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

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

ABSTRACT: The combination of high-end cryogenic transmission electron microscopes (cryo-EM), direct electron detectors, and advanced image algorithms allows researchers to obtain the 3D structures of much smaller macromolecules than years ago. However, there are still major challenges for the single-particle cryo-EM method to achieve routine structure determinations for macromolecules much smaller than 100 kDa, which are the majority of all plant and animal proteins. These challenges include sample characteristics such as sample heterogeneity, beam damage, ice layer thickness, stability, and quality, as well as hardware limitations such as detector performance, beam, and phase plate quality. Here, single particle data sets were simulated for samples that were ideal in terms of homogeneity, distribution, and stability, but with realistic parameters for ice layer, dose, detector performance, and beam characteristics. Reference data were calculated for human apo-ferritin using identical parameters reported for an experimental data set downloaded from EMPIAR. Processing of the simulated data set resulted in a value of 1.86 Å from 20 214 particles, similar to a 2 Å density map obtained from 29 224 particles selected from real micrographs. Simulated data sets were then generated for a 14 kDa protein, hen egg white lysozyme (HEWL), with and without an ideal phase plate (PP). Whereas we could not obtain a high-resolution 3D reconstruction of HEWL for the data set without PP, the one with PP resulted in a 2.78 Å resolution density map from 225 751 particles. Our simulator and simulations could help in pushing the size limits of cryo-EM.

INTRODUCTION

Determining 3D structures from sub-100-kDa protein molecules has been a long-standing goal of the cryo-EM community, for which the first successes have been reported. Analysis of the size distribution of proteins in plants, animals, and fungal and microbial species shows that 90% of the plant and animal proteins are smaller than 100 kDa, and more than 50% of the proteins are smaller than 50 kDa. However, resolving a small (monomeric) protein in the electron microscope is difficult for multiple reasons. Electrons scatter much more compared to X-rays and neutrons, resulting in high amounts of energy being deposited in the sample for a relative small number of electrons per Å² hitting the sample. Globular proteins of <50 kDa are smaller than 5 nm in size, whose signal can be easily swamped under thicker ice layers and ice imperfections. Protein molecules are likely to overlap when present in the bulk of the ice layer and can become (partly) damaged and preferentially oriented when attaching to the air–water interface. Sample heterogeneity, imperfect detectors, imperfect phase plates and optics, stage drift, and beam-induced motions all contribute to blurring of averages, which makes it challenging to achieve near-atomic-resolution structures of small macromolecules.

The development of direct electron detectors with more advanced hardware and advanced image algorithms allows researchers to obtain the 3D structures of much smaller macromolecules than years ago. However, there are still major challenges for the single-particle cryo-EM method to achieve routine structure determinations for macromolecules much smaller than 100 kDa, which are the majority of all plant and animal proteins. These challenges include sample characteristics such as sample heterogeneity, beam damage, ice layer thickness, stability, and quality, as well as hardware limitations such as detector performance, beam, and phase plate quality. Here, single particle data sets were simulated for samples that were ideal in terms of homogeneity, distribution, and stability, but with realistic parameters for ice layer, dose, detector performance, and beam characteristics. Reference data were calculated for human apo-ferritin using identical parameters reported for an experimental data set downloaded from EMPIAR. Processing of the simulated data set resulted in a value of 1.86 Å from 20 214 particles, similar to a 2 Å density map obtained from 29 224 particles selected from real micrographs. Simulated data sets were then generated for a 14 kDa protein, hen egg white lysozyme (HEWL), with and without an ideal phase plate (PP). Whereas we could not obtain a high-resolution 3D reconstruction of HEWL for the data set without PP, the one with PP resulted in a 2.78 Å resolution density map from 225 751 particles. Our simulator and simulations could help in pushing the size limits of cryo-EM.

