# Supplementary Figures, Tables, Methods, and Discussion

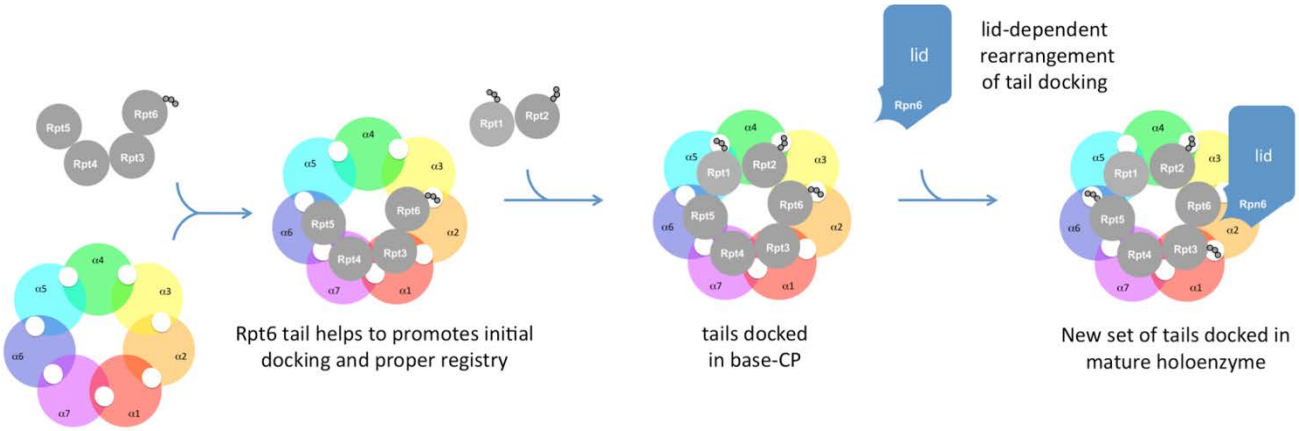
# for

# Reconfiguration of the proteasome during chaperone-mediated assembly

Soyeon Park[1,2,*], Xueming Li[3,*], Ho Min Kim3[,4,*], Chingakham Ranjit Singh[5], Geng Tian[1], Martin A. Hoyt[6], Scott Lovell[7], Kevin P. Battaile[8], Michal Zolkiewski[9], Philip Coffino[6], Jeroen Roelofs[5^], Yifan Cheng[3^], and Daniel Finley[1^]

[1]Dept. of Cell Biology, Harvard Medical School, 240 Longwood Ave., Boston, MA 02115
[2]MCD Biology, University of Colorado Boulder, Boulder, CO 80309
[3]The W.M. Keck Advanced Microscopy Laboratory, Department of Biochemistry and Biophysics, University of California San Francisco, 600 16th Street, San Francisco, CA 94158
[4]Present address: Graduate School of Medical Science and Engineering, Korea Advanced Institute of Science and Technology, Daejeon 305-701, Korea
[5]Division of Biology, Kansas State University, 338 Ackert Hall, Manhattan, KS 66506
[6]Department of Microbiology and Immunology, University of California San Francisco, 513 Parnassus Avenue, San Francisco, CA 94143
[7]Protein Structure Laboratory, Del Shankel Structural Biology Center, University of Kansas, Lawrence, KS 66047
[8]IMCA-CAT Hauptman-Woodward Medical Research Institute, 9700 South Cass Avenue, Building 435A, Argonne, Illinois, 60439
[9]Department of Biochemistry, Kansas State University, 176 Chalmers Hall, Manhattan Kansas 66506

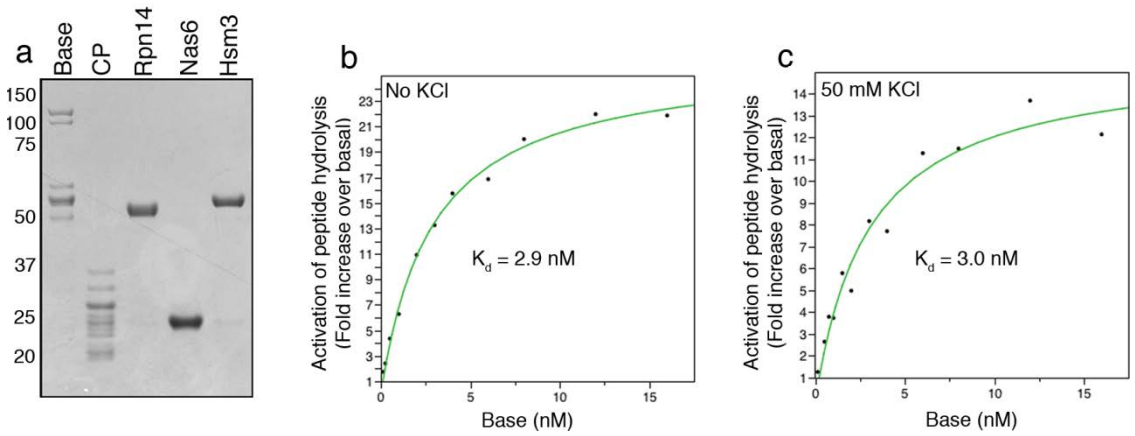
[*]These authors contributed equally to this work.
^Correspondence to jroelofs@ksu.edu, ycheng@ucsf.edu, daniel_finley@hms.harvard.edu



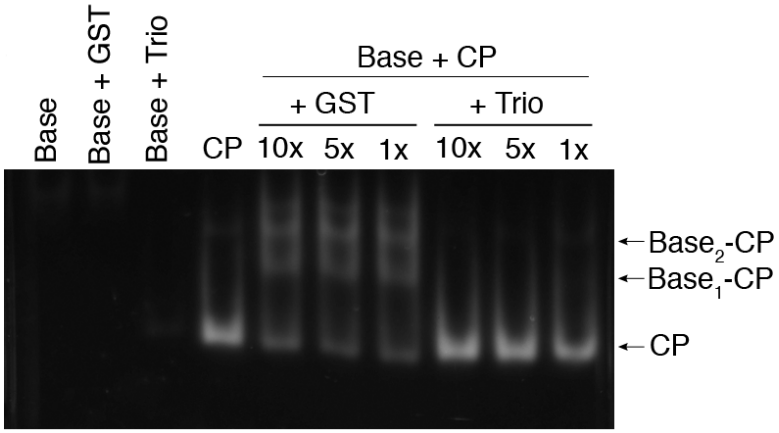

**Supplementary Figure 1. Proposed model of CP-dependent assembly of the proteasome regulatory particle.** While many details in proteasome assembly are still unclear, it is agreed that the formation of the AAA-ATPase ring by six paralogous Rpt proteins is a crucial step in the assembly pathway in yeast. Four chaperones – Nas2, Nas6, Hsm3, and Rpn14 – together with CP, play important roles in this process. The chaperones have been shown to regulate interactions of Rpt subunits with the CP and have also been suggested to ensure proper Rpt order. This diagram, however, focuses on the Rpt tails and their association with the CP. As reported previously, Rpt proteins first form pairs, and Rpt3-Rpt6 associate with Rpt4-Rpt5 early in assembly[17]. The tail of Rpt6 may promote the initiation of proper Rpt-CP docking, as this is the only Rpt tail found to have high specificity for a single α pocket (This does not imply that other tails are irrelevant to the pathway). Next, the incorporation of the Rpt1-Rpt2 module is proposed to result in formation of the base-CP complex (base subunits Rpn1, Rpn2 and Rpn13 have been omitted from the figure for clarity, but are also expected to associate in this step). Our EM structure of this base-CP complex shows the tails of Rpt1, Rpt2, and Rpt6 docked into specific α pockets of CP. While synthetic tail peptides of Rpt1 and Rpt2 lack specificity, their proximity to Rpt6 may promote proper register as they engage α pockets, according to the model.

Interestingly, ring formation may confer the capacity for ATP hydrolysis by the Rpt proteins[25]. While ATP does not seem to affect chaperone binding per se, it appears to prevent stable base-chaperone-CP complex formation (based on ATPγS data), and is thus likely to modulate assembly. Upon incorporation of lid, the base-CP interface appears to undergo dramatic rearrangements, as EM of the 26S proteasome[2] shows predominant docking of the Rpt2, Rpt3, and Rpt5 tails into CP-pockets. It is possible that these rearrangements reflect the influence of lid subunit Rpn6[2,3,23,26], which is positioned closely to Rpt6 and the CP subunit α2, suggesting it might play an anchoring role, perhaps with other lid subunits that project towards the CP, in the mature proteasome.

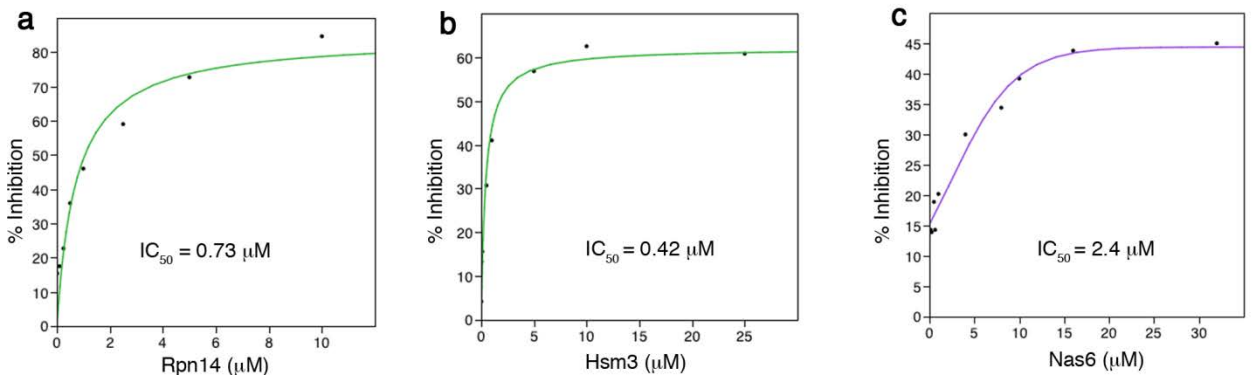
The model above is speculative and applies to the yeast proteasome. As mentioned in the main text, our model applies to CP-dependent assembly of the Rpt ring, but a parallel pathway of CP-independent assembly could also exist. In addition, our study examines the role of the chaperones in regulating the RP-CP interface, but is not meant to exclude additional roles of the chaperones[8,20]. Finally, we understand tail-pocket interactions visualized by cryoEM to represent dominant interactions, but they are unlikely to be the sole tail-pocket interactions in either the base-CP or holoenzyme complex. Both of these complexes are dynamic due to ongoing ATP hydrolysis, and consequently they may exist in a set of conformationally distinct states.



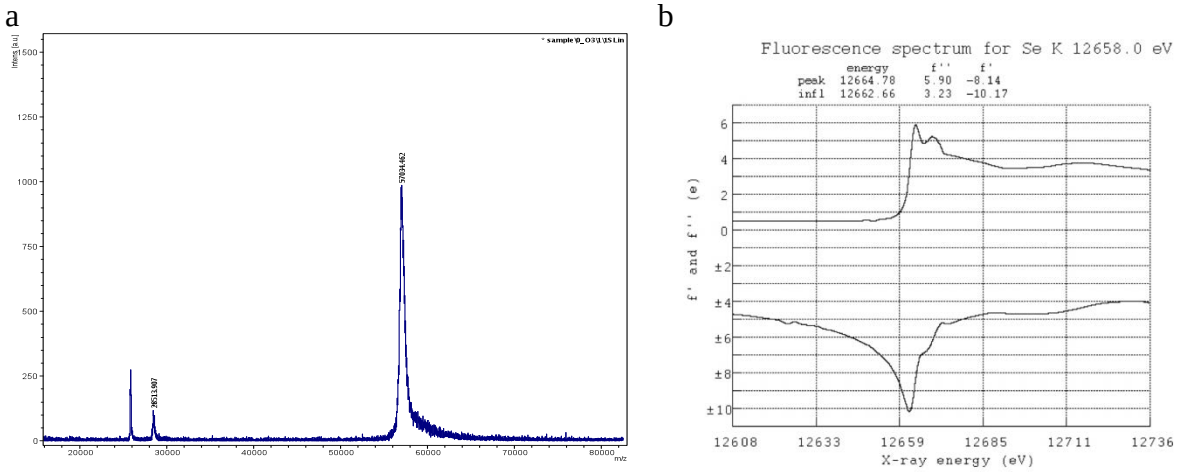

**Supplementary Figure 2. Purified base, CP and recombinant chaperones, and quantitative analysis of base-CP complex formation. a,** Affinity-purified base, CP, and recombinant chaperones (2 μg each) were resolved by 4-12% SDS-PAGE and stained with Coomassie blue. **b, c,** Increasing amounts of purified base were added to the purified CP (2 nM) to assess the formation of base-CP complexes. LLVY-AMC hydrolysis was measured real time for 30 min in 30 s intervals to obtain the rate of peptide cleavage at a given concentration of base. Assays were done in the presence of an ATP regenerating system and 2 mM ATP. The affinity of the base-CP complex was calculated using the velocities of LLVY-AMC hydrolysis: Fold activation = [Velocity of LLVY-AMC hydrolysis of a given base-CP association reaction]/[Velocity of the LLVY-AMC hydrolysis of CP alone]. The plots were fitted to Michaelis-Menten kinetics using JMP software (version 10). The high affinity of the base-CP complex is consistent with the slow loss of LLVY-AMC hydrolytic activity from base-CP after addition of CP trap (Fig. 2a).



