Supporting information

Could Egg White Lysozyme be Solved by Single Particle Cryo-EM?

Y. Zhang, R. Tammaro, P.J. Peters, R.B.G. Ravelli*

The Maastricht Multimodal Molecular Imaging Institute (M4I), Division of Nanoscopy,

Maastricht University, 6229 ER, Maastricht, the Netherlands

*Corresponding Author: Raimond Ravelli (rbg.ravelli@maastrichtuniversity.nl)

1. Theory

The expression of signal and noise is given by Henderson in 1995 [1]:

$$\text{Signal} = 2\sqrt{\frac{\langle I_{obs}\rangle}{I_0}}\sqrt{N_s} \tag{S1}$$

$$\text{Noise} = \frac{1}{\sqrt{N_e d^2/4}} \tag{S2}$$

Here, the term $\frac{\langle I_{obs}\rangle}{I_0}$ is the ratio between mean intensity $\langle I_{obs}\rangle$ and total incident beam intensity I_0. The physical meaning of this term is the information given by elastic scattering as the fraction of the incident beam.

$$\frac{\langle I_{obs}\rangle}{I_0} = \frac{f}{2N_s} \tag{S3}$$

where f is the fraction of electrons elastically to a resolution of d of the total elastically scattered electrons and N_s is the number of unique diffraction spots to resolution d.

$$f = \frac{\sigma_e(c)}{3}\left[\frac{N_c}{D^2}\right] \tag{S4}$$

$$N_s = \frac{\pi}{2}\left[\frac{D}{d}\right]^2 \tag{S5}$$

In here, σ_e is the elastic cross-section for carbon, $N_c = \frac{4}{3}\pi\left(\frac{D}{2}\right)^3\rho\frac{9.1}{110}$ is the number of carbon atom equivalents of a molecule, and ρ is the protein density which equals to $0.8\ \text{Da}/\text{Å}^3$. N_S is the number of unique diffraction spots to resolution d in projection. Combining equations (S3), (S4) and (S5), the ratio between mean intensity and total incident beam is:

$$\frac{\langle I_{obs}\rangle}{I_0} = \frac{\sigma_e(c)4}{3\ 3}\pi\left(\frac{D}{2}\right)^3\rho\frac{9.1}{110}\left(\frac{1}{D^2}\right)\frac{1}{\pi}\left(\frac{d}{D}\right)^2 \tag{S6}$$

As the signal calculated from the mean intensity of the diffracted beam as a fraction of the incident beam, we substitute d in equation (S6) to d_N to indicate the resolution at the Nyquist frequency.

$$\frac{\langle I_{obs}\rangle}{I_0} = \frac{\sigma_e(c)\ d_N^2}{272\ D} \tag{S7}$$

By substituting (S5) and (S7) into (S1), the signal from the particle is obtained as:

$$\text{Signal} = \sqrt{\frac{\sigma_e(c)\ \pi\ D\ d_N^2}{136\ d^2}} \tag{S8}$$

The expression of X_{sig} is given by equation (S9), which is the signal to noise ratio of the image. It describes whether the molecule can be detector not.

$$X_{sig} = \sqrt{\frac{\sigma_e\ \pi\ N_e\ D^3 d_N^2}{136\ d^2}} \tag{S9}$$

For given electron energy, electron dose, diameter of the particles and Nyquist frequency, the smaller the value of d, the larger the value of X_{sig}. The largest X_{sig} happens at the Nyquist frequency. So, X_{sig} at Nyquist is

$$X_{sig} = \sqrt{\frac{\sigma_e\ \pi\ N_e\ D^3}{136}} \tag{S10}$$

The number of particles that need to build a density map with a certain resolution is given by [2]:

$$\mathrm{N_{part}} = \mathrm{N_{inproj}}\ \left(\frac{\pi \mathrm{D}}{\mathrm{N_{asym}} \mathrm{d}}\right) \mathrm{e}^{\mathrm{B}/2\mathrm{d}^2} \tag{S11}$$

