (data not shown). Therefore, the Mpa OB ring is also the major stabilizing factor of the Mpa hexamer.

Thus, the OB ring as the core of Mpa serves multiple functions: the OB ring is a main factor that holds the Mpa hexamer together, provides the initial binding site for the substrate in conjunction with the coiled coil domain, is an entrance point into the ATPase central channel, and is a rigid platform against which the protein substrate can be pulled apart for translocation and ultimate degradation in the proteasome core. Our in vitro and in vivo mutagenesis studies support these functions.

Mpa-Facilitated Protein Degradation by the Mtb Proteasome

The system for protein tagging and degradation in Mtb consists at least of Pup-Mpa-20SCP, the organization of which is reminiscent of the prokaryotic SsrA-ClpX-ClpP and eukaryotic ubiquitin-19S-20SCP systems. Each uses a protease complex that requires an oligomeric ATPase for unfolding and translocating the protein substrate into the protease core for degradation. In the ClpXP system, the protein substrate unfolding and translocation involves the Ar- Φ pore loop of the ClpX ATPase (Martin et al., 2008a, 2008b). We found that the conserved Ar- Φ pore loop

in Mpa is also essential for substrate degradation in Mtb (Figure 5B), suggesting that Mpa might share a similar protein translocation mechanism.

In summary, our demonstration of the physical interaction between Mpa and the Mtb proteasome and the crystallographic revelation of the OB fold structure of Mpa-ID, in conjunction with our previous genetic and functional evidence (Darwin et al., 2003, 2005; Pearce et al., 2006), firmly establish Mpa as the bona fide Mtb proteasomal ATPase. Because the virulence of Mtb relies on proteasome function (Darwin et al., 2003, 2005; Gandotra et al., 2007; Pearce et al., 2006), and the function of the Mtb proteasome apparently depends on Mpa, it is possible to take advantage of the unique structure to design inhibitors for Mpa. Such an approach potentially circumvents the difficulty in developing species-specific proteasome inhibitors (Lin et al., 2008), given the high conservation of proteasome structure across all domains of life.

EXPERIMENTAL PROCEDURES

Purification of Mpa

Purification of the full-length Mpa was as described previously (Darwin et al., 2005).

Purification of Mpa-ID and Mpa_N97 Δ and the Triple Mutant Forms of Mpa and Mpa-ID

For all cloning, Pfu Ultra (Stratagene) was used for the polymerase chain reactions and the amplified products were cloned into pET24b(+) (Novagen) using enzymes from New England Biolabs. The inserts were sequenced by GENEWIZ. Plasmids were transformed into DH5a (GIBCO-BRL) and ER2566 (Chong and Garcia, 1994) for expression. Bacteria producing various Mpa constructs were grown in Luria-Bertani medium. Gene expression was induced by IPTG overnight at 20°C. The supernatant of the lysed bacteria was purified by a 5 ml QIAGEN Ni²⁺ affinity column. The imidazole eluate was buffer exchanged to 1 \times PBS and treated by thrombin for His tag removal. A second Ni²⁺ affinity chromatography was performed and the flow-through was collected and further purified by the Resource Q ion exchange column and the Superdex-200 gel filtration column. The purified proteins were dialyzed in buffer (10 mM Tris [pH 8.0], 10 mM NaCl, and 1 mM β -ME) and concentrated with a Vivaspin tube (WMC0; 10 kD). The blue native gel (~4–15% gradient Ready Gel) was purchased from Bio-Rad, run with the standard Tris-HCl Gel method, and stained by Coomassie blue. See Supplemental Experimental Procedures for details.

Mtb Wild-Type 20S Proteasome and the 20SOG Mutant Expression and Purification

The wild-type and the 20SOG were expressed and purified as described previously (Lin et al., 2006).

Electron Microscopy of Mpa-20SOG Complex

The purified Mpa ATPases and proteasome samples were mixed at a molar ratio of 4:1 in a buffer containing ATP γ S, incubated on ice briefly, and then used directly for negative stain EM grids preparation. Microscopy was carried out in a JEOL-1200EX TEM. See Supplemental Experimental Procedures.

