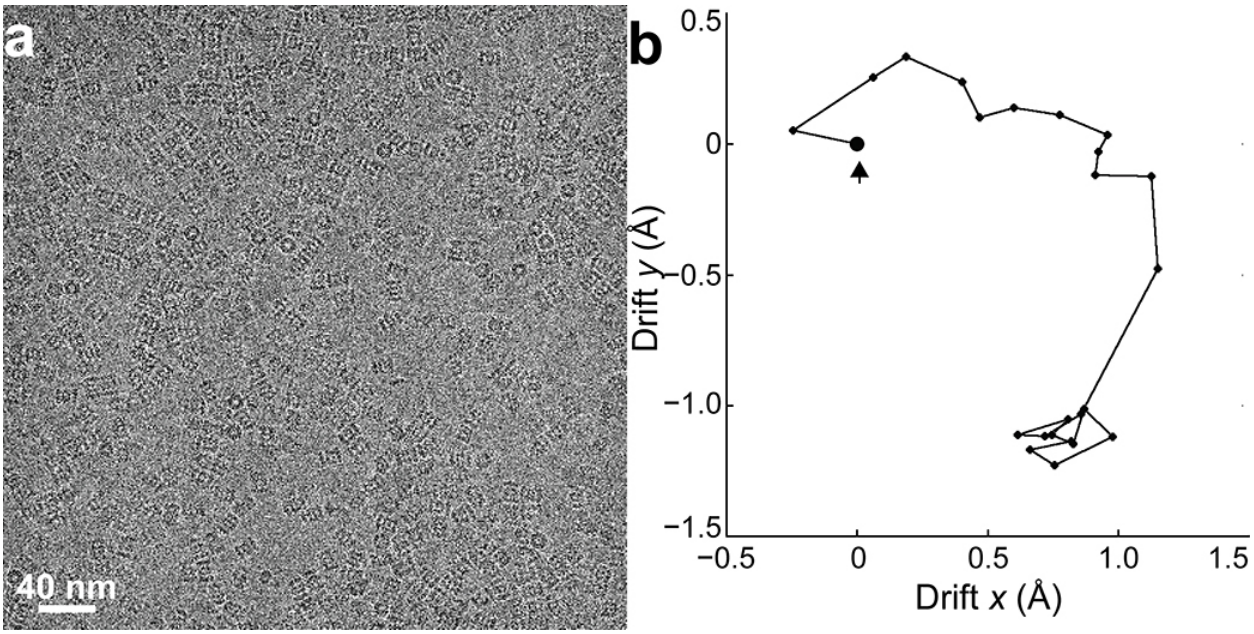
SUPPLEMENTARY FIGURES

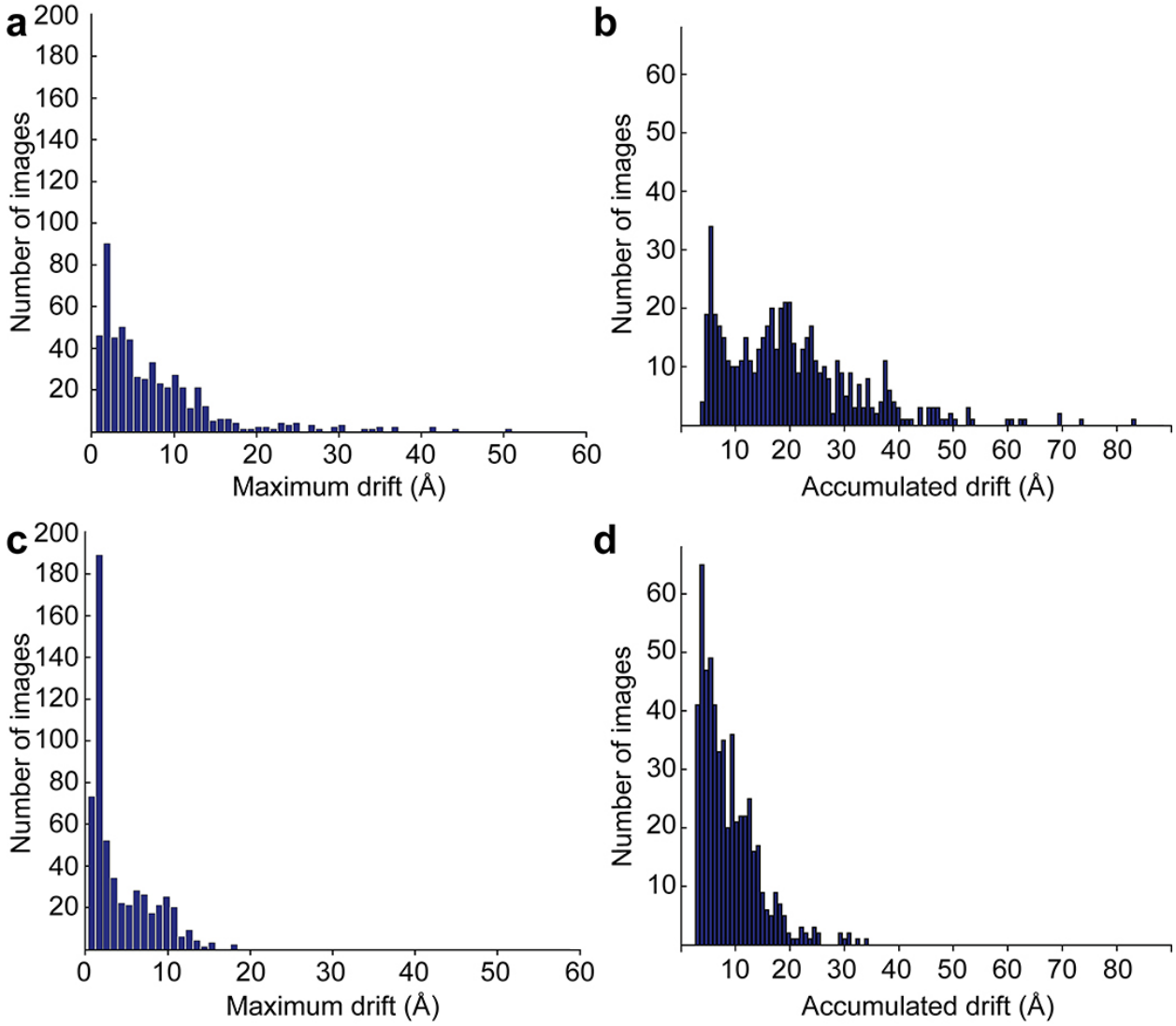
Supplementary Figure 1 *Image of frozen hydrated 20S proteasome and correction of small motion.*



