

Supporting Information

Schweitzer et al. 10.1073/pnas.1608050113

SI Materials and Methods

26S Proteasome Purification and Characterization. Fresh human blood from a healthy donor was collected under medical supervision in the presence of 1.8 mg/mL EDTA. The blood sample was washed three times with PBS containing 5 mM glucose (23). Cooled cells were lysed for 30 min with pure water containing ATP, MgCl₂, and β-mercaptoethanol in 8 vol of the initial cell pellet. After lysis the crude cell extract was filled up with water, glycerol, and Tris-HCl (pH 7.5–10) volumes of the initial cell pellet to a final concentration of 50 mM Tris-HCl pH 7.5, 10% (wt/vol) glycerol, 10 mM MgCl₂, 5 mM ATP, and 10 mM β-mercaptoethanol. The crude cell extract was centrifuged at 31,000 × *g* for 10 min to remove cell debris, and the supernatant was centrifuged again at 100,000 × *g* for 30 min to pellet cell membranes. Solid ammonium sulfate was added to the supernatant to 40% saturation, and the precipitated material was collected by centrifugation at 31,000 × *g* for 15 min. The pellet was resuspended in Buffer A [20 mM Tris-HCl, pH 7.5, at 21 °C, 10% (wt/vol) glycerol, 10 mM β-mercaptoethanol, 5 mM ATP, 10 mM MgCl₂] and centrifuged again at 31,000 × *g* for 5 min to remove insoluble precipitate. The 26S proteasome was then pelleted at 160,000 × *g* for 131 min. The supernatant was discarded and the pellet redissolved in Buffer A and pelleted again at 160,000 × *g* for 131 min. The supernatant was removed and the pellet was redissolved in a minimal volume of Buffer A and centrifuged at 20,000 × *g* for 1 min to remove insoluble protein. The sample was subjected to preparative sucrose density gradient centrifugation. Gradients were 20–40% (wt/vol) sucrose in Buffer B (20 mM Tris-HCl, pH 7.5, at 21 °C, 10 mM β-mercaptoethanol, 5 mM ATP, 10 mM MgCl₂; 10 mM creatine phosphate, 0.03 mg/mL creatine kinase) and centrifuged for 17 h at 208,000 × *g* in a Beckman SW41 rotor. Fractions with 26S proteasome activity were determined by hydrolysis of Suc-Leu-Leu-Val-Tyr-AMC, and 26S proteasome protein was determined by Coomassie blue staining of SDS/PAGE. To estimate the subunit abundance, the sample was subjected to mass spectrometry analysis and label-free quantification according to the iBAQ value (49, 50). Samples of ~0.5 mg/mL were quickly frozen for storage at –80 °C until use.

Data Acquisition. Data acquisition was performed essentially as described (51). In brief, the dataset was collected on a Titan Krios with a Falcon III camera using the FEI EPU software. Images were acquired at a pixel size of 1.35 Å at specimen level, a total dose of 45 electrons distributed over 50 frames, and a nominal defocus varying from 0.8 to 3 μm. Typically, the majority of particles on the micrographs were dc26S particles, but some sc26S as well as isolated CPs were additionally found.

Image Processing. In a first step, the acquired micrograph frames were translationally aligned and summed using an in-house implementation of the algorithm from (52). Both the aligned frame stacks and the summed images were saved for further use. In the next step, the summed images were used for contrast transfer function (CTF) estimation in CTFFIND3 (53). Only images with a CTF fit score above 0.05 and a defocus inside the range of 0.8–3.5 μm were retained for further analysis. This procedure resulted in a dataset of 40,211 images, subsequently subjected to automated particle localization implemented in the TOM package as described previously (13). Then reference-free 2D classification in RELION (44) was applied to filter out low-quality particles and to separate a total of 458,052 dc26S and

230,690 sc26S particles. These particles were extracted at a square size of 384 pixels at full size (pixel size: 1.35 Å) and at a reduced size of 256 pixels (pixel size: 2.03 Å). During all of the following processing steps the reduced sized particles were used except for the particle polishing and refinement.

The dc26S particles were aligned in RELION with applied C₂ symmetry. The result indicated an uneven angular distribution (Fig. S7). To decrease 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, which is a measure for the reliability of the (angular) class assignment of a particle. Evaluation on a subset of initially 183,000 particles showed only a minor decrease in resolution after data reduction. The remaining 267,660 particles were divided into two random subsets for refinement and particle polishing in RELION (54).

