Towards more dose efficient cryogenic electron microscopy of biological samples

Towards more dose-efficient cryogenic electron microscopy of biological samples

Towards more dose-efficient cryogenic electron microscopy of biological samples

DISSERTATION

to obtain the degree of Doctor at the Maastricht University,

on the authority of the Rector Magnificus,

Prof. dr. Pamela Habibović

in accordance with the decision of the Board of Deans,

to be defended in public

on Monday, 5th of June 2023, at 13.00 hours

by

Yue Zhang

Promoters:

Prof. dr. R.B.G. Ravelli

Prof. dr. P.J. Peters

Copromoter:

Dr. C. López-Iglesias

Assessment Committee:

Prof. dr. G.J. van Rooij (Chairman)

Dr. A. Jakobi, Delft University of Technology

Prof. dr. R.B. Jolivet

Prof. dr. M.E. Kooi

Prof. dr. H.W. Zandbergen, Delft University of Technology

For Raimond-Bernard Gabriël Ravelli

For my parents

For Rong

For Flora

Table of contents

CHAPTER 1

INTRODUCTION

1.1. HISTORY

"*We may soon have detailed images of life's complex machineries in atomic resolution.*" These were the words of the Nobel Prize committee members to Jacques Dubochet, Joachim Frank, and Richard Henderson, who won the Nobel Prize in Chemistry 2017 "*for developing cryo-electron microscopy for the high-resolution structure determination of biomolecules in solution.*" And three years later, in 2020, atomic resolution was achieved for imaging biological molecules using cryogenic-electron microscopy (cryo-EM)[1,2]. How long did it take for humans to see images of life at atomic resolution from the first observation of the world on microscale?

The answer is more than 400 years. Two Dutch spectacle-makers, Zacharias Jansen and Hans Lippershey, invented one of the earliest compound microscope in the late 16th century. Such microscope consisted of one concave lens and a convex lens which could magnify objects up to 20 or 30 times [3]. Then in the 17th century, Antoni van Leeuwenhoek developed his own high-powered microscopes that could magnify objects by up to 200 times [3]. The lenses are hand-grinded and he was the first person to observe and describe microscopic organisms, including bacteria, blood cells and single-celled organisms. Also in the 17th century, Robert Hooke published a book called *Micrographia*, in which he described his observation of various objects using his compound microscope. Numerous developments happened in the following 200 years to improve the resolution of the light microscope. Nevertheless, the resolution of light microscopes is fundamentally constrained by the wavelength of light, which poses a limitation. Ernst Abbe, a German physicist and mathematician, found out that the resolution of microscope is proportional to the wavelength:

$$d = \frac{\lambda}{2nsin(\alpha)} = \frac{\lambda}{2NA}$$
Eq. 1.1

where *d* is the smallest resolvable distance, *n* is the refractive index of the medium, α is the half-angle of angular aperture, and *NA* is the numerical aperture.

Visible light has wavelengths from ~380 to 750 nm. Intuitively, one could come up a method of improving the resolution by utilising particles with shorter wavelengths. Electrons, which were chosen as illumination source by Max Knoll and Ernst Ruska, have wavelengths at picometre scale, depending on their energy. In 1931, the first electron microscope (EM) was developed with significantly better resolution than light microscopes. However, imaging biological specimens using electron microscopes remained a challenge at that time due to biological specimens dehydrating in vacuum and suffering radiation damage caused by electrons. For decades, biological material have been stained by heavy metals before going in to the electron microscope. Imaging biological specimen with negative stain has some advantages, such as high contrast, radiation hardness and morphology preservation, and it has been used for studying

structures of biomolecule at low resolution [4]. However, it poses several drawbacks. Firstly, the resulting images are of the casting stain rather than the biological sample itself. Dehydration due to staining could cause structural changes in the biological sample. In addition, staining material, such as uranyl acetate, is only soluble at very low pH, which not all proteins can withstand, causing possible structural changes or even full denaturation. Secondly, the resolution is limited by the granularity of the stain. Staining introduces artefacts which degrades resolution. Fixation followed by gradual dehydration and resin embedding was developed for bulk specimens to stabilise them under the electron beam. Also, this approach allowed for sample thinning to make the electron beam able to cross the sample and form an image from it. However, artefacts induced by mechanical damage or chemical interactions can interfere with obtaining high resolution and interpretation of the specimen. As a result, scientists have been exploring ways to image biological specimens in their hydrated state using EM.

