

Supplemental data

Supplemental figure legends

Figure S1 Purification, EM data collection and class averages

(A) SDS gel obtained after size-exclusion chromatography of the *S. cerevisiae* exosome. The molecular weight marker is shown on the left and the Exo-10 subunits are shown on the right. (B) Negative stain (top) and cryo-EM (bottom) images of exosome particles. (C) Orthogonal views of the final 3D cryo-EM model of the apo exosome, reprojections of the maps and corresponding class averages.

Figure S2 Fourier shell correlation curves

(A) Fourier shell correlation (FSC) curves of the apo exosome cryo-EM map. The dashed lines indicate the resolution at 0.5 correlation. (B) FSC curve of the apo exosome negative stain map. (C) FSC curve of the RNA-bound exosome cryo-EM map. (D) FSC curve of the RNA-bound exosome negative stain map.

Figure S3 Fitting of *S.cerevisiae* 10-subunit pseudo-atomic model into the EM map

(A) Final fit of the RNase PH ring. Top: top view representation highlighting the pseudo-three fold symmetry. The dashed lines show the borders between the three RNase PH dimers. The numbering refers to the orientations of the three side views displayed at the bottom. Bottom: side views of the RNase PH like ring. Borders of the dimers are shown as dashed lines. (B) Top: the rigid fit of the Rrp44 crystal structure lacks density for CSD2 and the linker between CSD2 and the exoribonuclease domains suggesting that these regions are mobile. A small movement of CSD1 is also necessary for an optimal fit to the EM map. Bottom: final fit of Rrp44. (C) Top: the

rigid fit of KH/S1 cap proteins from the crystal structure of human Exo-9 reveals that movements are necessary to fit the EM map, suggesting the flexibility of these subunits. No density is present for the Rrp4 and Csl4 N-terminal domains (shown in grey). Bottom: final fit of *S. cerevisiae* pseudo-atomic model. As the N-terminal domains of Rrp4 and Csl4 do not display well-ordered density, they are excluded from the model.

Figure S4 Negative stain EM of RNA-bound exosomes

(A) Projection averages of negatively stained apo exosomes (left) with the 3D map in the corresponding orientation (right). (B) Projection averages of negatively stained RNA-bound exosomes (left) with the 3D map in the corresponding orientation (right).

Supplemental experimental procedures

Cryo-EM data collection and processing

Vitrified samples of *S. cerevisiae* apo and RNA-bound exosome were imaged on a Polara cryo-electron microscope (FEI) operated at 300 kV at a magnification of 121000 and a defocus range of 1.5-3.5 μm . Data were collected on a 4k x 4k charged coupled device (CCD) camera (Gatan). Micrographs were visually inspected before determining the defocus and astigmatism parameters with the MRC program CTFIND3 (Mindell & Grigorieff, 2003). Micrographs were binned to 2.6 \AA /pixel before particles were manually picked in Ximdisp and extracted in 128 by 128 pixel boxes with LABEL (Crowther et al, 1996). The particles were then corrected for the effect of the contrast transfer function by phase flipping, filtered between 150 and 8 \AA and normalized in SPIDER (Frank et al, 1996).

Particles were initially centred against a rotationally average total sum and classified using multivariate statistical analysis (MSA) as implemented in IMAGIC-5 (van Heel et al, 1996). Classes were inspected and selected (based on the visual match between the class average and the individual particles) for multi-reference alignment (MRA). After three rounds of MRA/MSA, anchor set refinement in IMAGIC-5 was used to assign angles to classes using projections of either a previously published negative stain structure of the yeast exosome (Wang et al, 2007) or the crystal structure of the human exosome (Liu et al, 2006). The model projections were in both cases filtered to 40 Å resolution to minimize model bias. A subset of good classes with angles covering the Euler sphere was selected for 3D reconstruction using the TRUE3D command in IMAGIC-5. This initial model was inspected, reprojected and used in MRA/MSA to improve alignment and classification of the particles. Classification, angle assignment and 3D reconstruction were iterated until the 3D model no longer improved. A final model was obtained using projection matching as implemented in SPIDER. The angular spacing between projections was reduced from 5° to 1° in successive alignments. When the alignment had stabilized, more than 95% of the images aligned to the same references in consecutive alignments. The final reconstructions were generated using 80% of the particles with the highest correlation coefficients. As a quality control, a map was also reconstructed using 20% of the particles with the highest correlation coefficients. The maps calculated with 20% or with 80% of the best particles are very similar, indicating that any degree of heterogeneity in the dataset does not significantly impact the resulting map (data not shown).

