Title
Architecture and assembly of the 19S proteasome regulatory particle

Author
Estrin, Eric

Publication Date
2014

Peer reviewed|Thesis/dissertation

Architecture and assembly of the 19S proteasome regulatory particle

By

Eric Estrin

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Molecular and Cell Biology

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Assistant Professor Andreas Martin, Chair
Professor James Berger
Professor Kathleen Collins
Professor David Wemmer

Spring 2014

Abstract

Architecture and assembly of the 19S proteasome regulatory particle

by

Eric Estrin

Doctor of Philosophy in Molecular and Cell Biology

University of California, Berkeley

Assistant Professor Andreas Martin, Chair

Abstract:
The 26S proteasome is the major ATP-dependent protease in eukaryotic cells. The 19S proteasome regulatory particle, consisting of the lid and base subcomplexes, recognizes and processes poly-ubiquitinated substrates. In this thesis I report the development of heterologous expression systems for the base and lid subcomplexes, and I use these new systems, along with biochemistry and electron microscopy, to map the architecture of the massive 19S regulatory particle, reveal the ordered self-assembly of the lid, and demonstrate structural and functional asymmetries in the AAA+ base unfoldase. I describe the spatial arrangement of ubiquitin receptors, deubiquitinating enzymes, and the unfolding machinery at subnanometer resolution, thereby outlining a substrate's path to degradation, providing the structural basis for the ability of the proteasome to degrade a diverse set of substrates and thus regulate vital cellular processes. Furthermore, I show that a helical bundle serves as a hub through which the last-added subunit, Rpn12, monitors proper lid assembly prior to incorporation into the proteasome. Finally, data presented here demonstrate that the six ATPases play distinct roles in degradation, highlighting how the 26S proteasome deviates from simpler, homomeric AAA+ proteases.

Table of Contents

Acknowledgements

There have been so many people that have made the last five years at Berkeley a unique and fulfilling experience. I would like to first thank the Martin lab – all of whom I have been tremendously lucky to work alongside and call my friends. The lab's powerful combination of intellect, motivation, good attitude, and sense of humor is rare and I deeply cherished my time in lab. In particular, those first few years of the lab starting out were a simultaneously exciting and draining time, and I want to especially recognize Charlene Bashore, Robyn Beckwith, Mary Matyskiela, Chris Padovani, and Kris Nyquist for their support through the early years in the lab. I also enjoyed working with Gabe Lander; I couldn't have had a more devoted and insightful EM collaborator, with a great sense of humor to boot. Also thanks to collaborators Pablo Chacon and José Ramón López-Blanco in developing a structural model of the helical bundle.

I also am indebted to Andy Martin. I can say without hesitation that he is one of the smartest people I know, and has taught me so many lessons, more than I can do justice to in this space. I cannot wait to see the amazing things he will continue to accomplish. Thanks to my committee members James Berger, Kathy Collins (her advice to forgo the vision quest was invaluable to Alex Muir and I), and Dave Wemmer for their guidance during my thesis work. I would also like to thank Susan Marqusee for her support throughout my grad school experience; I am inspired by her enthusiasm for teaching and her never-ending stream of invaluable guidance. My experience at Berkeley was also enriched by many lessons taught so long ago by Tom Alber and Jaime Fraser. I will never forget Tom's passion for science, best exemplified to me one evening as a first-year when we spent an hour ogling at a recently published ultra-resolution structure of lysozyme.

The last five years have been made enjoyable in part by friends who have kept me grounded and whom I have gone on so many adventures with. To name a few: Alex Muir, Soroush Khanlou, Damian Trujillo, Ellen Goodall (how lucky I am to have a labmate and climbing friend combined!), Jon McMurray, Dan Richter, Alec Sexton, Jacob Corn, Avi Flamholz, and members of the Ellsworth House all contributed to my sanity during the past five years, as well as reminded me of and introduced me to some of my passions outside of science. I owe special thanks to Carmen Olson who, through her kindness and generosity, supported me while I wasa completing much of the work presented here.

Finally, I would like to thank my family, only with whom I would have made it to and thrived at Berkeley. In particular I want to thank my Mom, Dad, grandparents, and Chase. Though it would be impossible for me to repay you for the support, love, and joy you have given me the past 27 years, I offer the following work as thanks to you.

Chapter 1: Introduction

Ubiquitin-Proteasome history and function

Proteins are the major basic functional building blocks in the cell, with thousands of unique proteins acting in a diverse range of processes, from catalyzing metabolic reactions or transducing signals across cell membranes, to serving as scaffolds for maintaining cellular shape. Maintaining precise protein concentrations and managing protein aggregation, together known as proteostasis, is essential in virtually all cellular processes. Protein levels are controlled through regulated synthesis and degradation. In the 1950s, our molecular understanding of these two processes began with the discovery of the ribosome as the site of protein synthesis and the lysosome as a major mediator of protein degradation (Palade, 1955). These discoveries occurred concomitantly with the emerging use of differential centrifugation. The lysosome, which houses a number of proteases, degrades proteins in a largely non-selective manner by fusing with vesicles and degrading their contents. This process results in the degradation of long-lived cytoplasmic proteins and, importantly, does not require energy (De Duve et al., 1955).

Curiously, reports of an ATP requirement for the degradation of several different proteins in a number of tissues emerged in the 1950s. A breakthrough in the mechanistic basis for these observations came in the 1970s when a number of seminal studies uncovered that an energy-dependent process, distinct from lysosomal degradation, required a number of step-wise reactions for specific degradation of cellular proteins (Etlinger and Goldberg, 1977; Hershko et al., 1979). This pathway for degradation turned out to be the ubiquitin-proteasome system, which serves as the eukaryotic mechanism for rapid, selective degradation of damaged, misfolded, and short-lived proteins.

Substrates destined for the proteasome are often regulatory proteins, such as the cyclin family, whose degradation ushers in new phases of the cell cycle. Some proteasome substrates are only partially degraded, yielding biologically active protein fragments. An example for such proteasomal activation is the NFκB-related transcription factor Spt23, which is anchored to the ER membrane and only free to diffuse into the nucleus where it can activate transcription of target genes upon partial proteolysis by the proteasome (Piwko and Jentsch, 2006). Additionally, the importance of regulated protein degradation is highlighted by its role in the breakdown of misfolded proteins, which in some cases can include up to 30% of newly synthesized proteins. Because of its critical role in diverse cellular processes such as the cell cycle and protein-quality control, the ubiquitin-proteasome system is seen as a promising therapeutic target for a number of diseases ranging from cancer to Alzheimer's disease. Indeed, Bortezomib, a proteasome inhibitor developed in the 1990s, is currently used to treat multiple myeloma (Richardson et al., 2003).

Covalent modification with a poly-ubiquitin chain targets proteins for ATP-dependent destruction by the 26S proteasome. Substrate tagging is executed by a multi-step process, beginning with activation of the 7kDa ubiquitin protein by covalent attachment to a so-called 'E1' enzyme (Ciechanover, 1994; Ciechanover et al., 1982). The activated

ubiquitin moiety is then transferred to a second enzyme, an E2, whereupon a third enzyme, known as an E3, recognizes both a substrate protein and the E2-linked ubiquitin to facilitate transfer of the ubiquitin to a lysine residue found on the substrate. The E2 and E3 enzymes are responsible for substrate specificity and the selection of the appropriate lysine for ubiquitination. Repetition of this process onto a previously-linked ubiquitin of a condemned substrate creates a chain of ubiquitin moieties. The ubiquitin chain formed can be linked through any one of the seven lysines on ubiquitin, with K48-, K11-, and potentially K63-linked chains thought to target substrates to the proteasome for degradation, and other chains leading to different consequences for the tagged protein (Mattiroli and Sixma, 2014).

In addition to containing the right linkage type for recognition by a proteasomal receptor, the ubiquitin chain on a substrate must also be at least four ubiquitin moieties long and the substrate must contain an unfolded segment of at least 35 amino acids to be efficiently degraded by the proteasome(Gregori et al., 1990). However, these rules may not be as strict as once thought: an *in vitro* study has shown that proteins can be degraded with less than four ubiquitin moieties attached, given that the unstructured segment of the substrate is more than the minimum of 35 amino acids (Inobe et al., 2011). Prior to the work presented here, the mechanistic and structural basis of these substrate requirements for degradation by the proteasome was not understood.

Proteasome structure and function

The 26S proteasome is a massive 1.5 MDa proteolytic machine composed of over 32 unique, essential proteins, many in multiple copies(Driscoll and Goldberg, 1990). The 26S proteasome can be stably and functionally separated into the 20S core particle (CP) and the 19S regulatory particle (RP) (Glickman et al., 1998a). The core particle is composed of two copies of 14 different subunits (α1-7 and β1-7), forming a barrel-shaped structure with two α-rings sandwiching two β-rings in a αββα topology, whose internal β-ring chamber houses the peptidase active sites. Three β subunits (β1, β2, and β5) contain active proteolytic sites of the N-terminal hydrolase superfamily, with each having mild specificity towards acidic, basic, and hydrophobic residues, respectively. Proteins do not get fully digested into their constituent amino acids when degraded by the proteasome, but rather into peptides of approximately 8-10 amino acids in length are thought to diffuse out of the CP chamber(Groll et al., 1997). The CP active sites exist at an extremely high local concentration due to the small volume within the barrel-shaped compartment, ensuring that any substrate that has been targeted to this space will be proteolyzed.

A gate formed by a meshwork of N-termini from the α subunits blocks access to the interior of the CP, providing essential protection of the cytoplasm from the 26S proteasome(Groll et al., 1997). Opening of the 'α-gate' is accomplished by the binding of the regulatory particle, which, as its name suggests, acts as a regulator of substrate entry(Groll et al., 2000). The regulatory particle itself can be separated into two subcomplexes, the lid and base, the domain architecture of which is summarized in Fig 1.1 (Glickman et al., 1998a). The base contains a heterohexameric ring of AAA+

(*A*TPases *a*ssociated with cellular *a*ctivities) ATPases (Rpt1-6, in the order Rpt1-2-6-3-4-5) that binds to the α-ring of the CP (Ahn et al., 1996). The interaction between the RP and the CP is thought to be primarily mediated by C-terminal tails of the Rpt subunits docking into pockets that are present in between each α-α subunit interface, leading to the opening of the aforementioned CP gate (Kohler et al., 2001). In addition to containing a C-terminal tail and an AAA+ domain, each ATPase subunit contains an OB-fold located in between an N-terminal helix and the AAA+ domain(Zhang et al., 2009b). The six OB-folds form a second ring on top of the AAA+ ring, and the N-terminal alpha helix pairs with a neighboring ATPase to form a coiled-coil dimer. Along with opening the CP gate, the ring of AAA+ ATPases play a crucial role in substrate processing: after engaging a substrate that is tethered to the proteasome through interactions of its ubiquitin chain with one of the proteasomal ubiquitin receptors, they unravel folded structures and translocate the polypeptide into the destructive CP chamber(Navon and Goldberg, 2001).

Much of what is known about the ATPases is derived from information on related, though less complex, homohexameric AAAs: the Proteasome-Activating Nuclease (PAN) found in archaea and ClpX found in bacteria. Previous studies on the ClpX homomeric machine have suggested that ATP hydrolysis occurs in one subunit at a time with a degree of coordination that allows subunits to contribute additively and equally to substrate processing (Martin et al., 2005). However, the homomeric nature of ClpX hinders an assessment of the precise coordination between subunits and whether all six subunits sequentially progress through the different stages of the ATP hydrolysis cycle. Previous work on various AAA+ unfoldases proposed that a conserved aromatic-hydrophobic (Ar-Φ) loop protrudes from every ATPase subunit into the central channel and undergoes nucleotide-dependent power strokes to drive substrate translocation(Martin et al., 2008a, b; Wang et al., 2001).

One particular structure that has been studied in ClpX are the IGF loops, which are functionally equivalent to the C-terminal tails of HslU (another bacterial AAA+ unfoldase) (Joshi et al., 2004). The IGF loops are thought to dock into the hydrophobic pockets of ClpP and induce a conformational change that raises the N-terminal loops termed the ‘pore-2 loops’ of ClpP towards loops in the pore of ClpX (Martin et al., 2007). Communication between these loops facilitate many phenomena, including the observed repression of ATPase rate of ClpX upon binding to ClpP. This communication can explain the observed increased unfolding ability of ClpXP, as a slower ATPase rate will allow the substrate to respond to unfolding force applied by the motor. The pore-2 loop directly proceeds the Walker B motif that plays the catalytic role in ATP hydrolysis, which may explain how a conformational change in the pore-2 loop effects the active-site geometry. Additionally, the pore-2 loop/N-terminal loop interaction is important, but not necessary, for binding of ClpX to ClpP and is thought to stabilize the open gate conformation of ClpP. Finally, the pore-2 loops are involved in generating force to unfold substrates as well as binding certain degradation tags. These loop interactions may further form a continuous channel to guide substrate towards the peptidase chamber as well as introduce a potential reduction of substrate ‘slippage’ due to interactions of substrate within the channel.

Claims that the C-terminal tails of the base are the only interacting partners between the base and 20S have precluded any search for activation loops in the base, though there is a known activation loop in another 20S CP binding partner, PA26 (Forster et al., 2005). A conserved pore-2 loop in the proteasomal AAA+s suggests that this loop may play an important functional role. For instance, it is located close to the N-terminal gates of the α-subunits and may play a role in stabilizing an open gate conformation of the core particle.

The base also contains the two known intrinsic ubiquitin receptors on the proteasome, Rpn10 and Rpn13, as well as two large (100+ kDa) non-ATPase subunits (Rpn1 and Rpn2) that are thought to be important for acting as a scaffold for the regulatory particle (Hanna et al., 2007). Both Rpn1 and Rpn2 contain HEAT-like sequence repeats that form helix-turn-helix structures (He et al., 2012). Both proteins have been suggested to form large alpha-helical toroid structures, and though this has been shown for Rpn2, direct evidence for Rpn1's ternary structure is limited to low resolution electron microscopy data (Effantin et al., 2009; He et al., 2012). Rpn1 is thought to serve as a binding platform for a number of transiently-associated ubiquitin 'shuttle' receptors, Ddi1, Dsk2, and Rad23 (Gomez et al., 2011). Additionally, Rpn1 binds Ubp6, a deubiquitinating enzyme (DUB) that is thought to 'trim' excessively long ubiquitin chains from proteasome-bound substrates and also delays substrate proteolysis by an unknown, catalytically-independent mechanism (Hanna et al., 2006).

The lid consists of nine non-ATPase proteins (Rpn3, 5-9, 11-12, and Sem1 in yeast), six of which (Rpn3,5,6,7,9,12) contain PCI domains and two of which contain MPN domains (Rpn8,11) (Glickman et al., 1998a). The PCI domain is composed of a N-terminal TPR-like (Tetracopeptide repeat) module followed by a winged-helix domain and is found in two other complexes, the eIF3 translation initiation factor and the COP9 signalosome (Hofmann and Bucher, 1998), while the MPN domain is found in the same three complexes and often has protease activity (Maytal-Kivity et al., 2002). Rpn11 is the only known enzyme in the lid and the best characterized of all lid subunits. It belongs to the JAMM/MPN+ family of zinc-dependent isopeptidases and functions in cleaving a substrate's ubiquitin chain *en bloc* at some point after substrate engagement to allow for ubiquitin recycling (Verma et al., 2002; Yao and Cohen, 2002). The second MPN-domain containing protein, Rpn8, lacks a intact Zn-binding site, though is thought to dimerize with Rpn11 and play a scaffolding role (Sanches et al., 2007). The PCI domain-containing subunits in the lid are characterized by a winged-helix domain sandwiched between an N-terminal TPR-like repeat domain and a short, C-terminal helical segment of unknown function. The 6PCI:2MPN domain architecture of the lid is also observed in two other protein complexes, the translation initiation factor eIF3 and the COP9 Signalosome (CSN), a cullin-RING ligase regulator, suggesting that these complexes diverged at one point from a common ancestor (Wei et al., 1998). Finally, Sem1, a 10kDa protein is present in the lid, though, as with most other lid subunits, its function and precise location is not known (Sone et al., 2004).

Proteasome holoenzyme and subcomplex assembly

As just described, the 26S proteasome is a large, complex proteinacious machine that can be separated into multiple distinct subcomplexes. The assembly of such a sizable and intricate complex poses a number of challenges. Large protein complexes often contain hetero-oligomeric assemblies of homologous subunits that in many cases are thought to have evolved by duplication and diversification from an ancient precursor within a homo-oligomer (Pereira-Leal et al., 2007). The proteasome is composed of a number of subassemblies that each harbor highly homologous subunits. For instance, the seven α and seven β subunits, which form the outer and inner rings of the core peptidase, exhibit high homology (Hughes, 1997). Similarly, the AAA+ unfoldase in the base is composed of a heterohexameric assembly of the highly homologous ATPase subunits Rpt1-6, and the lid structure is dominated by an assembly of six distinct, but related, PCI-domains (Serino and Pick, 2013). Though this diversification may allow for the specialization of individual subunits, it also might lead to more complicated assembly processes and a tendency of subunits to occupy the wrong position within a complex due to strong homology with other constituents. Subunit misassembly could lead to inactive complexes, or worse, to complexes that poison a cell. This burden may be especially high considering that one misplaced subunit has the potential of turning an otherwise useful assembly of numerous proteins into an inactive complex.

To prevent these assembly issues, large protein complexes often utilize a number of design principles to ensure their accurate and efficient maturation. In some cases, complex-specific assembly factors favor the association between appropriate partners. For example, proteasome-specific assembly factors ensure the proper arrangement of the α-subunits within the core particle (Kusmierczyk et al., 2008) and the correct order of Rpt subunits in the heterohexameric AAA+ ring of the base (Funakoshi et al., 2009). Though the AAA+ ring requires the presence of these assembly factors, it is still unclear whether these are sufficient for efficient assembly.

In addition to this strategy for correct subunit arrangement, many complexes rely on mechanisms that inhibit premature catalytic activity of assembly intermediates to prevent futile or improperly timed substrate processing. The proteasome utilizes such mechanisms to suppress its proteolytic activity until full maturation. The three proteolytically active β-subunits of the core particle are inhibited by N-terminal propeptides until the complete assembly of the core particle triggers the autocatalytic removal of these sequences. This process ensures that proteolysis is only activated once the catalytic sites are sequestered in the core particle internal chamber (Arendt and Hochstrasser, 1999; Chen and Hochstrasser, 1996). Another strategy is to strictly control substrate access to these active sites. In the proteasome core particle, the N-termini of the α-subunits form a gate that excludes folded proteins and large unfolded polypeptides from the proteolytic chamber (Groll et al., 2000). This gate is opened only upon binding of a properly assembled base that recognizes appropriate substrates and actively translocates them into the core (Smith et al., 2007). Moreover, the deubiquitinase activity of Rpn11 is inhibited until the lid is incorporated into the regulatory particle, though the mechanistic basis for this repression is unknown (Verma et al., 2002).

Large complexes are often assembled in parts, with individual subunits initially forming smaller subcomplexes that are combined into the holoenzyme only upon their proper completion (Sauer and Baker, 2011; Staley and Woolford, 2009). This strategy may be advantageous, as smaller pieces likely assemble more easily and controlled than one massive construction. Such modularity may also allow the quick regulation of complex levels, the ability to use interchangeable, functionally distinct parts, and the easier transport between organelles. The 26S proteasome seems to adhere to this strategy, as is appears to assemble from preformed subcomplexes (Murata et al., 2009; Tomko Jr and Hochstrasser, 2013), which for instance allows their separate import into the nucleus (Isono et al., 2007).

Outstanding Questions

A number of key mechanistic and structural details remain unresolved in regards to how the proteasome functions. For example, although the ATPase ring is known to bind atop the core, detailed information about the location of the lid, the essential deubiquitinase Rpn11, the non-ATPases in the base, and the two intrinsic ubiquitin receptor within the proteasome has remained elusive. Furthermore, the organization of the PCI/MPN-domain containing subunits of the lid remains unknown, as does how these subunits assemble without the aid of assembly factors. Uncovering the organization and assembly of the lid subcomplex would provide critical insight not only for proteasome function but also for such disparate cellular functions as translation initiation and regulation of E3 ligases, since the eIF3 and COP9 signalosome complexes each contain a similar domain architecture as the lid. Additionally, in part because of the paucity of structural information, we have only an at best hazy understanding of what role most regulatory particle subunits play in proteasome function. The arrangement of lid subunits within the lid and the holoenzyme may for instance provide crucial insight into how Rpn11 exhibits robust DUB activity only when it is incorporated into the holoenzyme. Also, despite years of study, a thorough understanding of subunit specialization and coordination within the base AAA+ ATPase ring has not been achieved. Furthermore, the unique heterohexameric architecture of the proteasomal ATPase ring has raised the fundamental question of whether the six Rpt subunits are functionally equivalent or have distinct roles in ATP hydrolysis and substrate processing. Base mutagenesis studies of complexes purified from yeast has revealed some insight into the mechanism of communication and subunit specialization, yet they have all been complicated by issues related to assembly and production of mutated proteasome particles from yeast.

Taken together, the lack of structural and functional information on the proteasome has limited our understanding of the mechanistic basis for substrate processing, such as the path a substrate takes during its final steps before destruction and how these steps are coordinated. Progress in answering these outstanding questions has been stifled by problems resulting from proteasome preparations from native sources. Proteasome holoenzyme and subcomplexes purified from these sources have two major issues: 1. Due to the importance of proteasome genes for cell viability, many manipulations of these

genes lead to cell death. 2. Proteasome preparations are heterogeneous, often containing sub-stoichiometric substrates and/or proteasome-interacting proteins.

Here, I present an approach for the heterologous production of the base and the lid, which, combined with structural and functional biochemical studies, has led to critical insight into these outstanding questions. Chapter 2 outlines the utilization of the novel heterologous expression system of the lid in combination with cryo-EM to map the entire architecture of the 26S proteasome. In Chapter 3 I describe the structure of an intricate helical bundle that is formed from helices present in each lid subunit and its role in dictating the step-wise assembly of the proteasome lid. Finally, in Chapter 4 I present the development of heterologous expression systems for the proteasome base in E. *coli*, and the use of this system to shed light on the distinct roles of the six AAA+ subunits in the base heterohexameric ring.

Figures

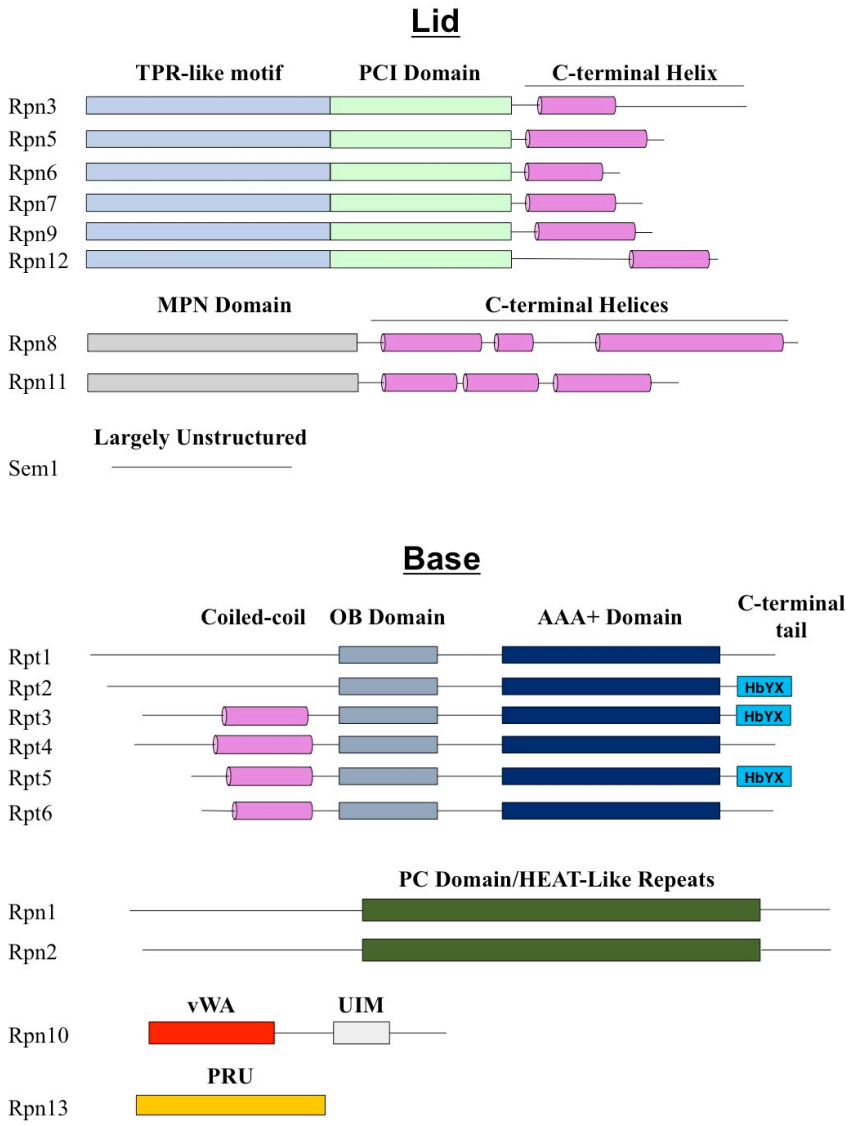


Fig. 1.1: Domain architecture of proteasome regulatory particle subunits. Helices depict predicted coiled-coils and boxes depict folded domains. All regions that are unstructured or form an unknown structure are depicted as a line.

Chapter 2: Complete subunit architecture of the proteasome regulatory particle

This chapter is based on work published in Lander et al, Nature, 2012. I was responsible for constructing the heterologous-expression system for the lid and all biochemistry using the recombinant system, Gabe Lander was responsible for the electron microscopy, Mary Matyskiela was responsible for the tagging and purification of complexes from yeast, and Charlene Bashore was responsible for atomic emission spectroscopy.

Introduction

While the proteolytic core has been well studied, there is only limited structural characterization of the regulatory particle (Bohn et al., 2010; da Fonseca and Morris, 2008; Forster et al., 2009; Nickell et al., 2009). None of the 13 non-ATPase subunits, including the ubiquitin receptors and deubiquitinating enzymes, have been localized within this assembly. It has been shown that efficient degradation depends on the length, linkage type, and placement of an ubiquitin chain, as well as the presence of an unstructured initiation site on a substrate (Inobe et al., 2011; Thrower et al., 2000; Xu et al., 2009), but the topological information needed to explain these requirements is missing. Furthermore, though they have been suggested to play a scaffolding role (Scheel, BMC Bioinformatics, 2005), the 6 PCI-containing subunits' precise function is not known, in part due to a paucity of structural information of the PCI domains in the holoenzyme context. Elucidating the architecture of the regulatory particle and the spatial arrangement of individual subunits is crucial to understanding the molecular mechanisms for substrate recognition and processing.

Here, we present the structure of the holoenzyme and individual subcomplexes solved by EM. A newly developed heterologous expression system for the lid facilitated the localization of all subunits within the regulatory particle, providing a complete architectural picture of the proteasome. The resulting structural understanding offers novel insight into the mechanisms of ubiquitin binding, deubiquitination, substrate unfolding, and translocation by this major eukaryotic proteolytic machine.

Methods

Recombinant lid construction and purification

Yeast Rpn5, Rpn6, Rpn8, Rpn9, and Rpn11-6xHis were cloned into pETDuet-1 (Novagen), yeast Rpn3, FLAG-Rpn7, and Rpn12 were cloned into pCOLADuet-1 (Novagen), and yeast Sem1 and Hsp90 were cloned into pACYCDuet-1 (Novagen). A T7 promoter preceded each gene and each plasmid contained a T7 terminator following the multiple cloning site. Genes for select rare tRNAs were included in the pACYCDuet-1 plasmid to account for codon-usage differences between yeast and *E.coli*. MBP tagged constructs contained between 2-6 amino acid linking segments between the fusion tag and lid protein. To ensure full-length of Rpn6 in lid particles used for biochemical experiments and the negative stain reconstruction of recombinant lid, we used a construct

with the FLAG tag moved from Rpn7 to Rpn6. *E. coli* BL21-star (DE3) cells were co-transformed with the three plasmids mentioned above. Lid proteins and the chaperone Hsp90 were coexpressed overnight at 18 °C after inducing cells with 0.5 mM isopropyl-β-D-thiogalactopyranoside at OD_{600} = 0.7. Cells were harvested by centrifugation (4,000 g for 30 min), resuspended in FLAG buffer (50 mM HEPES, pH 7.6, 100 mM NaCl, 100 mM KCl and 5 % glycerol) supplemented with protease inhibitors and 2 mg/mL lysozyme, and sonicated on ice for 2 min in 15-s bursts. The lysate was clarified by centrifugation (27,000 g for 30 min), and the complex was affinity-purified on anti-FLAG M2 resin (Sigma-Aldrich) using an N-terminal FLAG-tag on Rpn6 or Rpn7. The protein was concentrated in a 30,000 MWCO concentrator (Amicon) for further purification on a Superose 6 size-exclusion column (GE Healthcare) equilibrated in FLAG buffer. Intact, assembled lid particles eluted at 13.1 mL, similar to lid purified from yeast.
His_6-tagged yeast Rpn10 was expressed in *E. coli* and purified by Ni-NTA affinity and size-exclusion chromatography.

Yeast strain construction. Wild-type holoenzyme was purified from the strain YYS40 (*MAT***a** *ade2-1 his3-11,15 leu2-3,112 trp1-1 ura3-1 can1 RPN11::RPN11-3XFLAG (HIS3))(Sone et al., 2004)*. To generate *RPN10*, *RPN13*, and *UBP6* deletion strains, the kanMX6 sequence was integrated at the respective genomic locus, replacing the gene in YYS40. To generate the strains used to purify GST-Rpn2, GFP-Rpn5 and GFP-Rpn8 holoenzyme, sequences encoding the respective tags under the control of the P_{GAL1} promoter were integrated 5' of the respective genes in YYS40. To generate the strain used to purify Rpn1-FLAG holoenzyme, a sequence encoding the FLAG-tag was integrated 3' to *RPN1* in aW303 background strain (*MAT***a** *ade2-1 his3-11 leu2-3,112 trp1-1 ura3-1 can1-100 bar1*).

To generate the strains used to purify α2 mutant-containing core particle for the crosslinking experiments shown in Figure S11**,** pRS305 (*LEU2*) containing the mutant α2 and the genomic sequences found 500 nucleotides upstream and 100 nucleotides downstream of the gene was integrated at the *LEU2* locus of RJD1144 (*MAT***a***, his3*Δ*200 leu2-3,112 lys2-801 trp*Δ*63 ura3-52 PRE1-FLAG-6xHIS::Ylplac211* (*URA3*)) (Verma et al., 2000) , and the chromosomal copy of α2 was deleted. To generate the strain used to purify lid with Rpn6-3xHA for crosslinking, the 3xHA sequence was integrated 3' of *RPN6* in YYS40.

Endogenous proteasome holoenzyme and subcomplex purification.
Endogenous holoenzyme, core particle (Verma et al., 2000), and lid subcomplex (Leggett et al., 2005) were purified from *S. cerevisiae* essentially as described. The base subcomplex was purified according to protocols for the holoenzyme preparation. Details of yeast strain construction are provided in Table S1.

GFP degradation assay. Proteasome holoenzyme was reconstituted from 20S core, base, Rpn10, and recombinant or endogenous yeast lid in the presence of ATP. A GFP-titin-cyclin fusion protein was modified with a K48-linked poly-ubiquitin chain (Kim and Huibregtse, 2009) and degraded by reconstituted proteasome at 30°C in FLAG buffer

with an ATP-regeneration system (5 mM ATP, 16 mM creatine phosphate, 6 μg ml^{-1} creatine phosphokinase). Degradation was monitored by the loss of fluorescence using a QuantaMaster spectrofluorimeter (PTI).

Protein crosslinking. Sulfo-MBS (Thermo Scientific) is a short (7.3Å), heterobifunctional crosslinker, whose maleimide moiety reacts primarily with sulfhydryls between pH 6.5 and 7.5, and whose NHS ester reacts with primary amines between pH 7 and 9. We purified core particle from yeast strains in which the only copy of the core α2 subunit was either WT, a D245C mutant, or an A249C mutant. Other intrinsic cysteines of the core were found largely non-reactive towards sulfhydryl-modifying agents (not shown). 10 μM reduced core particle purified from strains containing WT, A249C, and D245C α2 was incubated with 150 μM sulfo-MBS for 15 min at pH 6.5, allowing conjugation of the crosslinker to cysteines. Core particle was buffer-exchanged to remove excess crosslinker and increase the pH to 7.5, activating the amine-reactive functional group on sulfo-MBS. This core particle was added at a final concentration of 2 μM to a proteasome reconstitution mixture, containing 2 μM purified base, 10 μM purified Rpn10, 0.5 mM ATP, and 2 μM lid purified from a yeast strain in which Rpn6 was C-terminally tagged with a 3x hemagglutinin (HA) tag. Crosslinking was allowed to proceed for 15 min before reactions were stopped by the addition of 0.5 mM glycine pH 7.5 and divided equally for separation by SDS-PAGE, followed by either coomassie staining or anti-HA western blotting.

Electron Microcopy:

Sample preparation for EM. Negative stain analysis of both the purified proteasome lid and holoenzyme complexes was performed using 400 mesh continuous carbon grids that had been plasma cleaned in a 75% Argon / 25% Oxygen atmosphere for 20 seconds using a Solarus plasma cleaner (Gatan, Inc). Due to the tendency for holoenzyme to adopt a preferential orientation on the carbon substrate, 5 μl of 0.1% poly L-lysine hydrobromide (Polysciences Inc. cat #09730) was placed onto the hydrophilized carbon grids and adsorbed for 90 seconds, washed twice with 5ul drops of water, and allowed to dry completely. This polylysine step was skipped when preparing grids containing the lid samples, as the lid does not adopt a preferred orientation on the carbon substrate. The remaining steps were identical for both holoenzyme and lid. A 4 μl drop of sample at a concentration of 25 μM was placed onto the grid and allowed to adsorb for one minute. The grid was blotted to near-dryness and a 4 μl drop of fresh 2% (w/v) uranyl formate was quickly placed onto the grid. In order to reduce the amount of glycerol remaining on the grids, they were subsequently floated on four successive 25 μl drops of the uranyl formate solution, waiting ten seconds on each drop. The grids were then blotted to dryness.

Preservation of both lid and holoenzyme complexes in vitreous ice was performed in the same manner. 400-mesh C-flats containing 2 μm holes with a spacing of 2 μm (Protochips Inc.) were plasma cleaned in a 75% Argon / 25% oxygen atmosphere for 8 s using a Solarus plasma cleaner (Gatan, Inc). The purified sample, at a concentration of 5 μM in a buffer containing 5% glycerol, was first diluted 1:5 from 60 mM HEPES, pH 7.6, 50 mM NaCl, 50 mM KCl, 5 mM $MgCl_2$, 0.5 mM EDTA, 10 % glycerol, 1 mM DTT, 0.5 mM ATP into a buffer containing 20 mM HEPES, pH 7.6, 50 mM NaCl, 50

mM KCl, 1 mM ATP, 1 mM DTT, and 0.05% NP40, and 4 µl aliquots were placed onto the grids. Grids were immediately loaded into a Vitrobot (FEI company) whose climate chamber had equilibrated to 4 ºC and 100% humidity. The grids were blotted for 3 s at an offset of -1 mm, and plunged into liquid ethane. The frozen grids were transferred to a grid box and stored in liquid nitrogen until retrieved for data collection.

EM data collection. Negative stain analysis of the lid and holoenzyme samples was performed using a Tecnai T12 Bio-TWIN and a Tecnai F20 TWIN transmission electron microscope operating at 120keV. Lid samples were imaged at a nominal magnification of 68,000X (1.57 Å/pixel at the specimen level) on the T12, and 80,000X (1.45 Å/pixel) on the F20. Holoenzyme samples were imaged at a magnification of 49,000X (2.18 Å/pixel) on the T12, and 50,000X (2.16 Å/pixel) on the F20. T12 data were acquired on a F416 CMOS 4Kx4K camera (TVIPS), F20 data were acquired on a Gatan 4Kx4K camera, and all micrographs were collected using an electron dose of 20 $e^-/Å^2$ with a randomly set focus ranging from -0.5 to -1.2 µm. The automatic rastering application of the Leginon data collection software was used for data acquisition. Between 300-500 micrographs were collected for each of the negatively stained datasets.
For cryoEM, individual grids were loaded into a 626 single tilt cryotransfer system (Gatan, Inc) and inserted into a Tecnai F20 TWIN transmission electron microscope operating at 120 keV. Data were acquired at a nominal magnification of 100,000X (1.08 Å/pixel) using an electron dose of 20 $e^-/Å^2$ with a randomly set focus ranging from -1.2 to -2.5 µm. A total of 9,153 micrographs were collected of the holoenzyme using the MSI-T application of the Leginon software. While the holoenzyme was remained intact during the freezing process, the isolated lid specimen became completely disassembled during the freezing process. In an attempt to overcome this, the isolated lid was also frozen using grids onto which a thin carbon film was floated. Due to the elevated background noise from the addition of a carbon substrate, the resulting images lacked the sufficient signal-to-noise ratio necessary to solve a cryoEM structure of the isolated lid to a better resolution than the negative stain structure.

Image processing of negative stain data. All image pre-processing and two-dimensional classification was performed in the Appion image processing environment (Lander et al., 2009). Due to the large number of datasets acquired for both the negatively stained lid and holoenzyme complexes, a generalized schema was utilized for image analysis. This schema also minimized user bias during comparison of tagged and deletion constructs with their wild type counterparts. The contrast transfer function (CTF) of each micrograph was estimated concurrently with data collection using the ACE2 and CTFFind programs (Mallick et al., 2005; Mindell and Grigorieff, 2003), providing a quantitative measurement of the imaging quality. Particle selection was also performed automatically concurrent with data collection. Negatively stained lid particles were selected from the micrographs using a difference of Gaussians (DoG) transform-based automated picker (Voss et al., 2009), and holoenzyme particles were selected using a template-based particle picker. Micrograph phases were corrected using ACE2, and both lid and holoenzyme particles were extracted using a 288x288-pixel box size. The data were then binned by a factor of two for processing. Each particle was normalized to

remove pixels whose values were above or below 4.5 sigma of the mean pixel value using the XMIPP normalization program (Sorzano et al., 2004).
In order to remove aggregation, contamination, or other non-particle selections, particle stacks were decimated by a factor of 2 and subjected to five rounds of iterative multivariate statistical analysis (MSA) and multi-reference alignment (MRA) using the IMAGIC software package (van Heel et al., 1996). 2D class averages depicting properly assembled complexes were manually selected, and the non-decimated particles contributing to these class averages were extracted to create a new stack for further processing. In order to include a larger range of holoenzyme views, particles contributing to doubly capped proteasome averages were removed. This stack of particles went through five rounds of MSA/MRA in IMAGIC (van Heel et al., 1996), and a final correspondence analysis and classification based on Eigen images using the SPIDER software package (Frank et al., 1996) was performed to generate 2D class averages of the complexes.
Initial models for reconstructions of both the holoenzyme and lid were determined using the established "C1 startup" routines in IMAGIC. 2D class averages were manually inspected to select three images representing orthogonal views of the complex, which were in turn used to assign Eulers in a stepwise fashion to the entire dataset of reference-free class averages. The resulting low-resolution models of the lid and holoenzyme were low-pass filtered to 60Å resolution, and these densities were used as starting points for refinement of the 3D structure.
3D reconstructions were all performed using an iterative projection-matching and back-projection approach using libraries from the EMAN2 and SPARX software packages (Hohn et al., 2007; Tang et al., 2007). Refinement of the starting models began using an angular increment of 25º, progressing down to 2º for the lid, and 1º for the holoenzyme. The refinement only continued to the subsequent angular increment once greater than 95% of the particles showed a pixel error of less than 1 pixel. The resolution was estimated by splitting the particle stack into two equally sized datasets, calculating the Fourier shell correlation (FSC) between the resulting back-projected volumes. The estimated resolutions for the final endogenous and recombinant lid structures based on their FSC curves at 0.5 were about 15Å.

