Supporting Information

Wehmer et al. 10.1073/pnas.1621129114

SI Materials and Methods

Endogenous Proteasome Purification. The 26S proteasome was purified by affinity purification using the 3× FLAG-tagged subunit Rpn11. *Saccharomyces cerevisiae* cells (YYS40: MATa *RPN11-3FLAG::HIS3*) were grown for 48 h in YPD medium at 30 °C. The cells were harvested by centrifugation, and the pellet was washed with water. The yeast cell pellet was diluted with buffer A [50 mM Tris·HCl (pH 7.4), 100 mM NaCl, 10% (vol/vol) glycerol, 10 mM $MgCl_2$, 4 mM ATP, 16 mM creatine phosphate, 0.03 mg/mL creatine phosphate kinase] in a ratio of 1 g of cells to 1 mL of buffer. Cells were disrupted using Zirkonia glass beads with a glass bead mill. The supernatant was separated from the beads and cell debris by subsequent centrifugation steps. The pH of the crude extract was adjusted to pH 7.4 using Tris base. FLAG-tagged proteasomes were purified by immunoprecipitation using M2 anti-FLAG beads (Sigma). The filtered crude extract was incubated with the beads for 1.5 h at 4 °C (1 mL of resin per 40 mL of crude extract). Unbound proteins were separated by washing with buffer A. The proteasome was eluted using 200 μg/mL 3× FLAG-peptide in buffer A. The eluate was concentrated by centrifugation in Amicon spin columns (10-kDa molecular weight cutoff). To purify double-capped proteasomes further, the concentrated eluate was applied to a sucrose gradient [15–45% sucrose (wt/vol), 20 mM Hepes (pH 7.4), 40 mM NaCl, 5 mM DTT, 10 mM $MgCl_2$, 4 mM ATP, 16 mM creatine phosphate, 0.03 mg/mL creatine phosphate kinase] and centrifuged for 17 h at 4 °C at 208,000 × *g* in a Beckman SW41 rotor. To obtain proteasome samples with different ATP analogs, ATP and the ATP regenerating system of the sucrose gradient were substituted with 2 mM ADP, 2 mM adenylyl-imidodiphosphate (AMP-PNP), or 2 mM ADP/2 mM beryllium fluoride (BeF_x), respectively. BeF_x was obtained by mixing NaF with $BeCl_2$ [2:1 (wt/wt)]. To get proteasomes with ATP and BeF_x, the proteasome sample containing ATP was incubated with 5 mM BeF_x after sucrose gradient centrifugation. Proteasome-containing fractions were detected by the degradation of the peptide *N*-succinyl-Leu-Leu-Val-Tyr-7-amino-4-methylcoumarin, SDS/PAGE analysis, Bradford assay, and negative-stain EM. The fraction with the highest specific activity was analyzed by MS to estimate the subunit abundances using label-free, intensity-based absolute quantification (47). The samples were stored at −80 °C after quickly freezing with liquid nitrogen.

ATPase Activity Assay. To estimate the inhibition of the proteasomale AAA^+ ATPase by different ATP analogs, the hydrolysis rate of ATP was measured using the colorimetric assay described by Bartolommei et al. (48). One hundred microliters of proteasome sample (20 μM), which was purified in the presence of ADP, was incubated with 900 μL of color solution [125 mM sulfuric acid, 0.5 mM ammonium molybdate tetrahydrate, 10 mM ascorbic acid, 0.04 mM potassium antimony(III) tartrate hydrate] for 35 min at room temperature, and the OD_{850} was determined. In the first step, the measurement was calibrated using different KPO_4 concentrations from 0 to 500 μM in 50 mM Hepes (pH 7.4) and 10% (vol/vol) glycerol. The proteasome ATPase activity was measured in ATP-containing reaction buffer [50 mM Hepes (pH 7.4), 10 mM $MgCl_2$, 4 mM ATP, 10% (vol/vol) glycerol] over 90 min at 30 °C. The phosphate concentration was determined at different time points using the same method as for the KPO_4 calibration curve. The inhibition of the AAA^+ ATPase by BeF_x was determined by measuring the ATP hydrolysis rate after addition of BeF_x to the reaction buffer. For AMP-PNP and adenosine 5′-[γ-thio]triphosphate tetralithium salt (ATP-γ-S), the ATPase activity was measured by using the sample purified in the presence of 2 mM AMP-PNP or 2 mM ATP-γ-S.

Data Acquisition. Data acquisition was performed with an FEI Titan Krios electron microscope. Proteasome samples were plunge-frozen on Quantifoil 2/1 or Lacey carbon-coated grids using a Vitrobot MK IV instrument. Datasets were collected with a Falcon II, Falcon III, or K2 camera using either FEI EPU or the in-house developed software TOM^2 (49). Movies were acquired at a pixel size of 1.35 Å at the specimen level for datasets collected with a Falcon camera and 1.38 Å for datasets collected with a K2 camera. A total dose of ∼45 electrons was distributed over 7 frames for the Falcon II camera, 24 frames for the Falcon III camera, and 20 frames for the K2 camera. The nominal defocus range of the acquisition varied from 1.5 to 3.5 μm. Movies acquired with the K2 camera were corrected for anisotropic magnification distortion using the software "mag_distortion_correct" (50).

Image Processing. Image processing was performed similar to the procedure described by Aufderheide et al. (30). In brief, all movie frames were aligned translationally and summed with an in-house implementation of the algorithm from the study by Li et al. (51). Next, the contrast transfer function (CTF) was estimated using CTFFIND3 (52), and micrographs with a defocus outside the range of 0.8–3.5 μm and a CTF fit score below 0.05 were discarded. After micrograph inspection, 66,462 micrographs of the ATP dataset (K2), 3,304 micrographs of the AMP-PNP dataset (Falcon II), 11,121 micrographs of the ATP-BeF_x dataset (Falcon II), 6,522 micrographs of the ADP-BeF_x dataset (Falcon III), and 5,732 micrographs (Falcon II) plus 6,136 micrographs (Falcon III) of the ADP dataset remained.

Single-Particle Analysis. Single-particle analysis was performed as described by Aufderheide et al. (30). In the first step, micrographs were subjected to automated localization of 26S proteasome particles using the TOM toolbox as described by Beck et al. (26). All further single-particle analysis steps were performed using the RELION software package (42). Proteasome particles were extracted with a box size of 384 × 384 for K2 datasets and 416 × 416 for Falcon datasets, resized to a box size of 256 × 256, and subjected to reference-free 2D classification. Only 2D classes containing high-quality particles were retained. Additionally, particles were classified into single-capped 26S (sc26S) and double-capped 26S (dc26S) proteasome particles according to their 2D class averages.