The RNA bound data set was processed as described for the apo data. The final apo 3D model (filtered to 40 Å resolution) was used as a starting model in anchor set refinement. To confirm the extra densities present at the entrance and the exit sites of the channel, the extra densities were deleted from the RNA-bound exosome map and projection matching carried out using this map as a starting point. After only a few rounds of projection matching extra densities reappeared strongly suggesting that they represent real features. Competitive alignment using projections of cryo-EM apo and RNA-bound maps as templates for projection matching do not yield to an apo class suggesting that RNA is bound to the vast majority of the exosome particles. The resolution was assessed by Fourier shell correlation (FSC) at 0.5 correlation and with a loose mask applied to the 3D maps giving 14 Å for the apo exosome structure and 12 Å for the RNA-bound exosome structure.

Negative stain EM data collection and processing

For the apo exosome, purified exosome at a concentration of 0.02 mg/ml was applied for 30 sec to a glow discharged continuous carbon grid, blotted and then stained for 30 s with 2% uranyl acetate. For the RNA-bound exosome, the RNA and the exosome were incubated for 30 min in a 1.2:1 molar ratio, subsequently diluted to reach a protein concentration of 0.02 mg/ml and stained as described above. Data was collected on a F20 cryo-electron microscope operated at 200 kV (FEI) equipped with a 4k x 4k CCD camera with a defocus range of 0.6 to 1 µm at a magnification of 66000. Data quality, defocus and astigmatism determination were determined as for the cryo-EM datasets. 5120 and 6605 particles were manually picked using Ximdisp for apo and RNA-bound exosomes, respectively. Particles were extracted in 150 x 150 pixel boxes with a pixel size of 2.3 Å using BOXER, CTF corrected by phase

flipping, band pass filtered between 150 and 15 Å and normalized in SPIDER. Apo and RNA-bound exosome datasets were aligned to projections of the apo exosome cryo-EM dataset and nine rounds of projection matching were performed in SPIDER. Extra densities attributed to RNA and exosome loops becoming ordered upon RNA binding appeared in the map for the RNA-bound exosome but not for the apo exosome. This is the case even though the negative-stain data of the RNA-bound exosome particles was aligned against projections of the apo cryo-EM map in which these features are absent. The resolution of the negative stain EM maps was assessed by Fourier shell correlation (FSC) at 0.5 correlation and with a loose mask applied to the 3D maps giving 16 Å for the apo exosome structure and 17 Å for the RNA-bound exosome structure.

To unambiguously establish that RNA is bound to the exosome in our EM experiments, we designed a gold-labeled RNA molecule that can be visualized in raw images of negatively stained exosomes. For this purpose, a 40 nt ssRNA labeled at the 5' end with biotin (synthesized by Dharmacon) was incubated for 30 min at 4°C with streptavidin covalently linked to one large 5 nm gold particle (Agar Scientific). The exosome was incubated with this gold-labeled RNA for 30 min at room temperature before final dilution and negative staining (as described above).

Fitting crystal structures into EM maps

Crystal structures of the *S. cerevisiae* ΔN-Rrp40, Rrp41, Rrp44 and Rrp45 proteins (Bonneau et al., 2009; Oddone et al., 2006) were used for fitting into the *S. cerevisiae* EM maps (PDB codes 2WP8 and 2JA9). Homology models were generated for the six other subunits and the N-terminal domain of Rrp40 with MODELLER (Sali &

Blundell, 1993) based on the alignment with their human exosome proteins (Liu et al, 2006) (pdb code 2NN6). The alignments were generated using MUSCLE (Edgar, 2004). The sequence identities between the human and yeast proteins are in the range of 23-43%. The N-terminal part of Csl4 (residues 1-135) could not be modelled with high confidence due to low sequence identity and long insertions compared to its human homolog and was thus omitted from the pseudo-atomic model.