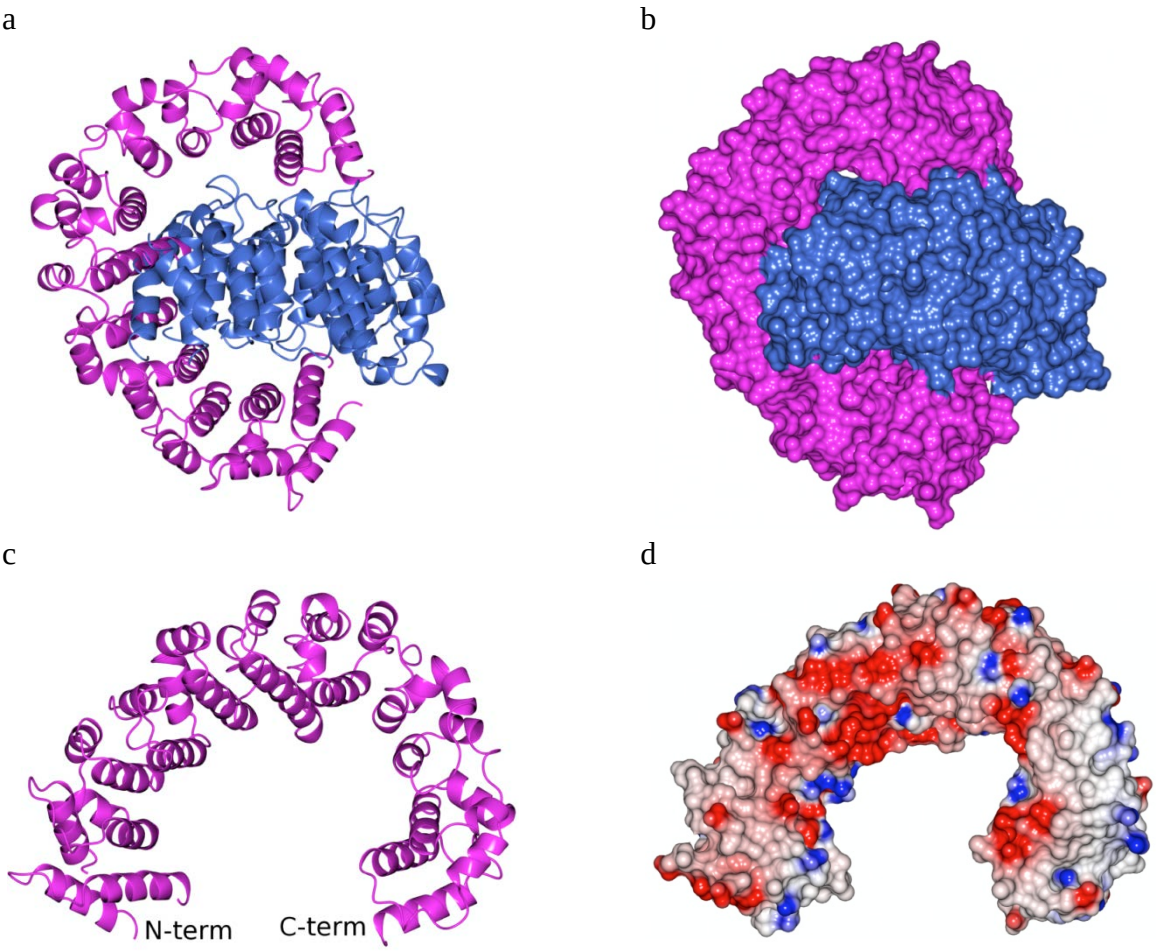

**Supplementary Figure 3. Chaperones inhibit base-CP assembly**
Base-CP complex formation was assessed in the presence of the chaperone trio, or non-specific protein, GST, at the indicated molar excess of base.




**Supplementary Figure 4. Quantitative analyses of chaperone effects on base-CP complex dissociation. a – c,** LLVY-AMC hydrolysis was measured in real time as increasing amounts of individual chaperones were added to the purified base (5nM) and CP (2nM). Percent inhibition = 100 x (1-[Velocity of the LLVY-AMC hydrolysis by a given Chaperone + base + CP sample]/[Velocity of the LLVY-AMC hydrolysis of base + CP]). Curve fitting was conducted as described in Supplementary Figures 2b and 2c.

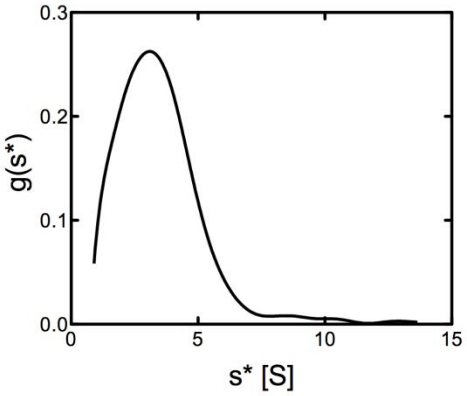



**Supplementary Figure 5. Incorporation of selenomethionine into Hsm3. a,** Unlabeled Hsm3 has a predicted molecular weight of 56597.96, and the incorporation of nine selenomethionine residues would result in a predicted increase in molecular weight of 422.1 daltons, giving an expected molecular weight of 57020.06. This value is very close to the observed molecular weight of 57034.462, indicating complete incorporation of the selenomethionine. **b,** X-ray fluorescence spectrum of a selenomethionine-labeled Hsm3 crystal.

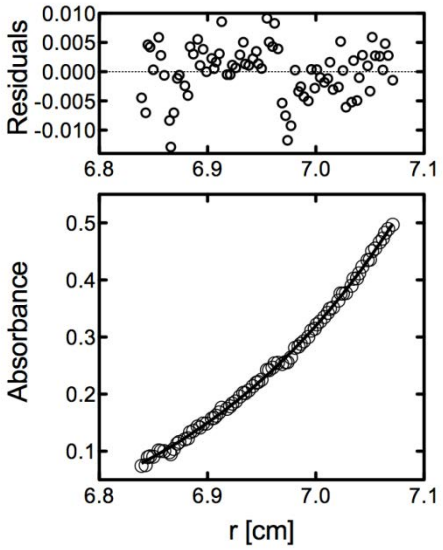



**Supplementary Figure 6. Structure of Hsm3.** **a,** Asymmetric unit of Hsm3 consisting of an NCS dimer showing chain A (magenta) and chain B (blue). **b,** Surface rendering of the figure at left. **c,** Chain A of the Hsm3 NCS dimer showing curvature of the subunit. **d,** Surface rendering of same view as the figure at left, but colored according to electrostatic potential. Acidic regions are displayed in red, basic regions in blue.

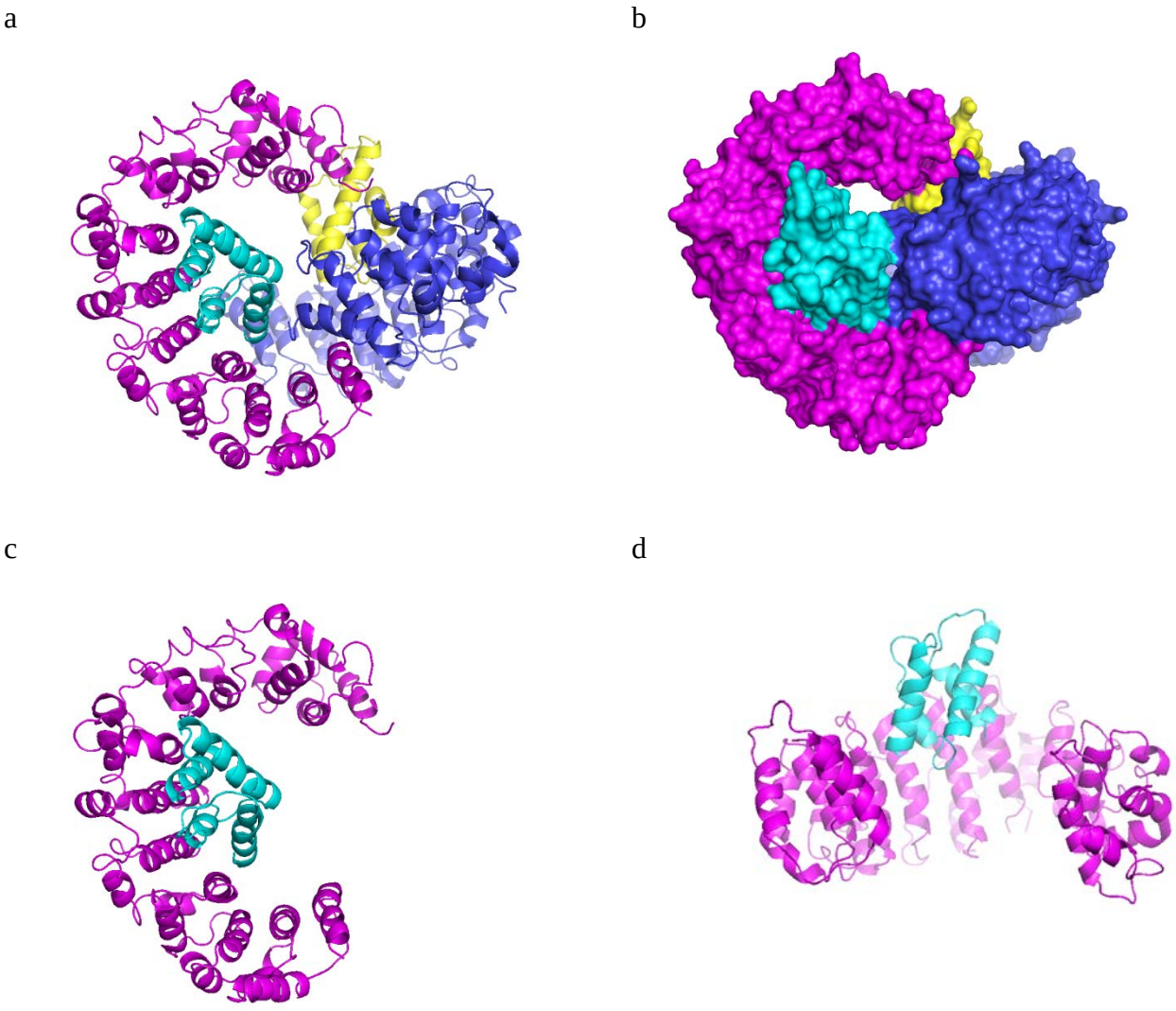
a




b

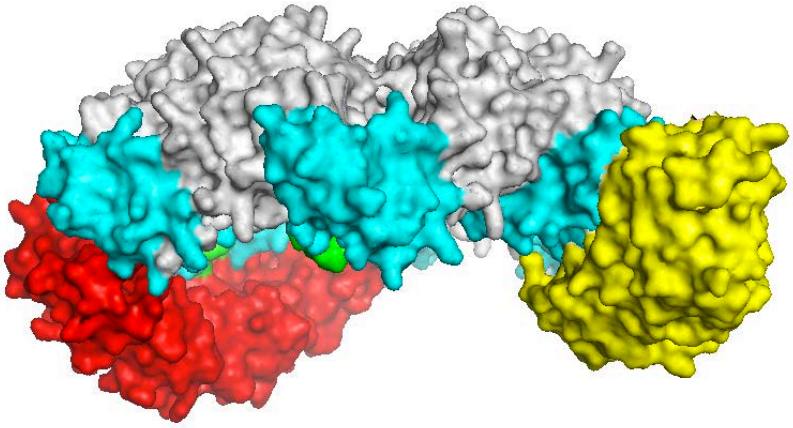



**Supplementary Figure 7. Hsm3 is a monomer in solution. a,** Analytical ultracentrifugation was used to determine the sedimentation velocity of Hsm3. Ultracentrifugation was performed at 58,000 rpm and 20°C. The absorption profiles of Hsm3 at 4.4 mg/ml were monitored at 290 nm. The apparent sedimentation coefficient distribution, g(s*), is shown as a function of the sedimentation coefficient in Svedbergs (S). **b,** Sedimentation equilibrium analysis of Hsm3. The Hsm3 absorption profile at 292 nm (lower panel, open circles) was measured after 48 h of centrifugation at 10,000 rpm and 4°C. A curve-fitting assuming a single component of molecular weight 54,200 is shown with the solid line. The upper panel shows the residuals ($A_{exp}$-$A_{model}$) for the single-species fit (open circles). The small and random residuals indicate that the experimental data are consistent with a single-species model. The empirically determined Hsm3 molecular weight of 54,200 is close to that predicted from its amino-acid composition, which indicates that Hsm3 is monomeric in solution even at high protein concentration (4 mg/ml).

