Structural basis for dynamic regulation of the human 26S proteasome

Shuobing Chen[a,b,c,1], Jiayi Wu[b,1], Ying Lu[d,1], Yong-Bei Ma[c], Byung-Hoon Lee[e], Zhou Yu[f], Qi Ouyang[a,b,g], Daniel J. Finley[e], Marc W. Kirschner[d,2], and Youdong Mao[a,b,c,h,i,2]

[a]Center for Quantitative Biology, Peking University, Beijing 100871, China; [b]State Key Laboratory for Artificial Microstructures and Mesoscopic Physics, Institute of Condensed Matter Physics, School of Physics, Peking University, Beijing 100871, China; [c]Intel Parallel Computing Center for Structural Biology, Dana-Farber Cancer Institute, Boston, MA 02215; [d]Department of Systems Biology, Harvard Medical School, Boston, MA 02115; [e]Department of Cell Biology, Harvard Medical School, Boston, MA 02115; [f]Department of Molecular and Cellular Biology, Graduate School of Arts and Sciences, Harvard University, Cambridge, MA 02138; [g]Peking-Tsinghua Joint Center for Life Sciences, Peking University, Beijing 100871, China; [h]Department of Cancer Immunology and Virology, Dana-Farber Cancer Institute, Boston, MA 02115; and [i]Department of Microbiology and Immunobiology, Harvard Medical School, Boston, MA 02115

Contributed by Marc W. Kirschner, September 9, 2016 (sent for review July 1, 2016; reviewed by Wah Chiu, Aaron Ciechanover, and Mark Hochstrasser)

The proteasome is the major engine of protein degradation in all eukaryotic cells. At the heart of this machine is a heterohexameric ring of AAA (ATPases associated with diverse cellular activities) proteins that unfolds ubiquitylated target proteins that are concurrently translocated into a proteolytic chamber and degraded into peptides. Using cryoelectron microscopy, we determined a near–atomic-resolution structure of the 2.5-MDa human proteasome in its ground state, as well as subnanometer-resolution structures of the holoenzyme in three alternative conformational states. The substrate-unfolding AAA-ATPase channel is narrowed by 10 inward-facing pore loops arranged into two helices that run in parallel with each other, one hydrophobic in character and the other highly charged. The gate of the core particle was unexpectedly found closed in the ground state and open in only one of the alternative states. Coordinated, stepwise conformational changes of the regulatory particle couple ATP hydrolysis to substrate translocation and regulate gating of the core particle, leading to processive degradation.

ubiquitin-proteasome system | AAA-ATPase | cyroelectron microscopy

The amount of each protein in the cell depends on its rates of synthesis and degradation. In eukaryotic cells, selective protein degradation is mostly carried out by a set of pathways of ubiquitylation that terminate in a 2.5-MDa protein proteolytic complex, called the 26S proteasome. The ubiquitin-proteasome pathways are essential parts of important biological processes, such as cell division, differentiation, innate immunity, adaptive immunity, regulation of gene expression, and the response to proteotoxic stress (1–4). The proteasome is also an important therapeutic target in multiple myeloma (5, 6).

The proteasome is composed of a 28-subunit barrel-shaped core particle (CP) in the center capped at the top and bottom by 19-subunit regulatory particles (RPs) (*SI Appendix*, Fig. S1) (7–10). The CP forming the catalytic chamber contains three proteolytically active threonine residues. Heptameric α-rings, positioned on each side of the catalytic chamber, control substrate entry into this space. Opening of a channel within the α-ring is thought to result from association of the RP and the CP (5, 11–13). However, the mechanism of the core regulatory step of the proteasome channel opening remains mysterious.

The RP is formed from two subcomplexes known as the lid and the base. Recognition of a substrate with the requisite number and configuration of ubiquitin is mediated principally by the base subunit Rpn13 and by another receptor, Rpn10 (4). To allow substrate degradation, ubiquitin is first removed by Rpn11, a metalloprotease subunit in the lid. The globular domains of a substrate are then unfolded mechanically by a ring-like heterohexameric complex consisting of six distinct subunits, Rpt1 to Rpt6, which belong to the ATPases-associated-with-diverse-cellular-activities (AAA) family. Both Rpn11-dependent deubiquitylation of the substrate and unfolding of substrate globular domains require prior engagement with the translocation machinery.

High-resolution structures of the isolated CP have been available for several species (5, 12, 13). The intact proteasome has not been resolved to a level at which a reliable Cα-backbone can be traced with spatial assignment of amino acids, although major advances have been made in recent years (2, 3, 7–11, 14–25). Several RP subunits have been resolved at high resolution by X-ray crystallography (14–16, 18, 21–24). Recent cryoelectron microscopy (cryo-EM) analysis of the complete proteasome at moderate resolution (6 to 10 Å) revealed an overall subunit organization of the RP (7–10, 25–28). The lid subcomplex, which consists of nine Rpn proteins (Rpn3, Rpn5 to Rpn9, Rpn11, Rpn12, and Dss1/Sem1), exhibits a horseshoe-like architecture and is organized around an elaborate bundle assembled from the C-terminal helices of each subunit but Dss1/Sem1 (20, 29). The six Rpt subunits of the base share a general domain organization,

Significance

The proteasome holoenzyme is an ATP-dependent protease in eukaryotes that degrades ubiquitylated substrates. It is involved in numerous important biological processes, such as cell division, differentiation, innate immunity, adaptive immunity, regulation of gene expression, and response to proteotoxic stress. Using cryoelectron microscopy, we have examined multiple conformational states of the human proteasome at medium to high resolution. Our results reveal that the substrate-conducting channel in the core particle is transiently opened and accompanied by dynamic changes in structure of the particle. These observations provide new insights into how the proteasome recognizes ubiquitylated substrates and translocates them through a channel and gate to degradation sites in the core particle.

Author contributions: Y.L. and Y.M. designed research; S.C., J.W., Y.L., and Y.M. performed research; Y.L., B.-H.L., and Y.M. contributed new reagents/analytic tools; S.C., J.W., Y.L., Y.-B.M., Z.Y., D.J.F., and Y.M. analyzed data; and S.C., Y.L., Q.O., D.J.F., M.W.K., and Y.M. wrote the paper.

Reviewers: W.C., Baylor College of Medicine; A.C., Technion–Israel Institute of Technology; and M.H., Yale University.

The authors declare no conflict of interest.

Data deposition: The single-particle reconstructions and atomic coordinates reported in this paper have been deposited in the Electron Microscopy Data Bank, www.emdatabank.org (accession nos. EMD-8332 to EMD-8337) and Protein Data Bank, www.wwpdb.org [PDB ID codes 5T0C (the doubly capped proteasome in the S_A state), 5T0G, 5T0H, 5T0I, and 5T0J (the half proteasome in the S_A, S_B, S_C, S_D states, respectively)]. The raw micrographs and particle data have been deposited in the Electron Microscopy Pilot Image Archive, www.ebi.ac.uk/pdbe/emdb/empiar (accession no. EMPIAR-10072).

See Commentary on page 12896.

[1]S.C., J.W., and Y.L. contributed equally to this work.

[2]To whom correspondence may be addressed. Email: marc@hms.harvard.edu or youdong_mao@dfci.harvard.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1614614113/-/DCSupplemental.

starting with a coiled-coil (CC) domain at the N terminus, followed by an oligonucleotide- and oligosaccharide-binding (OB) domain and an AAA domain toward the C terminus (7, 23). The yeast proteasome holoenzyme was previously observed to assume three distinct conformations, hypothetically considered the substrate-accepting (s1), commitment (s2), and translocating (s3) states (25–28). The limited resolution, however, precludes understanding of the critical molecular mechanisms underlying ATP-dependent degradation by the proteasome holoenzyme. The nucleotide-binding states of AAA-ATPases, architecture of the peptide-unfolding channel, and allosteric regulation of the substrate-translocation pathway remain particularly elusive.

We report here structures of the human proteasome in four conformational states by single-particle cryo-EM. These structures offer atomic-level details of key intersubcomplex interfaces in the human proteasome assembly. A detailed dissection of the AAA-ATPase channel provides key insights into the molecular mechanism of ATP-dependent, processive substrate translocation. Importantly, the four conformational states reveal dynamic gating of the CP channel, which is elegantly regulated through highly coordinated conformational changes among distinct subcomplexes.

Results

Cryoelectron Microscopy of the Proteasome. Cryo-EM data were collected on the proteasome holoenzyme purified from human embryonic kidney (HEK) 293 cells (30) using a 200-kV cryogenic electron microscope equipped with a direct electron detector (Fig. 1 *A* and *B* and *SI Appendix*, Fig. S2 *A* and *B*). We reconstructed the complete, doubly capped proteasome complex after imposing C2 symmetry using 237,083 single-particle images. The CP component was refined to a nominal resolution of 3.6 Å (Fig. 1*B* and *SI Appendix*, Fig. S2*C*), whereas the resolutions of its two RPs were significantly lower than that of the CP in this map, suggesting that the RPs fluctuate conformationally (*SI Appendix*, Fig. S2*F*). Maximum likelihood-based classification identified a dataset of 85,420 particles of improved structural homogeneity, which improved the reconstruction of AAA-ATPase to around 4-Å resolution (*SI Appendix*, Fig. S2 *C* and *D*). However, the lid subcomplex still demonstrates significantly lower resolution, suggesting it is the most flexible component in the proteasome holoenzyme.

To investigate the conformational dynamics, we boxed half of the holoenzyme, including half of the CP, in complex with a complete RP (hereafter referred to as the RP–CP subcomplex) using additional cryo-EM data (Fig. 1*C* and *SI Appendix*, Figs. S2 and S3). We conducted iterative maximum likelihood-based classification focusing on the RP structure by using single-particle images of the RP–CP subcomplex with the CP density subtracted (31, 32) (*SI Appendix*, Fig. S3 and *Materials and Methods*). After exhaustive computational purification, we obtained structurally homogeneous datasets of the RP–CP subcomplex corresponding to four distinct conformational states, namely a major state (S_A) and three alternative states (S_B, S_C, and S_D) (Fig. 1 *D*–*G* and *SI Appendix*, Fig. S3). We refined the overall RP–CP reconstructions of the S_A, S_B, S_C, and S_D states to nominal resolutions of 4.4, 6.8, 8.0, and 8.0 Å, respectively (*SI Appendix*, Fig. S2*E*).

The high-resolution features of the doubly capped proteasome map are consistent with those of the RP–CP map in the S_A state. In both maps, the ATPase subunits are better-resolved than the lid subcomplex (*SI Appendix*, Fig. S2 *F* and *G*). The ATPase density is best-resolved in the doubly capped map in the S_A state, whereas the lid subcomplex is best-resolved in the RP–CP map in S_A at a nominal resolution of 4.9 Å (*SI Appendix*, Fig. S2*C*). The cryo-EM maps of the S_A state enabled atomic modeling and refinement (Fig. 1*H* and *SI Appendix*, Figs. S4–S7 and Table S1). The final atomic model of the S_A state contains all lid and CP subunits, Rpn2, Rpn10, and Rpt1 to Rpt6 (Fig. 1*D*). The local resolution of Rpn1 in S_A is about 8 Å, which nonetheless suffices to build its pseudoatomic model. Based on the atomic model of the S_A state, we built pseudoatomic models for S_B, S_C, and S_D (Fig. 1 *E*–*G*). As expected, the ubiquitin-interacting motif of the ubiquitin receptor Rpn10 is disordered and missing in the cryo-EM densities of all conformational states (17, 19). The other ubiquitin receptor, Rpn13, was not observed (*SI Appendix*, Fig. S1).

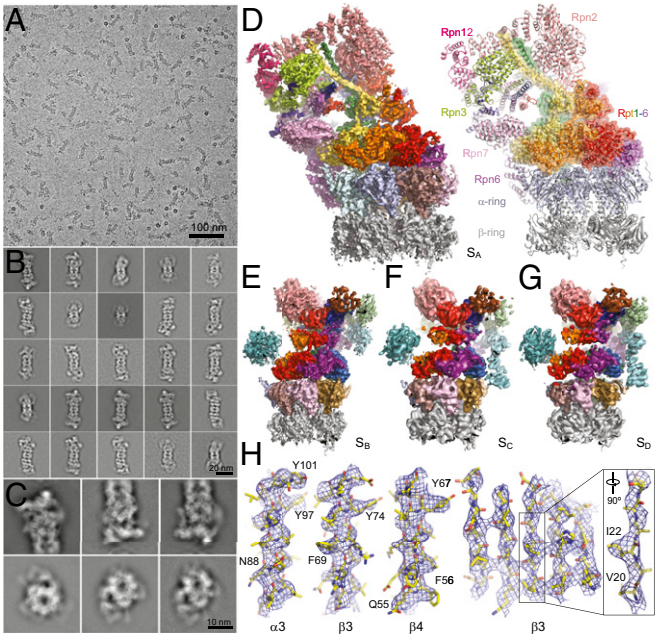


Fig. 1. Cryo-EM structure determination of the human proteasome in four conformational states. (*A*) A typical cryo-EM micrograph of the human proteasome imaged with a Tecnai Arctica and a Gatan K2 Summit direct detector camera. (*B*) Typical reference-free 2D class averages of the doubly capped proteasome computed by the ROME software (49). (*C*) Typical reference-free 2D class averages of the RP–CP subcomplex, showing great detail corresponding to secondary structures of the complex. (*D*) The cryo-EM density map of the RP–CP subcomplex in a surface representation (*Left*) and the atomic model built from the density map (*Right*). (*E*–*G*) The cryo-EM densities are shown as solid surfaces for the S_B (*E*), S_C (*F*), and S_D (*G*) states. (*H*) Representative cryo-EM densities of secondary structures of α-helices in the α3, β3, and β4 subunits (*Left* three panels) and β-strands in the β3 subunit (*Right* two panels) in the S_A state are superimposed with the fitted atomic model shown as a stick representation, showing that density quality is sufficient to allow side-chain fitting. The residue numbers of selected bulky side chains are labeled.

Overview of Conformational States of the Human Proteasome. The individual S_A, S_B, S_C, and S_D conformational states are represented in the particle populations at 76.2%, 10.1%, 5.8%, and 7.9%, respectively. The lid is rotated ∼40° clockwise in the S_B state relative to the S_A state but rotated ∼5° counterclockwise in S_C relative to S_B and translated ∼10 Å in S_D relative to S_C (Fig. 2 *A*–*D*). The AAA-ATPase heterohexamer movements on the α-ring in the S_C and S_D states exhibit stepwise axial rotation, lateral translation, and vertical rocking, which reduces the tilt of the ATPase ring relative to the α-ring (Fig. 2 *E*–*L*).

The structures of the CP in the S_A, S_B, and S_C states are identical up to their measured resolution. The crystal structure of free human CP with a closed CP channel can be fitted into the cryo-EM maps of these states as a rigid body, including the channel-blocking amino-terminal tails of the α-subunits (5) (Fig. 2 *M* and *N* and *SI Appendix*, Fig. S4 *J*–*M*). However, there is considerable structural change in the α-subunits of the S_D state, where the CP channel is open (Fig. 2 *O* and *P* and *SI Appendix*, Fig. S4*M*). A closed channel in the intact holoenzyme structure is unexpected (33), and suggests that association of the RP and CP subcomplexes does not open the CP channel by default in the S_A state of the human proteasome. Channel opening in the S_D state

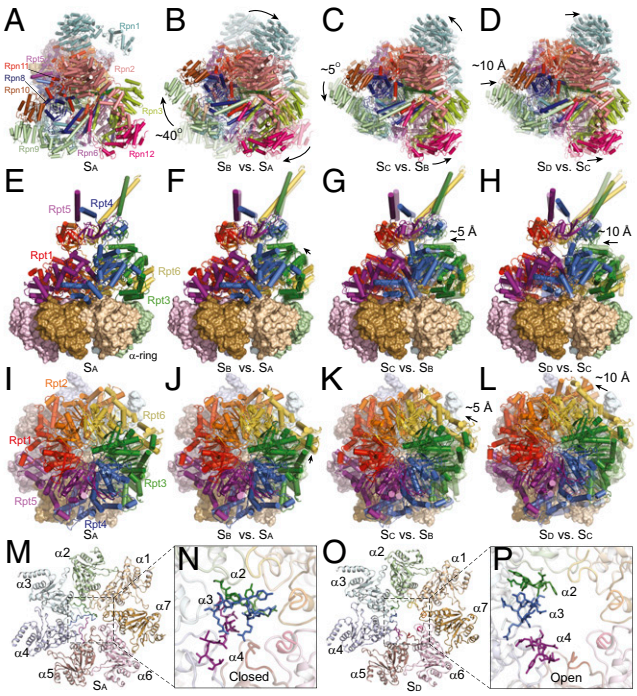


Fig. 2. Conformational changes of the human proteasome in four distinct states. (*A*) Top view of the lid of S_A. (*B–D*) Top views of the lid of S_B (*B*), S_C (*C*), and S_D (*D*) superimposed with transparent cartoons of S_A, S_B, and S_C, respectively. (*E*) Side view of the ATPase ring above the α-ring in S_A. (*F–H*) Side views of the ATPase ring above the α-ring in S_B (*F*), S_C (*G*), and S_D (*H*) superimposed with transparent cartoons of S_A, S_B, and S_C, respectively. (*I*) Top view of the ATPase ring in S_A. (*J–L*) Top views of the ATPase ring in S_B (*J*), S_C (*K*), and S_D (*L*) superimposed with transparent cartoons of S_A, S_B, and S_C, respectively. (*M*) The α-ring in a cartoon representation from the perspective of the AAA-ATPase or the RP–CP interface in S_A. (*N*) Close-up view of the central portion of the α-ring, showing that the CP channel is closed in this conformation. The amino-terminal tails of the α2, α3, and α4 subunits blocking the CP channel are shown in stick representation, whereas the rest of the structure is in cartoon representation. (*O*) The α-ring in a cartoon representation from the perspective of the AAA-ATPase or the RP–CP interface in S_D. (*P*) Close-up view of the central part of the α-ring, showing that the CP channel is open in this conformation. The amino-terminal tails of the α2, α3, and α4 subunits blocking the CP channel are shown in stick representation, whereas the rest of the structure is in cartoon representation.

is accompanied by coordinated conformational changes in the RP (Fig. 2 *A–L*).

Although multiple conformational states of the yeast proteasome were reported in previous studies (25–28), no human counterparts to these states have been observed before. Comparison between our human holoenzyme in four states and the yeast ones in three states (s1, s2, and s3) suggests considerable structural difference in both the lid and the base (*SI Appendix*, Fig. S8). However, the S_A and S_B states of the human holoenzyme generally correspond to the s1 and s2 states of the yeast holoenzyme (28) based on the general features of the Rpn11–OB and RC–CP interfaces (*SI Appendix*, Fig. S9*D*). In the human S_A and yeast s1 states, Rpn11 is positioned above the Rpt4–Rpt5 interface on the OB ring without blocking the substrate entry port (7, 9, 25, 28). By contrast, Rpn11 is moved over the center of the substrate entry port in both human S_B and yeast s2 states (28). In the human S_C/S_D states, there is prominent translation of the ATPase ring with respect to the CP, a feature shared with the yeast substrate-engaged and s3 states (26, 27).