ATP Dataset. For the ATP dataset, 3D structures of 419,464 sc26S and 269,854 dc26S particles were reconstructed with the RELION software package, using a down-filtered 3D reference of the 26S proteasome, reconstructed from particles from Unverdorben et al. (21) and filtered to 60 Å. Similar to a previous proteasome study (22), the reconstruction result indicated an uneven angular distribution (Fig. S1 *C* and *D*). To reduce the size of the dataset, angular classes with an above-average occupancy were reduced to the mean occupancy by discarding those particles that score worst in terms of the _rlnMaxValueProbDistribution value in RELION, as described in Schweitzer et al. (22). The remaining 196,475 sc26S and 194,803 dc26S particles were subjected to particle polishing. Polished sc26S (C1 symmetry) and dc26S (C2 symmetry) particles were refined using RELION auto-refinement. After 3D alignment of the polished sc26S particles, these particles were classified into 10 classes using a soft-edged mask focusing on

the RP and keeping the previously assigned angles constant. In line, polished double-capped particles were aligned, and after the broken C2 symmetry was addressed, the RPs of the dc26S (pseudo-sc26S) particles were classified into 10 classes using a soft-edged mask on the RP and keeping the previously assigned angles constant. The 3D class averages were compared with the previously identified proteasome states s1, s2, and s3 (21) using the UCSF Chimera (53) fit-in map. Among the sc26S particles, 83,007 were classified into classes showing the s1 state of the proteasome and 50,066 showing the s2 state. Particles from 3D class averages that did not reconstruct into a well-defined 26S proteasome were discarded ("broken" particles). Pseudo-sc26S particle classification resulted in 203,527 s1-like particles for the s1 state and 143,271 s2-like particles. The sc26S and pseudo-sc26S particles from the s1 state were combined and refined, using a soft-edged mask containing one RP and the CP, with a local angular search around the initial angles from the refinement of the polished particles. The resulting density was subjected to postprocessing in RELION for resolution determination and B-factor sharpening. Applying the mask used for refinement, an average resolution of 4.1 Å was obtained. The s2-like particles were refined and postprocessed applying the same procedure as for the s1-like particles, leading to a final reconstruction with an average resolution of 4.5 Å.

ATP/ADP-BeF$_x$ Datasets. A total of 20,325 sc26S and 119,023 dc26S particles from the 2D classification of the ATP-BeF$_x$ dataset and 75,323 sc26S and 109,021 dc26S particles from the 2D classification of the ADP-BeF$_x$ dataset were processed according to the procedure described for the ATP dataset, including 3D reconstruction, particle polishing, alignment of sc26S and pseudo-sc26S, and classification with a mask on the RP. In contrast to the ATP dataset, the BeF$_x$ datasets were not reduced by removing particles from angular classes with an above-average occupancy because the particle numbers were small enough to be processed by RELION. The classification into 12 classes of the ATP-BeF$_x$ particles resulted in 52,335 s1-like and 69,397 s2-like particles. Additionally, 60,819 particles adopted a conformation that could not be assigned to the previously known states. This state is referred to as s4. In a similar manner, the particles from the ADP-BeF$_x$ dataset were classified into 121,560 s1-like particles, 8,217 s2-like particles, 63,957 s3-like particles, and 38,661 s4-like particles. To analyze the conformational heterogeneity of the datasets further, s2-like, s3-like, and s4-like particles were combined, aligned, and further classified into 12 classes, again keeping the previously assigned angles constant. The final classification resulted in 125,903 s2-like particles, 69,723 s3-like particles, and 27,625 well-defined s4-like particles. The s4-like particles were refined and postprocessed, leading to a final reconstruction with an average resolution of 7.7 Å. In this final reconstruction, the density assigned to Ubp6 was not very well resolved, indicating heterogeneity of the particles in the s4 state. Therefore, the particles were further classified into four classes with a sphere mask around the Ubp6 density, using the focused classification described by Bohn et al. (54). Nearly half of the particles did not possess Ubp6 densities, whereas the rest of the particles showed a well-defined Ubp6 density.

AMP-PNP Dataset. From 3,304 micrographs, 5,680 sc26S and 40,709 dc26S particles remained after automated localization and 2D classification. The sc26S and dc26S particles were reconstructed separately. The resulting angles for dc26S were used for the generation of pseudo-sc26S particles, without prior polishing and dataset reduction. The resulting pseudo-sc26S and sc26S particles were merged and classified. Of six classes, two classes contained s3-like particles (67,522 particles). The particles were refined and postprocessed, leading to a final reconstruction with an average resolution of 7.8 Å.

ADP Dataset. The micrographs acquired with the Falcon II and Falcon III cameras were separately subjected to automated proteasome localization and 2D classification of the identified particles. For the Falcon II data, 22,621 sc26S and 35,254 dc26S particles were used for further analysis, and for the Falcon III data, 10,200 sc26S and 14,180 dc26S particles were used for further analysis. The sc26S and dc26S particles were reconstructed separately. The resulting angles for dc26S were used for the generation of pseudo-sc26S particles, without prior polishing and dataset reduction. The sc26S and pseudo-sc26S particles from both cameras were combined and classified into six classes. The class averages could not be clearly assigned to any state.

Model Building. The precise voxel size of the single-particle reconstruction was determined by a cross-correlation analysis using UCSF Chimera (53), maximizing the fit of the 2.3-Å resolution yeast CP crystal structure [PDB ID code 5cz4 (31)] for different voxel sizes. This atomic model of the yeast 20S proteasome was also used as an initial model for the CP of state s1. The opening of the gate in s4 was modeled based on PDB ID code 1z7q (32). The initial model for the RP of the human 26S proteasome was built through comparative modeling using Modeler (55). The RP of the human 26S proteasome [PDB ID code 5l4g (22)] was used as a template, and the Modeler-implemented alignment scheme was used for sequence alignment. For Rpn1 and Rpn2, the X-ray structure of the isolated yeast Rpn2 [PDB ID code 4ady (56)] was used as an additional template. Missing segments in the template or sequence regions with a sequence identity lower than 30% were predicted following the same strategy as described by Schweitzer et al. (22). The CP of the homology model was fitted into the density using UCSF Chimera and then fitted into the density using MDFF (27), which employs molecular dynamics to fit initial models into a density in real space, and thus permits protein flexibility while maintaining realistic protein conformations. MDFF runs were prepared with QwikMD (44) and analyzed using the MDFF graphical user interface of VMD (45). We used NAMD (46) with the correction map-corrected CHARMM22 force field (57) for conducting MDFF. During MDFF runs, restraints to preserve the secondary structure, chirality, and *cis*-peptide bonds were applied to avoid overfitting. We used cascade MDFF starting with a density low-pass-filtered to a resolution of 8 Å and a grid coupling of 0.3, followed by runs with the 4.1-Å resolution map. The coupling of the model to the density was weighted according to the local resolution of the density. The quality of fit of the secondary structure elements was checked through local cross-correlation calculation (58) implemented in VMD. Next, we refined the structure according to the density in real space by combining MDFF with de novo structure prediction as well as Monte Carlo-based backbone and side-chain rotamer search algorithms in an iterative manner following the strategy as described by Schweitzer et al. (22). For further refinement of the atomic coordinates in reciprocal space, the EM map and the Fourier shell correlation were converted into an MTZ file using the ban_mrc_to_mtz.py script (59). The atomic coordinates were then optimized against the Fourier coefficients using maximum-likelihood refinement in PHENIX (60). Differences between the refined model and the map were analyzed in Coot (61), and the model was adapted accordingly by the tools offered in Coot, followed by a second step of refinement in PHENIX, yielding the final model.

The initial structure for the yeast s1 state was the merged structure of PDB ID code 5cz4 (31) for the core and the homology model for the RP based on the human structure [PDB ID code 5l4g (22)]. The refined final structure of s1 was used to initiate an MDFF run into the density for the s2 state. The final refined

structure of the s2 state was fitted through MDFF into the density of the s3 state, and the final s3 state structure was then fitted into the s4 density.