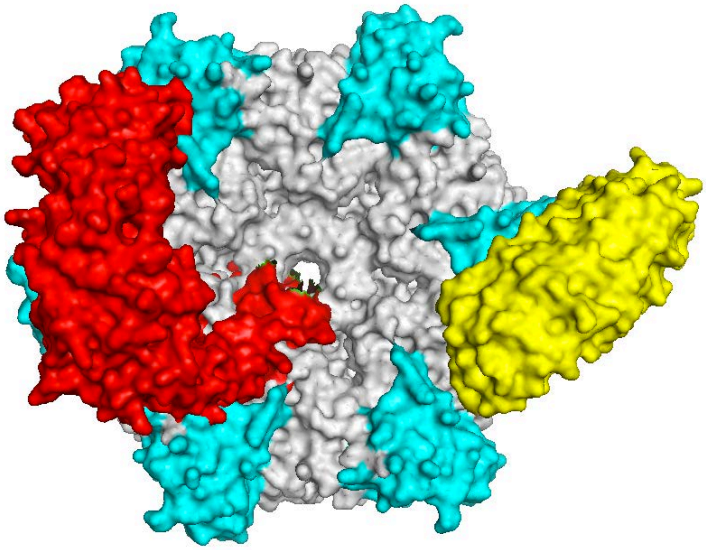



**Supplementary Figure 8. 5 Å resolution structure of the Hsm3-Rpt1 C domain. a,** The asymmetric unit of Hsm3-Rpt1 C-domain consists of a dimer. Shown are two chains of Hsm3, chain A (magenta) and chain B (blue), and two chains of Rpt1 C-domain, chain C (Cyan) and chain D (yellow). **b,** Surface rendering of the figure at left. **c,** Chain A of the Hsm3 with Chain C of Rpt1 C-domain. Hsm3 in solution is a monomer, as assessed in Supplementary Figure 7. Based on binding studies using truncated forms of Hsm3, we determined that the interaction as represented here by Chain A and C is the physiologically relevant interaction (data not shown). The interaction between chain A and chain D as seen in the top figures is most likely due to crystal packing. **d,** Rotated view of structure shown at left.

a

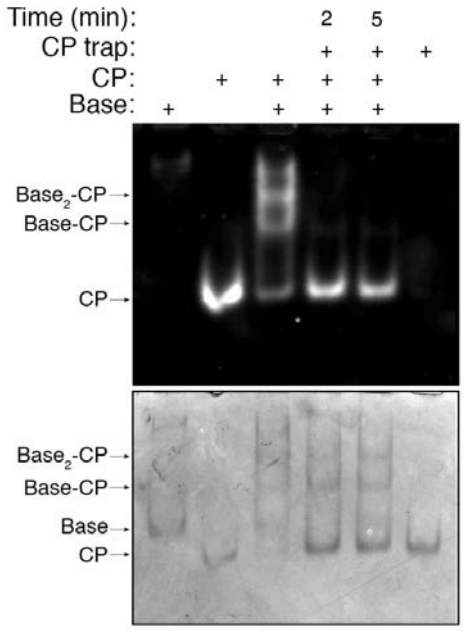


b

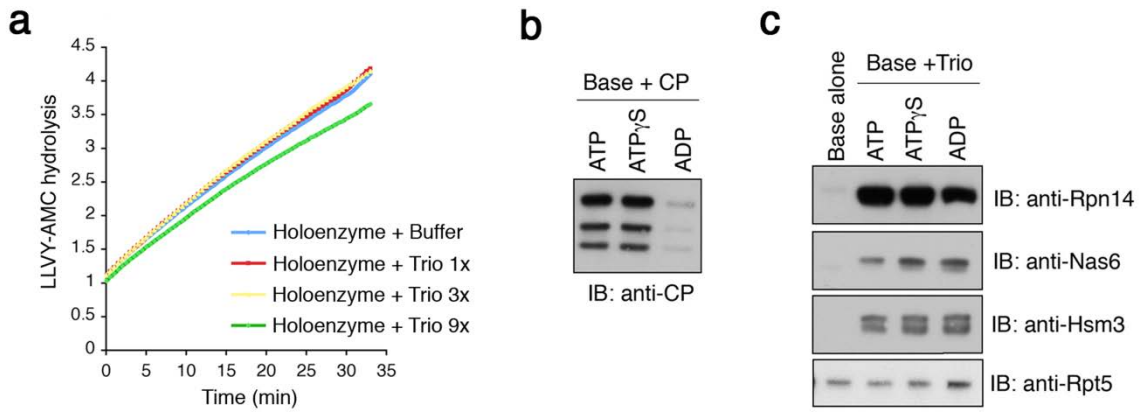


**Supplementary Figure 9. Alternative modeling of steric clashes between RP chaperones and the CP.** The modeling represented in Figure 1d suggests steric clashes between Nas6 and the CP as well as Hsm3 and the CP. To generate the figure above, we employed an alternative modeling approach, which leads to the same conclusion as that of Figure 1d, further supporting the steric clash hypothesis. Since the structure of the Rpt ring has not been determined, an important step in modeling is to build a theoretical Rpt ring model. In Figure 1d, this is achieved using the crystal structure of monomers of the related ATPase p97. These crystal structure data are fit into the EM map of the yeast proteasome holoenzyme[2]. In contrast, in the figure above we dock Nas6 and Hsm3 into the PAN hexamer model[27]. See supplementary methods for details of the modeling. **a,** side view where CP would bind from below. **b,** facing the surface of the ATPase that would interact with the CP. Red = Hsm3; Yellow = Nas6; gray = PAN-ATPase Ring; Cyan = PAN-C-domains.

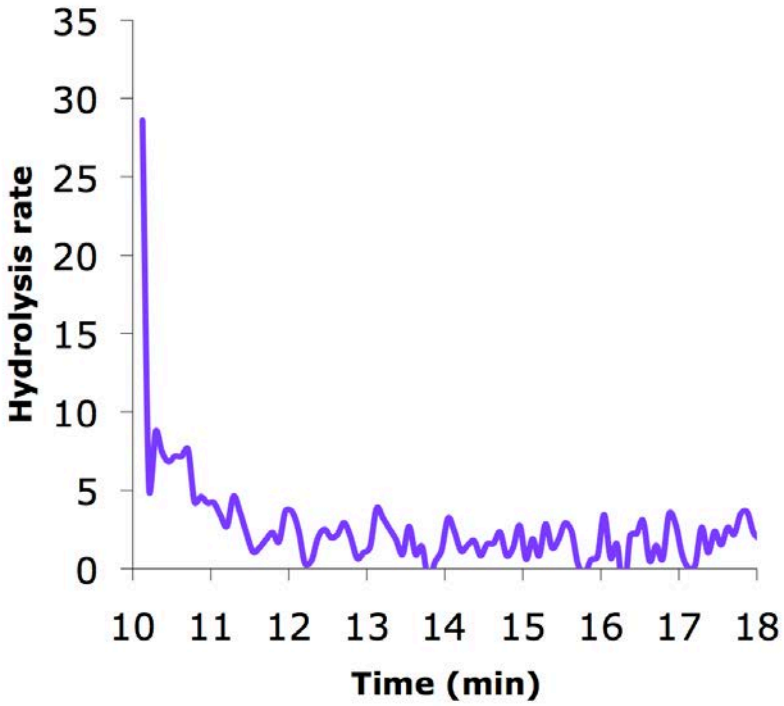
.



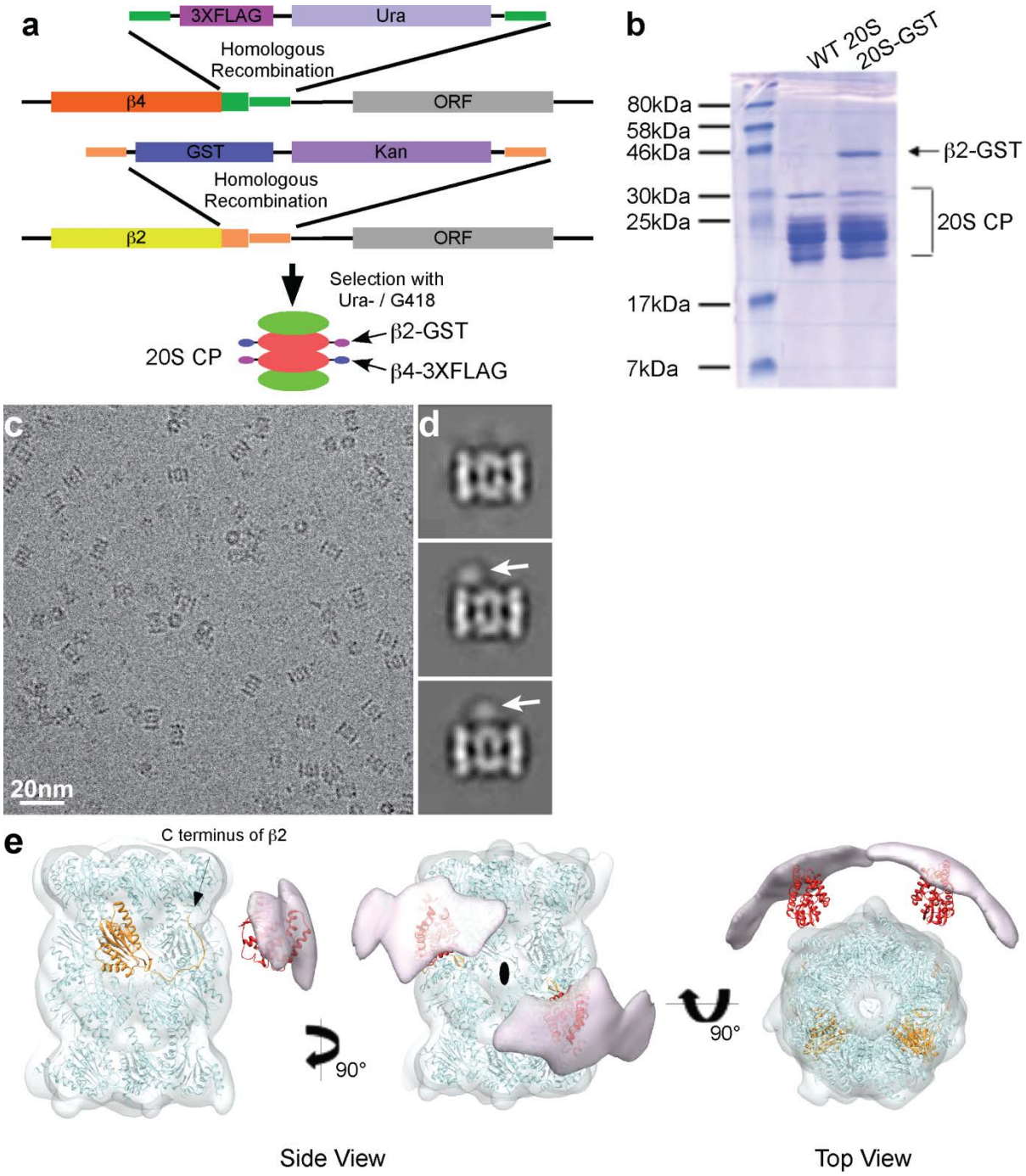

**Supplementary Figure 10. Dynamic and reversible association of base-CP complexes *in vitro.*** Purified base (200 nM) and CP (100 nM) were mixed for 5 min, then further incubated with CP trap at 3-fold molar excess for indicated time periods. Samples were resolved by 3.5% native PAGE, and subject to LLVY-AMC analysis in the presence of 0.02% SDS. The gels were stained with Commassie blue. Note that lanes 4 and 5 show the formation of new base-CP trap complexes lacking LLVY-AMC hydrolytic activities. Lane 6 confirms the lack of activity of epoxomicin-treated CP (CP trap).