Where $\mathrm{N_{inproj}}$ is the number of images that need per projection. The term $\frac{\pi \mathrm{D}}{\mathrm{d}}$ is followed by the Crowther criterion [3] which describes the minimum number of unique projections needed for reconstructing a particle of diameter D to a resolution of d. $\mathrm{N_{asym}}$ is the asymmetric unit that a molecule has. B is the temperature factor that describes the effect of contrast loss.

$$\mathrm{N_{inproj}} = \frac{\langle \mathrm{S} \rangle^2 \Big/ \langle \mathrm{N} \rangle^2}{\frac{\langle \mathrm{I_{obs}} \rangle}{\mathrm{I_0}} \mathrm{D}^2 \mathrm{N}_e} \tag{S12}$$

Based on gold-standard Fourier shell correlation [2,4], $\langle \mathrm{S} \rangle / \langle \mathrm{N} \rangle = 1/\sqrt{3}$ when $\mathrm{FSC} = 0.143$ and the map is interpretable. Thus, the number of particles needed to reach a certain resolution d is:

$$N_{part} = \frac{1}{N_{asym}} \left[\frac{90\,\pi}{\sigma_e\, N_e\, {d_N}^2\, d}\right] e^{B/2d^2} \tag{S13}$$

2. Supplementary figures

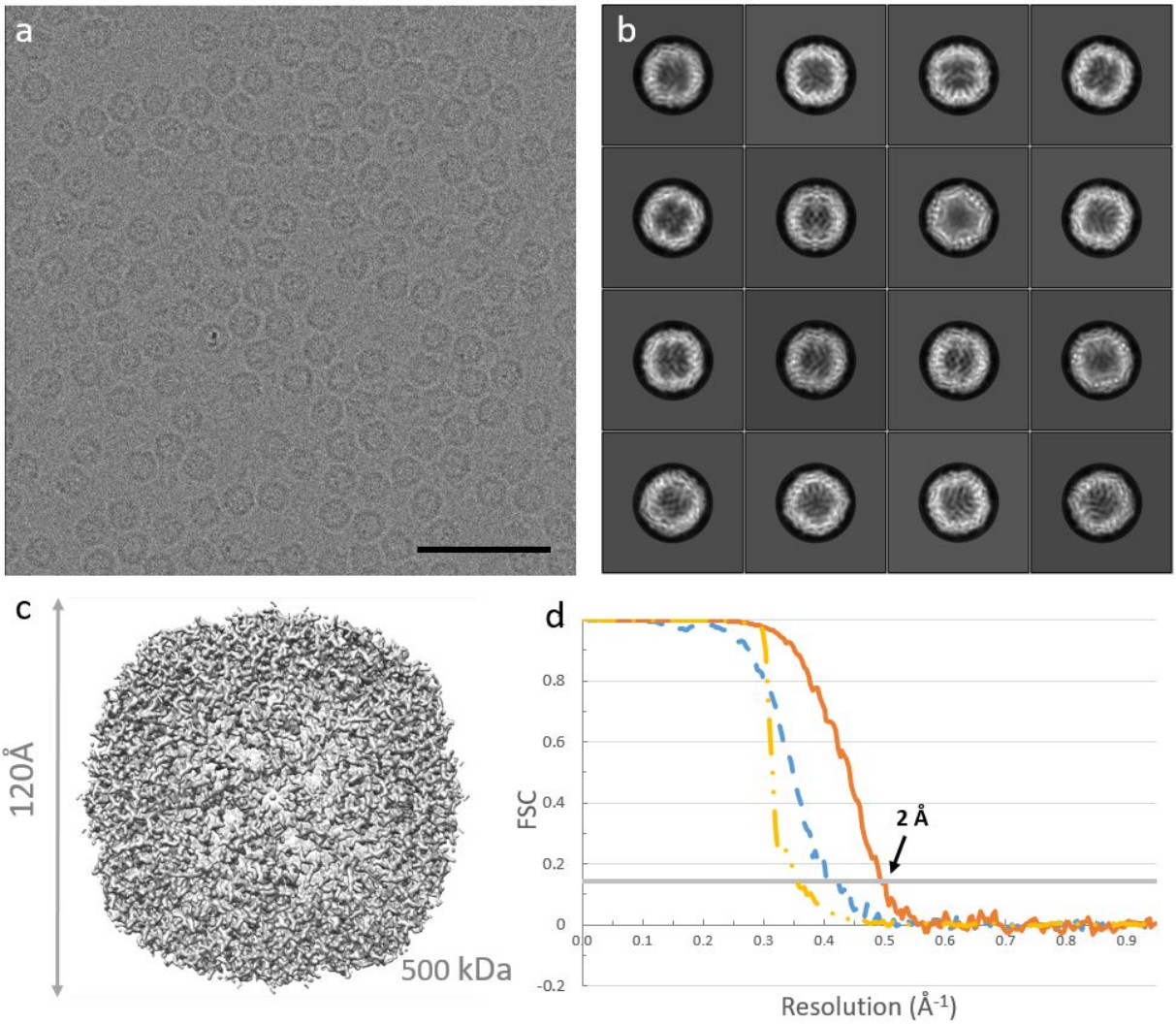

Figure S1. Single particle analysis of a subset of the experimental mouse apo-ferritin dataset from EMPIAR-10216. (a) A typical micrograph of apo-ferritin, the scale bar is 50 nm, (b) 2D class averages. (c) 3D reconstruction from 29224 particles at 2 Å resolution. (d) Gold-standard Fourier Shell correlation (FSC) before (blue line) and after (orange line) masking, and the phase randomized FSC (yellow line).

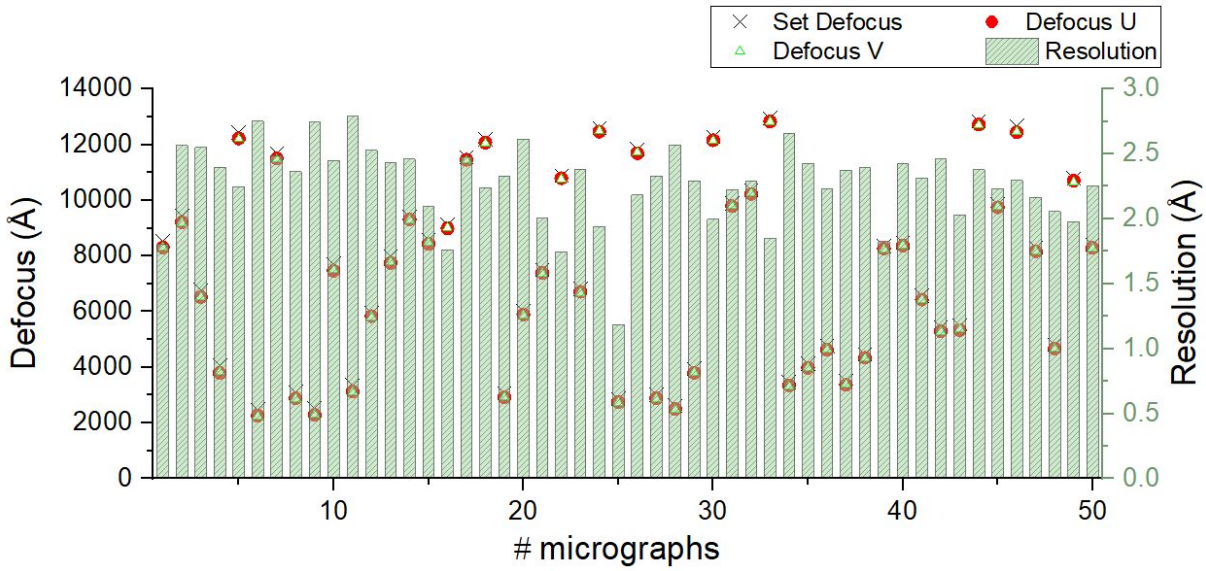

Figure S2. CTFs fitted with Gctf. 50 simulated micrographs for apo-ferritin have resolution higher than 3 Å.