Rpn13 Modeling. To determine the orientation of Rpn13, we performed an exhaustive 6D correlation scan, using the EM map of the best resolved class of a classification with a spherical mask on the RP (Fig. S2*A*), starting from the position on the yeast model of the s2 state (PDB ID code 4cr3) (21). Z-scores were calculated based on the cross-correlation values between the refined EM map and a density generated from the fitted atomic structure, as previously demonstrated (30) (Fig. S2 *C*–*F*). The angle and shifts found by the initial scan were refined in UCSF Chimera. With these refined alignment parameters, the scan was repeated (Fig. S2 *E* and *F*). To support the result, we analyzed the cross-linking MS data of the 26S proteasome, in which two confident linkages between Rpn2 and Rpn13 were identified (Table S2).

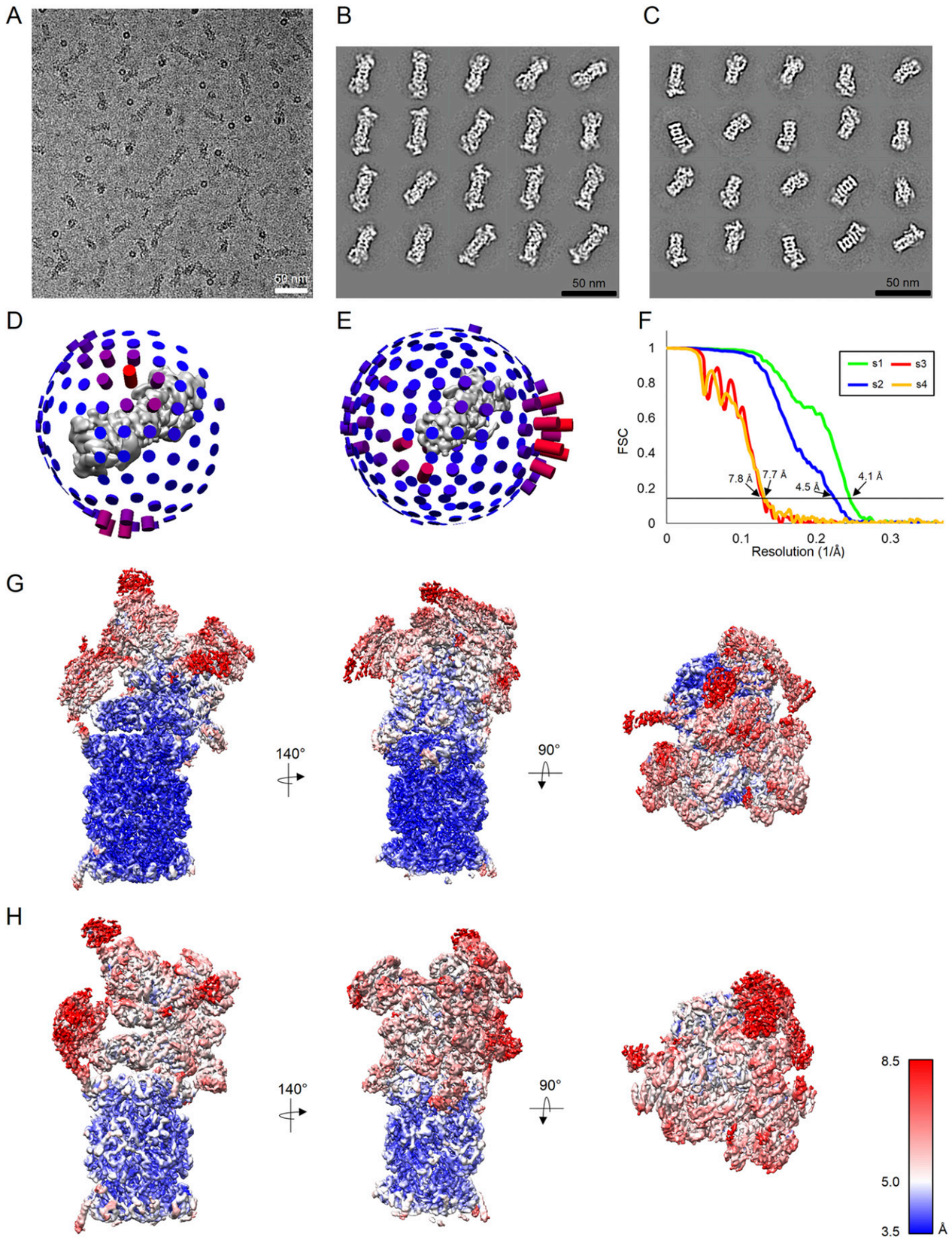

Fig. S1. Typical cryo-EM micrograph of the K2 dataset (*A*) and representative reference-free 2D class averages of high-quality dc26S (*B*) and sc26S (*C*) proteasome particles. The angular distribution of the 3D reconstruction of dc26S (*D*) and sc26S (*E*) proteasome particles from the K2 dataset is shown, demonstrating unequal occupancy of angular classes. (*F*) Resolution of the reconstructions of the s1, s2, s3, and s4 states on the basis of the gold-standard Fourier shell correlation criterion ($FSC_{0.143}$). Three-dimensional reconstructions of the yeast proteasome s1 (*G*) and s2 (*H*) states are colored according to the local resolution as given by the color gradient (angstrom units).