The Lid–Base Interfaces. The lid–base interface plays a critical role in coordinating substrate translocation with deubiquitylation (7–10, 26, 28). In the S_A state, the interface between the ATPase ring and lid buries ~3,900 Å^2 of interfacial area. Most of these interfaces are contributed by Rpt3 and Rpt6, whose CC–OB domains bury ~3,100 Å^2 of interface area with the lid (*SI Appendix*, Fig. S9*A* and Table S2). The CC domain of the Rpt3–Rpt6 heterodimer is encircled by the helical elements of Rpn2, Rpn3, Rpn8, and Rpn11 (*SI Appendix*, Fig. S9*A*) (7, 25, 27). This interfacial architecture constitutes the stabilizing core of the lid–base association and is largely invariant in all four conformational states. Using this interface as a pivot, Rpn11 rotates 30° to a position closer to the OB ring of AAA-ATPase in the S_B, S_C, and S_D states (*SI Appendix*, Fig. S9 *D–G*).

The lateral lid–base interface exhibits prominent conformational transitions. The AAA domains of Rpt3 and Rpt6 bury ~1,800 Å^2 of interface area with Rpn5 to Rpn7 in the S_A state (*SI Appendix*, Fig. S9*B* and Table S2). One side of the amino-terminal PCI (proteasome-cyclosome-initiation factor) domain of Rpn7, involving four helix-connecting loops, makes extensive interaction (~940 Å^2) with the Rpt6 AAA domain. This interface is translated clockwise for ~10 Å in S_B relative to S_A, and remains nearly invariant in S_C and S_D. The AAA domains of Rpt3 and Rpt6 make limited contact with Rpn5 and Rpn6, with only ~340 and ~480 Å^2 buried in the S_A state, respectively. Consistent with the small interface area, the cryo-EM densities of the Rpn5/Rpn6 amino-terminal PCI domain are substantially weakened in the S_B state, implying local dissociation between Rpn and Rpt subunits at these interfaces. In the S_D state, the Rpn5–Rpt3 interface is reestablished. This analysis suggests that the Rpn7–Rpt6 interface may be retained during conformational changes of the lid, whereas the smaller interfaces of Rpn5–Rpt3 and Rpn6–Rpt6 are labile (25–27).

The AAA-ATPase Heterohexamer. ATP binds the Walker A motif located next to a short linker between the small and large AAA subdomains of the ATPases. The nucleotide-binding state regulates the conformations of the Rpt subunits by modifying the geometric relationship between the small and large AAA subdomains (27, 34, 35). In the density map of the S_A state, we identified nucleotide densities in all Rpt subunits (Fig. 3*A*). We tentatively modeled six ATPs into the atomic structure of the S_A state. However, a caveat is noted. Because cryo-EM density is built from averaging many single-particle images of individual molecular copies, it does not sufficiently differentiate potential coexistence of ADP and ATP in certain Rpt subunits in different copies of single particles. Further work to establish the nucleotide composition in different states is needed.

The AAA domains of Rpt1 to Rpt6 form a staircase beneath the OB ring, with Rpt3 located at the top position, Rpt2 at the bottom, and Rpt6 bridging the two in the S_A state (7–10) (*SI Appendix*, Figs. S10 and S11*A*). The large AAA subdomain contacts both the large and small AAA subdomains of the

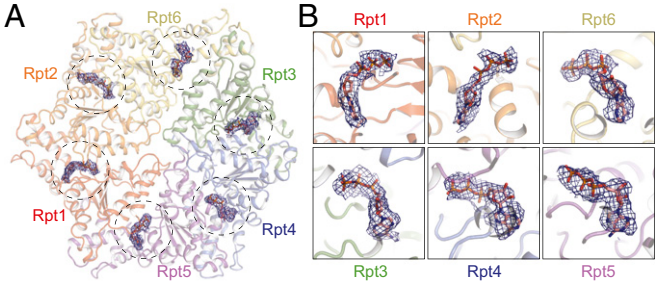


Fig. 3. Structure of the six nucleotide-binding sites of the proteasome. (*A*) Overview of six nucleotide-binding sites in the AAA-ATPase heterohexamer of the S_A state. Bound nucleotides are shown in stick representation superimposed with cryo-EM densities of the nucleotide that are shown in blue mesh. (*B*) Close-up views of nucleotide conformations in the six nucleotide-binding sites in S_A. ATP is tentatively modeled into the nucleotide density of each Rpt subunit in S_A.

neighboring subunit, forming an L-shaped intersubunit interface with a dihedral angle of about 120° (*SI Appendix*, Fig. S11*B*). The interfacial area between the two neighboring large AAA subdomains varies from 0.96 to 1.45 times that between the large and small AAA subdomains (*SI Appendix*, Fig. S11 *D* and *E*). The buried interface (~1,100 Å^2) between the AAA domains of Rpt3 and Rpt6 is half of the interfacial area between the other two adjacent Rpt subunits in S_A, suggesting that the former interface may be capable of greater structural rearrangement (26) (*SI Appendix*, Figs. S11*D* and S12). Indeed, Rpt3 demonstrates the greatest shift and rotation compared with other Rpt subunits during state transitions from S_A to S_C (*SI Appendix*, Fig. S12 and Table S3). In summary, the six AAA-ATPases are not arranged symmetrically; rather, a strong asymmetry in the intersubunit interfacial stability is a hallmark of the proteasomal AAA-ATPase heterohexamer (36).

During transitions from S_A to any of the alternative states, the OB ring behaves as a rigid body as its position slides above the AAA domains. By contrast, each Rpt AAA domain exhibits a distinct reconfiguration, manifested most obviously in different extents of hinge-like rotation between the small and large AAA subdomains (*SI Appendix*, Fig. S12 and Table S3). Both the OB and AAA rings translate and tilt to align axially with the CP channel (25–27). The current resolutions of the S_B, S_C, and S_D states are insufficient to resolve nucleotide densities. This leaves an open question as to whether and how ATPase repositioning is correlated with ATP hydrolysis cycles.

The Substrate-Translocation Pathway. The substrate-translocation channel in the center of the ATPase ring is mainly shaped by inward-facing pore loops (23, 36, 37). Each ATPase contains four pore loops, two on the OB domain and two on the AAA domain (Fig. 4). The right-handed helical architecture in the AAA channel gives rise to a constriction much narrower than the OB channel (Fig. 4 *E* and *F*). The interior of the AAA channel is largely negatively charged. By contrast, the interior of the OB channel is positively charged (*SI Appendix*, Fig. S11*C*). Both channels are dramatically enriched for tyrosine residues. The OB channel has six tyrosine residues (Tyr147 in Rpt1, Tyr72 and Tyr121 in Rpt6, Tyr111 in Rpt3, Tyr79 in Rpt4, and Tyr158 in Rpt5), with their oxygen atoms pointed toward the substrate-translocation pathway (Fig. 4*E*). The AAA channel has five tyrosine residues from Rpt1 to Rpt4 (Fig. 4*E* and *SI Appendix*, Fig. S13).

In the S_A state, pore-1 loops from Rpt3, Rpt4, Rpt5, Rpt1, and Rpt2 constitute a helical inner surface. These loops feature the conserved hydrophobic [Tyr/Phe]-[Val/Leu/Ile]-Gly sequence pattern that has been previously suggested to drive substrate translocation in many ATP-dependent unfoldases such as HslU, ClpX, ClpA, LonA, FtsH, and PAN (37, 38) (Fig. 4 *C* and *E* and *SI Appendix*, Figs. S13 *A* and *B* and S14). The pore-2 loops constitute a second helical inner surface running parallel to the pore-1 loops (Fig. 4 *B* and *D*). Note that these pore loops directly extend from the Walker B motif, which is involved in ATP hydrolysis (39). Unlike the hydrophobic pore-1 loops, the pore-2 loops are heavily populated with conserved, charged residues (seven glutamates and two aspartates) (Fig. 4*E* and *SI Appendix*, Fig. S13*C*).

The AAA channel in the S_A state has three narrow constrictions, defined by the axial area of overlap between the two pore-loop helices (Fig. 4*E*, *Inset*). In this region, Tyr207 and Ile208 from Rpt4, Phe260 and Ile261 from Rpt5, and Tyr249 and Val250 from Rpt1, all pore-1 loop residues, are paired with neighboring charged residues in the pore-2 loops of Rpt3, Rpt4, and Rpt5, respectively (Fig. 4 *E* and *F*). Interestingly, the phenotypic effects of pore-loop mutations are similar for the two members of each pore-loop pair. Thus, for the pore-1 loop, Rpt4 has the strongest phenotype, whereas for the pore-2 loop the partner Rpt3 has the strongest phenotype among the paired loops (36). Also, the strength of the phenotype diminishes with distance along the pore axis. As a result of pore-loop pairing, the channel radius is constricted to as little as ~2 Å (Fig. 4*F*). Among the pore loops, Rpt6 is uniquely displaced from the channel axis

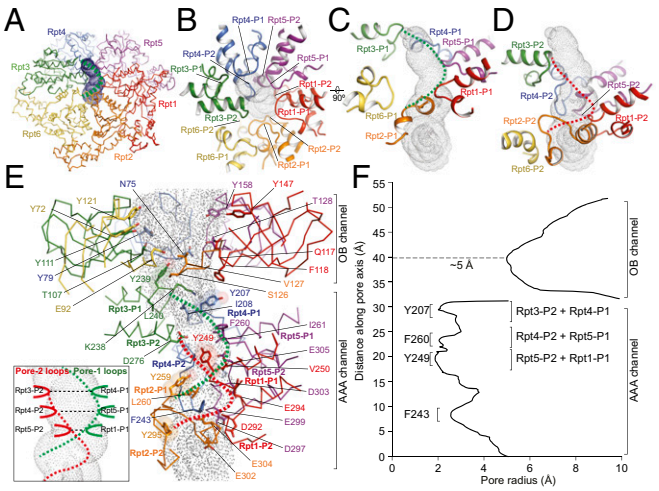


Fig. 4. Architecture of the substrate-translocation channel. (*A*) Overview of the AAA channel calculated by the HOLE program (50). The channel is rendered by surface dots. The dashed green curve indicates the spiral shape formed by the pore-1 loops. (*B*) Top view of the pore loops aligning along the channel axis from the perspective of the OB domain. (*C*) Close-up side views of the pore-1 loops from six Rpt subunits decorating the channel, which align along the channel in a spiral staircase formed from Rpt1 to Rpt5, with a backward recession in the Rpt6 pore-1 loop that is slightly away from the major channel pathway. The pore-1 loops form a helical part of the channel interior as illustrated by the dashed green line. (*D*) Close-up side view of the pore-2 loops from six Rpt subunits decorating the channel, which form a complete spiral staircase from Rpt1 to Rpt6. These pore-2 loops form another helical part of the channel interior, illustrated by the red dashed line. (*E*) Side view of the complete ATPase channel, including components from both the OB and AAA domains, calculated by the HOLE program (50). Side-chain patterns observed along the substrate-translocation pathway are highlighted. The five tyrosine residues, highlighted by transparent sphere representation, and a number of hydrophobic and negatively charged residues decorate the AAA channel; color codes for ATPase protomers match those shown in *A*. (*Inset*) A schematic cartoon showing that the pore-2 loops of Rpt3, Rpt4, and Rpt5 pair laterally with the pore-1 loops of Rpt4, Rpt5, and Rpt1, respectively, to form the three narrowest constrictions in the AAA channel. (*F*) The channel radius along the pore axis approximately estimated by HOLE (50), showing the three narrowest constrictions in the AAA channel but only one narrow constriction in the OB channel that is more than twice as wide as those of the AAA channel.

and does not contribute to these constrictions in the S_A and S_B states (Fig. 4 *C* and *D*). Consistent with the observed displacement, mutating the pore-2 loop of Rpt6 had little impact on protein-degradation rates (36).

RP–CP Interactions. The interface between the RP and CP is dominated by the associations of the hexameric AAA-ATPases and the heptameric α-ring (Fig. 5*A*). This symmetry-mismatched interface buries ~3,600 Å^2 of surface area in the S_A state (*SI Appendix*, Table S2). All Rpt subunits but Rpt6 are in direct contact with the α-ring (Fig. 5*A*). A large gap between Rpt6 and the α2-subunit is forced by the amino-terminal PCI domain of Rpn6, which itself contacts α2, preventing Rpt6 from directly contacting the α-ring (Fig. 5*A*). The amino-terminal PCI domain of Rpn6 extends down to the lateral surface of the α2-subunit, making an ~620-Å^2 contact, which is perhaps critical for stabilizing the lateral position of Rpn6 (*SI Appendix*, Fig. S15 *A* and *C*). The amino-terminal PCI domain of Rpn5 is also found lateral to the CP, although its contact area with α1 is small (~50 Å^2).

Previous biochemical studies on the yeast proteasome suggested that the C termini of Rpt2, Rpt3, and Rpt5 contain conserved C-terminal HbYX motifs that insert into pockets formed between adjacent α-subunits (40, 41). In the S_A state, the HbYX motifs of both Rpt3 and Rpt5 insert into the α1–α2 and

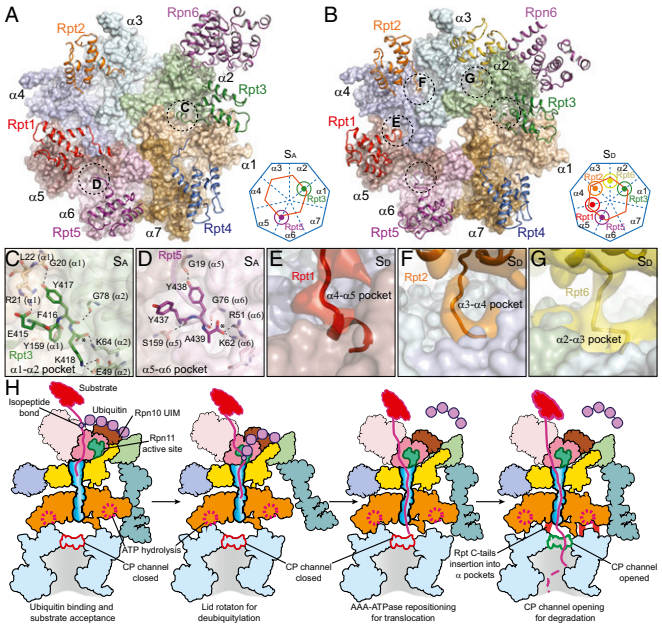


Fig. 5. RP–CP interface regulates gating of the CP channel. (*A*) Overview of the RP–CP interface in which the carboxyl-terminal tails of Rpt3 and Rpt5 are shown to insert into the α-pockets in the S_A state. Rpt6 is not shown, for clarity. The dashed circles mark the HbYX motifs of two Rpt subunits that insert into the α-pockets. (*B*) Overview of the RP–CP interface in which the carboxyl-terminal tails of Rpt1, Rpt2, Rpt6, Rpt3, and Rpt5 are shown to insert into the α-pockets in the S_D state. The dashed circles mark the HbYX motifs of five Rpt subunits that insert into the α-pockets. (*A* and *B*, *Insets*) Simplified illustrations of the tail–pocket interactions between the ATPase ring and CP. The orange hexagons and blue heptagons represent the ATPase and the α-ring, respectively. The small circles connected to the orange hexagon represent the carboxyl-terminal tails of the Rpt subunits inserted into the α-pockets. (*C*) Close-up view of interactions between the C-terminal HbYX motifs of Rpt3 and the α1–α2 pocket in the S_A state. (*D*) Close-up view of interactions between the C-terminal HbYX motifs of Rpt5 and the α5–α6 pocket in the S_A state. (*E*–*G*) Close-up view of interactions between the Rpt1 (*E*), Rpt2 (*F*), and Rpt6 (*G*) C termini and the α4–α5, α3–α4, and α2–α3 pockets in the S_D state, respectively. (*H*) Hypothetical structure-based model for substrate degradation by the human proteasome. UIM, ubiquitin-interacting motif.

α5–α6 pockets, respectively (Fig. 5 *C* and *D* and *SI Appendix*, Fig. S15 *G* and *H*). Consistent with previous structural data, Tyr417 in Rpt3 and Tyr438 in Rpt5 interact with Gly20-O of the α1-subunit and Gly19-O of the α5-subunit, respectively, through hydrogen bonding. The terminal carboxyl groups of Lys418 in Rpt3 and Ala438 in Rpt5 form hydrogen bonds with Lys64 of α2 (Fig. 5*C*) and Lys62 of α6 (Fig. 5*D*), respectively. These contacts are remarkably conserved, as they have previously been noted for homomeric proteasome precursors from archaea (41). By contrast, experimental density corresponding to the C-terminal HbYX tail (residues 434 to 440) of Rpt2 is missing, whereas there is no additional density in the α3–α4 pocket (*SI Appendix*, Fig. S15*D*). Similarly, the carboxyl-terminal tails of Rpt1 and Rpt4 do not insert into any α-pockets (*SI Appendix*, Fig. S15 *E* and *F*). These structural features may explain the previous finding that the extreme C termini of Rpt3 and Rpt5 are essential for assembly of the human proteasome (42, 43).

Gating of the CP Channel. In the states in which the CP channel is closed (i.e., S_A, S_B, and S_C), only Rpt3 and Rpt5 are found to have their C termini inserted into α-pockets (Fig. 5*A* and *SI Appendix*, Fig. S15 *G*–*J*). By contrast, in the S_D state, the carboxyl-terminal tails of Rpt1, Rpt2, and Rpt6 are inserted into the α4–α5, α3–α4, and α2–α3 pockets, respectively (Fig. 5 *E*–*G*). Thus, except for Rpt4, all Rpt carboxyl-terminal tails are inserted into α-pockets in the S_D state (*SI Appendix*, Fig. S15 *K*–*M*). The principal channel-blocking tails are from α2 and α4; the α3-tail, which behaves as a lynchpin of the gate (44), is controlled by Rpt2 (Fig. 2*N*) (40). The reorientation of these tails constitutes gating and controls substrate entry into the CP. Consistently, the AAA-ATPase heterohexamer is translated for ∼5 Å and rotates for ∼2° on the α-ring in the S_C state, and is further translated for ∼10 Å and rotated for ∼4° in the S_D state. Although the amino-terminal tails of α2, α3, and α4 are rotated over a large angle to roughly align along the heptameric axis to open the CP gate in the S_D state (Figs. 2 *O* and *P* and 5 *A* and *B*), the helical elements connected to the gate-blocking tails in the α-ring are nearly identical to those in other states. This is reminiscent of the observation of the open gate of the yeast CP in complex with the 11S regulator from *Trypanosoma brucei* (33, 41).

Comparison of the RP–CP interface among different conformational states reveals a structural mechanism for gating of the CP channel. The initial repositioning of the ATPase ring on the α-ring in the S_C state follows a large rotation of ∼40° in the lid in the S_B state. However, this movement is insufficient to allow the insertion of additional Rpt carboxyl-terminal tails into the α-pockets, leaving the CP channel closed in the S_C state. Further movement of the ATPase ring on the α-ring eventually allows the carboxyl-terminal tails of Rpt1, Rpt2, and Rpt6 to reach their nearest α-pockets. The insertion of these Rpt carboxyl-terminal tails into the α-pockets rotates the three gate-blocking tails of α2–α3–α4 to approximately align along the axial direction, which opens the CP channel in the S_D state. Importantly, our observations indicate that the open state of the CP channel in the human holoenzyme cannot be immediately achieved through RP–CP association. Instead, opening of the CP channel is primed through a series of coordinated, stepwise remodeling events around the RP–CP interface, including a lid reconfiguration, ATPase repositioning, and insertion of three additional Rpt carboxyl-terminal tails into the α-pockets (Fig. 5*H*).