Special Issue: Frontiers in Cryo-EM Modeling

Received: December 20, 2019
Published: March 23, 2020
Spatial and temporal incoherence of the electron source dampens the high-resolution signal under strong underfocus conditions, which impedes reconstruction schemes. Researchers are trying different ways to overcome these limitations, such as high protein concentration, minimal ice thickness, and the use of scaffolds. A particularly promising route for small proteins seems to be the use of a phase plate (PP). An ideal PP would offer a π/2 phase shift without introducing any post-sample scattering. The Volta Phase Plate has made phase plates accessible to a wide community; however, it cannot offer a constant phase shift and generates undesired postsample scattering. Ongoing research aims to develop next-generation phase plates, such as a laser PP or electrostatic PP. Lower energy electron microscopy was proposed by Henderson et al., as lower energy gives more information and has a somewhat better elastic/inelastic scattering ratio. Due to the faster damping of CTF at high spatial frequency at low energy, larger data sets and careful computational analysis would be required to recover high-resolution information.

Simulations could help identify and characterize the culprits preventing the 3D structural resolution of small proteins by single-particle cryo-EM. Furthermore, simulations facilitate the assessment of the new image processing methods and data collection techniques and could be used to evaluate the potential of new instrumentation improvements. Here, we adapted a TEM simulator developed by Vulović et al., which is based on physical principles and considers the interaction between solvent, ions, and molecules. The simulator considers electron dose, which is important for the TEM of biological samples; however, it cannot offer a constant phase shift and generates undesired postsample scattering. Ongoing research aims to develop next-generation phase plates, such as a laser PP or electrostatic PP. Lower energy electron microscopy was proposed by Henderson et al., as lower energy gives more information and has a somewhat better elastic/inelastic scattering ratio. Due to the faster damping of CTF at high spatial frequency at low energy, larger data sets and careful computational analysis would be required to recover high-resolution information.

Simulations could help identify and characterize the culprits preventing the 3D structural resolution of small proteins by single-particle cryo-EM. Furthermore, simulations facilitate the assessment of the new image processing methods and data collection techniques and could be used to evaluate the potential of new instrumentation improvements. Here, we adapted a TEM simulator developed by Vulović et al., which is based on physical principles and considers the interaction between solvent, ions, and molecules. The simulator considers electron dose, which is important for the TEM of biological samples; however, it cannot offer a constant phase shift and generates undesired postsample scattering. Ongoing research aims to develop next-generation phase plates, such as a laser PP or electrostatic PP. Lower energy electron microscopy was proposed by Henderson et al., as lower energy gives more information and has a somewhat better elastic/inelastic scattering ratio. Due to the faster damping of CTF at high spatial frequency at low energy, larger data sets and careful computational analysis would be required to recover high-resolution information.

Simulations could help identify and characterize the culprits preventing the 3D structural resolution of small proteins by single-particle cryo-EM. Furthermore, simulations facilitate the assessment of the new image processing methods and data collection techniques and could be used to evaluate the potential of new instrumentation improvements. Here, we adapted a TEM simulator developed by Vulović et al., which is based on physical principles and considers the interaction between solvent, ions, and molecules. The simulator considers electron dose, which is important for the TEM of biological samples; however, it cannot offer a constant phase shift and generates undesired postsample scattering. Ongoing research aims to develop next-generation phase plates, such as a laser PP or electrostatic PP. Lower energy electron microscopy was proposed by Henderson et al., as lower energy gives more information and has a somewhat better elastic/inelastic scattering ratio. Due to the faster damping of CTF at high spatial frequency at low energy, larger data sets and careful computational analysis would be required to recover high-resolution information.

Simulations could help identify and characterize the culprits preventing the 3D structural resolution of small proteins by single-particle cryo-EM. Furthermore, simulations facilitate the assessment of the new image processing methods and data collection techniques and could be used to evaluate the potential of new instrumentation improvements. Here, we adapted a TEM simulator developed by Vulović et al., which is based on physical principles and considers the interaction between solvent, ions, and molecules. The simulator considers electron dose, which is important for the TEM of biological samples; however, it cannot offer a constant phase shift and generates undesired postsample scattering. Ongoing research aims to develop next-generation phase plates, such as a laser PP or electrostatic PP. Lower energy electron microscopy was proposed by Henderson et al., as lower energy gives more information and has a somewhat better elastic/inelastic scattering ratio. Due to the faster damping of CTF at high spatial frequency at low energy, larger data sets and careful computational analysis would be required to recover high-resolution information.
The diameter of the particle is represented by $D$, and the resolution is expressed by $d$, both measured in Å.

