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).

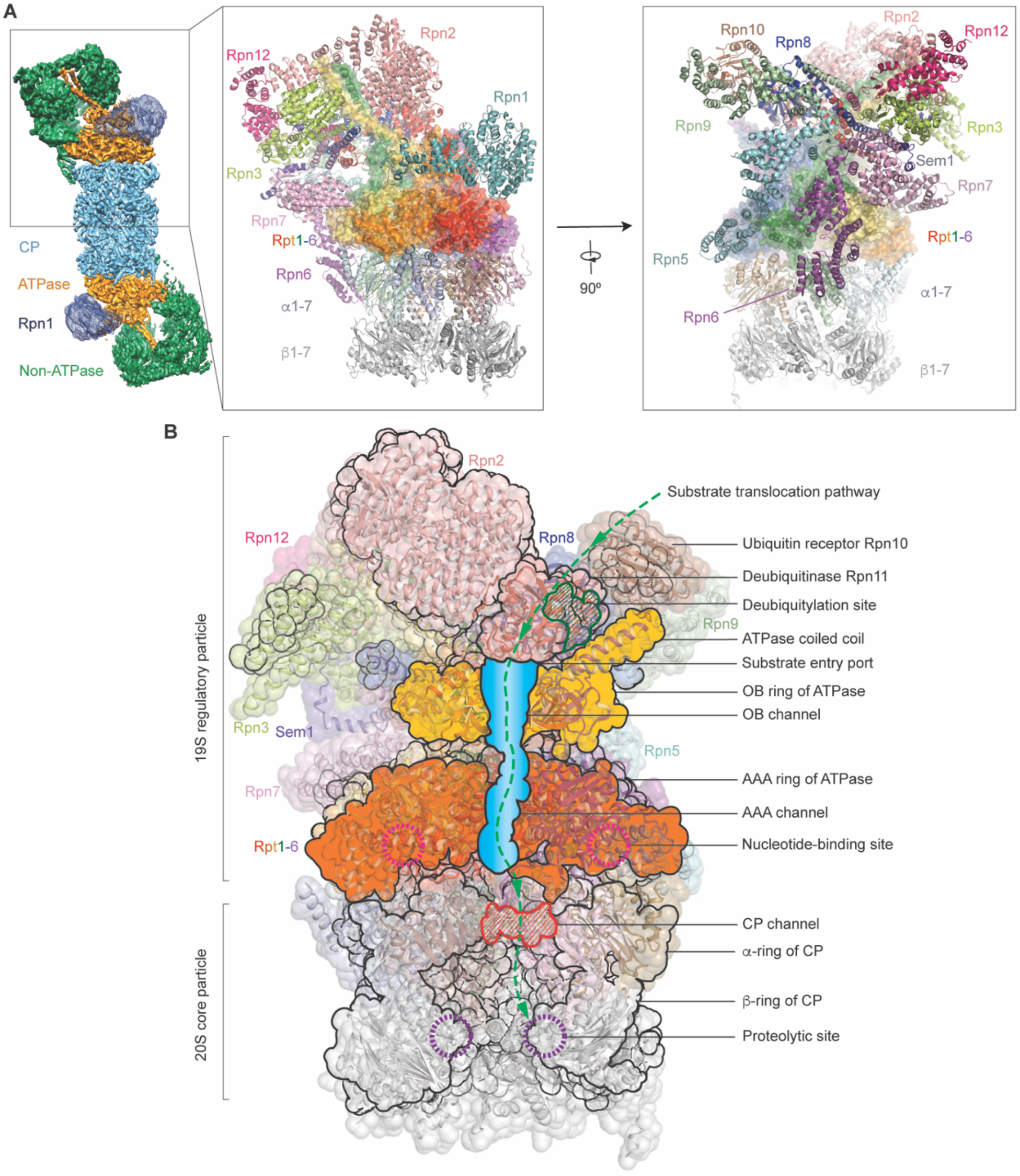


Fig. S1. Overview of the human 26S proteasome assembly. (**A**) Overview of the protein subunit organization in the proteasome holoenzyme from two orthogonal perspectives. The lid subcomplex includes Rpn3, Rpn5, Rpn6, Rpn7, Rpn8, Rpn9, Rpn11, Rpn12, Sem1; the base subunits include Rpn1, Rpn2, Rpt1, Rpt2, 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.

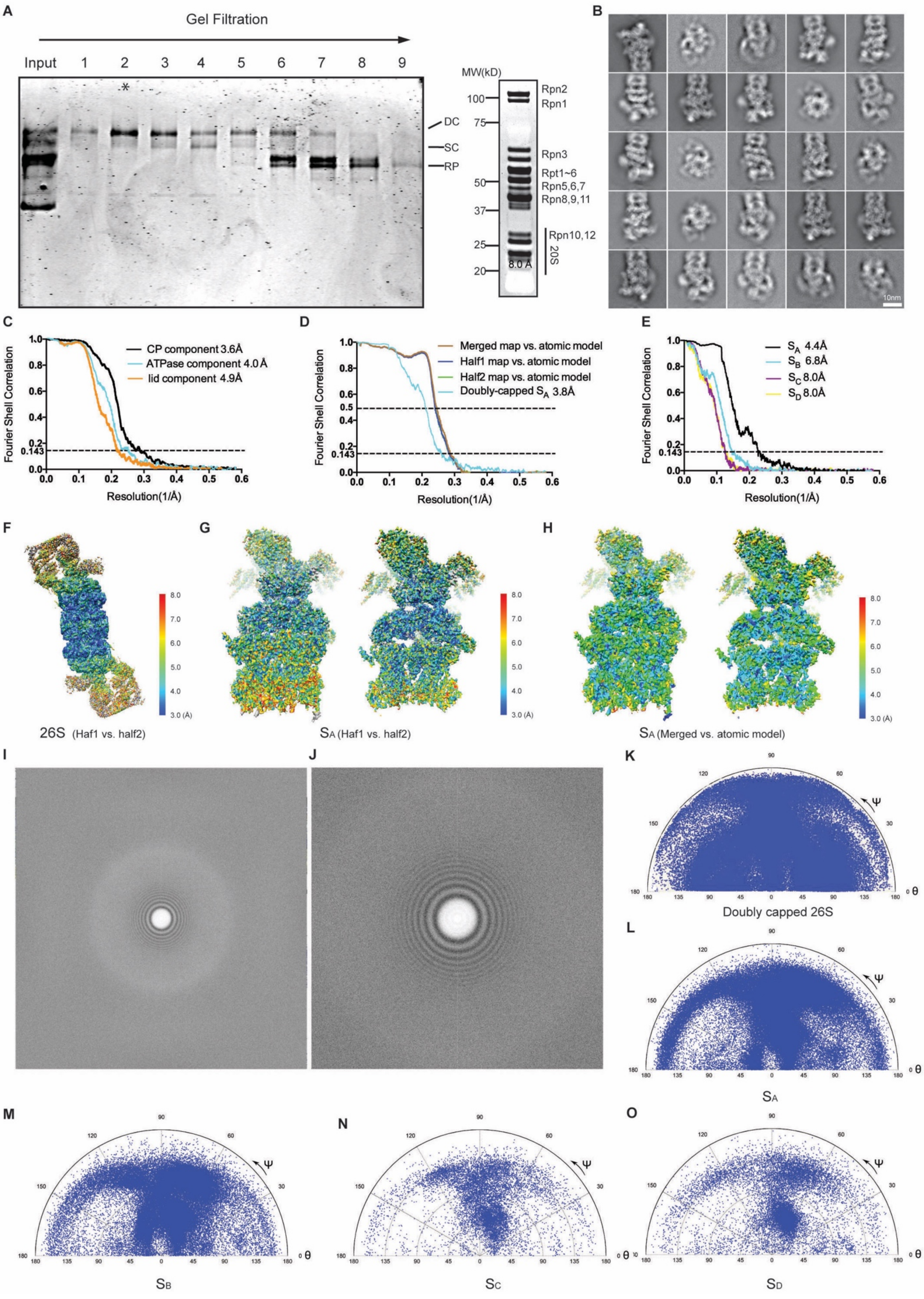

A
Gel Filtration
Input
1
2
3
4
5
6
7
8
9
*
DC
SC
RP
MW(kD)
100
75
50
37
25
20
Rpn2
Rpn1
Rpn3
Rpt1~6
Rpn5,6,7
Rpn8,9,11
Rpn10,12
20S
8.0 Å
B
10nm
C
Fourier Shell Correlation
CP component 3.6Å
ATPase component 4.0 Å
lid component 4.9Å
0.143
Resolution(1/Å)
D
Merged map vs. atomic model
Half1 map vs. atomic model
Half2 map vs. atomic model
Doubly-capped S_A 3.8Å
E
S_A 4.4Å
S_B 6.8Å
S_C 8.0Å
S_D 8.0Å
F
26S (Haf1 vs. half2)
G
SA (Haf1 vs. half2)
H
SA (Merged vs. atomic model)
8.0
7.0
6.0
5.0
4.0
3.0 (Å)
I
J
K
Doubly capped 26S
L
SA
M
SB
N
SC
O
SD
Ψ
θ
90
120
60
150
30
180
135
45
0

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.