Initially, the hexameric exosome ring (Rrp41/45, Rrp43/46, Rrp42/Mtr3) was fitted in both negative stain and cryo-EM maps using the *fit-in-map* module of the Chimera program (Goddard et al, 2007) starting from different initial positions. The best fits were then further refined within the cryo-EM map using simulated annealing rigid-body dynamics as implemented in the Flex-EM program (Topf et al, 2008). Next, fits for the cap proteins were analysed using both negative stain and cryo-EM maps, initially keeping contacts between the cap proteins and the hexameric ring as in the human crystal structure. The position of Rrp4 appeared to be mainly conserved with its C-terminal S1 and KH domains displaying well-ordered density, but the density for its smaller N-terminal domain could not be resolved in the map, suggesting that this domain is flexible. The EM density indicated a bigger movement of Rrp40 and its fit had to be optimized based on visual inspection and the Chimera *fit-in-map* module. With Rrp40 positioned in the map, some density remained unfilled that likely corresponds to the C-terminal part of Csl4 and a possible fit of this subunit is proposed in Figures 1 and S3C. The fit of the cap proteins was then refined in the cryo-EM map using Flex-EM. Next, Rrp44 was fitted into the map, keeping contacts between Rrp41/44/45 as in the *S. cerevisiae* crystal structure. Whereas the positions of the PIN, exoribonuclease and S1 domains appeared very conserved, the CSD1 fit

had to be optimized based on visual inspection and using the Chimera *fit-in-map* module. The fit of Rrp44 was then refined using Flex-EM dividing this protein into five domains. Finally, the entire model was refined using Flex-EM to maximise the fit to the EM map while keeping acceptable stereochemistry and preventing overlaps (Topf et al, 2008). The fitting procedure for the apo exosome improved the cross-correlation coefficient between the cryo-EM map and the pseudo-atomic model from 0.54 for a rigid body fit of the crystal structures to 0.62 for the final model after flexible fitting. Fitting of the RNA-bound exosome was carried out similarly to that of the apo exosome. This fitting improved the cross-correlation coefficient from 0.51 to 0.58 and showed an overall structure similar to that of the apo exosome.

Yeast complementation assay with Rrp4 mutant

The DNA fragment encompassing the Rrp4 ORF with adjacent 500 upstream and 495 downstream nucleotides was amplified by PCR and cloned into XhoI/XbaI sites of pRS415 vector generating the pADZ426 plasmid. R149E, R150D and K151D Rrp4 mutations were introduced by overlap PCR and cloned into the vector as described above for the pADZ426 plasmid. Constructs with wild type and mutated *RRP4* gene as well as an empty vector as negative control were transformed into the *S. cerevisiae* TH_6151 strain with the *RRP4* gene under control of tet- off promoter (PMID: 15242642). Transformed cultures were grown in synthetic complete medium without leucine at 30 °C overnight before spotting serial dilutions onto two plates in the absence or presence of doxycycline (10 µg/ml) to repress chromosomal WT *RRP4* gene expression. Cell growth was analyzed after 72 h of incubation at 30 °C.

Supplemental references

Crowther RA, Henderson R, Smith JM (1996) MRC image processing programs. *J Struct Biol* **116**(1): 9-16

Frank J, Radermacher M, Penczek P, Zhu J, Li Y, Ladjadj M, Leith A (1996) SPIDER and WEB: processing and visualization of images in 3D electron microscopy and related fields. *J Struct Biol* **116**(1): 190-199

Goddard TD, Huang CC, Ferrin TE (2007) Visualizing density maps with UCSF Chimera. *J Struct Biol* **157**(1): 281-287

Hernandez H, Dziembowski A, Taverner T, Seraphin B, Robinson CV (2006) Subunit architecture of multimeric complexes isolated directly from cells. *EMBO Rep* **7**(6): 605-610

Liu Q, Greimann JC, Lima CD (2006) Reconstitution, activities, and structure of the eukaryotic RNA exosome. *Cell* **127**(6): 1223-1237

Mindell JA, Grigorieff N (2003) Accurate determination of local defocus and specimen tilt in electron microscopy. *J Struct Biol* **142**(3): 334-347

Sali A, Blundell TL (1993) Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol* **234**(3): 779-815

Topf M, Lasker K, Webb B, Wolfson H, Chiu W, Sali A (2008) Protein structure fitting and refinement guided by cryo-EM density. *Structure* **16**(2): 295-307

van Heel M, Harauz G, Orlova EV, Schmidt R, Schatz M (1996) A new generation of the IMAGIC image processing system. *J Struct Biol* **116**(1): 17-24