Three-Dimensional Crystallization of Mpa-ID and Diffraction Data Collection

The protein was prone to aggregation, so the detergent cyclohexylpropanoyl-N-hydroxyethylglucamide (C-HEGA-9) was essential for crystallization. The hanging drop vapor diffusion method was used for crystallization of Mpa-ID. In brief, a 2 μ l droplet of Mpa-ID at 10.0 mg/ml in 10 mM Tris (pH 8.0), 10 mM NaCl, and 1 mM β -ME was mixed with 2 μ l of reservoir solution consisting of 0.1 M MES (pH 6.0), 50 mM MgCl₂, 50 mM NaCl, 20% PEG-2000, and 108 mM C-HEGA-9. We screened an array of heavy atom derivatives by native gel electrophoresis of the purified Mpa-ID and found that mercury acetate

(Hg(CH₃CO₂)₂) induced a clear upward shift of the protein on the native gel, so Hg(CH₃CO₂)₂ was chosen for derivatizing the crystals. The cryoprotectants used for flash freezing heavy atom derivatized crystals were 10% (v/v) glycerol and 10% (v/v) ethylene glycerol. Crystal screening and data set collection were carried out at X25 and X29 beam lines at National Synchrotron Light Source, Brookhaven National Laboratory. See [Supplemental Experimental Procedures](#) for more details.

Data Reduction and Structure Determination

See [Supplemental Experimental Procedures](#).

Mpa Mutant Analysis in Mtb

Mutations were introduced into pMV-mpa using overlap extension ("sewing") PCR or QuikChange Multi Site-Directed Mutagenesis kit (Stratagene). Plasmids were introduced into *mpa* null mutant Mtb as described elsewhere (Darwin et al., 2003). Equivalent Mtb cell numbers, determined by culture OD₅₈₀, were harvested for each strain and washed in 0.05% Tween-80 in PBS. Bacteria were resuspended in 250 μl of lysis buffer (100 mM Tris-Cl, 100 mM KCl, 1 mM EDTA, and 5 mM MgCl₂ [pH 8]) and transferred to bead beating tubes with 100–200 μl of zirconia silica beads (BioSpec Products, Inc). Each tube was bead beaten in a BioSpec Mini Bead Beater for 30 s, three times (tubes were chilled on ice between beatings). Lysates were not filtered due to significant loss of Mpa from the samples during filtration. Samples were boiled for 5 min prior to separation by 10% SDS-PAGE. For loading controls, immunoblots were stripped and incubated with anti-DlaT. SuperSignal West Pico or Femto Chemiluminescent Substrate (ThermoScientific) was used for detection.

ACCESSION NUMBERS

The Mpa-ID X-ray coordinates have been deposited in the Protein Data Bank (accession code 3FP9).

SUPPLEMENTAL DATA

Supplemental Data include Supplemental Experimental Procedures, Supplemental References, and four figures and can be found with this article online at [http://www.cell.com/structure/supplemental/S0969-2126\(09\)00333-5](http://www.cell.com/structure/supplemental/S0969-2126(09)00333-5).

ACKNOWLEDGMENTS

We thank A. Héroux and S. Eswaramoorthy for their expert advice on X-ray crystallography. We thank the staff of Beam lines X25 and X29 at the National Synchrotron Light Source, Brookhaven National Laboratory, for technical assistance in data collection. This work is supported by National Institutes of Health (NIH) grant AI070285 and Brookhaven National Laboratory LDRD grant 06-60 to H.L. and by NIH grants AI065437 and HL092774 to K.H.D. and AI064768 to C.N.

Received: August 18, 2009

Revised: August 18, 2009

Accepted: August 21, 2009

Published: October 13, 2009

REFERENCES

Alexandrescu, A.T., Gittis, A.G., Abeygunawardana, C., and Shortle, D. (1995). NMR structure of a stable "OB-fold" sub-domain isolated from staphylococcal nuclease. *J. Mol. Biol.* **250**, 134–143.

Alexandrescu, A.T., Jaravine, V.A., Dames, S.A., and Lamour, F.P. (1999). NMR hydrogen exchange of the OB-fold protein LysN as a function of denaturation: the most conserved elements of structure are the most stable to unfolding. *J. Mol. Biol.* **289**, 1041–1054.

Baker, T.A., and Sauer, R.T. (2006). ATP-dependent proteases of bacteria: recognition logic and operating principles. *Trends Biochem. Sci.* **31**, 647–653.

Benaroudj, N., Zwickl, P., Seemuller, E., Baumeister, W., and Goldberg, A.L. (2003). ATP hydrolysis by the proteasome regulatory complex PAN serves multiple functions in protein degradation. *Mol. Cell* **11**, 69–78.

Borissenko, L., and Groll, M. (2007). 20S proteasome and its inhibitors: crystallographic knowledge for drug development. *Chem. Rev.* **107**, 687–717.