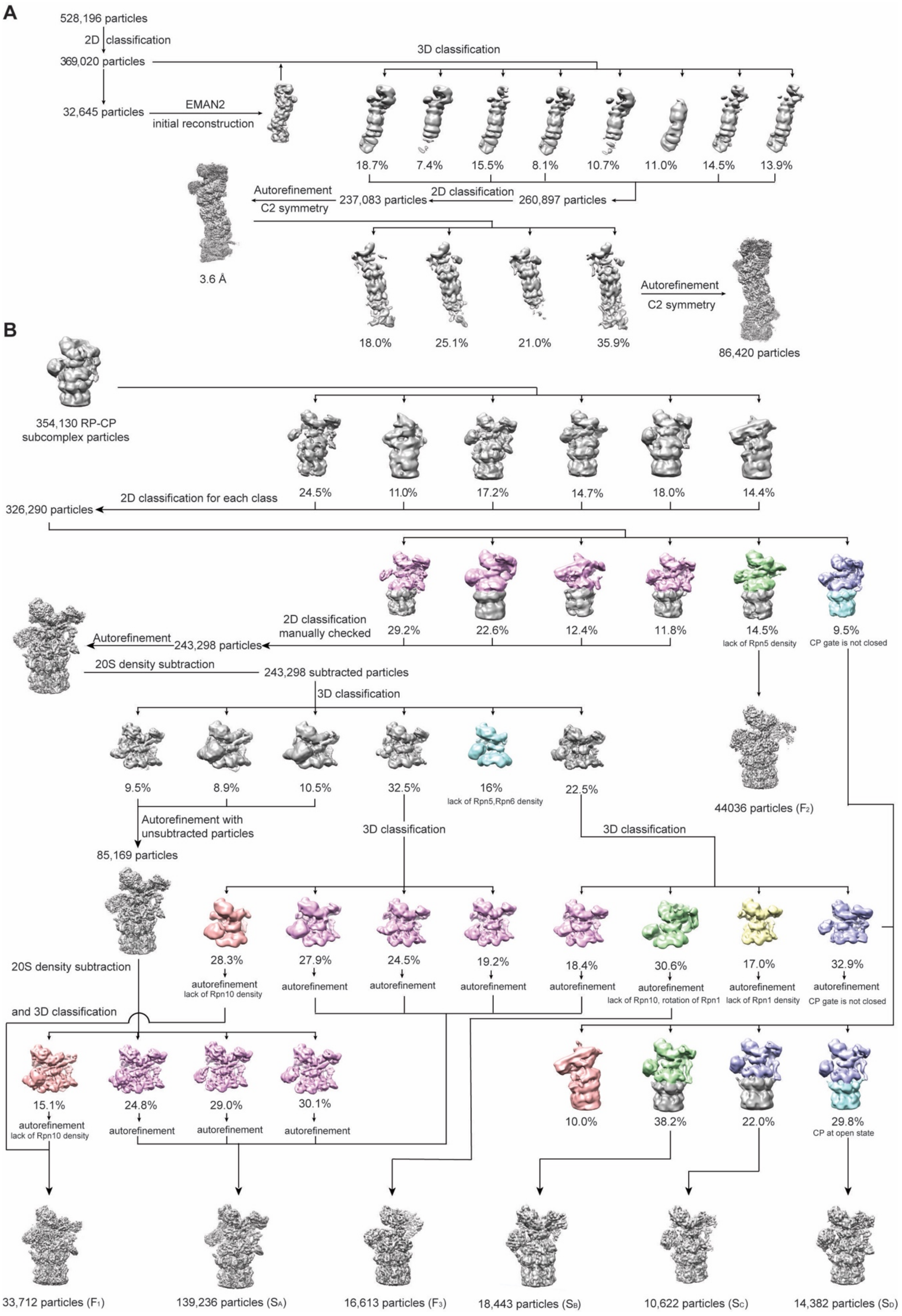

A
528,196 particles
2D classification
369,020 particles
3D classification
32,645 particles
EMAN2
initial reconstruction
18.7%
7.4%
15.5%
8.1%
10.7%
11.0%
14.5%
13.9%
Autorefinement
C2 symmetry
237,083 particles
2D classification
260,897 particles
3.6 Å
Autorefinement
C2 symmetry
18.0%
25.1%
21.0%
35.9%
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
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
20S density subtraction
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.

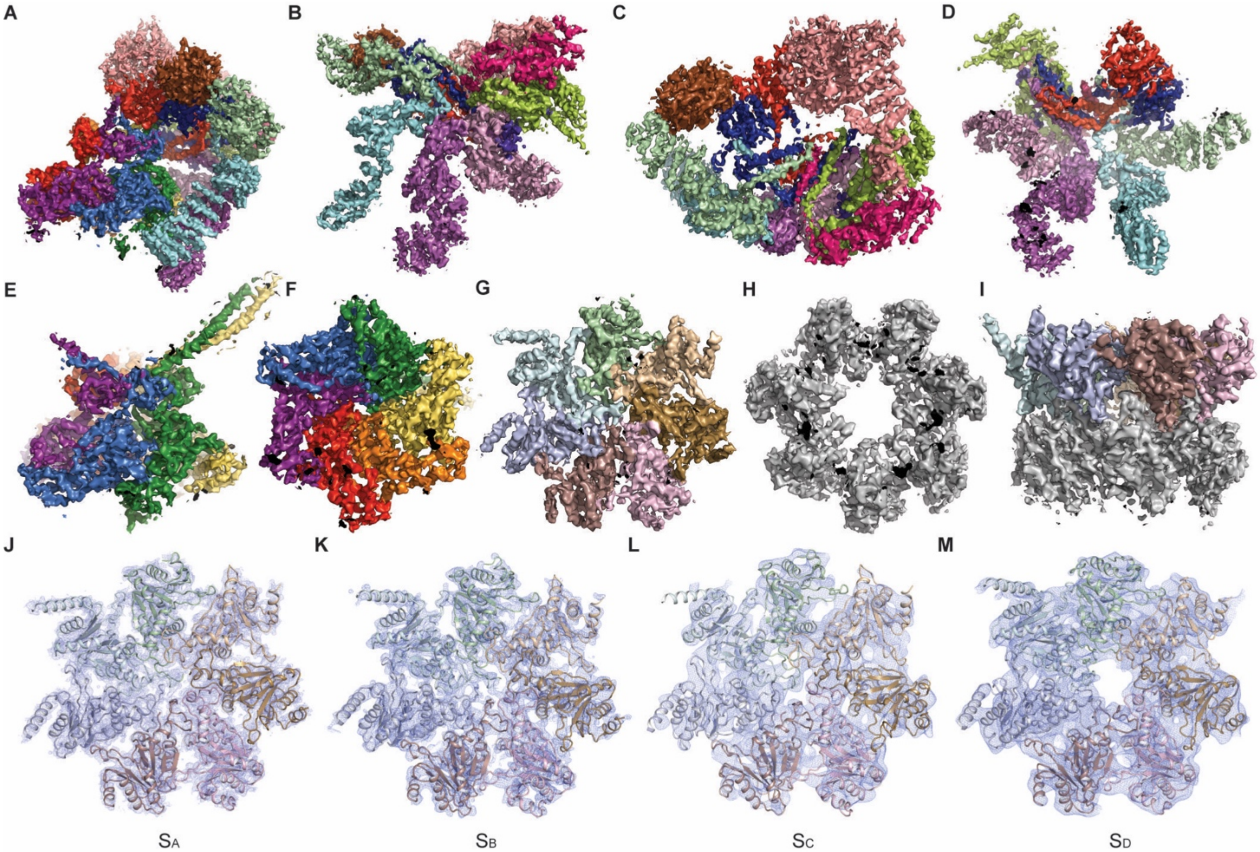


Fig. S4. Overview of cryo-EM density map. (**A**)-(**I**) The density map of the S_A state is shown in a solid surface representation for the side view of the complete RP (panel **A**), the side view of the non-ATPase RP components (panel **B**), the top view of the non-ATPase RP components (panel **C**), 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.