Wang HW, Wang J, Ding F, Callahan K, Bratkowski MA, Butler JS, Nogales E, Ke A (2007) Architecture of the yeast Rrp44 exosome complex suggests routes of RNA recruitment for 3' end processing. *Proc Natl Acad Sci U S A* **104**(43): 16844-16849

Crowther RA, Henderson R, Smith JM (1996) MRC image processing programs. *J Struct Biol* **116**(1): 9-16

Edgar RC (2004) MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* **5**: 113

Frank J, Radermacher M, Penczek P, Zhu J, Li Y, Ladjadj M, Leith A (1996) SPIDER and WEB: processing and visualization of images in 3D electron microscopy and related fields. *J Struct Biol* **116**(1): 190-199

Goddard TD, Huang CC, Ferrin TE (2007) Visualizing density maps with UCSF Chimera. *J Struct Biol* **157**(1): 281-287

Hernandez H, Dziembowski A, Taverner T, Seraphin B, Robinson CV (2006) Subunit architecture of multimeric complexes isolated directly from cells. *EMBO Rep* **7**(6): 605-610

Liu Q, Greimann JC, Lima CD (2006) Reconstitution, activities, and structure of the eukaryotic RNA exosome. *Cell* **127**(6): 1223-1237

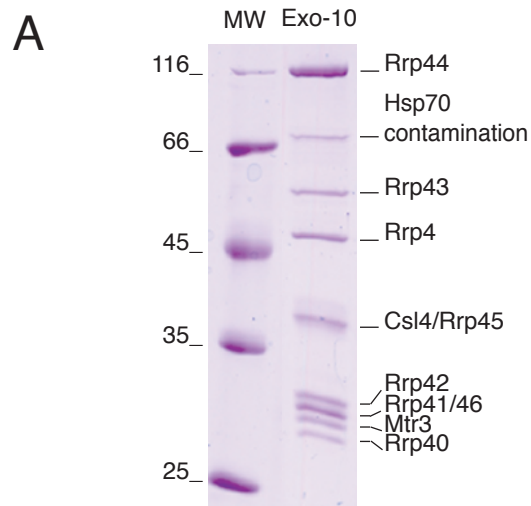
Mindell JA, Grigorieff N (2003) Accurate determination of local defocus and specimen tilt in electron microscopy. *J Struct Biol* **142**(3): 334-347

Sali A, Blundell TL (1993) Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol* **234**(3): 779-815

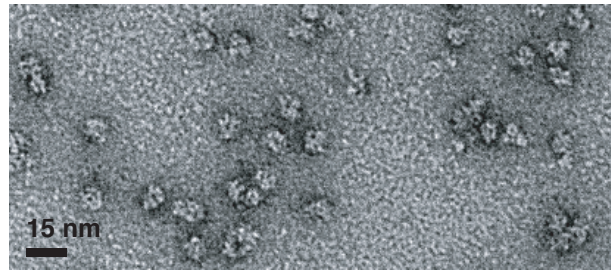
Topf M, Lasker K, Webb B, Wolfson H, Chiu W, Sali A (2008) Protein structure fitting and refinement guided by cryo-EM density. *Structure* **16**(2): 295-307

van Heel M, Harauz G, Orlova EV, Schmidt R, Schatz M (1996) A new generation of the IMAGIC image processing system. *J Struct Biol* **116**(1): 17-24

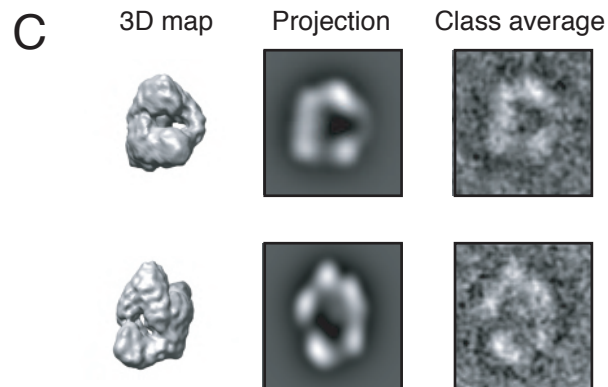
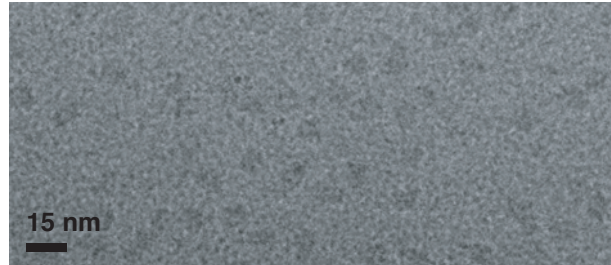
Wang HW, Wang J, Ding F, Callahan K, Bratkowski MA, Butler JS, Nogales E, Ke A (2007) Architecture of the yeast Rrp44 exosome complex suggests routes of RNA recruitment for 3' end processing. *Proc Natl Acad Sci U S A* **104**(43): 16844-16849