a. A "near-perfect" image of frozen hydrated *T. acidophilum* 20S proteasome embedded in vitreous ice and recorded with a defocus of ~0.9 μm. Thon rings calculated from this image extend to ~3 Å (Fig. 2a in the main text). 20S proteasome particles embedded in vitreous ice are clearly visible with sufficient contrast. **b**. Enlarged view of Figure 2b in the main text showing a trace of frame displacements determined by our motion correction method. The total accumulated motion is less than 2 Å, or ~ 1 physical pixel. The trace shows that the initial small movement settles into a random motion. Correction of such small motions slightly improves the Thon ring amplitudes at close to 3 Å resolution (Fig. 2c in the main text).

Nature Methods: doi:10.1038/nmeth.2472

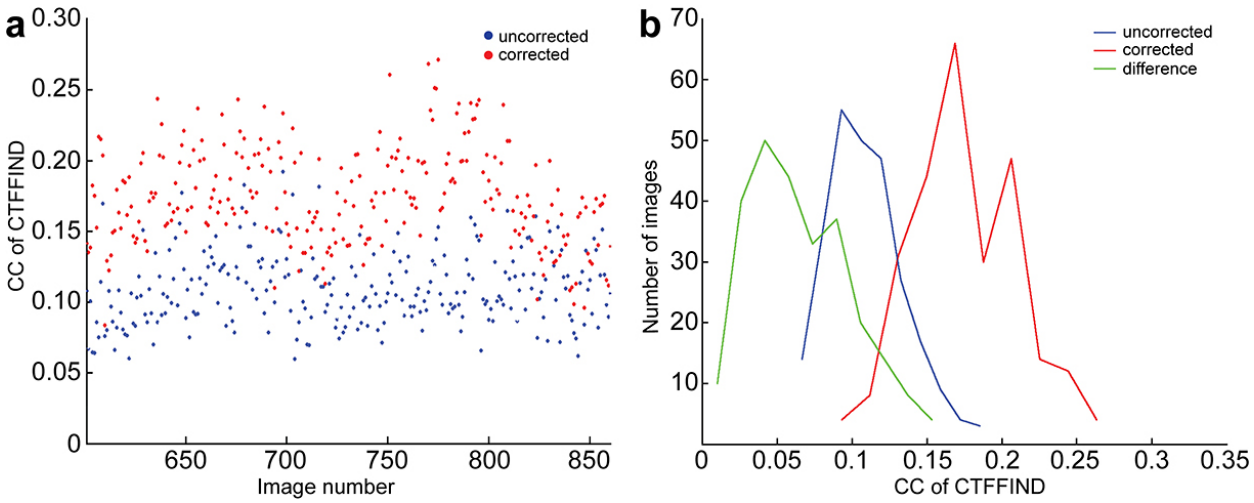
Supplementary Figure 2 *Histogram of maximum and total accumulated motion in each image.*



a. Histograms of the maximum image shift during the entire exposure time. **b**. Histogram of total accumulated motion during the entire exposure time. **c**. The same histogram as in a. after removing the first 4 subframes of each image. **d**. The same histogram as in b after removing the first 4 subframes from each image. Note that both maximum image shift and accumulated image motion are significantly reduced after removing the first 4 subframes indicating that the first 4 subframes contain the largest motion of the entire exposure time.

Nature Methods: doi:10.1038/nmeth.2472

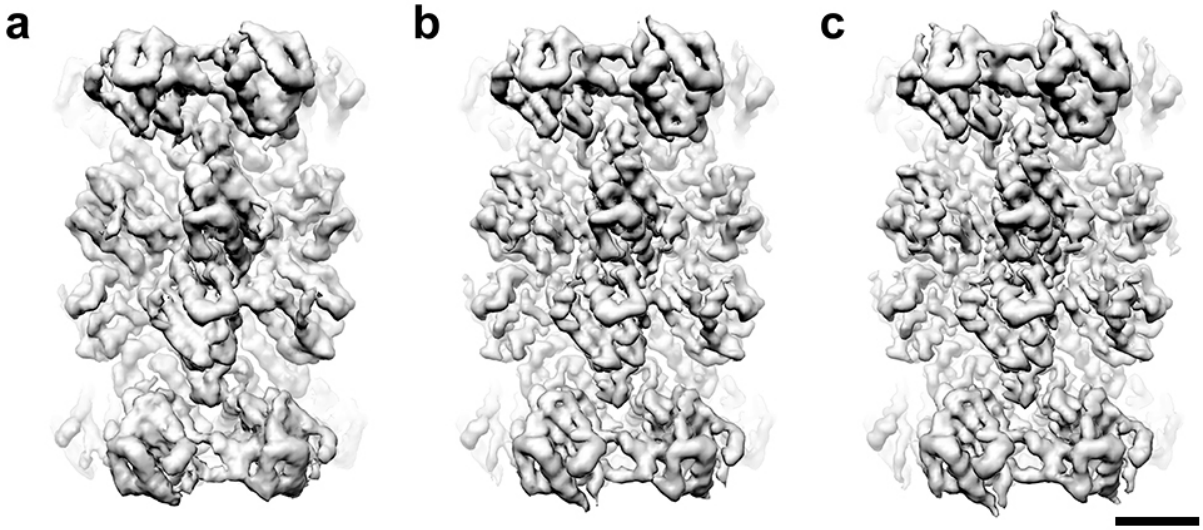
Supplementary Figure 3 *Evaluation of motion correction by Thon rings at high-resolution.*



We used the improvement of Thon ring at high-resolution to evaluate the benefits of motion correction. The cross correlation coefficient output from *CTFFIND*[1] was used to assess the Thon ring quality within the resolution range of 10 ~ 5 Å. A higher value indicates better Thon rings within this resolution range. **a**. Blue and red dots represent cross correlation coefficients of every image before and after motion correction, respectively. **b**. Histogram of cross correlation coefficients before (blue) and after (red) motion correction. The difference, which shows the amount of improvement, is shown in green. Note that the green histogram is everywhere positive, showing that all images are improved by motion correction. For "near perfect" images, the improvement by motion correction is small, but still beneficial.