1.2. Development of Cryo-EM

Humberto Fernández-Morán was a pioneer who tried to image cryo-cooled samples in the electron microscope to prevent water evaporation and reduce radiation damage [5]. Unfortunately, the formation of crystalline ice due to the low cooling rate strongly affected the signal from sample itself. Still, the use of cryogenic temperatures proved to be a good method for maintaining the sample hydration and obtain good resolution with crystalline samples [6,7]. In 1981, a method of cooling liquid water into a vitrified state was introduced by Jacques Dubochet and Alasdair McDowall [8]. This breakthrough allowed imaging close-to-native biological specimen in EM with much less radiation damage. Subsequently, better contrast, higher resolution micrographs were obtained for various biological specimens with this sample preparation method [9-13].

Beside sample preparation, developments in image processing were also indispensable for obtaining reliable 3D structure reconstructions of imaged objects [14,15]. Such methods involve aligning and averaging noisy particle images to boost their signal-to-noise ratio, then computationally sorting out their 3D orientation from averaged 2D projection classes [16,17]. In the early nineties, the first high resolution 3D structure of biomolecules was obtained using cryo-EM [18], and this approach has since been applied in many subsequent research studies [19-21]. Despite numerous developments and breakthroughs, cryo-EM was still referred to as "blobology" as reconstructions were obtained in which no secondary structures could be deciphered, making it difficult to objectively evaluate their correctness.

The year of 2013 marked a milestone in cryo-EM, as researchers achieved better than 4 Å resolution, allowing for *de novo* model building three-dimensional structures of biomolecules within the obtained reconstructions [22-25]. One of the key contribution to such "Resolution Revolution" is the development of direct electron detectors (DEDs) in the form of monolithic active-pixel sensors (MAPS) [26]. MAPS detectors showed much better sensitivity and better

modulation transfer functions (MTF) compared to its ancestor - charge-coupled device (CCD) cameras [27-31]. And importantly, MAPS detectors have fast readout speed allowing for the collection of "movies" with multiple frames per second, making it possible to correct for sample movements which occur during electron exposure. Sample movements during imaging have been a major bottleneck for obtaining high-resolution images in cryo-EM [24].

The development of better sample grids has also contributed to the acquisition of higher-resolution data. The cryo-cooling of the sample builds up compressive stress within the thin film of specimen due to the difference of thermal expansion coefficients between water, support foil and grid [32,33]. Relaxation of this stress, bond-breakage, and possible charging of the specimen, can all induce beam-induced sample movements [34]. UltrAuFoil grids minimise differences in thermal expansions/contractions as the same material was used for both the foil and grids [35], whereas its successor, HexAuFoil, was designed to have a better hole size-ice thickness ratio to further reduce stress relaxation during electron irradiation [33]. Graphene has also been used to reduce beam-induced motions, and its functionalisation treatments helped to resolve issues related to poor particle distribution, preferred orientation, and air-water interface [36-41].