Image processing of cryoEM holoenzyme. Processing of the holoenzyme cryo dataset proceeded in a very similar fashion to that of the negatively stained particle datasets. Only ACE2 was used to estimate CTF of the images and measure image quality, and particles were extracted using a box size of 576 pixels. Reference-free 2D classification was performed in order to remove particles that did not contribute to averages depicting a doubly capped proteasome. Three rounds of reference-free 2d classification, and particles were removed after each round. From an initial dataset of 312,483 automatically selected particles, 93,679 were kept for the 3D reconstruction. C2 symmetry was applied to one of the previously determined asymmetric negative stained reconstructions to serve as a starting model for structure refinement. The reconstruction began using an angular increment of 25º, and iterated down to 0.6º. C2 symmetry was imposed during the reconstruction. Low-resolution Fourier amplitudes of the final map were dampened matching the amplitudes of the density map to match those of an experimental GroEL SAXS curve using the SPIDER software package (Frank et al., 1996).

The estimated resolution based on the FSC of the half-volumes at 0.5 was ~9 Å, although a local resolution calculation using the "blocres" function in the Bsoft package (Heymann and Belnap, 2007) indicated a range of resolutions within the density. The majority of the core particle subunits and the AAA+ ATPases were resolved to between 7 and 8 Å resolution, while the non-ATPase subunits in the regulatory particle ranged from 8 to 12 Å resolution (Figure S7). Notably, Rpn1 and the ubiquitin receptors Rpn10 and Rpn13 were the lowest resolution features of the holoenzyme. In order to properly filter the low-resolution portions of the map, without destroying the details of the better-ordered features, a resolution-driven adaptive localized low-pass filter was applied to the final volume (G. Cardone, personal communication).
The segmentation analysis was manually performed using the "Volume Tracer" tool in the UCSF Chimera visualization software (Goddard et al., 2007). This software was additionally used to perform all rigid-body fitting of crystal structures into the holoenzyme cryoEM density, as well as to generate all renderings for figure images.

Availability. The cryoEM density map for the 26S proteasome can be found at the Electron Microscopy Data Bank under accession number EMD-1992. The negative stain reconstructions of the recombinantly expressed and yeast-purified lid have been assigned accession numbers EMD-1993 and EMD-1994, respectively

Results

Recombinant expression of yeast lid in *E. coli*
We developed a system for the heterologous coexpression of all nine lid subunits from *Saccharomyces cerevisiae* in *E. coli.* This system allowed us to generate truncations, deletions, and fusion constructs that were used to localize individual subunits and delineate their boundaries within the lid. The recombinant, purified lid was analyzed in its subunit composition and stoichiometry by SDS PAGE (Fig. 2.1, 2.2) and tandem mass spectrometry (MS). The small, non-essential subunit Sem1 could not be detected, neither for the recombinant nor the endogenous lid that was isolated from yeast. All other subunits were present with the expected stoichiometry, and gel filtration analyses showed indistinguishable elution profiles for the heterologously expressed lid and its endogenous counterpart. Furthermore, atomic emission spectroscopy confirmed that the essential Zn^{2+} ion was incorporated in Rpn11, suggesting proper folding in *E. coli*.
To compare the functionalities of recombinant and endogenous lid, we established conditions for their *in vitro* reconstitution with base and 20S core subcomplexes from yeast to yield 26S holoenzyme. These reassembled particles were assayed for their activity in ubiquitin-dependent substrate degradation by using a poly-ubiquitinated GFP-cyclin fusion protein and following the decrease in GFP fluorescence. Proteasome reconstituted with *E. coli*-expressed lid supported robust substrate degradation (Fig. 2.3). Importantly, the 3D EM reconstructions from negative-stained samples of both lid subcomplexes are practically identical (Fig. 2.4a, 2.5), establishing this recombinant system as an ideal tool for our structural studies of the regulatory particle.

Localization of regulatory particle subunits

As a first step in elucidating the architecture of the regulatory particle, we compared the single-particle EM reconstructions of the yeast holoenzyme and the isolated lid subcomplex obtained at 9 and 15 Å resolution, respectively (Fig. 2.4b, 2.6, 2.7, 2.8). Docking the five-lobed, hand-shaped structure of the lid into the electron density of the holoenzyme revealed the lid's position on one side of the regulatory particle, forming extensive interactions with the base subcomplex, but also contacting the 20S core. The lid subunits Rpn3, 5, 6, 7, 9, and 12 contain a C-terminal PCI (*P*roteasome-*C*SN-e*I*F3) domain that is assumed to play scaffolding roles and allow inter-subunit contacts (Finley, 2009). Our reconstruction provided sufficient resolution to unambiguously locate the winged-helix fold and the flanking TPR-like helical segments of individual PCIs (Fig. 2.1c). The C-terminal PCI domains of the six Rpn subunits thus interact laterally to form a horseshoe-shaped anchor from which the N-terminal domains radially extend. This arrangement demonstrates the scaffolding function of PCI domains in the lid, and we predict that similar interactions underlie the architecture of other PCI-containing complexes.

To determine the subunit topology of the lid, we utilized our heterologous *E. coli* expression system and fused the 40 kDa maltose-binding protein (MBP) to the N- or C-terminus of individual subunits (Fig. 2.1, 2.9a). We then localized the MBP within the tagged lid particles by negative-stain EM (Fig. 2.9b, Fig. 2.10a). Fortunately, none of the MBP fusions notably affected the lid structure, and we were able to identify the positions of all eight essential lid subunits together with the relative orientation of their N- and C-termini. As mentioned before, the peptide Sem1 alone could not be observed by Coommassie-staining, presumably since there are not enough residues that bind the stain to be discernible. However, when N-terminally tagged with the large MBP, a doublet band appeared on the SDS gel at the position of Rpn3, which was confirmed to be MBP-Sem1 by amylose pull-down and Western blot (Fig. 2.9a, data not shown). Thus, Sem1 binds with a strong affinity to the proteasome lid. MBP tags fused to the N terminus of Rpn3 and Rpn12 were largely overlapping in the EM reconstruction, and thus to delineate the envelope of Rpn3 from Rpn12 we produced mutant lid particles that lacked Rpn12. A difference map between wild-type lid and ΔRpn12 lid clearly distinguished the envelope of Rpn12 from Rpn3 (Fig. 2.10b). In combination with the PCI docking, the resolution of secondary structures in the cryo-electron density, and known molecular weights, this information allowed us to delineate approximate subunit boundaries. (Fig. 2.11a).

Overall, Rpn3, 7, 6, 5, and 9 form the fingers of the hand-shaped lid structure. We initially assigned Rpn8 to the density that connects Rpn3 and 9, closing the PCI horseshoe. Later work by the Baumeister group suggested that the majority of this density was instead composed of a helical bundle formed by helices found at the C-terminus of every lid subunit (Beck et al., 2012). I confirmed this latter model in a subsequently published paper, which will described in Chapter 3 of this thesis, though the figures presented in this chapter are as originally published in Lander, *et al*, Nature, 2012 (Estrin et al., 2013). Finally, Rpn11, the only essential DUB of the proteasome, lies in the palm of the hand and makes extensive contacts with Rpn8, 9, and 5.

Using the topology determined for the isolated lid subcomplex, we delineated the individual lid subunits in the context of the holoenzyme (Fig. 2.11b). To complete the

subunit assignment for the entire regulatory particle, the positions of Rpt1-6 in the base subcomplex were assigned according to established interactions with the core particle (Bohn et al., 2010; Tian et al., 2011), whose crystal structure could be docked unambiguously into the EM density (Fig. 2.12). We localized the two large non-ATPases Rpn1 and 2 of the base subcomplex by antibody-labeling of a C-terminal FLAG tag and N-terminal fusion of gluthathione-S-transferase (GST), respectively (Fig. 2.2, 2.13a-c). We found a high structural resemblance between Rpn1 and 2, both consisting of a strongly curled solenoid that transitions into an extended arm towards the C-terminus (Fig. 2.14a). Rpn1 contacts the C-terminal helix of the 20S core subunit α4 and, based on the variability observed in our EM images, is likely to be flexible or loosely attached to the side of the base. Previous crystallography studies of the archaeal proteasome homolog PAN revealed that the N-terminal domains of the ATPases form a separate hexameric ring (N-ring) that consists of OB domains and three protruding coiled-coil segments (Forster et al., 2009; Zhang et al., 2009a). Each coiled coil is formed by the far N-terminal residues of two neighboring ATPases in the hexamer. Although Rpt1 and 2 do not appear to form an extended coiled coil, we find that the N-terminal helical portion of Rpt1 interacts with the solenoid and the C-terminal arm of Rpn1. Rpn2 is located above the N-ring and mounted atop the longest of the protruding coiled coils, formed by Rpt3 and 6. These interactions strongly resemble those observed between Rpt1 and Rpn1 (Fig. 2.14a).

Localizing the ubiquitin receptors and DUBs within the regulatory particle is of particular interest. In addition to the DUB Rpn11 in the lid, we identified the positions of both intrinsic ubiquitin receptors, Rpn10 and 13, and of the base-associated DUB Ubp6 by imaging proteasome particles from yeast deletion strains (Fig. 2.13d-f, 2.14b). The ubiquitin receptor Rpn13 binds to Rpn2 as expected (Hamazaki et al., 2006) (Schreiner et al., 2008). The globular VWA domain of the second receptor Rpn10 had previously been shown to stabilize the lid-base interaction (Glickman et al., 1998a; Verma et al., 2004), however, we found that it does not directly contact the base. This domain bridges Rpn11 and 9, which might increase the lid-base affinity indirectly by stabilizing Rpn11 in its Rpn2-bound conformation (see below). The flexibly attached ubiquitin interacting motif (UIM) of Rpn10 likely contacts the coiled coil formed by Rpt4 and 5, helping to stabilize its position relative to other subunits and potentially communicate with the AAA motor. The DUB Ubp6 appears to be flexible and does not give rise to ordered density. Nonetheless, variance maps indicate that it interacts with the C-terminal arm of Rpn1, as suggested by immunoprecipitations (Leggett et al., 2002).

Inter-subcomplex contacts

The complete localization of subunits within the holoenzyme revealed unexpected contacts between the lid and core subcomplexes. Rpn5 and Rpn6 form fingers that touch the C-termini of the core subunits α1 and α2, respectively. We were able to confirm the interaction between Rpn6 and α2 by *in vitro* crosslinking, using an engineered cysteine in α2 and a 7-Å heterobifunctional crosslinker (Fig. 2.15). These previously unknown direct interactions between lid and core may stabilize the entire holoenzyme assembly. In addition, contacts between lid and core subunits may be part of an allosteric network that modulates the activities of either subcomplex.

Our holoenzyme structure shows that Rpn3, 7, 8, and 11 make extensive contacts with the base. Compared to their positions in the isolated lid, Rpn8 and 11 have undergone significant conformational changes in the holoenzyme (Fig. 2.16). The C-terminus of Rpn8 is detached from Rpn3 to interact with the coiled-coil of Rpt3/6, while the N-terminal MPN domain of Rpn11 extends towards the center of the regulatory particle to bind the solenoid portion of Rpn2. Similarly, the N-terminal region of Rpn3 is more elongated than in the isolated lid and also contacts the Rpn2 solenoid, but from the opposite side. In turn, the extended C-terminal arm of Rpn2 interacts with Rpn3 and 12, and thus forms a direct connection between the solenoid section of Rpn2, the coiled coil of Rpt3/6, and the lid (Fig. 2.14b).

Based on our holoenzyme structure, we speculate that Rpn2 stabilizes a lid conformation in which Rpn3, 8, and the DUB Rpn11 extend towards the base (Fig. 2.16b). Together, the lid, Rpn2, and the coiled coils of the N-ring appear to function as a scaffold that properly positions the two intrinsic ubiquitin receptors Rpn10 and 13, and the DUB Rpn11 for substrate binding, deubiquitination, and transfer to the subjacent central pore of the AAA motor (Fig. 2.14b). Interestingly, several lid subunits directly interact with AAA domains of the Rpts. Rpn7 contacts the AAA domains of Rpt2 and Rpt6, while Rpn6 and Rpn5 touch Rpt3. These interactions with contiguous motor domains are surprising, because current models for ATP-dependent unfoldases suggest significant conformational changes of individual subunits in the hexamer during ATP hydrolysis and substrate translocation (Aubin-Tam et al.; Glynn et al., 2009; Maillard et al., 2011). The observed contacts between lid and the motor domains might form only transiently; alternatively, the AAA ring of the proteasome may be much more static than previously assumed.

Lid conformational changes may regulate DUB activity

Comparing the structures of the lid in isolation and when bound to holoenzyme revealed major conformational changes that suggest an allosteric mechanism for the regulation of Rpn11 DUB activity (Fig. 2.16). In the isolated lid, the N-terminal MPN domain of Rpn11 forms extensive interactions with Rpn9 and the curled up Rpn5 finger. Upon lid binding to the holoenzyme, this Rpn5 finger swings down to contact the α1 subunit of the 20S core and thereby releases Rpn11, which then extends towards the Rpn2 solenoid. Docking the MPN domain of a related DUB (PDBid: 2znr) into the electron density of Rpn11 indicates the approximate location of the active site (Fig. 2.14b). The interactions of Rpn11 with Rpn9 and 5 in the free lid likely restrict access to this active site, which would prevent futile substrate deubiquitination in the absence of base and 20S core, and explain previous observations that the lid subcomplex has DUB activity only within the holoenzyme (Verma et al., 2002) (and our unpublished data).

Functional asymmetry in the AAA unfoldase

Our subnanometer structure of the holoenzyme provides exciting new insights into the architecture and potential mechanisms of the base AAA unfoldase. As suggested by previous EM studies (Bohn et al., 2010; Nickell et al., 2009), the ring of the base and the 20S core are slightly offset from a coaxial alignment, with the base shifted by ~10 Å towards the lid (Fig. 2.17a). Despite or perhaps due to this offset, the C-terminal tails of

Rpt2, Rpt3, and Rpt5 are docked into their cognate 20S binding pockets at the interfaces of the subunits α3 and α4, α1 and α2, and α5 and α6, respectively. Those three Rpt tails contain the terminal HbYX motif, which is critical for triggering gate opening in the 20S core (Gillette et al., 2008; Smith et al., 2007), and in fact our structure is consistent with an open gate conformation. The tails of Rpt1, 4, and 6 lack this motif and were not observed to statically interact with 20S in our holoenzyme structure.

Current mechanistic models for AAA+ unfoldases predict that ATPase subunits in the hexamer are in different nucleotide states and undergo significant conformational changes driven by coordinated ATP-hydrolysis (Glynn et al., 2009; Hersch et al., 2005; Martin et al., 2005). Because we determined the structure of wild-type proteasome in the presence of saturating ATP, we expected that different complexes would have any given Rpt subunit in different conformations, leading to reduced electron density or low resolution when averaging thousands of these unsynchronized motors. However, our reconstruction shows highly ordered density throughout the AAA domains of all six Rpts. While the C-terminal 'small AAA' subdomains (except for Rpt6) arrange in one plane above the 20S core, the 'large AAA' subdomains of Rpt1-5 are oriented in a spiral staircase around the hexameric ring, with Rpt3 at the highest and Rpt2 at the lowest position (Fig. 2.17b). The AAA domain of Rpt6 adopts a tilted orientation, bridging Rpt2 and 3. Similar staircase arrangements have been previously observed for helicases of the AAA+ and RecA superfamilies (Enemark and Joshua-Tor, 2006; Thomsen and Berger, 2009). It was suggested that during ATP hydrolysis, individual subunits progress through the different conformational stages of the staircase, thereby translocating substrate through the pore. The particular staircase orientation we observed identically for all proteasome particles may represent a low-energy state of the base, adopted under our experimental conditions. Alternatively, this staircase arrangement of Rpt1-6 may be static and reflect the functional state of the base, in which substrates are translocated by local motions of the pore loops while the relative positions of the motor subunits stay fixed. Future biochemical and structural studies will be required to distinguish between these two models.

Spatial arrangement of ubiquitin receptors and DUBs

Localizing all subunits of the regulatory particle enabled us to infer the requirements and potential mechanisms for the recognition and degradation of ubiquitin-tagged substrates (Fig. 2.18). After a substrate binds to an ubiquitin receptor, its poly-ubiquitin chain must be removed by Rpn11 cleavage at the proximal ubiquitin to permit subsequent fast degradation (Verma et al., 2002; Yao and Cohen, 2002). To allow cleavage without disengaging from the receptor, an ubiquitin chain must be long enough to span the distance between receptor and DUB. Based on our structure, both Rpn13 and the UIM of Rpn10 are located 70-80 Å from the predicted position of the Rpn11 MPN domain (Fig. 2.14b). The shuttle receptors Rad23, Ddi1, and Dsk2 are expected to reside ~80-120 Å away from Rpn11, depending on where they bind Rpn1 (Gomez et al., 2011). For receptor interaction, the ubiquitin chain has to be in an extended conformation with the hydrophobic patches exposed (Eddins et al., 2007; Riedinger et al.; Schreiner et al., 2008). Since a single ubiquitin moiety in an extended K48-linked chain contributes ~30 Å in length (Cook et al., 1992), it would take three ubiquitins to span the distance between Rpn10 or 13 and Rpn11. Moreover, both Rpn10 and Rpn13 bind between two

consecutive ubiquitin moieties (Riedinger et al.; Schreiner et al., 2008), such that at least a tetra-ubiquitin chain would be required on a substrate to allow interaction with a receptor and simultaneous deubiquitination by Rpn11 (Fig. 2.18). This model agrees with *in vitro* studies that indicate a minimum of four K48-linked ubiquitins is necessary for efficient substrate degradation (Thrower et al., 2000), although this number may differ for other chain types (Bremm et al., 2010). Given the spatial arrangement of Rpn10 and 13, an ubiquitin chain would have to be significantly longer to interact with both receptors. However, knockout studies have shown that ubiquitin chains are not required to bind to multiple receptors simultaneously (Husnjak et al., 2008).

In contrast to Rpn11, Ubp6 is known to cleave within poly-ubiquitin chains or trim them from their distal end (Hanna et al., 2006). Of all the ubiquitin-interacting subunits in the regulatory particle, we found Ubp6 to be the furthest away from the entrance to the pore, which may allow it to clip extended or unnecessary ubiquitin chains from substrates. Because Ubp6 is located closer to Rad23, Dsk2, or Ddi1 than to Rpn10 or 13, it may preferentially act on substrates delivered by these shuttle receptors.

To avoid dissociation upon deubiquitination, a substrate polypeptide must be engaged with the unfolding machinery of the base before or shortly after removal of its ubiquitin chain. Engagement by the base is known to depend on an unstructured initiation site or "tail" on the substrate (Prakash et al., 2004), which needs to be long enough to reach through the narrow N-ring and into the AAA pore (Fig. 2.18). In addition, this tail would have to be sufficiently spaced from the attachment point of the poly-ubiquitin chain to allow concurrent substrate engagement by the pore and deubiquitination by Rpn11. The distance between the predicted active site of Rpn11 and the AAA pore below the N-ring is ~60 Å, which could easily be bridged by 40-45 unstructured residues or a shorter tail combined with a folded structure.

Alternative to the above model for simultaneous receptor binding and deubiquitination, it has been proposed that commencing substrate translocation by the base might move the proximal ubiquitin from a receptor towards Rpn11 for cleavage (Verma et al., 2002). Our structure suggests for this model that efficient substrate processing would only require a mono- or di-ubiquitin for receptor binding and a 50-60 Å longer spacing between the ubiquitin and the flexible tail to reach the AAA pore. This length dependence of engagement is consistent with recent *in vitro* degradation studies, using model substrates with different lengths and ubiquitin modifications (Inobe et al., 2011). Future experiments will be required to assess whether substrates get deubiquitinated in a translocation-dependent or –independent manner.

Discussion

The work presented here defines the architecture of the entire proteasome regulatory particle and provides a much-needed structural framework for the mechanistic understanding of ubiquitin-dependent protein degradation. We localized Rpn11 directly above the entrance of the pore, surrounded by the ubiquitin receptors Rpn10 and 13. This insight into the spatial arrangement of subunits allows us to visualize the substrate's path towards degradation and will be critical in elucidating how the characteristics of ubiquitin modifications affect substrate recognition and processing. Moreover, our study significantly furthers the understanding of the heterohexameric AAA+ motor of the

proteasome. Individual ATPase subunits were found in a spiral staircase arrangement and may operate with more limited dynamics than previously assumed for AAA+ protein unfoldases.

Unexpectedly, the lid is bound to the side of the holoenzyme and interacts with both the base and core particle. These interactions induce major conformational changes in lid subunits that may allosterically activate the DUB Rpn11, allowing critical removal of ubiquitin chains during substrate degradation in the holoenzyme, while preventing futile deubiquitination by the isolated lid. In addition, contacts between the subcomplexes could have unexplored roles in coordinating individual substrate processing steps, for instance ubiquitin binding, deubiquitination, and the onset of translocation. The abundance of these contacts and the intricate architecture of the proteasome may highlight the complex requirements for this proteolytic machine, which must accommodate and specifically regulate a highly diverse set of substrates in the eukaryotic cell.

Figures

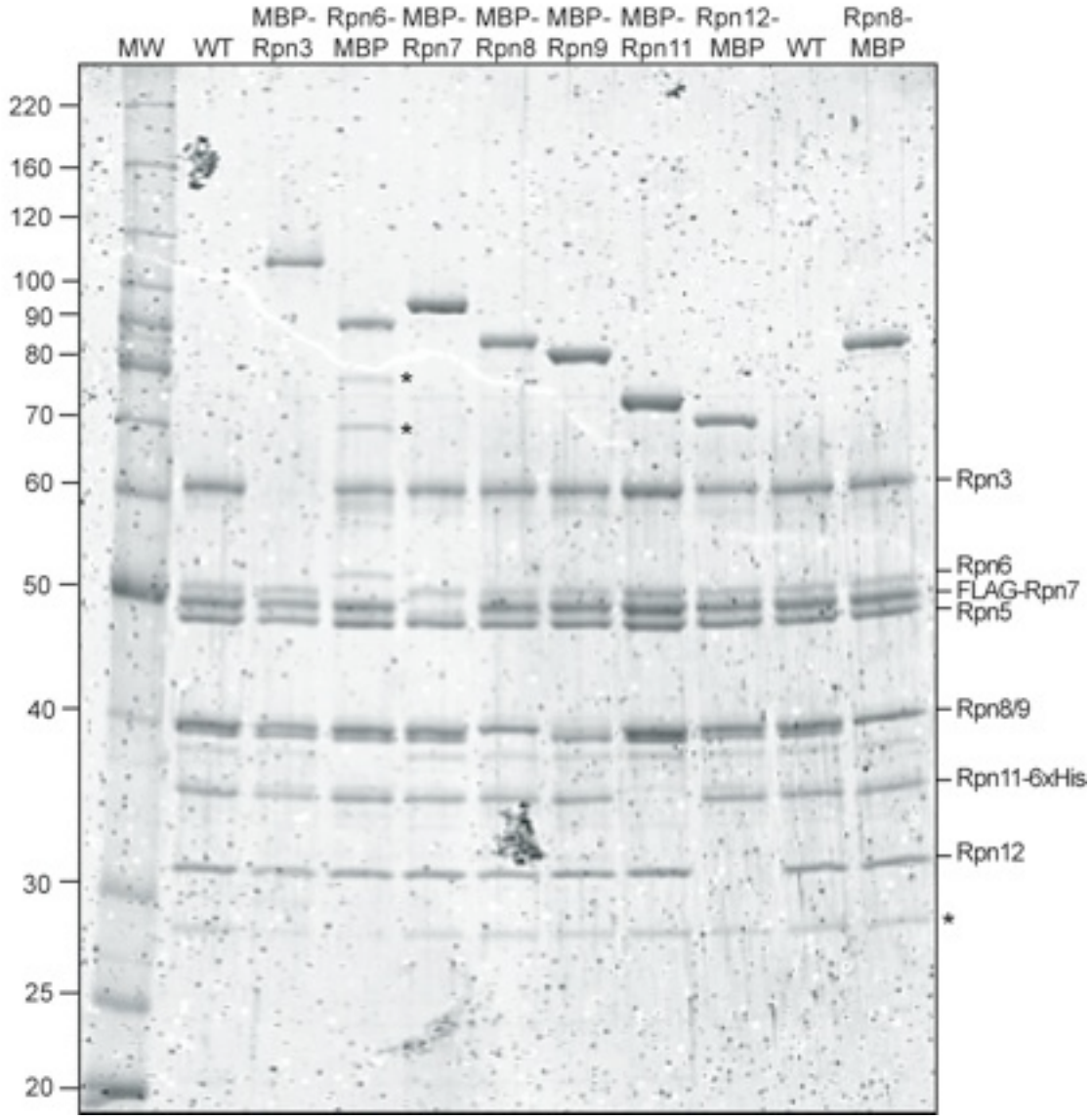


Fig. 2.1: Wild-type and MBP-fused recombinant lid. Sypro Ruby-stained SDS-PAGE gel showing purified recombinant lid with MBP-fused to the N- or C-terminus of individual subunits as indicated. Fused subunits are shifted relative to their wild-type counterpart. * denotes minor degradation products. Rpn6-MBP migrates at lower molecular weight than expected based on the untagged protein. This anomalous behavior might be a consequence of the MBP fusion or caused by proteolytic cleavage of the N-terminus of Rpn6. Class averages of recombinant lid display this N-terminal Rpn6 truncation. To select for lid subcomplexes with full-length Rpn6, the FLAG tag was transferred from Rpn7 to the N-terminus of Rpn6.

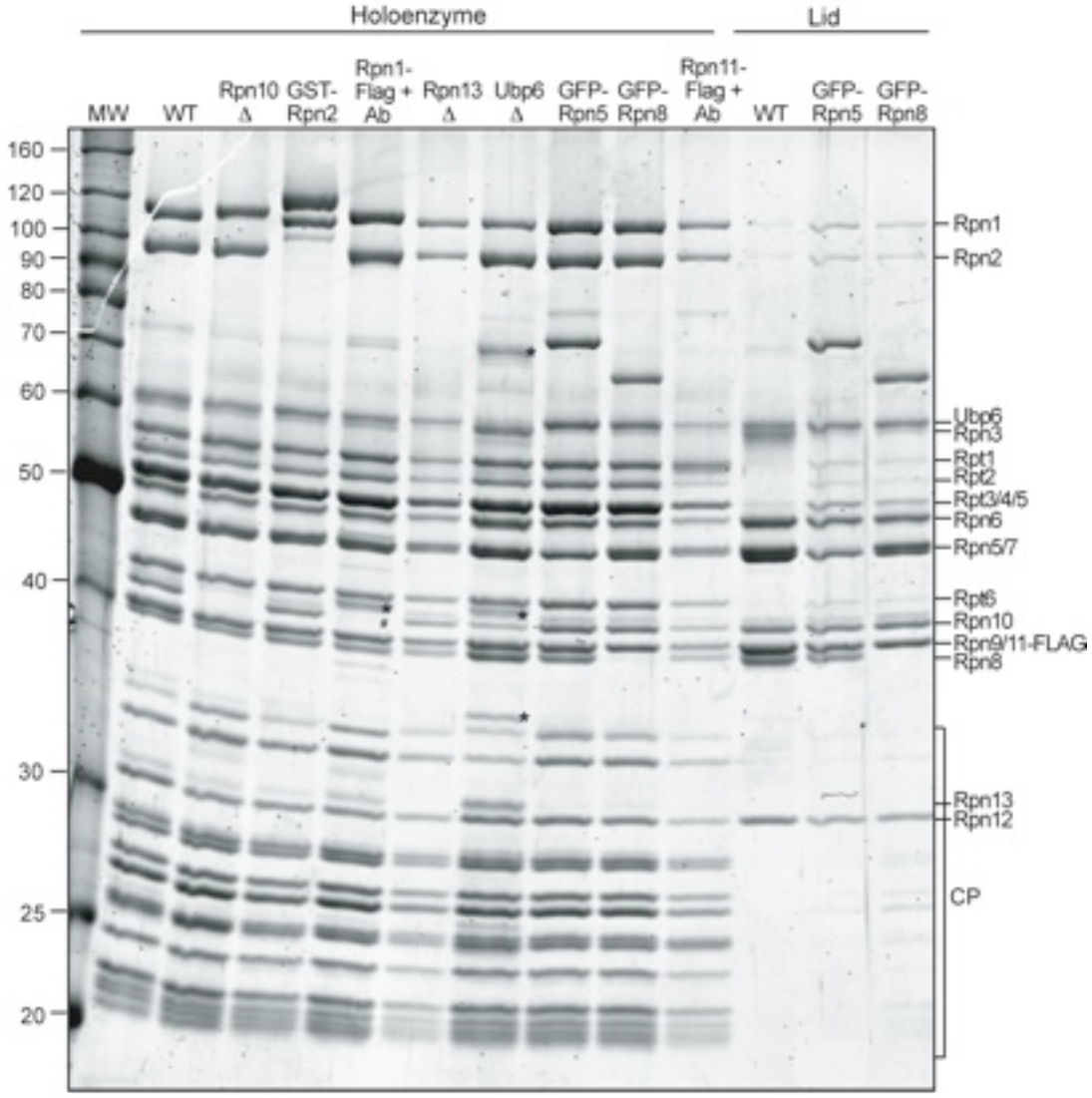


Fig. 2.2: Purified endogenous holoenzyme and lid used in EM analyses. Holoenzyme or lid was purified from yeast strains containing the indicated tag or deletions and a FLAG tag on Rpn11 (with the exception of Rpn1-FLAG) as described, separated by SDS PAGE, and stained with Sypro Ruby. +Ab indicates that sample was incubated with anti-FLAG antibody and purified by gel filtration. * indicates more prominent background bands. # indicates the loss of Rpn10 from Rpn1-FLAG holoenzyme after incubation with anti-FLAG antibody.

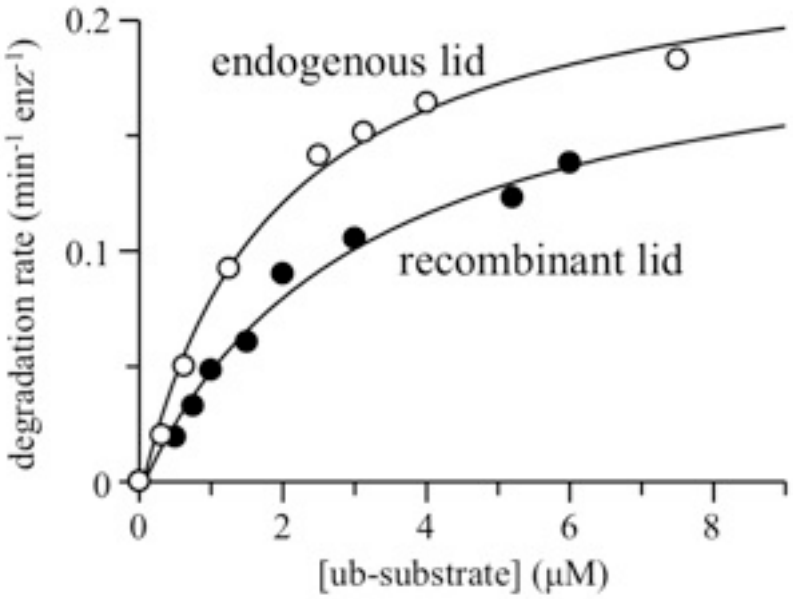


Fig. 2.3: Michaelis-Menten analyses for substrate degradation by holoenzymes reconstituted with recombinant and endogenous lid. Ubiquitinated GFP-titin-cyclin fusion substrate was degraded at 30 °C by proteasome holoenzyme (200 nM) reconstituted from base, 20S core, and recombinant or endogenous lid. Values for K_M and v_{max} were 1.8 μM and 0.24 min^{-1} enz^{-1} for holoenzyme with endogenous lid, and 3.0 μM and 0.21 min^{-1} enz^{-1} for holoenzyme with recombinant lid. The lack of post-translational modifications in *E. coli* might account for these differences.

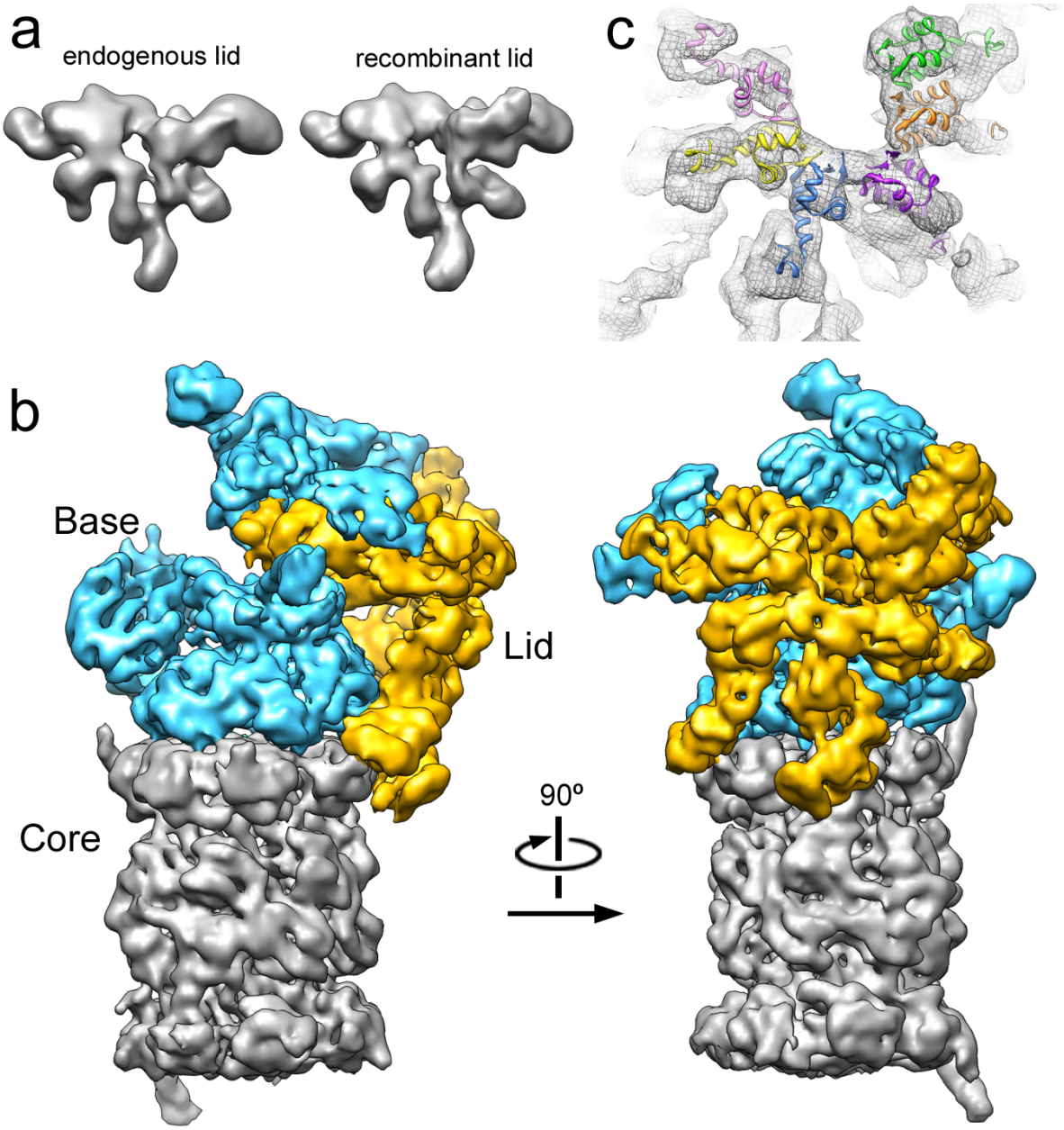


Fig. 2.4: Three-dimensional reconstructions of the recombinant lid subcomplex and the yeast 26S proteasome. a) Negative stain reconstruction of the isolated lid subcomplex at 15Å resolution, colored by subunit and shown from the exterior (left), the side (middle), and the interior, base-facing side (right). A dotted line (middle) indicates the highly variable electron density for the flexible N-terminal domains of Rpn5 and 11. b) Subnanometer cryoEM reconstruction of the holoenzyme, shown in three views corresponding to the isolated lid and colored as above, with the core particle in grey.

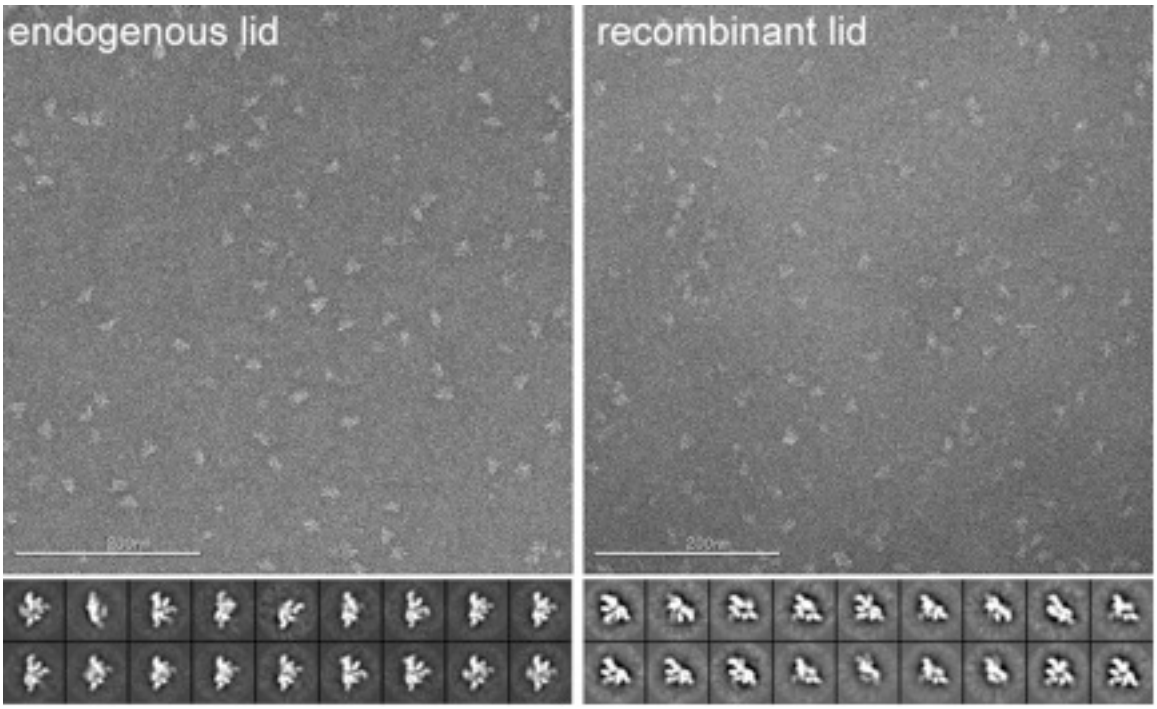


Fig. 2.5: Micrographs of negatively stained endogenous (left) and recombinantly expressed (right) yeast lid subcomplexes. Corresponding 2D class averages of the particles are shown directly below the micrograph, demonstrating that the recombinant lid exhibits the same overall morphology as the endogenous.