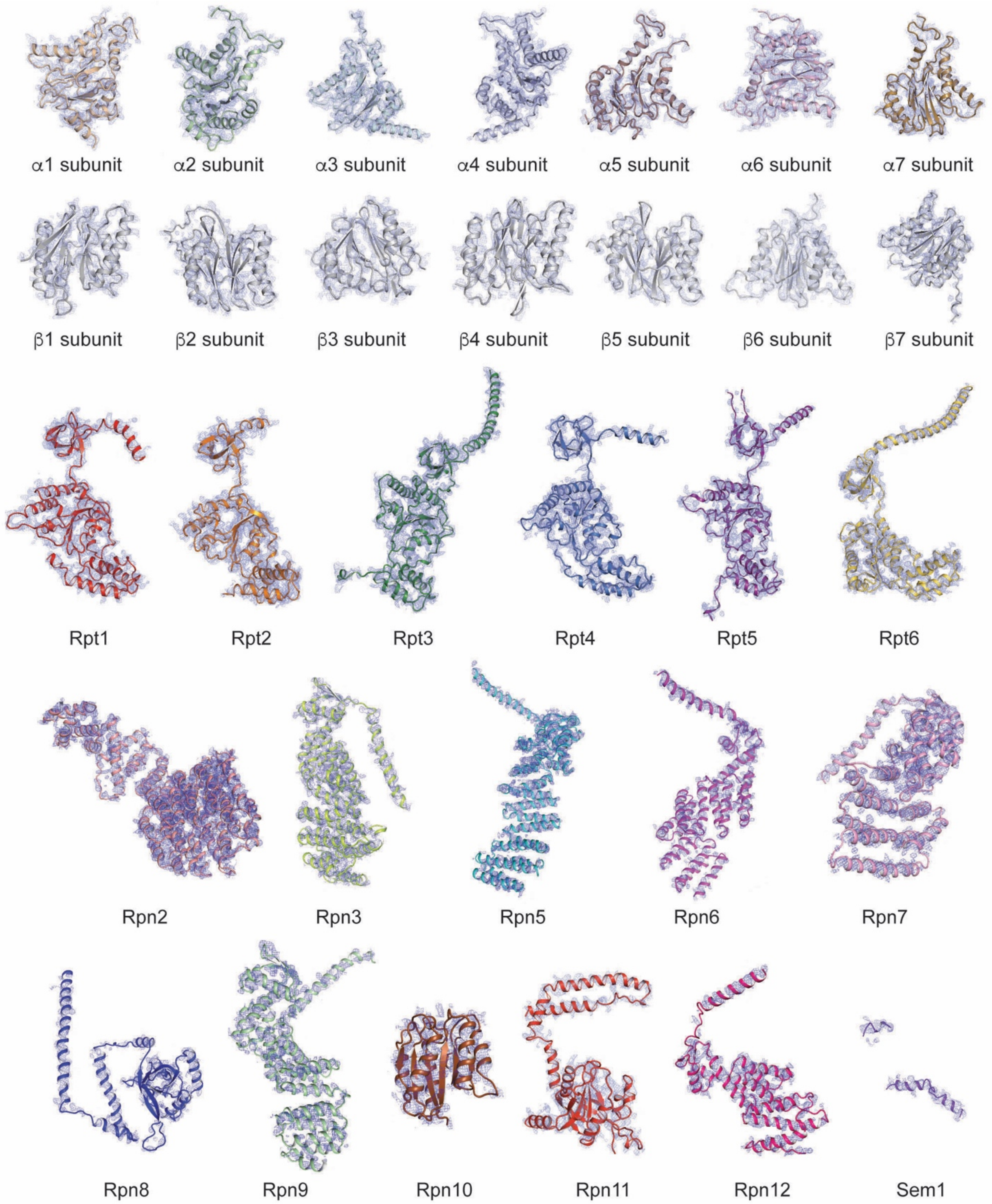


Fig. S5. Cryo-EM densities of subunit structures in the S_A state. The atomic models of each subunit structure in the S_A state, shown as a cartoon representation, are superimposed with the corresponding cryo-EM densities shown in blue meshes.

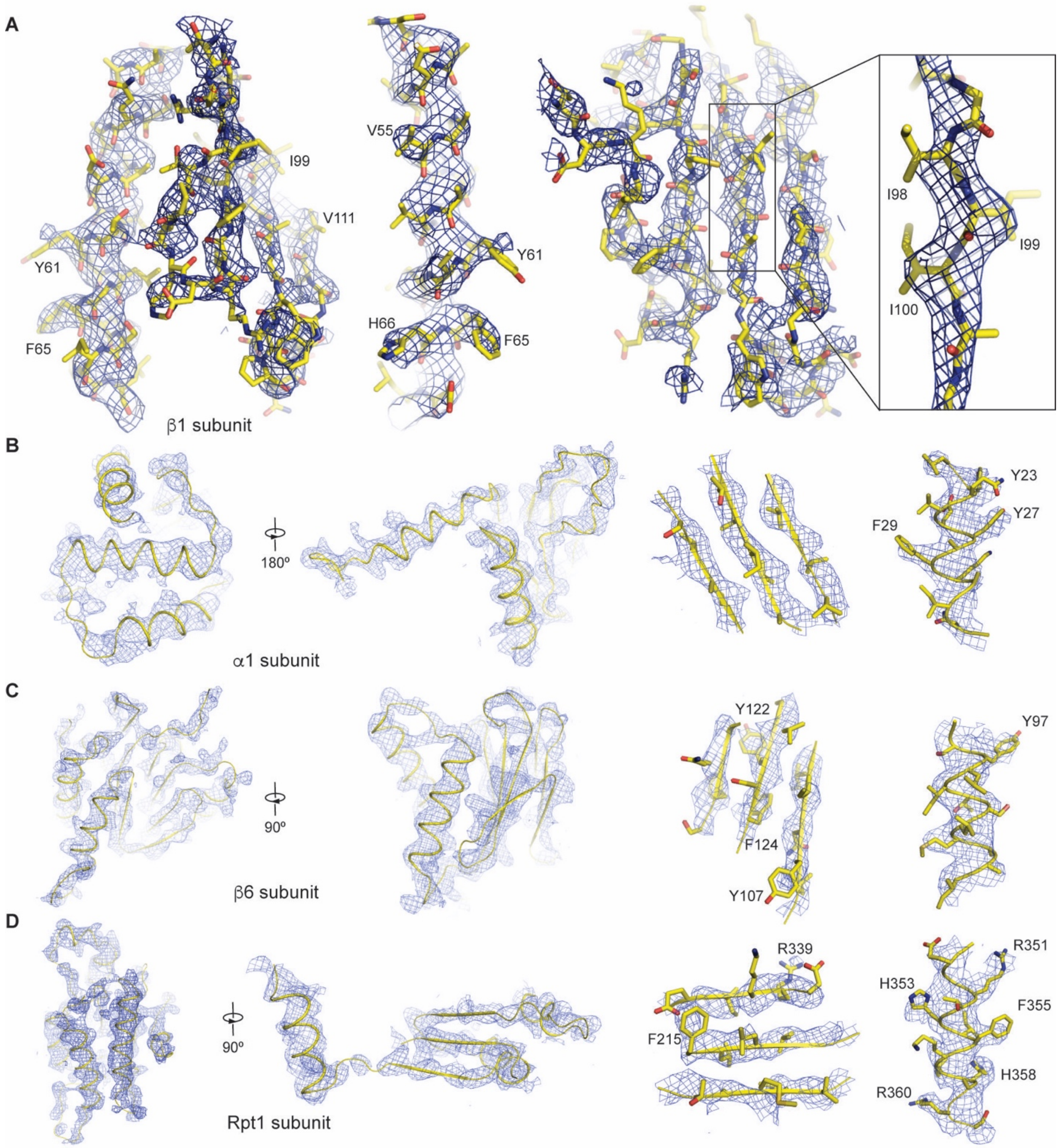


Fig. S6. Fitting of atomic structures into high-resolution cryo-EM densities of the 26S proteasome in the S_A state. The atomic models of subunit structures in the 26S proteasome, β1 subunit (panel **A**), α1 subunit (panel **B**), β6 subunit (panel **C**) and Rpt1 subunit (panel **D**), are superimposed 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.