The expression of $x$ is given by the inverse complementary error function, which represents the standard deviation in Gaussian distribution. In other words, it is the lowest signal-to-noise ratio required for particle alignment. The cross-correlation coefficient for translation (shifts $x$ and $y$) and rotation (Euler angles $\alpha$, $\beta$, and $\gamma$) are $(2D/0.2d)^2$ and $(2\pi D/0.2d)^2$, respectively. The value of $x$ will not change much. The key parameter that determines whether a single particle can be aligned is $X_{\text{sig}}$. The derivation of the equation according to Henderson can be found in the Supporting Information. Here, the expression of $X_{\text{sig}}$ is given by

$$X_{\text{sig}} = \sqrt{\frac{\alpha \pi N D^3 d_0^2}{136 d^2}}$$

For a given electron energy, electron dose, diameter of the particles, and Nyquist frequency, a smaller number of $d$ results in a larger value of $X_{\text{sig}}$. Therefore, the largest $X_{\text{sig}}$ happens at the Nyquist frequency. $X_{\text{sig}}$ at Nyquist is

$$X_{\text{sig}} = \sqrt{\frac{\alpha \pi N D^3}{136}}$$

This means for a microscope at a fixed energy and electron dose, the threshold of detecting a molecule is fixed. The value of $N_e$ used in the calculations of Henderson was $5 \text{ e}^{-/\text{Å}^2}$, with which he arrived at a 38 kDa theoretical size limit. Assuming a spherical protein with a density of 0.8 Da/Å$^3$, a 38 kDa protein has a diameter of 45 Å. With the incoming $5 \text{ e}^{-/\text{Å}^2}$ electron fluence, this protein is not able to be aligned as $X_{\text{sig}}$ equals 6.5, which is smaller than the value of $x$ which is 8.3. Henderson chose the value of $N_e$ based on radiation damage studies using electron diffraction. Recently, it was shown that the best microelectron diffraction data were obtained from lysozyme crystals at a fluence of 2.6 $\text{ e}^{-/\text{Å}^2}$. However, in most modern SPA cryo-EM studies, the fluence is much higher than $5 \text{ e}^{-/\text{Å}^2}$. Dose-correction schemes account for the loss of signal at higher spatial frequencies as a function of dose. The first frames typically contain less high-resolution information, as one would expect based on the relatively pristine state of the biomolecule. This is probably due to beam-induced motions, whereas the latter frames within a movie contain less high-resolution features due to radiation damage to the particle of interest. For the low frequencies, all the frames contribute more or less the same. Overall, for imaging, an electron fluence greater than $5 \text{ e}^{-/\text{Å}^2}$ may still contribute to the signal up to Nyquist and surely helps with determining particle orientation and translation.$^{25}$ In eq 12, when we use $50 \text{ e}^{-/\text{Å}^2}$ for $N_e$:

$$X_{\text{sig}} = 0.068 \times D^{3/2}$$

If we consider HEWL with a diameter of 32 Å, then $X_{\text{sig}}$ equals 12.3. Using a pixel size of 0.5 Å/pixel, $d$ at Nyquist frequency is 1 Å. So, $x$ equals 8 and $X_{\text{sig}} > x$. This means that in theory we should be able to align the 14 kDa particles with a perfect detector and perfect image contrast. If we consider the contrast $C$ of the micrographs,$^5$ which varies from 0 to 1, the signal in eq 9 needs to be multiplied by $C$. Then, the question arises: can we still align particles as small as 14 kDa with current TEMs and detectors? Or would one need a(n ideal) phase plate in order to increase contrast $C$ and obtain a meaningful alignment?
Number of Particles Needed to Reach a Certain Resolution. The number of particles required to build a density map with a certain resolution is given by Rosenthal and Henderson:

\[
N_{\text{part}} = N_{\text{inproj}} \left( \frac{\pi D}{N_{\text{sym}} d} \right)^{8/2d^2}
\]  