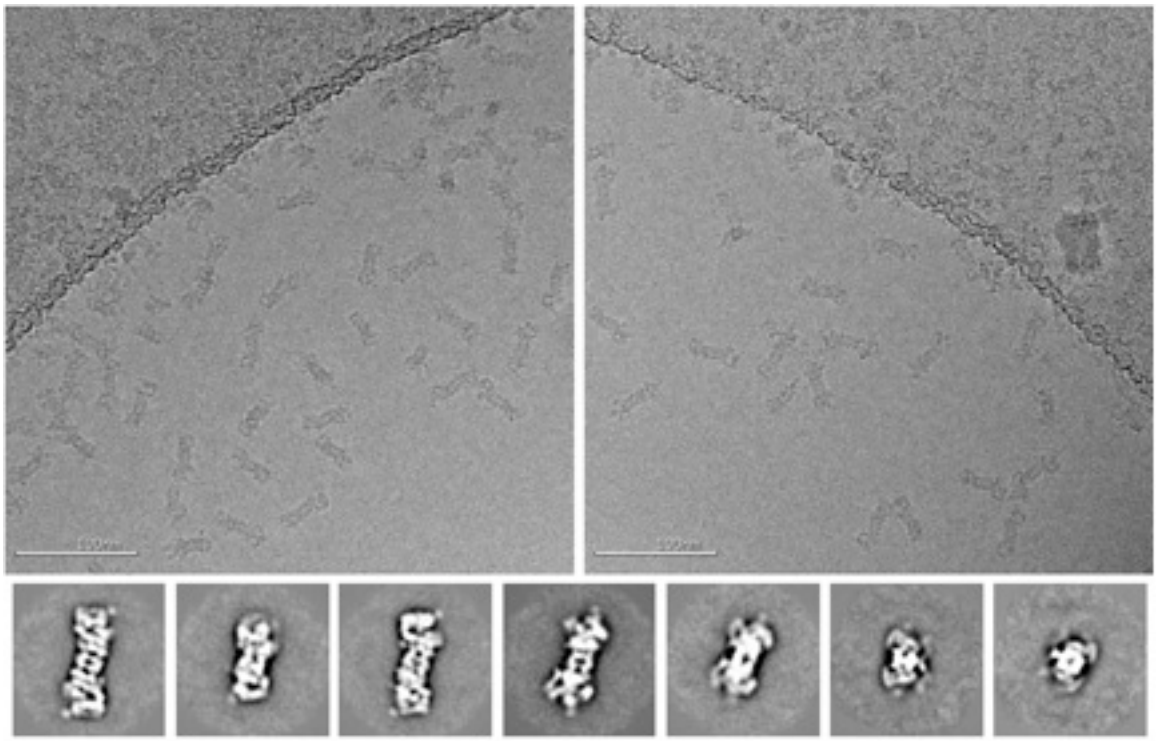


Fig. 2.6: Proteasome holoenzymes preserved in a frozen-hydrated state. Proteasome particles can be observed adopting a range of orientations, even a vertical position, when imaged through thick ice. Reference-free 2D class averages (beneath the micrographs) show a variety of views.

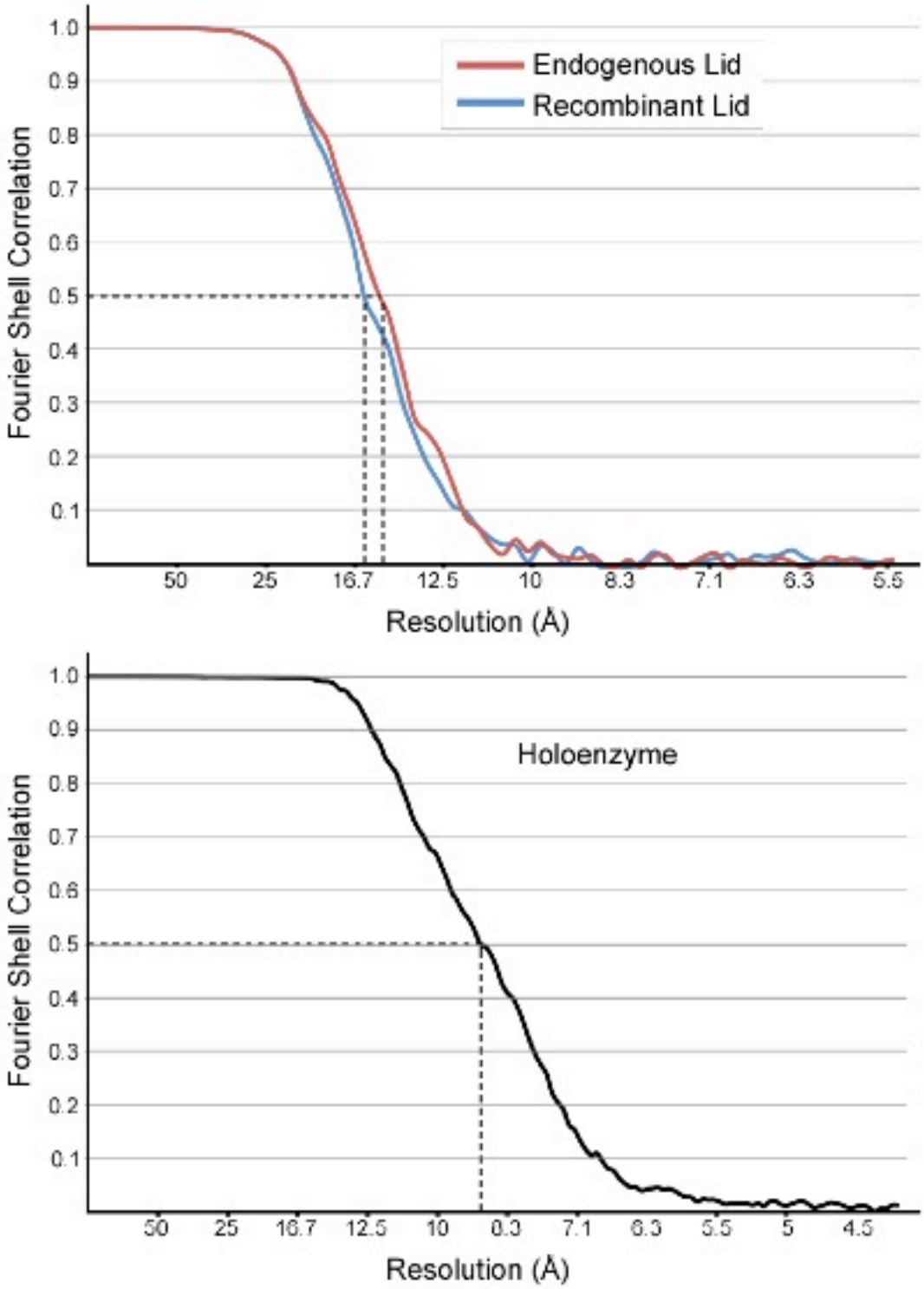


Fig. 2.7: Estimated resolutions. Resolutions of the reconstructions were estimated using a Fourier shell correlation of back projected even/odd datasets, using a criterion of 0.5 correlation. Reported resolutions for the endogenous and recombinant negative stain lid structures are 15 Å and 16 Å, respectively. The resolution for the cryoEM reconstruction is estimated to be 8.8 Å.

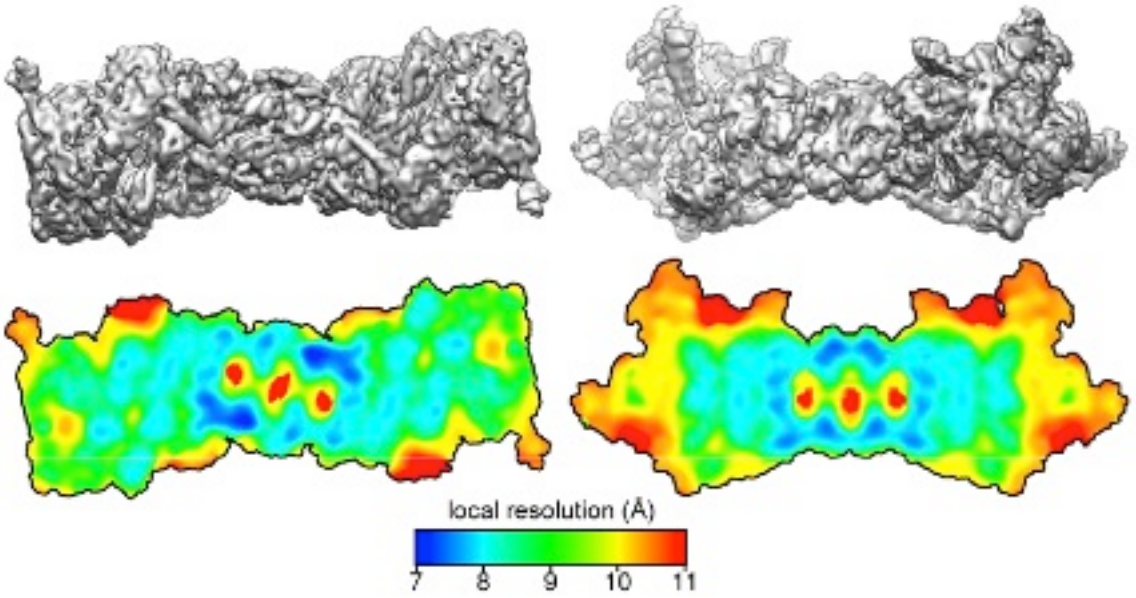


Fig. 2.8: Local resolution map of the holoenzyme cryoEM density. A local resolution calculation of the proteasome reconstruction shows a range of resolutions within the map. In grey are surface representations of the reconstruction, and shown below are cross-sections through the center of the density. The cross sections are colored according the map's calculated local resolution, with the highest resolution portions in dark blue, and the lowest resolution areas in red. Notably, the core particle and AAA+ ATPases contain the highest resolution data, and the ubiquitin receptors are the lowest in resolution.

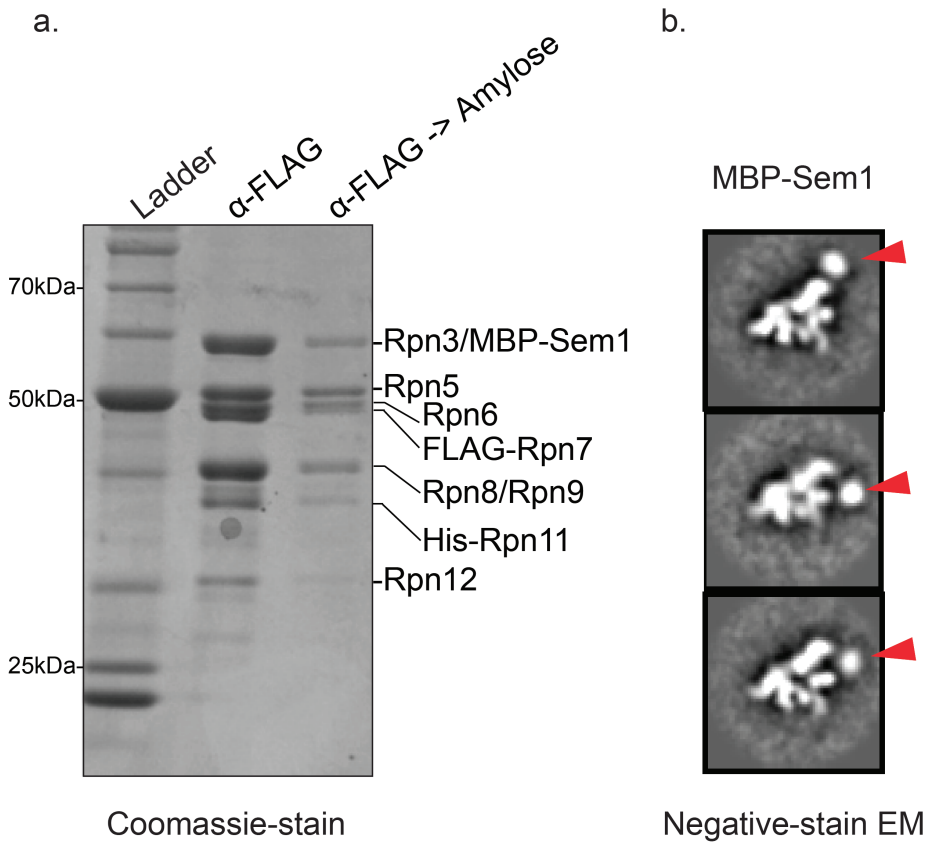
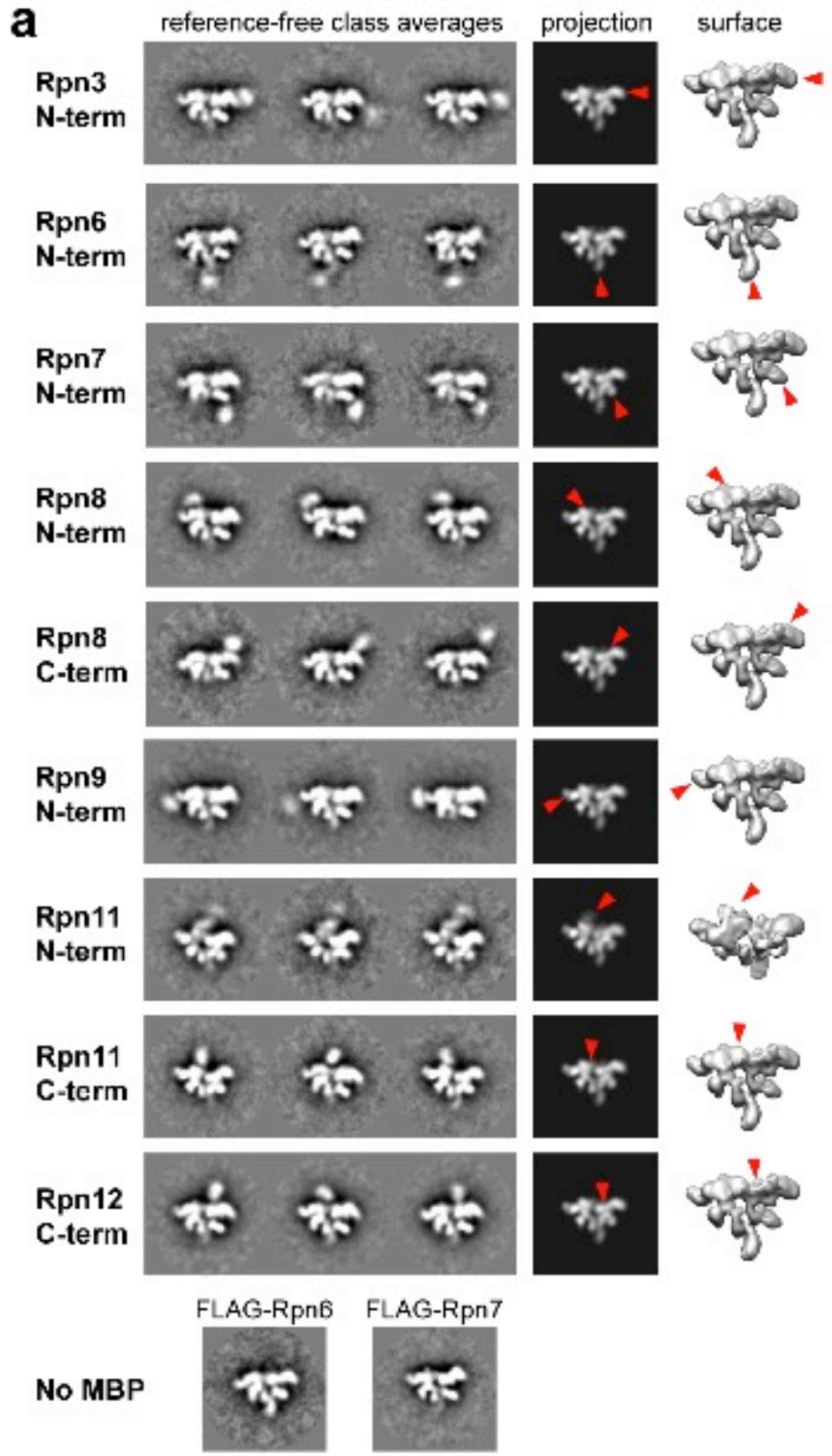


Fig. 2.9: Sem1 is present in the heterologously produced lid and is located between Rpn3 and Rpn7. a) Sem1's ability to pull-down all other lid subunits is verified using amylose resin and SDS-PAGE. b) The N-terminus of Sem1 is located between Rpn3 and Rpn7 based on negative-stain EM of MBP-Sem1.


a
reference-free class averages
projection
surface
Rpn3
N-term
Rpn6
N-term
Rpn7
N-term
Rpn8
N-term
Rpn8
C-term
Rpn9
N-term
Rpn11
N-term
Rpn11
C-term
Rpn12
C-term
FLAG-Rpn6
FLAG-Rpn7
No MBP

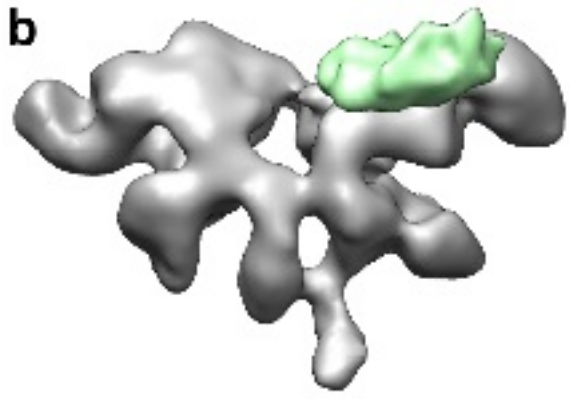


Fig. 2.10: Localization of lid subcomplex subunits by MBP labeling. a) Constructs bearing a FLAG tag on RPN7 and an MBP tag at either the N- or C-terminus of specific subunits were recombinantly expressed in *E. coli* and purified for EM analysis. MBP tags can be clearly observed as a small bright density attached to the subcomplex in reference-free 2D class averages. Three representative class averages for each of the analyzed constructs are shown in the leftmost column. Each MBP tag was unmistakably identifiable in the canonical front view of the lid particles, with the exception of the N-terminal Rpn11, which was only visible in a tilted view of the particle. The corresponding forward projection and surface representation of the recombinant lid reconstruction is shown to the right of each set of class averages, indicating the subunit localization with a red arrowhead. Notably, we see decreased density for the N-terminal portions of Rpn6, which is caused by a fraction of particles with N-terminal Rpn6 truncations (see also Fig. S1). We were able to select for a full-length Rpn6 by purifying the complex using a FLAG tag on Rpn6 (bottom panel) b) A Rpn12-deletion mutant clearly shows the location of Rpn12 in the lid complex. The difference density between the recombinant lid and recombinant Rpn12 delete lid is shown in green.

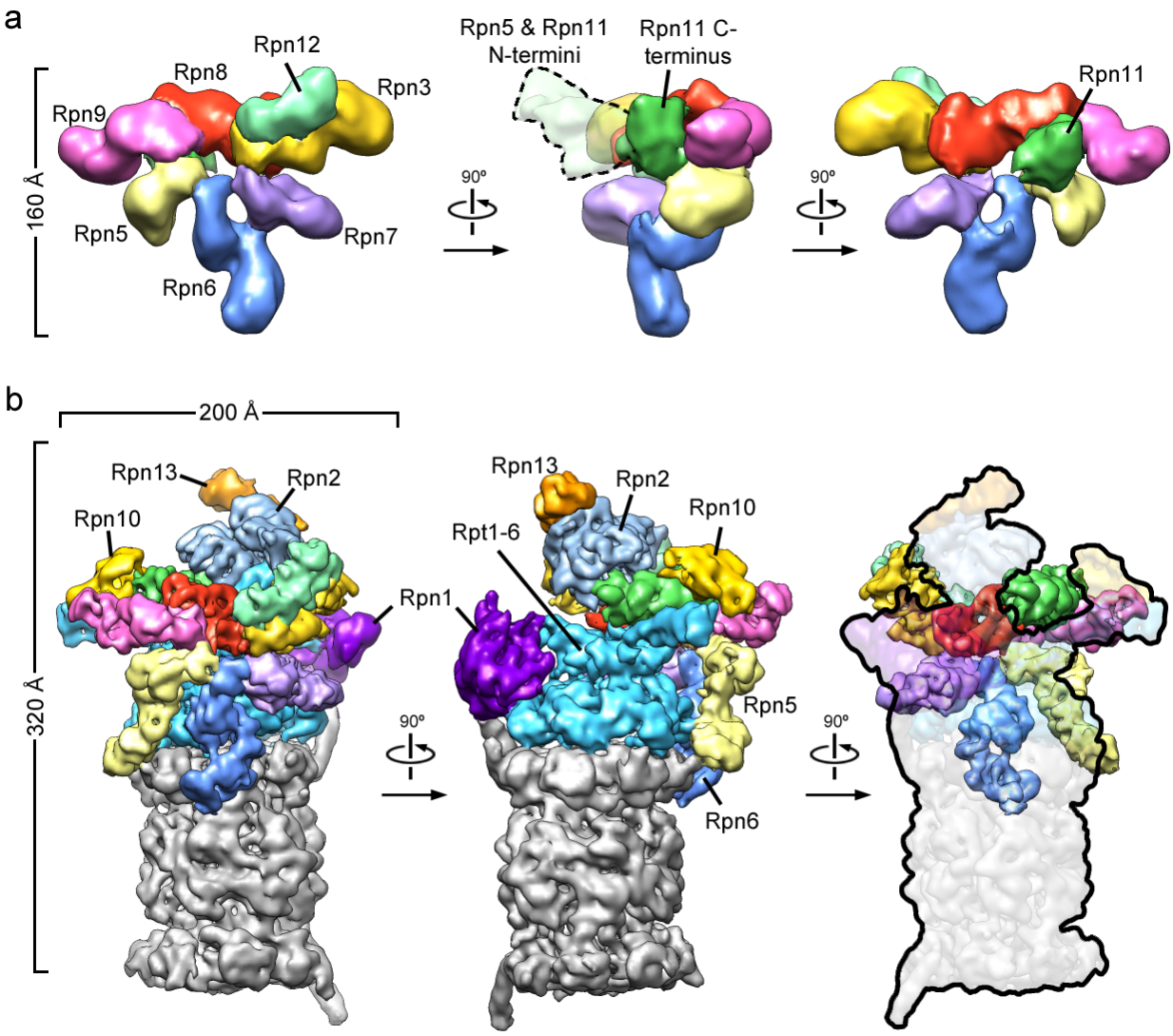


Fig. 2.11: Three-dimensional reconstructions of the recombinant lid subcomplex and the yeast 26S proteasome. a) Negative stain reconstruction of the isolated lid subcomplex at 15Å resolution, colored by subunit and shown from the exterior (left), the side (middle), and the interior, base-facing side (right). A dotted line (middle) indicates the highly variable electron density for the flexible N-terminal domains of Rpn5 and 11. b) Subnanometer cryoEM reconstruction of the holoenzyme, shown in three views corresponding to the isolated lid and colored as above, with the core particle in grey. Note that the position of Rpn8 was later revised as discussed in the text of Chapter 2 and Chapter 3.

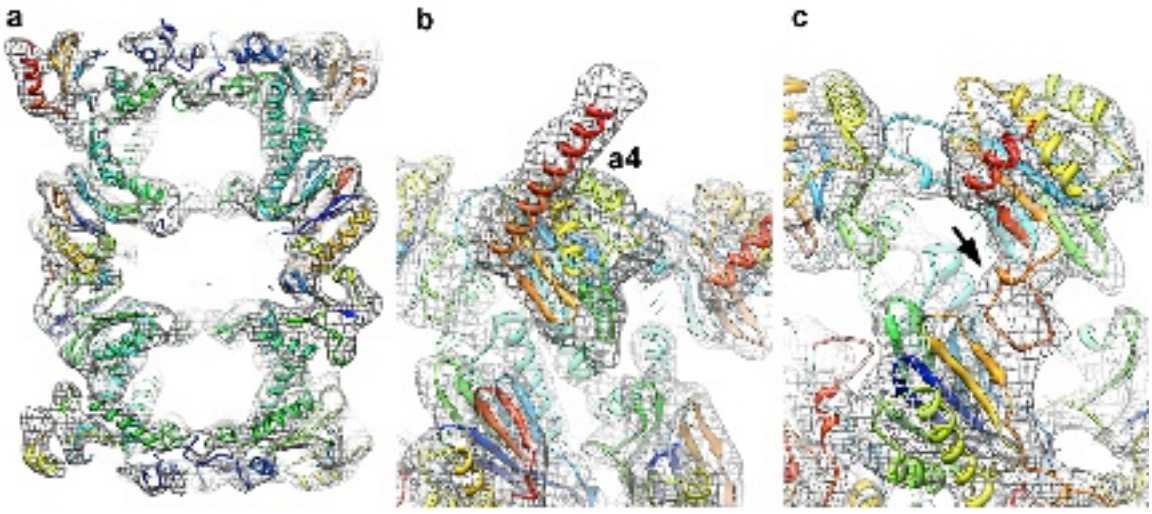


Fig. 2.12: Unambiguous docking of the crystal structure for yeast 20S core. Docking of the 20S core structure (PDBid: 1ryp) into the EM density provides an asymmetric orientation of the core relative to the base. a) Cross section of the crystal structure docked into the EM density, showing the high level of correlation between the molecular envelope of the electron density and the secondary structural elements of the atomic coordinates. b) Extended α-helix of the α4 subunit. The helix was extended to include the entire C-terminus. c) The insertion-loop of subunit α2 is obvious in the EM density.

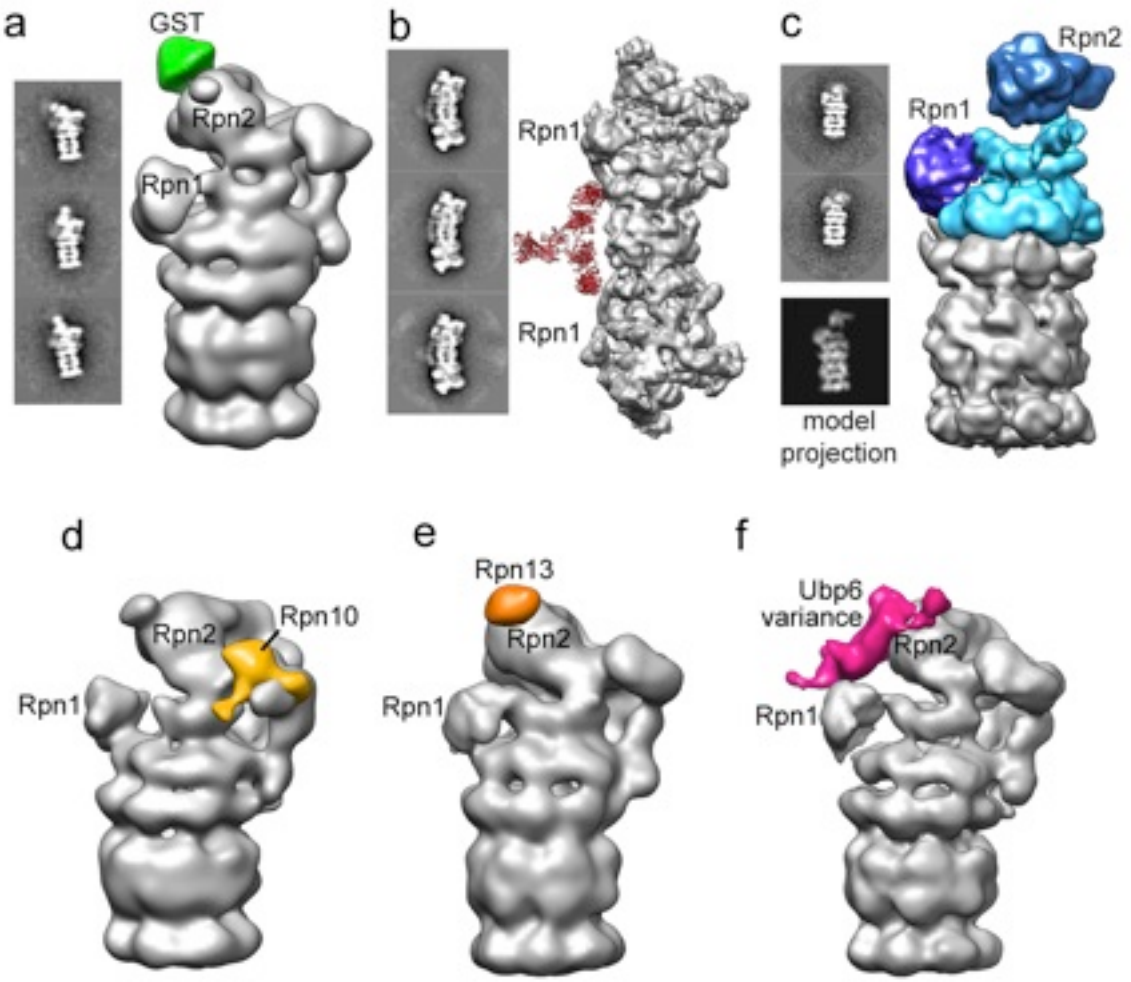


Fig. 2.13: Localization of Rpn1, Rpn2, and ubiquitin-interacting subunits. a) Reference-free class averages and reconstruction of an N-terminal GST fusion of Rpn2 revealed its location on the top of the holoenzyme. b) Antibody labeling of a C-terminal Rpn1 FLAG tag results in 2D class averages showing a dimeric antibody with the single-chain variable fragments attached to Rpn1. The view observed in the class averages is depicted using the holoenzyme reconstruction, with an antibody modeled alongside it. c) A small subset of all holoenzyme preparations resulted in particles that had lost the lid subcomplex. Although there were not sufficient views of these aberrant particles to generate a 3D reconstruction, a theoretical model generated by including only Rpn1, Rpn2, Rpt1-6, and core particle subunits accurately represents the observed class averages (shown as forward projection and surface rendering). d, e) Deletion mutants were used to generate difference maps (colored) that indicate the locations of the ubiquitin receptors Rpn10 and Rpn13, respectively. f) Due to the variability of Ubp6, this DUB was localized by subtracting the variance map of the wild-type holoenzyme from a variance map of an Ubp6-deletion mutant. Difference variance map is colored magenta.

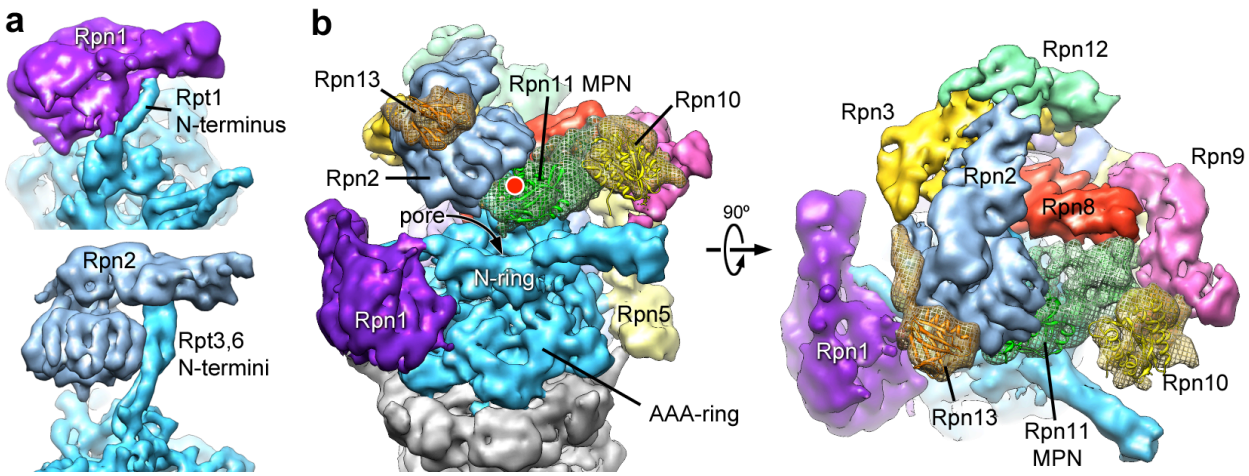


Fig. 2.14: Localization of Rpn1 and Rpn2, and ubiquitin-interacting subunits. a) Rpn1 (top) and Rpn2 (bottom) are oriented to emphasize similarities in their domain structure and solenoid attachment to the extended N-terminal helices of Rpt1 and Rpt3/6, respectively. b) Side and top views of the regulatory particle, showing the locations of the ubiquitin receptors Rpn10 and 13, and the DUB Rpn11 relative to the central pore. Crystal structures for Rpn10 (PDBid: 2x5n), Rpn13 (PDBid: 2r2y), and an MPN domain homologous to Rpn11 (AMSH-LP, PDBid: 2znr) are shown docked into the EM density. The predicted active site of Rpn11 is indicated (red dot).

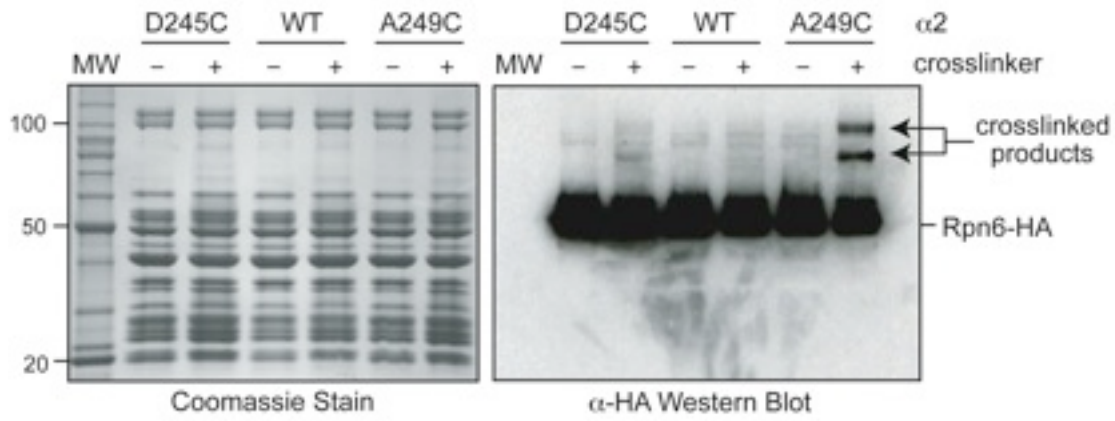


Fig. 2.15: Rpn6 contacts subunit α2 of the 20S core. Our cryoEM reconstruction of the proteasome indicates a direct contact between the α2 subunit of the core particle and Rpn6 of the lid. To confirm this contact by crosslinking, we engineered a cysteine in α2 either at the predicted point of contact (A249C) or nearby (D245C), and conjugated sulfo-MBS, a short (7.3Å spacer arm) heterobifunctional crosslinker, to this site. The core particle contains other cysteines, but those are relatively inaccessible to cysteine-reactive modifying agents (data not shown). Crosslinker-conjugated (or mock-conjugated) core particle purified from strains containing WT, A249C, and D245C α2 was incubated with purified base, Rpn10, 0.5 mM ATP, and lid purified from a yeast strain in which Rpn6 was C-terminally tagged with a 3x hemagglutinin (HA) tag. Reactions were divided equally for separation by SDS-PAGE followed by either coomassie staining or anti-HA western blotting. Rpn6 has a molecular weight of 50 kDa, α2 of 27 kDa, and a crosslink between them should create an anti-HA-reactive band above 77 kDa. This crosslinked band appears only when the cysteine is placed at A249C, closest to the predicted contact between α2 and Rpn6. The two different crosslinked products likely represent crosslinking to two different sites on Rpn6.

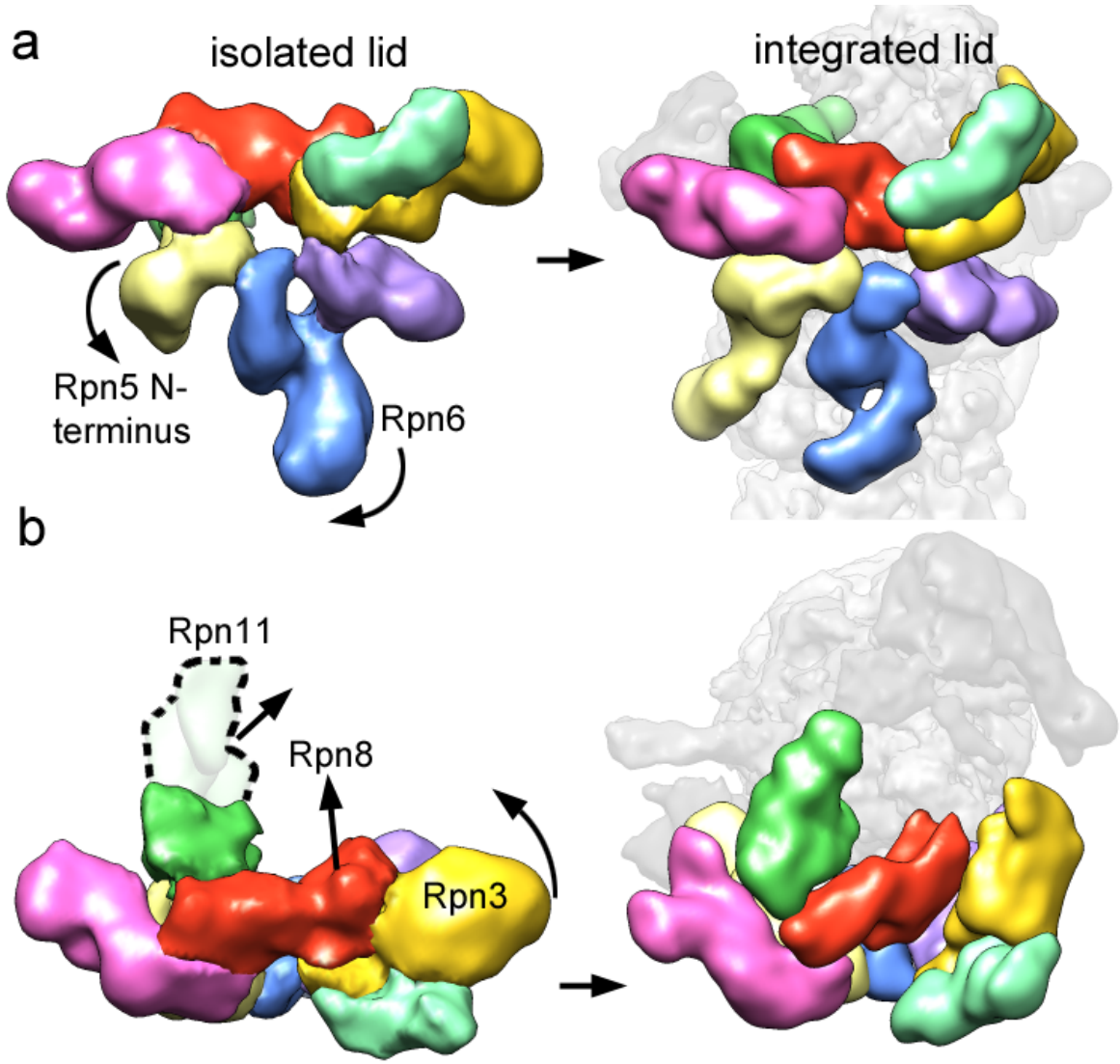


Fig. 2.16: Conformational rearrangements of the lid subcomplex upon integration into the holoenzyme. The lid complex in its isolated (left) and integrated (right) state is shown as viewed from the exterior (a) and top (b) of the regulatory particle. Major subunit rearrangements are depicted by arrows. The N-terminus of Rpn5 (light yellow) interacts with Rpn11 in the isolated complex, and swings down to contact the core particle upon incorporation into the holoenzyme. The N-terminal domain of Rpn6 swings to the left to interact similarly with the core particle. Rpn3, 8, and 11 undergo dramatic rearrangements, in which they move towards the center of the regulatory particle.