Discussion

The coexistence of four conformational states, one of which is open in its CP gate under common solution conditions, provides insights into the dynamic regulation of the proteasome holoenzyme (Fig. 5*H*). We may assume S_A, the most populated state, to represent a ground state, whereas S_D, the state most distinct from S_A, to represent the translocating or fully engaged state of the proteasome, in accordance with prior studies of the yeast proteasome (7, 25–28) (*SI Appendix*, Fig. S8 *E* and *F*). By structural criteria (*SI Appendix*, Table S3), the S_B and S_C states may represent intermediates between S_A and S_D, and may describe a progression of conformational transitions that is temporally ordered: S_A is converted to S_B, then to S_C and S_D (Fig. 5*H*). These transitions, although apparently inherent to the proteasome, are expected to be guided by interactions with substrate, particularly the engagement of substrate by the AAA pore loops. The transition from S_A to S_B involves local dissociation of the Rpn5–Rpt6 and Rpn6–Rpt6 interfaces. This may topologically destabilize the lid conformation in the S_A state, allowing it to rotate around ATPase with a large angle (∼40°). The key consequence of this lid rotation is the repositioning of Rpn11 above the OB ring, which was thought to free the Rpn11 catalytic loop and mobilize its Ins-1 region (15, 18, 25). This suggests that Rpn11 may be enhanced in its deubiquitinating activity in the S_B, S_C, and S_D states (25).

Rearrangement of the lid may not only reposition Rpn11 but also perturb the RP–CP interface, including the lateral contact of Rpn6 to the CP. The dissolution of this contact may release a constraint on movement of the ATPase heterohexamer on the α-ring in the S_C and S_D states. The stepwise repositioning of the ATPases on the α-ring aligns axially the OB ring and the AAA channel with the CP channel (25–27), allowing a major change in the ATPase–CP interface, where additional Rpt carboxyl-terminal tails are inserted into the α-pockets, to eventually open the CP gate in the S_D state.

In both the S_C and S_D states, a major rearrangement of the AAA channel would allow the pore loops from Rpt3 and Rpt4, the first Rpt pair interacting with an incoming substrate, to move

toward the CP (*SI Appendix*, Fig. S12). The contrasting side-chain patterns of the pore loops—the hydrophobic pore-1 versus the highly charged pore-2 loops—suggest that they might play distinct roles in processive substrate translocation (36, 45). We speculate that the pore-1 loops may propel substrates forward along the AAA channel through hydrophobic interactions, whereas the pore-2 loops translate their interactions with substrates into enhancement of ATP hydrolysis (46), possibly by controlling the configuration of the conserved glutamate residue in the Walker B motif (39) (*SI Appendix*, Fig. S16).

Taken together, our structural data lead to a model where stepwise, distributed conformational changes within the holoenzyme are coordinated so as to generate a series of enzymatic regulatory events: The deubiquitinating activity of Rpn11 is enhanced in an S_B-like state; the ATPase activity is likely heightened in an S_C-like state, which can initiate substrate translocation; and eventually the CP gate is opened for degradation in a translocation-competent S_D-like state. In the presence of a substrate in the translocation channel, we would imagine that the open gate would be stabilized until completion of substrate degradation.

Confirmatory evidence of our ground-state structure is as follows: Coincident with this study, two other groups determined the high-resolution cryo-EM structures of the human proteasome holoenzyme in the ground state (47, 48). Their structures are largely consistent with our proteasome structure in the S_A state. Key structural features of their ground-state structures are virtually identical to our S_A state, including (*i*) the closed CP gate, (*ii*) the insertion of the carboxyl-terminal tails of Rpt3 and Rpt5 into the α-pockets, and (*iii*) occupancy of all six nucleotide-binding sites.

Materials and Methods

Human proteasomes were affinity-purified on a large scale from a stable HEK293 cell line (30). Cryo-EM data were collected with the supercounting mode of a Gatan K2 Summit direct electron detector installed on an FEI Tecnai Arctica microscope operating at a nominal magnification of 21,000× and an acceleration voltage of 200 kV. Details of experimental procedures and cryo-EM data processing are provided in *SI Appendix, Materials and Methods*.

ACKNOWLEDGMENTS. This work was funded in part by a grant from the NIGMS (GM026875; to M.W.K.), a grant of the Thousand Talents Plan of China (to Y.M.), an Intel academic grant (to Y.M.), a grant from the National Natural Science Foundation of China (91530321; to Y.M. and Q.O.), and an NIH grant (GM43601; to D.J.F.). The cryo-EM experiments were performed in part at the Center for Nanoscale Systems at Harvard University, which is supported by the National Science Foundation under NSF Award 1541959. The cryo-EM facility was funded through NIH Grant AI100645 and the Center for HIV/AIDS Vaccine Immunology and Immunogen Design.

1. Finley D, Chen X, Walters KJ (2016) Gates, channels, and switches: Elements of the proteasome machine. *Trends Biochem Sci* 41(1):77–93.
2. Tomko RJ, Jr, Hochstrasser M (2013) Molecular architecture and assembly of the eukaryotic proteasome. *Annu Rev Biochem* 82:415–445.
3. Kish-Trier E, Hill CP (2013) Structural biology of the proteasome. *Annu Rev Biophys* 42: 29–49.
4. Finley D (2009) Recognition and processing of ubiquitin-protein conjugates by the proteasome. *Annu Rev Biochem* 78:477–513.
5. Harshbarger W, Miller C, Diedrich C, Sacchettini J (2015) Crystal structure of the human 20S proteasome in complex with carfilzomib. *Structure* 23(2):418–424.
6. Huber EM, Groll M (2012) Inhibitors for the immuno- and constitutive proteasome: Current and future trends in drug development. *Angew Chem Int Ed Engl* 51(35): 8708–8720.
7. Beck F, et al. (2012) Near-atomic resolution structural model of the yeast 26S proteasome. *Proc Natl Acad Sci USA* 109(37):14870–14875.
8. Lasker K, et al. (2012) Molecular architecture of the 26S proteasome holocomplex determined by an integrative approach. *Proc Natl Acad Sci USA* 109(5):1380–1387.
9. Lander GC, et al. (2012) Complete subunit architecture of the proteasome regulatory particle. *Nature* 482(7384):186–191.
10. da Fonseca PC, He J, Morris EP (2012) Molecular model of the human 26S proteasome. *Mol Cell* 46(1):54–66.
11. Löwe J, et al. (1995) Crystal structure of the 20S proteasome from the archaeon *T. acidophilum* at 3.4 Å resolution. *Science* 268(5210):533–539.
12. Groll M, et al. (1997) Structure of 20S proteasome from yeast at 2.4Å resolution. *Nature* 386(6624):463–471.
13. Unno M, et al. (2002) The structure of the mammalian 20S proteasome at 2.75 Å resolution. *Structure* 10(5):609–618.
14. Pathare GR, et al. (2012) The proteasomal subunit Rpn6 is a molecular clamp holding the core and regulatory subcomplexes together. *Proc Natl Acad Sci USA* 109(1):149–154.
15. Pathare GR, et al. (2014) Crystal structure of the proteasomal deubiquitylation module Rpn8-Rpn11. *Proc Natl Acad Sci USA* 111(8):2984–2989.
16. Riedinger C, et al. (2010) Structure of Rpn10 and its interactions with polyubiquitin chains and the proteasome subunit Rpn12. *J Biol Chem* 285(44):33992–34003.
17. Wang Q, Young P, Walters KJ (2005) Structure of S5a bound to monoubiquitin provides a model for polyubiquitin recognition. *J Mol Biol* 348(3):727–739.
18. Worden EJ, Padovani C, Martin A (2014) Structure of the Rpn11-Rpn8 dimer reveals mechanisms of substrate deubiquitination during proteasomal degradation. *Nat Struct Mol Biol* 21(3):220–227.
19. Zhang N, et al. (2009) Structure of the S5a:K48-linked diubiquitin complex and its interactions with Rpn13. *Mol Cell* 35(3):280–290.
20. Dambacher CM, Worden EJ, Herzik MA, Martin A, Lander GC (2016) Atomic structure of the 26S proteasome lid reveals the mechanism of deubiquitinase inhibition. *eLife* 5:e13027.
21. Boehringer J, et al. (2012) Structural and functional characterization of Rpn12 identifies residues required for Rpn10 proteasome incorporation. *Biochem J* 448(1):55–65.
22. He J, et al. (2012) The structure of the 26S proteasome subunit Rpn2 reveals its PC repeat domain as a closed toroid of two concentric α-helical rings. *Structure* 20(3):513–521.
23. Zhang F, et al. (2009) Structural insights into the regulatory particle of the proteasome from *Methanocaldococcus jannaschii*. *Mol Cell* 34(4):473–484.
24. Zhang F, et al. (2009) Mechanism of substrate unfolding and translocation by the regulatory particle of the proteasome from *Methanocaldococcus jannaschii*. *Mol Cell* 34(4):485–496.
25. Luan B, et al. (2016) Structure of an endogenous yeast 26S proteasome reveals two major conformational states. *Proc Natl Acad Sci USA* 113(10):2642–2647.
26. Matyskiela ME, Lander GC, Martin A (2013) Conformational switching of the 26S proteasome enables substrate degradation. *Nat Struct Mol Biol* 20(7):781–788.
27. Śledź P, et al. (2013) Structure of the 26S proteasome with ATP-γS bound provides insights into the mechanism of nucleotide-dependent substrate translocation. *Proc Natl Acad Sci USA* 110(18):7264–7269.
28. Unverdorben P, et al. (2014) Deep classification of a large cryo-EM dataset defines the conformational landscape of the 26S proteasome. *Proc Natl Acad Sci USA* 111(15):5544–5549.
29. Estrin E, Lopez-Blanco JR, Chacón P, Martin A (2013) Formation of an intricate helical bundle dictates the assembly of the 26S proteasome lid. *Structure* 21(9):1624–1635.
30. Wang X, et al. (2007) Mass spectrometric characterization of the affinity-purified human 26S proteasome complex. *Biochemistry* 46(11):3553–3565.
31. Scheres SH, et al. (2007) Disentangling conformational states of macromolecules in 3D-EM through likelihood optimization. *Nat Methods* 4(1):27–29.
32. Scheres SH (2012) A Bayesian view on cryo-EM structure determination. *J Mol Biol* 415(2):406–418.
33. Whitby FG, et al. (2000) Structural basis for the activation of 20S proteasomes by 11S regulators. *Nature* 408(6808):115–120.
34. Sauer RT, Baker TA (2011) AAA+ proteases: ATP-fueled machines of protein destruction. *Annu Rev Biochem* 80:587–612.
35. Ogura T, Wilkinson AJ (2001) AAA+ superfamily ATPases: Common structure–diverse function. *Genes Cells* 6(7):575–597.
36. Beckwith R, Estrin E, Worden EJ, Martin A (2013) Reconstitution of the 26S proteasome reveals functional asymmetries in its AAA+ unfoldase. *Nat Struct Mol Biol* 20(10):1164–1172.
37. Glynn SE, Martin A, Nager AR, Baker TA, Sauer RT (2009) Structures of asymmetric ClpX hexamers reveal nucleotide-dependent motions in a AAA+ protein-unfolding machine. *Cell* 139(4):744–756.
38. Iosefson O, Nager AR, Baker TA, Sauer RT (2015) Coordinated gripping of substrate by subunits of a AAA+ proteolytic machine. *Nat Chem Biol* 11(3):201–206.
39. Zhang X, Wigley DB (2008) The 'glutamate switch' provides a link between ATPase activity and ligand binding in AAA+ proteins. *Nat Struct Mol Biol* 15(11):1223–1227.
40. Tian G, et al. (2011) An asymmetric interface between the regulatory and core particles of the proteasome. *Nat Struct Mol Biol* 18(11):1259–1267.
41. Stadtmueller BM, et al. (2010) Structural models for interactions between the 20S proteasome and its PAN/19S activators. *J Biol Chem* 285(1):13–17.
42. Kim YC, DeMartino GN (2011) C termini of proteasomal ATPases play nonequivalent roles in cellular assembly of mammalian 26 S proteasome. *J Biol Chem* 286(30): 26652–26666.
43. Lee SH, Moon JH, Yoon SK, Yoon JB (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(12):9269–9279.
44. Groll M, et al. (2000) A gated channel into the proteasome core particle. *Nat Struct Biol* 7(11):1062–1067.
45. Erales J, Hoyt MA, Troll F, Coffino P (2012) Functional asymmetries of proteasome translocase pore. *J Biol Chem* 287(22):18535–18543.
46. Peth A, Kukushkin N, Bossé M, Goldberg AL (2013) Ubiquitinated proteins activate the proteasomal ATPases by binding to Usp14 or Uch37 homologs. *J Biol Chem* 288(11):7781–7790.
47. Schweitzer A, et al. (2016) Structure of the human 26S proteasome at a resolution of 3.9 Å. *Proc Natl Acad Sci USA* 113(28):7816–7821.
48. Huang X, Luan B, Wu J, Shi Y (2016) An atomic structure of the human 26S proteasome. *Nat Struct Mol Biol* 23(9):778–785.
49. Wu J, et al. (2016) Unsupervised single-particle deep classification via statistical manifold learning. arXiv:1604.04539.
50. Smart OS, Neduvelil JG, Wang X, Wallace BA, Sansom MS (1996) HOLE: A program for the analysis of the pore dimensions of ion channel structural models. *J Mol Graph* 14(6):354–360, 376.

Supplementary Information

Structural basis for dynamic regulation of the human 26S proteasome

Shuobing Chen*, Jiayi Wu*, Ying Lu*, Yong-Bei Ma, Byung-Hoon Lee, Zhou Yu, Qi Ouyang, Daniel Finley, Marc W. Kirschner†, Youdong Mao†

*Equal contribution.
†correspondence to: youdong_mao@dfci.harvard.edu and marc@hms.harvard.edu

SI Materials and Methods

Protein expression and purification. Human proteasomes were affinity-purified on a large scale from a stable HEK293 cell line harboring HTBH (hexahistidine, TEV cleavage site, biotin, and hexahistidine) tagged hRPN11 (a gift from L. Huang, Departments of Physiology and Biophysics and of Developmental and Cell Biology, University of California, Irvine, California 92697) (1). The cells were Dounce-homogenized in a lysis buffer (50mM NaH_2PO_4 [pH7.5], 100mM NaCl, 10% glycerol, 5mM $MgCl_2$, 0.5% NP-40, 5mM ATP and 1mM DTT) containing protease inhibitors. Lysates were cleared, then incubated with NeutrAvidin agarose resin (Thermo Scientific) overnight at 4°C. The beads were then washed with excess lysis buffer followed by the wash buffer (50mM Tris-HCl [pH7.5], 1mM $MgCl_2$ and 1mM ATP). Usp14 was removed from proteasomes using wash buffer +150mM NaCl for 30 min. 26S proteasomes were eluted from the beads by cleavage, using TEV protease (Invitrogen). The doubly-capped proteasome was enriched by gel filtration on a Superose 6 10/300 GL column at a flow rate of 0.15ml/minute in buffer (30mM Hepes pH7.5, 60mM NaCl, 1mM $MgCl_2$, 10% Glycerol, 0.5mM DTT, 0.8mM ATP). Gel-filtration fractions were concentrated to about 2mg/ml. Right before cryo-EM sample preparation, the proteasome sample was buffer-exchanged into 50mM Tris-HCl [pH7.5], 1mM $MgCl_2$, 3mM ATP, 0.5mM TCEP, to remove glycerol, and was supplemented with 0.005% NP-40.

Data collection. A 3-μl drop of 1.5 mg/ml proteasome solution was applied to a glow-discharged C-flat grid (R1/1 and R1.2/1.3, 400 Mesh, Protochips, CA, USA), blotted for 2 sec, then plunged into liquid ethane and flash-frozen using the FEI Vitrobot Mark IV. The cryo-plunging conditions combined with different NP-40 concentrations, ranging from 0.001% to 0.01%, were systematically screened to make an appropriate ice thickness in the carbon holes to mitigate the orientation preference of proteasome particles. The cryo-grid was imaged in an FEI Tecnai Arctica microscope, equipped with an Autoloader, at a nominal magnification of 21,000 times and an acceleration voltage of 200 kV. Coma-free alignment was manually conducted prior to data collection. Cryo-EM data were collected semi-automatically by the Leginon (2) version 3.1 on the Gatan K2 Summit direct detector camera (Gatan Inc., CA, USA) in a super-resolution counting mode, with 6.0 s of total exposure time and 200 ms per frame. This resulted in movies of 30 frames per exposure and an accumulated dose of 30 electrons/$Å^2$. The calibrated physical pixel size and the super-resolution pixel size are 1.72 Å and 0.86 Å, respectively. The defocus in data collection was set in the range of -1.0 to -3.0 μm. A total of 10,367 movies were collected, among which 8,075 movies were selected for further data analysis after screening and inspection of data quality.

Cryo-EM data processing and reconstruction. The raw movie frames were first corrected for their gain reference and each movie was used to generate a micrograph that was corrected for sample movement and drift with the MotionCorr program (3). These drift-corrected micrographs were used for the determination of actual defocus of each micrograph with the CTFFind3 program (4). Reference-free 2D classification were done in both RELION 1.3 and a program that was developed recently, ROME (5), which combines the maximum-likelihood based image alignment and statistical manifold learning-based classification (6). 3D classification and high-resolution refinement were conducted in RELION 1.3. 528,196 particles of doubly-capped proteasome were automatically selected using a deep-learning program before extracting them for structure determination (7). The initial model was generated in EMAN2 (8). A subset of 32,645 particles was used for reference-free classification performed by e2refine2d.py, generating 80 2D class averages. 36 class averages were selected for the generation of an initial model using e2initialmodel.py.

All 2D and 3D classifications were done at the pixel size of 1.72 Å. After the first round of reference-free 2D classification, bad particles were rejected as a whole class upon inspection of class average quality, which left 369,020 particles. The initial model, low-pass filtered to 60 Å, was used as the input reference to conduct unsupervised 3D classification into 8 classes without assumption of any symmetry, using an angular sampling of 7.5° and a regularization parameter T of 4. Five classes, counting to 67.4% particles, showed nearly identical conformations. These classes were further sorted by reference-free 2D classification separately, and 237,083 good particles were selected. Auto-refinement in RELION with imposing C2 symmetry were performed at the counting mode with a pixel size of 1.72 Å. A local mask focusing on the CP were applied in the local search step. Based on the in-plane shift and Euler angle of each particle, we reconstructed the two half-maps using single-particle images at the super-counting mode with a pixel size of 0.86 Å. The CP in this doubly-capped proteasome reconstruction yielded its best resolution, measured by a gold-standard FSC to 3.6 Å at FSC-0.143 cutoff (Fig. S2C), whereas the RP exhibited a much lower resolution, with the worst part in the lid subcomplex. To improve the local resolution of ATPase, we further classified the 237,083 particles into 4 classes by 3D maximum-likelihood method in RELION. A class containing 85,420 particles shows the same conformation of ATPase in both RPs. Auto-refinement of this good class in RELION with imposing C2 symmetry and local mask including the CP and ATPase components at the local search step were performed at the counting mode with a pixel size of 1.72 Å. Based on the in-plane shift and Euler angle of each particle, we reconstructed the two half-maps using single-particle images at the super-counting mode with a pixel size of 0.86 Å, which yielded a gold-standard resolution of 3.8 Å at FSC-0.143 cutoff (9) after adjustment using the high-resolution noise replacement method (10) (Fig. S2D). This map exhibited improved density quality for the RP, particularly for the ATPase component that was measured to 4 Å by a gold-standard FSC with a soft mask focusing on the ATPase component (Fig. S2C), whereas the lid subcomplex still presented a resolution no better than 6 Å.