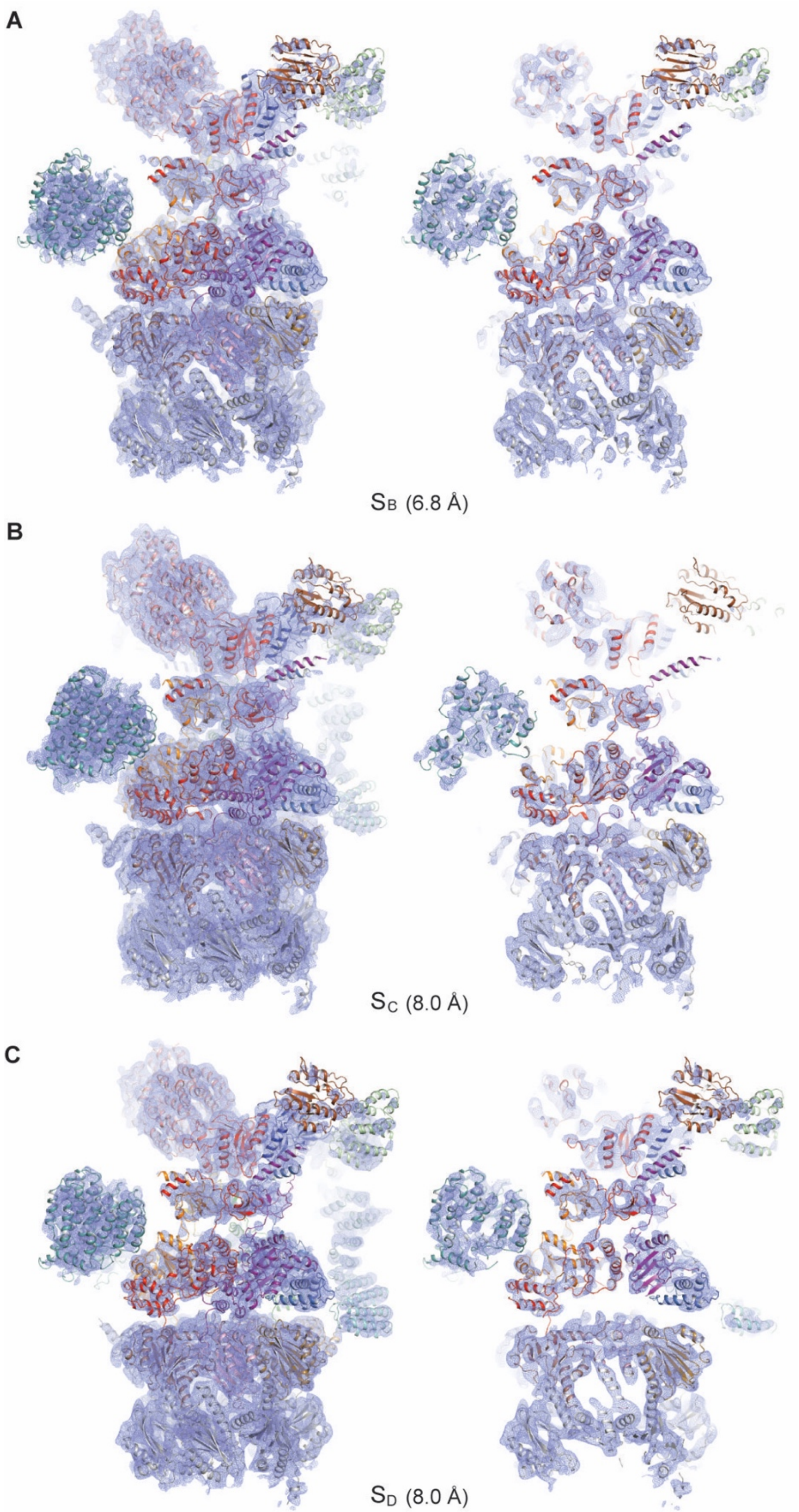
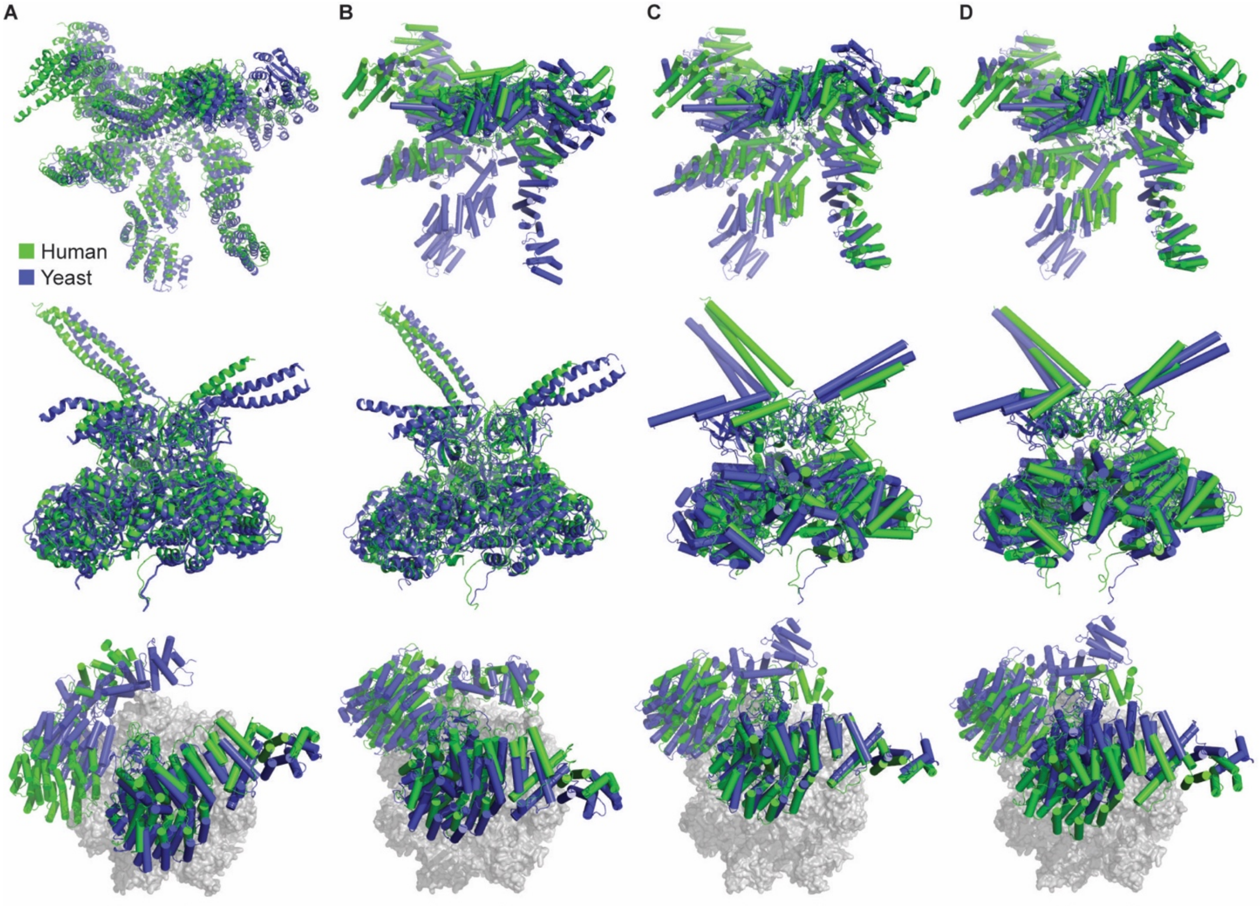
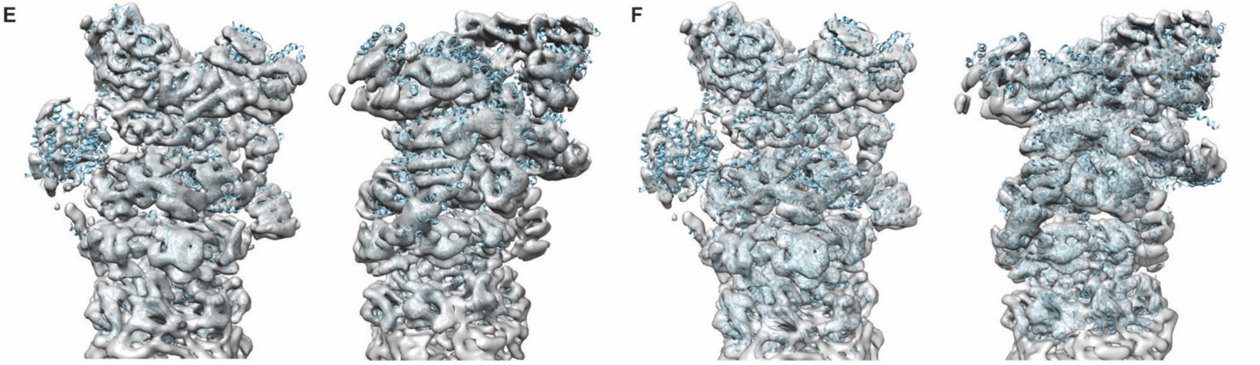