(14)

where \(N_{\text{inproj}}\) is the number of images needed per projection. The term \((\pi D)/d\) is given by the Crowther criterion,\(^32\) which describes the minimum number of unique projections needed for reconstructing a particle of diameter \(D\) to a resolution of \(d\). \(N_{\text{sym}}\) is the number of asymmetric units that a molecule has. \(B\) is the temperature factor that describes the effect of contrast loss. Using an expression for \(N_{\text{inproj}}\), we get (Supporting Information and ref 25):

\[
N_{\text{part}} = N_{\text{inproj}} \left[ \frac{90\pi}{N_{\text{sym}} d N_{\text{part}}^2} \right]^{8/2d^2}
\]  

(15)

In order to determine what resolution one might expect given a certain number of particles, we take the natural logarithm of both sides in eq 15 and rewrite it:

\[
\frac{1}{d^2} = \frac{2}{B} \times \left[ \ln(N_{\text{part}}) - \ln \left( \frac{90\pi}{N_{\text{sym}} d N_{\text{part}}^2} \right) + \ln(d) \right]
\]  

(16)

In eq 16, \(\ln(d) \ll \ln(N_{\text{part}})\), therefore, we ignore \(\ln(d)\). Thus, the final expression is

\[
\frac{1}{d^2} = \frac{2}{B} \times \left[ \ln(N_{\text{part}}) - \ln \left( \frac{90\pi}{N_{\text{sym}} d N_{\text{part}}^2} \right) \right]
\]  

(17)

### METHODS AND RESULTS

**Image Formation.** Interaction potentials of the specimen were built using IASA. Nonoverlapping particles were randomly oriented and positioned to simulate micrographs of 4096 × 4096 pixels. A thin ice layer of 20 nm was used, and the particles were randomly positioned in all three dimensions within the ice layer. The electron propagation through the specimen was simulated via the multislice method. We simulated a 300 kV FEG TEM with a Falcon III\(^33\) detector used in counting mode. For the spherical and chromatic aberrations, a value of 2.7 mm was used and 4.7 mm for the focal distance. These parameters are typical for a Thermo Fisher Krios microscope. The size of the illumination aperture was 0.03 mrad, and the diameter of the objective aperture was 100 \(\mu\)m. We did not include objective astigmatism in our simulations. We used the DQE and MTF of Falcon III electron counting (EC) mode at 300 keV as given by Kuijper et al.\(^33\) We simulated micrographs of human H-chain ferritin (PDB: \text{2fha})\(^34\) and HEWL (PDB: \text{1dpx}).\(^35\) For apo-ferritin, 166 micrographs (~20 000 particles) were simulated with an underfocus in the range of 0.2 to 1.3 \(\mu\)m. For HEWL, 501 micrographs with PP and 866 micrographs without PP were simulated, each with 800 to 1100 particles per micrograph. A fluence of 50 \(e^-/\AA^2\) was used for each of the data sets. Power spectra of the micrographs were calculated with Gctf\(^36\), and CTFs were fitted to these power spectra in the 30–2 Å resolution range. Data processing was done using Relion.\(^9\) The motion blur factor \(\sigma_M\) was 0.5 to approximate the beam induced movement of the specimen, which corresponds to a \(B\) factor of 19.7 Å\(^2\) according to eq 3. The quality of the simulated micrographs was compared to experimental ones from EMPIAR\(^37\) (Figure 1).
Data Processing. For the human apo-ferritin data set, 20,214 particles were picked from 166 micrographs (Figure 2a) with a pixel size of 0.5198 Å/pixel. Particles were extracted with a box size of $512 \times 512$ pixels. For 2D classification, we calculated 100 classes using a regularization parameter $T$ of 2 (Figure 2b). As the data were simulated homogeneously from one model (PDB: 2fha), only very few picked particles and 2D classes had to be discarded. A total of 18,062 particles were selected for the initial 3D model building, and 3D classification could be skipped. In 3D refinement, the 3D initial model was low-pass filtered to 50 Å and used as the reference map, with octahedral (O) symmetry, which resulted in a 2.14 Å resolution map. The map was subsequently sharpened, and CTF refinement was performed.24 Finally, Refine3D was run again; the map was postprocessed and sharpened (B factor of $-50 \text{ Å}^2$), yielding a final 1.86 Å resolution map (Figure 2c).