Nature Methods: doi:10.1038/nmeth.2472

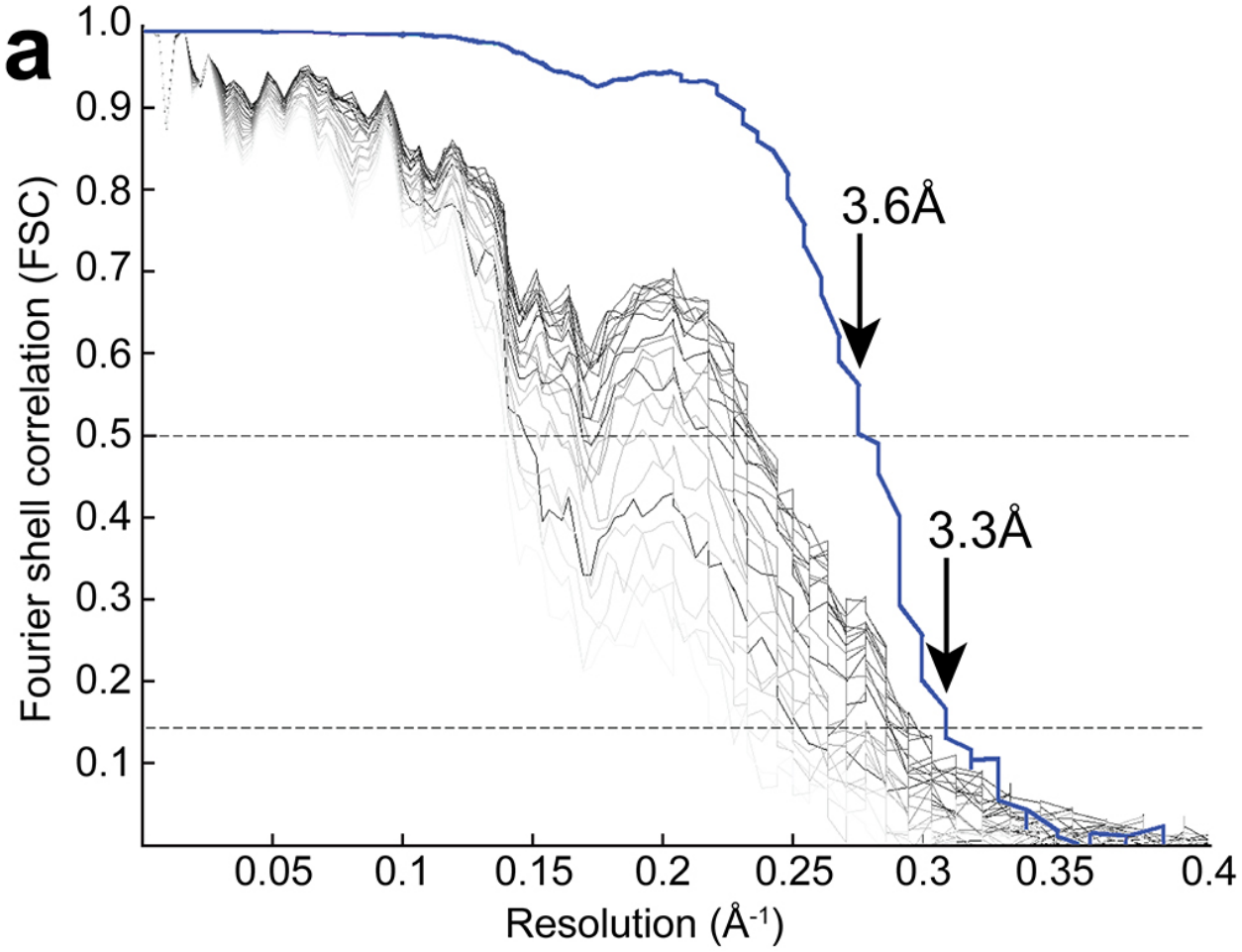
Supplementary Figure 4 *3D reconstructions of T. acidophilum 20S proteasome.*



a. 3D density map of *T. acidophilum* 20S proteasome determined from the uncorrected image dataset, corresponding to the red FSC curve of Fig. 3a. **b**. 3D density map of 20S proteasome after initial motion correction, corresponding to the green FSC curve. **c**. Final 3D density map, corresponding to the blue FSC curve. All density maps are shown here without filtering and/or sharpening. The scale bar is 3 nm.