In a next step the broken C₂ symmetry was addressed as described (11). In this procedure, the two RPs of one dc26S particle were effectively treated as separate particles during image processing; the resulting 535,320 pseudosc26S particles were processed separately by focusing the analysis on a single RP and restricting the angular search range. First, the pseudoparticles were classified into 10 classes in RELION for 34 iterations using a soft-edged spherical mask on the RP region, keeping the previously assigned angles constant. Of the 10 resulting reconstructions 4 were in a well-defined s1-like state, comprising 431,141 RPs. These pseudosc26S particles were subjected to a second round of 3D classification, yielding only one well-defined reconstruction among the 10 classes, which contained over 80% of the particles (361,475 pseudosc26S particles). These particles were chosen for refinement, performed in RELION with a soft-edged mask containing one RP and half of the CP and a local angular search around the initial angles provided as prior values. The resolution of the postprocessed map with the mask used for refinement was 4.1 Å. The sc26S particles were processed similarly as the dc26S particles. After an initial alignment the number of particles was reduced as described above for the dc26S particles. The remaining 128,741 particles were refined and polished in RELION, followed by two runs of 3D classification with a global angular search. The classification procedure identified 67,466 sc26S particles in an s1-like state and 16,840 previously wrongly classified dc26S particles (dc*). The s1-like sc26S particles were refined with a full angular search using the reconstruction resulting from the final refinement of the dc26S particles as an initial reference, resulting in a resolution of 4.8 Å. The dc* particles were refined by applying C₂ symmetry, followed by a separation of the RPs into 33,680 pseudosc26S particles. A subsequent step of classification under the same conditions as for the dc26S particles identified 32,461 good s1-like pseudosc26S particles, which were then refined locally onto the final reconstruction of the dc26S particles using a soft-edged mask containing one RP and half of the CP.

The 393,936 pseudosc26S particles from dc26S and dc26S* particles and the 67,466 sc26S particles were merged and locally refined around the input angles provided as prior values with a soft-edged mask on one RP and the CP, leading to a refined reconstruction with an average resolution of 3.9 Å according to Fourier shell correlation. Local resolution determination was determined by B-soft (55).

Model Building. The precise voxel size of the single-particle reconstruction was determined by a cross-correlation analysis in the University of California at San Francisco Chimera (56), maximizing

the fit of the 2.6-Å resolution human CP crystal structure [Protein Data Bank (PDB) ID code 4R3O (57)] for different voxel sizes. This atomic model of the human constitutive 20S proteasome was also used as an initial model for the CP.

The initial model for the RP of the human 26S proteasome was built through comparative modeling using Modeler (58). The RP of the yeast 26S proteasome [PDB ID code 4CR2 (11)] was used as a template, and in the Modeller-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 (59)] 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 de novo structure prediction protocol of Rosetta (60). These missing segments were furnished based on a local fragment library of solved models within the Protein Data Bank. A Metropolis Monte Carlo integration of simulated annealing, guided by a knowledge-based scoring function, was used to exchange and place these fragments into partial models (61).

The RP of the obtained homology model was aligned with the density by superimposing the CP of the yeast 26S proteasome (PDB ID code 4CR2) with the human CP crystal structure (PDB ID code 4R3O) already docked into the density. The human CP and RP models were combined and then fitted into the density using MDFF (45, 62), 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 (46) and ana-

lyzed using the MDFF graphical user interface of VMD (47). We used NAMD (48) with the correction map (CMAP)-corrected CHARMM22 force field (63) 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 filtered to 7-Å resolution and a grid coupling of 0.3 followed by runs with the 3.9-Å resolution map, increasing the grid coupling up to 1. 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 element was checked through local cross-correlation calculation (64) implemented in VMD. The identified segments of the model that clearly deviated from the density were adapted to the density by combining Rosetta's FastRelax algorithm (65) with MDFF in an iterative way following a similar strategy as described (66). The FastRelax algorithm optimizes the model by applying a gradient minimization to all torsional degrees of freedom (ϕ , ψ , ω , and χ) followed by a simulated annealing rotamer search (61).

For further refinement of the atomic coordinates in reciprocal space, the EM map and the Fourier shell correlation (FSC) were converted into an MTZ file using the `ban_mrc_to_mtz.py` script (67). The atomic coordinates were then optimized against the Fourier coefficients using maximum-likelihood refinement in PHENIX (68). Differences between the refined model and the map were analyzed in COOT (69), 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.