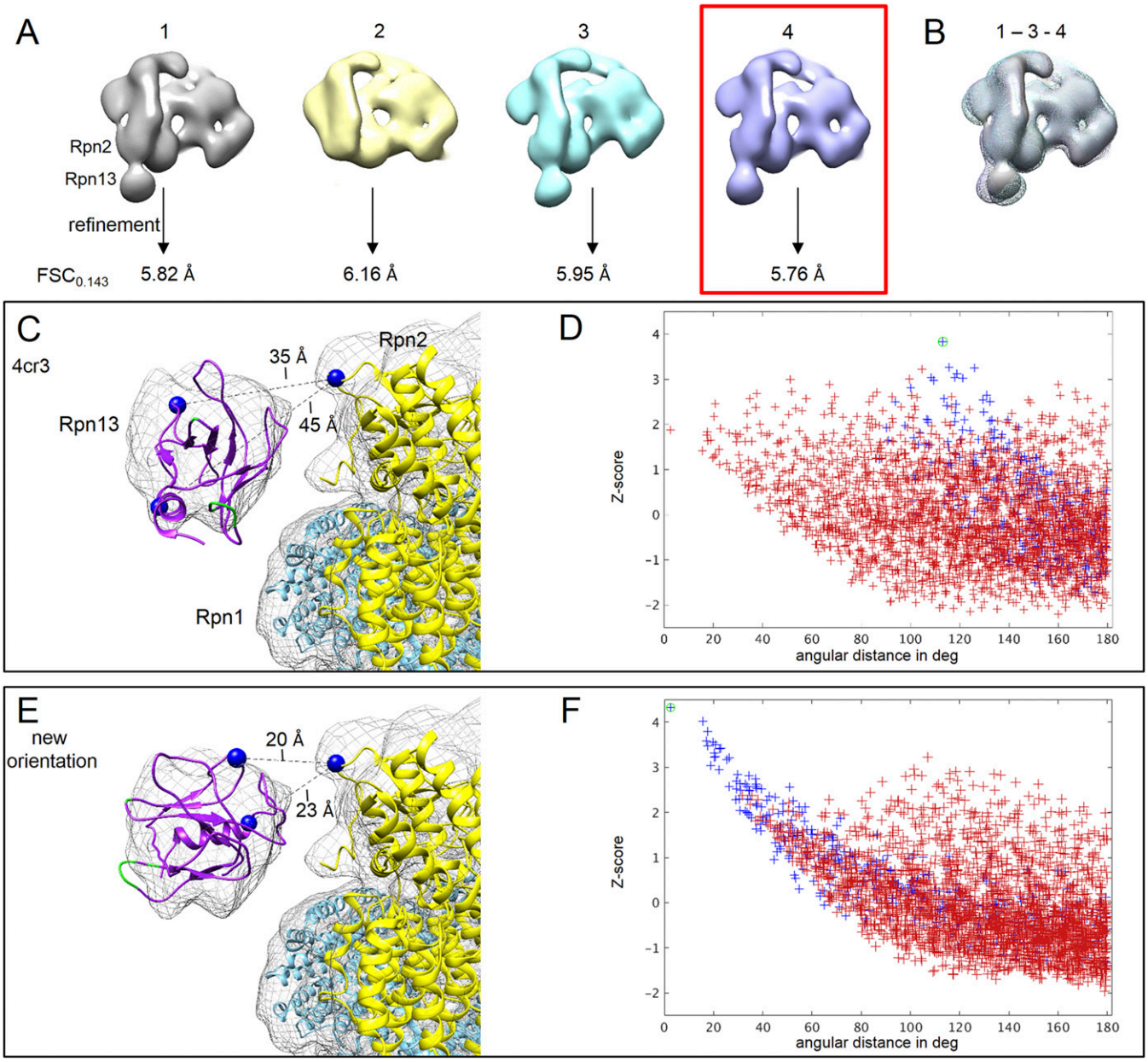

Fig. S2. Orientation of Rpn13 in the s2 state. (*A*) Class averages of the classification with a spherical mask on the lid and the corresponding resolution of the refined reconstructions. (*B*) Overlay of the class averages from *A* showing a density for Rpn13. Classes 3 and 4 are represented as meshes. (*C*) Orientation of the Rpn13 model (PDB ID code 4cr3; cyan; Rpn2, yellow; Rpn13, purple; and ubiquitin-binding residues of Rpn13, green; cross-linked lysine residues, blue spheres) fitted into the EM density of the s2 state before (*C*) and after (*E*) the assessment of orientation specificity of fitting. The distance between the two best-scoring cross-links is shown. (*D* and *F*) Z-scores of the correlation values are plotted against the angular distance to the positions shown in *C* and *E*, respectively. Value for the best fitting position (marked by the green circle) is more than 4 SDs above the mean value. deg, degrees.

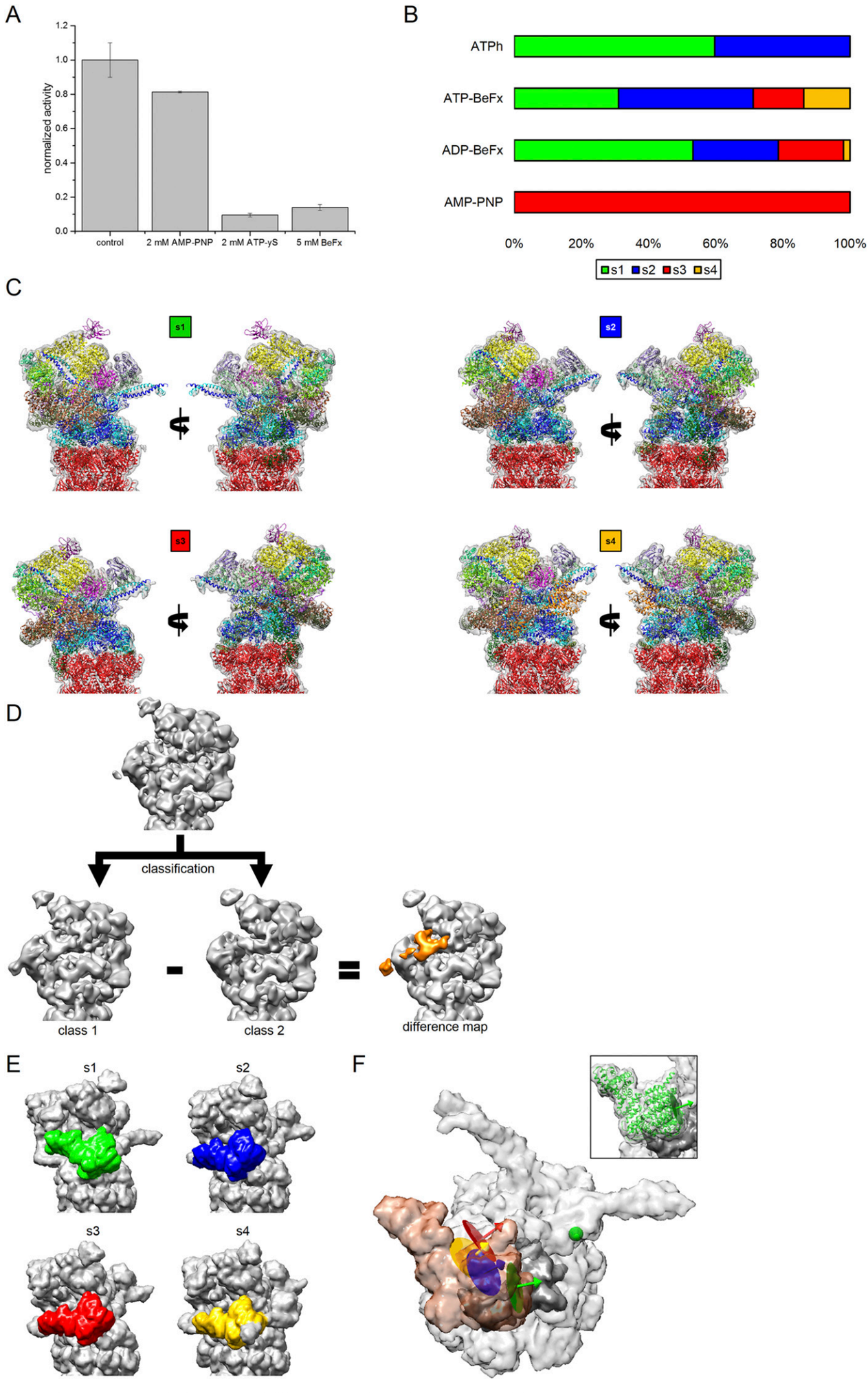

Fig. S3. The 26S proteasome in the presence of different ATP analogs. (*A*) ATPase activity of the 26S proteasome was monitored by measuring the hydrolysis of ATP in the presence of the ATP analogs AMP-PNP, ATP-γ-S, and BeF_x at different concentrations to determine the inhibition of the proteasomal AAA^+ ATPase for subsequent cryo-EM experiments. The activity was normalized to the control sample, which contained only ATP. (*B*) Percentages of state distribution of proteasome particles (sc26S and pseudo-sc26S) from the datasets after 3D classification. (*C*) Cryo-EM maps of the four states of the proteasome filtered to a resolution of 8 Å (s1, *Left Top*; s2, *Right Top*; s3, *Left Bottom*; and s4, *Right Bottom*) with the corresponding model depicted using the same color code as applied in Fig. 1. (*D*) Classification of s4-like particles. The difference map of the two classes is shown in orange, depicting the density of Ubp6. (*E*) Position of Rpn1 in the four states of the proteasome. Based on the atomic models, the simulated density of the proteasome at a resolution of 10 Å is shown in gray and the subunit Rpn1 is colored as given in *B*. (*F*) Simulated density of the AAA^+ ATPase, α-ring of the CP (gray), and Rpn1 (brown) in the s1 state, with superimposition of the torus position of Rpn1 (*Inset*), in the four different states (colors as in *C*). The binding site of Ubp6 at the OB ring is shown as a sphere (green), and the Rpt1 subunit that was used for alignment is shown in dark gray.