All the Relion processing procedures and parameters used for experimental data were the same as those for simulated data, except for an additional 3D classification step performed for the experimental data. In 3D classification, five classes were calculated using a regularization parameter $T$ equals to 4. After 3D refinement and postprocessing, we obtained a 2 Å resolution map from 29,224 selected particles (Figure S1c).

The pixel size was 0.5 Å/pixel for both lysozyme data sets. Without PP, data were simulated with a 2.5 to 4 μm underfocus (Figure 3a). From 866 micrographs, 525,053 particles were picked and subjected to 2D classification with $T$ equals to 2, using the “ignore CTFs until the first peak” option (Figure 3b). Hereafter, 484,137 particles were selected for 3D refinement using a reference map from the crystal structure (PDB: 1dpx), low-pass filtered to 20 Å, no symmetry. In our hands, building an initial 3D model de novo, as guided by Relion, failed. 3D refinement starting with a 15 Å low-pass filtered map from the known...
answer did not give any new information either, as it resulted in a 16 Å map (Figure 3c).

For HEWL with PP, the defocus ranged from 0.3 to 0.8 μm underfocus (Figure 4a). From 501 micrographs, 290 316 particles were picked and subjected to 2D classification with T equals to 2 with a mask diameter of 50 Å (Figure 4b). For this data, it was not needed to use the ‘ignore CTFs until the first peak’ option in 2D classification. Then 225 751 particles were extracted in a 256 \( \times \) 256 pixel box and used to build an initial 3D model. This initial model was used as a reference map in 3D re

\[\text{FIGURE 3. Single-particle analysis of a simulated HEWL data set without PP. (a) A micrograph of HEWL at 4.0 \, \mu\text{m} \text{ underfocus}; the scale bar is 50 nm. (b) 2D class averages. (c) The 3D reconstruction of HEWL without PP.}\]

Discussion and Conclusion

We simulated single-particle micrographs from which high-resolution 3D density maps could be reconstructed. The simulated micrographs of apo-ferritin were similar to experimental micrographs recorded at a similar defocus, both in terms of intensity, noise, power spectra, background-subtracted radial average, CTF fit, and sigma-to-noise spectra of extracted particles (Figure 1). The B factor estimated by Gctf for the simulated micrographs equals \(\sim\)40 Å\(^2\), slightly greater than what was obtained from the experimental micrographs (\(\sim\)30 Å\(^2\)). The experimental data showed a somewhat stronger signal compared to the simulated one up to 0.5 Å\(^{-1}\) (Figure 1e vs i, Figure 1g), whereas the simulated data showed a slightly stronger signal beyond the resolution at which a 3D structure was obtained. These differences might be due to differences in ice as well as the specific detector as we used generic models in our simulations.

We obtained a 1.86 Å resolution density map for human apo-ferritin using 166 simulated micrographs with 18 062 particles. Circa 10% of the particles were discarded as we still found “imperfect particles” in the micrographs. The particles were placed at random positions within each micrograph taking a minimum interparticle distance into account. Retrospectively, this minimum distance was a bit too small. The simulator relocates the interaction potential of one particle when placing a second one too nearby, resulting in “damaged particles” in the simulated micrographs.

The B factor we calculated using eq 17 from the Guinier plots (Figure S3) was smaller for the simulated data (43 Å\(^2\)) compared to the experimental data set (54 Å\(^2\)). This could relate to the absence of large conformational differences between different particles. With 4260 simulated particles, we achieved a 2.04 Å resolution map, just slightly better than the 2.18 Å map from 4405 experimental selected particles (Figure S3). The identical conformation of simulated particles improved the determination of the five parameters (shifts x and y and Euler angles \(\alpha\), \(\beta\), \(\gamma\)) and particle alignment during 3D re

\[\text{https://dx.doi.org/10.1021/acs.jcim.9b00176}\]

J. Chem. Inf. Model. 2020, 60, 2605−2613
crystal structure 1dpx, we could not obtain a higher resolution map (Figure 3c). For the simulated HEWL data set in the presence of a PP, more than half of the 2D class averages \((T = 2)\) converged into good 2D classes revealing clear details (Figure 4b), with high rotational and translational accuracy and high resolution \((\sim 3 \text{ Å})\). Postprocessing revealed a \(B\) factor of \(-117 \text{ Å}^2\), almost double the \(B\) factor used for apo-ferritin \((-50 \text{ Å}^2\)). The higher \(B\) factor for HEWL compared to ferritin is probably due to inaccuracies in orientation and translation alignments which are \(2.085^\circ\) and \(0.721\) pixels in 3D refinement.