To study the conformational dynamics of RP as well as improve the resolution of the lid subcomplex, 354,130 RP-CP subcomplex particles were manually selected from 8075 micrographs in EMAN2. We segmented the density of RP-CP subcomplex from the doubly-capped proteasome reconstruction and used it as the initial model for 3D classification in RELION. After an initial round of reference-free 2D classification, bad particles were rejected as a whole class upon inspection of class average quality. The initial model, low-pass filtered to 60

Å, was used as the input reference to conduct unsupervised 3D classification into 6 classes without assumption of any symmetry, using an angular sampling of 7.5° and a regularization parameter T of 4.0. All classes were further sorted by reference-free 2D classification separately, and 326,290 good particles were selected. We combined these particles for another round of 3D classification and obtained six classes. Particles in each class were manually examined after reference-free 2D classification to reject the projections of apparently incomplete proteasome assembly including free CP and free RP. All 3D classes were examined manually in Chimera for their structural integrity and were compared through volume alignment to identify any conformational difference. One of the classes was found with the CP gate opened; the second class containing 44,036 particles was named F_2 dataset that is lack of Rpn10 density. The rest four classes containing 235,037 particles showed the conformation of a closed CP gate and therefore were used for an auto-refinement procedure to prepare a high-quality consensus reconstruction for 20S density subtraction.

The auto-refinement in RELION resulted in a consensus reconstruction at a pixel size of 1.72 Å. Based on the refinement result, we re-centered the particle using the refined Euler angles and x/y-shifts and re-extracted particles with a pixel size of 0.86 Å at a super-counting mode. To improve the structural homogeneity of the single-particle dataset of RP-CP subcomplex, we conducted further 3D classification based on density subtraction. First, we low-pass filtered the two half maps of the refined consensus reconstruction to 8.0 Å separately and segmented the CP component of each filtered half map that was used to create the soft mask (with 9-pixel fall-off). We used SPIDER command "MMC" to apply mask with soft edge to each unfiltered half map. Next, we use these 2 masked unfiltered maps as input to use the RELION command "relion_project" to subtract the CP's intensity from each raw particle image. Thus, we obtained 235,037 density-subtracted particles. We used these subtracted particles to conduct 3D classification and then used raw images to perform auto-refinement. All the 3D classes were further refined separately for their density map by the auto-refinement procedure using un-subtracted particles, allowing detailed comparison of features that may reflect the conformational changes. In the first round of density subtracted 3D classification, two classes with the largest particle populations were further classified into four classes separately, and three classes of the smallest particle populations, containing 85,169 particles in total, were combined to do the auto-refinement using un-subtracted particles. One of the classes, lack of several RP subunits (Rpn5, Rpn6, Rpn9 and Rpn10), was discarded. We used the 85,169 un-subtracted particles to conduct the density subtraction 3D classification again. Finally, we obtained 13 classes based on density-subtracted 3D classification. Seven classes containing 139,236 particles were found to have the same conformation and were combined as the S_A dataset. Two classes containing 33,712 particles were designated the F_1 dataset, and another class containing 16,613 particles was designated the F_3 dataset. The F_1, F_2 and F_3 datasets all lack of Rpn10 density. One of the classes showed an open CP gate, thus was combined with the previously identified open-gate class together and then classified into four classes. Three classes containing 18,443, 10,622 and 14,382 particles were designated the S_B, S_C and S_D datasets, respectively.

Throughout the classification procedure, three principle criteria were used for manual decision-making: (1) the classes showing incomplete assemblies were discarded without further analysis; (2) the classes showing complete assemblies and similar conformation were combined for further analysis; (3) the final auto-refinement was performed when there was no more separable structural difference found in the selected, regrouped classes. The final auto-refinements of the S_A, S_B, S_C and S_D dataset resulted in the reconstruction with an overall

resolution of 4.4 Å, 6.8 Å, 8.0 Å and 8.0 Å measured by a gold-standard FSC at 0.143-cutoff and further adjusted by the high-resolution noise replacement method (10) (Fig. S2E). The lid subcomplex in this S_A map was measured to 4.9 Å by a gold-standard FSC with a soft mask on the lid (Fig. S2C). The map segmentation for the target region was conducted in Chimera (11), and the segmented target regions were used to calculate the soft masks in SPIDER (12). Prior to visualization, all density maps were sharpened by applying a negative B-factor (see Table S1). Local resolution variations were estimated using ResMap (13) and Bsoft (14) package.

To further validate the reconstructions from the complex 3D classification procedure, two different initial models, the S_A-state map and S_D-state map, both lower-pass filtered to 60 Å, were used to conduct the auto-refinement (Fig. S17). The cross correlation value between the refined RP-CP maps in the S_A, S_B, S_C and S_D states using different initial models is 0.989, 0.994, 0.986, 0.992, respectively. This suggests that the RP-CP reconstructions in four distinct conformational states are robust and free of initial model bias.

Atomic model building and refinement. To build the initial atomic model of the proteasome, we used homology modeling in Modeller (15) to generate a starting coordinate file whenever there was a homology structure that had been solved by X-ray crystallography or NMR, and then manually improved the main-chain and side-chain fitting in Coot (16). For non-ATPase subunits, the crystal structures of yeast Rpn2, Rpn6, Rpn8, Rpn10, Rpn11 and Rpn12 (17-22) were used to produce the initial coordinates for these subunits for manual fitting in Coot (16). The crystal structure of COP9 was used to generate the initial coordinates for Rpn3, Rpn5, Rpn7 and Rpn9 subunits (17). The crystal structure of PAN was used to generate the initial coordinates for all six Rpt subunits (18). The initial main chains of Rpn8, Rpn10 and Rpn11 were fit closely into the densities. A considerable fraction of the initial main chains of Rpn2, Rpn3, Rpn5, Rpn7 and six Rpt subunits were also fit well into the densities. The known RP subunit or homology structures used in the initial model building are summarized in Table S1B. Manual fitting adjustment was conducted to complete the initial model building in Coot prior to atomic model refinement (16). For the densities that are not accounted after the aforementioned fitting, the backbones were traced manually in Coot without any reference (16). The crystal structure of human CP (23) was found to directly fit well into our cryo-EM density map of S_A and was used as the initial atomic model for further atomic model refinement. The atomic model does not include Rpn1, which has a low resolution in its cryo-EM density insufficient for atomic modeling, and Rpn13, which was absent in our purified proteasome complex. The initially built atomic models were refined locally in real space in Coot, with the geometry quality monitored by the validation tools included in the Coot package (16).

Atomic model refinement was first carried out on nonATPase and ATPase/CP components separately in real space using Phenix (24) with secondary structure and geometry restraints to prevent over-fitting, using the 26S map with C2 symmetry imposed and the S_A map lowered-pass filtered at the nominal resolution and without B-factor sharpening. The two parts of density maps were segmented in Chimera. The cryo-EM map and the atomic model were placed into a pseudo-unit cell and the pseudo-crystallographic refinement was performed in Phenix in Fourier space using both amplitudes and phases. Multiple rounds of atomic model refinement in Phenix and model rebuilding in Coot were performed to improve the goodness of fit. The final refinement statistics are shown in Table S1A. The FSC curve between the map and the final refined atomic model gave an estimated resolution of 3.8 Å at FSC = 0.5. Overfitting of the overall model was monitored by refining the model in one of the two independent maps from the

gold-standard refinement approach, and testing the refined model against the other map. All refinements in Phenix were done with the cryo-EM maps that were low-pass filtered at the nominal resolution without B-factor sharpening or amplitude correction. The pseudo-atomic model of Rpn1 in S_A were built through homology modeling using Rpn2 structure as a reference in Modeller (15). The pseudo-atomic models of the S_B, S_C, S_D and Rpn1 of S_A were fitted in Coot and energetically optimized in Phenix.

Structural analysis and visualization. Structural comparison was conducted in Pymol (23) and Chimera. Interaction analysis between adjacent subunits was performed using PISA (25). The peptide-conducting channel calculated by HOLE (26) with OB ring and ATPase separately. All figures of the structures were plotted in Chimera and Pymol (27).

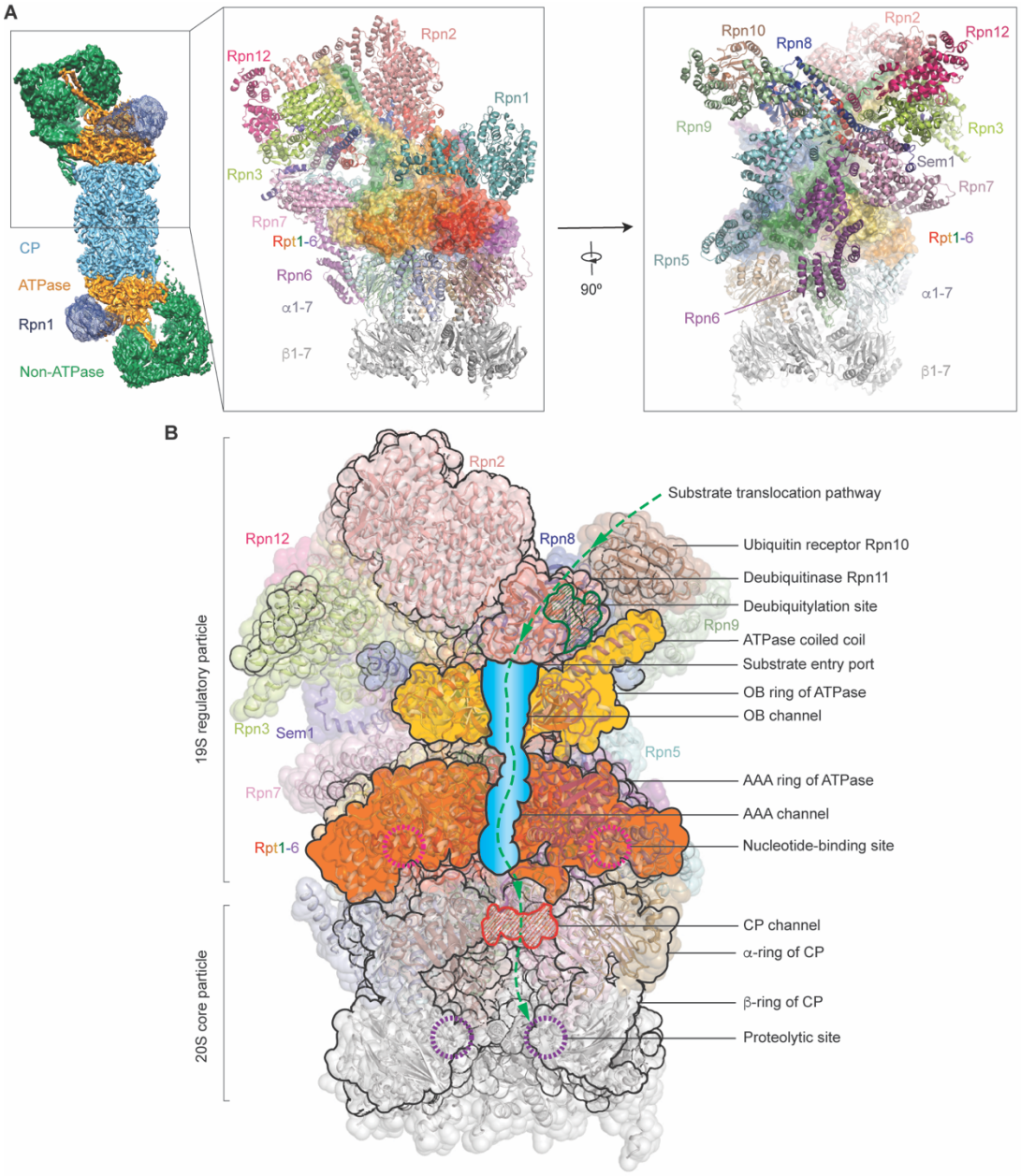


Fig. S1. Overview of the human 26S proteasome assembly. (**A**) Overview of the protein subunit organization in the proteasome holoenzyme from two orthogonal perspectives. The lid subcomplex includes Rpn3, Rpn5, Rpn6, Rpn7, Rpn8, Rpn9, Rpn11, Rpn12, Sem1; the base subunits include Rpn1, Rpn2, Rpt1, Rpt2, Rpt3, Rpt4, Rpt5 and Rpt6. The core particle is composed of seven α subunits α1-7 and seven β subunits β1-7. (**B**) Illustrative anatomy of the RP-CP subcomplex structure, with black silhouettes representing the central cross-section along the ATPase channel colored opaque blue. Dashed green curve illustrates the substrate-translocation pathway across the channel to the proteolytic sites in the CP chamber.

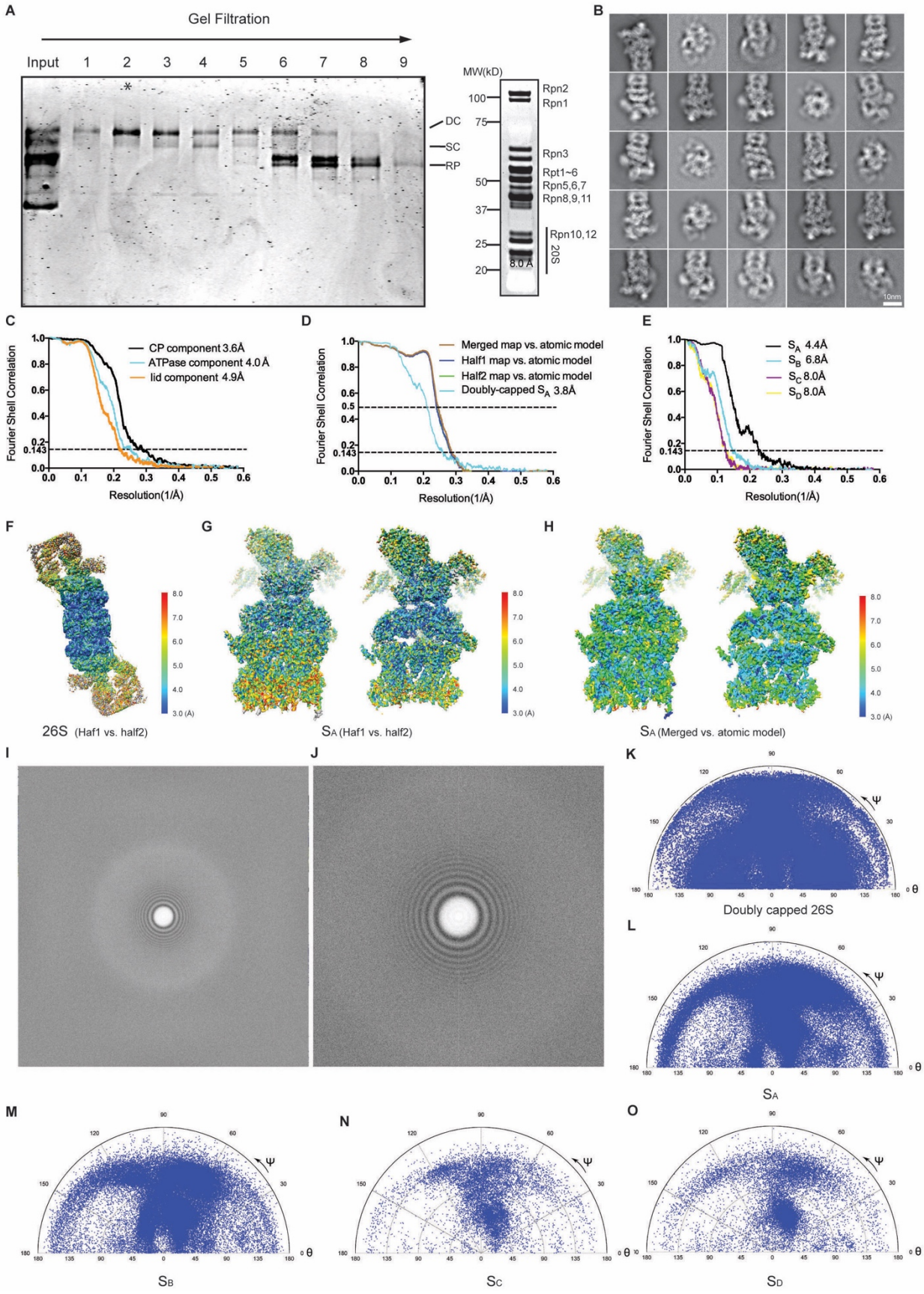

A
Gel Filtration
Input 1 2 3 4 5 6 7 8 9
DC
SC
RP
MW(kD)
100
75
50
37
25
20
Rpn2
Rpn1
Rpn3
Rpt1~6
Rpn5,6,7
Rpn8,9,11
Rpn10,12
20S
8.0 Å
B
10nm
C
Fourier Shell Correlation
CP component 3.6Å
ATPase component 4.0 Å
lid component 4.9Å
Resolution(1/Å)
D
Fourier Shell Correlation
Merged map vs. atomic model
Half1 map vs. atomic model
Half2 map vs. atomic model
Doubly-capped SA 3.8Å
Resolution(1/Å)
E
Fourier Shell Correlation
SA 4.4Å
SB 6.8Å
SC 8.0Å
SD 8.0Å
Resolution(1/Å)
F
26S (Haf1 vs. half2)
G
SA (Haf1 vs. half2)
H
SA (Merged vs. atomic model)
I
J
K
Doubly capped 26S
L
SA
M
SB
N
SC
O
SD

Fig. S2. Purification and cryo-EM structure determination of the human proteasome. (**A**) Nondenaturing PAGE analysis of fractions collected from a Superose 6 10/300 GL column (left), and SDS-PAGE analysis results of pooled fractions that were rich in doubly-capped proteasomes (right). (**B**) Typical reference-free 2D class averages of the RP-CP subcomplex computed by the ROME software package (5). (**C**) Gold-standard Fourier-shell correlation (FSC) plot that gives an estimate of the resolutions of the three components, the CP, the ATPase hexamer and the lid, each measured from its best resolved map. The resolution of CP was measured from the 237,083-particle double-cap map with C2 imposed; the resolution of the ATPase hexamer was measured from the 85,420-particle double-cap S_A map with C2 imposed; and the resolution of the lid was measured from the 139,236-particle RP-CP S_A map. (**D**) FSC plots give an estimate of the resolution of the doubly-capped map in the S_A state from the 85,420-particle dataset. Black curve shows the gold-standard FSC of the doubly-capped proteasome in the S_A state. This FSC curve was adjusted with the high-resolution noise replacement method (10) to exclude any effect of either overfitting in refinement or the effect of soft masking in FSC calculation. Orange, green and blue curves show the FSC cross-validation using the atomic model refined against the two half-maps and the merged map refined with 85,420-particle dataset. (**E**) The gold-standard, high-resolution noise replacement-adjusted FSC plots estimate the resolutions of refined cryo-EM maps of the RP-CP subcomplex in the S_A, S_B, S_C and S_D state. (**F**) Local resolution measurement of the doubly-capped proteasome map by ResMap (13) using the two half-maps separately refined in a gold-standard procedure. (**G**) Local resolution measurement of the RP-CP subcomplex map in the S_A state by ResMap between the two half-maps separately refined in a gold-standard procedure. Left, the complete view; right, the central cross-section of the map. (**H**) Local resolution measurement by Bsoft (program blocres) (14) between the refined map of the RP-CP subcomplex in the S_A state and the atomic model refined against the cryo-EM map. Left, the complete view; right, the central cross-section of the map. (**I**) and (**J**) The power spectrum of a typical drift-corrected cryo-EM micrograph at -1.8 μm defocus at the super-resolution mode (panel **I**) and the counting mode (panel **J**), respectively. (**K**)-(**O**) Angular distribution of the refined 3D reconstructions of the doubly-capped proteasome (panel **K**), the RP-CP subcomplex in the S_A (panel **L**), S_B (panel **M**), S_C (panel **N**) and S_D (panel **O**) states.