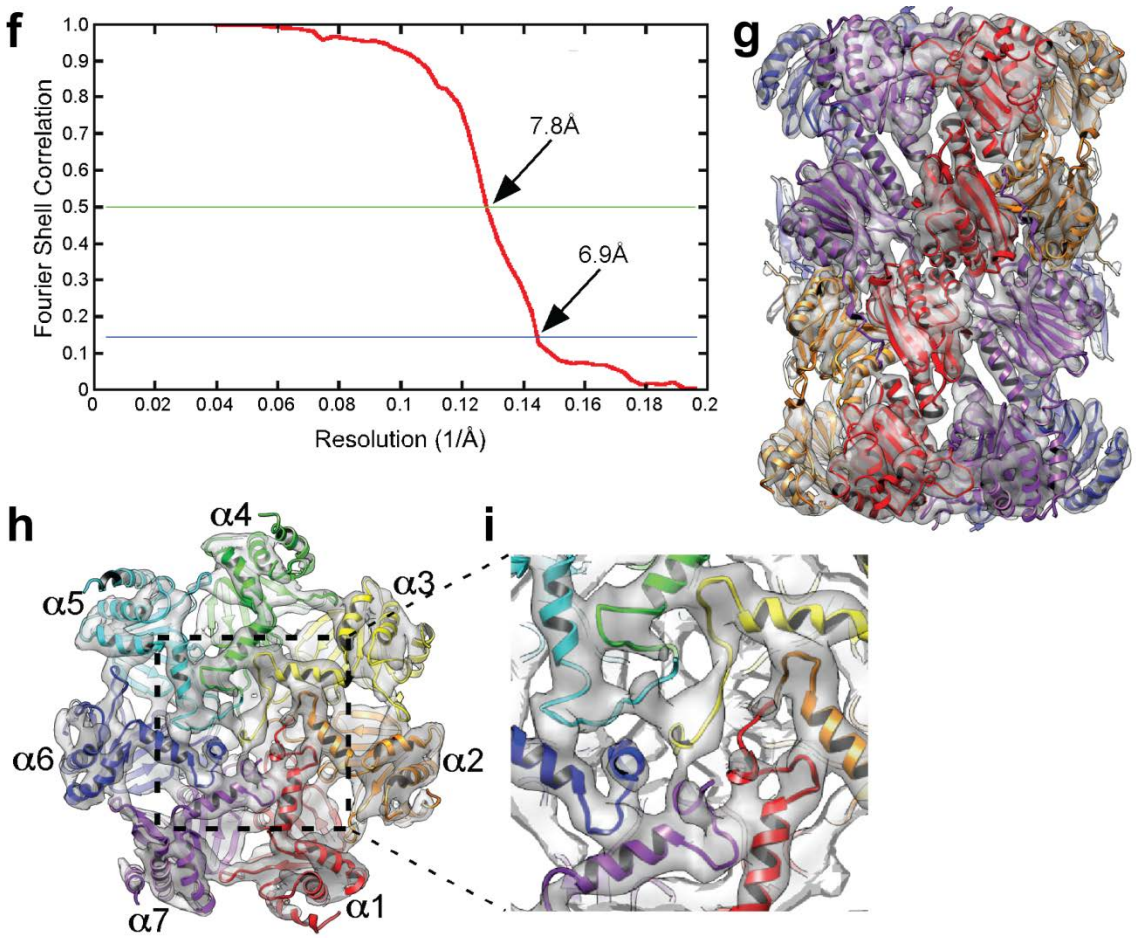



**Supplementary Figure 11. Requirements of RP-CP and base-CP for complex formation and stability. a,** Real-time assay of chaperone effects on the stability of holoenzymes. Affinity-purified holoenzymes (0.5 nM) via Protein A appended N-terminally to endogenous Rpt1 were mixed with chaperone trio at the indicated molar excess in the presence of 2 mM ATP and ATP regeneration system. LLVY-AMC hydrolysis (r.f.u.) was measured every 30s. **b,** Purified base (80 nM) was immobilized on IgG resin via a Protein A tag, then incubated with CP (approximately 15 nM) in the presence of indicated nucleotides (2 mM ATP, 5 mM ADP, 0.5 mM ATPγS) for 20 min at room temperature. Bound proteins were then eluted with TEV protease and resolved by SDS-PAGE and immunoblotting. Note that similar results were obtained when all nucleotides were each present at 2 mM. **c,** Nucleotide effects on chaperone association with base. Immobilized base (approximately 80 nM) was incubated with 2-fold molar excess of three recombinant chaperones, Rpn14, Nas6, and Hsm3, for 20 min in distinct nucleotides as in **b** at room temperature. Bound proteins were then eluted with TEV protease, and subject to SDS-PAGE followed by immunoblotting. Although this figure shows that free base is competent to bind chaperones in the presence of ADP, we could not identify a stable base-CP complex under these conditions, even in the absence of chaperones (Supplementary Figure 11b and data not shown). Consequently, possible ADP effects have not been studied.

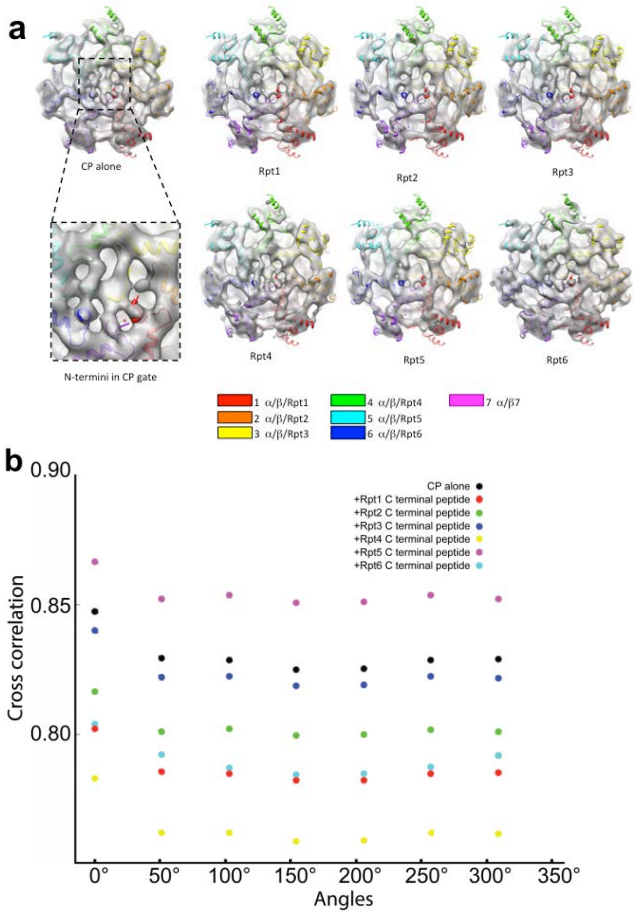



**Supplementary Figure 12. ATP activates chaperone-mediated disassembly of the base-CP complex.** This figure is a plot of reaction rates of LLVY-AMC hydrolysis, derived from the data of Figure 2d, and spanning time points from 10 to 18 min. The plot illustrates rapid achievement of a new steady-state hydrolytic rate in the chaperone-containing sample when ATP is added. Assembly of the base and CP was monitored in the presence of 0.1 mM ATPγS for the first 5 min. At 5 min, the chaperone trio (containing 0.1 mM ATPγS) was added to the reaction. At 10 min, buffer containing either ATPγS or ATP + ATPγS was added to the reaction. Final nucleotide concentrations were either 0.1 mM ATPγS or 10 mM ATP + 0.1 mM ATPγS. LLVY-AMC hydrolysis was measured every 5 seconds during the interval shown. The figure shows that a new steady-state hydrolytic rate is achieved within approximately 1.5 minutes. When the chaperone trio is added to preformed base-CP complex in the presence of ATP, complex disassembly, as indicated by decay of the rate of LLVY-AMC hydrolysis, returns to baseline far more slowly (Figure 2a, right). Therefore, preloading the base-CP complex with chaperones results in very efficient complex dissociation. In the continuous presence of ATP, association of the chaperones with the base-CP may be rate-limiting for base-CP disassociation. We note that the high apparent reaction rate observed in the first few seconds following ATP addition is artifactual, and represents a transient inaccuracy in the reading of the fluorimeter owing to the mixing of solutions in the ongoing reaction (see for example Figure 2d). This perturbation is quickly exhausted.

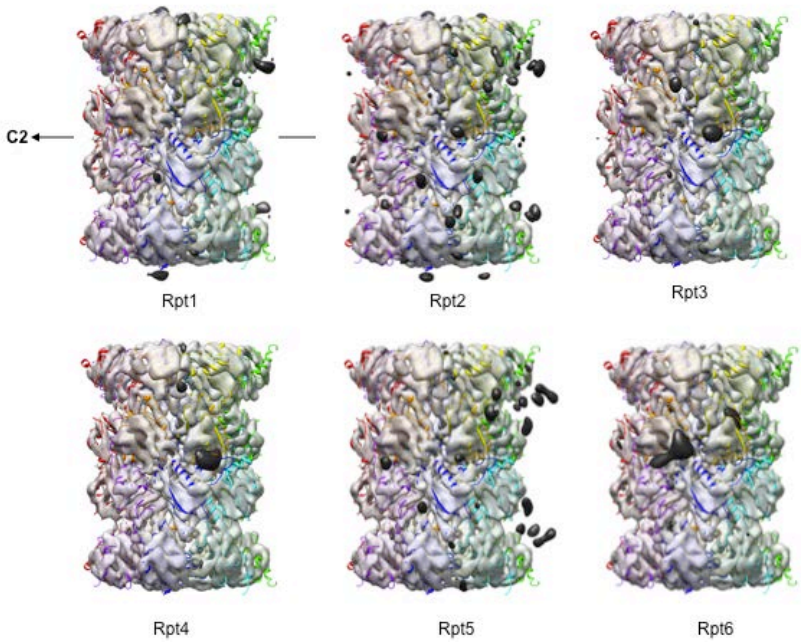

a
3XFLAG
Ura
Homologous Recombination
β4
ORF
GST
Kan
Homologous Recombination
β2
ORF
Selection with Ura- / G418
20S CP
β2-GST
β4-3XFLAG
b
WT 20S
20S-GST
80kDa
58kDa
46kDa
30kDa
25kDa
17kDa
7kDa
β2-GST
20S CP
c
20nm
d
e
C terminus of β2
90°
90°
Side View
Top View



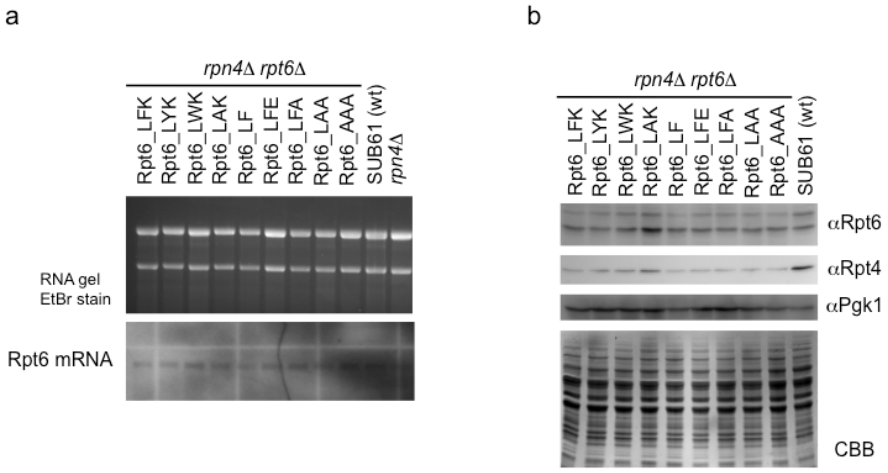

**Supplementary Figure 13. 3D reconstitution of GST-tagged *S. Cerevisiae* CP to a resolution of 6.9 Å. a,** a GST-tag was added to the C-terminus of β2 to break the pseudo 7-fold symmetry, and a 3X FLAG-tag was added to the C-terminus of β4 subunit for affinity purification. **b,** SDS-PAGE of the wildtype and GST-tagged 20S CP. The lane at left is a ladder of protein standards. The β2 subunit with a GST-tag is shifted to a higher molecular weight. **c,** A raw micrograph of GST-tagged free CP embedded in vitreous ice. **d**, 2D Class averages of the CP. Densities indicated by arrows correspond to the GST-tag. **e,** Three different views of the density map with a relatively low isosurface treshold. Densities corresponding to the GST-tags are visible. **f,** A Fourier Shell Correlation (FSC) curve of the final 3D reconstitution of 20S CP with a 2-fold symmetry. Resolutions corresponding to FSC = 0.5 and 0.143 are marked. **g,** Side view of the 3D reconstruction of the CP with the atomic structure docked. A negative temperature factor of -600 Å2 was used to sharpen the map. **h,** Top view of the 3D-reconstructed 20S CP with the atomic structure docked. A negative temperature factor of -600 Å was used to sharpen the map. **i,** An enlarged view of the gate region showing an asymmetric closed-gate conformation.



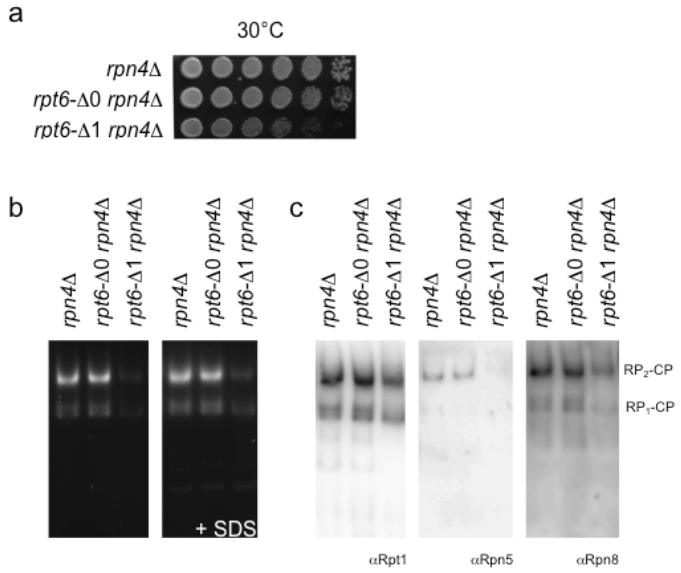

**Supplementary Figure 14. Asymmetrical features of 3D reconstructions. a,** Top views of 3D reconstructions of CP alone and CP-peptide complexes filtered to a resolution of 8 Å. The asymmetric features, especially of the N-termini of α-subunits in the gate region, were used to confirm the orientation of the α ring. **b**, Cross-correlation coefficients calculated from all the experimental maps with the map calculated from the atomic structure filtered to the same resolution. The calculated map was docked in seven different orientations. In all maps, the correct dockings (at 0°) always gave a higher correlation coefficient.