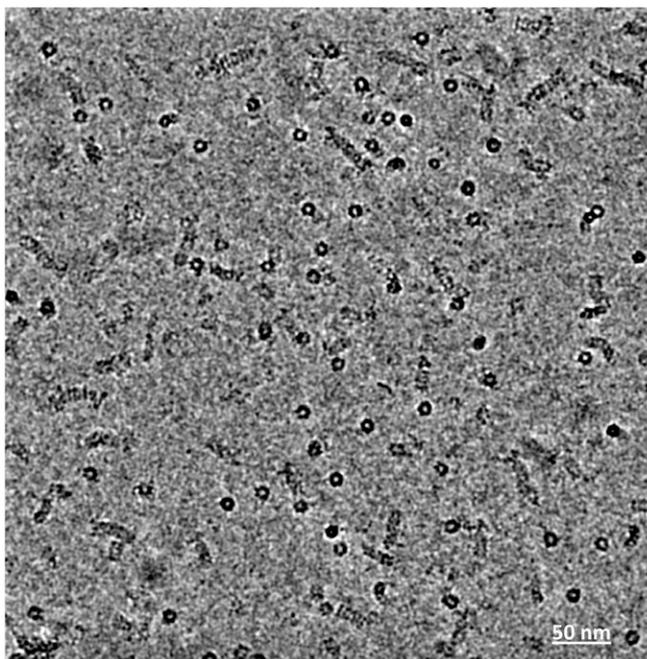


Fig. S1. Representative cryo-EM image of human 26S proteasome dataset. For better visibility the image has been low-pass-filtered to a 0.3-nm resolution. (Scale bar, 50 nm.)