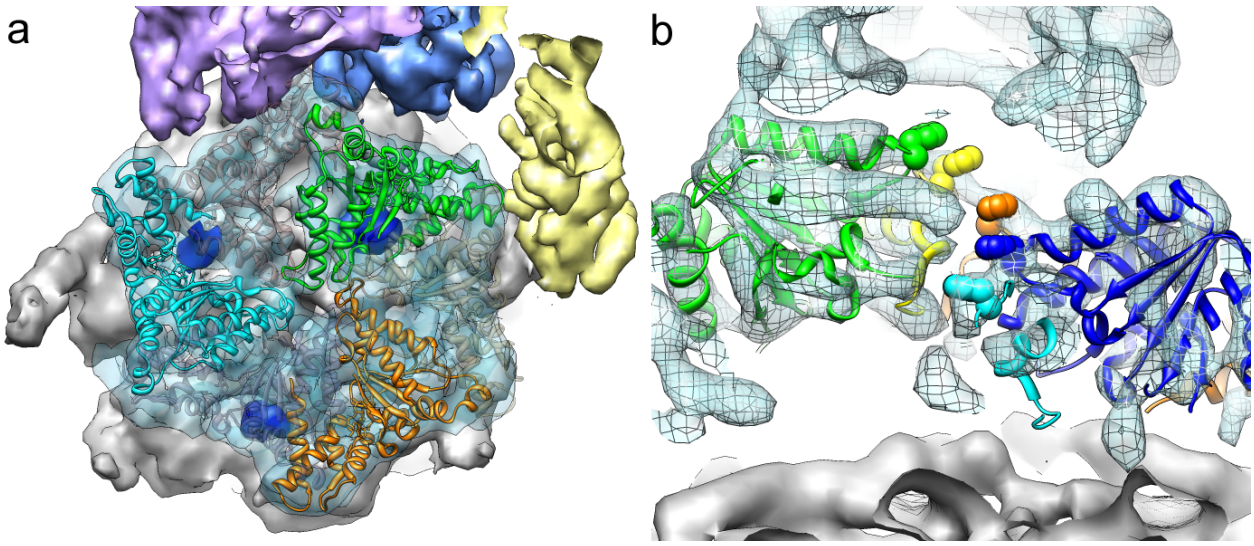


Fig. 2.17: Structural features of the base ATPase subunits. a) Positions of Rpt2 (cyan), 3 (green), and 5 (orange) within the base hexameric ring and relative to the 20S core (grey) are shown using fitted crystal structures of the homologous PAN AAA domain (PDBid: 3h4m). The EM density contains the molecular envelope of the C-terminal tails (dark blue), docked into their cognate binding sites on the 20S core. Corresponding densities were not found for the tails of the Rpt1, 4, and 6 (grey ribbon structure). b) Cutaway side view of the holoenzyme EM density with Rpt1-5 visible. Individually docked copies of the PAN crystal structure reveal a spiral staircase arrangement of the Rpts, emphasized by space-filling representations of the PAN pore-1 loop residues (not resolved in the Rpts).
ATP, and lid purified from a yeast strain in which Rpn6 was C-terminally tagged with a 3x hemagglutinin (HA) tag. Reactions were divided equally for separation by SDS-PAGE followed by either coomassie staining or anti-HA western blotting. Rpn6 has a molecular weight of 50 kDa, α2 of 27 kDa, and a crosslink between them should create an anti-HA-reactive band above 77 kDa. This crosslinked band appears only when the cysteine is placed at A249C, closest to the predicted contact between α2 and Rpn6. The two different crosslinked products likely represent crosslinking to two different sites on Rpn6.

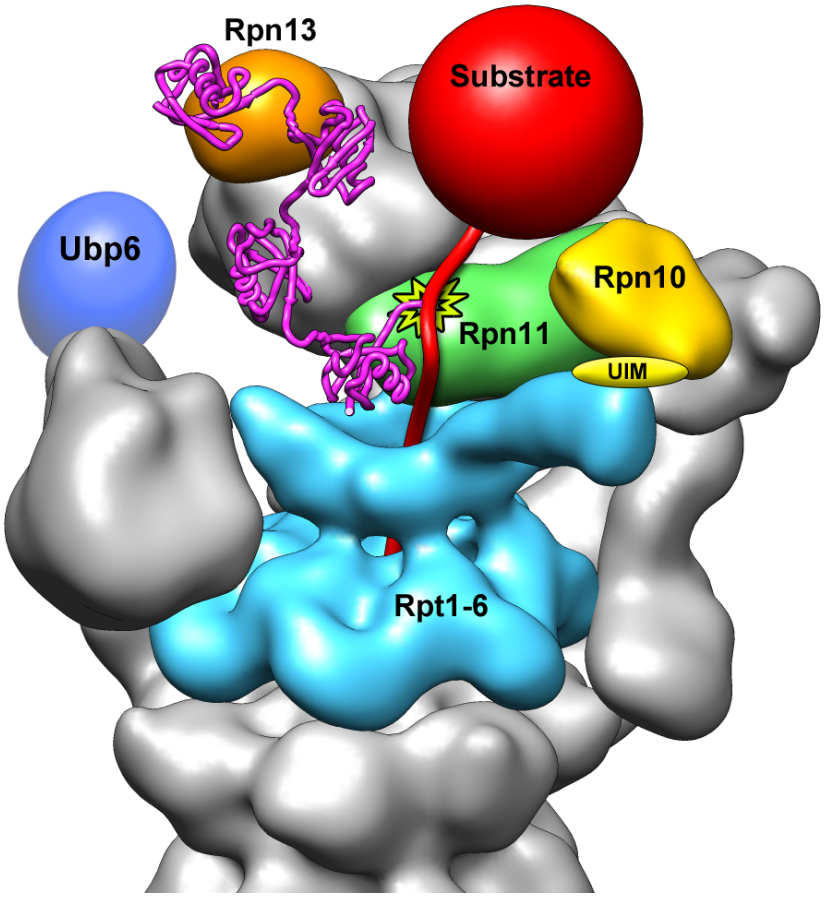


Fig. 2.18: Model for the recognition, deubiquitination, and engagement of a poly-ubiquitinated substrate by the 26S proteasome. A K48-linked tetra-ubiquitin chain (magenta, PDBid: 2kde) is conjugated to the unstructured initiation region of a substrate (red) and bound to the ubiquitin receptor Rpn13 (orange). The substrate is poised for deubiquitination by Rpn11 (green, active site indicated by star), and its unstructured initiation region is engaged by the translocation machinery of the base (cyan). A poly-ubiquitin chain could alternatively bind to the UIM of Rpn10 (yellow) or interact with both receptors simultaneously. The DUB Ubp6 is localized further from the central pore, in a position to trim excess ubiquitin chains.

of the PAN pore-1 loop residues (not resolved in the Rpts).

ATP, and lid purified from a yeast strain in which Rpn6 was C-terminally tagged with a 3x hemagglutinin (HA) tag. Reactions were divided equally for separation by SDS-PAGE followed by either coomassie staining or anti-HA western blotting. Rpn6 has a molecular weight of 50 kDa, α2 of 27 kDa, and a crosslink between them should create an anti-HA-reactive band above 77 kDa. This crosslinked band appears only when the cysteine is placed at A249C, closest to the predicted contact between α2 and Rpn6. The two different crosslinked products likely represent crosslinking to two different sites on Rpn6.

Chapter 3: Formation of an intricate helical bundle dictates the assembly of the 26S proteasome lid

This chapter is based on work published in Estrin et al., Structure, 2013. José Ramón Lopez-Blanco and Pablo Chacon are responsible for the hybrid approach used for the assignment of the C terminal helices contributed to the bundle by individual lid subunits.

Introduction

Though the assembly factors and the biogenesis for the base and the core particle have been well studied, the detailed mechanisms involved in lid assembly remain largely unknown. As discussed in the previous chapter, we have recently shown that *E. coli*-expressed lid resembles the endogenous yeast lid, both in structure and function, suggesting that there are no essential factors dedicated to lid assembly (Lander et al., 2012). Previous cryo-EM reconstructions revealed that the six PCI domains of the lid are arranged in a horseshoe-shaped structure that appears to act as a scaffold and stabilize the lid through substantial lateral interactions between PCI-containing subunits (da Fonseca et al., 2012; Lander et al., 2012; Lasker et al., 2012). Additionally, a number of subassemblies of the lid have been detected *in vivo* and upon subcomplex dissociation *in vitro*, and these observations have given rise to an assembly model wherein Rpn5/8/9/11 associate first, followed by the ordered additions of Rpn6, Rpn3/7, and finally Rpn12 (Fukunaga et al., 2010; Sharon et al., 2006; Tomko and Hochstrasser, 2011). Furthermore, Rpn12 has been shown to use a short C-terminal segment for binding the completed lid subcomplex, and this association is required for efficient incorporation of the lid into the regulatory particle (Tomko and Hochstrasser, 2011). Recent cryo-EM studies have located a helical bundle within the regulatory particle, and it has been suggested that this structure is composed of the lid subunits' C-termini (Beck et al., 2012). However, the functional relevance and detailed topology of the bundle are unknown.

Here, we investigated the assembly mechanism of the proteasome lid. We show that the helices found at the C-terminus of each lid subunit form a peculiar bundle that governs an ordered self-assembly process. Moreover, we obtained an *ab initio* atomic model of this bundle based on a new combinatorial search algorithm combined with previously reported crosslinking data. The resulting topology reveals how the helical bundle serves as an unusual structure that couples the completion of lid assembly with binding to the base and thus with the activation of Rpn11's deubiquitinase activity. Our data also suggest that this bundle represents an important architectural feature that likely enables the lid to tolerate conformational changes in the regulatory particle during substrate processing. Overall, these findings shed light on the mechanisms by which the related CSN may assemble.

Methods

Recombinant lid construction and purification.
Plasmids for recombinant lid production were the same as described in Chapter 2, with the exception of the addition of a N-terminal MBP (maltose binding protein) to Rpn6. However, the Rpn8 & Rpn11 Δ MPN construct lacked the MBP tag on Rpn6. Truncations of individual subunits were produced using conventional cloning techniques. Lid proteins were expressed in *E. coli* BL21-star (DE3) as described in Chapter 2. Cells were collected by centrifugation (4,000g for 30 min), resuspended in FLAG buffer (50 mM HEPES, pH 7.6, 100 mM NaCl, 100 mM KCl, 5% glycerol) supplemented with protease inhibitors and 2 mg/ml lysozyme, and sonicated on ice for 2 min in 15-s bursts. The lysate was clarified by centrifugation (27,000g for 30 min), and lid assemblies were affinity-purified either using anti-FLAG M2 resin (Sigma-Aldrich), Amylose resin (New England Biosciences), or Ni-NTA Agarose resin (Qiagen), selecting for FLAG-Rpn7, MBP-Rpn6, and His_6-Rpn11, respectively. The protein was concentrated in a 30,000 MWCO concentrator (Amicon) for further purification on a Superose 6 size-exclusion column (GE Healthcare) equilibrated in FLAG buffer plus 1mM DTT (Fig. 3.2a).

Hybrid approach for helical bundle assignment.
We developed a two-step integrative approach to unravel the structure of the proteasome helical bundle. In the first step, a combinatorial search is performed to assign the helix predictions of lid C-termini with the physical constraints observed in the cryo-EM map. Then, the compatible models found in the search are further screened with available cross-link information.

EM-map derived geometrical constraints
The helical bundle region was extracted from the experimental proteasome map (EMDB ID 2165 (Beck et al., 2012)) using the Situs package (Rusu et al., 2010). Twelve helical segments were automatically detected using a stochastic template-based search method (Rusu and Wriggers, 2012) (Table 3.2). The last residues of the flexibly-fitted globular structures for Rpn3, Rpn5, Rpn6, Rpn7, Rpn8, Rpn9, Rpn11, Rpn12 (PDB-ID 4B4T (Beck et al., 2012)) were taken as starting points for the combinatorial search. The anchor points and the positions of the traced helices were used as geometric constraints (Table 3.2).

Topological constraints
The secondary structure predictions have been carried out with *PSIPRED* (Buchan et al., 2010) from the corresponding lid C-terminal sequences (Table 3.1). The topological constraints (TC) were calculated from the predicted linker and helix length. The helix lengths were estimated by multiplying the corresponding number of predicted helix residues by 1.5Å. We employed a maximum inter-residue linker length of 2.9Å computed from a representative benchmark of ~200 loops (Chys and Chacon, 2013).

Combinatorial search.
The combinatorial search algorithm recursively matched all TCs and GCs while allowing a certain tolerance to account for inaccuracies in the modeling procedures. By default a

broad tolerance was employed. In fact, the linker lengths were extended by ~9Å (3 residues) and the length mismatch between predicted helices and helices observed in the EM map was set to ±30%. Since the number of helices present in the map must be greater or equal to the number of predicted helices, one predicted helix was split in two. It was apparent that only one of the longest predicted helices, Rpn8 (50 residues) or Rpn11 (40 residues), could match into the largest helix of the map (38 residues). Valid configurations were only obtained by splitting Rpn11 into two helices of 19 residues each, separated by a linker of 2 residues. Splitting the long helix of Rpn8 was an unsuccessful strategy and no valid configurations were obtained. Even using generous tolerances and relatively few constraints, the search yielded only four solutions (Table 3.3). All solutions shared the same assignment but differed in the orientation of the last helix of both Rpn8 and Rpn12. In general, the predicted helices fit very well into the assigned densities (Fig. 3.9b and Fig. 3.8). The discrepancy between the number of residues for the predicted versus EM-extracted helices was typically below 4 residues.

Cross-link evaluation

For further validation, the helix configurations found in the combinatorial search were first idealized by a linear arrangement of 1.5Å spaced C_α atoms and then centered into the corresponding assigned EM helix. The cross-links available for the helical bundle region were taken from Kao et al. (Kao et al., 2012) and Lasker et al. (Lasker et al., 2012) and detailed in Table 3.4. The cross-link distances of the models were measured between the C_α atoms of lysine residues from either the linear helices of the bundle or the fitted lid subunits. We defined a crosslink violation distance of 34Å. This cross-linking distance was calculated by adding 5Å to previous reported crosslink distances estimated with crystallographic structures (Chen et al., 2010; Seebacher et al., 2006) to account for modeling inaccuracies in tracing, prediction, and fitting procedures. All the computed crosslink distances of the four solutions found in the combinatorial search were compatible. However, incompatible distances of 41-42Å were found for solutions 1 and 2, which were therefore rejected. The remaining solutions 3 and 4 only differed in the orientation of Rpn12's helix. Previous cross-linking experiments have shown that the last residue of Rpn12 is in close proximity to the N-terminal coiled-coil domain of Rpt3 (Tomko and Hochstrasser, 2011), which is only in agreement with solution 3.

Modeling

The final model was further refined and extended to heavy atom representation with Chimera (Goddard et al., 2007). Moreover, the helices were manually rotated around their main axis to roughly expose the hydrophilic face to the solvent. Finally, this model was flexibly fitted into the helical bundle segmented map by using iMODFIT with default parameters.

Availability

The Protein Data Bank accession code for the predicted topological model of the helical bundle is 3J47.

Results

The C-terminus of each subunit in the lid is predicted to form one or more helices (Fig. 3.1a). These C-terminal helices are highly conserved (Fig. 3.1b) and have been suggested to form a helical bundle structure (Fig. 3.1c) (Beck et al., 2012). For several lid subunits, the structures of their individual MPN or PCI domain have been solved by X-ray crystallography, but in all cases their C-terminal helices were either truncated before crystallization or not resolved in the electron density (Boehringer et al., 2012; Pathare et al., 2011; Sanches et al., 2007) Based on the lack of ordered density and observed truncations during expression and purification, these C-terminal regions are likely flexible or flexibly attached to individual subunits when not incorporated in the lid.

In order to determine the function of the C-terminal helices as well as the architecture of the helical bundle, we utilized our recently developed heterologous *E. coli* expression system to produce yeast lid particles lacking C-terminal helices from individual subunits. These constructs contained N-terminal fusions to three different subunits: a His_6 tag on Rpn11, a FLAG epitope on Rpn7, and maltose binding protein (MBP) on Rpn6 (Fig. 3.2a). This allowed us to assess lid assembly and observe subassemblies using affinity purification steps followed by gel filtration chromatography (Fig. 3.2b-c).

The C-terminal helices are essential for lid assembly

First, we deleted the C-terminal helix of Rpn12 and found that only Rpn12 was absent from an otherwise complete lid complex (Fig. 3.3a, Fig. 3.4a). This is consistent with previous studies, showing that Rpn12 requires its C-terminus for incorporation into the lid (Tomko and Hochstrasser, 2011).

Next, we produced particles lacking the C-terminal helix of Rpn5. Surprisingly, Rpn12 was missing from an otherwise completely assembled lid subcomplex (Fig. 3.3b, Fig. 3.4a). Similarly, we found that truncating the C-terminal helix of Rpn9 prevented Rpn12 from assembling with the lid complex (Fig. 3.3c, Fig. 3.4a). Together, these data suggest that the tails of Rpn5 and Rpn9, which previously had no presumed functions, are important for Rpn12 binding to the lid. Rpn12 had been suggested to use its C-terminus for monitoring a complete lid assembly state, but the mechanism of this surveillance remained elusive (Tomko and Hochstrasser, 2011). The individual tail truncations of Rpn5 and Rpn9 now suggest that the mechanism of surveillance by Rpn12 may be mediated through the C-terminal tails of the lid subunits.

Unlike most other lid subunits, Rpn3 contains a 45 amino-acid extension past its conserved C-terminal helix. This extra region is neither well conserved nor confidently predicted in its secondary structure. Upon deletion of the extension, we did not observe any lid assembly defects, suggesting that this portion of Rpn3 is not essential for complex formation (Fig. 3.4b).

In contrast, deleting the conserved C-terminal helix of Rpn3 caused the lid to be separated into a Rpn5/6/8/9/11 subassembly, the Rpn3/7 heterodimer, and Rpn12, though for a small fraction the Rpn3/7 heterodimer was also associated with the Rpn5/6/8/9/11

assembly (Fig. 3.3d, Fig. 3.4b-c). Thus, Rpn3 and Rpn7 stably interact with each other through their PCI domains. We expect that the observed fractions of lid-bound and isolated Rpn3/7 heterodimer do not reflect their distribution under equilibrium conditions, given the extended purification procedure. Nevertheless, this partitioning indicates that the Rpn3/7 heterodimer has a reduced affinity for the Rpn5/6/8/9/11 subassembly when Rpn3's helix is missing.

We observed complementary results when Rpn7's C-terminal helix was deleted: the lid was separated into two subassemblies, Rpn3/7 and Rpn5/6/8/9/11, while Rpn12 did not associated with any members of the lid (Fig. 3.3e, Fig. 3.4b-c).

Deleting Rpn6's C-terminal helix led to the appearance of four components: Rpn5/8/9/11, Rpn3/7, Rpn6, and Rpn12 (Fig. 3.3f, Fig. 3.4d). Thus, Rpn6 relies on its own helix for lid binding, whereas Rpn3 and 7 depend not only on their own and each other's helices, but also on that of Rpn6 for incorporation into the subcomplex. This is in agreement with previous pulldown experiments, showing that Rpn6 requires its C-terminal helix for interaction with Rpn7 (Pathare et al., 2011), and with experiments that indicated lid-assembly defects when Rpn6 was C-terminally extended (Isono et al., 2007). Notably, both Rpn3 and Rpn7 were truncated in the Rpn3/7 heterodimer, suggesting that their C-terminal helices were susceptible to proteolytic cleavage in *E. coli* or in lysate when not assimilated into the lid (Fig. 3.4d).

To test the role of the MPN-domain containing subunits for lid assembly, we deleted all three predicted helices at the C-terminus of Rpn11. The truncation caused the majority of lid particles to be separated into Rpn5/6/8/9/11 and Rpn3/7, while the remainder of particles contained all subunits except for Rpn12 (Fig. 3.3g, Fig. 3.4e). We also wanted to assess whether the three predicted C-terminal helices of Rpn11 contribute differently to lid assembly, and therefore deleted only the last one. The phenotype for this variant resembled that for the deletion of all three helices, albeit with a larger fraction of particles only lacking Rpn12 (Fig. 3.4e).

Finally, deleting the three predicted helices at the C-terminus of Rpn8 caused the most severe assembly defect (Fig. 3.3h, Fig. 3.4f). Rpn11 and the truncated Rpn8 were found to be associated as a heterodimer. Similarly, Rpn3 and Rpn7 were bound to each other, whereas Rpn6 and Rpn12 were monomeric and not present in any subassemblies. With our experimental setup, we were unable to distinguish whether Rpn5 and Rpn9 are isolated or form a heterodimer.

Interestingly, during expression and purification of the lid variant with truncated Rpn8, the helices of Rpn11 were partially cleaved off, likely due to their increased accessibility when not assembled with other helices of lid subunits. Rpn8 and Rpn11 thus interacted solely through their MPN domains, making this the first reported observation of the Rpn8-Rpn11 MPN heterodimer. So far, this heterodimer had only been predicted based on domain docking into cryo-EM maps of the proteasome and based on homodimers that were observed in crystal structures of related MPN domains (Beck et al., 2012; Sanches et al., 2007).

Secondary structure prediction software infers a random coil between the end of each PCI domain and the start of the corresponding C-terminal helix (Fig. 3.1a). Many of these linkers between the helical bundle and the PCI domain of individual subunits contain at least one proline residue. We hypothesized that these prolines may rigidify the linkers in order to arrange the PCI domains and facilitate their association in a horseshoe-shaped structure as the helical bundle assembles. To test this hypothesis, we mutated the linker regions of Rpn6 and Rpn7 from ETPN to ASAS and RPDN to ASAS, respectively, and analyzed the assembly phenotypes of these mutant lid constructs. In both cases we observed complete assembly of the lid, suggesting that there is no sequence-specific assembly function of the linkers (Fig. 3.4g). These linkers of 5-20 residues may thus primarily function as flexible tethers between the helical bundle and the horseshoe-shaped structure of PCI domains.

Lid assembly is largely independent of individual PCI and MPN domains

Subnanometer cryo-EM structures indicated that six of the lid subunits form extensive contacts with lateral neighbors through their PCI domains. To determine the role of these interactions in lid assembly, we deleted the PCI domain of Rpn6, which is located in a central position within the horseshoe-shaped PCI arrangement. To our surprise, the presence of only the C-terminal helix of Rpn6 allowed a fraction of lid particles to fully assemble, with even Rpn12 attached (Fig. 3.5a, Fig. 3.4h). The other fraction was found in four pieces: Rpn3/7, Rpn5/8/9/11, Rpn12, and the C-terminal helix of Rpn6, which resembles the assembly defect observed when Rpn6's C-terminal helix is deleted. Based on these observations, we envision a scenario wherein the absence of Rpn6's PCI domain causes the PCI hexamer to lose its rigid horseshoe shape, with the two remaining halves now connected only through the helical bundle, able to swivel independently of one another. Our data suggest that the PCI-domain interactions contribute some of the binding energy between subunits, but lid assembly primarily relies on the bundle formation between their C-terminal helices.

In order to assess the role of the two MPN domains for lid assembly, we removed them simultaneously from Rpn8 and Rpn11. These domain deletions did not disrupt lid formation, demonstrating that the C-terminal helices of Rpn8 and Rpn11 are sufficient for complex assembly, whereas the MPN domains are dispensable for this process (Fig. 3.4h, 3.5b). This settles conflicting hypotheses in recent reviews suggesting that contacts between the MPN and PCI domains are important for lid assembly (Lander et al., 2013; Tomko Jr and Hochstrasser, 2013).

Previous studies have shown that Rpn9 is not essential for cell viability in yeast. The lid purified from an Rpn9 deletion strain had been found to lack only Rpn12 from an otherwise fully assembled subcomplex (Fukunaga et al., 2010). When we deleted the Rpn9 subunit in our recombinant system, the observed assembly phenotype paralleled the *in vivo* data, with lid lacking Rpn9 and only Rpn12 (Fig. 3.4h, Fig. 3.5c). Thus, our results support previous evidence that neither the C-terminal helix nor the PCI domain of Rpn9 is necessary for lid assembly with the exception of Rpn12.

Together, these data prove that the C-terminal helices of the lid subunits are essential and largely sufficient for lid self-assembly. Their association into the helical bundle as a well-defined three-dimensional structure provides most of the binding energy for complex formation. Furthermore, this bundle apparently forms a hub that allows Rpn12 to monitor the overall status of lid assembly. Our biochemical data suggest a number of likely contacts between helices within the bundle. For example, Rpn12 is sensitive to the absence of any other helix, suggesting that it directly interacts with several of them. Furthermore, the helix of Rpn6 is required for incorporation of the Rpn3/7 heterodimer and therefore likely contacts one or both helices of this dimer. Based on the severe assembly phenotype observed upon C-terminal truncation of Rpn8, its helices may occupy a central position within the bundle. Given these biochemical results on bundle organization, we aspired an independent determination of the bundle topology to provide an important structural framework.

Modeling reveals the topology of the lid helical bundle

Previous efforts to understand the bundle arrangement were only able to assign three helices within the twelve-helix bundle, two from Rpn11 and one from Rpn8 (Beck et al., 2012). Determining which helix belongs to which protein is not trivial. For most subunits, the cryo-EM maps did not allow the tracing between helices and the respective MPN or PCI domains. In addition, the number of potential helix configurations in the bundle without constraints is enormous – for 12 helices, there are $12! \times 2^{12}$ (2×10^{12}) possible non-redundant solutions.

In order to alleviate this massive computational problem, we developed a combinatorial search algorithm that recursively matches a set of geometric constraints (GCs) with a set of topology constraints (TCs) (Fig. 3.6). The TCs include the lengths of individual helices as well as the lengths of the linkers between the PCI domains and the bundle, as estimated from the secondary structure prediction of the lid subunits (Buchan et al., 2010) (Table 3.1). The GCs were extracted from a ~7Å resolution cryo-EM map (EMD: 2165 (Beck et al., 2012)), using a tracing algorithm to localize the individual helix densities within the bundle (Rusu et al., 2010) (Table 3.2). Additional GCs were provided by the flexibly fitted structures of the PCI and MPN domains of all lid subunits (PDB id. 4B4T (Beck et al., 2012)). The last Cα atoms of these structures were taken as anchor points for the linkers connecting individual globular domains with the respective helices in the bundle (Table 3.2). Essentially, the problem is reduced to an iterative search of models for which predicted helix and linker lengths are compatible with the physical constraints observed in the cryo-EM map.

Intriguingly, with relatively few constraints and even using generous tolerances, our hybrid approach gave only four solutions (Table 3.3). These solutions shared the same assignment of helices, but differed in the orientations of the last helix of both Rpn8 and Rpn12. To determine the relative directions of these two helices, we further validated the four solutions against all 13 previously observed inter-lysine crosslinks that include at least one residue within the bundle (Table 3.4) (Kao et al., 2012; Lasker et al., 2012). With these additional constraints, we were able to further narrow down the arrangement of helices to a single solution. The resulting topological model accommodated every

reported crosslink, strongly supporting its validity. Furthermore, our helical bundle configuration is consistent with the previously reported placement of three helices from Rpn8 and Rpn11 (Beck et al., 2012), with the linker lengths required to connect a given PCI domain with the respective C-terminal helix in the bundle (Table 3.5), and with the connecting electron densities observed for the linkers of Rpn5 and Rpn7 (Fig. 3.7). Based on our predicted topology, a heavy-atom model of the helical bundle was generated and flexibly fitted into the experimental map with iMODFIT (Fig. 3.9a,b). This model accounts for all density in the helical bundle region, and also fits well into the other previously published sub-nanometer cryo-EM reconstruction of the yeast proteasome (Lander et al., 2012) (Figure 3.8).

The bundle architecture explains assembly defects observed for helix deletions

The helical bundle is composed of twelve helices, with seven of them forming a structure reminiscent of a revolver cylinder (Fig. 3.9c). In this cylinder, a central helix (the 'center pin') is surrounded by six other helices (the 'chambers'), which are slightly angled relative to the central one. The center pin position is occupied by the terminal helix of Rpn8, the longest predicted helix in the bundle. The C-terminal helices of Rpn12, Rpn3, Rpn7, Rpn6, and the final two helices of Rpn11 surround this terminal helix of Rpn8. The remaining five helices of the bundle pack against this seven-helix cylinder and each other (Fig. 3.9a-b.). Importantly, this topology derived by our integrative approach is consistent with the biochemical data presented above.

In our recombinant system, truncation of Rpn8's helices results in a major assembly defect with only Rpn8/11 and Rpn3/7 still interacting through their MPN and PCI domains, respectively (Fig. 3h). Thus, Rpn5 and Rpn9 rely on Rpn8's helices for assembly into the lid. The seven-helix cylinder structure is hollow when missing Rpn8's terminal helix, which serves as the lynchpin for assembly of each of the six surrounding helices (Fig. 3.9c). Consistent with this model, we find that the C-terminal helix of Rpn8 exhibits high hydrophobicity over a ~15 amino acid segment, which is contacted by the six helices of the cylinder according to our derived bundle topology.

The fact that the C-terminal helices of both Rpn3 and Rpn7 are required for stable incorporation of these subunits into the bundle is consistent with the observed positions of these helices, occupying two adjacent "chambers" of the cylinder structure (Fig. 3.3d-e, 3.9c). Though the helix of Rpn7 alone provides enough binding energy to allow a fraction of the Rpn3/7 heterodimer to associate with Rpn5/6/8/9/11, this incorporation is strengthened by the presence of the Rpn3 helix. Our structural model suggests that the interaction of Rpn3's helix with the final helix of Rpn8 provides this additional contact.

The absence of Rpn6's helix causes a similar assembly defect as observed for the truncation of Rpn3 or Rpn7, with the difference that Rpn6 is no longer present in complex with Rpn5/8/9/11. Our structural model is in agreement with this observation, as the helix of Rpn6 makes direct contact with the helix of Rpn7, which itself is necessary for incorporation of the Rpn3/7 heterodimer (Fig. 3.3e).

Based on our topological model, the terminal helix of Rpn11 contacts Rpn8 and Rpn9. However, its truncation caused only a minor assembly defect, indicating that the interactions with Rpn8 and Rpn9 are not integral for the quaternary structure of the bundle (Fig. 3.3g). Deletion of all three C-terminal helices of Rpn11 results in the formation of two lid species. One fraction of subunits is completely assembled, with the exception of Rpn12, while the other fraction assembles only partially into the Rpn5/6/8/9/11 and Rpn3/7 subcomplexes (Fig. 3.3g). Thus, the antepenultimate and penultimate helices of Rpn11 play a role but are not crucial for the interaction between Rpn8's terminal helix and Rpn3/7. Notably, this observation also implies that Rpn6 is able to bind to the lid in the absence of the other five 'chamber' helices (Rpn12, Rpn3, Rpn7, and both helices of Rpn11). The incorporation of Rpn6 is apparently supported by the PCI-PCI interactions with its neighbor Rpn5. This is consistent with the destabilization observed for the lid complex upon deletion of Rpn6's PCI domain.

C-terminal truncations of Rpn5 or Rpn9 did not cause any major lid assembly defects, but prevented the association of Rpn12 (Fig. 3.3b,c). Based on our determined bundle topology, the helix of Rpn12 directly interacts with Rpn9 and thus is able to sense the presence of this subunit. However, there are no observed contacts between Rpn5 and Rpn12, suggesting that Rpn12 may detect Rpn5's presence indirectly through a repositioning of Rpn11's helices or through global conformational changes of the entire bundle induced by Rpn5.

Rpn9 does not rely on a C-terminal helix for incorporation into the lid, but instead seems to depend on interactions between its PCI domain and the helical bundle or the neighboring PCI domain of Rpn5. Interestingly, Rpn5's association with other subunits is not only independent of its C-terminal helix, but also not affected by individually eliminating either of its PCI-domain neighbors, Rpn6 and Rpn9. Additional interactions between the N-terminal portion of Rpn5 and the MPN domain of Rpn11 had been postulated based on our EM structure of the isolated lid (Lander et al., 2012). However, deletion of the Rpn8-Rpn11 MPN-domain dimer also did not disrupt Rpn5's association. Rpn5 thus appears to utilize several different interfaces to assemble with the lid.

The topology of the bundle suggests an elegant mechanism through which the incorporation of Rpn12's helix relies on the helices of every other subunit (Fig. 3.9d). Although the 15-residue helix of Rpn12 contacts the helices of Rpn3, Rpn8, Rpn9, and Rpn11, these interactions likely include no more than two side-chain contacts each. This relieves the dependence of the helical-bundle assembly on the presence of Rpn12, as all helices that contact Rpn12 form many substantial interactions with other helices. In contrast, the helix of Rpn12 makes fewer contacts and thus seems to rely on the presence of all other lid subunits, either through direct interactions or indirectly through the overall bundle conformation. This feature is especially important considering that the C-terminus of Rpn12 has previously been shown to be necessary for efficient association of the completed lid with the base subcomplex to form the regulatory particle (Tomko and Hochstrasser, 2011).

Discussion

Here, we identified a helical bundle as a device that ensures the correct self-assembly of the proteasome lid subcomplex, provides the majority of the binding energy between eight lid subunits, and functions as a hub through which the lid monitors its own completion. The bundle is formed by the conserved C-terminal helices of the MPN- and PCI-domain containing subunits, whose individual truncations revealed an ordered assembly mechanism. Furthermore, using hybrid methodology, we gained critical insights into the topology of the bundle, allowing us to interpret our biochemical data within a structural framework.

Assembly mechanism for the lid

Our findings reveal the striking ability of the large lid subcomplex to self-assemble as well as monitor its proper completion without the help of any specific assembly factors or chaperones. The assembly likely proceeds through an ordered addition of C-terminal helices into a bundle that thus allows a hierarchical complex construction (Fig. 3.10). First, Rpn8 and Rpn11 dimerize through their MPN domains, whose helices then interact with the C-termini of Rpn5 and Rpn9. Next, Rpn6 binds through its C-terminal helix to the nascent Rpn5/8/9/11 assembly, followed by incorporation of the Rpn3/7 heterodimer. As the final assembly piece, Rpn12 binds through its C-terminus to the bundle, allowing the lid to incorporate into the regulatory particle. Although we used an *E. coli*-expression system that lacks posttranslational modifications or other potential factors, our derived mechanistic assembly model is consistent with previously reported endogenous subassemblies isolated from yeast (Fukunaga et al., 2010; Isono et al., 2007; Tomko and Hochstrasser, 2011). Our structural model predicts direct contacts between most of the C-terminal helices, which thus are able to detect each other directly. However, some of the helices appear to indirectly sense others through conformational changes of mutual interaction partners or the entire bundle during lid formation.

The lid assembly mechanism is reminiscent of the four-helix bundle formation harnessed by the SNARE complex to induce vesicle fusion (Hanson et al., 1997). In the case of the SNARE bundle, the helices pack as a four-helix coiled-coil, a state predicted by the SNARE's heptad repeat (Sutton et al., 1998). For the lid subcomplex, Rpn9, 11, and 12 possess helices predicted to form coiled coils, and future structural work at atomic resolution will be required to define the detailed interactions within the lid helical bundle.

The helical bundle acts as a suspended assembly hub

The results presented here show that the helical bundle serves as a hub through which the ~15-residue C-terminal helix of Rpn12 can monitor the complete lid assembly (Fig. 3.9d). Individual truncations revealed that Rpn12 is sensitive to the absence of any single helix within the eight-subunit bundle. The C-terminal location of all the helices may be another important design principle, ensuring that only fully translated subunits are incorporated into the lid.

Macromolecular assemblies often prevent their catalytic activity until complex maturation. Given the importance and diverse roles of ubiquitin signaling in the cell, it

may be necessary to prevent the rogue deubiquitination of proteins by proteasome assembly intermediates outside the holoenzyme context. The deubiquitinase activity of Rpn11 is therefore inhibited until lid incorporation into the regulatory particle (Verma et al., 2002), presumably by allosteric interactions with the helical bundle or the N-terminus of Rpn5 (Beck et al., 2012; Lander et al., 2012). Since efficient lid incorporation depends on the presence of Rpn12, the activation of Rpn11 is elegantly linked to the complete assembly of the helical bundle (Fig. 3.10).

The helical bundle likely has additional functional properties. For instance, it links the MPN-domain dimer and the horseshoe-shaped arrangement of PCI domains through flexible tethers, allowing a certain degree of independent movements of these structural entities. This flexibility may be important for proteasome function, as different portions of the regulatory particle have been observed to undergo significant differential movements during substrate engagement and degradation (Matyskiela et al., 2013). Moreover, the AAA+ ATPases of the base are expected to transition through substantial conformational changes during their cycles of ATP binding, hydrolysis, and nucleotide release, which may require a flexible attachment of the lid. In fact, two significant contacts are made between the bundle and the base. The C-terminal helices of Rpn3 and Rpn8 contact the coiled-coil of Rpt6, while the terminal helix of Rpn11 contacts the OB-fold of Rpt3. Additional studies will be required to define the functional implications of these contacts.

The helix interactions within the bundle appear to provide most of the energy for the assembly of lid subunits, whereas the weaker PCI-PCI interactions may primarily contribute specificity and thus determine the subunit order within the horseshoe-shaped arrangement of PCI subunits. This specificity contribution might be important considering that highly homologous PCI-domain subunits are also present in other macromolecular complexes such as eIF3 and the CSN (Pick et al., 2009; Sun et al., 2011). The requirement of the bundle for lid assembly could thus prevent the incorporation of PCI-containing subunits from other complexes. Furthermore, weaker lateral PCI-PCI interactions may provide flexibility within the horseshoe arrangement and allow certain conformational changes within the regulatory particle during substrate degradation (Matyskiela et al., 2013).

The helical bundle is also interesting from a protein folding perspective, as its twelve helices are contributed by eight different polypeptides. To our knowledge, this is in contrast to all previously described helical bundles of similar size. The association of eight polypeptides into this structure likely imposes a large entropic cost, and how this burden is overcome may be revealed by future studies on the energetics of bundle formation. Additionally, the bundle may have potential for synthetic biology applications and the ordered assembly of designed circuits by attaching short C-terminal helices to protein components.

Lid, CSN, and eIF3 likely employ similar assembly strategies

The CSN regulates the activity of the cullin-RING family of E3 ligases through removal of Nedd8, an ubiquitin-like moiety, from cullin subunits (Cope et al., 2002). eIF3 is a

multisubunit complex that controls loading of mRNA onto the small ribosomal subunit. Given the strong structural and topological homologies between the lid, CSN, and eIF3, we predict that these three complexes assemble by very similar mechanisms using a helical bundle. Like the lid, the CSN and eIF3 each consist of MPN- and PCI-domain containing subunits that, according to secondary structure predictions, all contain C-terminal helices. We thus anticipate that these helices play a vital role in eIF3 and CSN assembly. Interestingly, two recent reports document the *in vivo* and *in vitro* assembly of the CSN strictly requiring the C-terminal helices but not the MPN domain of Csn6 (the paralog of the lid Rpn8) (Kotiguda et al., 2012; Pick et al., 2012). Moreover, deletion of Csn5 (the paralog of the lid Rpn11) from a recombinant expression system for the human CSN allowed an otherwise complete assembly of the complex (Enchev et al., 2012). This result parallels our observation that neither the MPN domain nor the C-terminal helices of Rpn11 are necessary for lid assembly. Interestingly, though, while a truncation of Rpn11 prevents Rpn12 from associating with the lid, the equivalent truncation in the CSN still allows the incorporation of Csn8 (the Rpn12 paralog). This suggests that the CSN may not use its Rpn12 paralog to monitor proper complex formation. Finally, the C-terminus of Csn7 (paralog of the lid Rpn9) has been shown to be necessary and sufficient for binding to Csn6 (the paralog of Rpn8) (Dessau et al., 2008), consistent with our results for the Rpn8/9 interactions. Similar to the lid, the CSN thus seems to utilize a helical bundle for the association of its subunits, and our lid assembly model may provide a valuable framework to explain and interpret previous results for this related complex.

In summary, our present study revealed several important features of the lid subcomplex that are critical for the assembly and function of the 26S proteasome. The helical bundle drives lid formation and acts as a scaffold for monitoring the presence of every constituent within the subcomplex. It functions analogously to assembly factors by establishing an ordered assembly process and ensuring the proper arrangement of subunits. Completion of the bundle allows the lid to be incorporated into the holoenzyme, which has been shown previously to activate proteasomal deubiquitinase activity. Furthermore, we hypothesize that the helical bundle allows the lid to tolerate conformational changes in the regulatory particle during substrate processing. The helical bundle of the proteasome lid therefore exhibits a number of features that are employed by other large protein assemblies, and the related PCI-containing assemblies, CSN and eIF3, likely share these features.