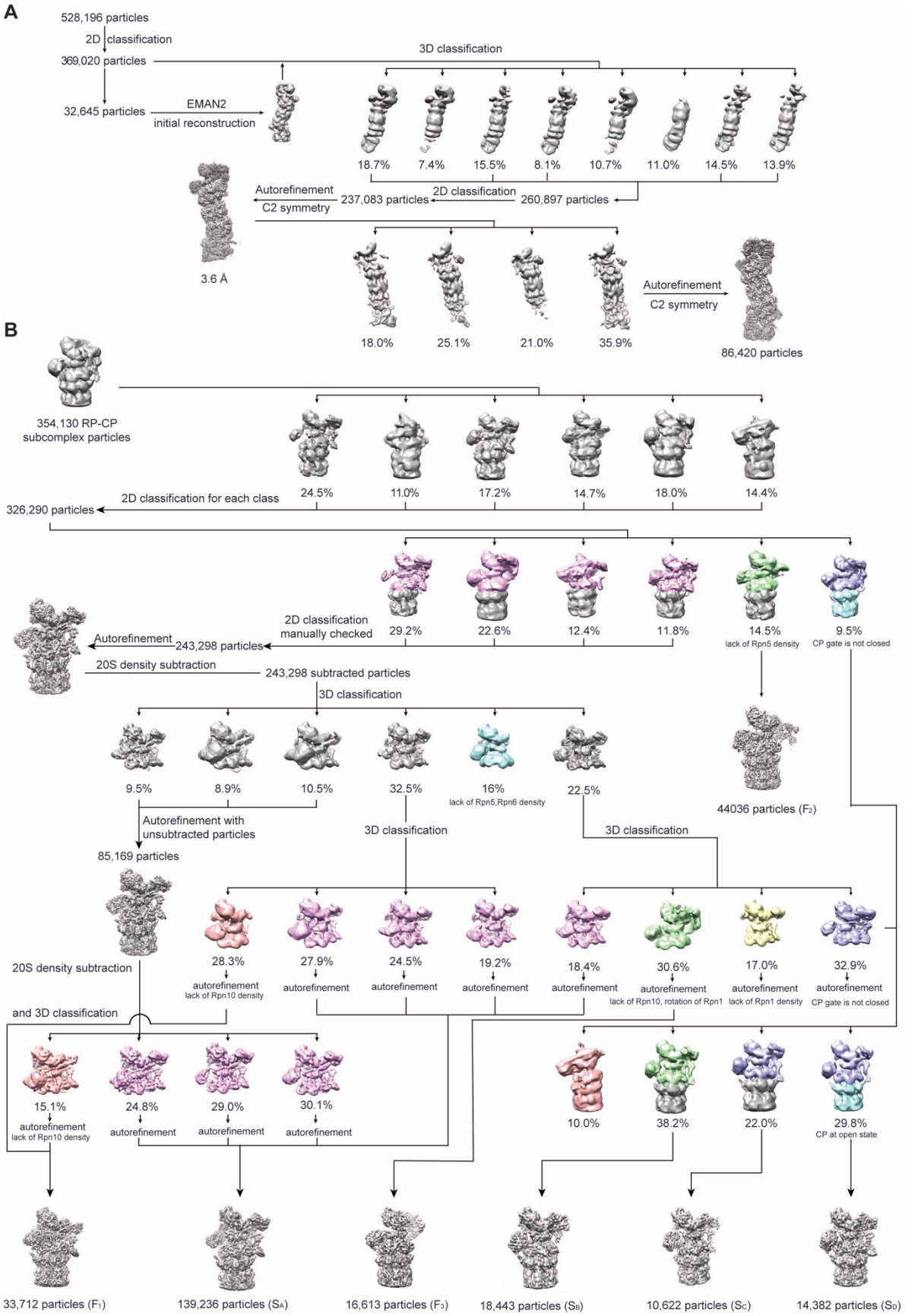

A
528,196 particles
2D classification
369,020 particles
32,645 particles
EMAN2
initial reconstruction
3D classification
18.7%
7.4%
15.5%
8.1%
10.7%
11.0%
14.5%
13.9%
2D classification
260,897 particles
237,083 particles
Autorefinement
C2 symmetry
3.6 Å
18.0%
25.1%
21.0%
35.9%
Autorefinement
C2 symmetry
86,420 particles
B
354,130 RP-CP
subcomplex particles
24.5%
11.0%
17.2%
14.7%
18.0%
14.4%
2D classification for each class
326,290 particles
29.2%
22.6%
12.4%
11.8%
14.5%
lack of Rpn5 density
9.5%
CP gate is not closed
2D classification
manually checked
Autorefinement
243,298 particles
20S density subtraction
243,298 subtracted particles
3D classification
9.5%
8.9%
10.5%
32.5%
16%
lack of Rpn5,Rpn6 density
22.5%
44036 particles (F2)
Autorefinement with
unsubtracted particles
85,169 particles
3D classification
3D classification
20S density subtraction
28.3%
autorefinement
lack of Rpn10 density
27.9%
autorefinement
24.5%
autorefinement
19.2%
autorefinement
18.4%
autorefinement
30.6%
autorefinement
lack of Rpn10, rotation of Rpn1
17.0%
autorefinement
lack of Rpn1 density
32.9%
autorefinement
CP gate is not closed
and 3D classification
15.1%
autorefinement
lack of Rpn10 density
24.8%
autorefinement
29.0%
autorefinement
30.1%
autorefinement
10.0%
38.2%
22.0%
29.8%
CP at open state
33,712 particles (F1)
139,236 particles (SA)
16,613 particles (F3)
18,443 particles (SB)
10,622 particles (SC)
14,382 particles (SD)

Fig. S3. 3D classification and refinement of a homogeneous dataset for high-resolution reconstruction. Schematic procedure of the maximum-likelihood based 2D and 3D classification that helped to improve the homogeneity of the particle dataset of the doubly-capped proteasome (**A**) and the RP-CP subcomplex (**B**). After initial rounds of 3D classification interleaved with 2D classification, the CP-subtracted particles were used for 3D classification, but the original, non-subtracted particle images were used in high-resolution refinement. During the classification, the classes showing incomplete assemblies were discarded. The classes of complete assemblies with similar or same conformations were combined for either further classification or auto-refinement.

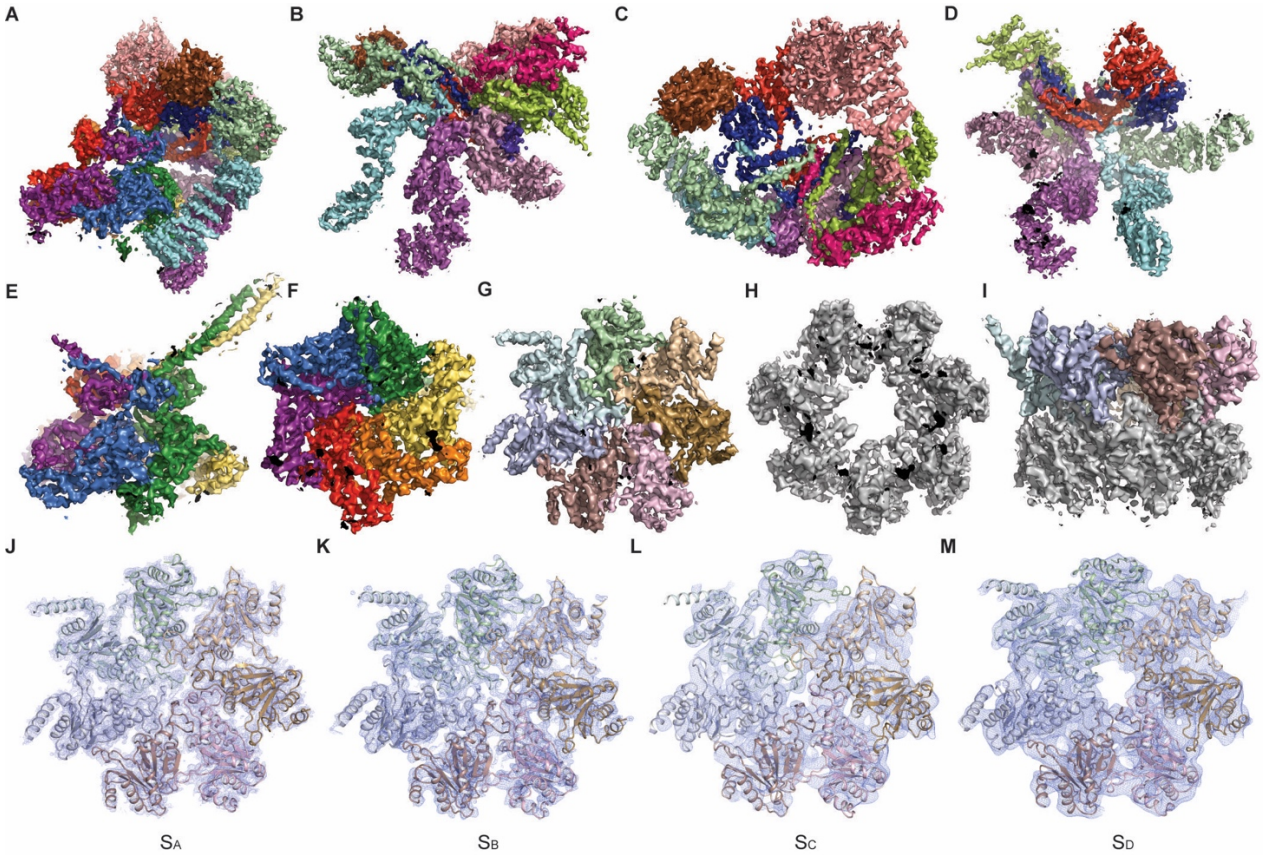


Fig. S4. Overview of cryo-EM density map. (**A**)-(**I**) The density map of the S_A state is shown in a solid surface representation for the side view of the complete RP (panel **A**), the side view of the non-ATPase RP components (panel **B**), the top view of the non-ATPase RP components (panel **C**), the bottom view of the lid (panel **D**), the side view of the ATPase ring (panel **E**), the bottom view of the ATPase ring (panel **F**), the top view of the α ring (panel **G**), the top view of the β ring (panel **H**) and the side view of the half CP (panel **I**). (**J**)-(**M**) The density map of the CP from the perspective of ATPase ring in the S_A (panel **J**), S_B (panel **K**), S_C (panel **L**) and S_D (panel **M**), showing that the CP channel is closed in S_A, S_B and S_C and is open in S_D.

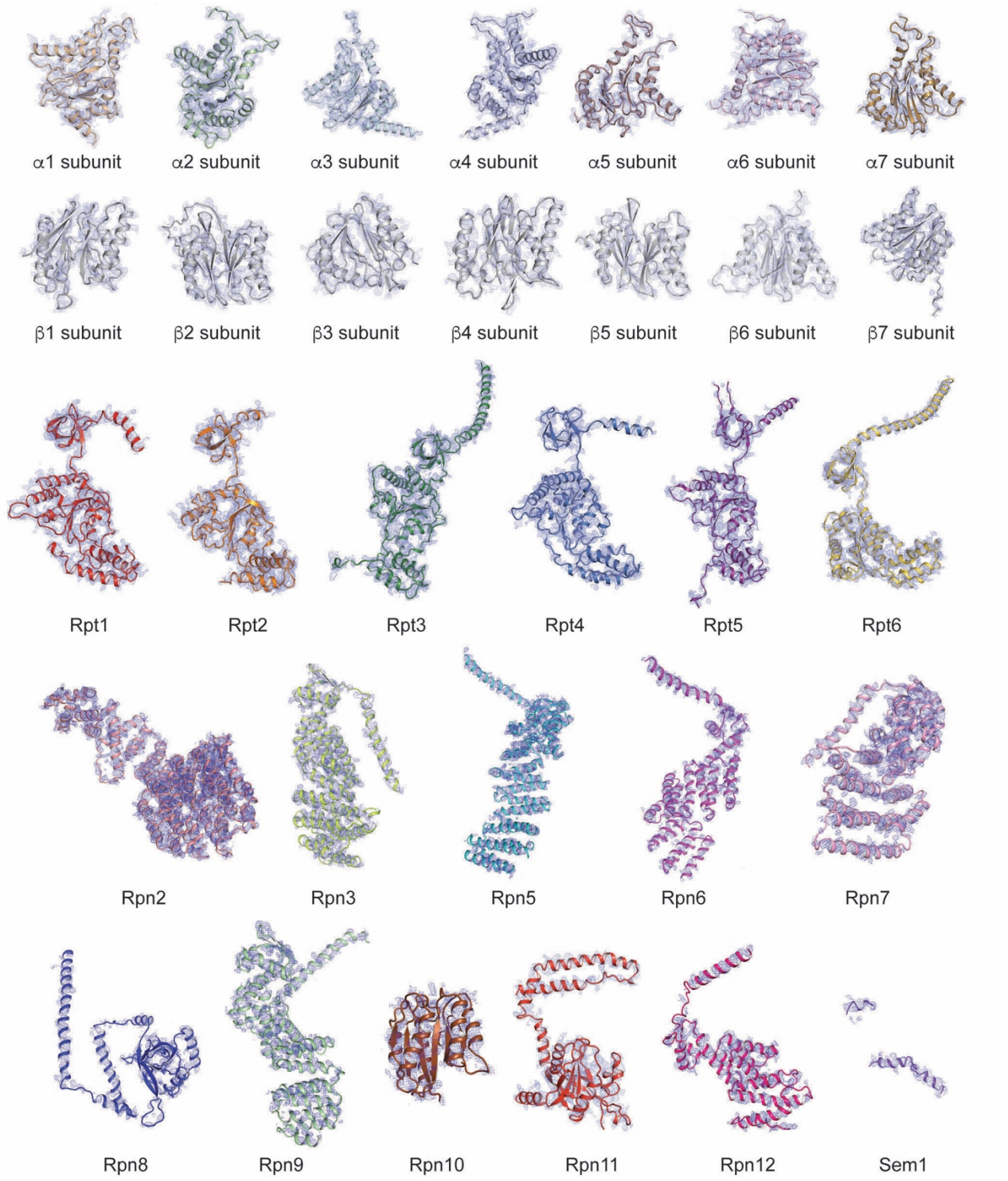


Fig. S5. Cryo-EM densities of subunit structures in the S_A state. The atomic models of each subunit structure in the S_A state, shown as a cartoon representation, are superimposed with the corresponding cryo-EM densities shown in blue meshes.

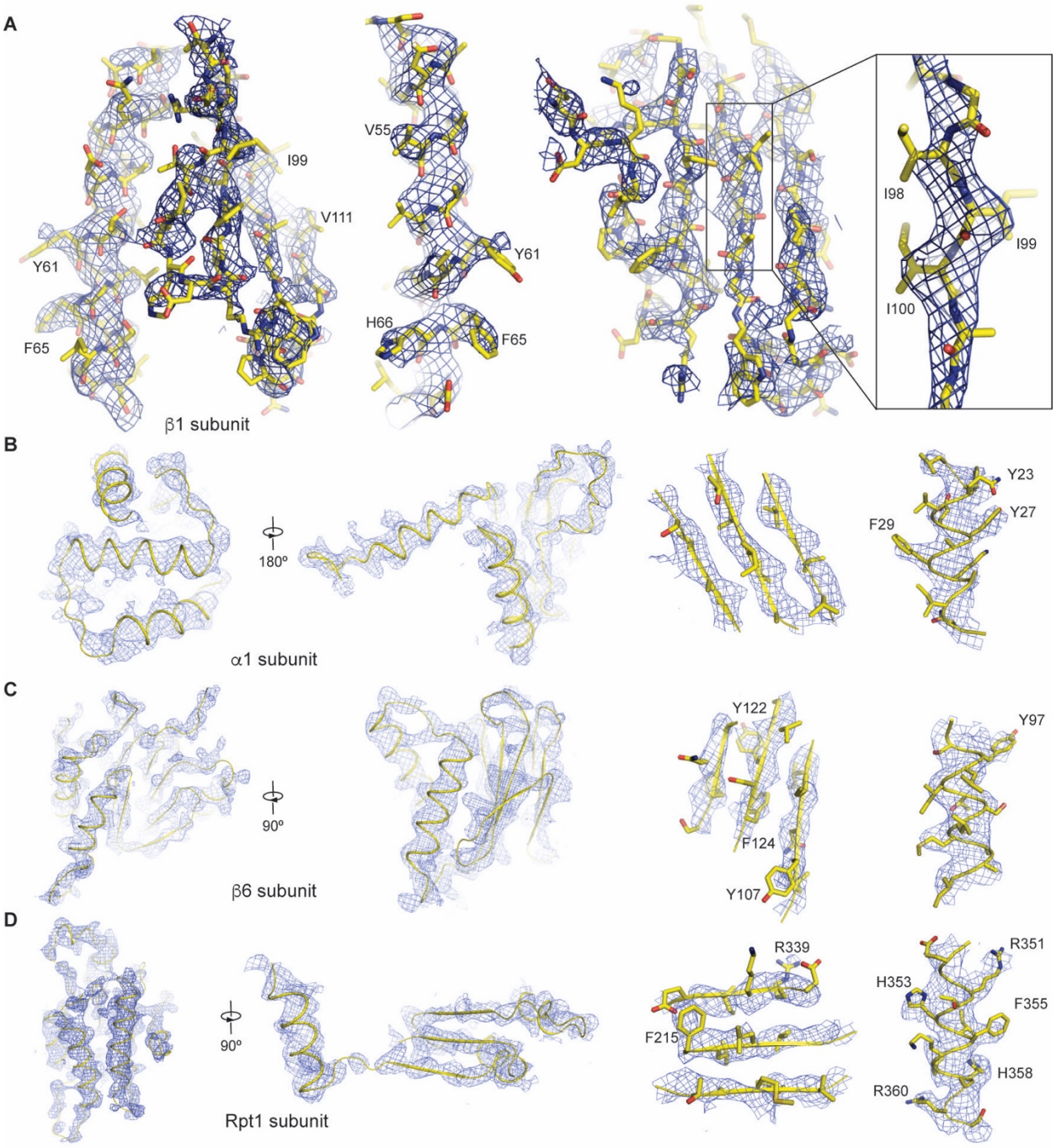


Fig. S6. Fitting of atomic structures into high-resolution cryo-EM densities of the 26S proteasome in the S_A state. The atomic models of subunit structures in the 26S proteasome, β1 subunit (panel **A**), α1 subunit (panel **B**), β6 subunit (panel **C**) and Rpt1 subunit (panel **D**), are superimposed with the corresponding cryo-EM densities shown in blue meshes. For each subunit, a few secondary structures including both α helices and β sheets are shown. Residues of large side chains are labelled.