Another important development in cryo-EM is the use of maximum likelihood methods, which deal with cryo-EM structure determination from a Bayesian approach, allowing for a statistical evaluation of unknown parameters (shift and Euler angles) of individual particles [42]. Recent advancements in image processing techniques, such as contrast transfer function (CTF) parameter refinement, Ewald sphere curvature correction, as well as Bayesian polishing, has further pushed the resolution that can be obtained from a given single particle analysis (SPA) dataset [43-46]. Improvement on instrumentation also plays a crucial role in pushing the resolution of cryo-EM. The development of Volta phase plate led to the successful reconstruction of a 64 kDa haemoglobin at 3.2 Å [47], however, its use remained limited thus far. The development of new faster detectors, cold field emission gun (CFEG), spherical aberration (C_s) correction, better energy filters, and the use of a monochromator have all contributed to the achievement of atomic resolution in cryo-EM [1,2].

However, despite these numerous spectacular development and achievement in cryo-EM, there are still important challenges, in particular when dealing with small proteins or with highly heterogeneous samples. It was estimated by Henderson [48] that 38 kDa is the smallest particle that can be determined by cryo-EM. Glaeser [49] later proposed that such estimation was too conservative, and that particles as small as 17 kDa could be imaged and reconstructed using cryo-EM. Despite the progress made in cryo-EM, there is still a significant disparity between the current demands for achieving high-resolution reconstructions and the estimated theoretical limits, especially when working with small proteins. According to the analysis of the size distribution of proteins in plants, animals, fungal and microbial species [50], the majority of plant and animal proteins (90%) are, in monomeric form, smaller than 100 kDa, with more than 50% of them being smaller than 50 kDa. Meanwhile, many targets for drug design are also small proteins. One of the main constraints in routinely achieving high resolution of small proteins is

the low signal-to-noise ratio (SNR) of images recorded with the limited electron irradiation. All of the above brings me to the topic of this thesis: **Towards more dose-efficient cryogenic electron microscopy of biological samples**.

1.3. SCOPE OF THIS THESIS

In this thesis, studies have been carried out to develop new cryo-EM approaches, which should enable scientists to overcome existing challenges and allow for higher dose efficiency imaging and diffraction techniques.

As I introduced above, software and hardware developments have enabled a resolution revolution in cryo-EM. A standard protein is needed to characterise the performance of any new developments. So far, ferritin has allowed for the highest resolution 3D SPA reconstruction. It has been chosen due to its stability, suitable size, and high symmetry. **Chapter 2** describes the use of ferritin from the bacterium *Mycobacterium tuberculosis* (BfrB) as a model protein for developing and optimising cryo-EM techniques. Unlike other ferritins, BfrB can be easily obtained by a simple and high-yield protocol. In this thesis, I have shown that it is a suitable workhorse protein for cryo-EM development, for which we obtained a near-atomic 1.9 Å resolution map.

Ferritin is far from being the smallest protein that can be determined by cryo-EM. Henderson [48] pointed out proteins smaller than 38 kDa cannot be resolved by cryo-EM, nevertheless, a 32 kDa haemoglobin heterodimer was solved by cryo-EM at 6.4 Å [51]. What is the size limitation of cryo-EM? Could egg white lysozyme (14 kDa) be solved by single particle cryo-EM? **Chapter 3**, the work for which was done during one of the COVID-19 pandemic lockdowns, provides insights into the theoretical limits of cryo-EM resolution for smaller proteins by using realistically simulated single-particle micrographs to describe the size limitation of cryo-EM.

Sample charging is one of the fundamental issues of imaging biological specimen embedded in suspended vitreous ice. Charging can also be one of the main problems preventing us from reaching higher resolutions and imaging smaller proteins in cryo-EM. **Chapter 4** characterises charging of vitreous samples, and presents a novel approach to mitigate charging of vitreous samples in cryo-EM through the deposition of a layer of graphene on top of regular EM grids.

Chapter 5 describes the integration of a Timepix3 hybrid pixel detector into a cryo-EM workflow for high-resolution structure determination of biological macromolecules. Such event-driven detector has high time resolution, and can operate in both imaging and diffraction modes, paving the way towards more dose-efficient single-particle techniques discussed in **chapter 6**. **Chapter 6** (written during another COVID19 lockdown) discusses alternative strategies for improving the dose efficiency of single-particle cryo-EM, including laser phase plate, multi-pass transmission electron microscopy, off-axis holography, ptychography, and a quantum sorter.