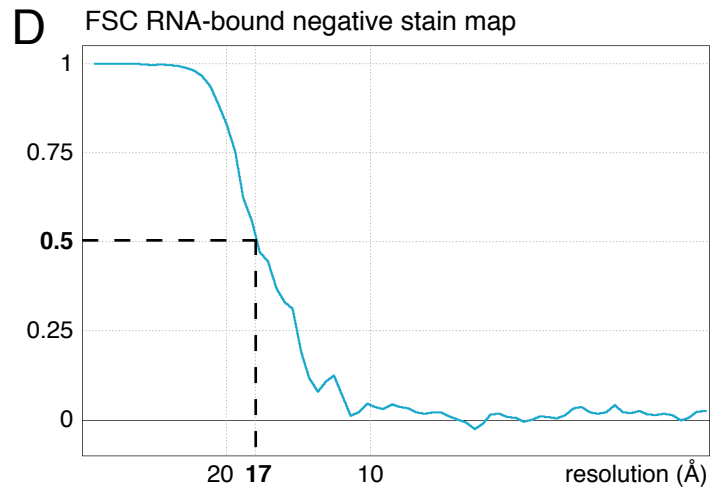
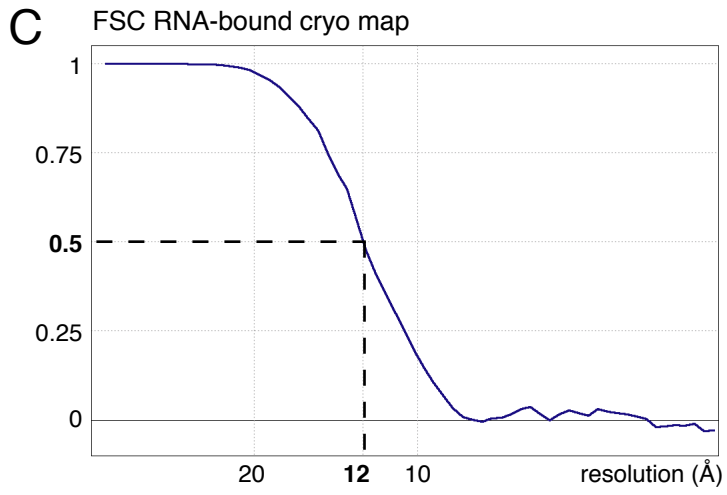
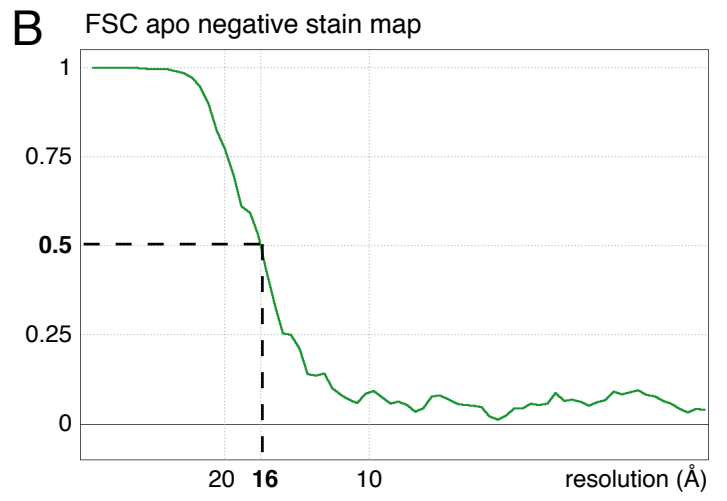
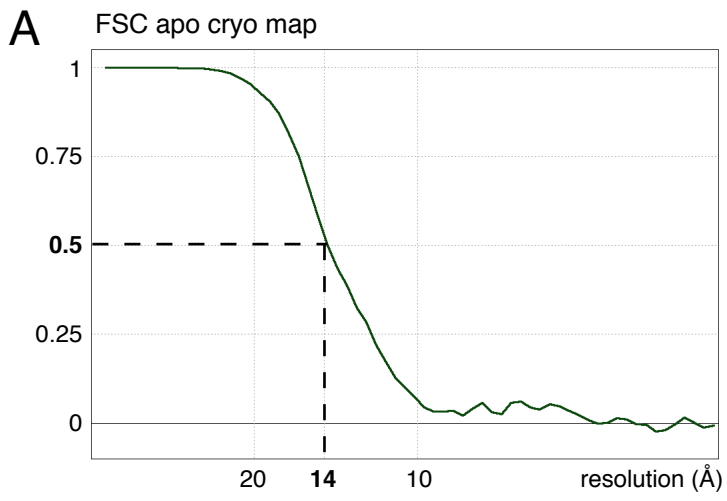
B Negative stain EM micrograph