Figures

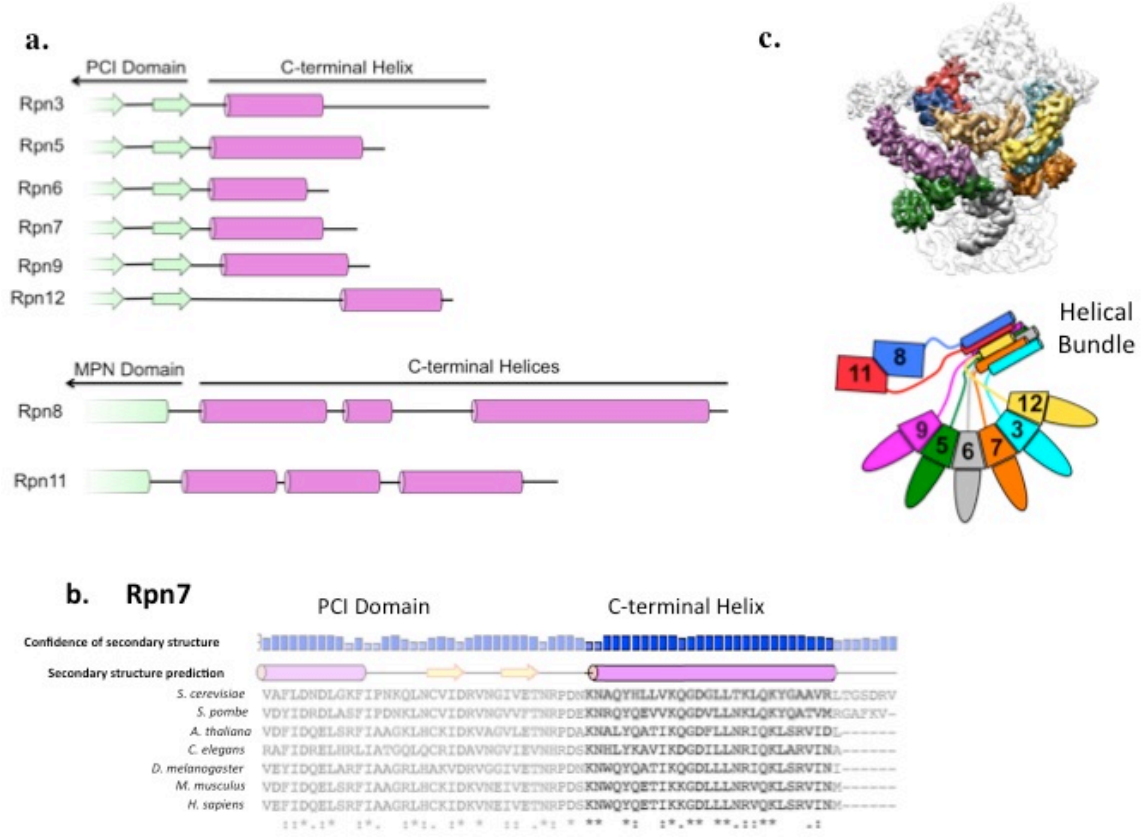


Fig. 3.1: Conserved C-terminal helices of individual lid subunits form a helical bundle. (a) Secondary structure elements at the C-termini of individual lid subunits. β-sheets, random coils, and α-helices are depicted as arrows, lines, and cylinders, respectively. MPN- and PCI-domain structures are indicated above the secondary structure and colored in green, the C-terminal helices are colored in pink. (b) The C-terminal helices are well conserved. As an example, a sequence alignment of Rpn7's C-terminus is shown together with the secondary structure prediction and the prediction confidence. (c) Proposed location of the lid helical bundle (beige) within the proteasome holoenzyme EM reconstruction (Beck et al., 2012). Lid subunits are individually colored and labeled as followed: Rpn3 = cyan, Rpn5 = green, Rpn6 = gray, Rpn7 = orange, Rpn8 = blue, Rpn9 = pink, Rpn11 = red, Rpn12 = gold. A similarly colored cartoon model of the lid is shown below the EM reconstruction.

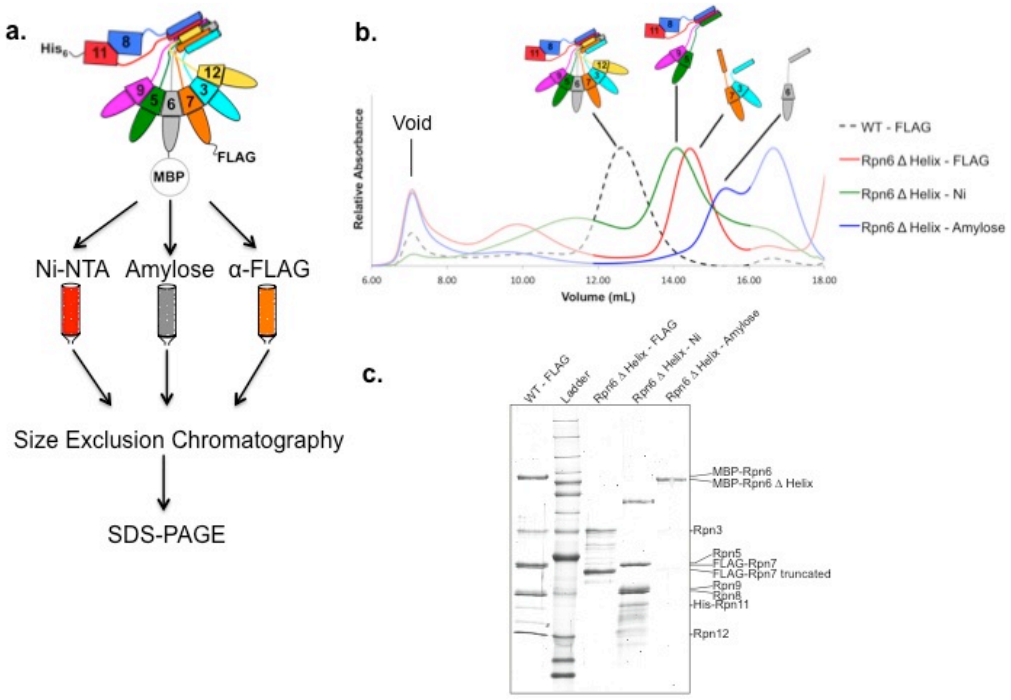


Fig. 3.2: Experimental design for assessing lid assembly. (a) Schematic of the experimental setup. Individual subunits were pulled down from recombinant lid expressions using affinity chromatography (anti-FLAG, Ni-NTA, or amylose), and their assembly status was assessed by size-exclusion chromatography and SDS-PAGE. The schematic representation of the lid shows subunits individually colored as in Fig. 3.1c. (b) Representative size-exclusion chromatography traces after pulling down Rpn11, Rpn7, or Rpn6 from a recombinant lid expression including Rpn6 with a C-terminal helix deletion. Schematics above elution peaks indicate the compositions of subassemblies. (c) Sypro-stained SDS gel showing the lid-subunit compositions present in the main elution peaks of the size-exclusion chromatography after pulling down Rpn11, Rpn7, or Rpn6 as described in (b).

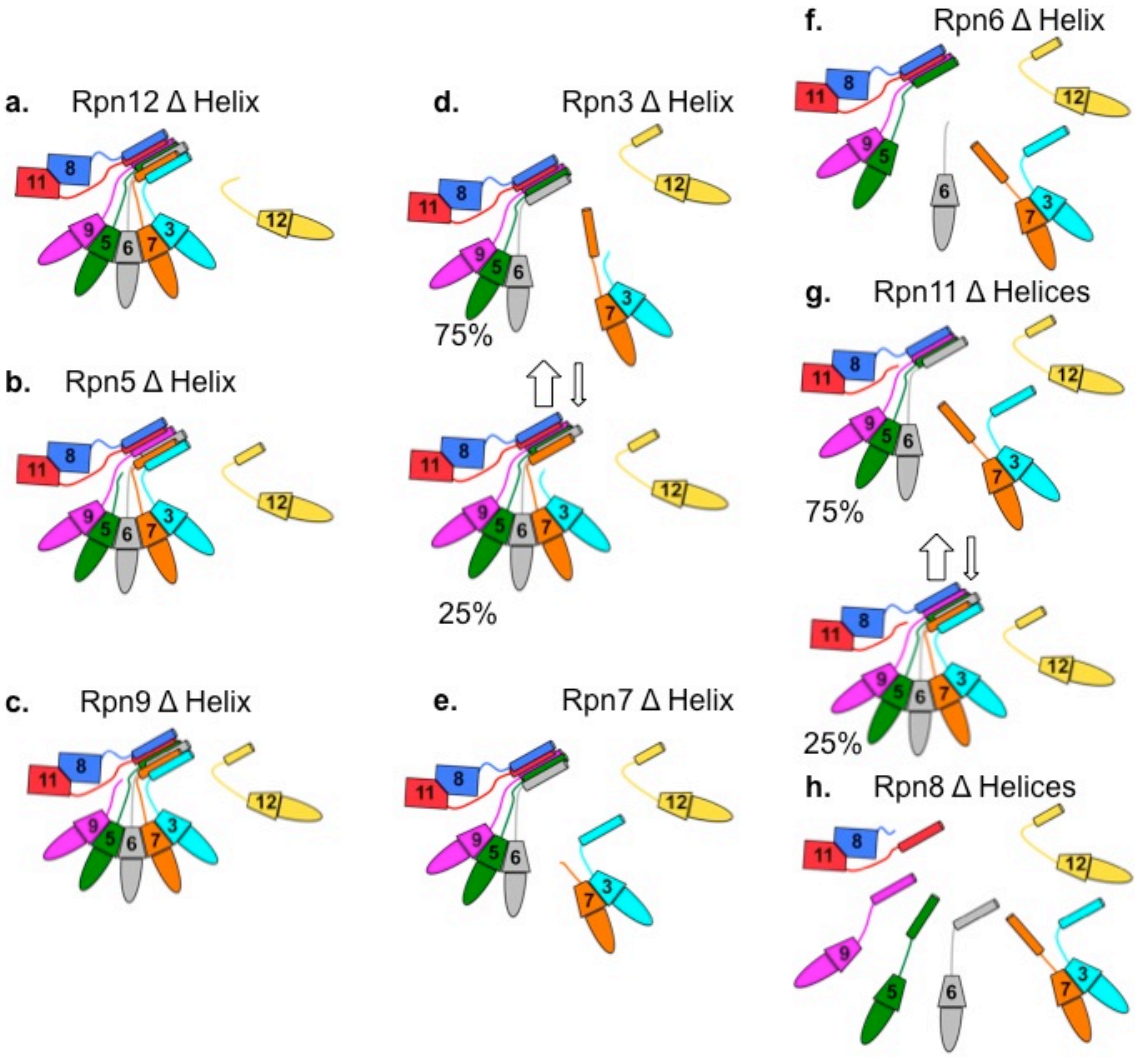
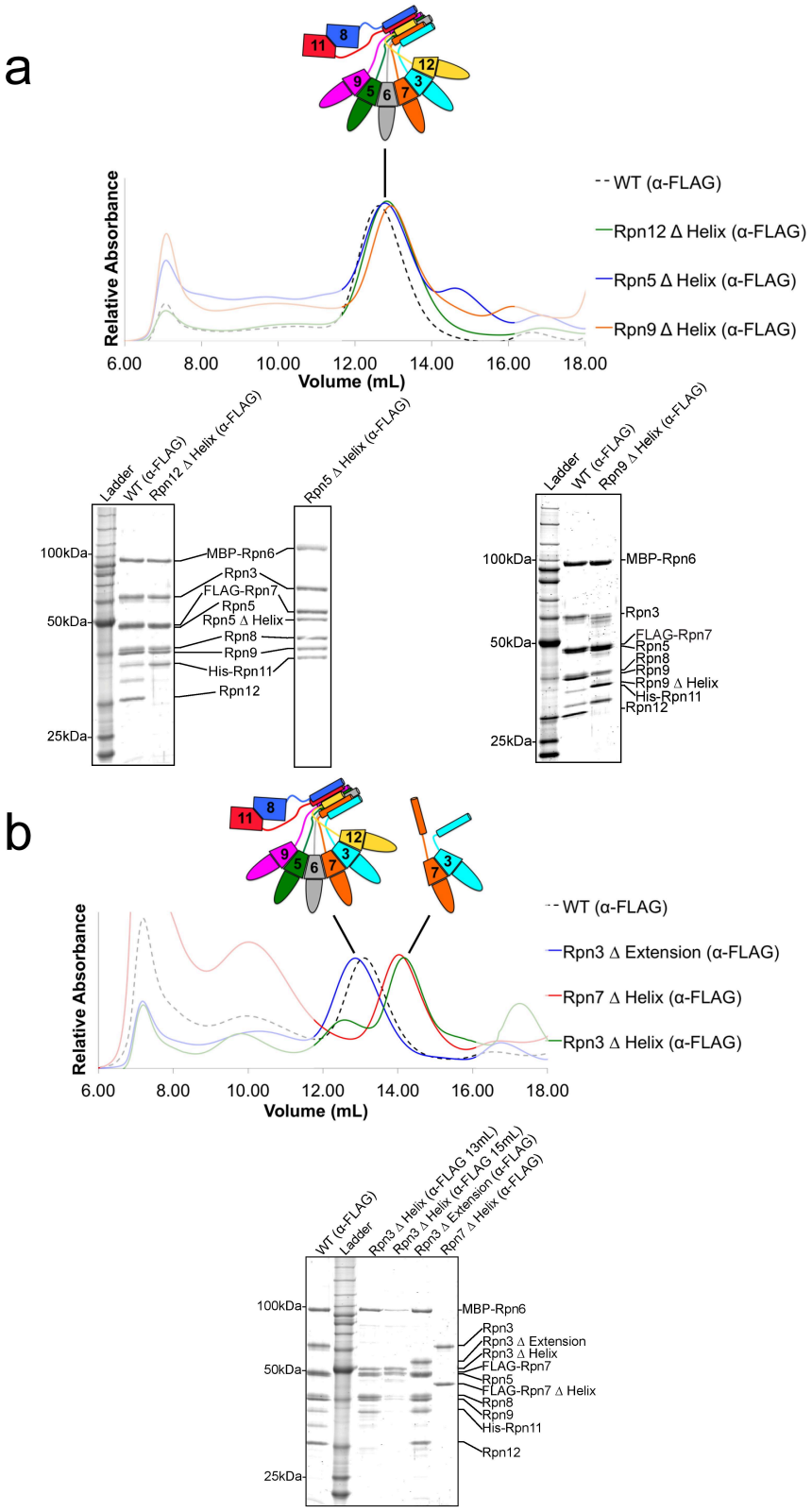
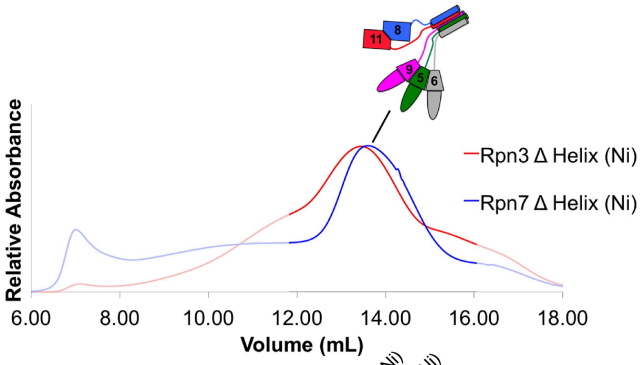
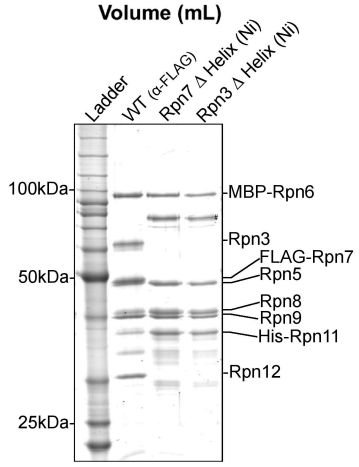


Fig. 3.3: The C-terminal helices of the lid subunits are essential for lid assembly. Panels (a-h) show the schematic representations of the lid subassemblies that were observed for the C-terminal helix truncations of individual subunits as indicated for the detailed analyses of subassemblies see Figure 3.4. Where noted, percentages reflect the relative abundance of different subassemblies. (h) Given our experimental setup and epitope placement, we are unable to distinguish whether Rpn5 and Rpn9 are separated (as shown) or form a heterodimer.

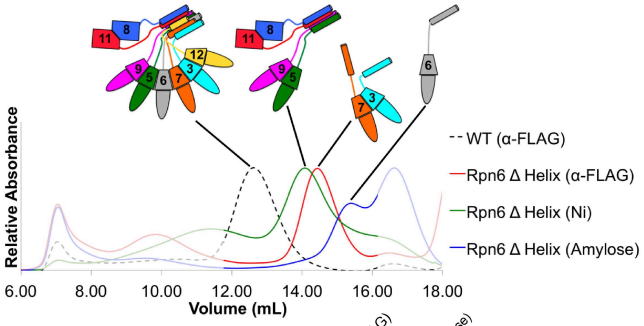

a
11
8
12
9
5
6
7
3
Relative Absorbance
6.00
8.00
10.00
12.00
14.00
16.00
18.00
Volume (mL)
WT (α-FLAG)
Rpn12 Δ Helix (α-FLAG)
Rpn5 Δ Helix (α-FLAG)
Rpn9 Δ Helix (α-FLAG)
Ladder
WT (α-FLAG)
Rpn12 Δ Helix (α-FLAG)
Rpn5 Δ Helix (α-FLAG)
100kDa
50kDa
25kDa
MBP-Rpn6
Rpn3
FLAG-Rpn7
Rpn5
Rpn5 Δ Helix
Rpn8
Rpn9
His-Rpn11
Rpn12
Ladder
WT (α-FLAG)
Rpn9 Δ Helix (α-FLAG)
100kDa
50kDa
25kDa
MBP-Rpn6
Rpn3
FLAG-Rpn7
Rpn5
Rpn8
Rpn9
Rpn9 Δ Helix
His-Rpn11
Rpn12
b
11
8
12
9
5
6
7
3
7
3
Relative Absorbance
6.00
8.00
10.00
12.00
14.00
16.00
18.00
Volume (mL)
WT (α-FLAG)
Rpn3 Δ Extension (α-FLAG)
Rpn7 Δ Helix (α-FLAG)
Rpn3 Δ Helix (α-FLAG)
WT (α-FLAG)
Ladder
Rpn3 Δ Helix (α-FLAG 13mL)
Rpn3 Δ Helix (α-FLAG 15mL)
Rpn3 Δ Extension (α-FLAG)
Rpn7 Δ Helix (α-FLAG)
100kDa
50kDa
25kDa
MBP-Rpn6
Rpn3
Rpn3 Δ Extension
Rpn3 Δ Helix
FLAG-Rpn7
Rpn5
FLAG-Rpn7 Δ Helix
Rpn8
Rpn9
His-Rpn11
Rpn12

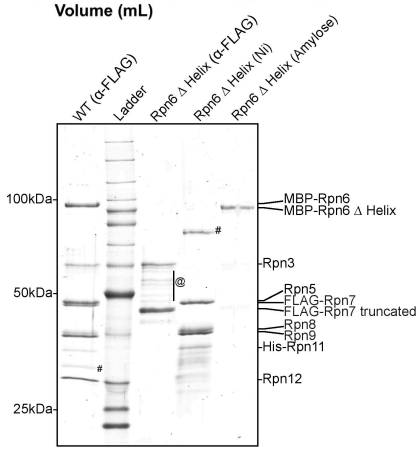
c


11
8
9
5
6
Relative Absorbance
Rpn3 Δ Helix (Ni)
Rpn7 Δ Helix (Ni)
6.00
8.00
10.00
12.00
14.00
16.00
18.00
Volume (mL)

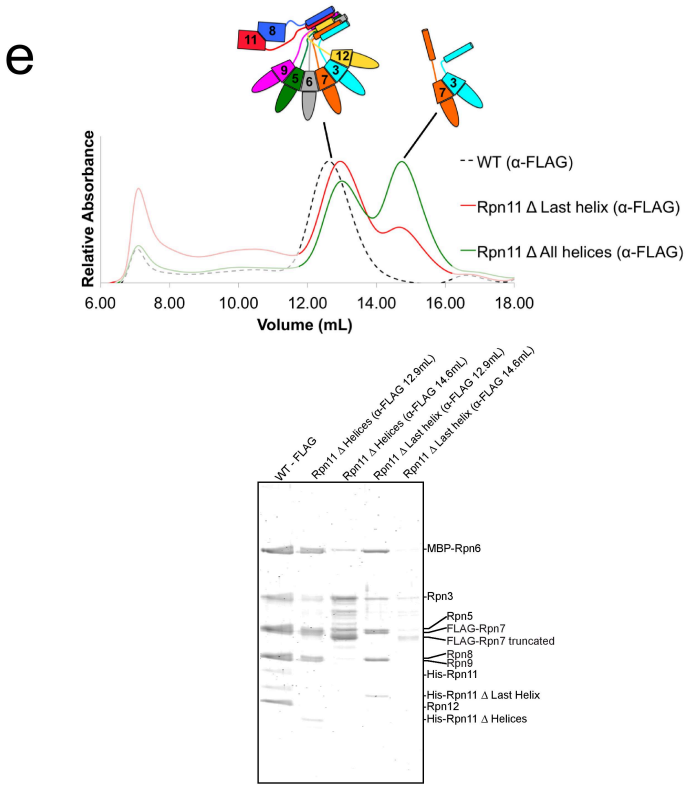

Ladder
WT (α-FLAG)
Rpn7 Δ Helix (Ni)
Rpn3 Δ Helix (Ni)
100kDa
50kDa
25kDa
MBP-Rpn6
Rpn3
FLAG-Rpn7
Rpn5
Rpn8
Rpn9
His-Rpn11
Rpn12

d

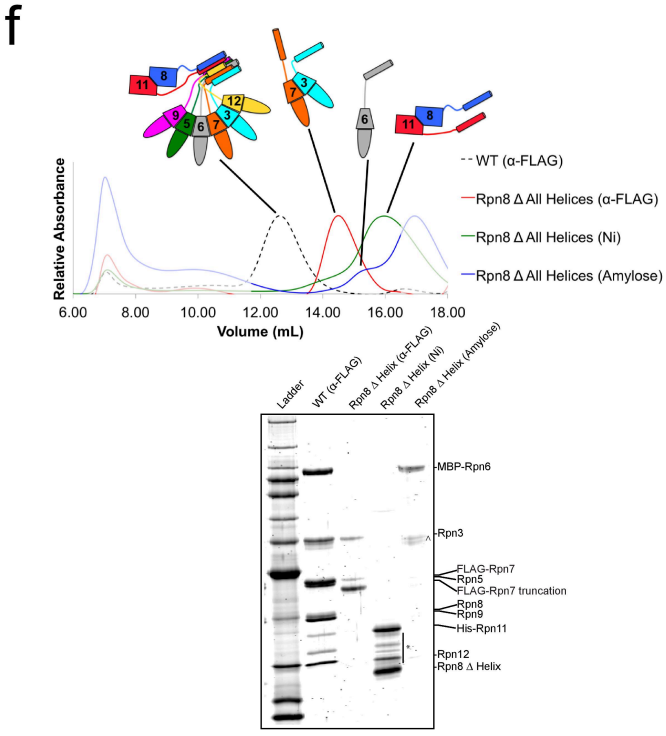

11
8
12
3
9
5
6
7
Relative Absorbance
WT (α-FLAG)
Rpn6 Δ Helix (α-FLAG)
Rpn6 Δ Helix (Ni)
Rpn6 Δ Helix (Amylose)
6.00
8.00
10.00
12.00
14.00
16.00
18.00
Volume (mL)


WT (α-FLAG)
Ladder
Rpn6 Δ Helix (α-FLAG)
Rpn6 Δ Helix (Ni)
Rpn6 Δ Helix (Amylose)
100kDa
50kDa
25kDa
MBP-Rpn6
MBP-Rpn6 Δ Helix
Rpn3
Rpn5
FLAG-Rpn7
FLAG-Rpn7 truncated
Rpn8
Rpn9
His-Rpn11
Rpn12

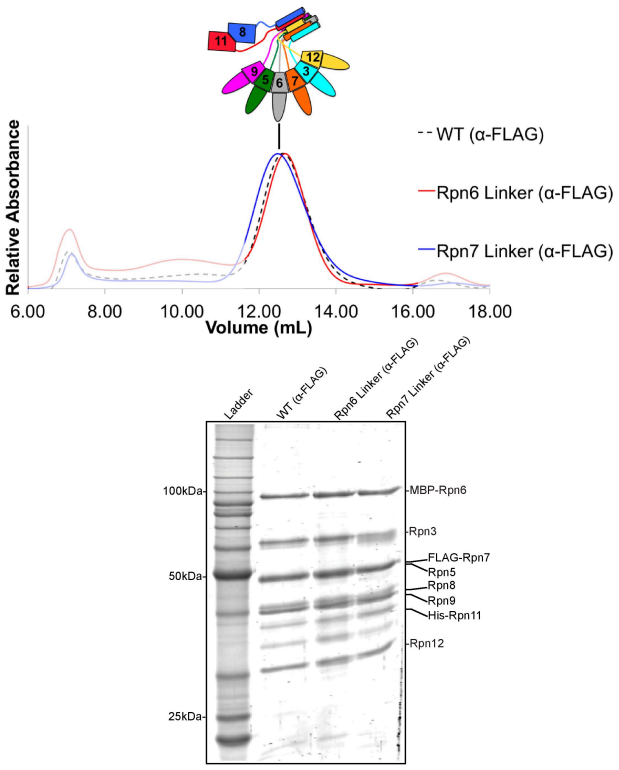
e


Relative Absorbance
6.00
8.00
10.00
12.00
14.00
16.00
18.00
Volume (mL)
WT (α-FLAG)
Rpn11 Δ Last helix (α-FLAG)
Rpn11 Δ All helices (α-FLAG)
WT - FLAG
Rpn11 Δ Helices (α-FLAG 12.9mL)
Rpn11 Δ Helices (α-FLAG 14.6mL)
Rpn11 Δ Last helix (α-FLAG 12.9mL)
Rpn11 Δ Last helix (α-FLAG 14.6mL)
MBP-Rpn6
Rpn3
Rpn5
FLAG-Rpn7
FLAG-Rpn7 truncated
Rpn8
Rpn9
His-Rpn11
His-Rpn11 Δ Last Helix
Rpn12
His-Rpn11 Δ Helices

f


Relative Absorbance
6.00
8.00
10.00
12.00
14.00
16.00
18.00
Volume (mL)
WT (α-FLAG)
Rpn8 Δ All Helices (α-FLAG)
Rpn8 Δ All Helices (Ni)
Rpn8 Δ All Helices (Amylose)
Ladder
WT (α-FLAG)
Rpn8 Δ Helix (α-FLAG)
Rpn8 Δ Helix (Ni)
Rpn8 Δ Helix (Amylose)
MBP-Rpn6
Rpn3
FLAG-Rpn7
Rpn5
FLAG-Rpn7 truncation
Rpn8
Rpn9
His-Rpn11
Rpn12
Rpn8 Δ Helix

g


11
8
12
9
5
6
7
3
Relative Absorbance
WT (α-FLAG)
Rpn6 Linker (α-FLAG)
Rpn7 Linker (α-FLAG)
6.00
8.00
10.00
12.00
14.00
16.00
18.00
Volume (mL)
Ladder
WT (α-FLAG)
Rpn6 Linker (α-FLAG)
Rpn7 Linker (α-FLAG)
100kDa
50kDa
25kDa
MBP-Rpn6
Rpn3
FLAG-Rpn7
Rpn5
Rpn8
Rpn9
His-Rpn11
Rpn12

h

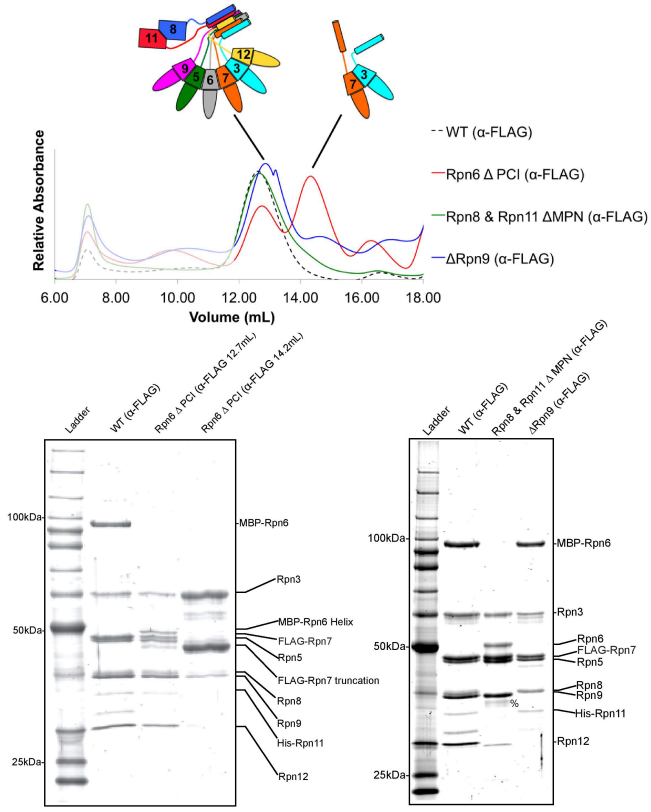

11
8
12
9
5
6
7
3
7
3
Relative Absorbance
WT (α-FLAG)
Rpn6 Δ PCI (α-FLAG)
Rpn8 & Rpn11 ΔMPN (α-FLAG)
ΔRpn9 (α-FLAG)
6.00
8.00
10.00
12.00
14.00
16.00
18.00
Volume (mL)
Ladder
WT (α-FLAG)
Rpn8 Δ PCI (α-FLAG 12.7mL)
Rpn6 Δ PCI (α-FLAG 14.2mL)
100kDa
50kDa
25kDa
MBP-Rpn6
Rpn3
MBP-Rpn6 Helix
FLAG-Rpn7
Rpn5
FLAG-Rpn7 truncation
Rpn8
Rpn9
His-Rpn11
Rpn12
Ladder
WT (α-FLAG)
Rpn8 & Rpn11 Δ MPN (α-FLAG)
ΔRpn9 (α-FLAG)
100kDa
50kDa
25kDa
MBP-Rpn6
Rpn3
Rpn6
FLAG-Rpn7
Rpn5
Rpn8
Rpn9
His-Rpn11
Rpn12

Fig. 3.4: Assembly status of the lid upon various subunit truncations. Gel-filtration traces (top) and Sypro-Ruby stained SDS-PAGE gels (bottom) showing purified recombinant lid with individual subunit truncations. The regions corresponding to the size range of relevant lid subassemblies are highlighted in individual traces. Schematics above the elution peaks indicate the general composition of subassemblies. (a) Assembly defects resulting from the deletion of the helix from Rpn5, Rpn9, or Rpn12. (b,c) Assembly defects resulting from the C-terminal helix truncation of Rpn3 or Rpn7. For the Rpn3 helix deletion in panel (b), the two gel-filtration peaks slightly overlap, which on the SDS gel leads to a weak appearance of bands for subunits present in the neighboring peak. (d) Assembly defect resulting from deletion of Rpn6's helix. e) Assembly defect resulting from deletion of Rpn11's helices. For the Rpn11 helix deletions, the two gel-filtration peaks slightly overlap and thus lead to minor SDS-gel bands corresponding to subunits from the neighboring peak. (f) Assembly defect resulting from deletion of Rpn8's helices. (g) Mutation of the linker regions of Rpn6 and Rpn7 from ETPN to ASAS and RPDN to ASAS, respectively. (h) Deletion of the MPN domains of Rpn8 and Rpn11, PCI-domain deletion of Rpn6, and complete deletion Rpn9. The lid construct lacking the two MPN domains harbors an untagged copy of Rpn6 instead of MBP-Rpn6. # denotes a Ni-NTA specific contaminant. @ and * denote truncated species of Rpn3 and Rpn11, respectively, produced during expression and purification. ^ denotes a truncation product of Rpn6 that is observed during expression in *E. coli*, as we have described previously (Lander et al., 2012).

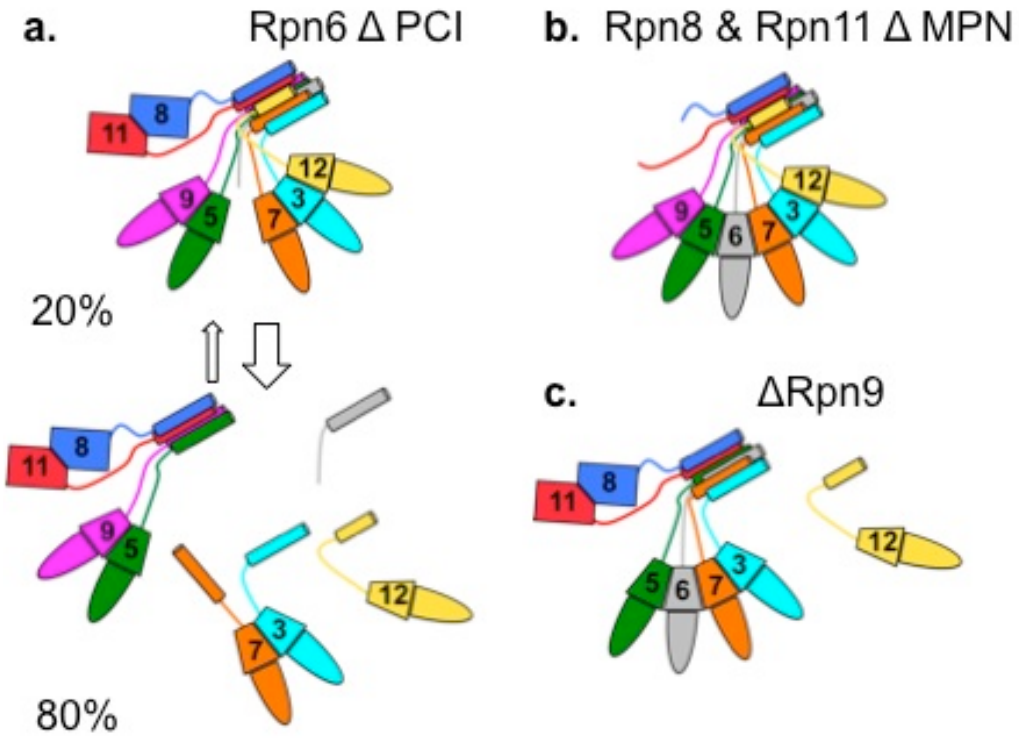


Fig. 3.5: Role of MPN- and PCI domains for lid assembly. (a) The lid is able to assemble even upon deletion of the PCI domain of Rpn6, which is located in the middle of the horseshoe-shaped PCI arrangement. Percentages reflect the relative abundance of two different assembly species for this Rpn6 PCI deletion. (b) The MPN domains are not involved in lid assembly, as indicated by the phenotype observed for the deletion of MPN domains from Rpn8 and Rpn11. (c) The PCI domain of Rpn9 is neither required for assembly of the remaining PCI-arrangement nor the association of the neighboring MPN-domain containing subunit Rpn8.

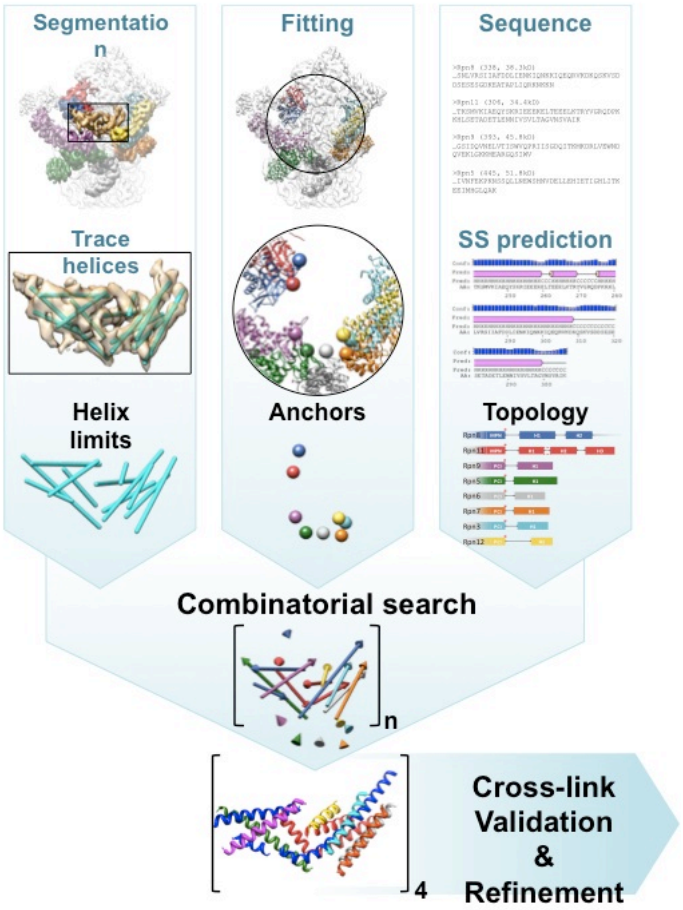


Fig. 3.6: Modeling workflow for determining helical bundle topology. (a) The helical-bundle region (yellow) was segmented from a ~7Å resolution proteasome map (grey transparency, EMDB 2165 (Beck et al., 2012)), with lid subunits individually colored. Helix limits (cyan sticks) were automatically traced by the voltrac tool (Rusu and Wriggers, 2012) from the segmented map (yellow transparency). (b) The anchor points for the C-terminal helix linkers were extracted from the available fitted models of lid subunits (PDB 4B4T (Beck et al., 2012)) and are represented by colored spheres. (c) The topologies of individual *S. cerevisiae* lid subunits were predicted based on their sequences using the PSIPRED server (Buchan et al., 2010). (d) Colored arrows represent one example of the helical-bundle configurations tested during the combinatorial search. Note that the anchor points are now shown as cones pointing towards the connected helix. (e) Resulting models for the helical bundle with individual helices colored corresponding to the respective lid subunits. These models were further validated with existing cross-linking data (Cohn and Hirsch, 1960; Kao et al., 2012; Lasker et al., 2012; Tomko and Hochstrasser, 2011) and finally refined by flexible fitting.

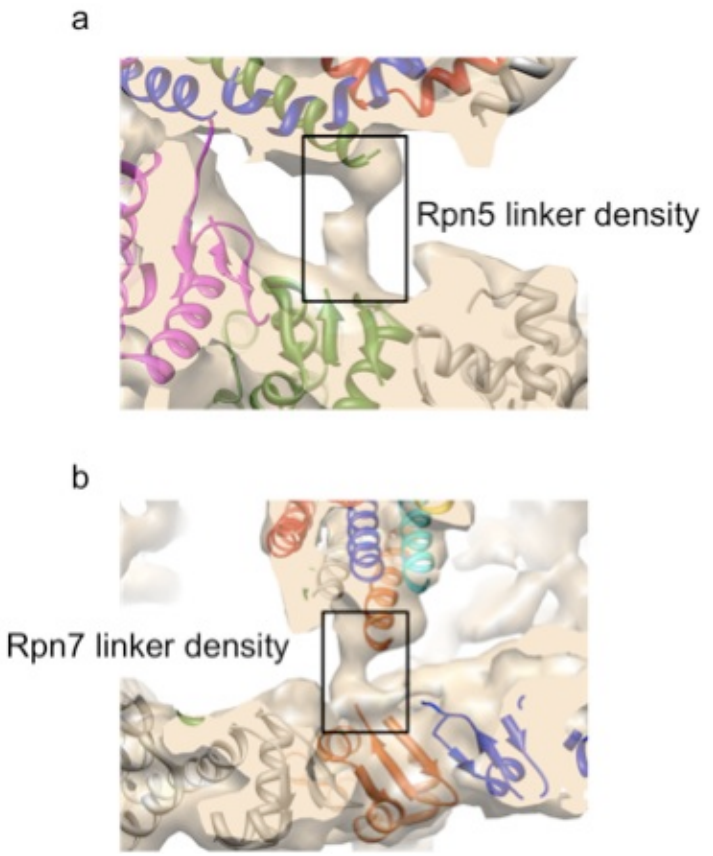


Fig. 3.7: Linker densities observed in the cryo-EM proteasome structure. Predicted heavy atom model for the helical bundle (PDB ID 3J47) and flexibly fitted atomic models of lid subunits (PDB ID 4B4T (Beck et al., 2012)), colored as in Fig. 1c, were docked into the cryo-EM electron density (EMD-2165 (Beck et al., 2012)). This docking revealed the location of linkers between the helix and respective PCI domain of (a) Rpn5 and (b) Rpn7 (linkers indicated by black boxes). Images were generated using the Chimera software package (Goddard et al., 2007).