Fig. S7. Cryo-EM densities of the RP-CP subcomplexes in the S_B, S_C and S_D states. (**A**)-(**C**), The density maps of the S_B (panel **A**), S_C (panel **B**) and S_D (panel **C**) states at 6.8, 8.0 and 8.0 Å resolution are shown in blue meshes and superimposed with corresponding pseudo-atomic models.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



S_A vs. s1 S_B vs. s2 S_C vs. s3 S_D vs. s3



Human S_C vs. yeast substrate-bound map Human S_D vs. yeast substrate-bound map

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.

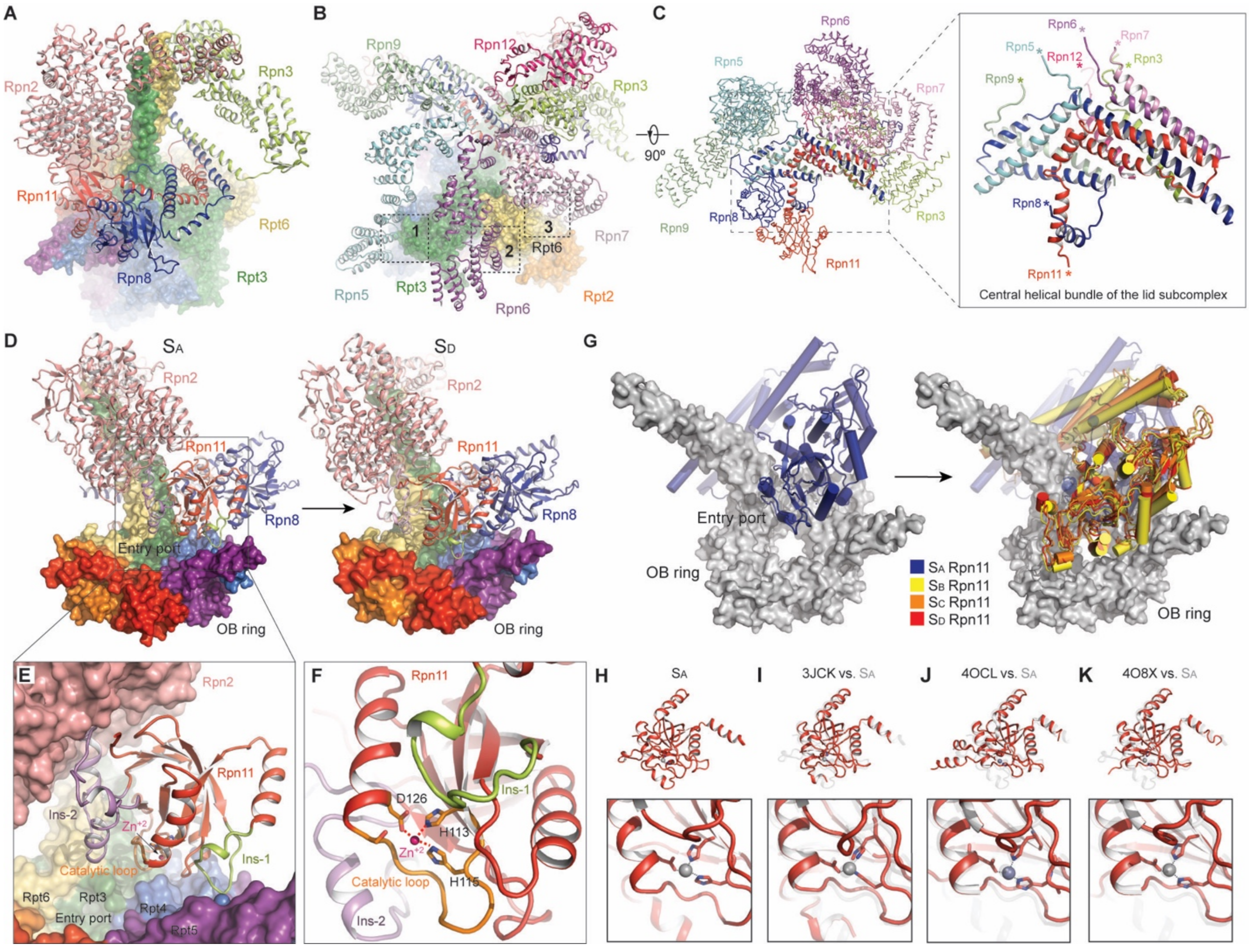


Fig. S9. Key structural features of the RP at the S_A state. (**A**) The lid-base interface surrounding the Rpt3-Rpt6 heterodimer involves Rpn2, Rpn3, Rpn8 and Rpn11, as well as Rpt4. (**B**) The lid-base interface lateral to the ATPase ring is composed of three local inter-subunit interfaces of Rpn5-Rpt3, 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.