Butler, S.M., Festa, R.A., Pearce, M.J., and Darwin, K.H. (2006). Self-compartmentalized bacterial proteases and pathogenesis. *Mol. Microbiol.* **60**, 553–562.

Chong, S., and Garcia, G.A. (1994). A versatile and general prokaryotic expression vector, pLACT7. *Biotechniques* **17**, 686, 688, 690–691.

Dar, A.C., and Sicheri, F. (2002). X-ray crystal structure and functional analysis of vaccinia virus K3L reveals molecular determinants for PKR subversion and substrate recognition. *Mol. Cell* **10**, 295–305.

Darwin, K.H., Ehrt, S., Gutierrez-Ramos, J.C., Weich, N., and Nathan, C.F. (2003). The proteasome of *Mycobacterium tuberculosis* is required for resistance to nitric oxide. *Science* **302**, 1963–1966.

Darwin, K.H., Lin, G., Chen, Z., Li, H., and Nathan, C.F. (2005). Characterization of a *Mycobacterium tuberculosis* proteasomal ATPase homologue. *Mol. Microbiol.* **55**, 561–571.

DeLaBarre, B., Christianson, J.C., Kopito, R.R., and Brunger, A.T. (2006). Central pore residues mediate the p97/VCP activity required for ERAD. *Mol. Cell* **22**, 451–462.

Djuranovic, S., Hartmann, M.D., Habeck, M., Ursinus, A., Zwickl, P., Martin, J., Lupas, A.N., and Zeth, K. (2009). Structure and activity of the N-terminal substrate recognition domains in proteasomal ATPases. *Mol. Cell* **34**, 580–590.

Festa, R.A., Pearce, M.J., and Darwin, K.H. (2007). Characterization of the proteasome accessory factor (paf) operon in *Mycobacterium tuberculosis*. *J. Bacteriol.* **189**, 3044–3050.

Forster, A., Whitby, F.G., and Hill, C.P. (2003). The pore of activated 20S proteasomes has an ordered 7-fold symmetric conformation. *EMBO J.* **22**, 4356–4364.

Gandotra, S., Schnappinger, D., Monteleone, M., Hillen, W., and Ehrt, S. (2007). In vivo gene silencing identifies the *Mycobacterium tuberculosis* proteasome as essential for the bacteria to persist in mice. *Nat. Med.* **13**, 1515–1520.

Goldberg, A.L. (2007). Functions of the proteasome: from protein degradation and immune surveillance to cancer therapy. *Biochem. Soc. Trans.* **35**, 12–17.

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.

Hershko, A., and Ciechanover, A. (1998). The ubiquitin system. *Annu. Rev. Biochem.* **67**, 425–479.

Hu, G., Lin, G., Wang, M., Dick, L., Xu, R.M., Nathan, C., and Li, H. (2006). Structure of the *Mycobacterium tuberculosis* proteasome and mechanism of inhibition by a peptidyl boronate. *Mol. Microbiol.* **59**, 1417–1428.

Jaravine, V.A., Rathgeb-Szabo, K., and Alexandrescu, A.T. (2000). Microscopic stability of cold shock protein A examined by NMR native state hydrogen exchange as a function of urea and trimethylamine N-oxide. *Protein Sci.* **9**, 290–301.

Kloks, C.P., Spronk, C.A., Lasonder, E., Hoffmann, A., Vuister, G.W., Grzesiek, S., and Hilbers, C.W. (2002). The solution structure and DNA-binding properties of the cold-shock domain of the human Y-box protein YB-1. *J. Mol. Biol.* **316**, 317–326.

Lin, G., Hu, G., Tsu, C., Kunes, Y.Z., Li, H., Dick, L., Parsons, T., Li, P., Chen, Z., Zwickl, P., et al. (2006). *Mycobacterium tuberculosis* prcBA genes encode a gated proteasome with broad oligopeptide specificity. *Mol. Microbiol.* **59**, 1405–1416.

Lin, G., Tsu, C., Dick, L., Zhou, X.K., and Nathan, C. (2008). Distinct specificities of *Mycobacterium tuberculosis* and mammalian proteasomes for N-acetyl tripeptide substrates. *J. Biol. Chem.* **283**, 34423–34431.

Martin, A., Baker, T.A., and Sauer, R.T. (2008a). Diverse pore loops of the AAA+ ClpX machine mediate unassisted and adaptor-dependent recognition of ssrA-tagged substrates. *Mol. Cell* **29**, 441–450.