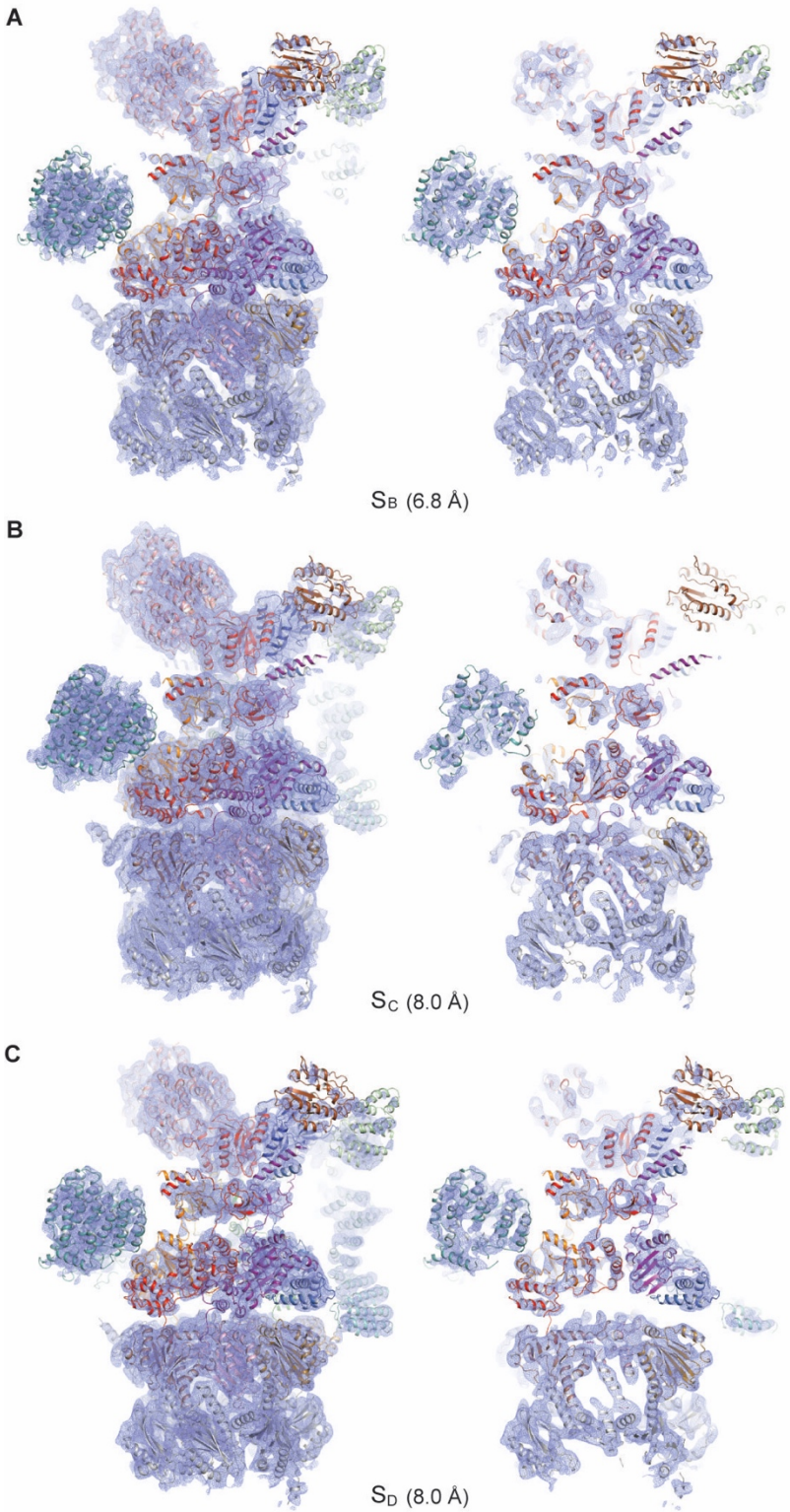


Fig. S7. Cryo-EM densities of the RP-CP subcomplexes in the S_B, S_C and S_D states. (**A**)-(**C**), The density maps of the S_B (panel **A**), S_C (panel **B**) and S_D (panel **C**) states at 6.8, 8.0 and 8.0 Å resolution are shown in blue meshes and superimposed with corresponding pseudo-atomic models. For each panel, the left shows the complete perspective, whereas the right shows the central slice.

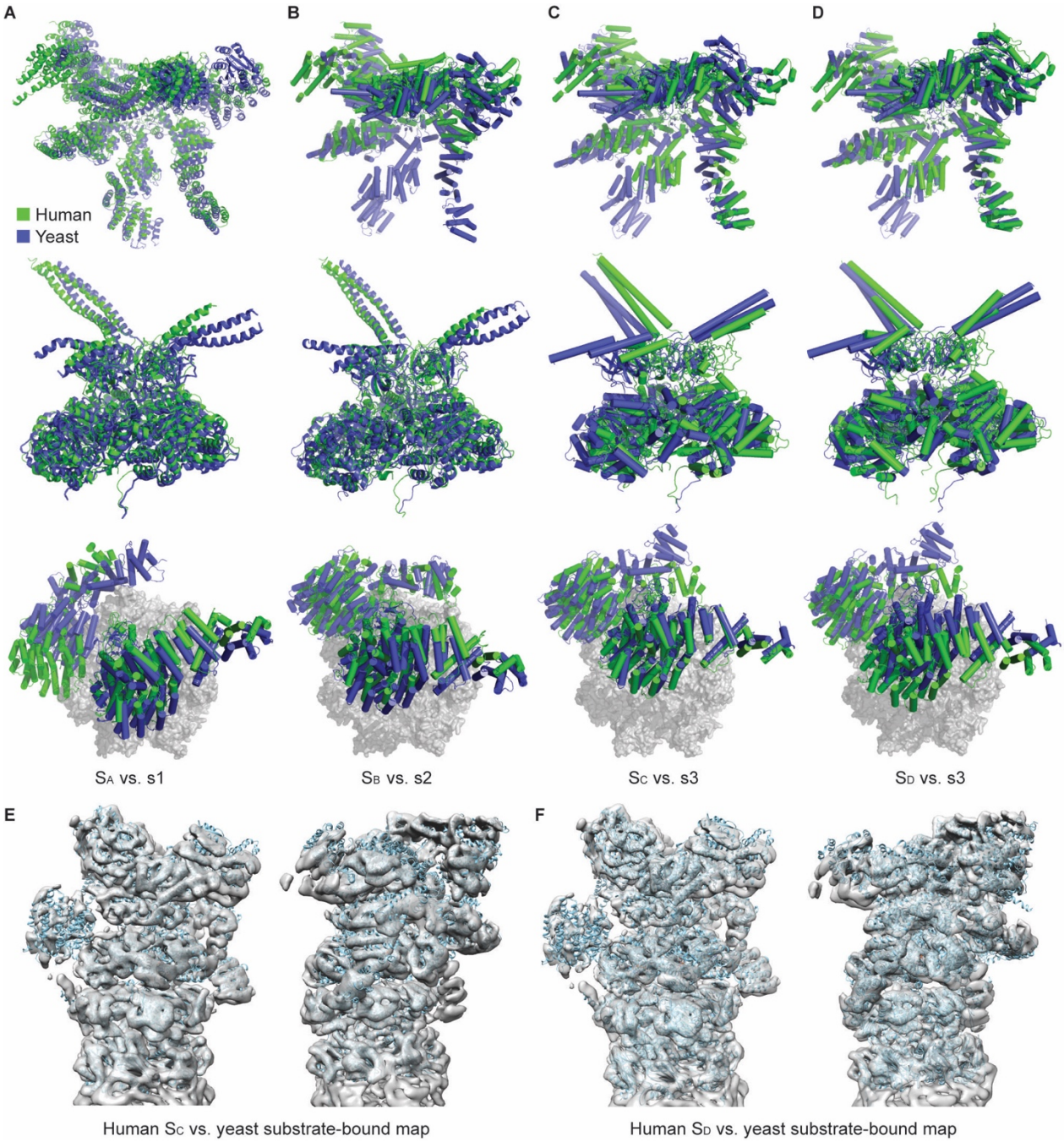


Fig. S8. Comparison of the human 26S structures in four states with those from the yeast 26S. (**A**)-(**D**) The structures of human holoenzyme are compared to the yeast one in the S_A vs. s1 (panel **A**), S_B vs. s2 (panel **B**), S_C vs. s3 (panel **C**), and S_D vs. s3 (panel **D**). Top, the lid comparison. Middle, the ATPase comparison. Bottom, the Rpn1 and Rpn2 comparison and CP is rendered grey surface for clarify. (**E**) and (**F**) The human S_C (panel **E**) and S_D (panel **F**) structures are superimposed with the cryo-EM density of the yeast substrate-bound proteasome.

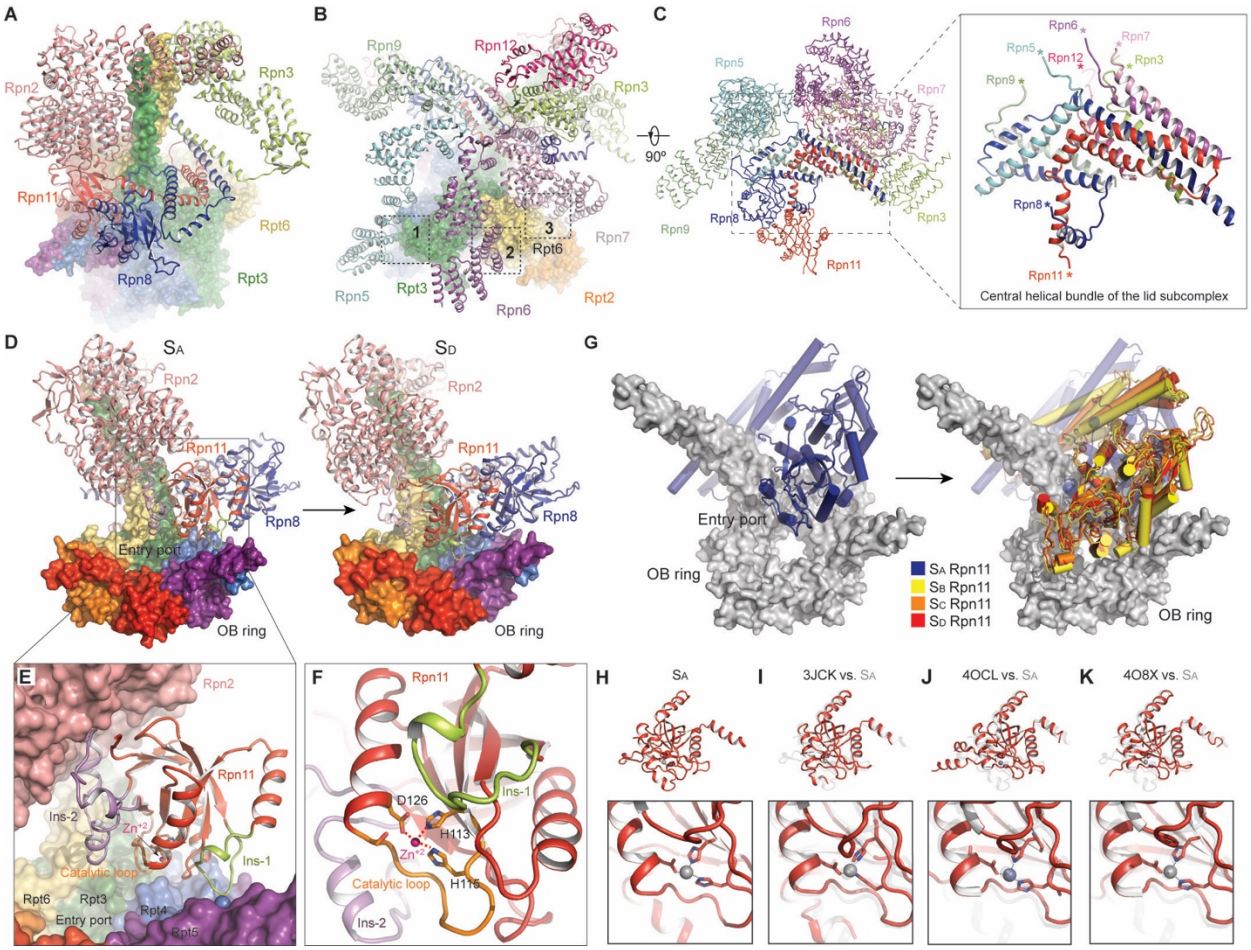


Fig. S9. Key structural features of the RP at the S_A state. (**A**) The lid-base interface surrounding the Rpt3-Rpt6 heterodimer involves Rpn2, Rpn3, Rpn8 and Rpn11, as well as Rpt4. (**B**) The lid-base interface lateral to the ATPase ring is composed of three local inter-subunit interfaces of Rpn5-Rpt3, Rpn6-Rpt6 and Rpn7-Rpt6, which are highlighted by dashed boxes 1, 2 and 3, respectively. (**C**) The atomic model of the lid subcomplex, highlighting its central helical bundle shown in cartoon representation, whereas the rest of the lid structure is shown in a ribbon representation. The right inset zooms into the central helical bundle. The asterisk symbols label the N-terminal ends of the lid subunits. (**D**) Overview showing the inter-subunit relationship between Rpn2-Rpn11-Rpn8 and the CC-OB domains of the ATPases in the S_A (left) and S_D (right) states. Black box indicates the location to which the close-up view in panel (E) is zoomed. (**E**) Close-up view of Rpn11 in cartoon representation. The adjacent Rpt subunits and Rpn2 are shown in surface representation. (**F**) Close-up view of the Zn^{+2} active site of Rpn11. The viewing angle is rotated clockwise ~90° relative to that of panel (E). (**G**) Top view of Rpn11 (blue) sitting on the OB ring (grey) in S_A (left), which moves about 15 Å in S_B, S_C and S_D (right). (**H**)-(**K**) Comparison of the Rpn11 structure in the S_A state (panel **H**) with those from cryo-EM structure of the yeast lid subcomplex (28), PDB ID 3JCK (panel **I**), and two crystal structures, PDB ID 4OCL (29) (panel **J**) and 4O8X (19) (panel **K**). Lower insets, the close-up views of the Zinc active site.

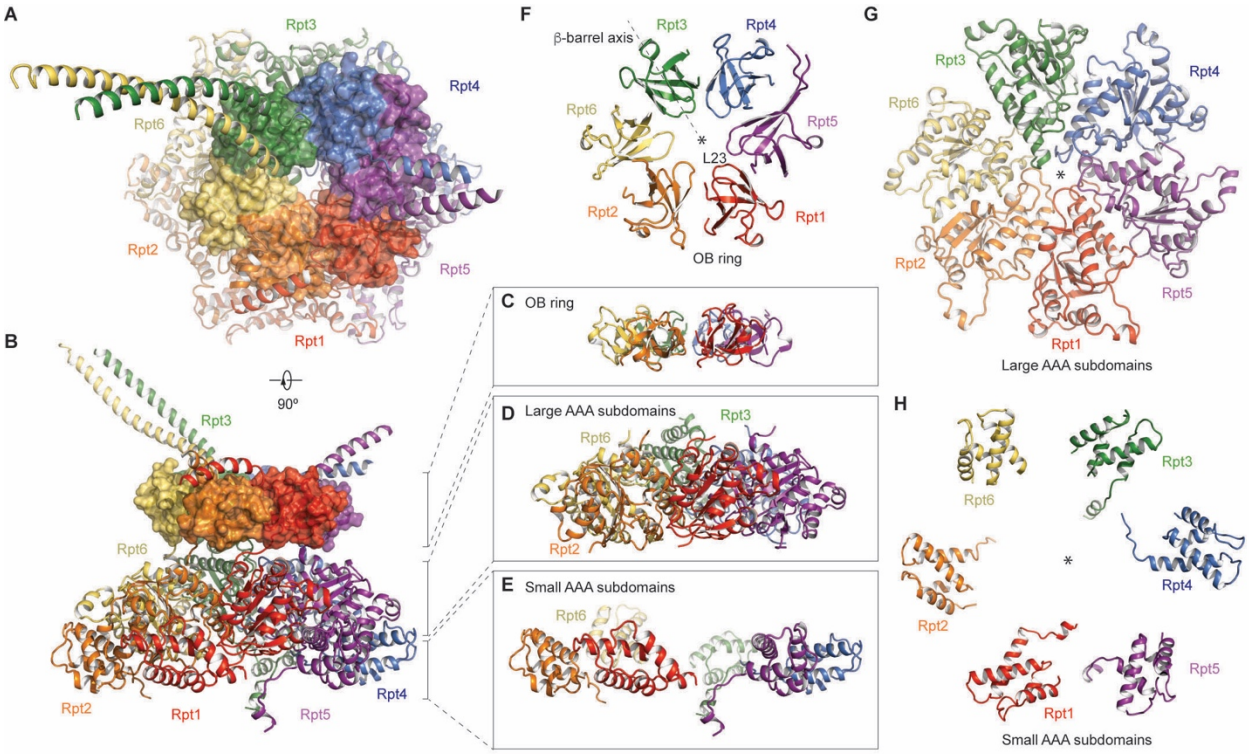
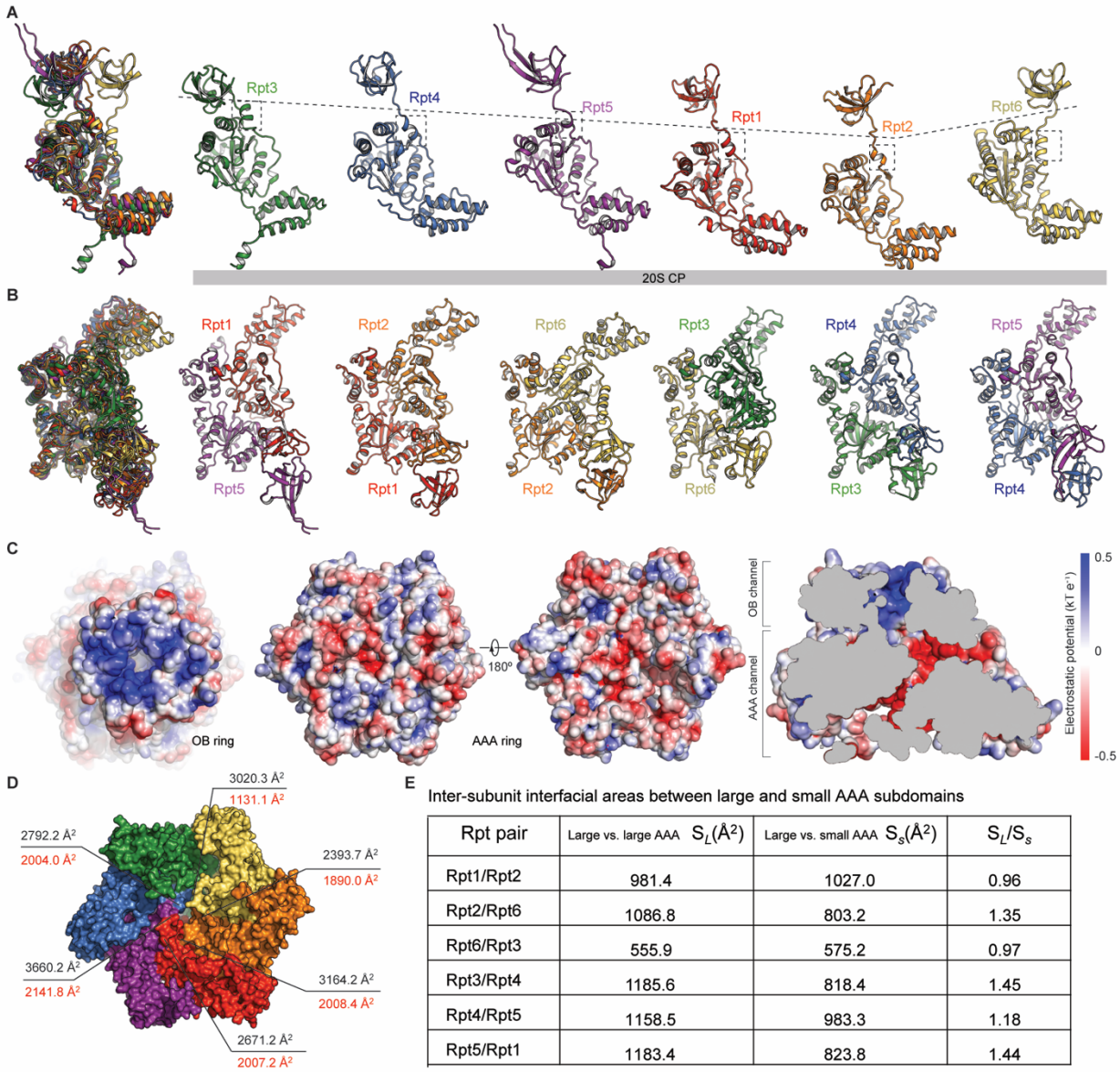