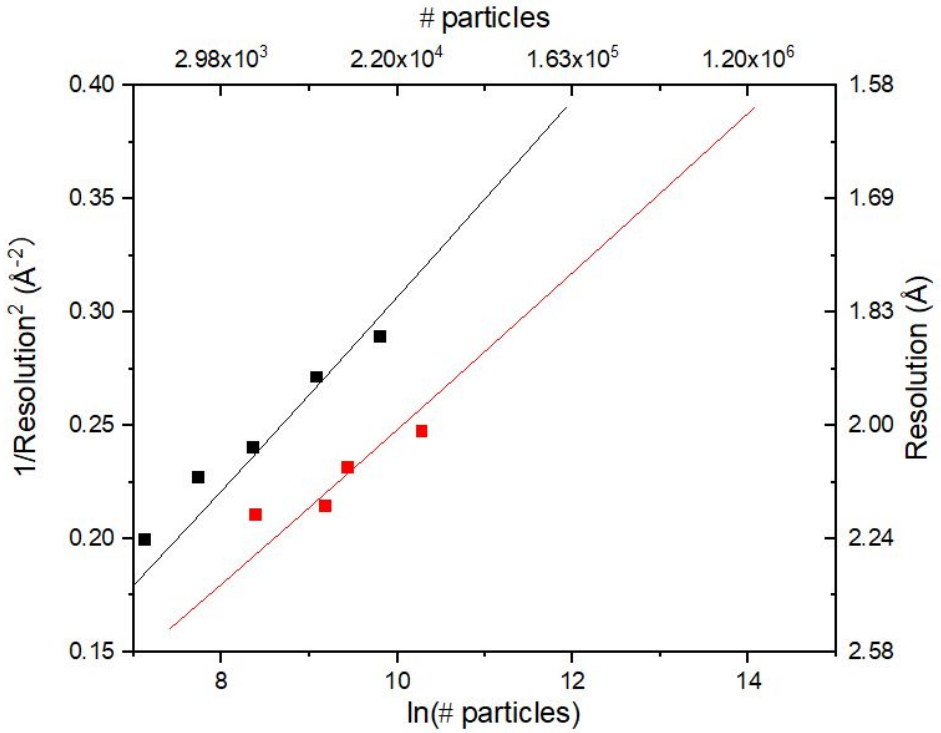

Figure S3. Number of particles and resolution plot based on equation (13). The B-factor for simulated data is 43 Å^2(black), according to least squares regression line fitting. The fitted B-factor for the experimental data is 54 Å^2 (red).

References

(1) Henderson, R. The Potential and Limitations of Neutrons, Electrons and X-Rays for Atomic Resolution Microscopy of Unstained Biological Molecules. *Q. Rev. Biophys.* **1995**, *28*, 171–193. https://doi.org/10.1017/S003358350000305X.

(2) Rosenthal, P. B.; Henderson, R. Optimal Determination of Particle Orientation, Absolute Hand, and Contrast Loss in Single-Particle Electron Cryomicroscopy. *J. Mol. Biol.* **2003**. https://doi.org/10.1016/j.jmb.2003.07.013.

(3) R . A . Crowther, D. . J. . D. and A. . K. S. The Reconstruction of a Three-Dimensional Structure from Projections and Its Application to Electron Microscopy. *Proc. R. Soc. London* **1970**, 319–340.

(4) Chen, S.; McMullan, G.; Faruqi, A. R.; Murshudov, G. N.; Short, J. M.; Scheres, S. H. W.; Henderson, R. High-Resolution Noise Substitution to Measure Overfitting and Validate Resolution in 3D Structure Determination by Single Particle Electron Cryomicroscopy. *Ultramicroscopy* **2013**. https://doi.org/10.1016/j.ultramic.2013.06.004.