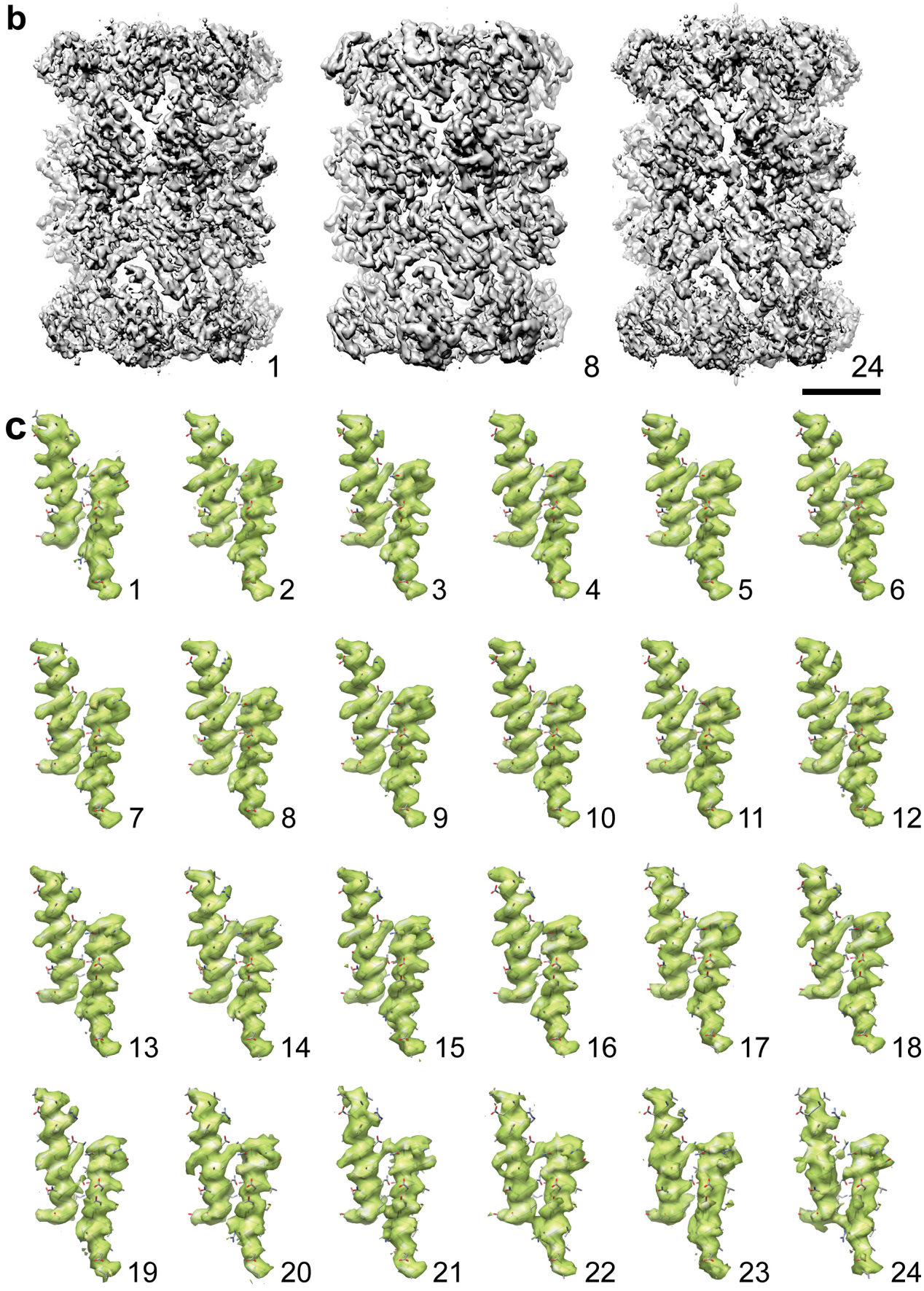
Nature Methods: doi:10.1038/nmeth.2472

Supplementary Figure 5 *3D reconstructions of T. acidophilum 20S proteasome calculated from individual subframes*.



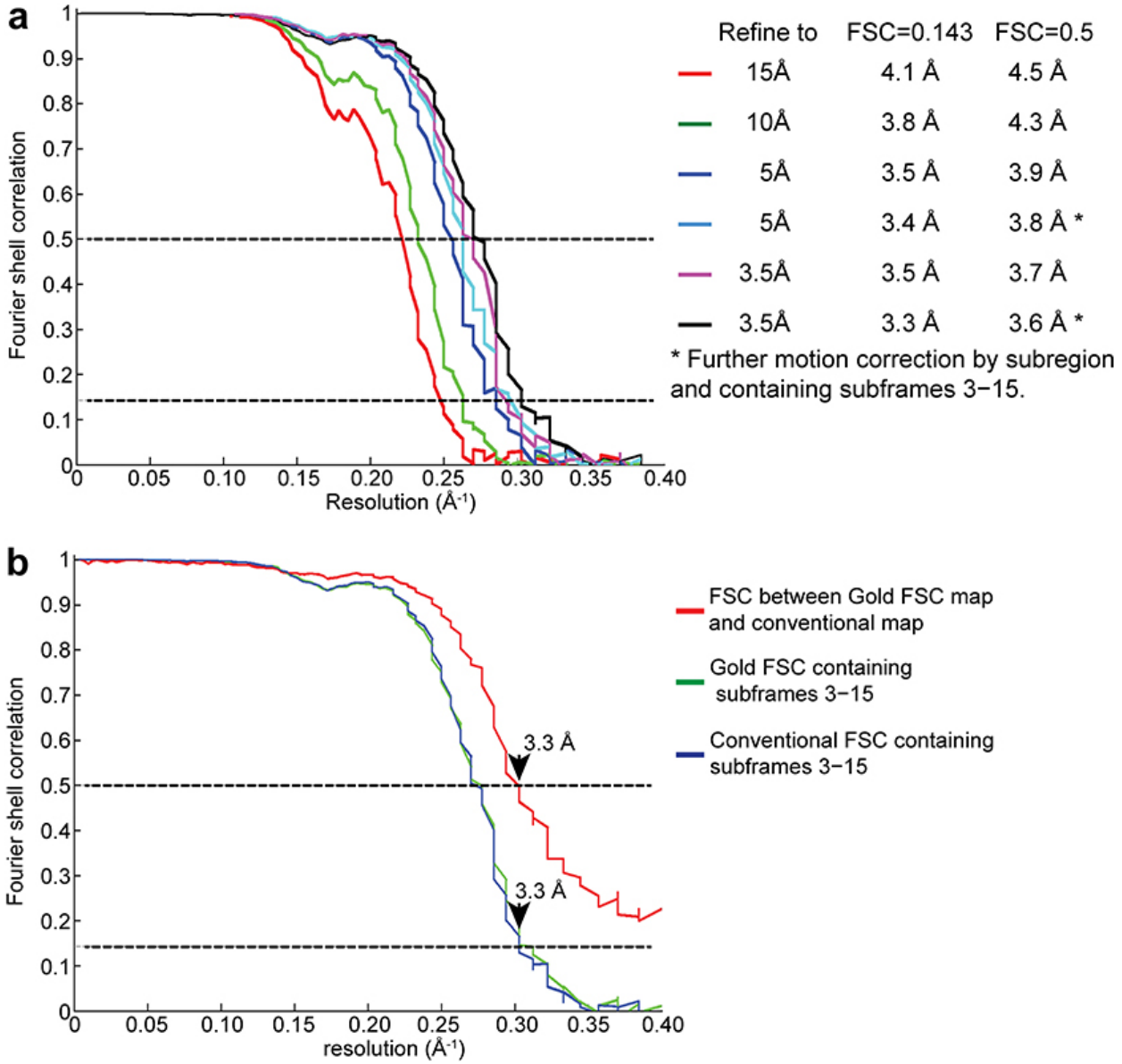
We calculated 24 3D reconstructions, each of which was calculated from a single subframe of every image. **a**. FSC curves of the 24 subframe 3D reconstructions. The gradually varied curves from the darkest to the lightest gray correspond to subframes ranging from 1 to 24. The blue FSC curve is from the final 3D reconstruction that uses subframes 3 – 15. **b**. Three selected density maps from subframes 1, 8 and 24, without amplitude scaling. The scale bar is 3nm. **c**. The amplitudes for each 3D reconstruction were scaled to match those calculated from the docked atomic structure. The same two α-helices were segmented from each 3D reconstruction and displayed sequentially. Despite the large motion, some side chain densities are visible in subframe 1. In the last few subframes, these side chain densities have disappeared due to radiation damage. Together, these figures show how motion induced image blurring and radiation damages captured by the set of dose fractionated subframes influence the 3D reconstruction.