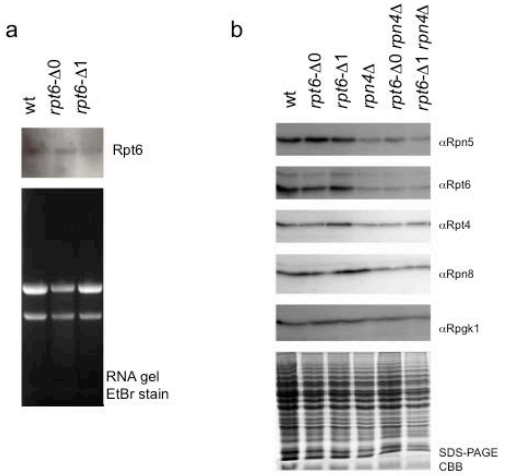

**Supplementary Figure 15. Difference maps calculated between CP-peptide complex and CP alone.** The entire difference maps between each CP-peptide complex and CP alone are shown.



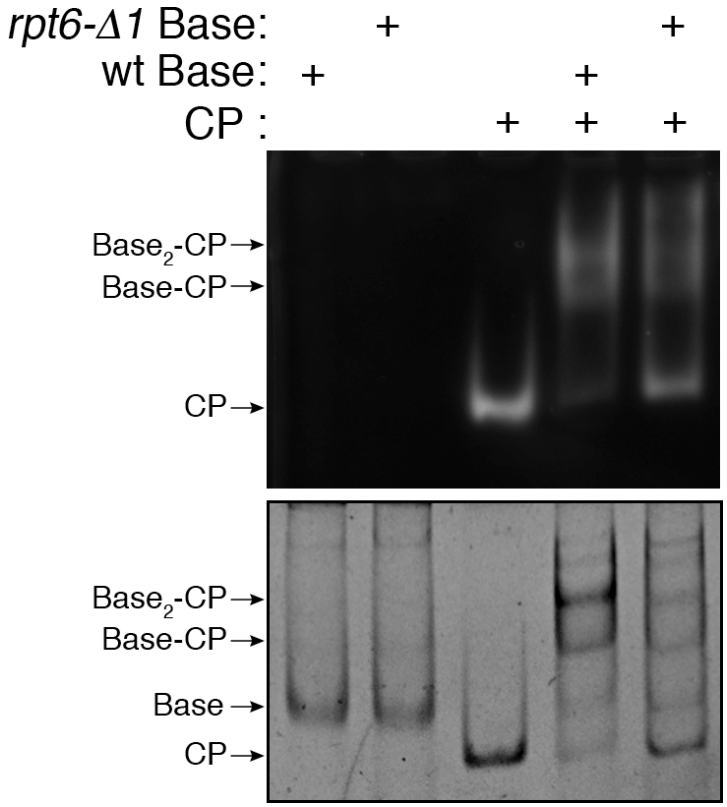

**Supplementary Figure 16. mRNA and Rpt6 protein levels.** In these strains, Rpt6 is expressed from its own promoter, on a low-copy (*CEN*) plasmid. **a,** Total RNA was isolated from yeast strains used in Figure 4. Levels of *RPT6* mRNA were determined by Northern blotting using the DIG detection system (Roche) with a single stranded RNA probe complementary to the *RPT6* mRNA. **b,** Total lysates of indicated strains were resolved by SDS-PAGE and immunoblotted for indicated proteins. Lower panel shows Coomassie Brilliant Blue staining of the SDS-PAGE gel.



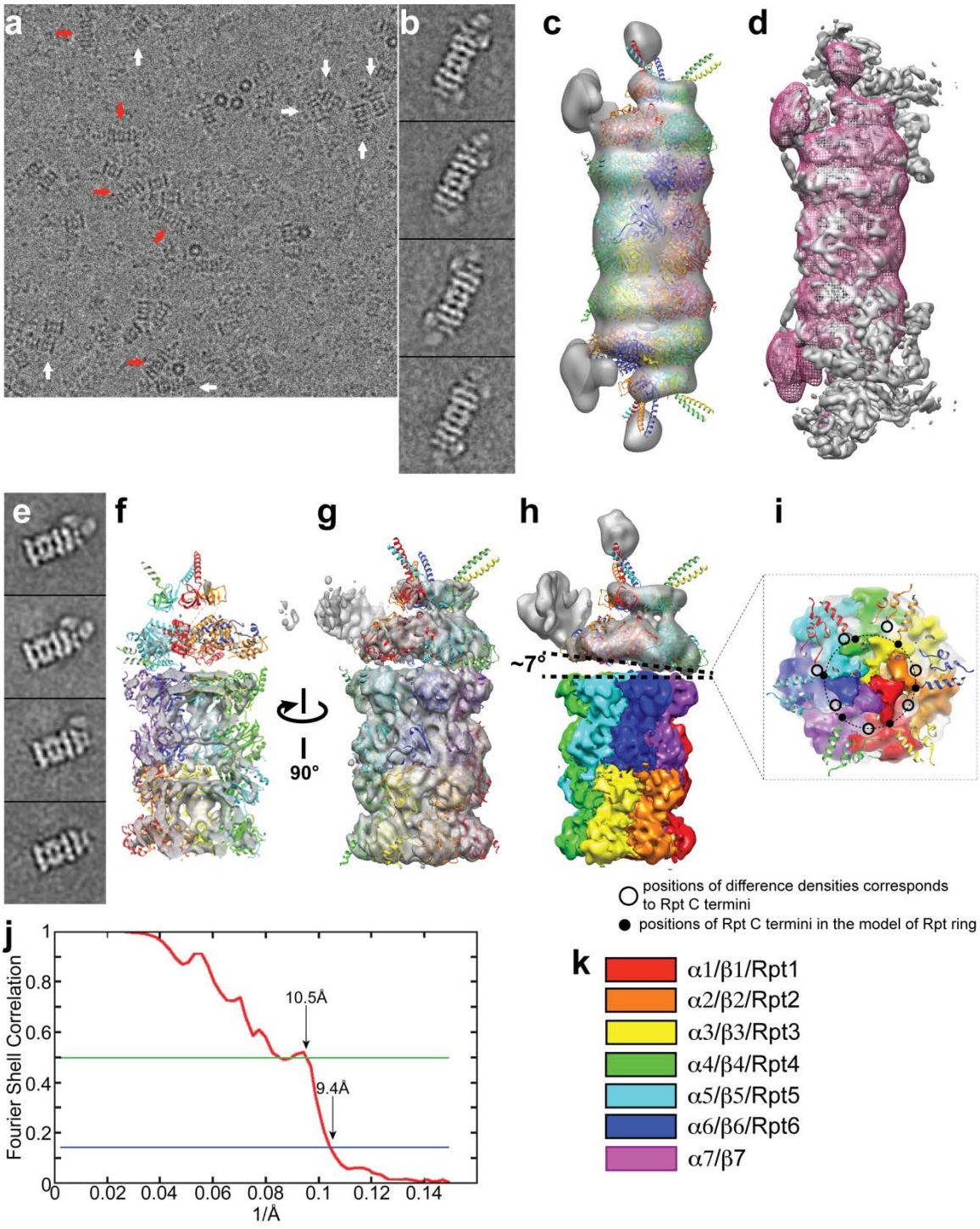

**Supplementary Figure 17. Phenotypic analysis of *rpt6–Δ0* and r*pt6–Δ1* mutants.** In these strains, the chromosomal *RPT6* locus is mutated. **a,** Dilution assay similar to Figure 4a. *rpt6–Δ1* contains a deletion of the codon of the *RPT6* ORF, followed by an *ADH1* terminator and a CloneNAT selection cassette. *rpt6–Δ0* is identical to *rpt6–Δ1* only has no deletion in the ORF. **b,** Native gel of total cell lysates stained for LLVY-AMC hydrolytic activity in absence (left) or presence (right) of 0.02% SDS. **c,** Immunoblotting of gel from (b) for indicated proteins.




**Supplementary Figure 18. Rpt6 mRNA and Protein levels for strains with modifications at *RPT6* locus.** Strains as in Supplementary Figure 17. **a,** Total RNA was isolated from yeast strains indicated strains. Levels of *RPT6* mRNA were determined by Northern blotting using the DIG detection system (Roche) with a single stranded RNA probe complementary to the *RPT6* mRNA. **b,** Total lysate of indicated strains were resolved by SDS-PAGE and immunoblotted for indicated proteins. Lowest panel shows Coomassie Brilliant Blue staining of the SDS-PAGE gel.

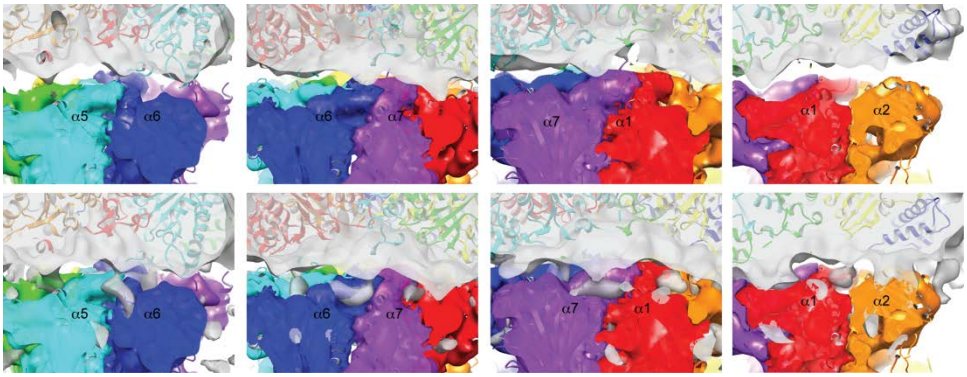



**Supplementary Figure 19. *rpt6-Δ1* base is deficient in associating with CP.** Affinity-purified base (180 nM) from wild-type or *rpt6-Δ1* strain was reconstituted with CP (90 nM) for 5 min at room temperature. Samples were analyzed via 3.5% native gel followed by in-gel peptidase assay with LLVY-AMC in the presence of 0.02% SDS. The gel was then stained with Coomassie blue.




**Supplementary Figure 20. Single particle cryoEM 3D reconstruction of base-CP complex. a,** Raw cryoEM image of frozen hydrated base-CP complex. Red arrows mark representative doubly-capped particles, and white arrows mark representative singly-capped particles. Most of the complexes are singly-capped. **b,** Representative 2D class averages of doubly-capped particles. **c,** A 3D reconstruction of doubly-capped base-CP complex at a resolution of ~24Å calculated from a total 1755 particles. A 2-fold

symmetry was applied to the 3D reconstruction. **d,** The density map of the doubly-capped base-CP complex (red mesh) is overlaid with the density map of 26S holoenzyme. The resolution of the doubly-capped base-CP complex is relatively low, preventing direct subunit assignments based on the density map. However, the match between the base-CP complex and holoenzyme provides unambiguous assignments of all subunits in the reconstruction of the base-CP complex. **e,** Representative 2D class averages of singly-capped base-CP complex. **f,** The density map of base-CP complex at a higher threshold, showing that the secondary structural features of CP are resolved. **g,** 3D reconstruction of the singly-capped base-CP complex without applying any symmetry. Comparison with the doubly-capped base-CP complex provides unambiguous assignment of all subunits in the 3D reconstruction. **h,** A hybrid 3D reconstruction, in which the CP (colored according to subunits) is from the 3D reconstruction of the base-CP complex and base (colored grey) is from a separated 3D reconstruction refined and reconstructed independently. When base is reconstructed alone, the structural features are more complete than when base is reconstructed together with CP. This indicates that base is mobile when bound to the CP, which is consistent with partial binding of the Rpt C-termini to the α pockets of the CP. The orientation of base is also tilted, similarly to holoenzyme. **i,** The interface between base and CP, where the location of the pockets and the position of the Rpt C-termini are marked. **j,** A Fourier shell correlation curve of the base-CP complex reconstruction. The resolution is close to 10 Å. **k,** Color code used in the figure.




**Supplementary Figure 21. Weak binding of Rpt3, Rpt4, and Rpt5 tails in the base-CP complex.** Using the same contour level as in Figure 5b, four of the α-ring pockets are empty (upper row). However, densities from the base are visible in these pockets with a reduced contour level (lower row), indicating possible weaker binding of Rpt C termini to these pockets.