The distinct differences between data sets with and without an ideal PP would argue for the necessity of a PP for particle alignment in order to solve the structures of small proteins. Without PP, a density map of HEWL was not obtainable, at least in our hands. Simulations with and without PP of other small proteins, in particular those smaller than 50 kDa, would provide additional insight into the value of phase plates for single-particle reconstructions. An ideal phase plate would enhance the contrast of the low spatial frequency signal while maximizing the signal at high resolution as the envelope function will have less damping at lower defocus. The contrast in the images will strongly affect the quality of the alignment of particles and the number of particles that are needed to obtain a high-resolution 3D structure. If we take into account the contrast factor, the signal in eq 9 is degraded by a factor of \(C\). Then, the total number of particles required to achieve a certain resolution will increase by a factor of \(1/C^2\). Having an ideal phase plate would therefore decrease the number of particles needed to reach a certain resolution. The PP should be able to give a stable phase shift, preferably \(\pi/2\), and introduce no or a small amount of postsample scattering. It has been shown that the variable phase shift provided by the Volta Phase Plate can be computationally accounted for; however, its postsample scattering will unavoidably dampen all signals, which will be detrimental, in particular at higher resolution. A laser PP or electrostatic PP holds the promise of constant phase shift with minimal postsample scattering.

While our simulation studies gave promising results, we note several caveats. A number of crucial factors in SPA cryo-EM were not accounted for in our simulations. We simulated particles with minimal overlap in a thin ice layer. Such layers have been described in the literature, however, these were most likely obtained by proteins attaching to an air–water interface, at which proteins can partly unfold, contributing to (an increase of) sample heterogeneity. In the simulations presented here, sample heterogeneity was not included. One could, for example, introduce heterogeneity within the biological assembly of ferritin, by having 24 slightly different copies per oligomer. Furthermore, each oligomer itself could be slightly different from the other ones. This would increase the estimated accuracy of angles and offsets reported during the refinements as well as the \(B\)-factor obtained from the Guinier plot. Another caveat is the way we modeled radiation

---

**Figure 4.** Single-particle analysis of a simulated HEWL data set with PP. (a) A typical micrograph of HEWL with PP; the scale bar is 50 nm. (b) 2D class averages. (c) The 3D reconstruction of HEWL at 2.78 Å resolution from 225 751 particles. (d) Gold-standard Fourier shell correlation (FSC) before (blue line) and after (orange line) masking, and the phase randomized FSC (yellow line).
damage, for which we employed a motion-blur factor. More advanced radiation damage models could certainly be envisioned, as well as higher motion-blur factors, than the one we used. An ideal phase plate does not exist yet: it should also be possible to integrate an existing phase plate in the simulator. We deposited all our data sets in EMPIAR and distribute the source code of the simulator via GitHub\(^6\) and hope that some of these caveats will be tackled in later versions.

Nevertheless, our simulations demonstrate that it should be possible to solve sub-50 kDa proteins with current image processing algorithms. We expect that with the development of better detectors, improved phase plates, and optimized sample preparation,\(^9\) one should be able to study a much larger percentage of all the known plant and animal proteins by SPA cryo-EM compared to what is possible nowadays.

Our simulations and simulation software can also be used for other purposes. First, it could help novel cryo-EM users in data processing training with the unique feature that all parameters are known \textit{a priori}. It could help image processing developers to test novel algorithms, e.g., for improving initial 3D model algorithms for fewer numbers of particles. One could simulate focal pairs, to check procedures for combining high-resolution particle information collected close to focus with low-resolution information collected afterward at larger defocus. The potential benefits of better detectors, better beam source characteristics, and the use of different electron beam energies could all be explored computationally. Combined, it could help in pushing forward the already growing field of cryo-EM.