Nature Methods: doi:10.1038/nmeth.2472



Nature Methods: doi:10.1038/nmeth.2472

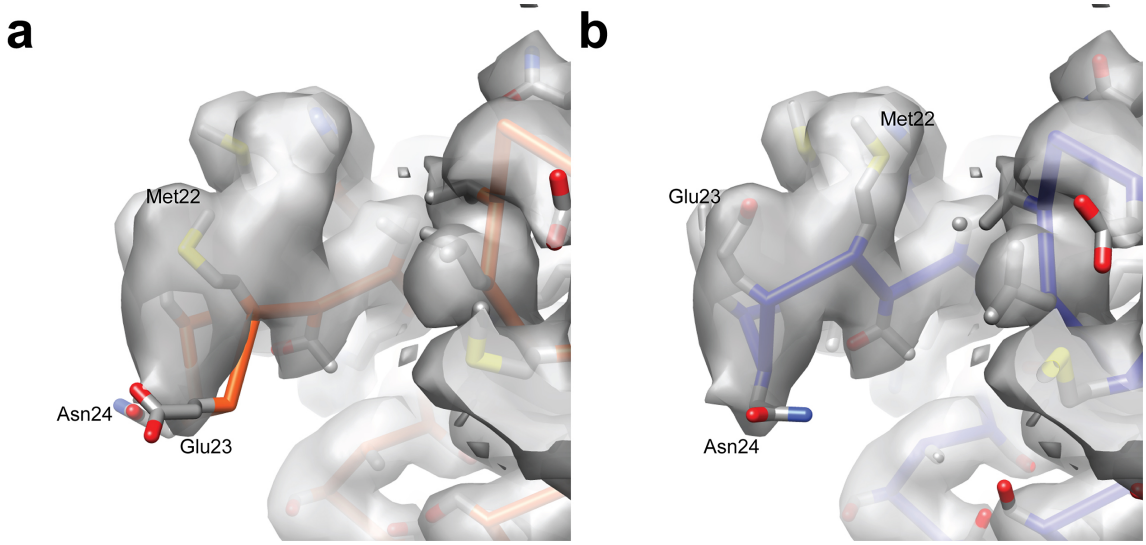
Supplementary Figure 6 *Resolution validation of the final 3D reconstruction.*



a. In each cycle, Euler angles and in-plane shifts of every particle in the entire dataset were determined by global search and/or local refinement against the same 3D reference map using GeFREALIGN[2] to the resolution limit as indicated. The new 3D reconstruction was always calculated to the Nyquist resolution and used for the next cycle without sharpening or low-pass filtering. These FSC curves show how the resolution of 3D reconstruction was improved as the parameters were refined towards higher resolution. **b**. We also processed the data following the "gold standard procedure"[3], by splitting the dataset into two half datasets that were then processed completely independently. The resolution estimated from the gold standard FSC is 3.3 Å at FSC = 0.143. The gold standard FSC curve (green) is very similar to the conventional FSC curve (blue, and black in a) indicating that the data is sufficiently good that there is a minimal tendency to overfitting. The gold standard map was calculated by averaging the two 3D reconstructions from two half datasets. The FSC curve calculated between the gold standard map and conventional map (red) shows that the two maps are the same up to a resolution of 3.3 Å. This comparison confirms the resolution estimated from the conventional FSC curve.