# Supplementary Tables

**Supplementary Table 1.** Crystallographic data for Hsm3 (PDB ID: 4FP7) and Hsm3-Rpt1 C-domain complex (PDB ID: 4JPO).

| | Hsm3 (Se-Met) | Hsm3 (Native) | Hsm3-Rpt1 |
|---|---|---|---|
| **Data Collection** | | | |
| Unit-cell parameters (Å, °) | *a*=84.32, *b*=94.47, *c*=129.93 | *a*=85.15, *b*=94.47, *c*=129.93 | *a*=185.30, c=357.38 |
| Space group | $P2_12_12_1$ | $P2_12_12_1$ | $P6_522$ |
| Resolution (Å)[1] | 129.59-2.20 (2.32-2.20) | 129.93-2.20 (2.32-2.20) | 48.80-5.00 (5.59-5.00) |
| Wavelength (Å) | 0.97898 | 1.00000 | 1.00000 |
| Temperature (K) | 100 | 100 | 100 |
| Observed reflections | 708,088 | 357,957 | 316,697 |
| Unique reflections | 53,402 | 53,064 | 16,422 |
| $\langle I/\sigma(I)\rangle$[1] | 15.9 (3.9) | 15.1 (3.0) | 20.1 (3.6) |
| Completeness (%)[1] | 100 (100) | 98.6 (97.8) | 99.9 (100) |
| Multiplicity[1] | 13.3 (13.3) | 6.7 (6.9) | 19.3 (19.2) |
| $R_{merge}$ (%)[1, 2] | 12.4 (80.3) | 9.7 (65.7) | 8.3 (97.1) |
| $R_{meas}$ (%)[1, 4] | 14.3 (87.2) | 10.5 (71.0) | 8.5 (99.7) |
| $R_{pim}$ (%)[1, 4] | 3.9 (23.9) | 4.0 (26.9) | 1.9 (22.6) |
| **Refinement** | | | |
| Resolution (Å) | | 44.39-2.20 | 46.87-5.00 |
| Reflections (working/test) | | 47,900 / 5,101 | 15,526/829 |
| $R_{factor}$ / $R_{free}$ (%)[3] | | 18.8 / 23.9 | 25.3/27.3 |
| No. of atoms (protein/ /water) | | 7,025 / 261 | 8,391 / 0 |
| **Model Quality** | | | |
| R.m.s deviations | | | |
| Bond lengths (Å) | | 0.010 | 0.009 |
| Bond angles (°) | | 1.125 | 1.030 |
| Average *B* factor (Å$^2$) | | | |
| All Atoms | | 37.0 | 273.0 |
| Protein | | 37.0 | 273.0 |
| Water | | 38.0 | - |
| Coordinate error, maximum likelihood (Å) | | 0.32 | 2.055 (Luzzati Plot) |
| Ramachandran Plot | | | |
| Most favored (%) | | 97.9 | 88.9 |
| Additionally allowed (%) | | 2.1 | 8.1 |

1) Values in parenthesis are for the highest resolution shell.
2) $R_{merge} = \Sigma_{hkl}\Sigma_i |I_i(hkl) - \langle I(hkl)\rangle| / \Sigma_{hkl}\Sigma_i I_i(hkl)$, where $I_i(hkl)$ is the intensity measured for the *i*th reflection and $\langle I(hkl)\rangle$ is the average intensity of all reflections with indices hkl.
3) $R_{factor} = \Sigma_{hkl} ||F_{obs}(hkl)| - |F_{calc}(hkl)|| / \Sigma_{hkl} |F_{obs}(hkl)|$; $R_{free}$ is calculated in an identical manner using 5% of randomly selected reflections that were not included in the refinement.
4) $R_{meas}$ = redundancy-independent (multiplicity-weighted) $R_{merge}$ [28,29] $R_{pim}$ = precision-indicating (multiplicity-weighted) $R_{merge}$ [30,31].

**Supplementary Table 2.** 3D reconstructions of yeast 20S in complex with C-terminal peptides

| Name | Number of particles | Resolution (FSC=0.14) | EMDB entry number |
|---|---|---|---|
| CP | 321,296 | 6.9 Å | EMD-5593 |
| CP + Rpt1 C-terminal peptide | 78,994 | 7.3 Å | EMD-5611 |
| CP + Rpt2 C-terminal peptide | 88,013 | 7.7 Å | EMD-5612 |
| CP + Rpt3 C-terminal peptide | 74,177 | 7.7 Å | EMD-5613 |
| CP + Rpt4 C-terminal peptide | 36,303 | 7.8 Å | EMD-5614 |
| CP + Rpt5 C-terminal peptide | 58,953 | 8.1 Å | EMD-5615 |
| CP + Rpt6 C-terminal peptide | 51,135 | 7.9 Å | EMD-5616 |
| CP + $base_1$ | 51,000 | 9.6 Å | EMD-5617 |
| CP + $base_1$ (focused on Rpt-ring) | 51,000 | 10.0 Å | |
| CP + $base_2$ | 1,755 | 24.0 Å | |

**Supplementary Table 3. Yeast strains used in this study**

| strain | Relevant Genotype[a] | | figures | reference |
|---|---|---|---|---|
| SUB62 | *MAT***a** | *lys2-801 leu2-3, 2-112 ura3-52 his3-Δ200 trp1-1* | | [32] |
| SUB61 | *MATα* | *lys2-801 leu2-3, 2-112 ura3-52 his3-Δ200 trp1-1* | S16, S18 | [32] |
| sDL135 | *MAT***a** | *pre1::PRE1-TEV-ProA(HIS3)* | 1,2,4, S2, S4, S9, S10,S19 | [33] |
| sDL133 | *MAT***a** | *rpn11::RPN11-TEV-ProA(HIS3)* | 4c, 4d | [33] |
| TG931 | *MAT***a** | *rpt1::ProA-TEV-RPT1 (HIS3)* | 1a, S3 | This study |
| SP1734 | *MAT***a** | *rpn10::LEU2 rpt1::PROA-TEV-RPT1(HIS3)* | 1b, 1c, 2, S2, S4, S10, S11 | This study |
| SY36 | *MATα* | *rpt1::HIS3, ProA-TEV-RPT1 in pEL36 (TRP)* | S11, 4c, 4d, S19 | [34] |
| SP1737A | *MAT***a** | *rpt1::ProA-TEV-RPT1 (HIS3), rpt6::rpt6-Δ1 (KAN)* | 4c, 4d, S19 | This study |
| RJD1144 | *MAT***a** | *PRE1-Flag-His6::Ylpac211(URA3)* | | [35] |
| HM018 | *MAT***a** | *PRE1-Flag-His6::Ylpac211(URA3) PUP1-GST:: pFA6a-GST-KanMX6* | 3, S13 | This Study |
| DY30 | *MAT***a** | *rpt6::HIS3 [RPT6-URA3]* | | This Study |
| HM085 | *MAT***a** | *rpn4::KAN rpt6::HIS3 [RPT6_WT-URA3]* | | This Study |
| HM087 | *MAT***a** | *rpn4::KAN rpt6::HIS3 [RPT6_WT-LEU2]* | 4a, b, S16 | Rpt6_wt_LFK; this Study |
| HM088 | *MAT***a** | *rpn4::KAN rpt6::HIS3 [RPT6_F404Y-LEU2]* | 4a, b, S16 | Rpt6_LYK; this Study |
| HM089 | *MAT***a** | *rpn4::KAN rpt6::HIS3 [RPT6_F404W-LEU2]* | 4a, b, S16 | Rpt6_LWK; this Study |
| HM090 | *MAT***a** | *rpn4::KAN rpt6::HIS3 [RPT6_F404A-LEU2* | 4a, b, S16 | Rpt6_LAK; this Study |
| HM091 | *MAT***a** | *rpn4::KAN rpt6::HIS3 [RPT6_ΔK405-LEU2]* | 4a, b, S16 | Rpt6_LF; this Study |
| HM092 | *MAT***a** | *rpn4::KAN rpt6::HIS3 [RPT6_K405E-LEU2]* | 4a, b, S16 | Rpt6_LFE; this Study |
| HM093 | *MAT***a** | *rpn4::KAN rpt6::HIS3 [RPT6_K405A-LEU2]* | 4a, b, S16 | Rpt6_LFA; this Study |
| HM094 | *MAT***a** | *rpn4::KAN rpt6::HIS3 [RPT6_LAA-LEU2]* | 4a, b, S16 | Rpt6_LAA; this Study |
| HM095 | *MAT***a** | *rpn4::KAN rpt6::HIS3 [RPT6_AAA-LEU2]* | 4a, b, S16 | Rpt6_LAA; this Study |
| MHY297 | *MAT***a** | *pre1::PRE1-3XFLAG-HIS3* | | This Study |
| MHY298 | *MAT***a** | *pre1::PRE1-3XFLAG-HIS3 α3ΔN (2GSRRYDSRT10)* | | This Study |
| HM082 | *MAT***a** | *pre1::PRE1-3XFLAG-HIS3 α5ΔN (2FLTRSEYDRGV12)* | | This Study |
| HM083 | *MAT***a** | *pre1::PRE1-3XFLAG-HIS3 α6ΔN (2FRNNYDGDT10)* | | This Study |
| HM084 | *MAT***a** | *pre1::PRE1-3XFLAG-HIS3 α7ΔN (2TSIGTGYDLSN12)* | | This Study |
| sJR287 | *MATα* | *rpn4::KAN* | S16, S17, S18 | This Study |
| sJR730A | *MAT***a** | *rpn4::KAN rpt6::rpt6–Δ0 (NAT)* | S17, S18 | This Study |
| sJR732 | *MATα* | *rpn4::KAN rpt6::rpt6–Δ1 (NAT)* | S17, S18 | This Study |
| sJR719 | *MAT***a** | *rpt6::rpt6–Δ0 (NAT)* | S18 | This Study |
| sJR733 | *MAT***a** | *rpt6::rpt6–Δ1 (NAT)* | S18 | This Study |

[a] *All strains above carry the lys2-801, leu2-3, 2-112, ura3-52, his3-Δ200, and trp1-1 markers, and are congenic to SUB62*

# Supplementary Methods

**Yeast techniques, plamids and reagents**

All yeast manipulations were conducted according to standard protocols[36-38]. The generation of strains sJR730 and sJR719 utilized targeted integration of a PCR fragment generated using primers pRL159 (TAAGGTTATG AACAAGAACC AAGAAACGGC CATTTCTGTC GCCAAGCTGT TCAAGTGAGG CGCGCCACTT CTAAATAAGC) and pRL141 (ACACTAAGTA ACATATACAA TGAGCCAAGT GAAACGAATA GAATTCGAGC TCGTTTAAAC) and template pFa6-3HA-NATMX6. For strains sJR732 and sJR733, primer pRL159 was replaced by pRL137 (TAAGGTTATG AACAAGAACC AAGAAACGGC CATTTCTGTC GCCAAGCTGT TCTGAGGCGC GCCACTTCTA AATAAGC). Antibodies to Rpn14, Nas6, and Hsm3 were previously described[5,6]. Rpn1 rabbit polyclonal antibody was produced using the N-terminal 378 amino acids of Rpn1 (Proteintech). Anti-CP antibody (MCP231) and anti-Rpt5 antibodies are from ENZO Life Sciences. Rpt6 antibody is a gift from Carl Mann. ATP (A3377), ADP (A2754), and ATPγS (A1388) were purchased from Sigma.

**Affinity purification of base, CP, and chaperones**

To obtain proteasome base that is competent for *in vitro* reconstitution with CP and lid as well as analysis via native gel, we found that all buffers during the purification should contain 1-2 mM ATP, 5 mM $MgCl_2$, and 10% glycerol. Proteasome base was purified using a protein A affinity tag appended to the 5' end of the *RPT1* open reading frame in the endogenous chromosomal locus. Unless otherwise specified, the strain used for purification was SP1734, which is an *rpn10* null. Typically, a yeast culture of 4-6 L was grown in YPD overnight to an $OD_{600}$ of 15-20. Yeast cells were harvested and ground in liquid nitrogen as described previously[6]. Ground powders were then hydrated in buffer A (50 mM Tris-HCl [pH 7.5], 5 mM $MgCl_2$, 1 mM EDTA, 10% glycerol, and 1 mM ATP) supplemented with protease inhibitors[6,15], with the volume of buffer A corresponding to that of the cell pellet upon harvest. Lysates were then centrifuged at 20,000 x *g* for 30 min at 4°C. Supernatants were incubated with IgG resin for 1-2 hrs at 4°C. Resins were washed with 20 bed volumes of buffer A containing 100 mM NaCl, and incubated in 5 bed volumes of buffer A containing 1 M NaCl for 1 hr at 4°C. Resins were further washed in 50 bed volumes of buffer A containing 1 M NaCl at 4°C, followed by a final wash in 5 bed volumes of buffer A only. Base was then eluted by incubating the resin with 1 bed volume of buffer A containing TeV protease (Invitrogen) for 1 hr at 30°C. Following purification, base was stored at -80°C in small aliquots to minimize repeated freeze-thaw cycles.

CP was purified as described previously[33,34] except that lysis was performed in liquid nitrogen. Recombinant chaperones were purified as described previously, using pGEX6P-1 derived plamids pJR40 (expressing GST-Nas6), pJR56 (expressing GST-Rpn14) or pJR89 (expressing GST-HSm3)[5]. The GST tag was removed from each chaperone in the course of purification, using Prescission protease[5].

**Analysis of base-CP association by native gel and solution assays**

To obtain base that is functional in reconstitution, buffers during affinity purification contain 10% glycerol, 5 mM $MgCl_2$, and 1 mM ATP[6]. For native gels, typically 150-300 nM of affinity-purified base was mixed with CP at 2:1 ratio in 10-20 μl of buffer A at 30°C for 15-20 min unless indicated otherwise. The reactions were immediately loaded on 3.5% polyacrylamide native gels[39] and run for 2.5-3 hr at 100V at 4°C. In-gel peptidase assays were conducted as described previously, using an LSA3000 imaging system (Kodak) or AlphaImager system (ProteinSimple).