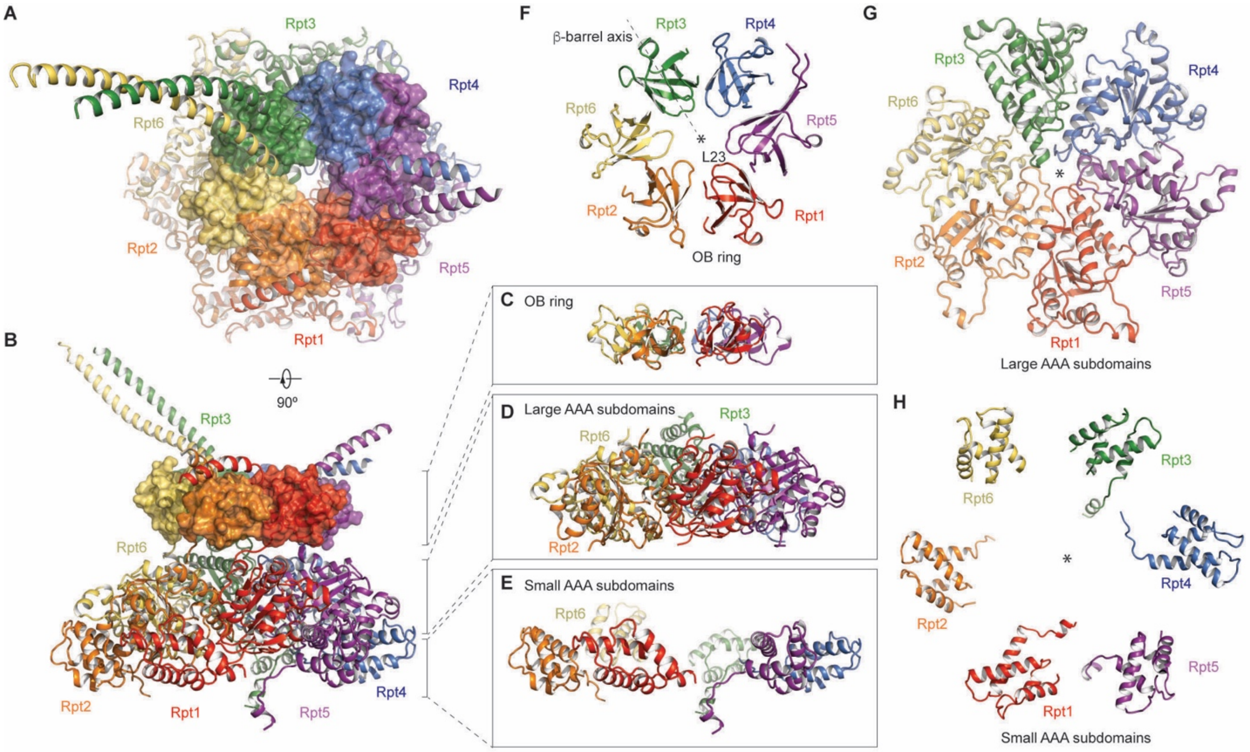
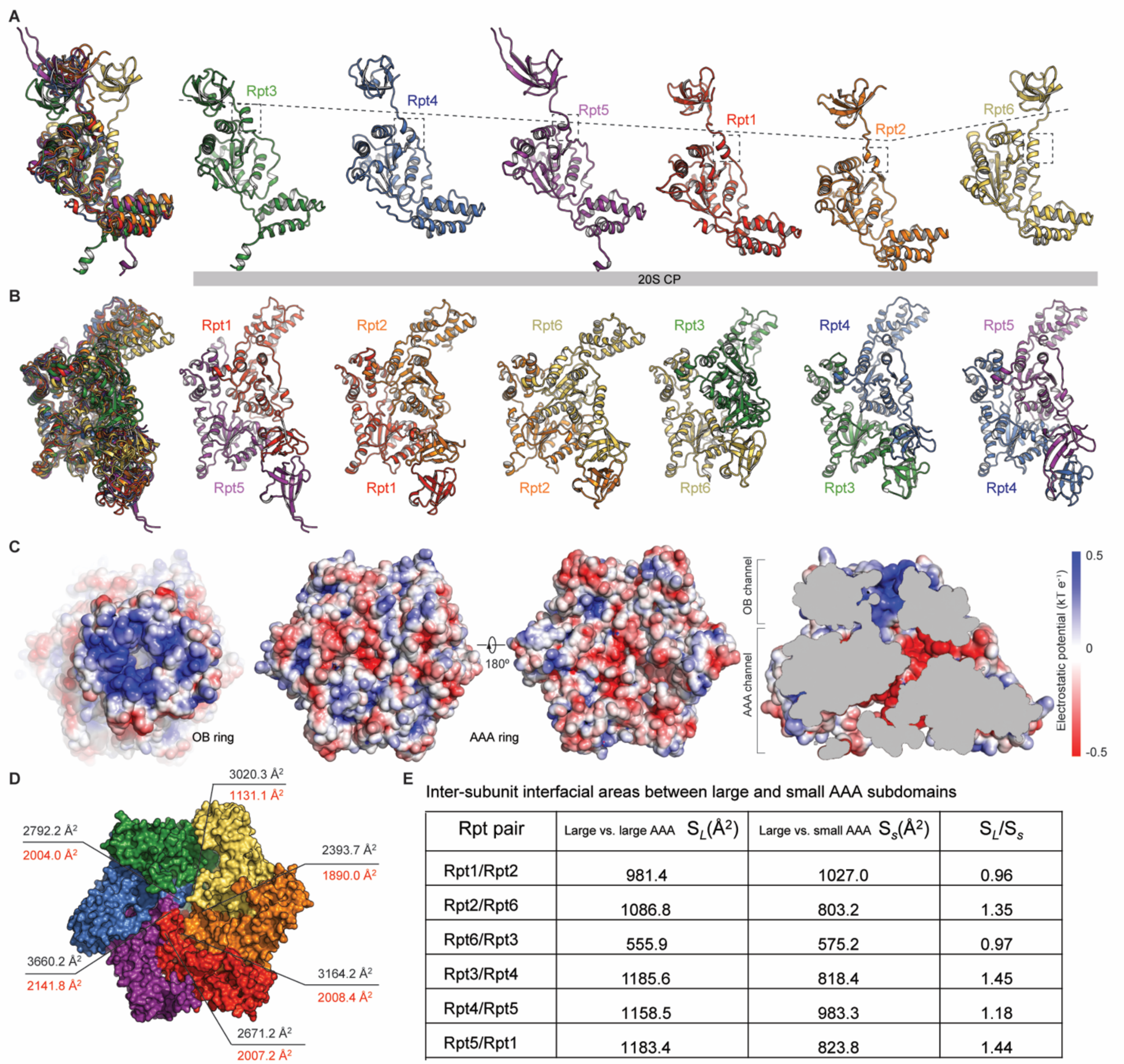


Fig. S10. The atomic structure of the ATPase heterohexamer in the S_A state. (**A**) The overall structure of the ATPase viewed from the perspective of the substrate entry port. The OB ring is in the foreground and is shown with transparent surface superimposed on a cartoon representation of the AAA 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).



Inter-subunit interfacial areas between large and small AAA subdomains

Rpt pair	Large vs. large AAA $S_L(Å^2)$	Large vs. small AAA $S_s(Å^2)$	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.