Nature Methods: doi:10.1038/nmeth.2472

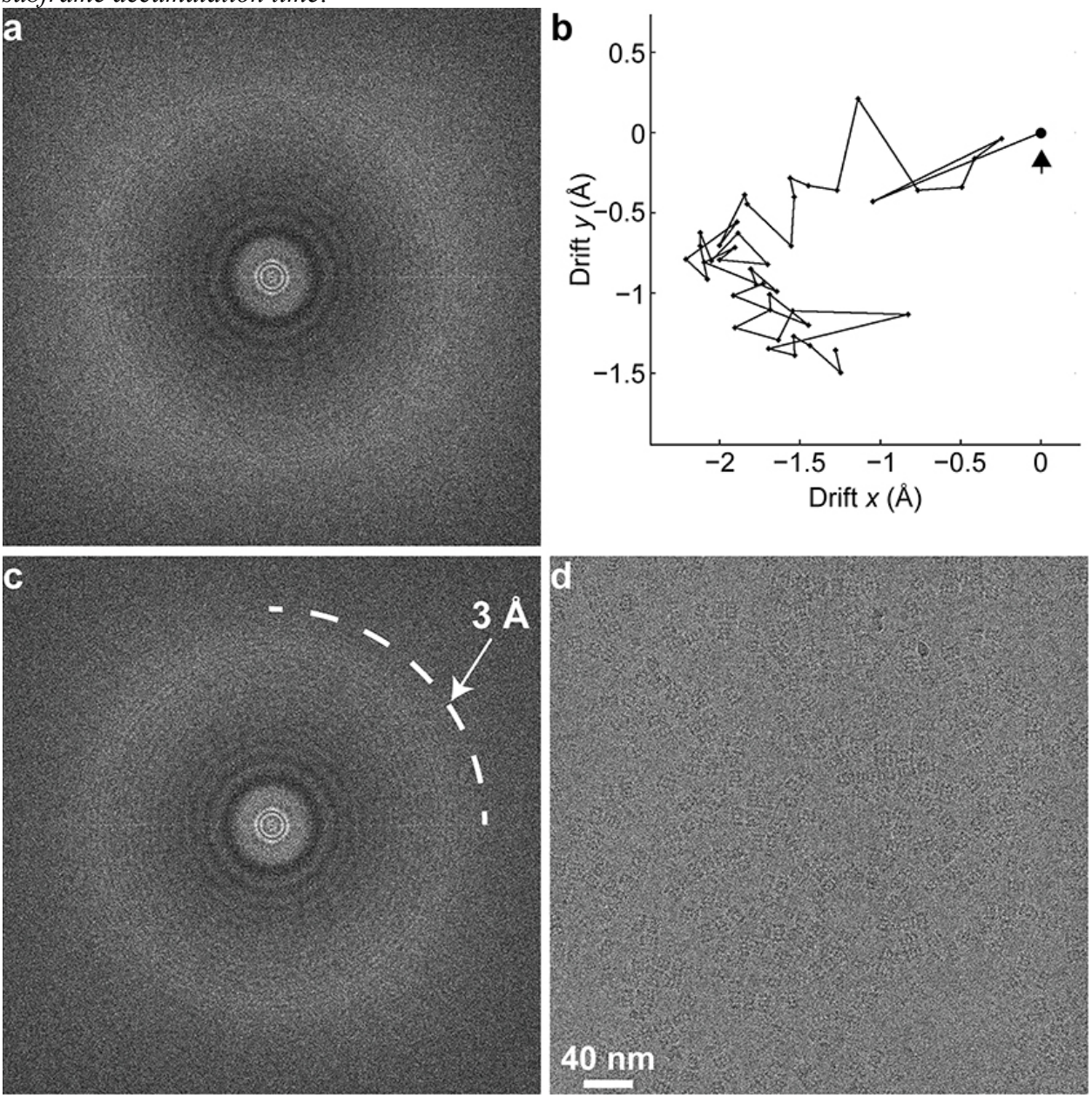
Supplementary Figure 7 *Refitting the loop (Met22 ~ Asn24) from the crystal structure into our 3.3 Å cryoEM density map.*



a. A short loop (Met22 ~ Asn24) in the β subunit from the crystal structure does not fit well into the final 3D density map, but is easily corrected. While not in a crystal contact, these residues do have much higher temperature factors in the 3.4 Å crystal structure than the average. **b**. We remodeled this loop to fit better into our 3.3 Å density map.

Nature Methods: doi:10.1038/nmeth.2472

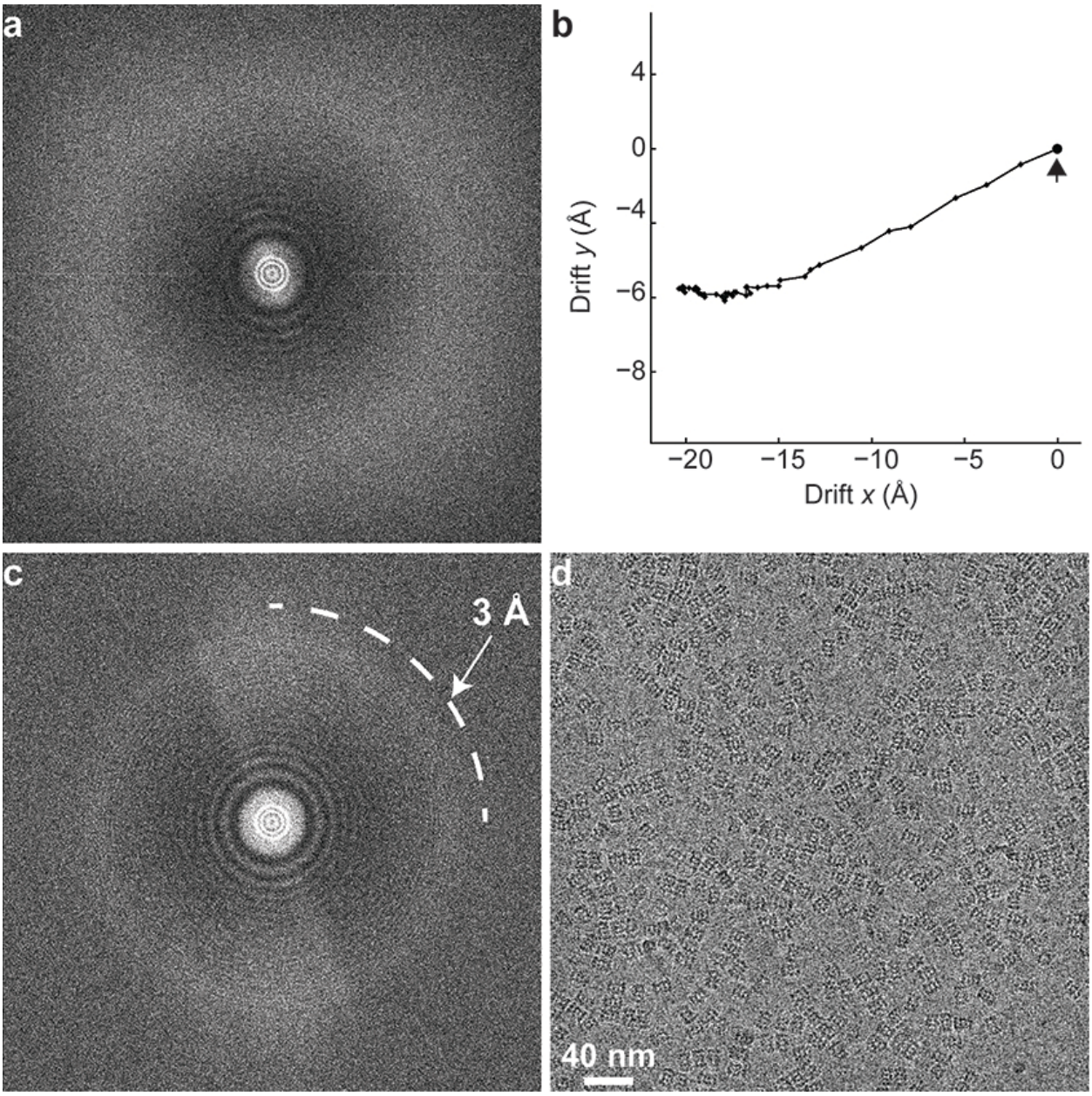
Supplementary Figure 8 *Accurate correction of small image motions with a 0.1 second subframe accumulation time.*