Real-time assays of the peptidase activity of base-CP complexes were conducted in solution. Purified base and CP (generally 5 nM and 2 nM, respectively, or otherwise as specified) were mixed in buffer containing

50 mM Tris-HCl [pH 7.5], 5 mM $MgCl_2$, 1 mM ATP, and 5 μM LLVY-AMC in 20 μL. An ATP regeneration system (52 mM creatine phosphate and 0.5 mg/ml creatine phosphokinase) was used when base-CP complex formation was assayed in the presence of ATP. The reactions were prepared in triplicate in 384-well plates[40] (Corning, low-volume plate,). AMC signals were measured using EnVision plate reader (Perkin Elmer)[40] in r.f.u. x $10^6$. For the experiments in Figures 1a-c, base, CP, and chaperones were added at the same time. For experiments in Figures 2a, b, d and Figures 4c, d, the addition of buffer, trap or chaperones was done manually. For the hydrolysis rate plots of Figures 2a and 2d, the raw LLVY-AMC hydrolysis curves were first smoothed by a running average of three adjacent values. The running average was applied twice, first to the raw curve and then to the smoothed curve. The first derivatives were then calculated from the smoothed curves.

**Preparation of CP trap**

Affinity-purified CP was incubated with epoxomicin (Boston Biochem) at 200-500 μM final concentration for 20 min at room temperature. Epoxomicin-treated CP was then dialyzed in buffer A in the absence of ATP overnight at 4°C.

**Modeling of base-CP-chaperone co-complexes**

For the modeling of Fig 1d, Modeller[41] was used initially to build the hexamer-ring model, using as a template PBD:1R7R (p97 D1 domain hexamer). Subsequent fitting to the EM map of the yeast Rpt hexamer was carried out with Chimera[42]. Superimposition of Hsm3 and Nas6 on the map of Rpt ring employed the structures of Hsm3-Rpt1C (PDB: 3VLF) and Nas6-Rpt3C (PDB: 2DZN). The clipped view of the Rpt ring with bound chaperones and CP (Fig. 1d) was created using Pymol[43].

**Expression and purification of Hsm3**

Recombinant Hsm3 was expressed as GST fusion in BL21(DE3)[5]. Cells were inoculated in LB supplemented with ampicillin and incubated at 37°C with shaking at 200 rpm. At $OD_{600}$ of 0.6, IPTG was added to 0.1 mM final concentration and the culture was incubated overnight at 21°C with shaking. Cells were harvested by centrifugation (4,000 *g*), washed in lysis buffer (50mM Tris-HCl [pH 7.5], 75 mM NaCl, 1mM EDTA, 1 mM DTT and 10% Glycerol), and resuspended in lysis buffer containing protease inhibitor cocktail tablet (Roche). Lysis was performed at 900 psi using a French Press system. After lysis, 0.1% Triton X-100 was added and lysate was rotated for 30 min at 4°C. Upon clearing the cell lysate (12K rpm, 30 min, 4°C), glutathione-agarose was added to the lysate and samples were incubated at 4°C for 1 hour. The resin was washed batch style four times with 50 bed volumes of PBS. To elute Hsm3, resin was washed once and resuspended in PreScission cleavage buffer (50 mM Tris-HCl [pH 6.8], 150 mM NaCl, 1 mM EDTA, 1 mM DTT, 0.01% Triton x-100) and incubated overnight in the presence of GST-PreScission protease. Eluted Hsm3 was further purified using an AKTA purifier with a HiLoad Superdex 200 column equilibrated with running buffer (50 mM Tris-HCl [pH 6.8], 50 mM NaCl, 1 mM EDTA, 1 mM DTT). Fractions containing monomeric Hsm3 protein were pooled and concentrated to 8 mg/ml in a buffer containing 50 mM Tris-HCl [pH 6.8], 50 mM NaCl, 1 mM EDTA, 1 mM DTT.

**Expression and purification of selenomethionine labeled Hsm3**

To produce Seleno-methionine labeled Hsm3 we expressed Hsm3 in BL21(DE3). 2 liters of cells were inoculated in LB supplemented with ampicillin at an $OD_{600}$ of 0.1 and shaken at 37°C at 200 rpm. At $OD_{600}$ of 0.6, cells were collected by centrifugation at 4,000 x *g* for 10 minutes at room temperature, washed in M9 salt minimal medium and resuspended in 2 liters of M9 medium without methionine[44]. After 30 minutes incubation (37°C, 200 rpm) 2 ml of seleno-methionine (50 mg/ml) (Acros, AC25996-0010) was added and cells were

incubated for an additional 30 min (37°C, 200 rpm). Next, IPTG was added to 0.1 mM final concentration and culture was incubated over night at room temperature (21°C, 200 rpm). Cells were harvested and purifications were performed as described above for unlabeled Hsm3. Successful incorporation of seleno-methionine was confirmed by mass spectrometry (Supplementary Figure 5a).

**Expression and purification of Hsm3-Rpt1 C-domain complex**

To prepare a complex of Hsm3 with the C-domain of Rpt1, we co-expressed GST-Hsm3 with His-tagged Rpt1 in *E.coli* as described previously[5]. Protein expression was induced with 0.2 mM IPTG at OD 0.8 and cells were grown overnight at room temperature. After harvesting, cells were resuspended in lysis buffer (20 mM Tris-HCl [pH 8.0], 100 mM NaCl, 20% glycerol, supplemented with protease inhibitor cocktail (EDTA free, Roche)). Cells were disrupted at 20 psi using a cell disruptor. Lysates were cleared at 20,000 x *g* for 30 min at 4°. Cleared lysate was loaded on a 5 ml His-Trap™ HP column and eluted using an imidazole gradient up to 500 mM in lysis buffer. Fractions containing the Hsm3-Rpt1 C-domain pair were pooled and loaded onto a GSTrap™ HP column. After washing (25 mM Sodium Phosphate buffer [pH 7.4], 50 mM NaCl, 1 mM EDTA, 1 mM DTT), proteins were eluted from the column using PreScission protease in elution buffer (50 mM Tris-HCl [pH 8.0], 10 mM glutathione, 1 mM DTT) that cleaves between GST and Hsm3. Fractions containing the protein pair were pooled, concentrated and loaded on a Superdex 200 10/300 GL column equilibrated with 50 mM Tris-HCl [pH 6.8], 50 mM NaCl, 1 mM EDTA, 1 mM DTT. The final sample was concentrated to 8.5 mg/ml and used for crystallization screening.

**Crystallization and data collection**

All crystallization screening was conducted in Compact Jr. (Emerald Biosystems) sitting drop vapor diffusion plates at 20°C using 0.5µL of protein and 0.5µL of crystallization solution equilibrated against 100 µL of the latter. Crystals of purified seleno-methionine labeled Hsm3, concentrated to 9 mg/mL in 150 mM NaCl, 50 mM Tris-HCl (pH 6.8), 1 mM DTT, 1 mM EDTA, were obtained overnight from the Wizard 4 screen (Emerald Biosystems) condition #19 (25%(w/v) PEG 1500, 100 mM MMT Buffer/NaOH [pH 6.5]). MTT buffer consists of D,L-malic acid, MES and Tris base in a 1:2:2 molar ratio. Crystals were transferred to a fresh drop containing 80% crystallization and 20% PEG 400 before flash freezing in liquid nitrogen. Crystals of native Hsm3 were obtained overnight from the Proplex HT screen (Molecular Dimensions) condition C6 (20%(w/v) PEG 4000, 100 mM Na-Hepes [pH 7.0], 150 mM ammonium sulfate). Crystals were transferred to a fresh drop containing 75% crystallization and 25% PEG 400 and frozen for data collection. Crystals of the Hsm3-Rpt1 C-domain complex were obtained overnight from 3 M Sodium Formate, 100 mM Bis-Tris-HCl (pH 7.0),10 mM calcium chloride. Crystals were transferred to a fresh drop containing 7 M sodium formate and frozen for data colletion. Data were collected at the Advanced Photon Source beamline 17-ID using a Dectris Pilatus 6M pixel array detector.

**Structure solution and refinement for Hsm3**

Intensities were integrated using XDS[45] and the Laue class check and data scaling were performed with Pointless and Scala[29]. The highest probability Laue class was *mmm* and space group $P2_12_12_1$. The Matthew's coefficient (Vm)[46] and solvent content were estimated to be Vm=2.3 / 46.7% solvent for 2 molecules in the asymmetric unit. Data for phasing were collected, for a Se-Met labeled sample, at the absorption edge λ=0.97898 Å (12664.78 eV) as determined from an X-ray fluorescence scan (Supplementary Figure 5b). Structure solution was conducted using the SAD method with Autosolve via the Phenix[47] interface which yielded a figure of merit of 0.39 and a Bayes-CC of 0.449. Following automated model building with Phenix, 811 residues out of 982 were placed which converged at $R$= 0.24, $R_{free}$= 0.28 following refinement. The resulting model, which consisted of a non-crystallographic dimer, was used for molecular replacement with

Phaser [48] against a native data set. The top solution was found in the space group $P2_12_12_1$. Structure refinement and manual model building were conducted with Phenix and Coot[49] respectively. Structure validation was conducted with Molprobity[50]. Disordered side chain atoms were truncated to the point where electron density could be observed. Figures were prepared using the CCP4MG package[51]. Relevant crystallographic data are provided in Supplementary Table 1.

**Sedimentation centrifugation**

Beckman XL-I analytical ultracentrifuge was used in sedimentation velocity and sedimentation equilibrium experiments with two-channel analytical cells. The Hsm3 sample was dialyzed against 50 mM Tris-HCl (pH 6.8), 50 mM NaCl, 1 mM EDTA, 1 mM β-mercaptoethanol. The sedimentation velocity data were analyzed using the time-derivative method in the Origin software distributed with the instrument. For the analysis of the sedimentation equilibrium data, the protein partial specific volume (0.747 ml/g) and the buffer density (1.004 g/ml) were calculated with Sednterp ver. 1.09 (D.B. Hayes, T. Laue, J. Philo, the University of New Hampshire). Fitting of the single-species model was performed using the Origin software distributed with the instrument.

**Cryo-EM data acquisition and image processing**

Purified GST-tagged *S. cerevisiae* 20S CP was diluted to a concentration of 1.6 μM, which is optimal for preparing cryoEM grids. For peptide-CP reconstruction, all six peptides were used at a concentration of 0.5 mM and were incubated with 20S CP for 1 hour at 37°C. For base-CP reconstruction, the base was used at a molar ratio of 2:1 against CP, and was incubated with 20S CP, 6 mM ATP and 5 mM $MgCl_2$ for 20 min at 30°C. A drop of 2 μL of the sample was applied to glow-discharged Quantifoil holey carbon grids (Quantifoil, Micro Tools GmbH, Germany), and plunge-frozen by using a Vitrobot Mark III (FEI company, USA). Grids of frozen hydrated samples were imaged using a Tecnai TF20 electron microscope equipped with a field-emission gun (FEI Company, USA) and operated at an acceleration voltage of 200 kV. Images at a defocus ranging from 1.2 to 2.1 μm were recorded at a nominal magnification of 100kX on a TemCam-F816 camera with 8196×8196 pixels (TVIPS GmbH, Germany). The final image has a pixel size of 0.737 Å on the specimen level. Total dose used to image was ~25$e^-/Å^2$. All data were collected in a semi-automated manner following strict low-dose procedure using the program UCSFImage[52].

Images were 2×2 binned before further processing. Binning was calculated by a procedure in which a forward Fourier transform of 8K×8K image was calculated, then the central 4K×4K region was cropped out and inverse Fourier transformed into a 4K×4K image with a pixel size of 1.474 Å. The CTFFIND3[53] was used to determine the defocus values. The particles in side view were picked automatically by using FindEM[54]. All particles were first subject to 2D classification by using standard procedure implemented in SPIDER[55], and particles of bad 2D classes were removed. More bad particles were removed by manual inspection. The total numbers of particles used for calculating each 3D reconstruction are listed in Supplementary Table 2.

Images of frozen hydrated CP with the GST-tag showed mostly side-view particles, which are needed for calculating a 3D reconstruction (Fig. 3a). Two-dimensional (2D) class averages show an extra density on one side of the CP, corresponding to GST (Supplementary Fig. 13c,d). 3D reconstructions were calculated and refined using GPU-enabled FREALIGN[56,57]. The crystal structure of the *S. cerevisiae* 20S CP[58] filtered to a resolution of 15 Å was used as the initial model for image alignments of CP alone and peptide-CP complex. Initial models of base-CP complexes were generated by adding a half Gauss ball to either one end or both ends of the initial model of CP alone to mimic the overall shape of the base-CP complexes. A 2-fold symmetry was applied to the 3D reconstructions of CP alone, peptide-CP and $base_2$-CP complexes. No symmetry was applied to the 3D reconstruction of $base_1$-CP complex. Resolutions of final 3D reconstructions were estimated using Fourier shell correlation (FSC = 0.143) criterion[59]. Resolutions of 3D reconstructions of CP and CP-peptide are

at subnanometer (Supplementary Fig. 13g and Supplementary Table 2). Secondary structural features clearly resolved. Docking of the atomic structure of the yeast CP3 confirmed the nominal resolution (Supplementary Fig. 13h), with the placement of densities corresponding to GST indicating correct alignment (Supplementary Figs. 13i and 14). Fourier amplitudes of the density maps with peptides were scaled to the level of 20S CP alone, and filtered to a resolution of 12 Å before calculating difference maps. 3D density maps of the CP incubated with peptides were first normalized against the reference map, and then subtracted by it to calculate difference maps (Supplementary Fig. 15). All difference maps are shown in Supplementary Figure 15. The contour levels of difference maps were set so that there is no specific binding in the pocket between α7 and α1 subunit, which lacks the required lysine residue for binding the C terminus of any activators. The volumes of each peptide densities within the pockets were estimated from the total voxel number of the difference densities. As controls, difference maps between two independent 3D reconstructions from two separately collected datasets of the same sample showed no significant difference density. Based on these and other controls (Supplementary methods), we conclude that the distinct difference densities assigned to each peptide were not generated by image misalignment or random noise. Density maps of single particle 3D reconstructions of yeast CP alone and in complex with Rpt peptides have been deposited in the Electron Microscopy Data Bank (EMDB) with accession code EMD-5593, and EMD-5511 to EMD-5516 (Supplementary Table 2).