Fig. S10. The atomic structure of the ATPase heterohexamer in the S_A state. (**A**) The overall structure of the ATPase viewed from the perspective of the substrate entry port. The OB ring is in the foreground and is shown with transparent surface superimposed on a cartoon representation of the AAA domains. (**B**) Overall structure of the ATPase ring viewed from a perspective rotated 90 degrees from that in panel (**C**). (**C**)-(**E**) The component structures of the OB ring (panel **C**), large AAA subdomains (panel **D**) and small AAA subdomains (panel **E**) viewed from the same perspective as that in panel (B**)**. (**F**)-(**H**) The component structures of the OB ring (panel **F**), large AAA subdomains (panel **G**) and small AAA subdomains (panel **H**) viewed from the same perspective as that in panel (A).



E

Inter-subunit interfacial areas between large and small AAA subdomains

Rpt pair	Large vs. large AAA S_L(Å²)	Large vs. small AAA S_s(Å²)	S_L/S_s
Rpt1/Rpt2	981.4	1027.0	0.96
Rpt2/Rpt6	1086.8	803.2	1.35
Rpt6/Rpt3	555.9	575.2	0.97
Rpt3/Rpt4	1185.6	818.4	1.45
Rpt4/Rpt5	1158.5	983.3	1.18
Rpt5/Rpt1	1183.4	823.8	1.44

Fig. S11. Asymmetric structure of the ATPase heterohexamer in the S_A state. (**A**) The OB-AAA domain structures of Rpt subunits in a cartoon representation from a perspective showing an L-shape. The leftmost panel shows the six Rpt structures superimposed, with each structure separately shown on the right. (**B**) The dimeric AAA domain structure of two adjacent Rpt subunits. The leftmost panel shows the six AAA dimer structures superimposed together, with each structure separately shown on the right. (**C**) The electrostatic structure of the ATPase of OB ring (left), AAA ring (middle), and the ATPase channel (right). (**D**) The inter-subunit interfacial areas between two adjacent Rpt subunit. The black numbers show the total inter-subunit interfacial areas, whereas the red numbers show the interfacial areas between the adjacent AAA domains. (**E**) A table shows the inter-subunit interfacial areas between large and small AAA subdomains from two adjacent Rpt subunits. Note that the two adjacent small AAA subdomains do not contact each other, as shown in Fig. S10H.

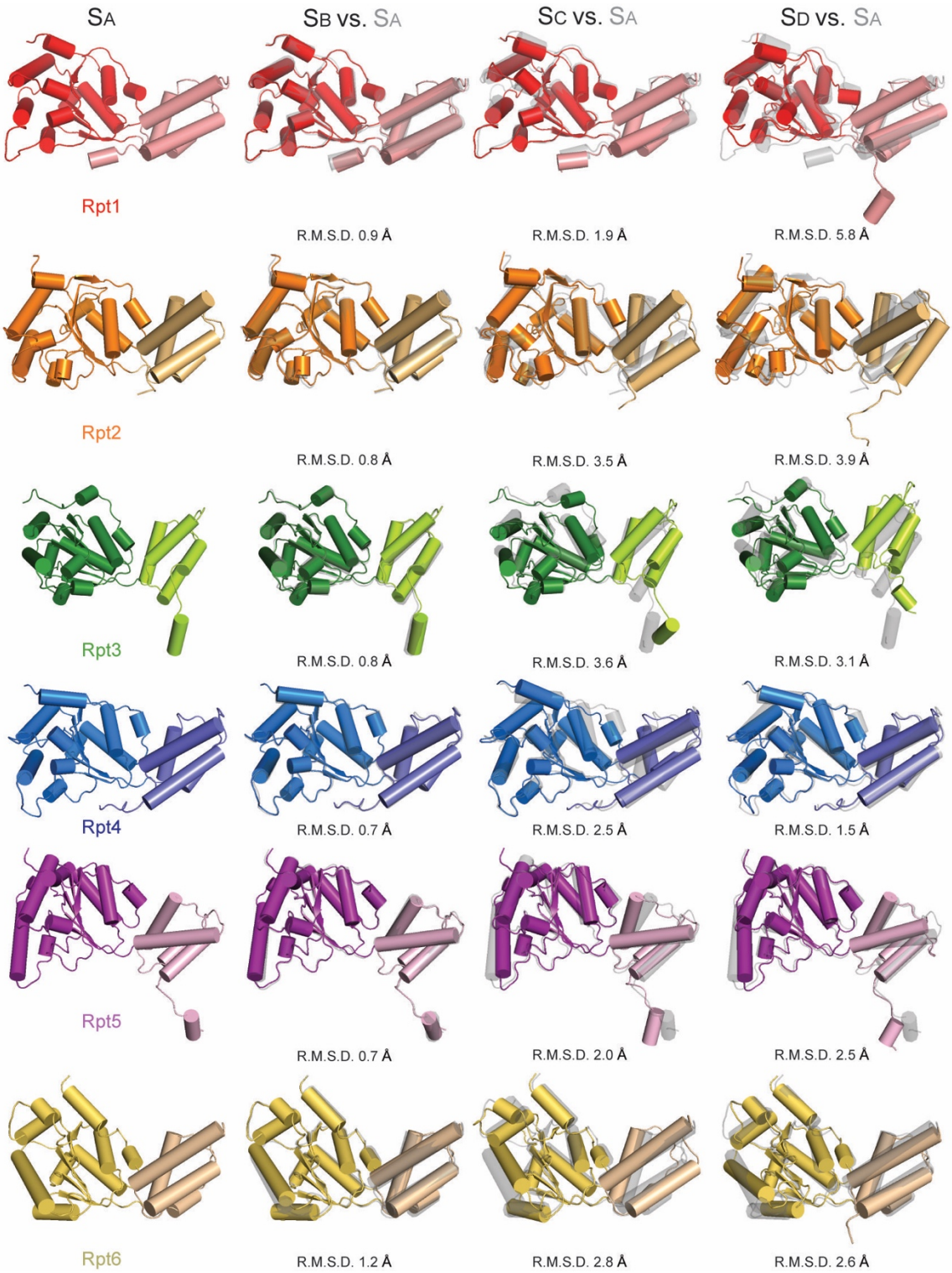


Fig. S12. Subunit conformational changes in the ATPase heterohexamer. Each Rpt subunit is compared for its conformations in the S_A, S_B, S_C and S_D. The root-mean-square deviation (R.M.S.D) is shown for each of alternative states (S_B, S_C and S_D) versus S_A. Detailed statistics of the R.M.S.D. between any two states are shown in Table S3.

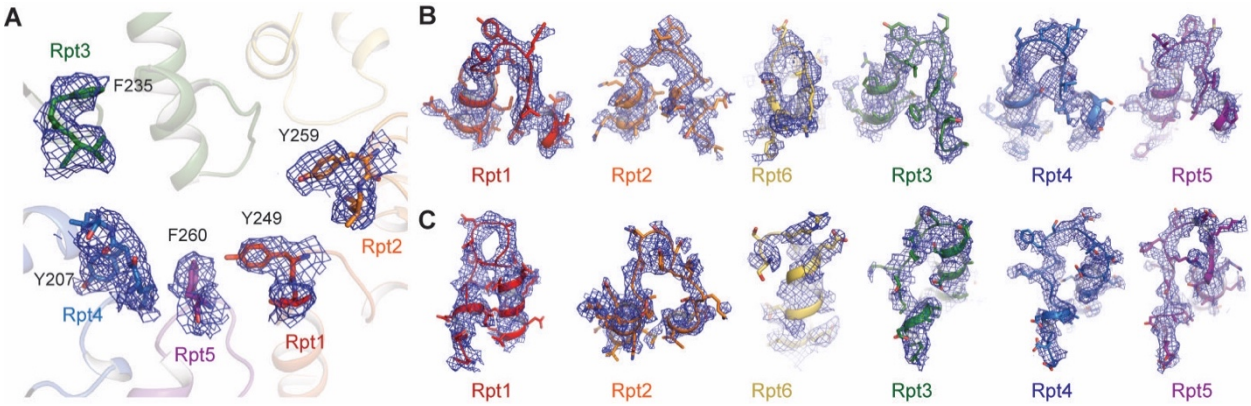


Fig. S13. ATPase pore loops in the S_A state. (**A**) The aromatic side chain density in the pore-1 loop lining in the interior of the AAA channel. (**B**) and (**C**) The densities of pore-1 loops (panel **B**) and pore-2 loops (panel **C**) from Rpt subunits are superimposed with the atomic model. Side chains are shown in stick representation, whereas main chains are shown in cartoon representation.

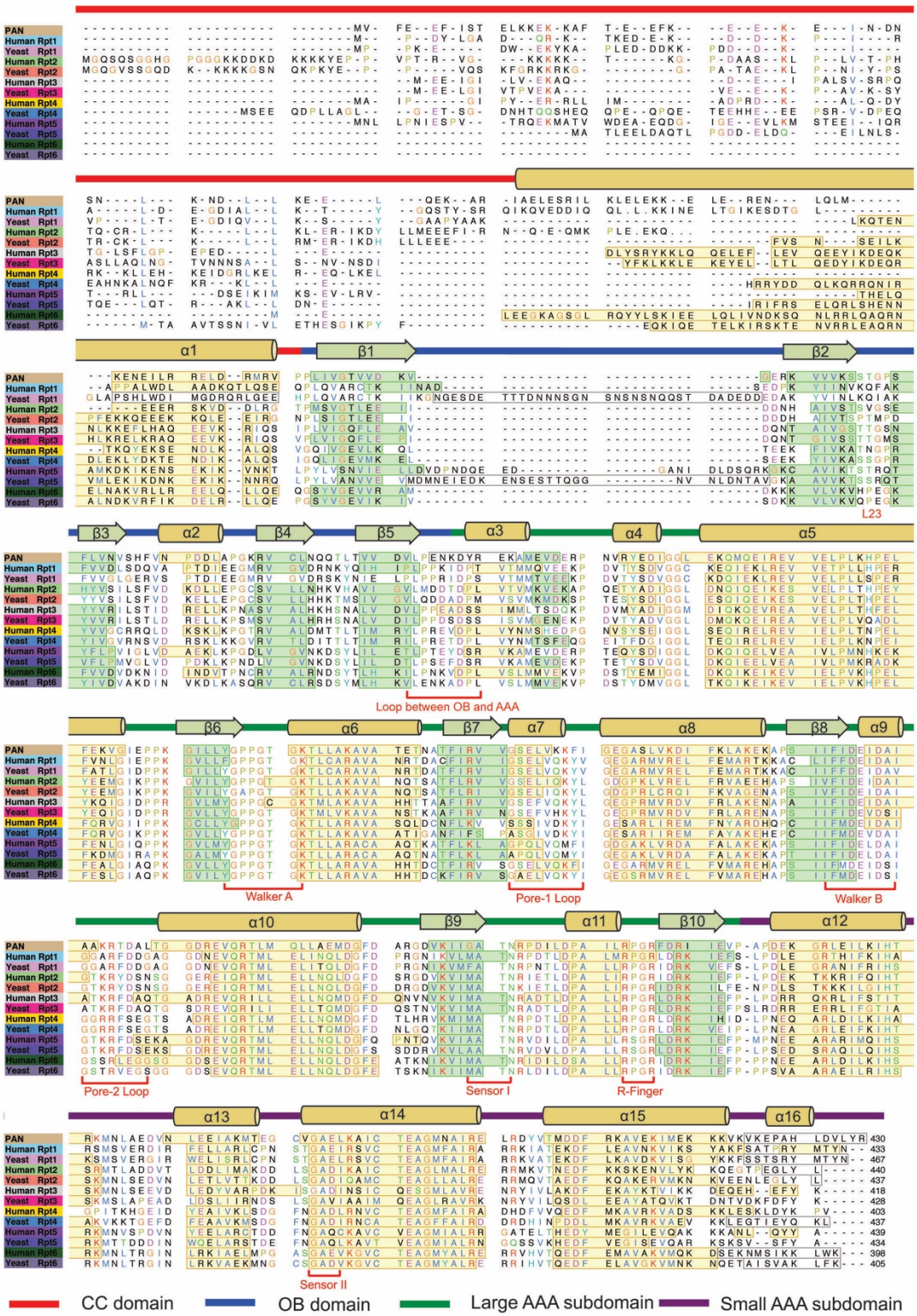


Fig. S14. Structure-based sequence alignment. Sequences from the yeast and human proteasomal ATPases are aligned with each other and with that of the archaeal PAN complex, based on their structures. The yeast structure used in this alignment is based on the pseudo-atomic model previously built from a 7-8 Å cryo-EM map (30) (PDB ID: 4CR2).

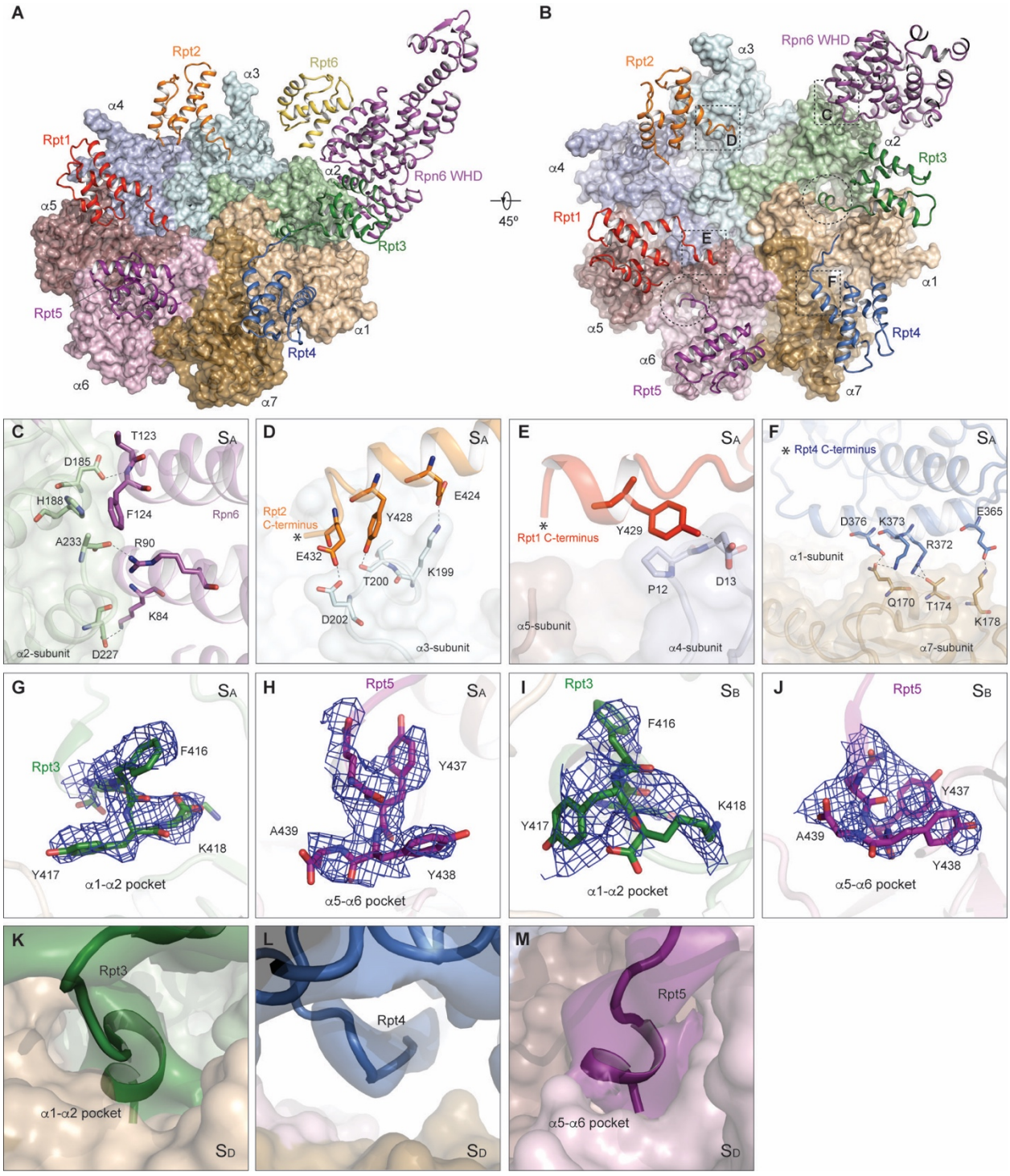


Fig. S15. RP-CP interface in different states. (**A**) Overview of the RP-CP interface of the S_A state in which the α ring is shown in surface representation and the small AAA domains of Rpt1-6 subunits and Rpn6 are shown in cartoon representation. From this perspective, one can see that Rpt6 is the only ATPase subunit that does not directly interact with the CP. (**B**) Rotated overview of the RP-CP interface of the S_A state in which the C-terminal tails of Rpt3 and Rpt5 are shown to insert into the α pockets. Rpt6 is not shown in this panel for clarity. The dashed circle marks the HbYX motifs of two Rpt subunits that insert into the α pockets. The dashed

boxes indicate the locations of the C-terminal tails of Rpt1, Rpt2 and Rpt4, and the Rpt6-α2 interface, which are shown as close-up views in subsequent panels. (**C**) Close-up view of the interaction between the Rpn6 N-terminus and α2. (**D**) Close-up view of the interactions between the Rpt2 C-terminal and α3 in the S_A state. (**E**) Close-up view of the interactions between the Rpt1 C-terminal helix and α4 in the S_A state. (**F**) Close-up view of the interactions between the Rpt4 C-terminal helix and α7 in the S_A state. The asterisk in each panel labels the extreme C-terminus of the Rpt subunit. Hydrogen bonds are illustrated as gray dashed lines. (**G**) and (**H**) Close-up view of the density of the extreme C-terminal HbYX motif of Rpt3 (panel **G**) and Rpt5 (panel **H**) in the S_A state. The atomic model is shown in stick representation for the C-terminal segment but in transparent cartoon form for the rest. (**I**) and (**J**) Close-up view of the density of the extreme C-terminal HbYX motif of Rpt3 (panel **I**) and Rpt5 (panel **J**) in the S_B state. (**K**)-(**M**) Close-up view of the density map of the extreme C-terminal segment of Rpt3 (panel **K**), Rpt4 (panel **L**) and Rpt5 (panel **M**) in the S_D state. Because the density of its HbYX motif is not observed, the atomic model of the Rpt2 C-terminus is terminates at residue Gly433.