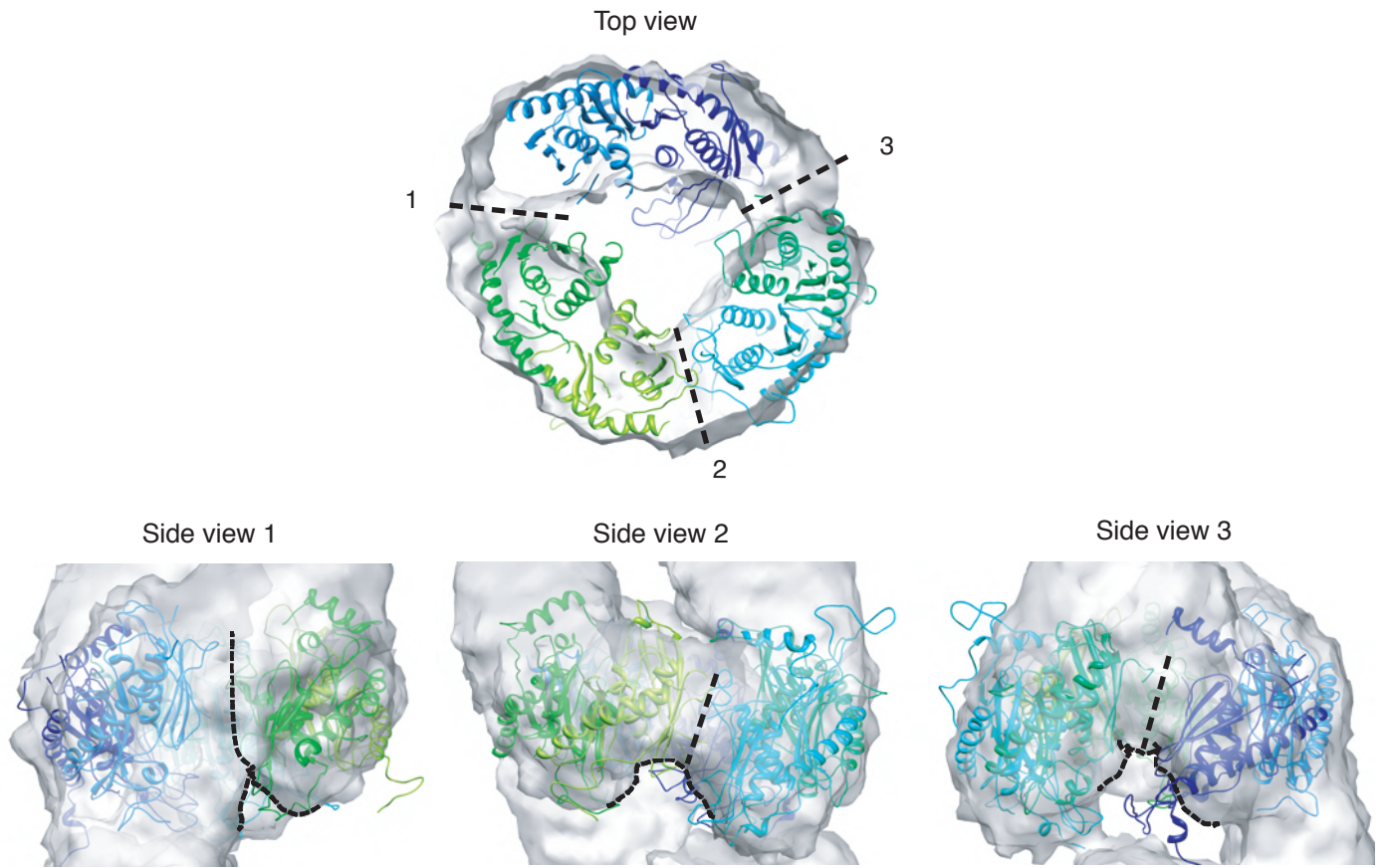
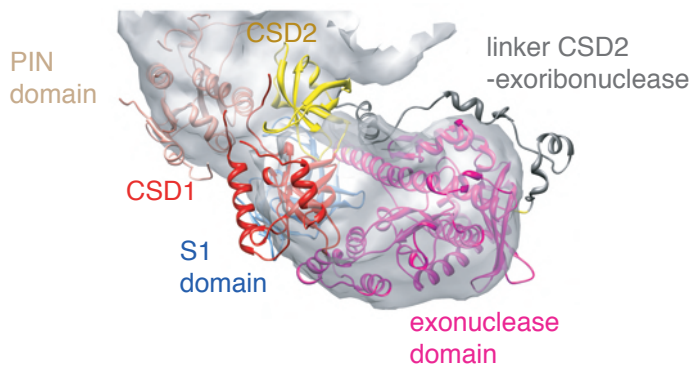
Cryo-EM micrograph