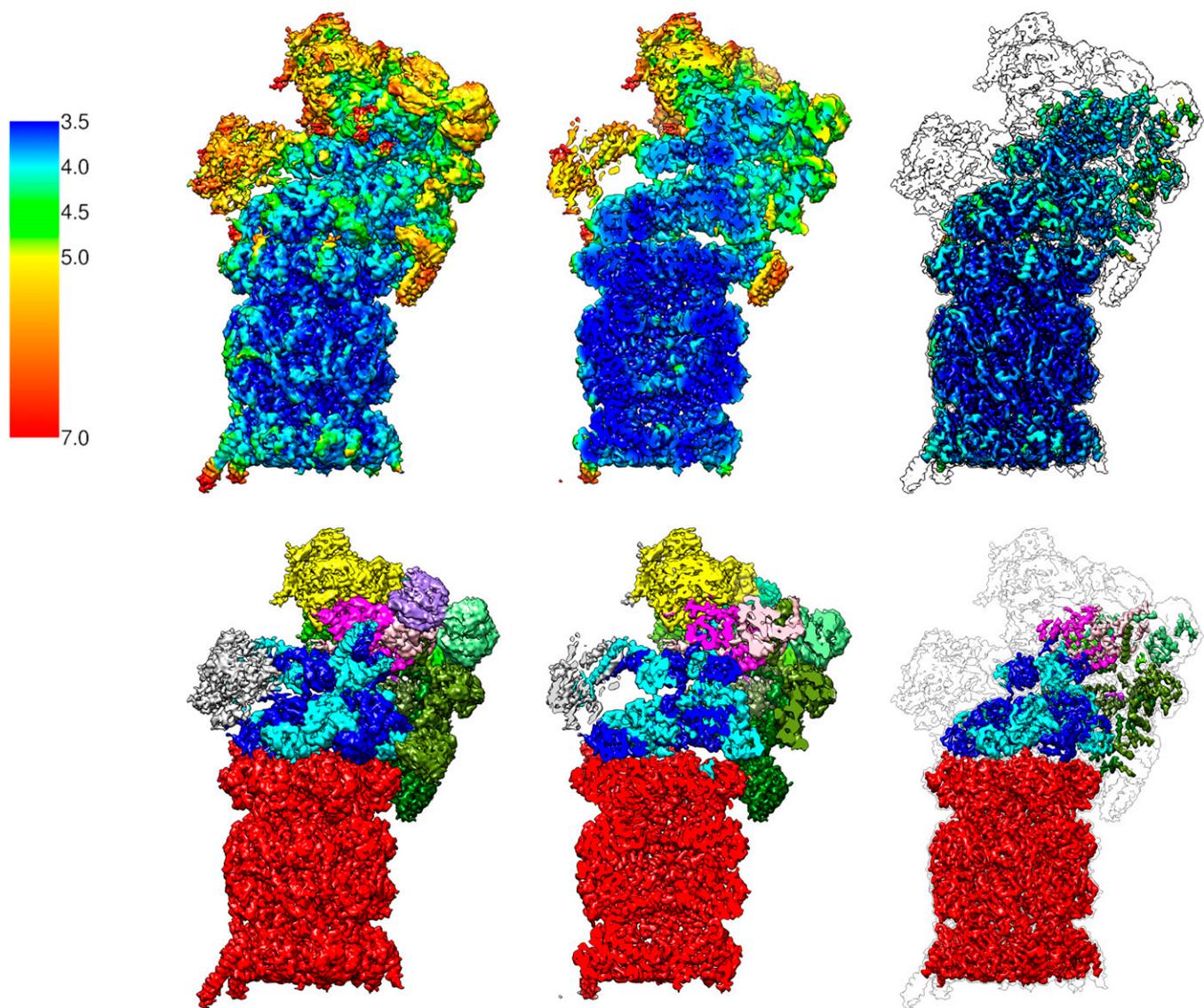


Fig. 54. Local resolution of EM map of human 26S proteasome. (*Top, Left*) The 26S proteasome isosurface rendering colored according to the local resolution (color bar: local resolution in Å). (*Top, Center*) Same as cut-open view. (*Top, Right*) Isosurface rendering at higher density value encapsulates only the highest resolution parts. The surroundings of the isosurface at lower density value are indicated. (*Bottom, Left*) Isosurface colored by subunit (red: CP; blue: AAA-ATPase heterohexamers; gray: Rpn1; yellow: Rpn2; green: Rpn3, -5, -6, -7, -9, -12; magenta: Rpn8, -11; purple: Rpn10). (*Bottom, Center*) Cut-open view. (*Bottom, Right*) Isosurface rendering at higher density value.