Motion correction was tested with images recorded a 5 sec exposure fractionated into 50 subframes (0.1 sec accumulation time). On average subframe contains ~0.8 count/pixel. A total of 44 subframes were used for motion correction. **a**. Fourier transform of un-corrected image. **b**. Trace of motion determined by our method, with the arrow pointing to the position of the first subframe. **c**. Fourier transform of corrected image. Thon rings at a high-resolution range (near ~3 Å) become clearer and stronger after motion correction. **d**. Corrected image of frozen hydrated archaeal 20S proteasome recorded at a defocus of -0.9 μm.

Nature Methods: doi:10.1038/nmeth.2472

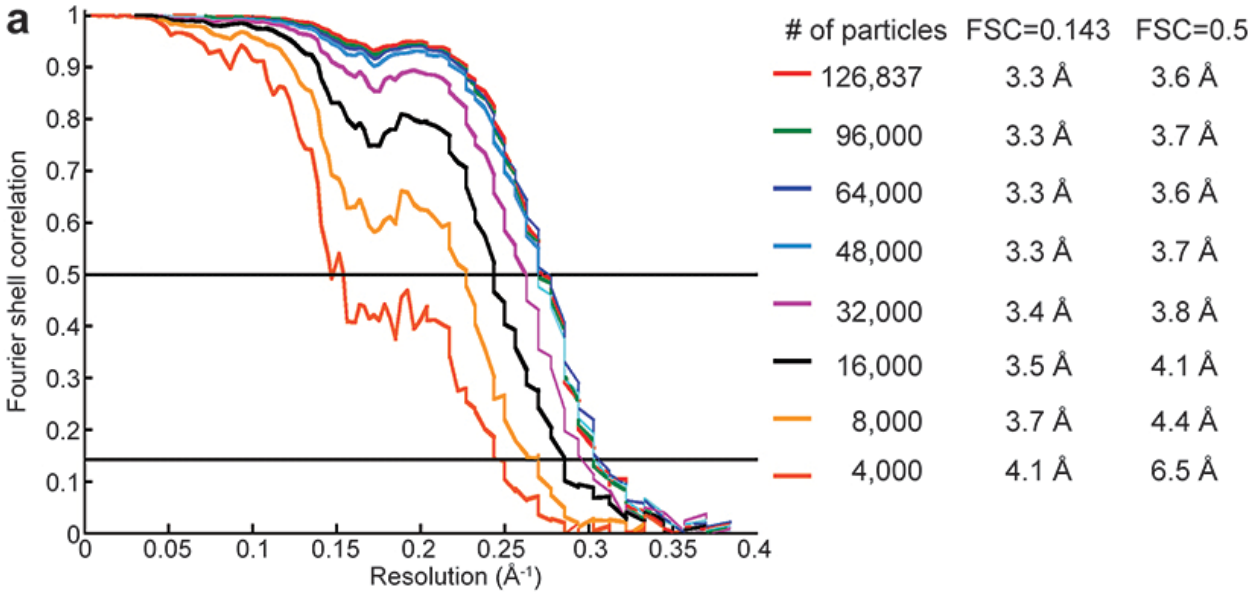
Supplementary Figure 9 *Correction of large image motions with a 0.1 second subframe accumulation time.*



Another example of motion correction using a 0.1 second subframes accumulation time, but this time for an image showing much larger beam induced motion. Same as in Supplementary Figure 9, this image was recorded with 5 sec exposure and fractionated into 50 subframes. **a**. Fourier transform of the uncorrected image showing a more significant degradation than in Supplementary Figure 9. **b**. Trace of motion determined by our method, with the arrow pointing to the position of the first subframe. **c**. Fourier transform of corrected image showing that high-resolution range (near ~3 Å) Thon rings are recovered. **d**. Corrected image of frozen hydrated archaeal 20S proteasome recorded at a defocus of -1.4μm.

Nature Methods: doi:10.1038/nmeth.2472

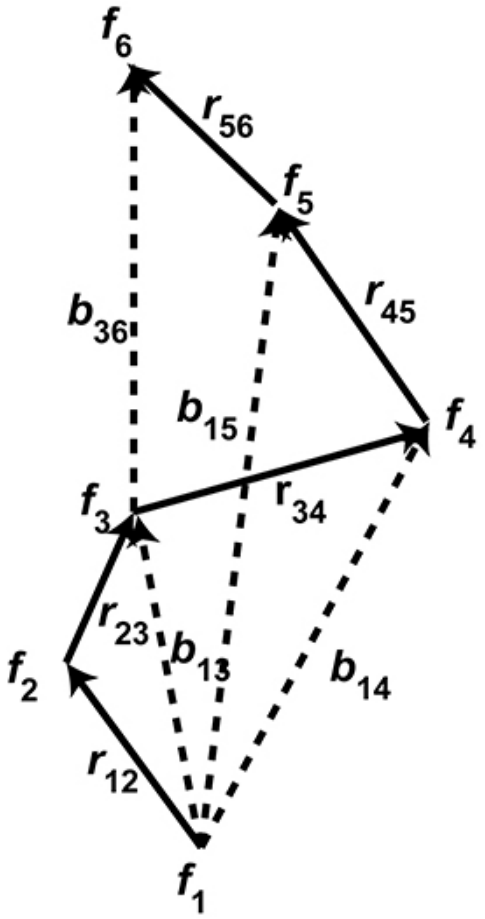
Supplementary Figure 10 *FSC curves of 3D reconstructions calculated from different number of particles.*



3D reconstructions were calculated using different number of particles taken from first portions of the original dataset. Note that in calculating each 3D reconstruction, the Euler angles and in-plane shifts for each particle are based on the final refinement using all particles.