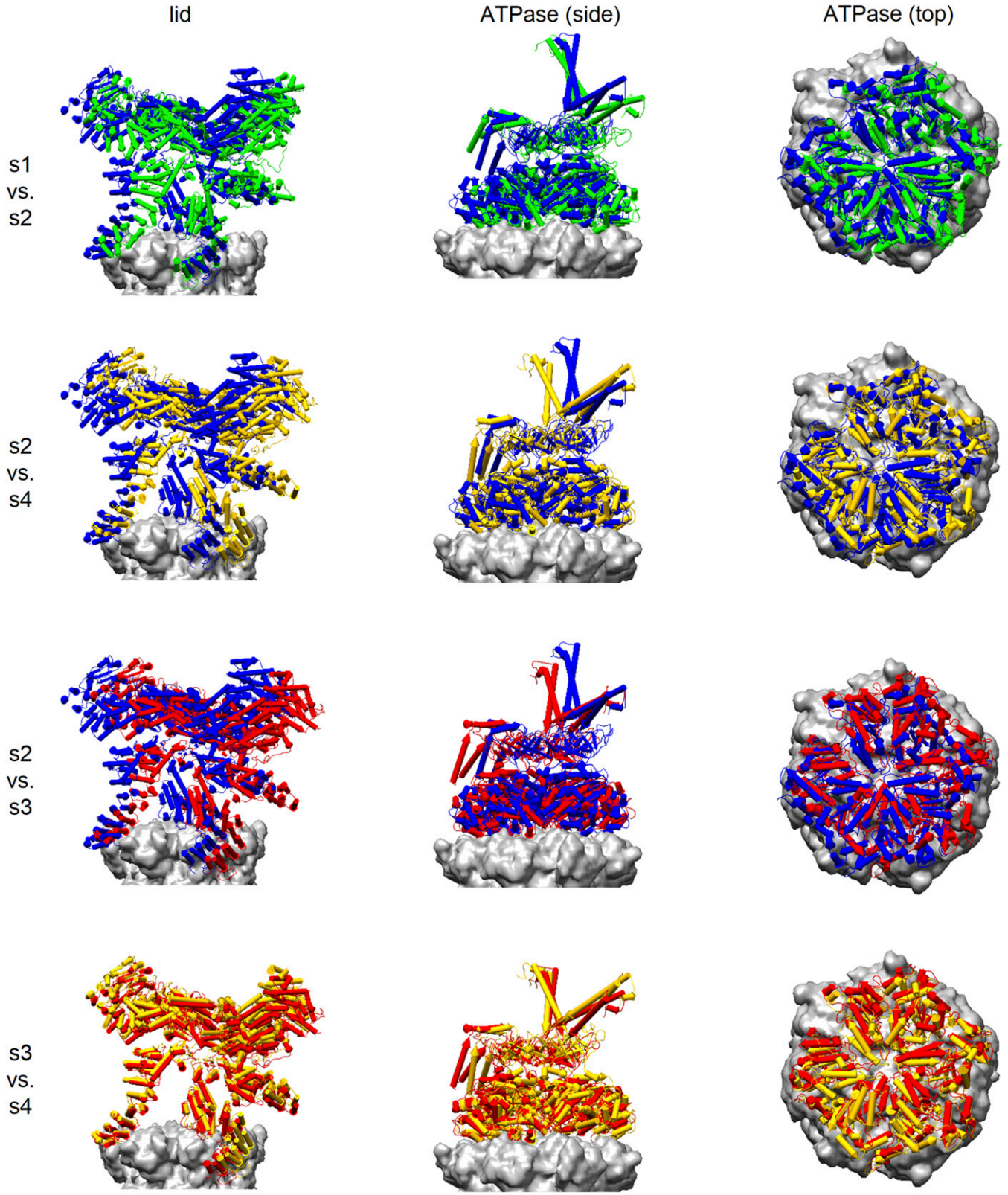

Fig. S4. Superimposition of the model of the 26 proteasome lid (*Left*) and ATPase (*Center* and *Right*) depicted in a pipes-and-planks representation (s1, green; s2, blue; s3, red; and s4, yellow). The EM density of the CP is shown in gray.