**ASSOCIATED CONTENT**

#### Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jcim.9b01176.

Equations supporting the explanation in theory and extra figures and diagrams to support the validation of simulated data (PDF)

**AUTHOR INFORMATION**

Corresponding Author

R.B.G. Ravelli – The Maastricht Multimodal Molecular Imaging Institute (M4I), Division of Nanoscopy, Maastricht University, 6299ER Maastricht, The Netherlands; \[orcid.org/0000-0001-6056-5888; Email: rbg.ravelli@maastrichtuniversity.nl\]

**Authors**

Y. Zhang – The Maastricht Multimodal Molecular Imaging Institute (M4I), Division of Nanoscopy, Maastricht University, 6299ER Maastricht, The Netherlands; \[orcid.org/0000-0003-4482-1866\]

R. Tammaro – The Maastricht Multimodal Molecular Imaging Institute (M4I), Division of Nanoscopy, Maastricht University, 6299ER Maastricht, The Netherlands

P.J. Peters – The Maastricht Multimodal Molecular Imaging Institute (M4I), Division of Nanoscopy, Maastricht University, 6299ER Maastricht, The Netherlands; \[orcid.org/0000-0002-2964-5684\]

Complete contact information is available at: https://pubs.acs.org/doi/10.1021/acs.jcim.9b01176

**Notes**

The authors declare no competing financial interest.

**ACKNOWLEDGMENTS**

The authors acknowledge funding from the European Union’s Horizon 2020 Research and Innovation Programme under Grant Agreement No. 766970 Q-SORT (H2020-FETOPEN-1-2016-2017). We like to thank Hang Nguyen for critical reading of the manuscript.

**REFERENCES**

(1) Herzik, M. A.; Wu, M.; Landor, G. C. High-Resolution Structure Determination of Sub-100 KDa Complexes Using Conventional Cryo-EM. Nat. Commun. 2019, 10, DOI: 10.1038/s41467-019-08991-8.

(2) Fan, X.; Wang, J.; Zhang, X.; Yang, Z.; Zhang, J.-C.; Zhao, L.; Peng, H.-L.; Lei, J.; Wang, H.-W. Single Particle Cryo-EM Reconstruction of 52 KDa Streptavidin at 3.2 Angstrom Resolution. Nat. Commun. 2019, 10, 2386.

(3) Liu, Y.; Huynh, D. T.; Yeates, T. O. A 3.8 Å Resolution Cryo-EM Structure of a Small Protein Bound to an Imaging Scaffold. Nat. Commun. 2019, 10, 1–7.

(4) Tiessen, A.; Pérez-Rodríguez, P.; Delaye-Arredondo, L. Mathematical Modeling and Comparison of Protein Size Distribution in Different Plant, Animal, Fungal and Microbial Species Reveals a Negative Correlation between Protein Size and Protein Number, Thus Providing Insight into the Evolution of Proteomes. BMC Res. Notes 2012, 5, DOI: 10.1186/1756-0500-5-85.

(5) 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.

(6) D’Imprima, E.; Floris, D.; Joppe, M.; Sánchez, R.; Graninger, M.; Kühnbrandt, W. Protein Denaturation at the Air-Water Interface and How to Prevent It. eLife 2019, 8, DOI: 10.7554/eLife.42747.

(7) Scheres, S. H. w. Beam-Induced Motion Correction for Sub-Megadalton Cryo-EM Particles. eLife 2014, 3, No. e03665.

(8) Li, X.; Mooney, P.; Zheng, S.; Booth, C. R.; Braunfeld, M. B.; Gubbens, S.; Agard, D. A.; Cheng, Y. Electron Counting and Beam-Induced Motion Correction Enable near-Atomic-Resolution Single-Particle Cryo-EM. Nat. Methods 2013, 10, 584–590.

(9) Zivanov, J.; Nakane, T.; Forsberg, B. O.; Kimanius, D.; Hagen, W. J. H.; Lindahl, E.; Scheres, S. H. W. New Tools for Automated High-Resolution Cryo-EM Structure Determination in RELION-3. eLife 2018, 7, DOI: 10.7554/eLife.42166.