Nature Methods: doi:10.1038/nmeth.2472

Supplementary Figure 11 *Illustration of redundant and accurate motion measurement algorithm.*



$$b_{13} = r_{12} + r_{23}$$

$$b_{14} = r_{12} + r_{23} + r_{34}$$

$$b_{15} = r_{12} + r_{23} + r_{34} + r_{45}$$

$$b_{16} = r_{12} + r_{23} + r_{34} + r_{45} + r_{56}$$

.....

$$b_{35} = r_{34} + r_{45}$$

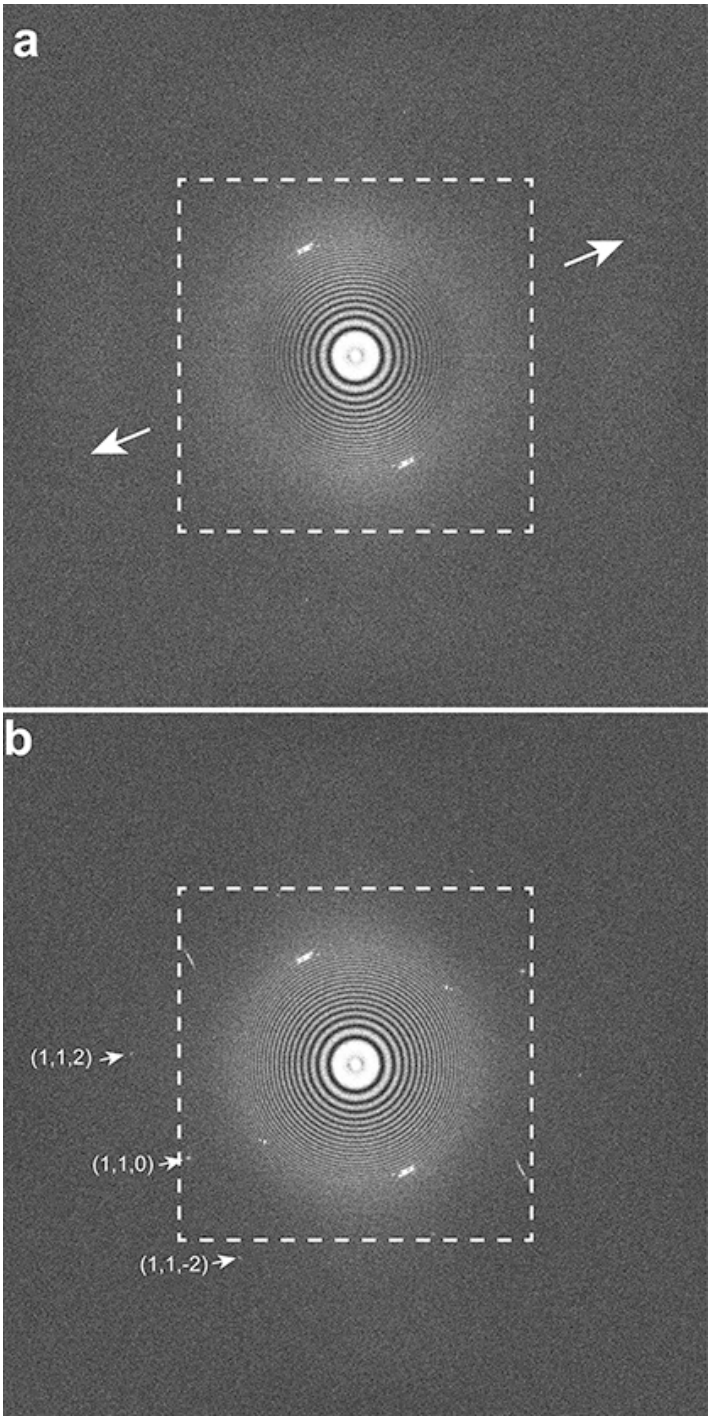
$$b_{36} = r_{34} + r_{45} + r_{56}$$

.....

When images are recorded as a stack of dose fractionated subframes, image motion during the total exposure can be described as a combination of sequential subframe displacements between every pair of adjacent subframes. Shown on the left is a diagram illustrating the principle that image motion between any two non-adjacent subframes f_i and f_j is the vector summation of all adjacent displacements in between these two subframes. The formula on the right illustrates that applying this relationship to all non-adjacent frames generates a set of over-determined linear equations, in which the total number of unknowns is smaller than the total number of equations.

Nature Methods: doi:10.1038/nmeth.2472

Supplementary Figure 12 *Motion correction restores high-resolution information.*



a. Fourier transform calculated from an electron counting super-resolution image taken from the carbon film of a Quantifoil grid containing frozen hydrated archaeal 20S proteasomes. A small contaminating ice crystal gives the sharp diffraction spots in the Fourier transform. The dashed square indicates the physical Nyquist limit. Note that there is weak direction in the diffraction pattern (indicated by arrows) in which a slight dampening of Thon ring amplitude was caused by a small image motion. **b**. After motion correction, Thon rings in the weak direction are restored and a number of diffraction spots from the hexagonal ice crystal at or beyond physical Nyquist (marked with indices) are also restored.

Nature Methods: doi:10.1038/nmeth.2472

Reference:

1 Mindell, J. A. & Grigorieff, N. Accurate determination of local defocus and specimen tilt in electron microscopy. *Journal of structural biology* 142, 334-347 (2003).

2 Li, X., Grigorieff, N. & Cheng, Y. GPU-enabled FREALIGN: accelerating single particle 3D reconstruction and refinement in Fourier space on graphics processors. *Journal of structural biology* 172, 407-412 (2010).

3 Scheres, S. H. & Chen, S. Prevention of overfitting in cryo-EM structure determination. *Nature methods* 9, 853-854 (2012).

Nature Methods: doi:10.1038/nmeth.2472