The difference map of base was calculated by subtracting base$_1$-CP reconstruction from the same reconstruction but after 180$^o$ rotation around 2-fold symmetry axis of the CP. This difference map was used to further align the base portion independently from the CP. A hybrid map was generated by combining the CP portion from the 3D reconstruction of base$_1$-CP complex with the base portion determined by aligning only the base of the complex. The density maps in Figure 5 are low-pass filtered to 10 Å with a soft cosine edge of 10-pixel width, and sharpened by a negative b-factor, -600 Å$^2$. We used the "fit-in-map" option in UCSF Chimera to dock atomic structures into cryoEM density maps[42]. For CP alone and CP-peptide complexes, seven different orientations of docking the atomic structure into the density maps were tested and the correct docking gives the highest correlation value (Supplementary Figure 14).

**Subunit assignment of base-CP complex**

Particles of base$_1$- and base$_2$-CP complexes were separated by standard 2D multi-reference alignment and classification procedure using SPIDER[55] (Supplementary Figure 20b and e). Only less than 5% of base-CP complex particles had two bases bound to both ends of the CP (referred as base$_2$-CP), resulted a 3D reconstruction with a lower resolution, which is not sufficient to allow us assign subunits unambiguously. We then overlaid the 3D reconstruction of base$_2$-CP complex with the holoenzyme[2] with their 2-fold symmetry axes aligned with each other (Supplementary Figure 20d). Because all proteasomal subunits were assigned unambiguously in the holoenzyme, the overlaps of the two density maps enabled us to transfer all subunit assignments into the base$_2$-CP complex unambiguously (Supplementary Figure 20c and d). Asymmetrical features shown in the reconstructions, including the position of Rpn1, off-axis position of the OB domain in Rpt ring and the N-termini densities of three CP α subunits (α1, α3 and α5), enabled us to align the base$_1$-CP and base$_2$-CP complexes and thus to assign all the subunits to the base$_1$-CP complex unambiguously (Supplementary Figure 20f).

**Base is relatively mobile in the base-CP complex**

In the 3D reconstruction of the base$_1$-CP complex (EMD-5517), the density of the base portion is less complete than the CP portion. In comparison, the density of base is more complete when this part was refined independently from the CP. This indicates that the base is mobile in the base-CP complex. The base is also tilted with one side (Rpt3/Rpt6/Rpt2) closer to the CP than the other side.

# Supplementary Discussion

**Structure of Hsm3**

The structure of Hsm3, determined to 2.2 Å resolution (PDB ID: 4FP7; Supplementary Figure 6 and Supplementary Table 1), contains two molecules of Hsm3 in the asymmetric unit. Structure predictions suggest Hsm3 consists of HEAT or Armadillo like repeat, which are evolutionary related. HEAT repeats consist of two anti-parallel alpha helixes linked by a flexible loop, while ARM repeats contain three alpha helixes, where the two shorter helixes are the equivalent to the bended first helix of the HEAT repeat. The structure shows that Hsm3 is built up exclusively out of alpha helixes connected by loops, forming a total of 7 HEAT repeats and 4 ARM repeats. Alignment of the two Hsm3 molecules in the dimer shows they have a similar structure using secondary structure matching with an overall root mean square deviation of 0.48 Å between C-alpha atoms (418 residues). In subunit A an additional alpha helix could be traced at the N-terminus, but this was disordered in subunit B. The determined structure could be modeled from TYR 8 to GLY 466 for subunit A, and from ASN 30 to GLY 466 for Subunit B (amino acid numbering based on SGD for the translated protein).

The Hsm3 molecule forms a structure with a large curvature, creating a C-shaped protein. This unusually strong curvature explains why molecular replacement attempts were unsuccessful using homologous structures as the search model and therefore experimental phasing using seleno-methionine labeled protein had to be used. The two Hsm3 molecules are oriented nearly orthogonally to one another and adopt an orientation in which the N- and C-terminal regions of each respective subunit interact with the middle region of the other. This intricate interaction has also been observed by others[19,20] and might suggest Hsm3 exists as a dimer *in vivo* or *in vitro*. However, size exclusion chromatography of purified Hsm3 shows this protein migrates at a molecular weight consistent with a monomer (data not shown). Because Hsm3 concentration will be reduced in the course of size exclusion chromatography, we addressed whether Hsm3 in solution at high concentrations (>4 mg/ml) forms dimers using analytical ultracentrifuge centrifugation. Sedimentation velocity analysis (Supplementary Figure 7a) of Hsm3 at 4.4 mg/ml showed a single molecular species with a sedimentation coefficient of ~3 S. Sedimentation equilibrium results (Supplementary Figure 7b, lower panel) were also consistent with a single-species model, as shown by small and randomly distributed residuals of the fit (Supplementary Figure 7b, upper panel), and yielded a molecular weight of 54.2 kDa. Altogether, the analytical ultracentrifugation results indicate that Hsm3 is monomeric in solution over a broad concentration range.

We have previously shown that Hsm3 interacts directly with Rpt1 and this binding mapped to the C-domain of Rpt1[5]. Expression of Rpt1-C-domain in the absence of Hsm3 resulted in the majority of Rpt1 being insoluble after lysis. Therefore, we co-expressed Hsm3 with Rpt1 and used affinity purification to purify the complex. The purified complex yielded crystals that diffracted to ~5 Å resolution and belonged to the space group $P6_522$ with $a$ = 185.4 Å, $c$ = 357.1 Å. The structure of the Hsm3-Rpt1 complex was solved by molecular replacement using the Hsm3 structure as a search model. Following initial refinement, difference electron density (Fo-Fc) consistent with the Rpt1-C-domain was observed, which allowed us to assign the backbone trace of Rpt1 in this structure and to determine the global binding mode of these two proteins (Supplementary Figure 8). The structure of the Hsm3-Rpt1 complex was refined with Buster[60] with *B*-factors set to a constant value as determined from the Wilson plot. During the preparation of this manuscript, two structures (PDB: 3VLF and 4AV3) of the Hsm3-Rpt1 domain complex, refined to 3.8 Å resolution were published[19,20]. These structures were similar to the 5Å resolution structure that we had determined independently and also were solved in the same space group ($P6_522$). However, the unit cell constants for the structure reported here differ from these higher resolution structures. The structure of Hsm3-Rpt1 C-domain can provide insight into how Hsm3

functions as a chaperone by aligning the structure with the PAN ATPases as done previously for Nas6 (Ref. [5] and Supplementary Figure 9), or by aligning the structure to the recent EM-structure of the proteasome holoenzyme, as described in the main text (Figure 1d). Both of these approaches suggest substantial physical clashing between the CP and the chaperones, consistent with the steric interference hypothesis[5,6]. In all alignments by us as well as others[19,20], the Hsm3 structure shows a steric clash with Rpt2. As all literature consistently shows that Hsm3 can bind to Rpt1-Rpt2 complexes as well as the assembled base and RP, some structural rearrangements may occur for Hsm3-Rpt1 C-domain as compared to Hsm3 in larger assemblies.

## Supplementary References (continue from main text)


26 da Fonseca, P. C. A. & Morris, E. P. Structure of the Human 26S Proteasome: subunit radial displacements open the gate into the proteolytic core. . *J. Biol. Chem.* **283**, 23305-23314 (2008).

27 Zhang, F. *et al.* Structural insights into the regulatory particle of the proteasome from Methanocaldococcus jannaschii. *Mol. Cell* **34**, 473-484 (2009).

28 Evans, P. Scaling and assessment of data quality. *Acta Crystallogr. D* **62**, 72-82 (2006).

29 Evans, P. R. An introduction to data reduction: space-group determination, scaling and intensity statistics. *Acta Crystallogr. D* **67**, 282-292 (2011).

30 Diederichs, K. & Karplus, P. A. Improved R-factors for diffraction data analysis in macromolecular crystallography. *Nat. Struct. Biol.* **4**, 269-275 (1997).

31 Weiss, M. S. Global indicators of X-ray data quality. *J. Appl. Crystallogr.* **34**, 130-135 (2001).

32 Finley, D., Ozkaynak, E. & Varshavsky, A. The yeast polyubiquitin gene is essential for resistance to high temperatures, starvation, and other stresses. *Cell* **48**, 1035-1046 (1987).

33 Leggett, D. S. *et al.* Multiple associated proteins regulate proteasome structure and function. *Mol. Cell* **10**, 495-507 (2002).

34 Leggett, D. S., Glickman, M. H. & Finley, D. Purification of proteasomes, proteasome subcomplexes, and proteasome-associated proteins from budding yeast. *Methods Mol. Biol.* **301**, 57-70 (2005).

35 Verma, R. *et al.* Proteasomal proteomics: identification of nucleotide-sensitive proteasome-interacting proteins by mass spectrometric analysis of affinity-purified proteasomes. *Mol. Biol. Cell* **11**, 3425-3439 (2000).

36 Guthrie, C. & Fink, G. R. *Guide to yeast genetics and molecular biology*. (Academic Press, San Diego, 1991).

37 Wach, A., Brachat, A., Pohlmann, R. & Philippsen, P. New heterologous modules for classical or PCR-based gene disruptions in Saccharomyces cerevisiae. *Yeast* **10**, 1793-1808 (1994).

38 Longtine, M. S. *et al.* Additional modules for versatile and economical PCR-based gene deletion and modification in Saccharomyces cerevisiae. *Yeast* **14**, 953-961(1998).

39 Elsasser, S., Schmidt, M. & Finley, D. Characterization of the Proteasome Using Native Gel Electrophoresis. *Methods Enzymol.* **398**, 353-363 (2005).

40 Lee, B. H. *et al.* Enhancement of proteasome activity by a small-molecule inhibitor of USP14. *Nature* **467** (2010).

41 Sali, A. Comparative protein modeling by satisfaction of spatial restraints. *Mol. Med. Today* **1**, 270-277 (1995).

42 Pettersen, E. F. *et al.* UCSF Chimera--a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605-1612 (2004).

43 Schrodinger, L. The Pymol Molecular Graphics System, Version 1.3r1. (2010).

44 Doublie, S. Production of selenomethionyl proteins in prokaryotic and eukaryotic expression systems. *Methods Mol. Biol.* **363**, 91-108 (2007).

45 Kabsch, W. Automatic indexing of rotation diffraction pattens. *J. appl. Crystallogr.* **21**, 67-72 (1988).

46 Matthews, B. W. Solvent content of protein crystals. *J. Mol. Biol.* **33**, 491-497 (1968).

47 Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66**, 213-221 (2010).

48 McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658-674 (2007).

49 Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **66**, 486-501 (2010).

50 Chen, V. B. *et al.* MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **66**, 12-21 (2010).

51 Potterton, L. *et al.* Developments in the CCP4 molecular-graphics project. *Acta Crystallogr D Biol Crystallogr* **60**, 2288-2294 (2004).

52 Avila-Sakar, A., Li, X. M., Zheng, S. Q. & Cheng, Y. Recording high-resolution images of two-dimensional crystals of membrane proteins. *Methods Mol. Biol.* **955**, 129-52. (2013).

53 Mindell, J. A. & Grigorieff, N. Accurate determination of local defocus and specimen tilt in electron microscopy. *J. Struct. Biol.* **142**, 334-347 (2003).

54 Roseman, A. M. FindEM--a fast, efficient program for automatic selection of particles from electron micrographs. *J. Struct. Biol.* **145**, 91-99 (2004).

55 Shaikh, T. R. *et al.* SPIDER image processing for single-particle reconstruction of biological macromolecules from electron micrographs. *Nat. Protoc.* **3**, 1941-1974 (2008).

56 Li, X., Grigorieff, N. & Cheng, Y. GPU-enabled FREALIGN: accelerating single particle 3D reconstruction and refinement in Fourier space on graphics processors. *J. Struct. Biol.* **172**, 407-412(2010).

57 Grigorieff, N. FREALIGN: high-resolution refinement of single particle structures. *J. Struct. Biol.* **157**, 117-125 (2007).

58 Groll, M. *et al.* Structure of 20S proteasome from yeast at 2.4 A resolution. *Nature* **386**, 463-471 (1997).

59 Rosenthal, P. B. & Henderson, R. Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *J. Mol. Biol.* **333**, 721-745 (2003).

60 Blanc, E. *et al.* Refinement of severely incomplete structures with maximum likelihood in BUSTER-TNT. *Acta Crystallogr D Biol Crystallogr* **60**, 2210-2221, (2004).