(10) Tan, Y. Z.; Aiyer, S.; Mietzsch, M.; Hull, J. A.; McKenna, R.; Griebler, J.; Samulski, R. J.; Baker, T. S.; Agbandje-McKenna, M.; Lyumkin, D. Sub-2 Å Ewald Curvature Corrected Structure of an AAV2 Capsid Variant. Nat. Commun. 2018, 9, DOI: 10.1038/s41467-018-06076-6.

(11) Bartesaghi, A.; Aguerrebere, C.; Falconieri, V.; Banerjee, S.; Earl, L. A.; Zhu, X.; Grigorieff, N.; Milne, J. L. S.; Sapiro, G.; Wu, X.; Subramaniam, S. Atomic Resolution Cryo-EM Structure of β-Galactosidase. Structure 2018, 26, 848–856.

(12) Hamaguchi, T.; Maki-Yonekura, S.; Naitow, H.; Matsuura, Y.; Ishikawa, T.; Yonekura, K. A New Cryo-EM System for Single Particle Analysis. J. Struct. Biol. 2019, 207, 40–48.

(13) Daney, R.; Buijse, B.; Khoshouei, M.; Plitzko, J. M.; Baumeister, W. Volta Potential Phase Plate for In-Focus Phase Contrast Transmission Electron Microscopy. Proc. Natl. Acad. Sci. U. S. A. 2014, 111, 15635–15640.

(14) Schwartz, O.; Axelrod, J. J.; Campbell, S. L.; Turnbaugh, C.; Glaser, R. M.; Müller, H. Laser Phase Plate for Transmission Electron Microscopy. Nat. Methods 2019, 16, 1016–1020.

(15) Müller, H.; Jin, J.; Daney, R.; Spence, J.; Padmore, H.; Glaser, R. M. Design of an Electron Microscope Phase Plate Using a Focused Continuous-Wave Laser. New J. Phys. 2010, 12, 073011.

(16) Tavabi, A. H.; Beleggia, M.; Mignov, V.; Savenko, A.; Öktem, O.; Dunin-Borkowski, R. E.; Pozzi, G. Tunable Ampere Phase Plate
CryoEM at 100 keV: A Demonstration and Prospects. C.; Chen, S.; Leahy, E.; Scotcher, S.; Henderson, R.; Russo, C. J.

Dependence of Contrast and Damage in Electron Cryomicroscopy for Low Dose Imaging of Biomolecular Complexes. J. Chem. Inf. Model. 2020, 60, 2605–2613.

Preparation by Pin-Printing and Jet Vitrification. BioRxiv 2019, DOI: 10.5281/ZENODO.3693834.

Image Formation Modeling in Cryo-Electron Microscopy. J. Struct. Biol. 2013, 183, 19–32.

Image Formation Modeling in Cryo-Electron Microscopy. J. Struct. Biol. 2013, 183, 19–32.

Optimal Determination of Particle Orientation, Absolute Hand, and Contrast Loss in Single-Particle Electron Cryomicroscopy. J. Mol. Biol. 2003, 333, 721–745.

Use B-Factor Related Features for Accurate Classification between Protein Binding Interfaces and Crystal Packing Contacts. BMC Bioinf. 2014, 15, 53.

How Large B-Factors Can Be in Protein Crystal Structures. BMC Bioinf. 2018, 19, DOI: 10.1186/s12859-018-2083-8.

Advanced Computing in Electron Microscopy, Second ed.; Springer: Boston, MA, 2010; DOI: 10.1007/978-1-4419-6533-2.

The Reconstruction of a Three-Dimensional Structure from Projections and Its Application to Electron Microscopy. Proc. R. Soc. London 1970, 317 (1530), 319–340.

Direct Electron Detector Developments: Embarking on a Revolution in Cryo-TEM. J. Struct. Biol. 2015, 192 (2), 179–187.

Crystallography, Structure Solution and Refinement of Hen Egg-White Lysozyme at PH 8.0 in the Presence of MPD. Acta Crystallogr., Sect. D: Biol. Crystallogr. 2000, 56 (8), 952–958.