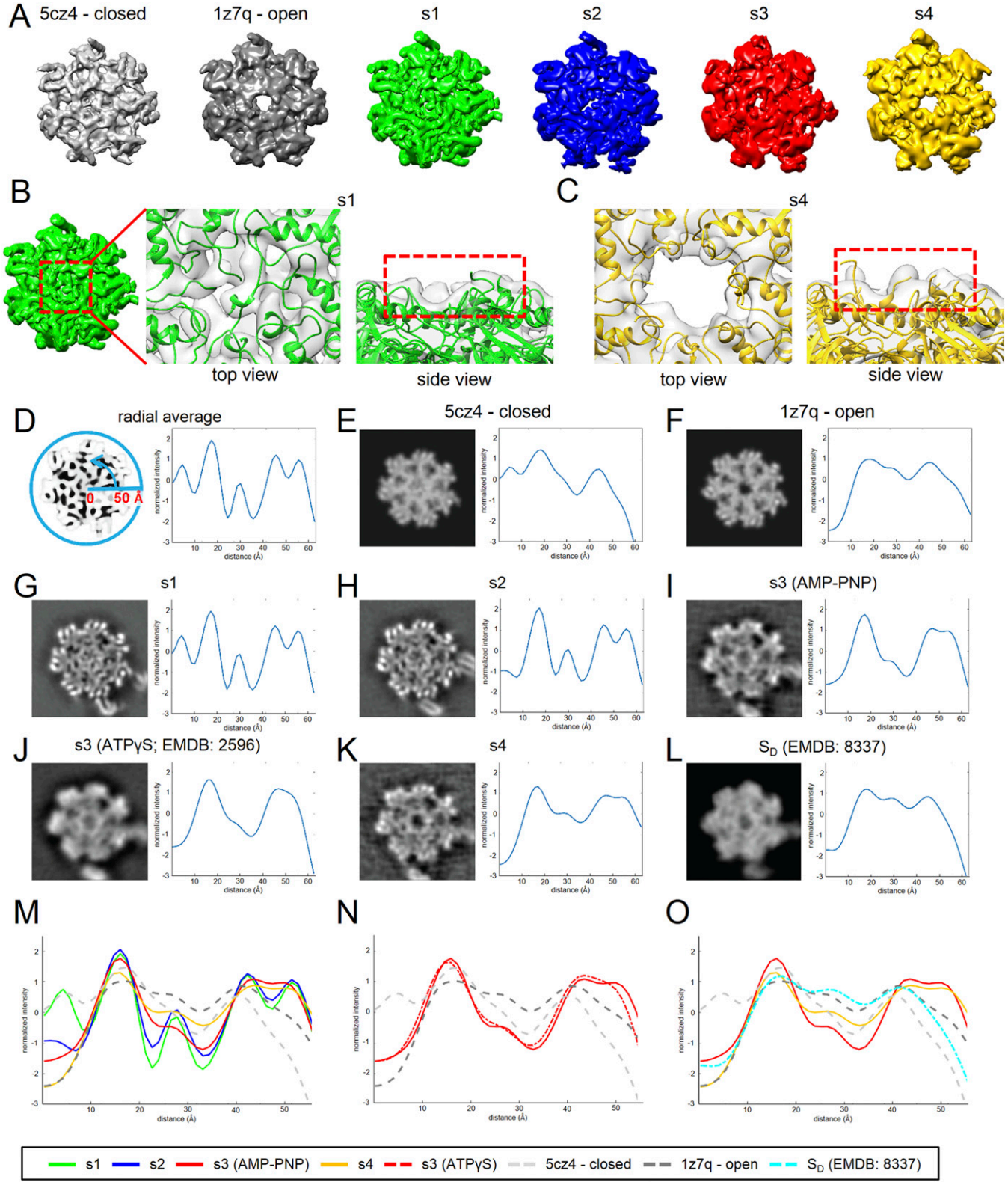

Fig. S5. (*A*) Density maps simulated from crystal structures of the 20S proteasome showing a closed gate (PDB ID code 5cz4; *Left*, gray) and open gate (PDB ID code 1z7q; *Right*, dark gray) at a resolution of 8 Å and segmented densities of the CP α-rings of the four different 26S proteasome states (s1, green; s2, blue, s3, red; and s4, yellow). (*B* and *C*) Close-up view of the EM density and model of the gate in the s1 and s4 states. (*D*) Schematic representation of the radial average of the α-ring of the CP; the radius of the CP gate is ∼6 Å (7). Determination of the radial average of the simulated maps (8 Å) of the crystal structure of a closed 20S (*E*; PDB ID code 5cz4) and open 20S (*F*; PDB ID code 1z7q) is illustrated. (*G–I* and *K*) Radial average of the EM maps of the proteasome structures determined in this study. The radial average of the proteasome CP in the presence of ATP-γ-S (*J*) (21) and of the human S_D state (*L*) (25). (*M–O*) Comparison of the radial averages of different EM maps.

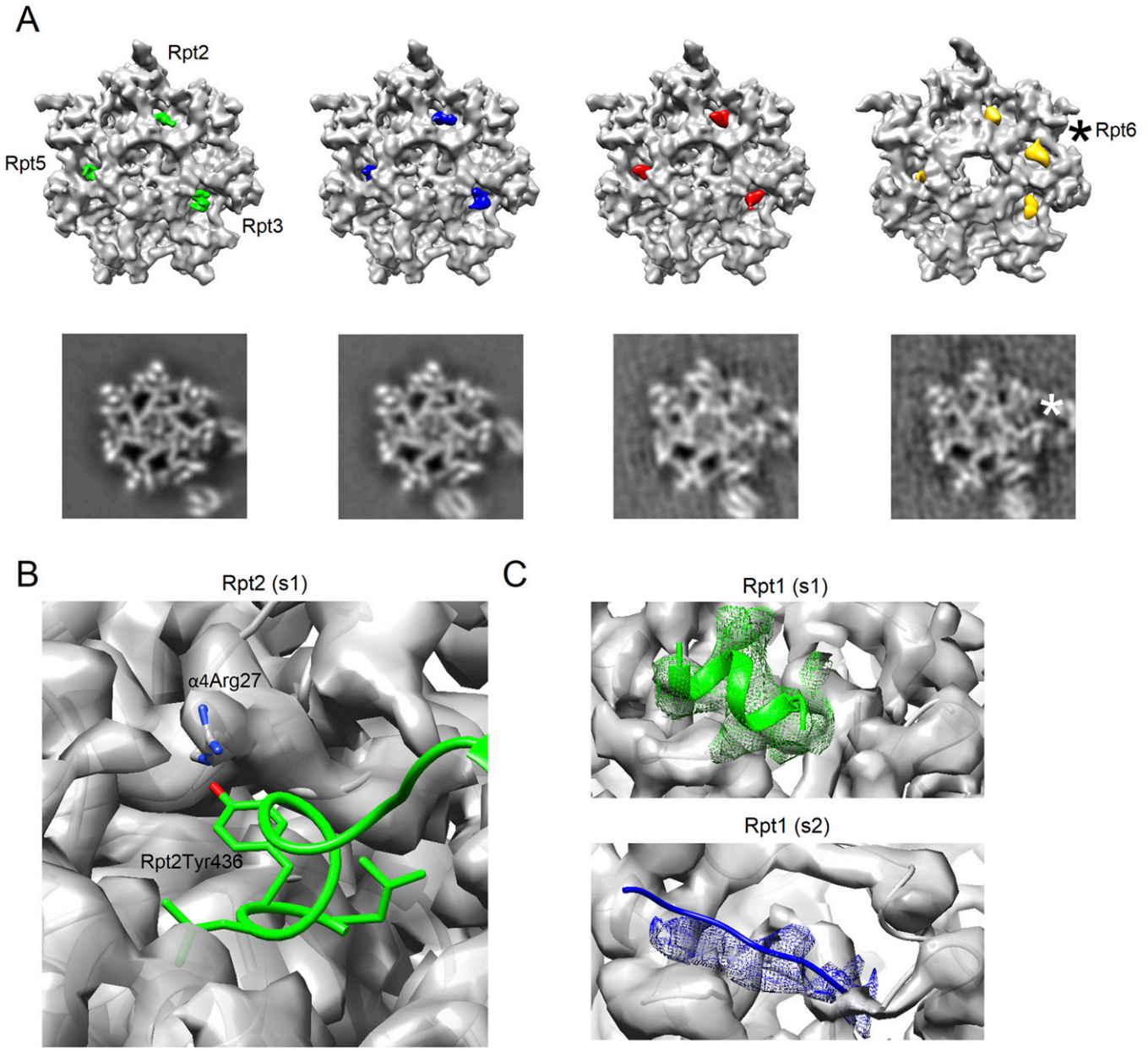

Fig. S6. HbYX motifs of the 26S proteasome. (*A*, *Top*) Difference map resulting from the subtraction of the simulated map of the α-ring of the CP and the experimental map for the four states (s1, green; s2, blue; s3, red; and s4, yellow) is depicted as an overlay with a simulated map (gray) of the closed 20S (PDB ID code 5cz4; s1, s2, and s3) and open 20S (PDB ID code 1z7q; s4). (*A*, *Bottom*) A slice through the EM density at a similar position. (*B*) HbYX motif of the Rpt2 subunit in the s1 state. The EM density is depicted in gray. (*C*) Comparison of the C-terminal tail of the Rpt1 subunit in the s1 (green) and s2 (blue) states. The EM map of the CP is shown in gray, and the EM map of the C-terminal tail of Rpt1 is shown as mesh.