- Martin, A., Baker, T.A., and Sauer, R.T. (2008b). Pore loops of the AAA+ ClpX machine grip substrates to drive translocation and unfolding. *Nat. Struct. Mol. Biol.* *15*, 1147–1151.
- Medalia, N., Beer, A., Zwickl, P., Mihalache, O., Beck, M., Medalia, O., and Navon, A. (2009). Architecture and molecular mechanism of PAN, the archaeal proteasome regulatory ATPase. *J. Biol. Chem.* *284*, 22952–22960.
- Murzin, A.G. (1993). OB(oligonucleotide/oligosaccharide binding)-fold: common structural and functional solution for non-homologous sequences. *EMBO J.* *12*, 861–867.
- Nagy, I., Tamura, T., Vanderleyden, J., Baumeister, W., and De Mot, R. (1998). The 20S proteasome of *Streptomyces coelicolor*. *J. Bacteriol.* *180*, 5448–5453.
- Nathan, C., Gold, B., Lin, G., Stegman, M., de Carvalho, L.P., Vandal, O., Venugopal, A., and Bryk, R. (2008). A philosophy of anti-infectives as a guide in the search for new drugs for tuberculosis. *Tuberculosis (Edinb.)* *88 (Suppl 1)*, S25–S33.
- Pearce, M.J., Arora, P., Festa, R.A., Butler-Wu, S.M., Gokhale, R.S., and Darwin, K.H. (2006). Identification of substrates of the *Mycobacterium tuberculosis* proteasome. *EMBO J.* *25*, 5423–5432.
- Pearce, M.J., Mintseris, J., Ferreyra, J., Gygi, S.P., and Darwin, K.H. (2008). Ubiquitin-like protein involved in the proteasome pathway of *Mycobacterium tuberculosis*. *Science* *322*, 1104–1107.
- Schmidt, M., Hanna, J., Elsasser, S., and Finley, D. (2005). Proteasome-associated proteins: regulation of a proteolytic machine. *Biol. Chem.* *386*, 725–737.
- Smith, D.M., Kafri, G., Cheng, Y., Ng, D., Walz, T., and Goldberg, A.L. (2005). ATP binding to PAN or the 26S ATPases causes association with the 20S proteasome, gate opening, and translocation of unfolded polypeptides. *Mol. Cell* *20*, 687–698.
- Smith, D.M., Benaroudj, N., and Goldberg, A. (2006). Proteasomes and their associated ATPases: a destructive combination. *J. Struct. Biol.* *156*, 72–83.
- Smith, D.M., Chang, S.C., Park, S., Finley, D., Cheng, Y., and Goldberg, A.L. (2007). Docking of the proteasomal ATPases' carboxyl termini in the 20S proteasome's alpha ring opens the gate for substrate entry. *Mol. Cell* *27*, 731–744.
- Striebel, F., Imkamp, F., Sutter, M., Steiner, M., Mamedov, A., and Weber-Ban, E. (2009). Bacterial ubiquitin-like modifier Pup is deamidated and conjugated to substrates by distinct but homologous enzymes. *Nat. Struct. Mol. Biol.* *16*, 647–651.
- Voges, D., Zwickl, P., and Baumeister, W. (1999). The 26S proteasome: a molecular machine designed for controlled proteolysis. *Annu. Rev. Biochem.* *68*, 1015–1068.
- Wolf, S., Nagy, I., Lupas, A., Pfeifer, G., Cejka, Z., Muller, S.A., Engel, A., De Mot, R., and Baumeister, W. (1998). Characterization of ARC, a divergent member of the AAA ATPase family from *Rhodococcus erythropolis*. *J. Mol. Biol.* *277*, 13–25.
- Zhang, C., and Kim, S.H. (2000). A comprehensive analysis of the Greek key motifs in protein beta-barrels and beta-sandwiches. *Proteins* *40*, 409–419.
- Zhang, F., Hu, M., Tian, G., Zhang, P., Finley, D., Jeffrey, P.D., and Shi, Y. (2009). Structural insights into the regulatory particle of the proteasome from *Methanocaldococcus jannaschii*. *Mol. Cell* *34*, 473–484.
- Zhang, X., Stoffels, K., Wurzbacher, S., Schoofs, G., Pfeifer, G., Banerjee, T., Parret, A.H., Baumeister, W., De Mot, R., and Zwickl, P. (2004). The N-terminal coiled coil of the *Rhodococcus erythropolis* ARC AAA ATPase is neither necessary for oligomerization nor nucleotide hydrolysis. *J. Struct. Biol.* *146*, 155–165.