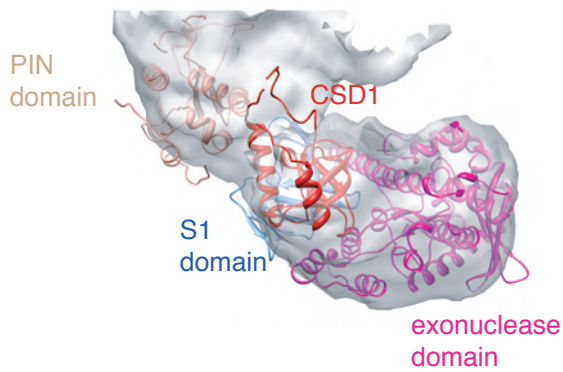
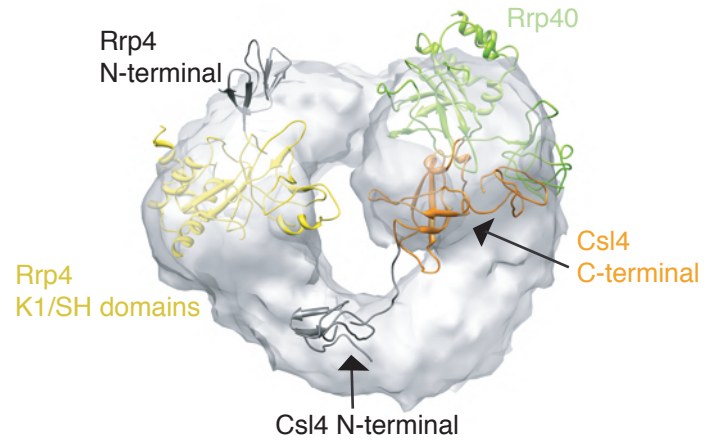
Supplementary Figure 1