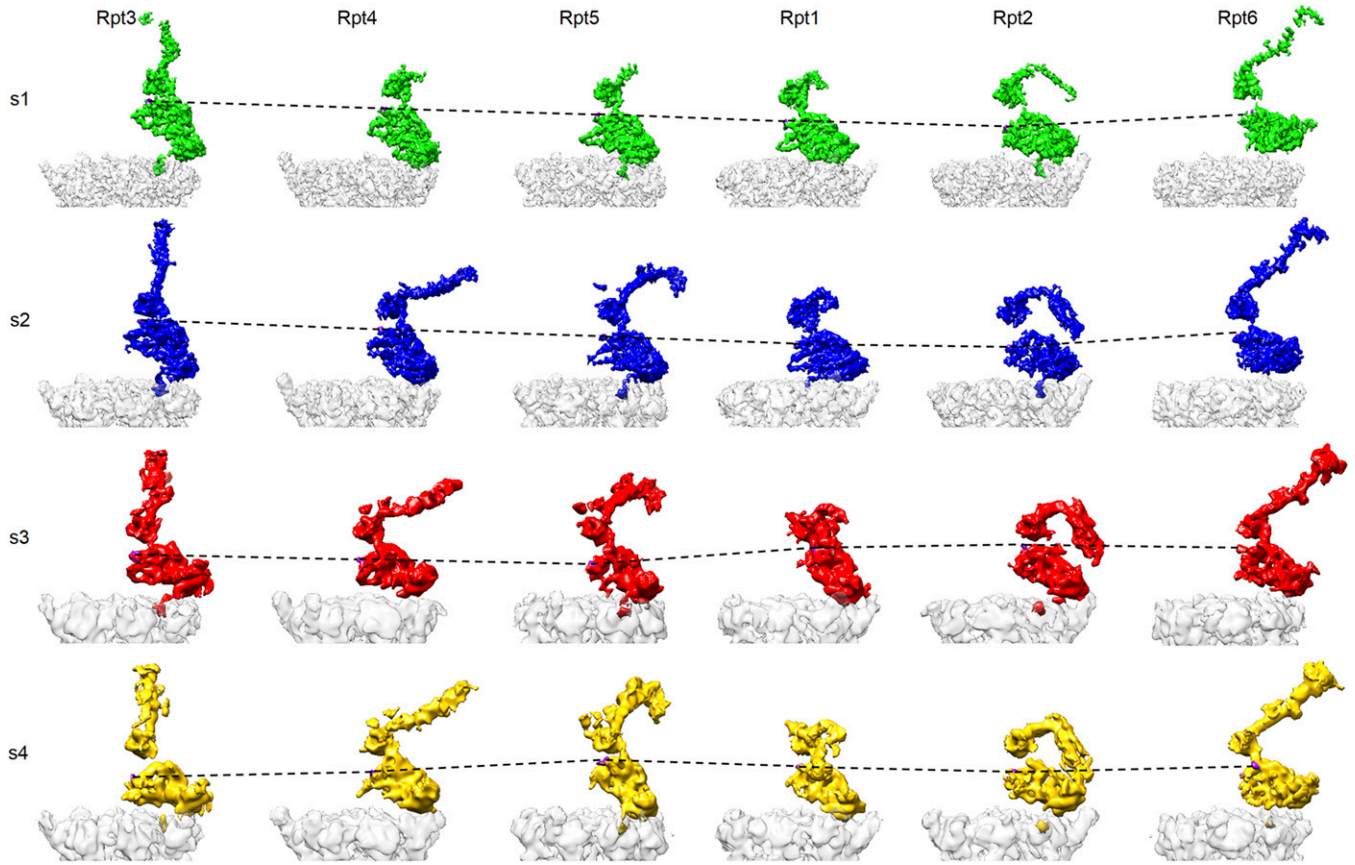

Fig. S7. Organization of the AAA$^+$ ATPase and interaction with the CP. Segmented densities of the Rpt subunits of the four states rotated around the CP axis (60° steps) are shown. The dotted lines represent the position of the conserved tyrosine residues of the pore-1 loops.