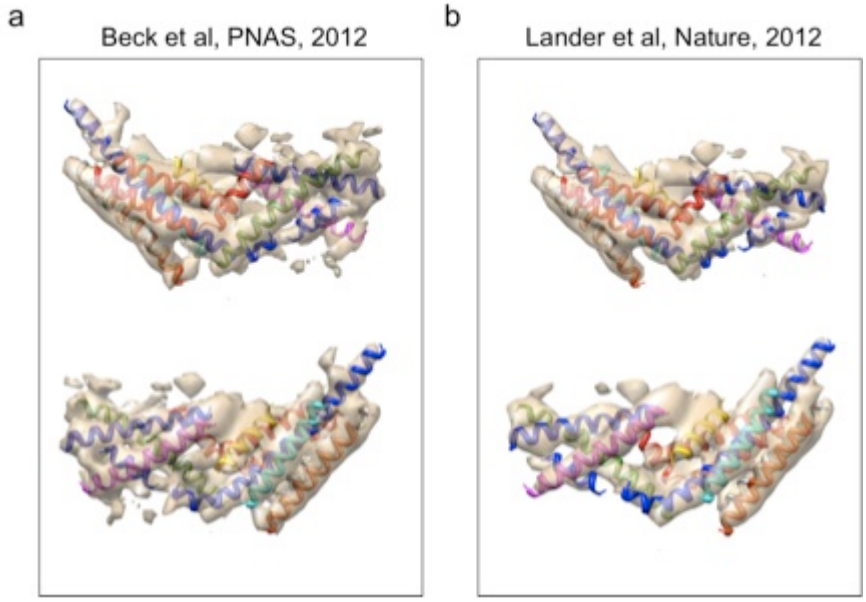


Fig. 3.8: Helical bundle model fit into the high-resolution cryo-EM maps of the *S. cerevisiae* proteasome. (a) The different views of the resulting model reported here flexibly fitted into the EMD 2165 map (Beck et al., 2012). (b) This model also fits well into an independent reconstruction at subnanometer resolution (Lander et al., 2012), except for some missing density for the last helix of Rpn8 and the antepenultimate helix of Rpn11.

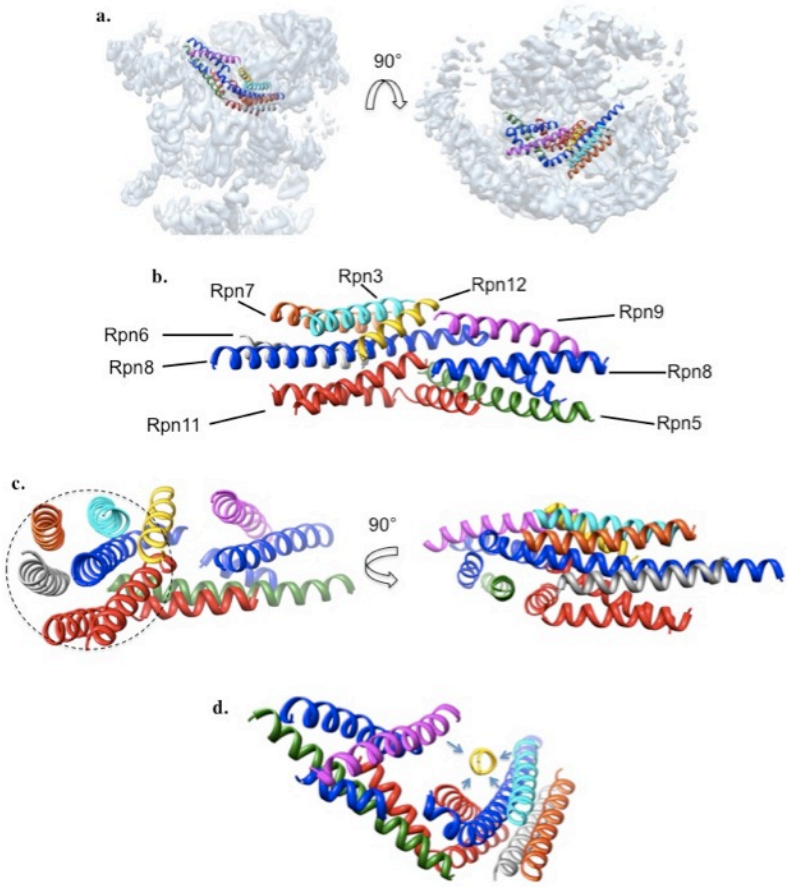


Fig. 3.9: Topological organization of the lid helical bundle. (a) Predicted backbone model for the helical bundle (PDB ID 3J47) fitted into the regulatory particle EM density (Beck et al., 2012), with individual helices colored according to the respective subunit. (b-d) Model for the helical bundle viewed from five different angles, with the left representation in panel (b) showing the fit into the EM density. (c) The helices of Rpn3, Rpn6, Rpn7, Rpn8, Rpn11, and Rpn12 form a shape that resembles the cylinder of a revolver (indicated by a circle), with Rpn8 occupying the 'center pin' position. The axes of the surrounding helices are slightly tilted relative to the central helix of Rpn8. (d) The helix of Rpn12 directly contacts the helices of Rpn3, Rpn8, Rpn9, and Rpn11, as indicated by arrows.

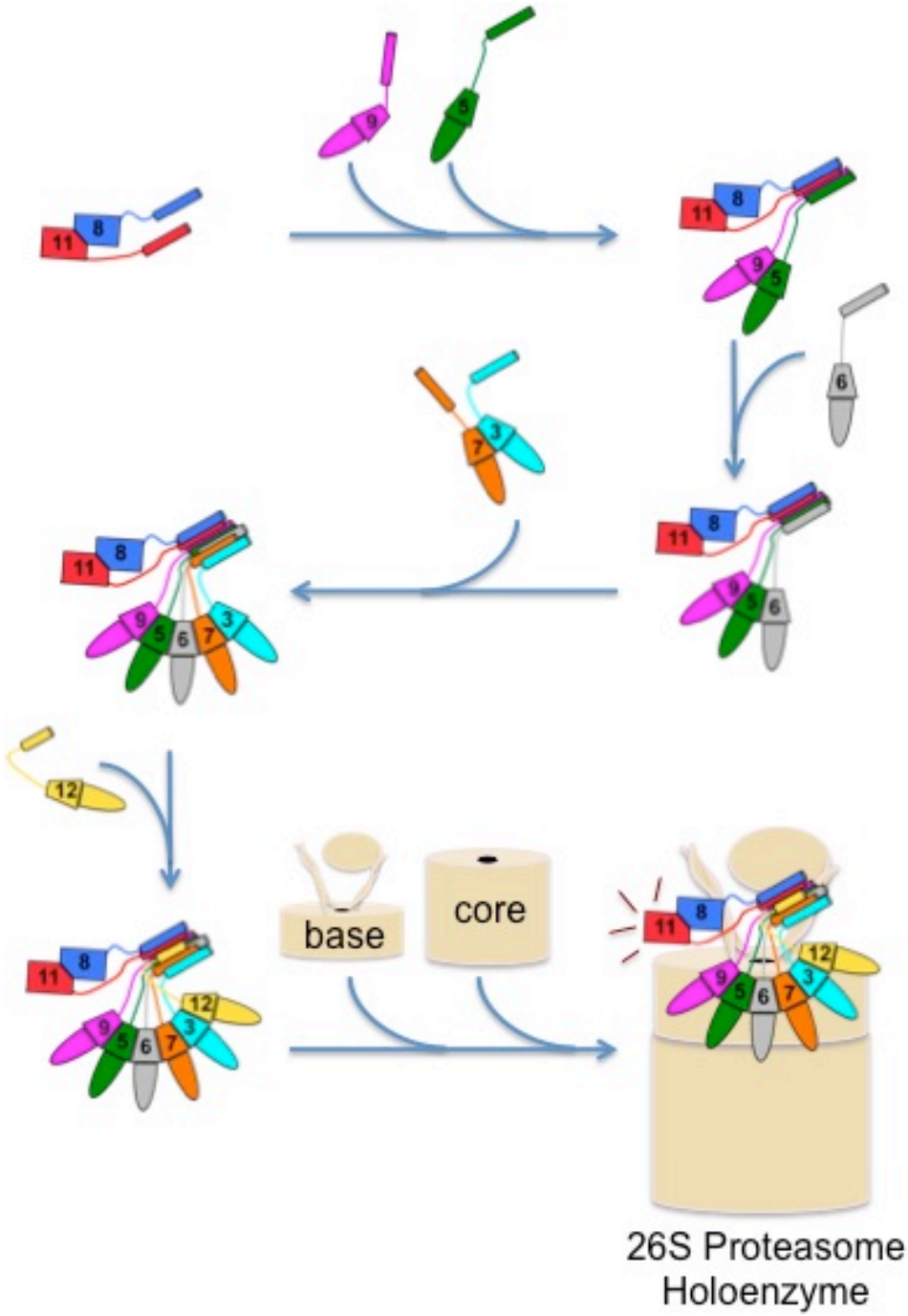


Fig. 3.10: Model for lid assembly and activation of Rpn11. Cartoon models depict the steps of lid assembly, which follows an ordered process guided by the C-terminal helices of individual subunits. Efficient association of the lid, the base, and the core to form the 26S proteasome holoenzyme requires complete lid assembly and activates the deubiquitinase of Rpn11 (indicated by red lines).

Tables

		Linker			Helix		
	Rpn	From	To	Length	From	To	Length
1	8	179	187	9	188	215	28
2	8	216	222	7	223	235	13
3	8	236	258	23	259	308	50
4	11	220	229	10	230	248	19
5	11	249	250	2	251	269	19
6	11	270	275	6	276	298	23
7	9	351	359	9	360	387	28
8	5	403	408	6	409	442	34
9	6	401	406	6	407	430	24
10	7	391	396	6	397	422	26
11	3	446	454	9	455	478	24
12	12	236	255	20	256	272	17

Table 3.1: Topology constraints based on secondary-structure predictions. The secondary structure constraints were obtained from the corresponding subunit sequence using the *PSIPRED* server (Buchan et al., 2010). The first predicted helix for Rpn11 (40 amino acids) has been split in two helices as detailed in the text.

Map Helix #	End A [Å]			End B [Å]			Length	
	x	y	z	x	y	z	[Å]	[#aa]
1	432.082	210.234	301.858	451.664	198.287	340.418	44.9	30
2	413.717	191.747	312.962	451.184	207.548	339.625	48.6	33
3	428.686	192.546	322.749	442.333	200.379	341.679	24.6	16
4	440.181	187.297	336.845	443.026	197.314	307.313	31.3	21
5	433.946	193.675	312.102	413.149	218.219	281.772	44.2	30
6	427.924	218.712	268.572	422.996	186.050	313.982	56.1	38
7	433.934	188.517	302.051	431.624	205.635	290.928	20.5	14
8	410.822	182.971	299.090	427.547	205.428	267.233	42.4	28
9	411.394	195.961	297.613	416.536	210.753	268.405	33.1	22
10	410.781	216.119	273.439	413.078	203.282	304.462	33.7	23
11	415.654	202.827	307.613	436.051	209.583	319.316	24.5	16
12	427.972	204.929	278.863	422.801	184.890	300.925	30.3	20

Table 3.2: Helix limits used for search input.
Individual helices were extracted by tracing the helix bundle region of the cryo-EM map (EMDB 2165 (Beck et al., 2012)) using the *voltrac* utility tool (Rusu et al., 2010). The ends A and B correspond to the end points of the detected helices, except for Rpn9 and Rpn12, whose helices were extended at one end to account for evident addditional helical density. The number of amino acids (#aa) was estimated considering 1.5Å per helix residue.

Lid subunit	Topology constraint		Map Helix Assignment											
			Solution 1			Solution 2			Solution 3			Solution 4		
			#	D	M	#	D	M	#	D	M	#	D	M
8	H1	1	1	F	+2	1	F	+2	1	F	+2	1	F	+2
8	H2	2	3	B	+3	3	B	+3	3	B	+3	3	B	+3
8	H3	3	6	F	-12	6	F	-12	6	B	-12	6	B	-12
11	H1	4	11	B	-3	11	B	-3	11	B	-3	11	B	-3
11	H2	5	10	B	+4	10	B	+4	10	B	+4	10	B	+4
11	H3	6	5	B	+7	5	B	+7	5	B	+7	5	B	+7
9	H1	7	4	F	-7	4	F	-7	4	F	-7	4	F	-7
5	H1	8	2	F	-1	2	F	-1	2	F	-1	2	F	-1
6	H1	9	9	F	-2	9	F	-2	9	F	-2	9	F	-2
7	H1	10	8	F	+2	8	F	+2	8	F	+2	8	F	+2
3	H1	11	12	B	-4	12	B	-4	12	B	-4	12	B	-4
12	H1	12	7	F	-3	7	B	-3	7	F	-3	7	B	-3

Table 3.3: Combinatorial search results.
These results have been obtained using a length tolerance of 30% for the helix and 3 residues for linkers (see methods). The map helix number (#) corresponds to the number given in supplementary table 2. The direction of the assignment (D) is denoted forward (F) when it goes from end A to end B, and backward (B) when it goes from B to A. The helix length mismatch (M) is computed by subtracting the helical number of residues detected in the map from the number of helix-predicted residues.

					Distances [Å]				
					Solution #				Final
#		Cross-links[a]		Type[b]	1	2	3	4	Model[c]
1	Sc	Rpn8:K198	Rpn9:K382/3	HH	11.7	11.7	**11.7**	11.7	**11.9**
2	Sc	Rpn8:K300	Rpn11:K267	HH	41.3	41.3	**18.4**	18.4	**18.1**
3	Sc	Rpn6:K421	Rpn11:K253	HH	19.0	19.0	**19.0**	19.0	**21.5**
4	Sc	Rpn8:K299	Rpn3:K299	HP	41.8	41.8	**20.1**	20.1	**17.3**
5	Sc	Rpn8:K195	Rpn11:K208	HP	25.6	25.6	**25.6**	25.6	**24.9**
6	Sc	Rpn6:K421	Rpn7:K225	HP	25.6	25.6	**25.6**	25.6	**22.6**
7	Sc	Rpn7:K416	Rpn3:K389	HP	34.1	34.1	**34.1**	34.1	**31.0**
8	Sc	Rpn7:K416	Rpn3:K299	HP	8.8	8.8	**8.8**	8.8	**9.6**
9	Sp	Rpn7:K416	Rpn3:S336	HP	20.6	20.6	**20.6**	20.6	**18.0**
10	Sp	Rpn11:S243	Rpt3:K87	HP	33.0	33.0	**33.0**	33.0	**32.9**
11	Sp	Rpn11:K278	Rpt3:K87	HP	24.4	24.4	**24.4**	24.4	**21.9**
12	Sp	Rpn11:R273[d]	Rpt3:K87	HP	24.5	24.5	**24.5**	24.5	**23.3**
13	Sp	Rpn11:R273[d]	Rpt3:K132	HP	17.1	17.1	**17.1**	17.1	**13.3**

Table 3.4: Experimental crosslink validation of combinatorial models from Table 3.3.

[a] Crosslinks observed for either *S. cerevisiae* (Sc) or *S. pombe* (Sp) proteasomes. Crosslinks 1-7 were taken from Kao et al. (Kao et al., 2012) and the rest from Lasker et al. (Lasker et al., 2012). Residue numbers for Sp subunits were adjusted to the equivalent Sc residue numbers using standard multi-sequence alignment.

[b] HH stands for a crosslink between helices of the bundle, while HP refers to a crosslink between a residue of the bundle and a Rpn or Rpt residue outside the bundle.

[c] Crosslink distances after refinement and fitting of the solution model 3.

[d] Since the residue 273 was outside the modeled regions of Rpn11, the closest modeled residue 276 was considered instead.

Rpn	Last residue modeled in PCI or MPN	First residue in C-terminal helix	Linker length [Å]	
			Estimated	Measured
8	178	188	26.1	25.4
11	219	230	29.0	10.5
9	350	360	26.1	25.5
5	402	409	17.4	18.7
6	400	407	17.4	22.2
7	390	397	17.4	17.6
3	445	455	26.1	22.8
12	235	256	58.0	31.8

Table 3.5: Comparison between the predicted and measured linker lengths connecting the C-terminal helices with the respective globular domains of individual lid subunits. The estimated linker length for each Rpn subunit was calculated based on the last modeled residue in the globular domain (PDB ID 4B4T, (Beck et al., 2012)) and the first residue of the C-terminal helix as determined by secondary-structure prediction. A distance of 2.9 Å per residue was considered based on previously used loop-closure algorithms (Chys and Chacon, 2012). Except for Rpn6, the estimated linker lengths are equal to or larger than the measured distances between the globular domains and the C-terminal helices in our topological model of the bundle. The 5-Å deviation for the linker of Rpn6 may arise from uncertainties in the prediction of the starting residue for its C-terminal helix and from a polypeptide length that likely exceeds 2.9 Å per residue for an extended linker.

Chapter 4: Reconstitution of the 26S proteasome reveals functional asymmetries in its AAA+ unfoldase

This chapter is based on work published in Beckwith et al, NSMB, 2013. Robyn Beckwith and I constructed the heterologous expression system for the base, I constructed the pore-2 loop mutants and tested them biochemically, and Robyn Beckwith built and tested the Walker B mutations along with many more mutants not presented here.

Introduction

Previous studies of the related homohexameric unfoldase ClpX suggest that ATP hydrolysis occurs in one subunit at a time with a certain degree of coordination, such that subunits may contribute additively and equally to substrate processing (Martin et al., 2005). However, the homomeric nature of ClpX hinders assessment of whether all six subunits indeed sequentially progress through the different stages of the ATP-hydrolysis cycle. The unique heterohexameric architecture of the proteasomal ATPase ring has thus prompted the fundamental question whether the six Rpts are functionally equivalent or play distinct roles in ATP hydrolysis and substrate processing. While Rpt1–6 share highly homologous AAA+ ATPase domains, they differ substantially in their N-terminal coiled-coil domains, which interact with the laterally bound lid, and in their C-terminal unstructured tails, which mediate interaction with the core particle and trigger gate opening to the proteolytic chamber (Gillette et al., 2008; Kim and DeMartino; Rabl et al., 2008; Smith et al., 2007). Furthermore, as described in Chapter 2, our recent EM structure of the apo 26S proteasome revealed distinct vertical asymmetries within the base ATPase ring (Beck et al., 2012; Lander et al., 2012; Matyskiela et al., 2013; Sledz et al., 2013). In the absence of substrate, the large AAA+ subdomains of Rpt1–6 adopt a pronounced spiral-staircase configuration, with Rpt3 at the top and Rpt2 at the bottom position. Since that work, the lab has determined another sub-nanometer cryo-EM structure of the proteasome, though this time with substrate bound. Strikingly, upon substrate engagement, the base switches to a more planar ring conformation that is characterized by a spiral staircase with Rpt1 at the top and Rpt4 at the bottom (Matyskiela et al., 2013). However, the functional significance of these staircase configurations, potentially manifested as differential subunit contributions to substrate degradation or subcomplex interactions within the holoenzyme, has yet to be determined.

As mentioned in Chapter 1, ClpX contains pore-2 loops that are involved in communicating between the unfoldase and peptidase, inducing a repression of ATPase rate of ClpX upon binding to ClpP (Martin et al., 2007). This loop on ClpX also interacts with substrate, potentially both binding degradation tags and generating force in unfolding. Finally, the pore-2 loop is thought to stabilize the open gate conformation of ClpP. Claims that the C-terminal tails of the base are the only interacting partners between the base and 20S have precluded any search for activation loops in the base, though there is a known activation loop in PA26 (Forster et al., 2005). A conserved pore-

2 loop in the proteasomal AAA+s suggests that this loop may play an important functional role.

Endogenous 26S proteasome has been used to investigate the role of individual Rpts, revealing functional differences in their contributions to base assembly, 26S holoenzyme formation, ATP hydrolysis, peptidase gate opening, and substrate degradation (Erales et al., 2012; Kim et al., 2012; Kohler et al., 2001; Kumar et al., 2010; Lee et al., 2012; Park et al., 2009; Rubin et al., 1998; Smith et al., 2007; Thompson et al., 2009). Despite these results, limitations in working with endogenous proteasome, in part due to *in vivo* assembly problems or lethal degradation defects, have largely prevented extensive systematic studies and a quantitative mechanistic understanding of the individual processes involved in substrate degradation.

Here, we investigated the mechanisms underlying ATP-dependent substrate processing by the heterohexameric unfoldase of the 26S proteasome. To define the differential contributions of individual Rpts to ATP hydrolysis, substrate degradation, peptidase binding, and gate opening, we developed systems for the heterologous expression of the base subcomplex and the *in vitro* reconstitution of partially recombinant proteasomes, and performed systematic mutational analyses of key catalytic and structural motifs.

Methods

Recombinant base expression and purification

Thirteen subunits were cloned into three Novagen vectors including pCOLA-1 (FLAG-Rpt1, Rpt2, His_6-Rpt3, Rpt4, Rpt5, Rpt6), pETDuet-1 (Rpn1, Rpn2, Rpn13), and pACYCDuet-1 (Nas2, Nas6, Hsm3, Rpn14). Each subunit was preceded by a T7 promoter and all plasmids contained one T7 terminator at the end of the multiple cloning sites. Genes for rare tRNAs were also included in the pACYCDuet-1 plasmid to account for differences in codon usage between yeast and *E. coli.* Point mutations in individual ATPase subunits were generated by PCR using pETDuet-1 plasmids containing individual Rpt subunits, which were then used for amplification and substitution into the wild type hexamer in pCOLADuet-1. Base expression strains were generated by co-transforming the pETDuet-1, pCOLA-1 and pACYCDuet-1 plasmids into *E. coli* BL21-star (DE3) cells. The base subcomplex was produced by growing the expression strain to OD_{600}=0.6–0.8 and inducing with 1mM isopropyl-b-D-thiogalactopyranoside overnight at 18°C. Cells were harvested by centrifugation at 5000 rpm for 15 min, resuspended in nickel buffer (25 mM HEPES pH 7.6, 100 mM NaCl, 100 mM KCl, 10 % glycerol, 10 mM $MgCl_2$, 0.5 mM EDTA, 20 mM imidazole) supplemented with 2 mg ml^{-1} lysozyme, protease inhibitors and benzonase. Cells were lysed by freeze-thaw and sonication on ice for 1 min 30 sec in 15 sec bursts. Lysate was clarified by centrifugation at 15,000 rpm at 4°C for 30 min. A two-step affinity purification of the base subcomplex was performed using Ni-NTA agarose (Qiagen) to select for His_6-Rpt3 and anti-FLAG M2 resin (Sigma-Aldrich) selecting for FLAG-Rpt1. 0.5 mM ATP was present in all purification buffers. The Ni-NTA and anti-FLAG M2 columns were eluted with nickel buffer containing 250 mM imidazole or 0.15 mg ml^{-1} 3xFLAG peptide, respectively. The Flag column eluate

was concentrated using a 30,000 MWCO concentrator (Amicon) and run on a Superose6 gel filtration column (GE Healthcare) equilibrated with gel filtration buffer (60 mM HEPES pH 7.6, 50 mM NaCl, 50 mM KCl, 10 % glycerol, 5 mM $MgCl_2$, 0.5 mM EDTA, 1 mM DTT, 0.5 mM ATP).

Purification of yeast holoenzyme and subcomplexes

Yeast holoenzyme, core particle, base and lid subcomplexes were purified from *S. cerevisiae* essentially as previously described(Leggett et al., 2005). Frozen yeast cells were lysed using a Spex SamplePrep 6870 Freezer/Mill. Holoenzyme was purified from a yeast strain containing FLAG-Rpn11. Lysed cells were resuspended in lysis buffer containing 60 mM HEPES pH 7.6, 100 mM NaCl, 100 mM KCl, 10 % glycerol, 5 mM $MgCl_2$, 0.5 mM EDTA, 0.2 % NP-40 and ATP regeneration mix (5 mM ATP, 0.03 mg ml^{-1} creatine kinase, 16 mM creatine phosphate). Holoenzyme was bound to anti-FLAG M2 resin and washed with wash buffer (60 mM HEPES pH 7.6, 100 mM NaCl, 100 mM KCl, 10 % glycerol, 5 mM $MgCl_2$, 0.5 mM EDTA, 0.1 % NP-40, 0.5 mM ATP). Holoenzyme was eluted with 0.15 mg ml^{-1} 3xFLAG peptide and further purified by gel filtration using a Superose-6 column with gel filtration buffer (see above). Lid and base subcomplexes were isolated from FLAG-Rpn11 or FLAG-Rpn2 yeast strains, respectively, and purified by exposure to a 1 M NaCl wash while bound to anti-FLAG M2 resin. Base purification buffers included 0.5 mM ATP. Core particle was purified from a 3xFLAG-Pre1 yeast strain using a 500 mM salt wash. All subcomplexes underwent size exclusion chromatography using a Superose-6 column as described above.

Yeast strains

Yeast lid and holoenzyme were purified from strain YYS40 (genotype *MAT***a** *ade2-1 his3-11,15 leu2-3,112 trp1-1 ura3-1 can1 Rpn11::Rpn11-3XFLAG(HIS3), source Y. Saeki)*. Core particle was prepared either from strain RJD1144 (genotype *MAT***a** *his3Δ200 leu2-3,112 lys2-801 trpΔ63 ura3-52 PRE1-FLAG-6xHIS::Ylpac211(URA3) source R. Deschaies*) or strain yAM14 (genotype *MAT***a** *ade2-1 his3-11,15 leu2-3,112 trip1-1 ura3-1 can1-100 bar1 PRE1::PRE1-3XFLAG(KanMX*), this study).

Native gel electrophoresis

Analysis of proteasome holoenzyme and subcomplexes by native gel was performed as described previously(Leggett et al., 2005). Assembly reactions were incubated for 15 min at 23°C with 5 mM ATP, followed by electrophoresis on a 3.5% native polyacrylamide gel. Electrophoresis was conducted at 4°C with stirring and running buffer containing 0.5 mM ATP. The gel was overlaid with developer solution (running buffer with 100 nM Suc-LLVY-AMC peptide and 0.02% SDS) and incubated at 30°C for 10 minutes prior to imaging. Fluorescence imaging was performed using a Typhoon scanner (GE Healthcare) and followed by Coomassie staining.

ATPase and peptidase stimulation assays

ATPase activity was quantified using an NADH-coupled ATPase assay. 500 nM base was incubated with 1x ATPase mix (3 U ml^{-1} pyruvate kinase, 3 U ml^{-1} lactate

dehydrogenase, 1 mM NADH, 7.5 mM phosphoenol pyruvate) at 30°C. Absorbance at 340 nm was monitored for 900 sec at 10 sec intervals using a UV-Vis Spectrophotometer (Agilent). Peptidase stimulation was monitored by following the increase in fluorescence resulting from cleavage of a fluorogenic peptide substrate(Glickman et al., 1998b), Suc-LLVY-AMC (Boston Biochem), using a QuantaMaster spectrofluorimeter (PTI). 50 nM core particle was incubated with saturating base subcomplex in the presence of an ATP regeneration system (5 mM ATP, 16 mM creatine phosphate, 6 mg ml^{-1} creatine phosphokinase) and 50 μM Suc-LLVY-AMC. Titration experiments were conducted using either 25 nM core particle and increasing amounts of base (peptidase stimulation) or 100 nM base and increasing amounts of core particle (ATPase activity). K_D values were extracted by fits to a simple binding curve using Grafit (Erithacus Software).

Multiple and single turnover GFP substrate degradation

Proteasome holoenzyme was reconstituted from core particle, lid, base and Rpn10. A GFP-titin^{V15P}-cyclin-PY fusion protein was modified *in vitro* with a polyubiquitin chain using Uba1, Ubc1, Rsp5 and wild-type ubiquitin. Degradation reactions were performed at 30°C in gel filtration buffer (60 mM HEPES pH 7.6, 50 mM NaCl, 50 mM KCl, 10 % glycerol, 5 mM $MgCl_2$, 0.5 mM EDTA, 1 mM DTT, 0.5 mM ATP) supplemented with an ATP regeneration system. Details for the gel-based degradation assay can be found in the legend of Supplementary Fig. 2. Single and multiple turnover degradation activities were monitored by the loss of GFP fluorescence (excitation 467 nm; emission 511 nm) using a QuantaMaster spectrofluorimeter (PTI). Multiple turnover degradation experiments were performed with 50 nM reconstituted holoenzyme under V_{max} conditions (saturating base, lid and Rpn10) with 2 μM substrate. Excess base, lid, and Rpn10 did not affect the observed degradation rate (see Supplementary Fig. 3 for lid and Rpn10). For single turnover degradation, 2 μM holoenzyme was reconstituted from equimolar concentrations of subcomplexes and mixed with 100 nM substrate. Single turnover degradation traces fit with double exponential decay curves using Grafit (Erithacus Software)

Affinity pulldowns

Base subcomplexes were mixed with core particle, lid and Rpn10 in nickel buffer with 1x ATP regeneration system at 23°C for 15 min. All components were present at a concentration of 900 nM in the reaction, which was then incubated with 5 μl of magnetic Dynabeads (Invitrogen) at 23°C for 15 min. The beads were washed three times with nickel buffer supplemented with 0.05% NP-40 and 0.5 mM ATP. Bound proteins were eluted with nickel buffer containing 500 mM imidazole. Pulldown samples were run on a 10% SDS-PAGE gel, stained with Sypro Ruby and imaged using a Typhoon scanner (GE Healthcare).

Results

Heterologous expression of the base subcomplex

The base subcomplex of the proteasome from *Saccharomyces cerevisiae* was produced in *Escherichia coli* by co-expression of thirteen yeast proteins, including nine integral base subunits (Rpt1–6, Rpn1, Rpn2, Rpn13) and four proteasome assembly chaperones (Nas2,

Nas6, Rpn14, Hsm3 (Funakoshi et al., 2009; Kaneko et al., 2009; Roelofs et al., 2009; Saeki et al., 2009)). We isolated assembled base by tandem-affinity purification, using tags on two different subunits, followed by gel-filtration chromatography. The purified base exhibited appropriate stoichiometry and no subunit truncations, as revealed by SDS-PAGE (Fig. 1a) and mass spectrometry. We observed Nas6, Hsm3, and Rpn14 stably associated with the recombinant base, whereas these chaperones were not present in the base purified from yeast, as indicated by SDS-PAGE, native PAGE (Fig. 1b), and size-exclusion chromatography (Supplementary Fig. 1). This result is consistent with studies of *in vivo* proteasome assembly, indicating that Nas6, Hsm3, and Rpn14 are displaced upon base binding to the core particle and lid, whereas Nas2 dissociates at an earlier stage of base assembly (Funakoshi et al., 2009; Park et al., 2013; Park et al., 2009; Roelofs et al., 2009). One model for *in vivo* base assembly proposes that the core particle might act as a template to facilitate the proper arrangement of Rpts in the hexameric ring(Park et al., 2009). However, our successful constitution of the base subcomplex in *E. coli* rules out a strict requirement for such templated assembly.

We compared the activities of the recombinant base to endogenous yeast base. Both base subcomplexes hydrolyzed ~ 51 ATP enz^{-1} min^{-1} in the absence of substrate (Table 1). The ability of the ATP-bound base to interact with core particle and induce gate opening was determined by monitoring the fluorescence increase upon peptidase cleavage of the fluorogenic peptide Suc-LLVY-AMC. In the presence of ATP, recombinant base stimulated core-particle activity approximately 20-fold, similar to endogenous yeast base. In agreement with previous reports, we measured about two-fold higher peptide hydrolysis with the non-hydrolysable analog ATPγS compared to ATP (Liu et al., 2006; Smith et al., 2005), which may be due to potential differences in the ATPase-ring conformation (Sledz et al., 2013) or the dynamics of base-core interactions.

Importantly, we also reconstituted 26S holoenzyme, using either endogenous or recombinant base and the lid and core particle purified from yeast. Successful reconstitution was assessed by native PAGE (Fig. 1b) and *in vitro* degradation of a polyubiquitinated model substrate (Supplementary Fig. 2), a green fluorescent protein (GFP)-titin^{V15P}-cyclin-PY fusion, whose degradation could be measured through the decrease of GFP fluorescence (Fig. 2). Proteasomes reconstituted with saturating recombinant or endogenous base degraded substrate at a maximal rate of 0.3 enz^{-1} min^{-1}, comparable to 0.32 enz^{-1} min^{-1} observed for holoenzyme purified from yeast (Table 1). Substrate degradation by reconstituted proteasomes strictly required addition of recombinant Rpn10, an intrinsic ubiquitin-receptor that does not co-purify with isolated lid or base subcomplexes. Consistent with previously described degradation defects in the absence of Rpn10 (Verma et al., 2004), we found that omitting Rpn10 or deleting its ubiquitin-interacting motif resulted in 40-fold slower degradation (Fig. 2a, Supplementary Fig. 3), despite the presence of the second ubiquitin receptor, Rpn13. Since proteasome formation did not depend on Rpn10 and degradation was not facilitated by Rpn10 lacking its ubiquitin-interacting motif, this result suggests that Rpn10 is either the primary receptor for our model substrates or required in combination with Rpn13 for multivalent ubiquitin-chain binding.

Functional asymmetry of the heterohexameric AAA+ unfoldase

To examine the roles of Rpt1–6 in nucleotide-dependent substrate processing, we individually abolished their ATP hydrolysis by systematically introducing a catalytic mutation in the recombinant base. In the homohexameric bacterial unfoldase ClpX, mutation of the conserved Walker-B glutamate prevents hydrolysis and induces a permanently ATP-bound state in the mutated subunit (Hersch et al., 2005), but other AAA+ unfoldases require distinct Walker-B mutations to eliminate ATP-hydrolysis activity (Gomez et al., 2002; Weibezahn et al., 2004). We therefore tested the effects of various substitutions of the conserved Walker-B aspartate and glutamate residues by simultaneously placing them in all six Rpts. Ultimately, mutation of glutamate to glutamine (EQ) allowed proper assembly of base that exhibited wild-type levels of peptidase-binding and gate-opening activities despite being inactive in ATP hydrolysis (Table 1), indicating that this mutation in fact traps Rpt subunits in a permanently ATP-bound state. We next introduced a single EQ mutation per hexamer to fix individual Rpts in the ATP-bound state and test their contributions to ATP hydrolysis as well as core-gate opening. Depending on which subunit was mutated, we observed considerable differences in activities (Table 1), indicating that the Rpt subunits are functionally non-equivalent.

ATP hydrolysis by the isolated base decreased by more than 60% upon mutation of Rpt3, whereas inactivating other subunits had either minor effects (Rpt1, Rpt4, Rpt5) or notably increased the hydrolysis rate (Rpt2, Rpt6). Peptidase stimulation by these base mutants also varied: mutated Rpt5, Rpt2, and Rpt6 caused 20%, 30%, and 60% stronger gate opening compared to wild-type base, mutated Rpt4 did not lead to noticeable changes, while mutations in Rpt3 or Rpt1 decreased gate-opening activity by 20% or 30% despite proper complex formation with the core particle. It is important to consider that the measured ATPase and gate-opening activities reflect an average of the unmutated five Rpt subunits and are influenced by subunit communication. The increase in ATPase activity resulting from mutation of Rpt2 and Rpt6, for instance, is likely caused by the response of neighboring subunits to a permanently ATP-bound state. Some of these stimulated ATP-hydrolysis events may be non-productive and thus not result in increased substrate-degradation rates. This is supported by the fact that most base variants containing a mutant Rpt showed no stimulation or even slight repression in ATP-hydrolysis activity upon the addition of substrate, whereas the ATPase rate of wild-type base approximately doubled (Supplementary Fig. 4).

To explore the distinct roles of individual Rpt subunits in substrate processing, we reconstituted proteasomes with base variants containing single-subunit EQ mutations and compared their rates of ubiquitin-mediated substrate degradation under multiple-turnover conditions (Fig. 3a). An EQ mutation in either Rpt3 or Rpt4 completely eliminated substrate degradation, and a mutated Rpt6 resulted in a 90% decrease in degradation rate. Mutations in Rpt1 and Rpt5 lowered the degradation rate by 73% and 56%, respectively, whereas the Rpt2 mutant showed no defect. Importantly, the observed degradation defects were not simply the result of compromised proteasome assembly, as all mutants stimulated peptidase-gate opening (Table 1) and bound lid and core particle (Supplementary Fig. 5a,b). Considering the order of ATPase subunits within the base

(Rpt1–2–6–3–4–5), it is strikingly evident that mutants with severe degradation defects (Rpt6, Rpt3, and Rpt4) map to one half of the ring.

Next we investigated the mechanistic role of individual Rpts at different stages of substrate processing by performing single-turnover degradation experiments (Fig. 3a, Supplementary Fig. 5c). In contrast to steady-state analyses, these measurements discriminate between potential defects in substrate engagement versus translocation and unfolding. The resulting data for GFP-substrate turnover were best fit by a double-exponential decay. Since GFP loses fluorescence in a single unfolding step, this double-exponential behavior was likely due to two types of substrates that probably varied in their ubiquitin tagging and were degraded at different rates. As expected, the two rates averaged to roughly match the multiple-turnover rate, and the base mutants exhibited the same ranking of activities as in multiple-turnover degradation (Fig. 3a, Table 1). Even under these single-turnover conditions, Rpt3 or Rpt4 mutants did not show any measurable degradation, whereas the Rpt6 mutant degraded substrate at 5% of the wild-type rate. This Rpt6 mutant exhibited a short lag preceding the exponential fluorescence decay (Supplementary Fig. 5c), which may indicate delayed substrate translocation and unfolding due to defects in engagement. The 95% decrease in degradation rate may thus originate from slower engagement instead of or in addition to compromised unfolding and translocation. The complete lack of substrate degradation for ATPase-deficient Rpt3 or Rpt4 may similarly be a consequence of severe engagement defects.

Spiral staircase configurations of the base

EM reconstructions of the ATP-bound 26S proteasome in the absence of substrate revealed that the large AAA+ subdomains of the Rpts adopt a pronounced spiral-staircase configuration around the ring(Lander et al., 2012; Lasker et al., 2012) (Fig. 3b). Rpt3 occupies the highest and Rpt2 the lowest position relative to the core particle, with Rpt6 bridging the vertical gap between the two subunits. The degradation defects of EQ mutants and the inferred contributions of individual Rpts to substrate degradation largely correlate with the vertical positions of subunits in this pre-engaged staircase. Conformational changes in subunits at the top of the apo spiral (Rpt3, Rpt4 and Rpt6) thus appear to be critical to engage a substrate and initiate translocation, whereas subunits in lower positions are less important.

Our recent EM reconstruction of the translocating proteasome demonstrated that substrate engagement induces substantial conformational changes in the regulatory particle, leading to a more planar ATPase-ring with a rearranged staircase of pore loops, in which Rpt1 is at the top and Rpt4 at the bottom position (Fig. 3b) (Matyskiela et al., 2013). Based on the static appearance of the spiral, in combination with single-molecule data for the related protease ClpXP, we previously proposed that the substrate-engaged staircase represents a default dwell state that is adopted by the base before or after coordinated ATP-hydrolysis events progress around the ATPase ring (Matyskiela et al., 2013). Interestingly, in the present study we observed an approximately 70% reduction in degradation activity when ATP hydrolysis was eliminated in Rpt1, located at the top of the substrate-engaged staircase. Rpt1 may therefore play a special role in processive

substrate translocation, possibly by triggering the coordinated ATP hydrolysis of subunits.

Differential contributions of pore loops

Previous work on various AAA+ unfoldases proposed that a conserved aromatic-hydrophobic (Ar-Φ) loop protrudes from every ATPase subunit into the central channel and undergoes nucleotide-dependent power strokes to drive substrate translocation (Aubin-Tam et al., 2011; Erales et al., 2012; Hinnerwisch et al., 2005; Maillard et al., 2011; Martin et al., 2008a, b; Park et al., 2005; Wang et al., 2001; Yamada-Inagawa et al., 2003).