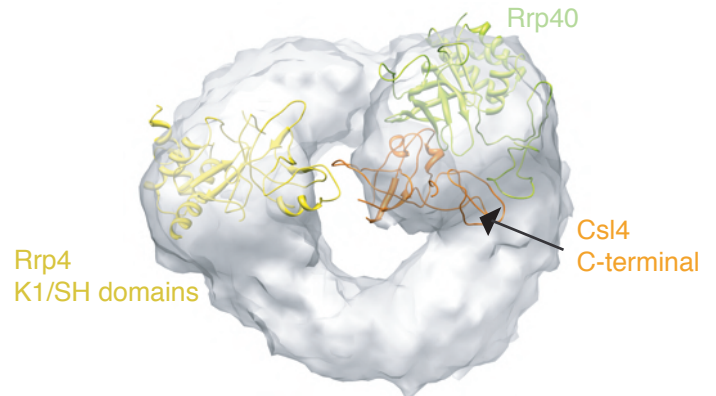
Supplementary figure 2

A**B** Rigid fit of Rrp44 crystal structure

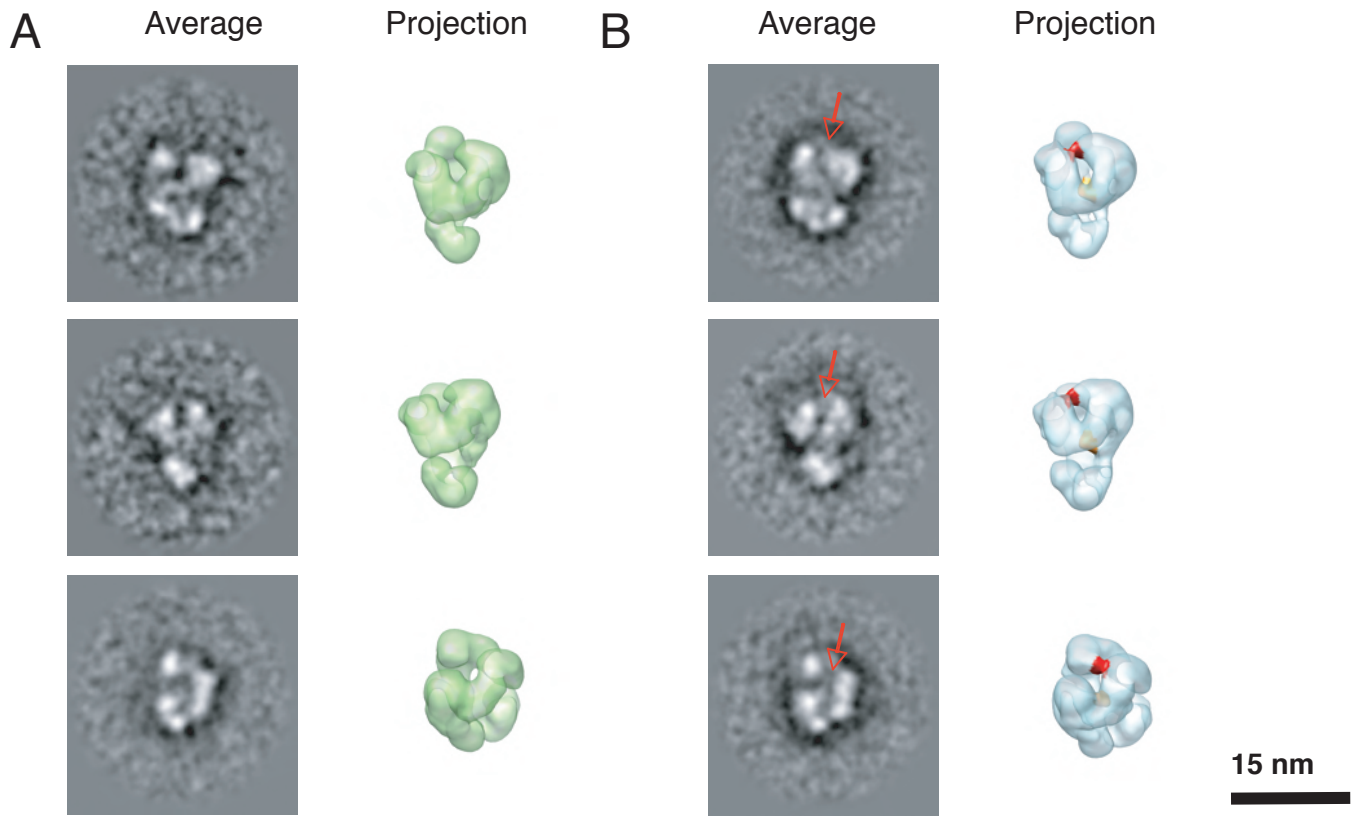
Final fit of the EM map

**C** Rigid fit of KH/S1 cap subunits

Final fit of KH/S1 cap subunits



Supplementary figure 3



Supplementary Figure 4