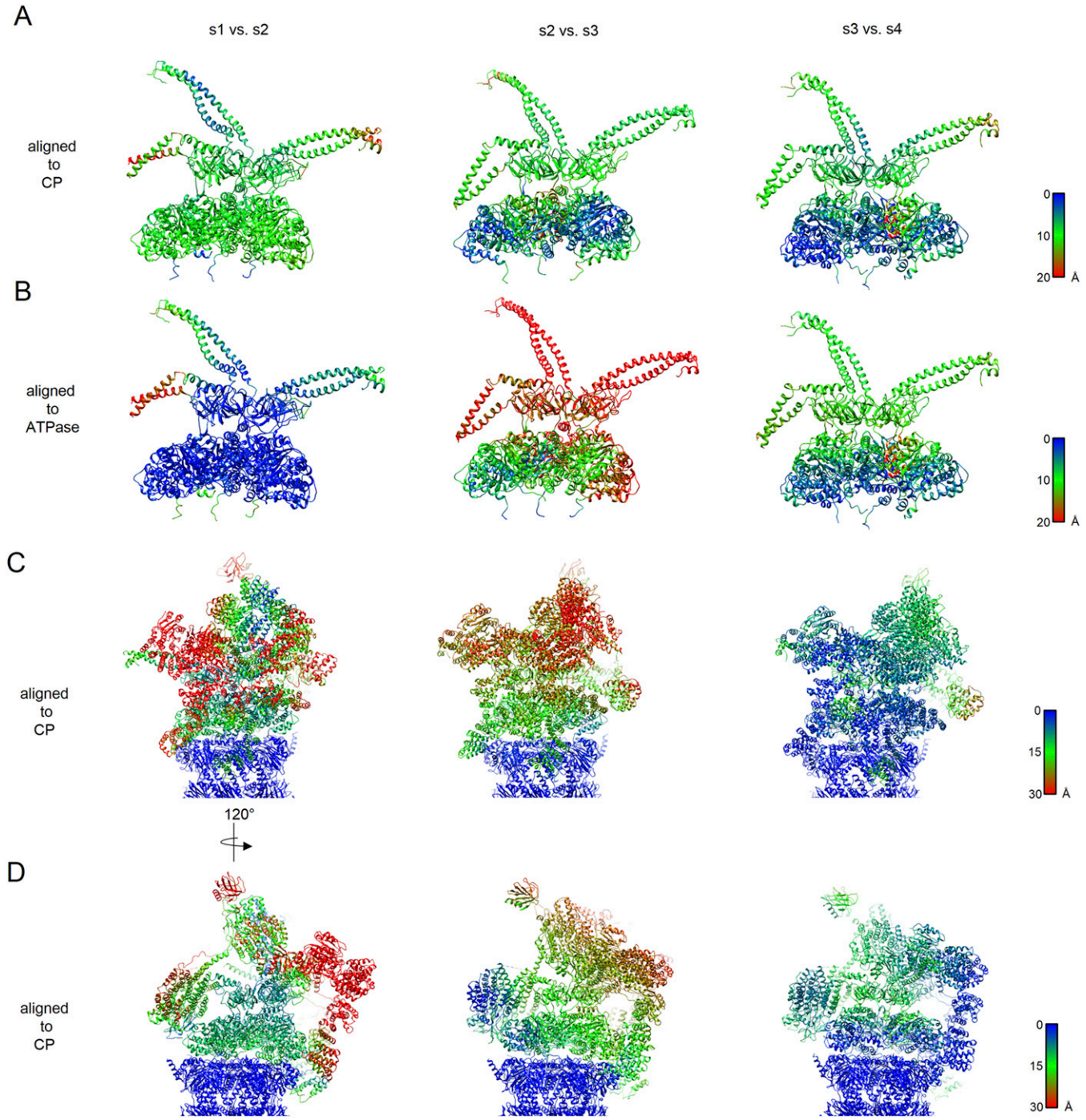

Fig. S8. Residue-wise root mean square deviation (rmsd, in angstroms) of the 26S proteasome. The residue-wise rmsds of the AAA$^+$ ATPase in the s1 vs. s2 state (*Left*), s2 vs. s3 state (*Center*), and s3 vs. s4 state (*Right*) are shown, aligned to the CP (*A*) and to the ATPase (*B*). (*C* and *D*) Residue-wise rmsds of the whole proteasome aligned to the CP. The figure is colored according to the rmsd, as given by the respective color gradient (in angstroms).

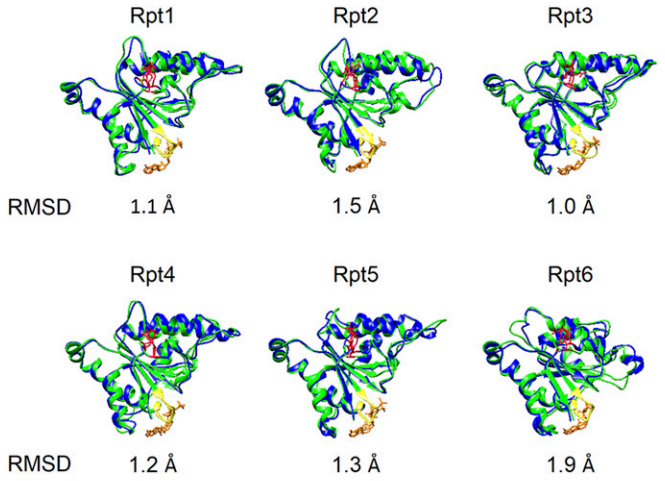

Fig. S9. Superimposition of the large domains of the Rpt subunits of AAA$^+$ ATPase between the s1 (green) and s2 (blue) states.

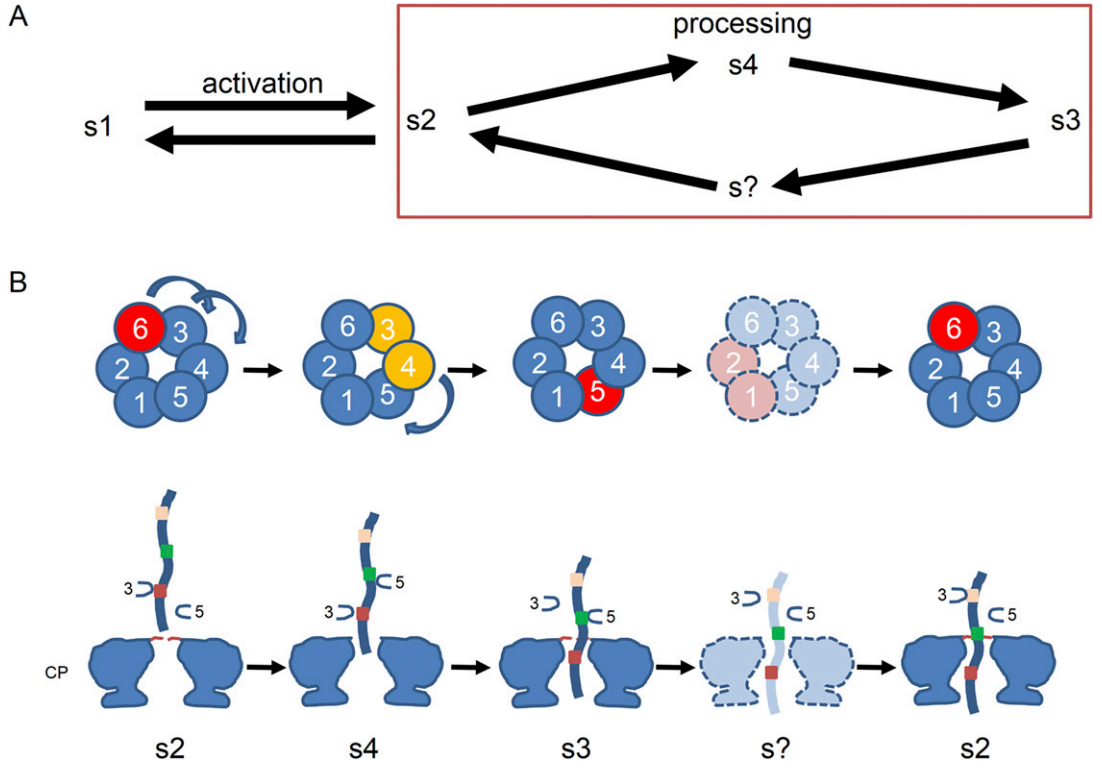

Fig. S10. Model of the coupling between ATP hydrolysis and substrate translocation of the 26S proteasome. (*A*) Upon activation, the proteasome shifts from the substrate-accepting state (s1) to the substrate-processing states (s2, s3, and s4). (*B*, *Top*) ATP hydrolysis cycle occurs in a sequential order around the AAA$^+$ ring. The conformational state of the nucleotide-binding pockets is shown in different colors (open, red; engaged, blue; and intermediate, yellow) as seen in Fig. 4. (*B*, *Bottom*) Concerted conformational change of the AAA$^+$ ATPase with the CP gate. s?, hypothetical states.

Table S1. Summary of single-particle analysis

Dataset	Micrographs	sc	dc	s1	s2	s3	s4	Not well defined
ATP	66,462	196,475	194,803	286,534	193,337	—	—	106,210
ATP-BeF$_x$	11,121	20,325	119,023	52,335	67,747	25,376	23,228	89,685
ADP-BeF$_x$	6,522	75,323	109,021	121,560	58,156	44,347	4,397	73,135
AMP-PNP	3,304	5,680	40,709	—	—	67,522	—	19,576
ADP	11,868	32,821	49,434	—	—	—	—	131,689

dc, double-capped; sc, single-capped.

Table S2. Cross-linking data for Rpn13

Cross-linked peptide	Protein 1	Protein 2	Type	AbsPos1	AbsPos2	ΔS	Id score	FDR
DREPKEPVALIETVR-SSKSGR-a5-b3	RPN2	RPN13	Interlink	911	82	0.37	43.18	0
DREPKEPVALIETVR-DKEIYNK-a5-b2	RPN2	RPN13	Interlink	911	119	0.93	36.38	0.029

AbsPos, absolute position of the cross-linked lysine residues within the UniProt or construct sequence; ΔS, delta score of the respective crosslink; Id score, weighted sum of different scores used to assess the quality of the composite second-stage mass spectrometry spectrum.