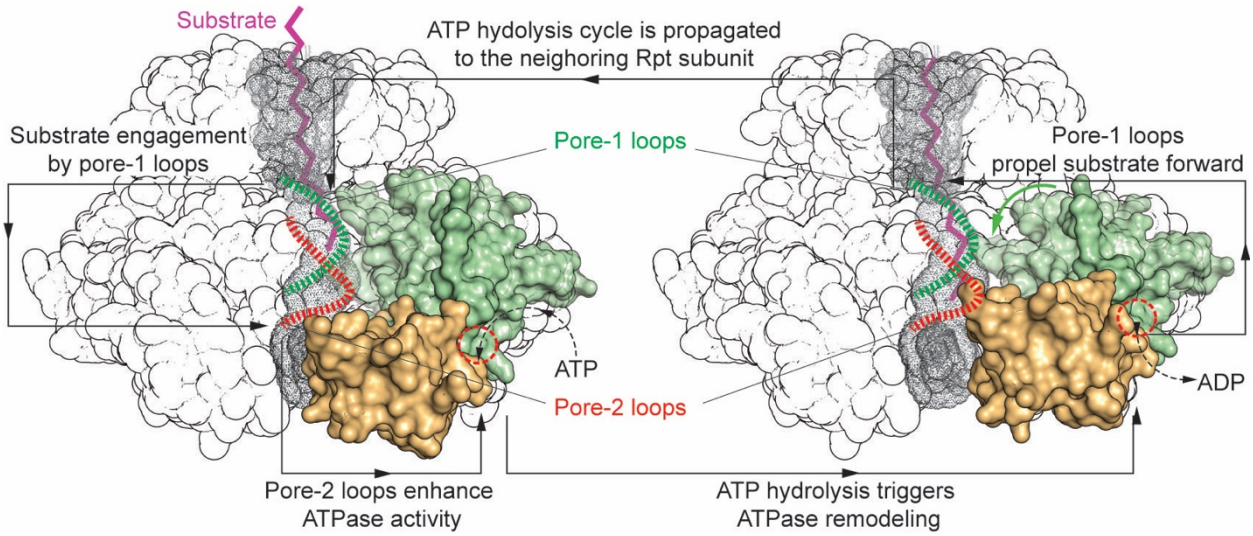


Fig. S16. A hypothetical working model for processive substrate translocation by the ATPase heterohexamer. The ATPase channel is represented by gray dots, and substrate by a purple zigzag chain. The AAA domain of one Rpt subunit is illustrated in a surface representation. The nucleotide-binding site is illustrated as a dashed red circle. The pore-1 and pore-2 loops are illustrated as dashed green and red curves, respectively. The energy transduction cycle in the Rpt subunit propagates to the adjacent one at the end of the cycle. The scheme shown in the illustration is not quantitatively modeled, but only serves to illustrate hypothetical scenarios in the proposed mechanism. In the S_A state, an unfolded segment of a substrate could diffuse into the OB channel and be captured by the hydrophobic residues of the pore-1 loops (31). The pore-2 loops opposite the pore-1 loops translate the passive interactions with the substrate into activation or enhancement of ATP hydrolysis, possibly by controlling the interacting distance and orientation of the conserved glutamate residue in the Walker B motif (32). The energy of ATP hydrolysis would now be converted into mechanical work that drives a hinge-like movement between the small and large AAA subdomains (33). These structural changes in the Rpt proteins generate forces through the pore-1 loops that act back on the substrate and propel it through the AAA channel. Such an energy transduction cycle may propagate to the adjacent AAA domain, allowing a processive action on the substrate.

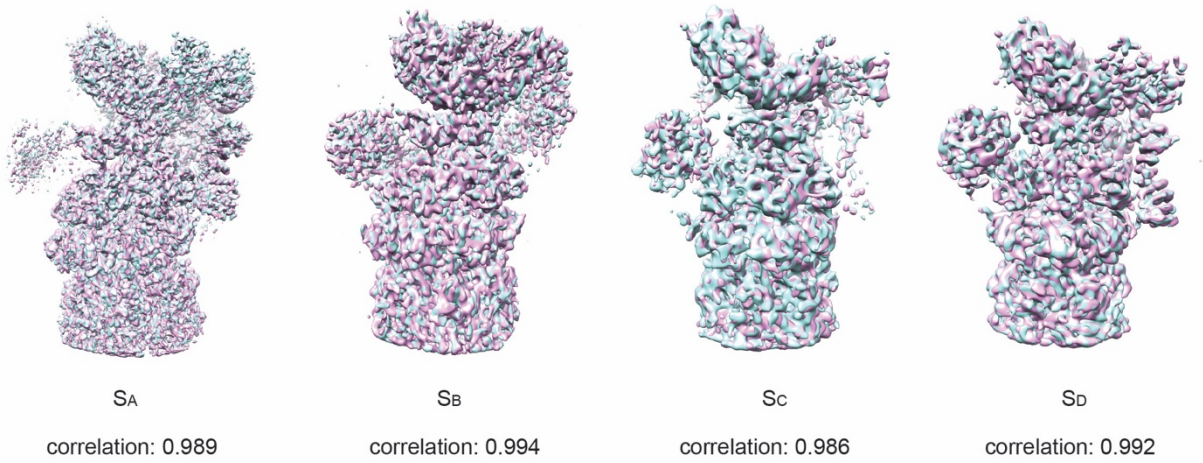


Fig. S17. Validation of RP-CP complex classification using different initial models. Auto-refinement using the S_A state map (cyan) and the S_D state map (pink) as an initial model. Cross correlation value is shown for each RP-CP complex between two independent auto-refinement using S_A versus S_D map as an initial model.

Table S1. (A) Statistics of the human 26S proteasome structure determination by single-particle cryo-EM.

Electron energy(kV)					200kV
Electron dose(e-/ Å2)					30
Pixel size corresponding to the physical detector sensor (Å)					1.72
Pixel size in the super-resolution counting mode of K2 Summit (Å)					0.86
Defocus range(μm)					-3.0 ~ -1.0
Number of micrographs					10,367
Number of particles for 26S proteasome					237,083
Resolution of 26S proteasome complex with C2 symmetry (Å) (applied CP mask)					3.6
	S_A state Doubly capped with C2 applied	S_A state RP-CP	S_B state RP-CP	S_C state RP-CP	S_D state RP-CP
Particle number	85,420	139,236	18,443	10,622	14,382
Resolution (Å)	3.8	4.4	6.8	8.0	8.0
B-factor ($Å^2$)	-80	-70	-150	-200	-150
Pseudo-crystallographic refinement of atomic models					
Cell dimension a,b,c (Å)	215,215,215	309.6,309.6,309.6			
Cell angle α, β, γ (Å)	90,90,90	90,90,90			
Space group	P1	P1			
Resolution range	215-3.6	309.6-4.0			
Number of atoms	39,636	72,467			
R_{work} factor	0.213	0.254			
Geometric Parameters (RMSD)					
Bond length (Å)	0.0043	0.0042			
Bond angles (°)	0.873	0.865			
Ramachandran plot statistics					
Favored(%)	91.3	92.3	88.3	88.0	89.6
Allowed (%)	7.9	6.9	11.2	11.0	9.9
Outliers(%)	0.8	0.8	0.5	1.0	0.5
MolProbity validation					
Rotamer outliers (%)	0.1	0.1			
Clashscore	20.4	15.8			
MolProbity score	2.32	2.18			
C-beta outliers	0	0			

(B) Atomic model building.

RP Subunits	Homology structure to generate the initial coordinates	PDB ID of the homology structure	Residue range of the built atomic model	References
Rpn2	Yeast Rpn2	4ADY	7-270, 320-814, 836-844, 880-917	(34)
Rpn3	Human COP9, PCI domain	4D10	18-497	(17)
Rpn5	Human COP9, PCI domain	4D10	1-456	(17)
Rpn6	Yeast Rpn6 PCI domain	3TXN	43-422	(35)
Rpn7	COP9, PCI domain	4D10	12-389	(17)
Rpn8	Yeast Rpn8, MPN domain	4O8X	5-290	(19)
Rpn9	Yeast Rpn9, PCI domain	2MR3	4-376	(20)
Rpn10	Yeast Rpn10, VWA domain	2X5N	1-191	(21)
310Rpn11	Yeast Rpn11, MPN domain	4O8X	24-310	(19)
Rpn12	Yeast Rpn12, PCI domain	4B0Z	1-257	(22)
Dss1/Sem1	N/A	N/A	1-9, 40-70	
Rpt1	PAN	3H4M, 3H43	73-433	(18)
Rpt2	PAN	3H4M, 3H43	93-433	(18)
Rpt3	PAN	3H4M, 3H43	39-418	(18)
Rpt4	PAN	3H4M, 3H43	37-390	(18)
Rpt5	PAN	3H4M, 3H43	63-439	(18)
Rpt6	PAN	3H4M, 3H43	11-394	(18)

Table S2. Inter-subunit interfacial areas between the lid and the base, and between RP and CP in the S_A states.

Interface Area ($Å^2$)		Base							
		Rpt1	Rpt2	Rpt3	Rpt4	Rpt5	Rpt6	Rpn2	Rpn10
Lid	Rpn3						513.8	532.5	
	Rpn5			265.4	185.6				
	Rpn6			76.9			483.9		
	Rpn7		82.1	44.9			937.8		
	Rpn8			191.6	72.5	65.4		29.6	608.2
	Rpn9								554.9
	Rpn11			445.4	376.9		172.8	802.1	71.3
	Rpn12							569.8	
CP	α1-subunit			329.0	163.1		31.1		
	α2-subunit			336.5			0.6		
	α3-subunit	7.0	395.4						
	α4-subunit	215.9	111.4						
	α5-subunit	468.5				344.8			
	α6-subunit	73.0				630.3			
	α7-subunit	60.6			263.6	163.0			

Interface Area ($Å^2$)	Lid-complex	
	Rpn5	Rpn6
α1-subunit	53.2	
α2-subunit		617.9

Table S3. Root-mean-square-deviation (Å) between subunit or domain structures in different states.

	S_B vs. S_A	S_C vs. S_A	S_D vs. S_A	S_C vs. S_B	S_D vs. S_B	S_D vs. S_C
Rpt1	1.0	2.9	11.0	2.6	10.6	9.3
Rpt2	1.4	4.3	11.3	3.9	10.5	9.1
Rpt3	2.7	6.2	9.9	5.0	9.5	7.0
Rpt4	1.7	6.2	9.5	5.1	8.4	5.7
Rpt5	1.4	4.8	9.3	4.0	8.5	6.4
Rpt6	3.4	4.9	10.1	3.4	10.1	8.2
Rpn1	50.4	42.6	48.0	10.1	9.1	6.5
Rpn2	16.6	14.7	20.7	4.3	12.4	11.0
Rpn3	24.4	17.4	14.1	9.3	18.4	10.1
Rpn5	36.7	24.2	25.6	10.5	15.7	7.2
Rpn6	31.2	21.1	18.7	13.2	20.5	7.7
Rpn7	32.2	21.6	19.0	15.3	23.7	9.1
Rpn8	30.4	27.0	26.7	5.6	11.8	8.1
Rpn9	47.7	39.25	40.0	10.5	14.7	7.9
Rpn10	45.7	37.7	40.0	8.9	12.2	9.2
Rpn11	22.1	20.0	23.5	6.3	12.3	8.2
Rpn12	34.0	23.2	18.1	13.6	21.9	10.1
Sem1	35.2	21.0	16.6	14.2	23.4	9.8
OB domain of Rpt subunit						
Rpt1	1.2	2.6	10.7	1.7	9.8	8.6
Rpt2	1.6	3.7	12.2	2.4	10.7	8.9
Rpt3	2.2	5.5	10.7	3.7	9.8	7.2
Rpt4	2.2	6.2	11.3	4.3	9.9	6.7
Rpt5	1.9	5.7	10.3	4.1	9.0	7.3
Rpt6	2.0	3.5	10.1	2.2	9.0	8.0
OB ring	1.9	4.7	10.9	3.2	9.7	7.8
AAA domain of Rpt subunit						
Rpt1	1.0	3.0	11.2	2.8	10.9	9.5
Rpt2	1.3	4.4	11.0	4.2	10.4	9.1
Rpt3	2.1	6.5	9.8	5.6	9.0	6.3
Rpt4	1.5	6.3	8.9	5.3	7.8	5.2
Rpt5	1.0	4.5	8.5	3.9	7.9	5.2
Rpt6	2.6	5.1	10.4	3.6	9.8	7.8

AAA ring	1.8	5.1	10.0	4.4	9.4	7.4

References

1. Wang X, *et al.* (2007) Mass spectrometric characterization of the affinity-purified human 26S proteasome complex. *Biochemistry* 46(11):3553-3565.
2. Suloway C, *et al.* (2005) Automated molecular microscopy: the new Leginon system. *J Struct Biol* 151(1):41-60.
3. Li X, *et al.* (2013) Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. *Nat Methods* 10(6):584-590.
4. Mindell JA & Grigorieff N (2003) Accurate determination of local defocus and specimen tilt in electron microscopy. *J Struct Biol* 142(3):334-347.
5. Wu J, *et al.* (2016) Unsupervised single-particle deep classification via statistical manifold learning. *arXiv*:1604.04539 [physics.data-an].
6. Scheres SH (2012) RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J Struct Biol* 180(3):519-530.
7. Zhu Y, Ouyang Q, & Mao Y (2016) A deep learning approach to single-particle recognition in cryo-electron microscopy. *arXiv*:1605.05543 [physics.data-an].
8. Tang G, *et al.* (2007) EMAN2: an extensible image processing suite for electron microscopy. *J Struct Biol* 157(1):38-46.
9. Scheres SH & Chen S (2012) Prevention of overfitting in cryo-EM structure determination. *Nat Methods* 9(9):853-854.
10. Chen S, *et al.* (2013) High-resolution noise substitution to measure overfitting and validate resolution in 3D structure determination by single particle electron cryomicroscopy. *Ultramicroscopy* 135:24-35.
11. Pettersen EF, *et al.* (2004) UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* 25(13):1605-1612.
12. Shaikh TR, *et al.* (2008) SPIDER image processing for single-particle reconstruction of biological macromolecules from electron micrographs. *Nat Protoc* 3(12):1941-1974.
13. Kucukelbir A, Sigworth FJ, & Tagare HD (2014) Quantifying the local resolution of cryo-EM density maps. *Nat Methods* 11(1):63-65.
14. Heymann JB & Belnap DM (2007) Bsoft: image processing and molecular modeling for electron microscopy. *J Struct Biol* 157(1):3-18.
15. Marti-Renom MA, *et al.* (2000) Comparative protein structure modeling of genes and genomes. *Annu Rev Biophys Biomol Struct* 29:291-325.
16. Emsley P & Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60(Pt 12 Pt 1):2126-2132.
17. Lingaraju GM, *et al.* (2014) Crystal structure of the human COP9 signalosome. *Nature* 512(7513):161-165.
18. Zhang F, *et al.* (2009) Structural insights into the regulatory particle of the proteasome from Methanocaldococcus jannaschii. *Mol Cell* 34(4):473-484.
19. Worden EJ, Padovani C, & Martin A (2014) Structure of the Rpn11-Rpn8 dimer reveals mechanisms of substrate deubiquitination during proteasomal degradation. *Nat Struct Mol Biol* 21(3):220-227.
20. Hu Y, Wu Y, Li Q, Zhang W, & Jin C (2015) Solution structure of yeast Rpn9: insights

into proteasome lid assembly. *J Biol Chem* 290(11):6878-6889.
21. Riedinger C*, et al.* (2010) Structure of Rpn10 and its interactions with polyubiquitin chains and the proteasome subunit Rpn12. *J Biol Chem* 285(44):33992-34003.
22. Boehringer J*, et al.* (2012) Structural and functional characterization of Rpn12 identifies residues required for Rpn10 proteasome incorporation. *Biochem J* 448(1):55-65.
23. Harshbarger W, Miller C, Diedrich C, & Sacchettini J (2015) Crystal structure of the human 20S proteasome in complex with carfilzomib. *Structure* 23(2):418-424.
24. Adams PD*, et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66(Pt 2):213-221.
25. Krissinel E & Henrick K (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol* 372(3):774-797.
26. Smart OS, Neduvelil JG, Wang X, Wallace BA, & Sansom MS (1996) HOLE: a program for the analysis of the pore dimensions of ion channel structural models. *J Mol Graph* 14(6):354-360, 376.
27. Schrödinger L (2015) The PyMOL Molecular Graphics System, Version 1.8.
28. Dambacher CM, Worden EJ, Herzik MA, Martin A, & Lander GC (2016) Atomic structure of the 26S proteasome lid reveals the mechanism of deubiquitinase inhibition. *Elife* 5.
29. Pathare GR*, et al.* (2014) Crystal structure of the proteasomal deubiquitylation module Rpn8-Rpn11. *Proceedings of the National Academy of Sciences of the United States of America* 111(8):2984-2989.
30. Unverdorben P*, et al.* (2014) Deep classification of a large cryo-EM dataset defines the conformational landscape of the 26S proteasome. *Proceedings of the National Academy of Sciences of the United States of America* 111(15):5544-5549.
31. Prakash S, Tian L, Ratliff KS, Lehotzky RE, & Matouschek A (2004) An unstructured initiation site is required for efficient proteasome-mediated degradation. *Nat Struct Mol Biol* 11(9):830-837.
32. Zhang X & Wigley DB (2008) The 'glutamate switch' provides a link between ATPase activity and ligand binding in AAA+ proteins. *Nat Struct Mol Biol* 15(11):1223-1227.
33. Sledz P*, et al.* (2013) Structure of the 26S proteasome with ATP-gammaS bound provides insights into the mechanism of nucleotide-dependent substrate translocation. *Proceedings of the National Academy of Sciences of the United States of America* 110(18):7264-7269.
34. He J*, et al.* (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(3):513-521.
35. Pathare GR*, et al.* (2012) The proteasomal subunit Rpn6 is a molecular clamp holding the core and regulatory subcomplexes together. *Proceedings of the National Academy of Sciences of the United States of America* 109(1):149-154.