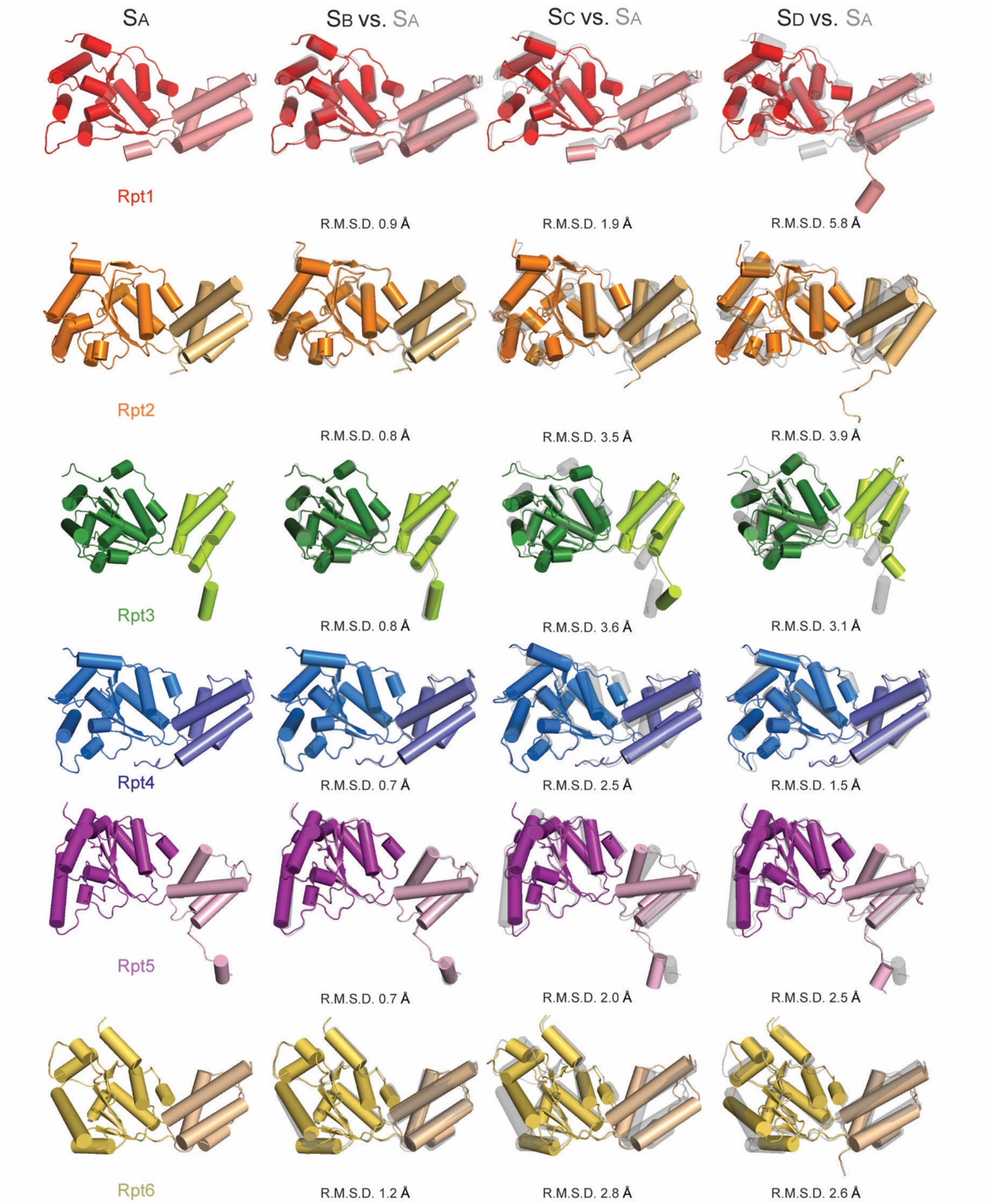


Fig. S12. Subunit conformational changes in the ATPase heterohexamer. Each Rpt subunit is compared for its conformations in the S_A, S_B, S_C and S_D. The root-mean-square deviation (R.M.S.D) is shown for each of alternative states (S_B, S_C and S_D) versus S_A. Detailed statistics of the R.M.S.D. between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

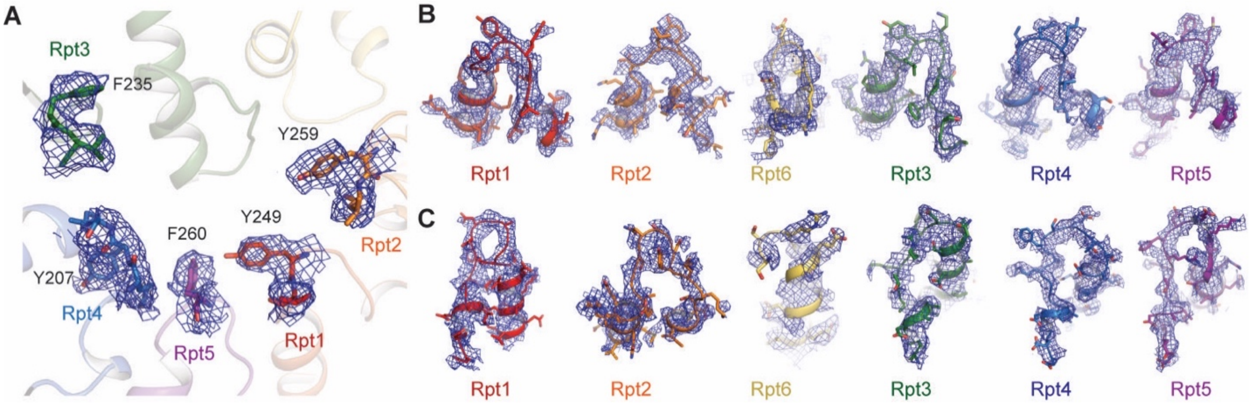


Fig. S13. ATPase pore loops in the S_A state. (**A**) The aromatic side chain density in the pore-1 loop lining in the interior of the AAA channel. (**B**) and (**C**) The densities of pore-1 loops (panel **B**) and pore-2 loops (panel **C**) from Rpt subunits are superimposed with the atomic model.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

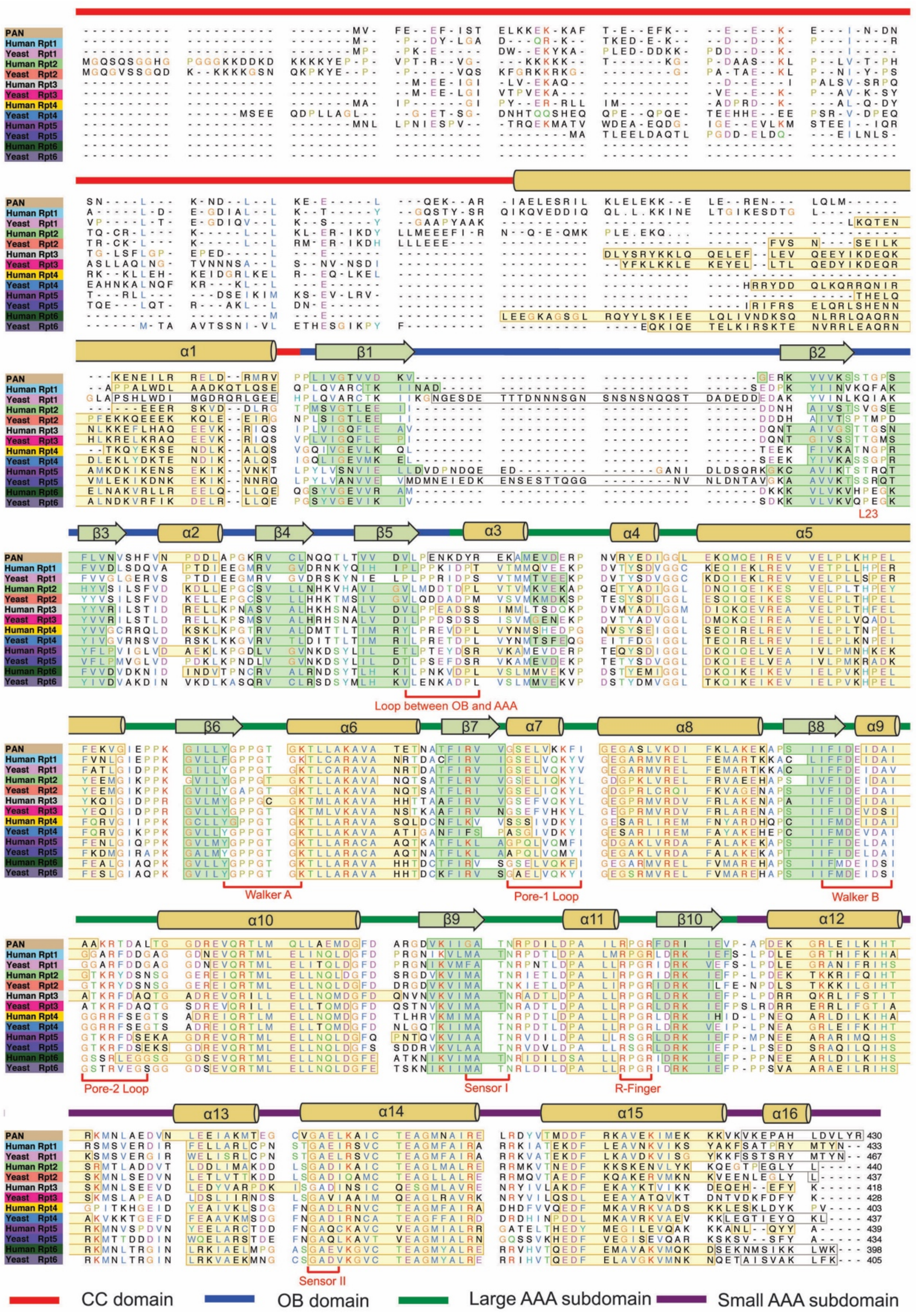


Fig. S14. Structure-based sequence alignment. Sequences from the yeast and human proteasomal ATPases are aligned with each other and with that of the archaeal PAN complex, based on their structures. The yeast structure used in this alignment is based on the pseudo-atomic model previously built from a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