The pore-2 loop, located below the Ar-Φ loop in the central channel and directly adjacent to the Walker-B motif, has been implicated in substrate binding and unfolding as well as peptidase interaction and the control of ATPase activity in ClpX (Glynn et al., 2009; Martin et al., 2007, 2008a). A recent study on the archaeal homohexameric unfoldases Cdc48 and PAN suggested that the pore-2 loop also interacts with the proteasomal core particle (Barthelme and Sauer, 2013). As the importance of the pore-2 loop for 26S-proteasome function is still unknown, we replaced the highly conserved aspartate with asparagine (DN) in the pore-2 loops of individual Rpts, with Rpt2 requiring a glutamate to asparagine mutation in this position. Base variants with a single pore-2-loop mutation hydrolyzed ATP up to three-fold faster than wild type, with the exception of Rpt1, which showed a 25% decrease (Table 1). The increased ATPase rates are reminiscent of previous observations for ClpX (Martin et al., 2007) and may result from induced structural changes in the pore-2 loop that affect the adjacent Walker-B portion of the ATPase active site. Proteasomes reconstituted with pore-2-loop mutants exhibited a wide range of degradation rates (Fig. 4b), with the most severe defect observed for Rpt3, located at the top of the substrate-free staircase. Changes in substrate degradation did not directly correlate with changes in ATP hydrolysis, indicating that pore-2-loop mutations may interfere with substrate interactions and thus result in futile ATP-hydrolysis events. The largest discrepancies between ATPase and degradation activities were observed for pore-2-loop mutations in Rpt3 and Rpt4, indicating that Rpt subunits at the top of the pre-engaged staircase are particularly important for substrate processing, presumably by enabling efficient engagement. The degradation defects do not appear to be a consequence of compromised gate opening, as all pore-2-loop mutants show robust peptidase-gate opening activity, ranging from 84% to 192% compared to wild-type base (Table 1). The substantial variations in gate opening indicate that pore-2 loops are involved in unfoldase-peptidase communication and the Rpt C-terminal tails are not the sole determinants of the interaction between these subcomplexes. Our results thus emphasize the functional differences between individual Rpts.

Discussion

Our mutational studies using heterologously expressed base subcomplex and *in vitro* reconstituted 26S holoenzymes revealed that substrate degradation by the proteasome relies on distinct functional asymmetries with strongly non-equivalent contributions of individual Rpts. While a static three-fold symmetry determines the interactions between

the base ATPase ring and the core particle, substrate degradation seems to depend on asymmetric spiral-staircase arrangements of the ATPases in which subunits close to the pore entrance play crucial roles in substrate engagement.

Spiral-staircase arrangements have been previously observed in several homohexameric motors of the AAA+ and RecA families, and it has been suggested that individual ATPase subunits successively transition through the different vertical registers of the spiral to drive substrate translocation (Costa et al., 2011; Enemark and Joshua-Tor, 2006; Glynn et al., 2009; Singleton et al., 2000; Thomsen and Berger, 2009). However, the apparently rigid spiral-staircase configurations observed for the proteasomal ATPase ring in the absence and presence of substrate, together with the biochemical data presented here, contradict stepwise successive progression of individual Rpts through the different registers. Instead, the proteasome base relies on differential subunit contributions prescribed by the regulatory particle's asymmetric architecture. Conformational changes in the top subunits of a pre-engaged staircase thread an incoming substrate, before the ATPase ring switches to an engaged staircase arrangement, in which subunits contribute more equally to protein translocation.

Although all mutant base complexes in this study were characterized by *in vitro* assays that may or may not wholly recapitulate the complexity of *in vivo* conditions, our results agree with and substantially extend earlier findings, suggesting non-equivalent contributions of individual Rpt subunits to substrate processing by the 26S proteasome (Erales et al., 2012; Kim et al., 2012; Rubin et al., 1998). Protein degradation by the proteasome deviates from models previously proposed for other AAA+ unfoldases, pointing toward major differences between the operating principles of homo- and heterohexameric AAA+ motors. Future experiments will have to address to what extent these motor designs rely on different strategies for substrate engagement, generation of mechanical force for unfolding, and translocation of the polypeptide chain into a peptidase for degradation.

Figures

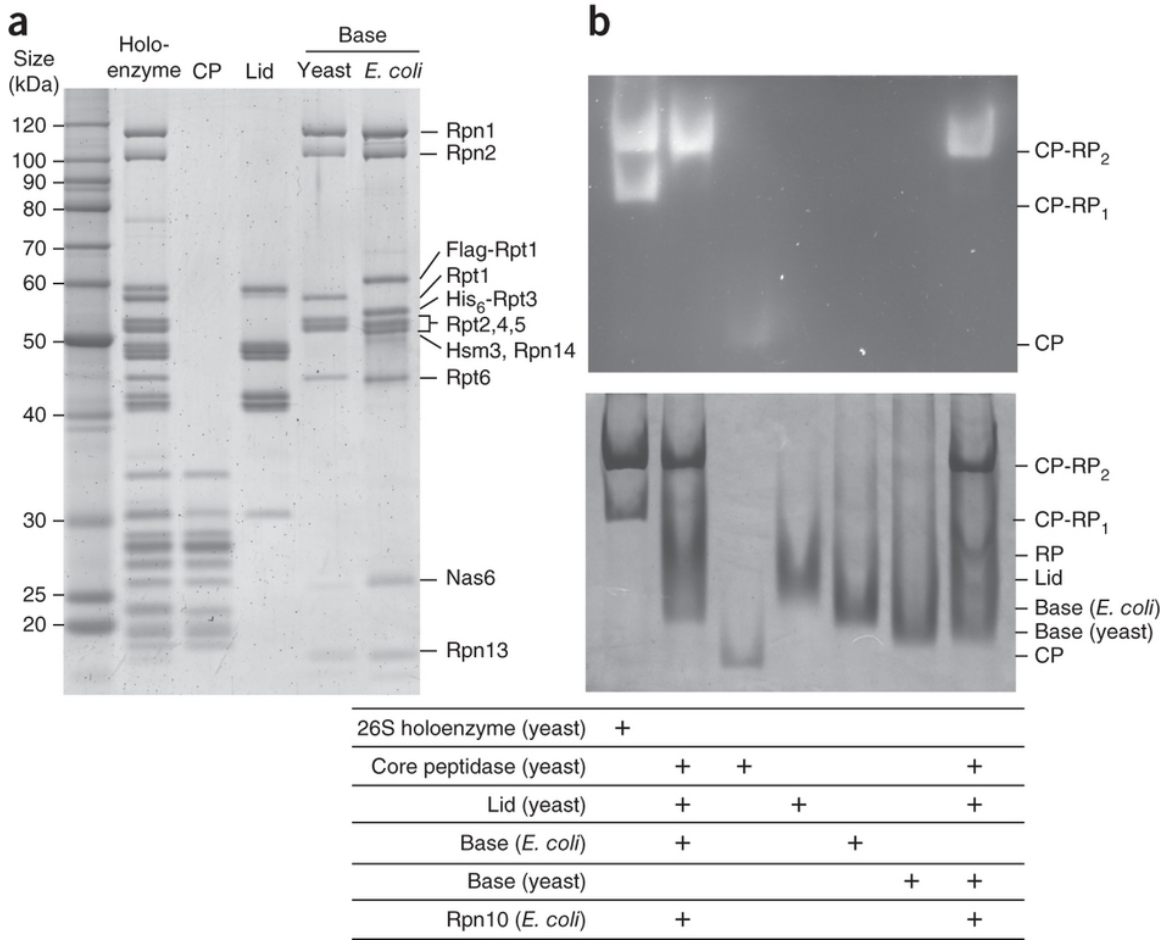


Fig. 4.1: Expression of yeast base subcomplex in *E. coli* and reconstitution of 26S proteasome. a) Sypro-Ruby stained SDS PAGE of the purified proteasomal subcomplexes used in this study. Endogenous complexes were isolated from yeast using FLAG tags on Rpn11 for holoenzyme and lid, on Pre1 for core particle (CP), and on Rpn2 for base preparations. Recombinant base expressed in *E. coli* was purified using a FLAG tag on Rpt1 and a His_6 tag on Rpt3. Proteasome chaperones Nas6, Hsm3 and Rpn14 only co-purify with the base subcomplex produced in *E. coli*. **b)** 26S holoenzyme reconstituted with CP, lid, Rpn10, and either endogenous or recombinant base was analyzed by native gel electrophoresis. Endogenous yeast 26S holoenzyme and individual CP, lid, and base subcomplexes were also analyzed for comparison. Yeast holoenzyme migrated as two bands corresponding to proteasomes singly ($CP\text{-}RP_1$) and doubly ($CP\text{-}RP_2$) capped with regulatory particles (RP). Excess lid and base was used for reconstituted proteasome samples, which therefore migrated only as doubly capped holoenzyme.

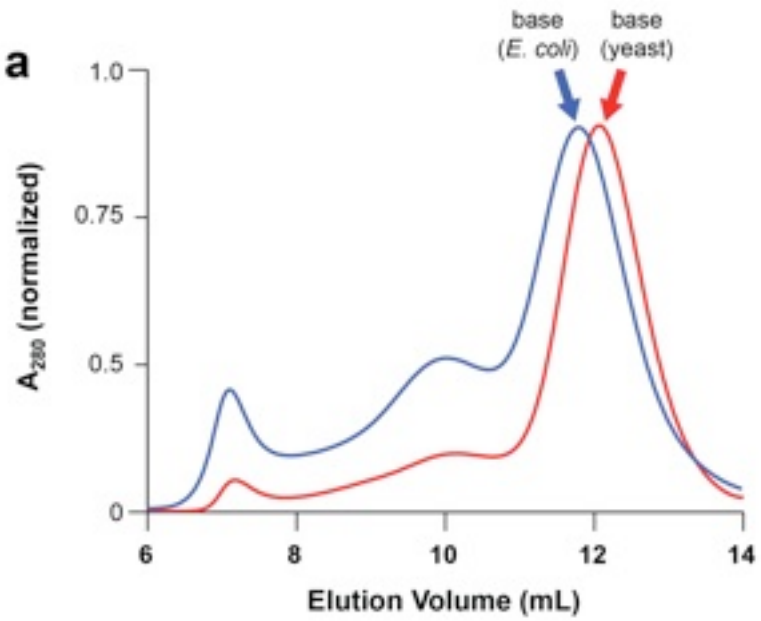


Fig. 4.2: Purified recombinant base subcomplex resembles endogenous base. Endogenous (red) and *E. coli*-expressed, recombinant (blue) base subcomplexes show similar elution profiles from a Superose6 size-exclusion column. The slightly smaller elution volume for recombinant base is attributed to the co-purification of proteasome-specific chaperones that stably associate with the complex when heterologously expressed in the absence of core particle in *E. coli*. The absorbance at 280 nm is normalized for comparison. For equal cell mass, recombinant base expression yields approximately 10-fold more protein than the purification of endogenous base from yeast.

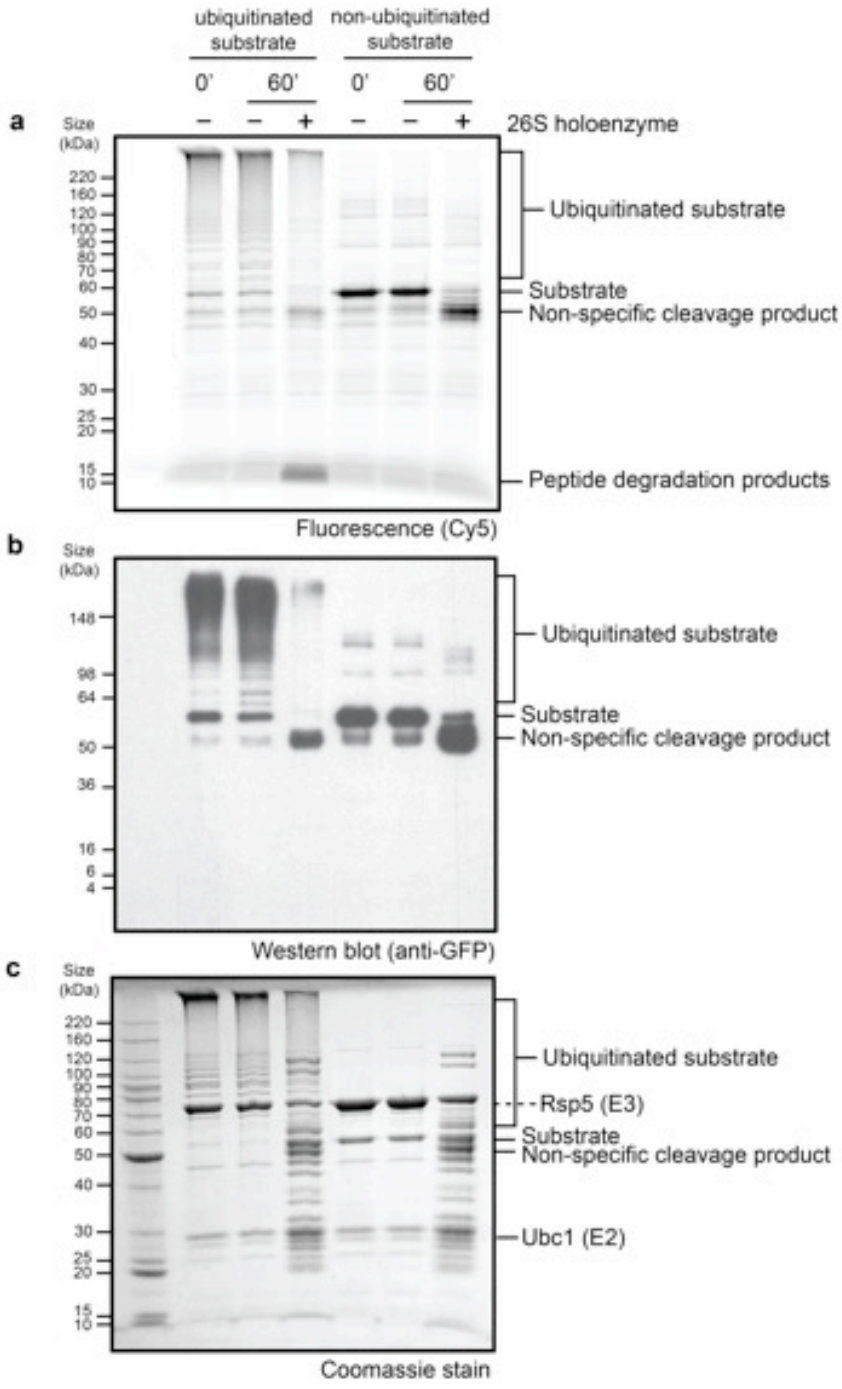

ubiquitinated substrate
non-ubiquitinated substrate
0'
60'
0'
60'
a
Size (kDa)
– – + – – +
26S holoenzyme
220
160
120
100
90
80
70
60
50
40
30
25
20
15
10
Ubiquitinated substrate
Substrate
Non-specific cleavage product
Peptide degradation products
Fluorescence (Cy5)
b
Size (kDa)
148
98
64
50
36
16
6
4
Ubiquitinated substrate
Substrate
Non-specific cleavage product
Western blot (anti-GFP)
c
Size (kDa)
220
160
120
100
90
80
70
60
50
40
30
25
20
15
10
Ubiquitinated substrate
Rsp5 (E3)
Substrate
Non-specific cleavage product
Ubc1 (E2)
Coomassie stain

Fig. 4.3: Ubiquitinated GFP-fusion substrate is degraded by the 26S proteasome. The model substrate, a green fluorescent protein (GFP)-titin^{V15P}-cyclin-PY fusion, was labeled at the N-terminus with Cy5 dye and subsequently modified with a polyubiquitin chain *in vitro* using Uba1, Ubc1, Rsp5 and wild-type ubiquitin. The non-ubiquitinated substrate was prepared similarly except wild-type ubiquitin was omitted from the reaction. Degradation was assessed by incubating substrates with 26S holoenzyme purified from yeast in the presence of an ATP regeneration system at 30°C for one hour. Substrate degradation was then assessed by running samples on a SDS-PAGE gel followed by **a**) fluorescence scanning to detect the Cy5-labeled substrate (670 nm band-pass 30 filter), **b**) western blotting using an anti-GFP antibody, or (**c**) Coomassie staining for total protein. The fluorescence scan clearly shows the accumulation of small peptide degradation products only for the ubiquitinated substrate in the presence of holoenzyme. Some level of ubiquitin-independent partial cleavage of an unstructured region of the GFP model substrate was detectable in all three assays. Additional ubiquitination of the substrate was visible in the absence of holoenzyme both by fluorescence scan and anti-GFP western blot, which was not unexpected as the enzymes used for *in vitro* ubiquitination of the substrate were still present.

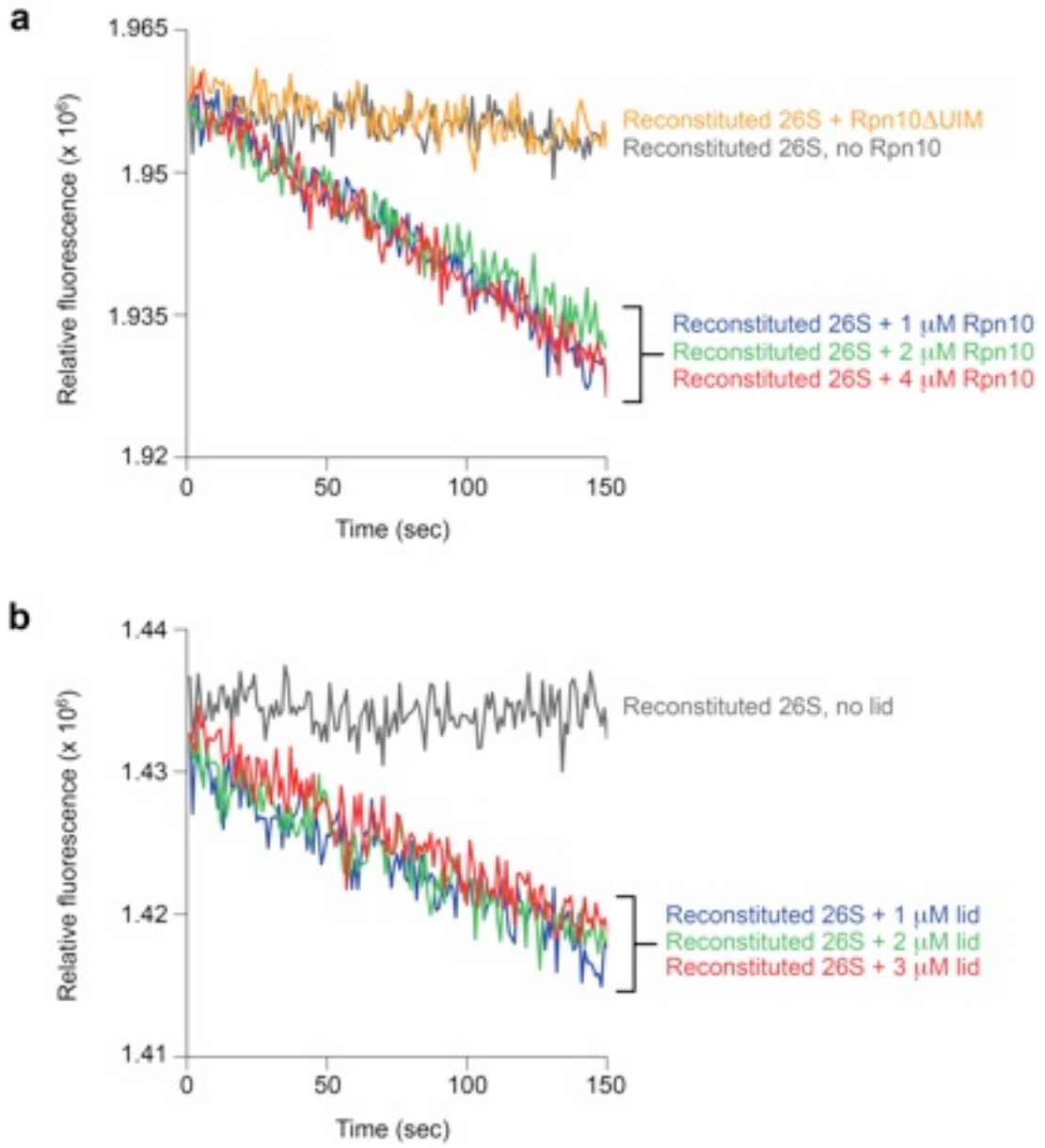


Fig. 4.4: Degradation rate for reconstituted 26S proteasome is not affected by excess amounts of lid or Rpn10. Proteasomal degradation was monitored by the decrease in fluorescence of a polyubiquitinated GFP-fusion substrate (excitation 467 nm, emission 511 nm) upon incubation with reconstituted 26S proteasome. Degradation reactions contained limiting amounts of core particle (yeast) and saturating concentrations of base (*E. coli*-expressed), lid (yeast), and 1 μM Rpn10 (*E. coli*-expressed). To establish that excess amounts of free lid and Rpn10 did not interact with our ubiquitinated substrate and adversely affect the measured degradation rates, we added increasing amounts of **a**) Rpn10 or **b**) lid and observed that the degradation rate remained constant.

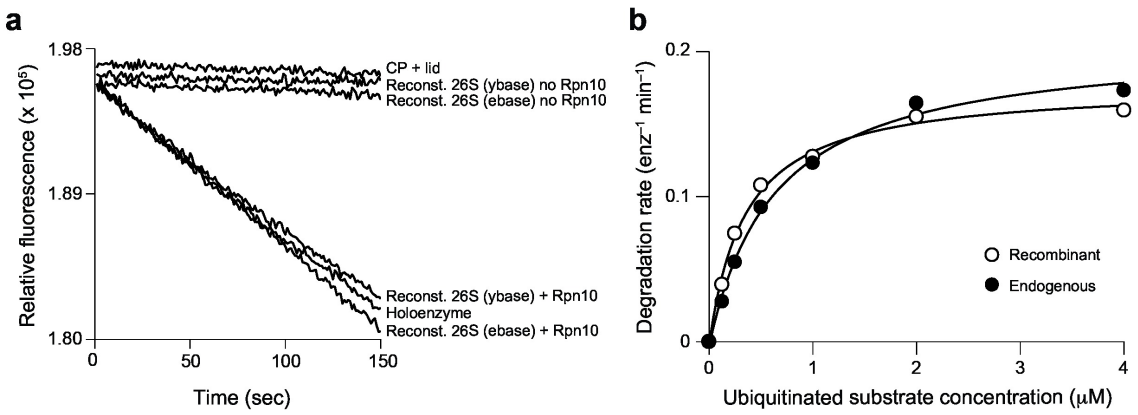
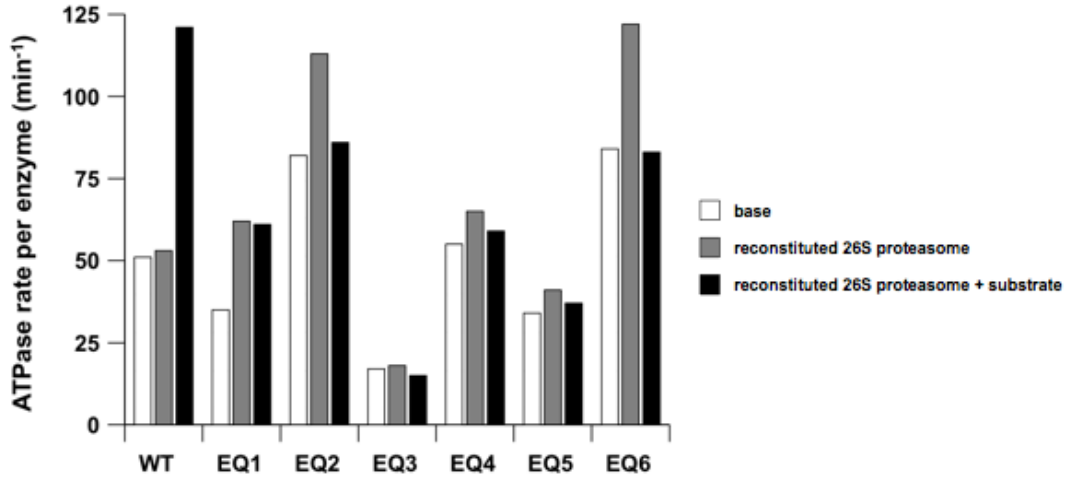


Fig. 4.5: Proteasomes reconstituted with endogenous or heterologously expressed base exhibit similar degradation activities for a polyubiquitinated substrate. a) Degradation of a polyubiquitinated GFP fusion substrate by endogenous yeast holoenzyme or 26S proteasomes reconstituted with saturating recombinant base (ebase) or endogenous yeast base (ybase). Substrate degradation was monitored by the loss of GFP fluorescence and strictly required the addition of Rpn10, despite the presence of Rpn13. **b)** Michaelis-Menten analyses of substrate degradation by proteasomes reconstituted with endogenous or recombinant base. Degradation reactions were performed using limiting base and excess core particle, lid and Rpn10 to ensure that reconstituted proteasome particles were singly capped. K_M and V_{max} values were 0.63 μM and 0.21 enz^{-1} min^{-1} for holoenzyme with endogenous base, and 0.35 μM and 0.18 enz^{-1} min^{-1} for holoenzyme with recombinant base.

a



b

	basal ATPase rate (base)		basal ATPase rate (reconstituted 26S)		working ATPase rate (reconstituted 26S proteasome + ubiquitinated substrate)	
	min^{-1}	% WT	min^{-1}	% WT	min^{-1}	% WT
WT	51	100	53	103	121	237
EQ1	35	68	62	122	61	120
EQ2	82	161	113	221	86	169
EQ3	17	33	18	36	15	28
EQ4	55	108	65	127	59	116
EQ5	34	67	41	81	37	72
EQ6	83	164	122	240	83	162

Fig. 4.6: ATP hydrolysis rates for Walker-B EQ mutant base subcomplexes are not stimulated by ubiquitinated substrate. Stimulation of ATPase activity by the base in the presence of ubiquitinated substrate was determined using a NADH-coupled ATPase assay. **a**) Basal rates of ATP hydrolysis per enzyme (base hexamer) were determined both for base subcomplexes alone (white) and reconstituted 26S proteasomes containing the base variants (gray). Working ATPase rates were measured by adding ubiquitinated substrate to reactions containing reconstituted 26S proteasomes (black). **b**) Table expressing the data from (a) in terms of ATP hydrolyzed per enzyme per minute and as a percentage of the rate observed for wild-type (WT) base alone. Errors for the ATPase assay were estimated to be ±10% of the WT mean value.

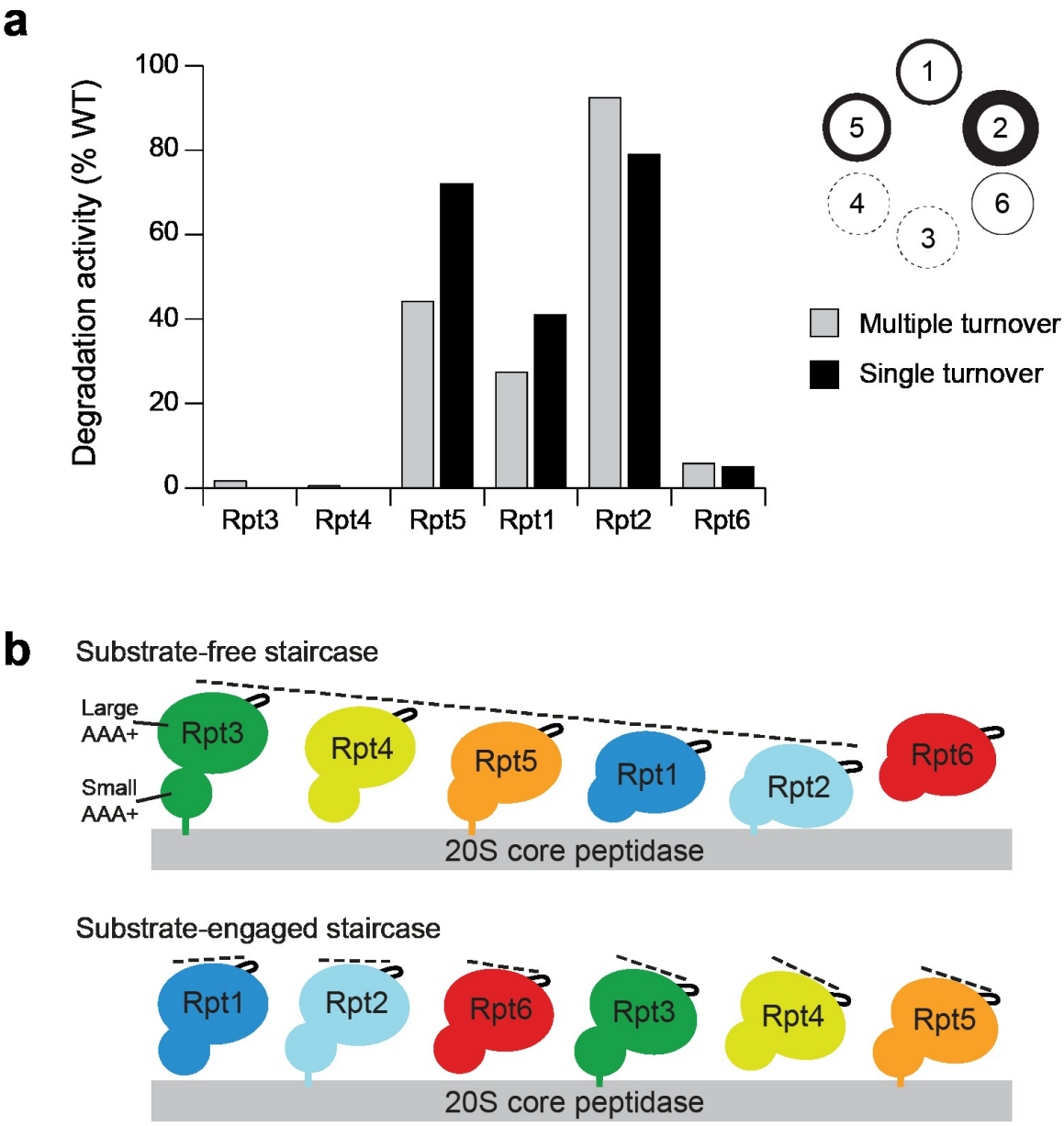

a
Degradation activity (% WT)
100
80
60
40
20
0
Rpt3
Rpt4
Rpt5
Rpt1
Rpt2
Rpt6
1
2
3
4
5
6
Multiple turnover
Single turnover
b
Substrate-free staircase
Large AAA+
Small AAA+
Rpt3
Rpt4
Rpt5
Rpt1
Rpt2
Rpt6
20S core peptidase
Substrate-engaged staircase
Rpt1
Rpt2
Rpt6
Rpt3
Rpt4
Rpt5
20S core peptidase

Fig. 4.7: Degradation activities for base variants with a Walker-B EQ mutation in individual Rpt subunits correlate with the subunit's position in the spiral staircase arrangement of the base. a) In vitro degradation rates for reconstituted proteasomes containing base variants with single Rpt subunits fixed in a permanent "ATP-bound" state by the EQ mutation. Degradation under multiple-turnover (gray) and single-turnover (black) conditions was monitored by the loss of fluorescence resulting from degradation of a polyubiquitinated GFP fusion substrate. Degradation activities were measured relative to reconstituted proteasome containing wild-type (WT) recombinant base. Errors for multiple-turnover degradation rates were estimated to be ±10% (s.d.) of the mean WT value based on repeat measurements (n = 3 technical replicates). The circular diagram is an alternative representation of the multiple-turnover data, with the line thickness corresponding to the observed degradation activities for a mutation in the respective subunit. **b)** The large AAA+ subdomains of Rpt1–6 adopt distinct spiral staircase arrangements in the absence and presence of substrate. Shown are cartoon representations of the pre-engaged (top) and substrate-engaged (bottom) staircases based on cryoEM reconstructions(Lander et al., 2012; Matyskiela et al., 2013), with the individual Rpt subunits splayed out and their pore-facing side pointing to the right. In the pre-engaged spiral, the small AAA+ subdomains of Rpt1–6 are arranged in a relatively planar fashion, while the large AAA+ subdomains are differentially lifted out of the ring plane, resulting in a pronounced spiral staircase with Rpt3 in the highest and Rpt2 in the lowest position. In the substrate-engaged spiral, the small and large AAA+ subdomains are mostly level, and the staircase orientation of pore loops primarily originates from differential rotations of subdomains in the plane of the ring.

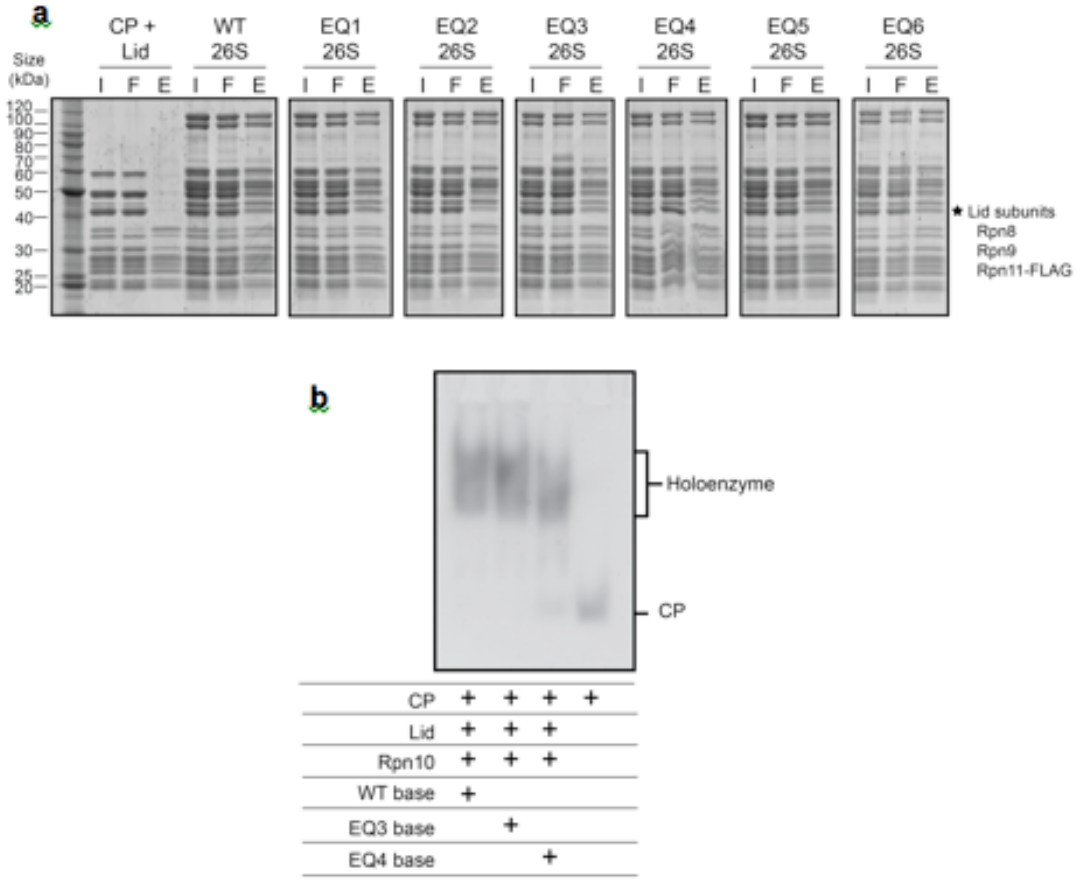

a
CP + Lid
WT 26S
EQ1 26S
EQ2 26S
EQ3 26S
EQ4 26S
EQ5 26S
EQ6 26S
Size (kDa)
I F E
120
100
90
80
70
60
50
40
30
25
20
★ Lid subunits
Rpn8
Rpn9
Rpn11-FLAG
b
Holoenzyme
CP
CP + + + +
Lid + + +
Rpn10 + + +
WT base +
EQ3 base +
EQ4 base +
c
Fraction native substrate (%)
100
80
60
40
20
0
EQ6
EQ1
EQ5
WT
EQ2
0
250
500
2000
4000
6000
Time (sec)

Fig. 4.8: Base variants with single-subunit Walker-B EQ mutations assemble into holoenzymes with varying degradation activities. a) Nickel affinity pulldown assay followed by SDS-PAGE analysis with Sypro Ruby staining to examine the assembly state of 26S proteasomes reconstituted with either wild-type (WT) or EQ base mutants. Untagged lid subcomplex was pulled down using His_6 tags on Rpt3 (base), Pre1 (core particle, CP) and Rpn10. Lanes for each sample are labeled I (input), F (flow through) and E (elution). Equivalent amounts of lid were observed for 26S proteasomes reconstituted with WT or EQ mutant base variants as indicated by the strong band containing lid subunits Rpn8, Rpn9 and Rpn11-FLAG (★). **b**) Native gel analysis demonstrating that WT, EQ3 and EQ4 base variants are competent for assembly into 26S holoenzyme. Assembly of proteasomes was performed by incubating constituent subcomplexes with ATP, followed by native polyacrylamide gel electrophoresis as described in the methods. (**c**) Single-turnover degradation traces and curve fits for EQ base mutants. Degradation under single-turnover conditions was monitored by the decrease in fluorescence of 100 nM polyubiquitinated GFP-fusion substrate (excitation 467 nm, emission 511 nm) upon incubation with 2 mM 26S proteasome reconstituted with either WT base or EQ base variants. Proteasomes reconstituted with EQ3 or EQ4 base variants did not exhibit any measurable degradation even under single turnover conditions. Curves were best fit with a double exponential decay, likely reflecting degradation of two subpopulations of the substrate. These classes of substrate probably differ in the number or location of conjugated polyubiquitin chains but their affinity for the proteasome is expected to be similar.

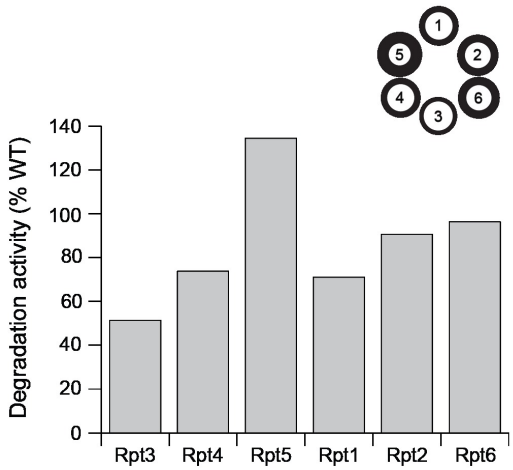


Fig. 4.9: Degradation activities for base variants containing single-subunit pore-2 loop mutations. Degradation of a polyubiquitinated GFP fusion substrate was monitored by the loss of fluorescence upon addition of proteasomes reconstituted with base variants containing DN mutations in the pore-2 loop of single Rpt subunits. Degradation activities under saturating conditions were measured relative to reconstituted proteasome containing wild-type (WT) recombinant base. Errors were estimated to be ±10% (s.d.) of the mean WT value based on repeat measurements (n = 3 technical replicates). The circular diagrams are alternative representations, with the line thickness corresponding to the observed degradation activities for a mutation in the respective subunit.

Tables

	Residue Mutated	basal ATPase rate		peptidase stimulation		degradation rate (k_{deg})	
		min^{-1}	% WT	fold increase	% WT	(enz^{-1} min^{-1})	% WT
Holoenzyme	-	107	-	-	-	0.32	-
WT (*E. coli*)	-	51	100	21	100	0.30	100
WT (yeast)	-	54	106	22	103	0.29	97
EQ hexamer	-	2	4	24	110	0.00	0.00
EQ1	310	35	68	15	68	0.08 / 0.12	27 / 40
EQ2	284	82	161	29	134	0.28 / 0.24	92 / 79
EQ3	273	17	33	19	89	0.005 / 0.00	2 / 0
EQ4	282	55	108	20	95	0.001 / 0.00	0.41 / 0
EQ5	282	34	67	25	118	0.13 / 0.21	44 / 72
EQ6	249	83	164	35	161	0.02 / 0.01	6 / 5
DN1	327	39	76	23	106	0.21	71
EN2	300	59	115	24	110	0.27	91
DN3	289	65	127	20	91	0.16	52
DN4	298	141	278	41	192	0.22	74
DN5	298	90	177	32	149	0.40	134
DN6	265	58	113	18	84	0.29	96

nd = not determined
** ATPase and degradation activities for the WA-KS mutants represent lower bounds due to varying degrees of misassembly observed for these base mutants.*

Table 4.1: Biochemical data for base variants with individual ATPase mutations. Summary of ATP hydrolysis rates (base), peptidase stimulation (CP and base), and degradation rates (reconstituted 26S proteasomes) for all base variants included in this study. Values for each assay are expressed both in absolute terms and as a normalized percentage of wild-type base activity. Base mutants are designated as EQ (Walker-B) and DN (pore-2 loop), followed by a number indicating which Rpt subunit contained the mutation (1–6). Degradation activities for WT base and EQ mutants are included for both multiple turnover (left) and single turnover (right) conditions. Errors were estimated to be ±10% (s.d.) based on repeat measurements ($n \geq 3$ technical replicates).