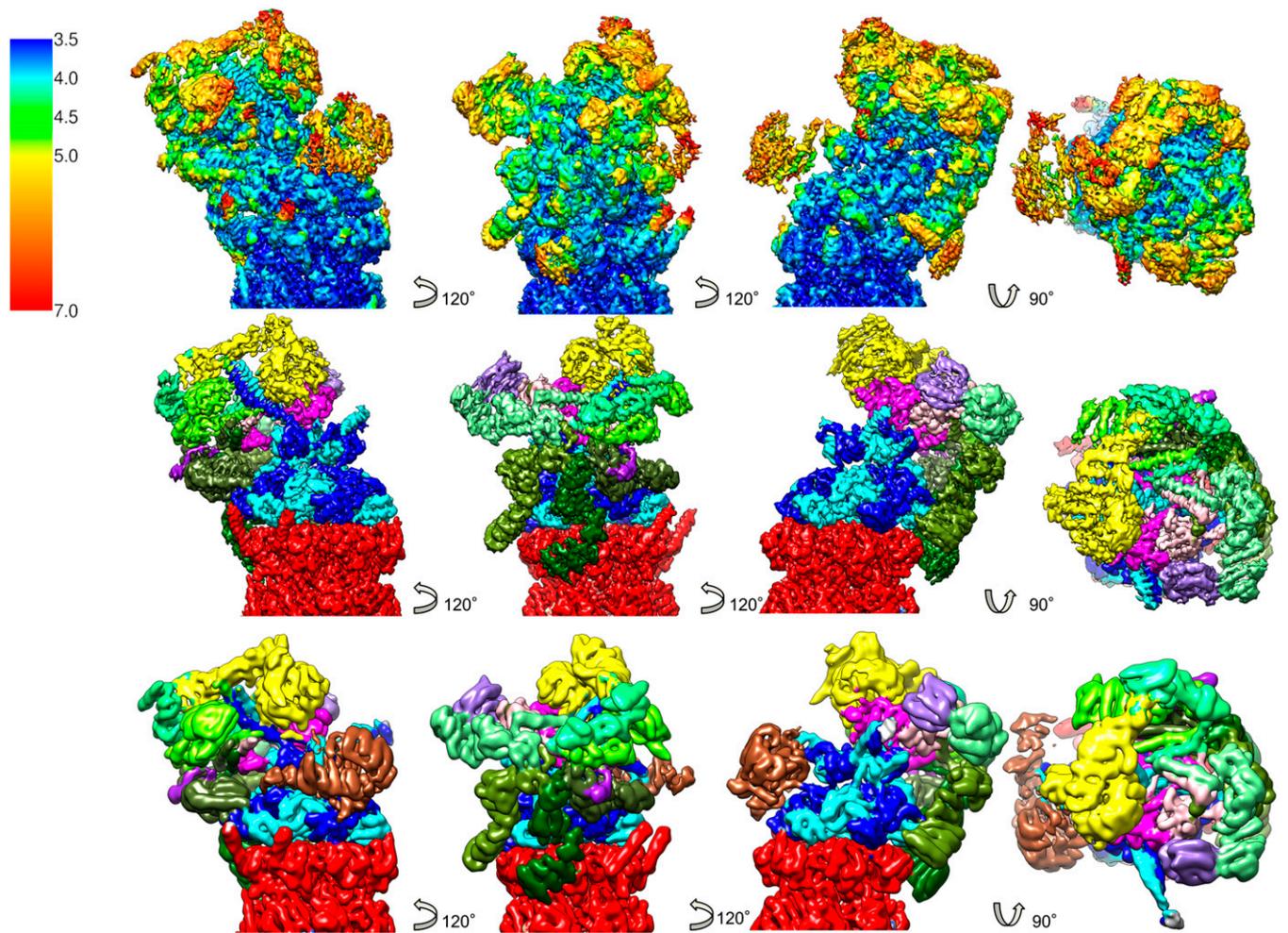


Fig. S5. Reconstruction at different resolutions. (Top) Reconstruction of the human 26S proteasome rendered at 3.9 Å resolution and colored according to local resolution. (Center) Same colored by subunits (see Fig. S4 for color code). (Bottom) Reconstruction rendered at 7.7 Å.

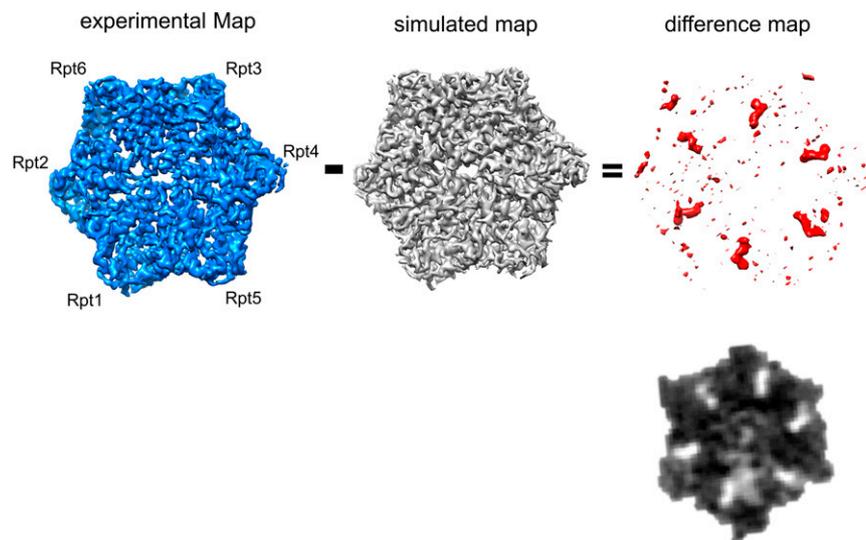


Fig. S6. Visualization of nucleotides in EM density. (Top) Subtraction of the polypeptide density (simulated from atomic model) from the EM data yields the density corresponding to the nucleotides shown as isosurface and gray values (Bottom).

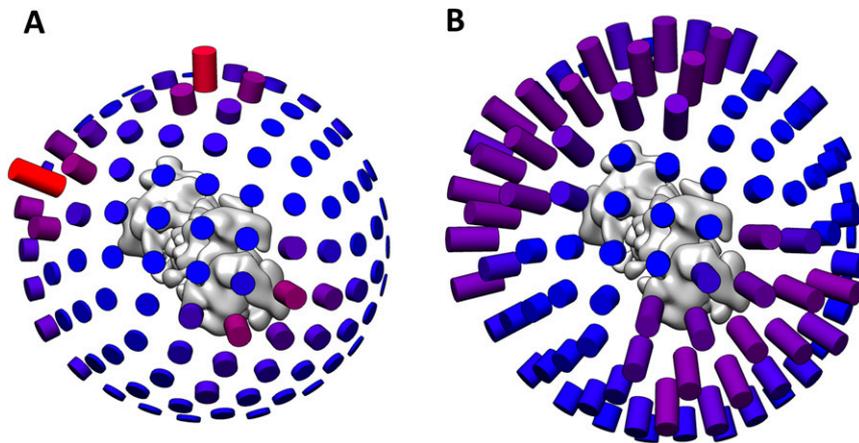


Fig. 57. Orientation distribution of 265 particles. (A) Initial angular distribution. (B) Angular distribution after reduction of particle numbers in highly populated orientation classes.