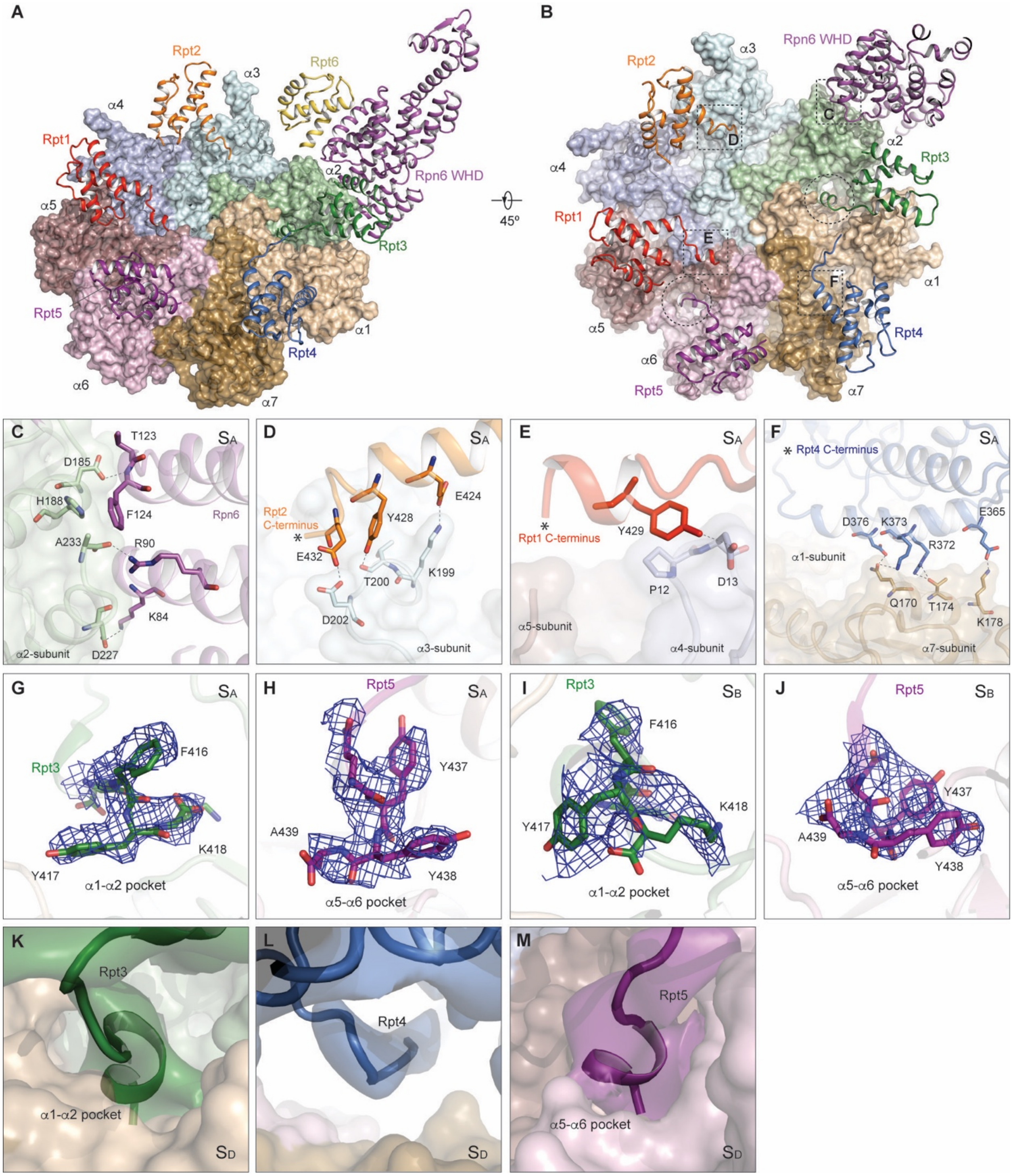


Fig. S15. RP-CP interface in different states. (**A**) Overview of the RP-CP interface of the S_A state in which the α ring is shown in surface representation and the small AAA domains of Rpt1-6 subunits and Rpn6 are shown in cartoon representation. From this perspective, one can see that Rpt6 is the 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.

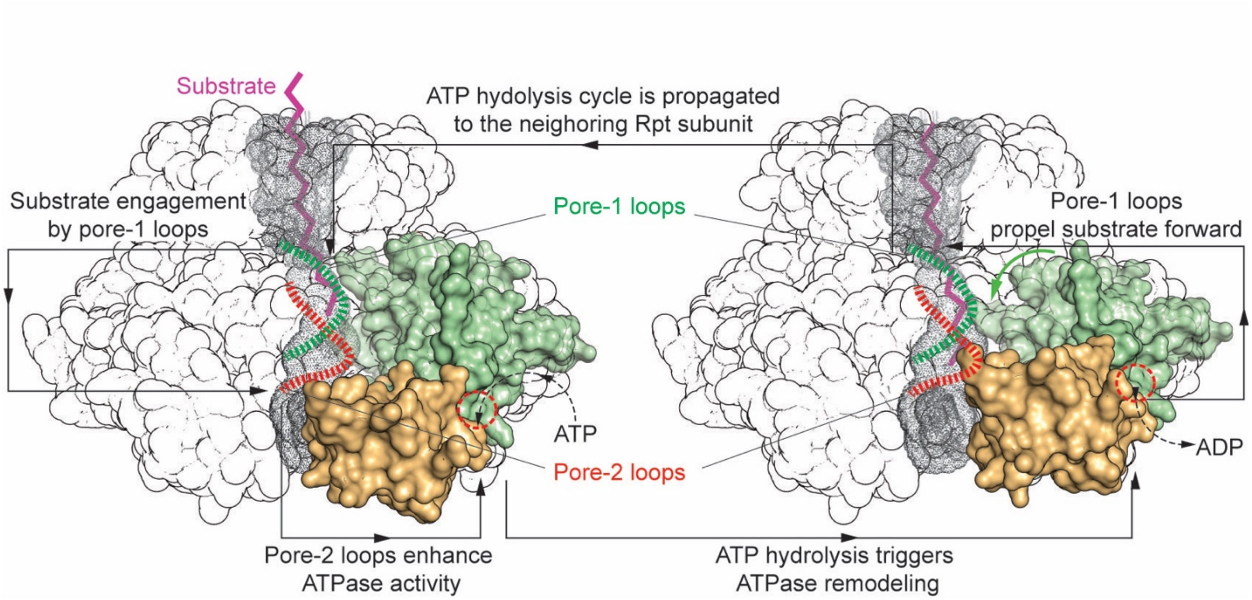


Fig. S16. A hypothetical working model for processive substrate translocation by the ATPase heterohexamer. The ATPase channel is represented by gray dots, and substrate by a purple zigzag chain. The AAA domain of one Rpt subunit is illustrated in a surface representation. The nucleotide-binding site 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.

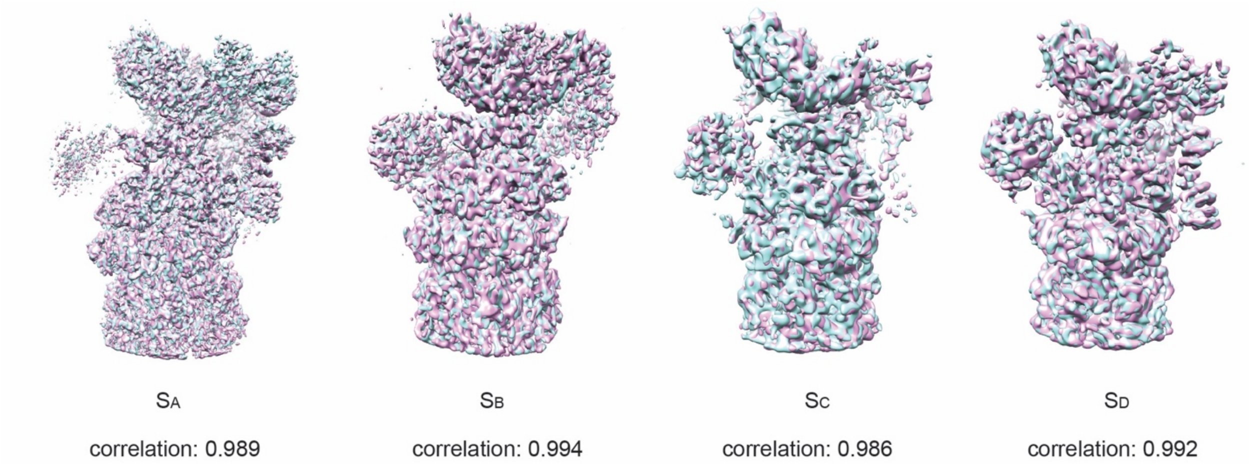


Fig. S17. Validation of RP-CP complex classification using different initial models. Auto-refinement using the S_A state map (cyan) and the S_D state map (pink) as an initial model. Cross correlation value is shown for each RP-CP complex between two independent auto-refinement using S_A versus S_D map 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	S_A state	S_B state	S_C state	S_D state
	Doubly capped with C2 applied	RP-CP	RP-CP	RP-CP	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.