Chapter 5: Concluding remarks

Conclusions

The 26S proteasome is responsible for the essential duty of selective protein degradation in eukaryotic cells. This role has apparently induced a considerable amount of evolved complexity (for both selective protein degradation in general as well as for the ubiquitin-proteasome system in particular), with multiple different caps (e.g. 19S regulatory particle, 11S regulator, Blm10) able to bind to and alter the function of the modular 20S core particle. Furthermore, the canonical 26S proteasome, with the 19S RP as the cap, contains three large subcomplexes, each of which has its unique assembly strategies as well as mechanisms for regulating the enzymatic activities within. The thesis presented here has revealed further complexities of the proteasome, increasing our understanding of an elegantly regulated and coordinated protein destruction machine involved in practically every cellular process.

The subnanometer structure and complete subunit assignment of the proteasome was presented in chapter 2. Unexpectedly, the lid binds to the side of the base *and* the core particle, making it more of a side-car than a lid, and hinting at a potential functional network between the lid and core. Additionally, this thesis describes that the horseshoe shape organization of the six PCI subunits, including Rpn5, which apparently undergoes a large conformational change upon lid incorporation into the holoenzyme that may regulate the activity of the DUB Rpn11. The AAA+ hexameric unfoldase in the proteasome forms a spiral staircase arrangement, with specific subunits occupying the same position in each particle averaged from EM micrographs. This staircase arrangement of Rpt1-6 is a pre-engaged state that is competent to accept an incoming substrate. Upon substrate engagement, the ring changes conformations to an arrangement in which subunits contribute more equally to protein translocation.

In chapter 3, I describe the identification of a helical bundle (to our knowledge the most complex helical bundle known) composed of helices from each subunit in the lid. This helical bundle guides the ordered self-assembly process of the lid. The bundle also serves as a hub through which the C-terminal helix of Rpn12 can monitor complete lid assembly. Since efficient lid incorporation depends on the presence of Rpn12, the incorporation-coupled activation of Rpn11 is elegantly linked to complete assembly of the helical bundle. The helical bundle also connects the MPN dimer and the PCI horseshoe through flexible tethers, allowing a certain degree of independent movement of these two entities. This flexibility may be important during the conformational changes associated with cycles of ATP binding, hydrolysis, and nucleotide release in the AAA+ ATPases during substrate processing. Finally, I describe how the helical bundle is likely a common structure, found in the evolutionarily related eIF3 complex and COP9 Signalosome.

In chapter 4, I provide evidence for functional asymmetries among the Rpt ATPases, which we revealed by using a novel heterologous expression system for the base. The apparently rigid spiral-staircase configurations observed for the proteasomal ATPase ring

in the absence and presence of substrate, together with the biochemical data presented here, contradict stepwise successive progression of individual Rpts through the different registers. Instead, the proteasome base relies on differential subunit contributions prescribed by the regulatory particle's asymmetric architecture.

Future directions

The work presented here forms a framework for future structural, biochemical, and cellular studies on the 26S proteasome. Indeed, since this work has been published, our lab has solved the crystal structure of the Rpn8-11 MPN dimer (Worden et al., 2014), further sharpening our understanding of the events that directly precede degradation. Work from the Baumeister group has used a massive EM data set along with 'deep' classification to identify a third, 'intermediate' state of the proteasome that has been suggested to be important for substrate commitment to degradation before translocating substrate (Unverdorben et al., 2014). This study, though, surely is not the end for structural studies of the proteasome, and I look forward to news of even further deep-classification of proteasomes in the apo-state as well as during substrate processing, and I am especially excited for a clearer understanding of the mechanism of coordination of the heterohexameric ATPases resulting from this work. This will be aided by newly outfitted electron microscopes that make use of a direct-electron detector to yield substantially higher resolution models with fewer particles by increasing the efficiency of electron detection and by enabling the computational correction of electron beam-induced particle movement (Li et al., 2013).

Furthermore, I look forward to combining the deeper understanding presented here with the explosion of genome sequencing to determine the molecular basis of mysregulation of the 26S proteasome in humans and model organisms. Even higher resolution structures of the proteasome will aid in this goal, and hopefully a crystal structure of all or part of the regulatory particle is not too far away. A crystal structure would also provide crucial insight into how specificity is achieved between PCI-PCI contacts as well as confirming the topology of the helical bundle in the lid, which could aid in the use of the helical bundle as a synthetic biology tool for guiding protein complex assembly (mentioned in the discussion of chapter 3).

Already, Bortezomib, an FDA-approved drug that targets active sites in the core particle, is used to treat multiple myeloma. The findings in this thesis are an exciting starting point for the development of new drug targets on the proteasome that could have therapeutic benefits for diseases such as cancer or Alzheimer. For example, protein-protein interactions between PCI domains within the lid, PCI contacts to the base or core subunits, or interactions between helices within the helical bundle are all potential targets for disruption, which could lead to altered proteasome levels and slight changes in proteasome function. Finally, I look forward to the proteasome field continuing work on such long-held questions as how the ubiquitination state (e.g. different chain length or linkage type) of a substrate leads to differential proteasomal processing and whether there is any functional importance to a singly- vs. doubly-capped proteasome.

Final remarks

It has been over 35 years since the discovery of an ATP-dependent mechanism of protein degradation. From the seminal discovery of two discrete subcomplexes (the base and the lid) in the 19S regulatory particle, to the more recent work using a simple yet powerful method to determine the order of ATPases in the base unfoldase ring, it has been fascinating to learn the history of the development of our understanding of such a basic molecular process as selective protein degradation (Glickman et al., 1998a; Tomko et al., 2010). However small, I am happy that I was able work on and add my own fingerprint to the cannon of work on such a fascinating, complex proteinacious machine.

References

Ahn, K., Erlander, M., Leturcq, D., Peterson, P.A., Fruh, K., and Yang, Y. (1996). In vivo characterization of the proteasome regulator PA28. J Biol Chem *271*, 18237-18242.

Arendt, C.S., and Hochstrasser, M. (1999). Eukaryotic 20S proteasome catalytic subunit propeptides prevent active site inactivation by N-terminal acetylation and promote particle assembly. EMBO J *18*, 3575-3585.

Aubin-Tam, M.E., Olivares, A.O., Sauer, R.T., Baker, T.A., and Lang, M.J. (2011). Single-molecule protein unfolding and translocation by an ATP-fueled proteolytic machine. Cell *145*, 257-267.

Barthelme, D., and Sauer, R.T. (2013). Bipartite determinants mediate an evolutionarily conserved interaction between Cdc48 and the 20S peptidase. Proc Natl Acad Sci U S A *110*, 3327-3332.

Beck, F., Unverdorben, P., Bohn, S., Schweitzer, A., Pfeifer, G., Sakata, E., Nickell, S., Plitzko, J.M., Villa, E., Baumeister, W.*, et al.* (2012). Near-atomic resolution structural model of the yeast 26S proteasome. Proc Natl Acad Sci U S A *109*, 14870-14875.

Boehringer, J., Riedinger, C., Paraskevopoulos, K., Johnson, E.O., Lowe, E.D., Khoudian, C., Smith, D., Noble, M.E., Gordon, C., and Endicott, J.A. (2012). Structural and functional characterization of Rpn12 identifies residues required for Rpn10 proteasome incorporation. Biochem J *448*, 55-65.

Bohn, S., Beck, F., Sakata, E., Walzthoeni, T., Beck, M., Aebersold, R., Forster, F., Baumeister, W., and Nickell, S. (2010). Structure of the 26S proteasome from Schizosaccharomyces pombe at subnanometer resolution. Proc Natl Acad Sci U S A *107*, 20992-20997.

Bremm, A., Freund, S.M., and Komander, D. (2010). Lys11-linked ubiquitin chains adopt compact conformations and are preferentially hydrolyzed by the deubiquitinase Cezanne. Nat Struct Mol Biol *17*, 939-947.

Buchan, D.W., Ward, S.M., Lobley, A.E., Nugent, T.C., Bryson, K., and Jones, D.T. (2010). Protein annotation and modelling servers at University College London. Nucleic Acids Res *38*, W563-568.

Chen, P., and Hochstrasser, M. (1996). Autocatalytic subunit processing couples active site formation in the 20S proteasome to completion of assembly. Cell *86*, 961-972.

Chen, Z.A., Jawhari, A., Fischer, L., Buchen, C., Tahir, S., Kamenski, T., Rasmussen, M., Lariviere, L., Bukowski-Wills, J.C., Nilges, M.*, et al.* (2010). Architecture of the RNA polymerase II-TFIIF complex revealed by cross-linking and mass spectrometry. EMBO J *29*, 717-726.

Chys, P., and Chacon, P. (2012). Spinor product computations for protein conformations. J Comput Chem *33*, 1717-1729.

Chys, P., and Chacon, P. (2013). Random Coordinate Descent with Spinor-matices and Geometric Filters for Efficient Loop Closure. Journal of Chemical Theory and Computation *9*, 1821-1829.

Ciechanover, A. (1994). The ubiquitin-proteasome proteolytic pathway. Cell *79*, 13-21.

Ciechanover, A., Elias, S., Heller, H., and Hershko, A. (1982). "Covalent affinity" purification of ubiquitin-activating enzyme. J Biol Chem *257*, 2537-2542.

Cohn, Z.A., and Hirsch, J.G. (1960). The isolation and properties of the specific cytoplasmic granules of rabbit polymorphonuclear leucocytes. The Journal of experimental medicine *112*, 983-1004.
Cook, W.J., Jeffrey, L.C., Carson, M., Chen, Z., and Pickart, C.M. (1992). Structure of a diubiquitin conjugate and a model for interaction with ubiquitin conjugating enzyme (E2). J Biol Chem *267*, 16467-16471.
Cope, G.A., Suh, G.S., Aravind, L., Schwarz, S.E., Zipursky, S.L., Koonin, E.V., and Deshaies, R.J. (2002). Role of predicted metalloprotease motif of Jab1/Csn5 in cleavage of Nedd8 from Cul1. Science *298*, 608-611.
Costa, A., Ilves, I., Tamberg, N., Petojevic, T., Nogales, E., Botchan, M.R., and Berger, J.M. (2011). The structural basis for MCM2-7 helicase activation by GINS and Cdc45. Nat Struct Mol Biol *18*, 471-477.
da Fonseca, P.C., He, J., and Morris, E.P. (2012). Molecular model of the human 26S proteasome. Mol Cell *46*, 54-66.
da Fonseca, P.C., and Morris, E.P. (2008). Structure of the human 26S proteasome: subunit radial displacements open the gate into the proteolytic core. J Biol Chem *283*, 23305-23314.
De Duve, C., Pressman, B.C., Gianetto, R., Wattiaux, R., and Appelmans, F. (1955). Tissue fractionation studies. 6. Intracellular distribution patterns of enzymes in rat-liver tissue. Biochem J *60*, 604-617.
Dessau, M., Halimi, Y., Erez, T., Chomsky-Hecht, O., Chamovitz, D.A., and Hirsch, J.A. (2008). The Arabidopsis COP9 signalosome subunit 7 is a model PCI domain protein with subdomains involved in COP9 signalosome assembly. Plant Cell *20*, 2815-2834.
Driscoll, J., and Goldberg, A.L. (1990). The proteasome (multicatalytic protease) is a component of the 1500-kDa proteolytic complex which degrades ubiquitin-conjugated proteins. J Biol Chem *265*, 4789-4792.
Eddins, M.J., Varadan, R., Fushman, D., Pickart, C.M., and Wolberger, C. (2007). Crystal structure and solution NMR studies of Lys48-linked tetraubiquitin at neutral pH. J Mol Biol *367*, 204-211.
Effantin, G., Rosenzweig, R., Glickman, M.H., and Steven, A.C. (2009). Electron microscopic evidence in support of alpha-solenoid models of proteasomal subunits Rpn1 and Rpn2. J Mol Biol *386*, 1204-1211.
Enchev, R.I., Scott, D.C., da Fonseca, P.C., Schreiber, A., Monda, J.K., Schulman, B.A., Peter, M., and Morris, E.P. (2012). Structural basis for a reciprocal regulation between SCF and CSN. Cell Rep *2*, 616-627.
Enemark, E.J., and Joshua-Tor, L. (2006). Mechanism of DNA translocation in a replicative hexameric helicase. Nature *442*, 270-275.
Erales, J., Hoyt, M.A., Troll, F., and Coffino, P. (2012). Functional asymmetries of proteasome translocase pore. J Biol Chem *287*, 18535-18543.
Estrin, E., Lopez-Blanco, J.R., Chacon, P., and Martin, A. (2013). Formation of an intricate helical bundle dictates the assembly of the 26S proteasome lid. Structure *21*, 1624-1635.
Etlinger, J.D., and Goldberg, A.L. (1977). A soluble ATP-dependent proteolytic system responsible for the degradation of abnormal proteins in reticulocytes. Proc Natl Acad Sci U S A *74*, 54-58.

Finley, D. (2009). Recognition and processing of ubiquitin-protein conjugates by the proteasome. Annu Rev Biochem *78*, 477-513.
Forster, A., Masters, E.I., Whitby, F.G., Robinson, H., and Hill, C.P. (2005). The 1.9 A structure of a proteasome-11S activator complex and implications for proteasome-PAN/PA700 interactions. Mol Cell *18*, 589-599.
Forster, F., Lasker, K., Beck, F., Nickell, S., Sali, A., and Baumeister, W. (2009). An atomic model AAA-ATPase/20S core particle sub-complex of the 26S proteasome. Biochem Biophys Res Commun *388*, 228-233.
Frank, J., Radermacher, M., Penczek, P., Zhu, J., Li, Y., Ladjadj, M., and Leith, A. (1996). SPIDER and WEB: processing and visualization of images in 3D electron microscopy and related fields. J Struct Biol *116*, 190-199.
Fukunaga, K., Kudo, T., Toh-e, A., Tanaka, K., and Saeki, Y. (2010). Dissection of the assembly pathway of the proteasome lid in Saccharomyces cerevisiae. Biochem Biophys Res Commun *396*, 1048-1053.
Funakoshi, M., Tomko, R.J., Jr., Kobayashi, H., and Hochstrasser, M. (2009). Multiple assembly chaperones govern biogenesis of the proteasome regulatory particle base. Cell *137*, 887-899.
Gillette, T.G., Kumar, B., Thompson, D., Slaughter, C.A., and DeMartino, G.N. (2008). Differential roles of the COOH termini of AAA subunits of PA700 (19 S regulator) in asymmetric assembly and activation of the 26 S proteasome. J Biol Chem *283*, 31813-31822.
Glickman, M.H., Rubin, D.M., Coux, O., Wefes, I., Pfeifer, G., Cjeka, Z., Baumeister, W., Fried, V.A., and Finley, D. (1998a). A subcomplex of the proteasome regulatory particle required for ubiquitin-conjugate degradation and related to the COP9-signalosome and eIF3. Cell *94*, 615-623.
Glickman, M.H., Rubin, D.M., Fried, V.A., and Finley, D. (1998b). The regulatory particle of the Saccharomyces cerevisiae proteasome. Mol Cell Biol *18*, 3149-3162.
Glynn, S.E., Martin, A., Nager, A.R., Baker, T.A., and Sauer, R.T. (2009). Structures of asymmetric ClpX hexamers reveal nucleotide-dependent motions in a AAA+ protein-unfolding machine. Cell *139*, 744-756.
Goddard, T.D., Huang, C.C., and Ferrin, T.E. (2007). Visualizing density maps with UCSF Chimera. J Struct Biol *157*, 281-287.
Gomez, E.B., Catlett, M.G., and Forsburg, S.L. (2002). Different phenotypes in vivo are associated with ATPase motif mutations in Schizosaccharomyces pombe minichromosome maintenance proteins. Genetics *160*, 1305-1318.
Gomez, T.A., Kolawa, N., Gee, M., Sweredoski, M.J., and Deshaies, R.J. (2011). Identification of a functional docking site in the Rpn1 LRR domain for the UBA-UBL domain protein Ddi1. BMC Biol *9*, 33.
Gregori, L., Poosch, M.S., Cousins, G., and Chau, V. (1990). A uniform isopeptide-linked multiubiquitin chain is sufficient to target substrate for degradation in ubiquitin-mediated proteolysis. J Biol Chem *265*, 8354-8357.
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., Ditzel, L., Lowe, J., Stock, D., Bochtler, M., Bartunik, H.D., and Huber, R. (1997). Structure of 20S proteasome from yeast at 2.4 A resolution. Nature *386*, 463-471.

Hamazaki, J., Iemura, S., Natsume, T., Yashiroda, H., Tanaka, K., and Murata, S. (2006). A novel proteasome interacting protein recruits the deubiquitinating enzyme UCH37 to 26S proteasomes. Embo J *25*, 4524-4536.
Hanna, J., Hathaway, N.A., Tone, Y., Crosas, B., Elsasser, S., Kirkpatrick, D.S., Leggett, D.S., Gygi, S.P., King, R.W., and Finley, D. (2006). Deubiquitinating enzyme Ubp6 functions noncatalytically to delay proteasomal degradation. Cell *127*, 99-111.
Hanna, J., Meides, A., Zhang, D.P., and Finley, D. (2007). A ubiquitin stress response induces altered proteasome composition. Cell *129*, 747-759.
Hanson, P.I., Roth, R., Morisaki, H., Jahn, R., and Heuser, J.E. (1997). Structure and conformational changes in NSF and its membrane receptor complexes visualized by quick-freeze/deep-etch electron microscopy. Cell *90*, 523-535.
He, J., Kulkarni, K., da Fonseca, P.C., Krutauz, D., Glickman, M.H., Barford, D., and Morris, E.P. (2012). The structure of the 26S proteasome subunit Rpn2 reveals its PC repeat domain as a closed toroid of two concentric alpha-helical rings. Structure *20*, 513-521.
Hersch, G.L., Burton, R.E., Bolon, D.N., Baker, T.A., and Sauer, R.T. (2005). Asymmetric interactions of ATP with the AAA+ ClpX6 unfoldase: allosteric control of a protein machine. Cell *121*, 1017-1027.
Hershko, A., Ciechanover, A., and Rose, I.A. (1979). Resolution of the ATP-dependent proteolytic system from reticulocytes: a component that interacts with ATP. Proc Natl Acad Sci U S A *76*, 3107-3110.
Heymann, J.B., and Belnap, D.M. (2007). Bsoft: image processing and molecular modeling for electron microscopy. J Struct Biol *157*, 3-18.
Hinnerwisch, J., Fenton, W.A., Furtak, K.J., Farr, G.W., and Horwich, A.L. (2005). Loops in the central channel of ClpA chaperone mediate protein binding, unfolding, and translocation. Cell *121*, 1029-1041.
Hofmann, K., and Bucher, P. (1998). The PCI domain: a common theme in three multiprotein complexes. Trends in biochemical sciences *23*, 204-205.
Hohn, M., Tang, G., Goodyear, G., Baldwin, P.R., Huang, Z., Penczek, P.A., Yang, C., Glaeser, R.M., Adams, P.D., and Ludtke, S.J. (2007). SPARX, a new environment for Cryo-EM image processing. J Struct Biol *157*, 47-55.
Hughes, A.L. (1997). Evolution of the proteasome components. Immunogenetics *46*, 82-92.
Husnjak, K., Elsasser, S., Zhang, N., Chen, X., Randles, L., Shi, Y., Hofmann, K., Walters, K.J., Finley, D., and Dikic, I. (2008). Proteasome subunit Rpn13 is a novel ubiquitin receptor. Nature *453*, 481-488.
Inobe, T., Fishbain, S., Prakash, S., and Matouschek, A. (2011). Defining the geometry of the two-component proteasome degron. Nat Chem Biol *7*, 161-167.
Isono, E., Nishihara, K., Saeki, Y., Yashiroda, H., Kamata, N., Ge, L., Ueda, T., Kikuchi, Y., Tanaka, K., Nakano, A*., et al.* (2007). The assembly pathway of the 19S regulatory particle of the yeast 26S proteasome. Mol Biol Cell *18*, 569-580.
Joshi, S.A., Hersch, G.L., Baker, T.A., and Sauer, R.T. (2004). Communication between ClpX and ClpP during substrate processing and degradation. Nat Struct Mol Biol *11*, 404-411.

Kaneko, T., Hamazaki, J., Iemura, S., Sasaki, K., Furuyama, K., Natsume, T., Tanaka, K., and Murata, S. (2009). Assembly pathway of the Mammalian proteasome base subcomplex is mediated by multiple specific chaperones. Cell *137*, 914-925.
Kao, A., Randall, A., Yang, Y., Patel, V.R., Kandur, W., Guan, S., Rychnovsky, S.D., Baldi, P., and Huang, L. (2012). Mapping the structural topology of the yeast 19S proteasomal regulatory particle using chemical cross-linking and probabilistic modeling. Mol Cell Proteomics *11*, 1566-1577.
Kim, H.C., and Huibregtse, J.M. (2009). Polyubiquitination by HECT E3s and the determinants of chain type specificity. Mol Cell Biol *29*, 3307-3318.
Kim, Y.C., and DeMartino, G.N. (2011). C termini of proteasomal ATPases play nonequivalent roles in cellular assembly of mammalian 26 S proteasome. J Biol Chem *286*, 26652-26666.
Kim, Y.C., Li, X., Thompson, D., and Demartino, G.N. (2012). ATP-binding by proteasomal ATPases regulates cellular assembly and substrate-induced functions of the 26S proteasome. J Biol Chem.
Kohler, A., Cascio, P., Leggett, D.S., Woo, K.M., Goldberg, A.L., and Finley, D. (2001). The axial channel of the proteasome core particle is gated by the Rpt2 ATPase and controls both substrate entry and product release. Mol Cell *7*, 1143-1152.
Kotiguda, G.G., Weinberg, D., Dessau, M., Salvi, C., Serino, G., Chamovitz, D.A., and Hirsch, J.A. (2012). The organization of a CSN5-containing subcomplex of the COP9 signalosome. J Biol Chem *287*, 42031-42041.
Kumar, B., Kim, Y.C., and DeMartino, G.N. (2010). The C terminus of Rpt3, an ATPase subunit of PA700 (19 S) regulatory complex, is essential for 26 S proteasome assembly but not for activation. J Biol Chem *285*, 39523-39535.
Kusmierczyk, A.R., Kunjappu, M.J., Funakoshi, M., and Hochstrasser, M. (2008). A multimeric assembly factor controls the formation of alternative 20S proteasomes. Nat Struct Mol Biol *15*, 237-244.
Lander, G.C., Estrin, E., Matyskiela, M.E., Bashore, C., Nogales, E., and Martin, A. (2012). Complete subunit architecture of the proteasome regulatory particle. Nature *482*, 186-191.
Lander, G.C., Martin, A., and Nogales, E. (2013). The proteasome under the microscope: the regulatory particle in focus. Curr Opin Struct Biol.
Lander, G.C., Stagg, S.M., Voss, N.R., Cheng, A., Fellmann, D., Pulokas, J., Yoshioka, C., Irving, C., Mulder, A., Lau, P.W.*, et al.* (2009). Appion: an integrated, database-driven pipeline to facilitate EM image processing. J Struct Biol *166*, 95-102.
Lasker, K., Forster, F., Bohn, S., Walzthoeni, T., Villa, E., Unverdorben, P., Beck, F., Aebersold, R., Sali, A., and Baumeister, W. (2012). Molecular architecture of the 26S proteasome holocomplex determined by an integrative approach. Proc Natl Acad Sci U S A *109*, 1380-1387.
Lee, S.H., Moon, J.H., Yoon, S.K., and Yoon, J.B. (2012). Stable incorporation of ATPase subunits into 19 S regulatory particle of human proteasome requires nucleotide binding and C-terminal tails. J Biol Chem *287*, 9269-9279.
Leggett, D.S., Glickman, M.H., and Finley, D. (2005). Purification of proteasomes, proteasome subcomplexes, and proteasome-associated proteins from budding yeast. Methods Mol Biol *301*, 57-70.

Leggett, D.S., Hanna, J., Borodovsky, A., Crosas, B., Schmidt, M., Baker, R.T., Walz, T., Ploegh, H., and Finley, D. (2002). Multiple associated proteins regulate proteasome structure and function. Mol Cell *10*, 495-507.
Li, X., Mooney, P., Zheng, S., Booth, C.R., Braunfeld, M.B., Gubbens, S., Agard, D.A., and Cheng, Y. (2013). Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. Nat Methods *10*, 584-590.
Liu, C.W., Li, X., Thompson, D., Wooding, K., Chang, T.L., Tang, Z., Yu, H., Thomas, P.J., and DeMartino, G.N. (2006). ATP binding and ATP hydrolysis play distinct roles in the function of 26S proteasome. Mol Cell *24*, 39-50.
Maillard, R.A., Chistol, G., Sen, M., Righini, M., Tan, J., Kaiser, C.M., Hodges, C., Martin, A., and Bustamante, C. (2011). ClpX(P) generates mechanical force to unfold and translocate its protein substrates. Cell *145*, 459-469.
Mallick, S.P., Carragher, B., Potter, C.S., and Kriegman, D.J. (2005). ACE: automated CTF estimation. Ultramicroscopy *104*, 8-29.
Martin, A., Baker, T.A., and Sauer, R.T. (2005). Rebuilt AAA + motors reveal operating principles for ATP-fuelled machines. Nature *437*, 1115-1120.
Martin, A., Baker, T.A., and Sauer, R.T. (2007). Distinct static and dynamic interactions control ATPase-peptidase communication in a AAA+ protease. Mol Cell *27*, 41-52.
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.
Mattiroli, F., and Sixma, T.K. (2014). Lysine-targeting specificity in ubiquitin and ubiquitin-like modification pathways. Nat Struct Mol Biol *21*, 308-316.
Matyskiela, M.E., Lander, G.C., and Martin, A. (2013). Conformational switching of the 26S proteasome enables substrate degradation. Nat Struct Mol Biol *20*, 781-788.
Maytal-Kivity, V., Reis, N., Hofmann, K., and Glickman, M.H. (2002). MPN+, a putative catalytic motif found in a subset of MPN domain proteins from eukaryotes and prokaryotes, is critical for Rpn11 function. BMC biochemistry *3*, 28.
Mindell, J.A., and Grigorieff, N. (2003). Accurate determination of local defocus and specimen tilt in electron microscopy. J Struct Biol *142*, 334-347.
Murata, S., Yashiroda, H., and Tanaka, K. (2009). Molecular mechanisms of proteasome assembly. Nat Rev Mol Cell Biol *10*, 104-115.
Navon, A., and Goldberg, A.L. (2001). Proteins are unfolded on the surface of the ATPase ring before transport into the proteasome. Mol Cell *8*, 1339-1349.
Nickell, S., Beck, F., Scheres, S.H., Korinek, A., Forster, F., Lasker, K., Mihalache, O., Sun, N., Nagy, I., Sali, A.*, et al.* (2009). Insights into the molecular architecture of the 26S proteasome. Proc Natl Acad Sci U S A *106*, 11943-11947.
Palade, G.E. (1955). A small particulate component of the cytoplasm. The Journal of biophysical and biochemical cytology *1*, 59-68.
Park, E., Rho, Y.M., Koh, O.J., Ahn, S.W., Seong, I.S., Song, J.J., Bang, O., Seol, J.H., Wang, J., Eom, S.H.*, et al.* (2005). Role of the GYVG pore motif of HslU ATPase in protein unfolding and translocation for degradation by HslV peptidase. J Biol Chem *280*, 22892-22898.

Park, S., Li, X., Kim, H.M., Singh, C.R., Tian, G., Hoyt, M.A., Lovell, S., Battaile, K.P., Zolkiewski, M., Coffino, P.*, et al.* (2013). Reconfiguration of the proteasome during chaperone-mediated assembly. Nature *497*, 512-516.
Park, S., Roelofs, J., Kim, W., Robert, J., Schmidt, M., Gygi, S.P., and Finley, D. (2009). Hexameric assembly of the proteasomal ATPases is templated through their C termini. Nature *459*, 866-870.
Pathare, G.R., Nagy, I., Bohn, S., Unverdorben, P., Hubert, A., Korner, R., Nickell, S., Lasker, K., Sali, A., Tamura, T.*, et al.* (2011). The proteasomal subunit Rpn6 is a molecular clamp holding the core and regulatory subcomplexes together. Proc Natl Acad Sci U S A *109*, 149-154.
Pereira-Leal, J.B., Levy, E.D., Kamp, C., and Teichmann, S.A. (2007). Evolution of protein complexes by duplication of homomeric interactions. Genome Biol *8*, R51.
Pick, E., Golan, A., Zimbler, J.Z., Guo, L., Sharaby, Y., Tsuge, T., Hofmann, K., and Wei, N. (2012). The minimal deneddylase core of the COP9 signalosome excludes the Csn6 MPN- domain. PLoS One *7*, e43980.
Pick, E., Hofmann, K., and Glickman, M.H. (2009). PCI complexes: Beyond the proteasome, CSN, and eIF3 Troika. Mol Cell *35*, 260-264.
Piwko, W., and Jentsch, S. (2006). Proteasome-mediated protein processing by bidirectional degradation initiated from an internal site. Nat Struct Mol Biol *13*, 691-697.
Prakash, S., Tian, L., Ratliff, K.S., Lehotzky, R.E., and Matouschek, A. (2004). An unstructured initiation site is required for efficient proteasome-mediated degradation. Nat Struct Mol Biol *11*, 830-837.
Rabl, J., Smith, D.M., Yu, Y., Chang, S.C., Goldberg, A.L., and Cheng, Y. (2008). Mechanism of gate opening in the 20S proteasome by the proteasomal ATPases. Mol Cell *30*, 360-368.
Richardson, P.G., Hideshima, T., and Anderson, K.C. (2003). Bortezomib (PS-341): a novel, first-in-class proteasome inhibitor for the treatment of multiple myeloma and other cancers. Cancer control : journal of the Moffitt Cancer Center *10*, 361-369.
Riedinger, C., Boehringer, J., Trempe, J.F., Lowe, E.D., Brown, N.R., Gehring, K., Noble, M.E., Gordon, C., and Endicott, J.A. (2010). Structure of Rpn10 and its interactions with polyubiquitin chains and the proteasome subunit Rpn12. J Biol Chem *285*, 33992-34003.
Roelofs, J., Park, S., Haas, W., Tian, G., McAllister, F.E., Huo, Y., Lee, B.H., Zhang, F., Shi, Y., Gygi, S.P.*, et al.* (2009). Chaperone-mediated pathway of proteasome regulatory particle assembly. Nature *459*, 861-865.
Rubin, D.M., Glickman, M.H., Larsen, C.N., Dhruvakumar, S., and Finley, D. (1998). Active site mutants in the six regulatory particle ATPases reveal multiple roles for ATP in the proteasome. Embo J *17*, 4909-4919.
Rusu, M., Starosolski, Z., Wahle, M., Rigort, A., and Wriggers, W. (2010). Automated tracing of filaments in 3D electron tomography reconstructions using Sculptor and Situs. J Struct Biol *178*, 121-128.
Rusu, M., and Wriggers, W. (2012). Evolutionary bidirectional expansion for the tracing of alpha helices in cryo-electron microscopy reconstructions. J Struct Biol *177*, 410-419.
Saeki, Y., Toh, E.A., Kudo, T., Kawamura, H., and Tanaka, K. (2009). Multiple proteasome-interacting proteins assist the assembly of the yeast 19S regulatory particle. Cell *137*, 900-913.

Sanches, M., Alves, B.S., Zanchin, N.I., and Guimaraes, B.G. (2007). The crystal structure of the human Mov34 MPN domain reveals a metal-free dimer. J Mol Biol *370*, 846-855.
Sauer, R.T., and Baker, T.A. (2011). AAA+ proteases: ATP-fueled machines of protein destruction. Annu Rev Biochem *80*, 587-612.
Schreiner, P., Chen, X., Husnjak, K., Randles, L., Zhang, N., Elsasser, S., Finley, D., Dikic, I., Walters, K.J., and Groll, M. (2008). Ubiquitin docking at the proteasome through a novel pleckstrin-homology domain interaction. Nature *453*, 548-552.
Seebacher, J., Mallick, P., Zhang, N., Eddes, J.S., Aebersold, R., and Gelb, M.H. (2006). Protein cross-linking analysis using mass spectrometry, isotope-coded cross-linkers, and integrated computational data processing. J Proteome Res *5*, 2270-2282.
Serino, G., and Pick, E. (2013). Duplication and familial promiscuity within the proteasome lid and COP9 signalosome kin complexes. Plant Sci *203-204*, 89-97.
Sharon, M., Taverner, T., Ambroggio, X.I., Deshaies, R.J., and Robinson, C.V. (2006). Structural organization of the 19S proteasome lid: insights from MS of intact complexes. PLoS Biol *4*, e267.
Singleton, M.R., Sawaya, M.R., Ellenberger, T., and Wigley, D.B. (2000). Crystal structure of T7 gene 4 ring helicase indicates a mechanism for sequential hydrolysis of nucleotides. Cell *101*, 589-600.
Sledz, P., Unverdorben, P., Beck, F., Pfeifer, G., Schweitzer, A., Forster, F., and Baumeister, W. (2013). Structure of the 26S proteasome with ATP-gammaS bound provides insights into the mechanism of nucleotide-dependent substrate translocation. Proc Natl Acad Sci U S A *110*, 7264-7269.
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.
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 proteins. Mol Cell *20*, 687-698.
Sone, T., Saeki, Y., Toh-e, A., and Yokosawa, H. (2004). Sem1p is a novel subunit of the 26 S proteasome from Saccharomyces cerevisiae. J Biol Chem *279*, 28807-28816.
Sorzano, C.O., Marabini, R., Velazquez-Muriel, J., Bilbao-Castro, J.R., Scheres, S.H., Carazo, J.M., and Pascual-Montano, A. (2004). XMIPP: a new generation of an open-source image processing package for electron microscopy. J Struct Biol *148*, 194-204.
Staley, J.P., and Woolford, J.L., Jr. (2009). Assembly of ribosomes and spliceosomes: complex ribonucleoprotein machines. Curr Opin Cell Biol *21*, 109-118.
Sun, C., Todorovic, A., Querol-Audi, J., Bai, Y., Villa, N., Snyder, M., Ashchyan, J., Lewis, C.S., Hartland, A., Gradia, S.*, et al.* (2011). Functional reconstitution of human eukaryotic translation initiation factor 3 (eIF3). Proc Natl Acad Sci U S A *108*, 20473-20478.
Sutton, R.B., Fasshauer, D., Jahn, R., and Brunger, A.T. (1998). Crystal structure of a SNARE complex involved in synaptic exocytosis at 2.4 A resolution. Nature *395*, 347-353.
Tang, G., Peng, L., Baldwin, P.R., Mann, D.S., Jiang, W., Rees, I., and Ludtke, S.J. (2007). EMAN2: an extensible image processing suite for electron microscopy. J Struct Biol *157*, 38-46.

Thompson, D., Hakala, K., and DeMartino, G.N. (2009). Subcomplexes of PA700, the 19 S regulator of the 26 S proteasome, reveal relative roles of AAA subunits in 26 S proteasome assembly and activation and ATPase activity. J Biol Chem *284*, 24891-24903.
Thomsen, N.D., and Berger, J.M. (2009). Running in reverse: the structural basis for translocation polarity in hexameric helicases. Cell *139*, 523-534.
Thrower, J.S., Hoffman, L., Rechsteiner, M., and Pickart, C.M. (2000). Recognition of the polyubiquitin proteolytic signal. Embo J *19*, 94-102.
Tian, G., Park, S., Lee, M.J., Huck, B., McAllister, F., Hill, C.P., Gygi, S.P., and Finley, D. (2011). An asymmetric interface between the regulatory and core particles of the proteasome. Nat Struct Mol Biol *18*, 1259-1267.
Tomko Jr, R.J., and Hochstrasser, M. (2013). Molecular Architecture and Assembly of the Eukaryotic Proteasome. Annu Rev Biochem.
Tomko, R.J., Jr., Funakoshi, M., Schneider, K., Wang, J., and Hochstrasser, M. (2010). Heterohexameric ring arrangement of the eukaryotic proteasomal ATPases: implications for proteasome structure and assembly. Mol Cell *38*, 393-403.
Tomko, R.J., Jr., and Hochstrasser, M. (2011). Incorporation of the Rpn12 subunit couples completion of proteasome regulatory particle lid assembly to lid-base joining. Mol Cell *44*, 907-917.
Unverdorben, P., Beck, F., Sledz, P., Schweitzer, A., Pfeifer, G., Plitzko, J.M., Baumeister, W., and Forster, F. (2014). Deep classification of a large cryo-EM dataset defines the conformational landscape of the 26S proteasome. Proc Natl Acad Sci U S A *111*, 5544-5549.
van Heel, M., Harauz, G., Orlova, E.V., Schmidt, R., and Schatz, M. (1996). A new generation of the IMAGIC image processing system. J Struct Biol *116*, 17-24.
Verma, R., Aravind, L., Oania, R., McDonald, W.H., Yates, J.R., 3rd, Koonin, E.V., and Deshaies, R.J. (2002). Role of Rpn11 metalloprotease in deubiquitination and degradation by the 26S proteasome. Science *298*, 611-615.
Verma, R., Chen, S., Feldman, R., Schieltz, D., Yates, J., Dohmen, J., and Deshaies, R.J. (2000). Proteasomal proteomics: identification of nucleotide-sensitive proteasome-interacting proteins by mass spectrometric analysis of affinity-purified proteasomes. Mol Biol Cell *11*, 3425-3439.
Verma, R., Oania, R., Graumann, J., and Deshaies, R.J. (2004). Multiubiquitin chain receptors define a layer of substrate selectivity in the ubiquitin-proteasome system. Cell *118*, 99-110.
Voss, N.R., Yoshioka, C.K., Radermacher, M., Potter, C.S., and Carragher, B. (2009). DoG Picker and TiltPicker: software tools to facilitate particle selection in single particle electron microscopy. J Struct Biol *166*, 205-213.
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.
Wei, N., Tsuge, T., Serino, G., Dohmae, N., Takio, K., Matsui, M., and Deng, X.W. (1998). The COP9 complex is conserved between plants and mammals and is related to the 26S proteasome regulatory complex. Current biology : CB *8*, 919-922.
Weibezahn, J., Bukau, B., and Mogk, A. (2004). Unscrambling an egg: protein disaggregation by AAA+ proteins. Microb Cell Fact *3*, 1.

Worden, E.J., Padovani, C., and Martin, A. (2014). Structure of the Rpn11-Rpn8 dimer reveals mechanisms of substrate deubiquitination during proteasomal degradation. Nat Struct Mol Biol *21*, 220-227.
Xu, P., Duong, D.M., Seyfried, N.T., Cheng, D., Xie, Y., Robert, J., Rush, J., Hochstrasser, M., Finley, D., and Peng, J. (2009). Quantitative proteomics reveals the function of unconventional ubiquitin chains in proteasomal degradation. Cell *137*, 133-145.
Yamada-Inagawa, T., Okuno, T., Karata, K., Yamanaka, K., and Ogura, T. (2003). Conserved pore residues in the AAA protease FtsH are important for proteolysis and its coupling to ATP hydrolysis. J Biol Chem *278*, 50182-50187.
Yao, T., and Cohen, R.E. (2002). A cryptic protease couples deubiquitination and degradation by the proteasome. Nature *419*, 403-407.
Zhang, F., Hu, M., Tian, G., Zhang, P., Finley, D., Jeffrey, P.D., and Shi, Y. (2009a). Structural insights into the regulatory particle of the proteasome from Methanocaldococcus jannaschii. Mol Cell *34*, 473-484.
Zhang, F., Wu, Z., Zhang, P., Tian, G., Finley, D., and Shi, Y. (2009b). Mechanism of substrate unfolding and translocation by the regulatory particle of the proteasome from Methanocaldococcus jannaschii. Mol Cell *34*, 485-496.