These techniques have potential in further pushing the limits of size, structural heterogeneity and resolution at which one can study the building blocks of life.

Last but not the least, out of my personal interest, a discussion about cryo-EM sample preparation, more accessible and affordable microscopes, and artificial intelligence in structural biology is provided as **chapter 7**.

CHAPTER 2

MYCOBACTERIUM TUBERCULOSIS FERRITIN: A SUITABLE WORKHORSE PROTEIN FOR CRYO-EM DEVELOPMENT

Yue Zhang[*], Abril Gijsbers[*], Ye Gao, Peter J. Peters and Raimond B. G. Ravelli

[*]both authors contributed equally
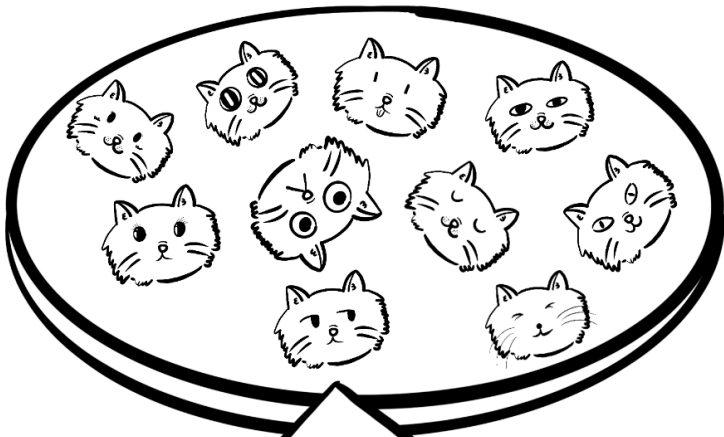
ABSTRACT

The use of cryo-EM continues to expand worldwide and calls for good-quality standard proteins with a simple protocol to produce it. We present a straightforward expression and purification protocol that provides high yield and purity of an apo-ferritin from *Mycobacterium tuberculosis*, BfrB. We report a 2.12 Å resolution cryo-EM structure of BfrB, showing the typical cage-like oligomer constituting of 24 monomers related by 432 symmetry. However, it also contains a unique C-terminal extension (164 - 181), which loops into the cage-region of the shell and provides extra stability to the protein. Part of this region was ambiguous in previous crystal-structures but could be built within the cryo-EM map. Our findings and protocol could serve the growing cryo-EM community to characterise and push the limits of their electron microscopes and workflows.

2.1. Introduction

Single particle (SPA) cryogenic electron microscopy (cryo-EM) has become an indispensable tool for structural biology. The combination of direct electron detectors, motion correction [24,52], high-end electron microscopes and advanced imaging process algorithms, allowed for the "Resolution Revolution" [26] from whereon an increasing number of SPA structures with resolution below 2 Å could have been determined [43,53-55]. Hardware improvements such as a monochromator, spherical aberration (C_S) corrector, energy filter and new generation of direct electron detector improved the signal-to-noise ratio even further and has brought SPA to real atomic resolution [1,2]. The success of SPA attracted many new scientists into the field of cryo-EM and led to an impressive growth of its community as well as the number of instruments that can deliver high-resolution cryo-EM structures. Ferritin is often used as a benchmark to commission and validate these machines, as it is relatively straightforward to get good-resolution datasets of its 24-mers. In fact, as of March 2021, 11 out of 25 SPA structures deposited in the Protein Data Bank (PDB) with resolution better than 2 Å are ferritins, from different organisms. Ferritin has thus become a gold standard for cryo-electron microscopists to evaluate their setups and for pushing the limits of sample preparation [35,56-58], imaging [1,2] and data processing [43,59,60].

Ferritins are protein complexes involved in iron homeostasis and DNA repair by storing it in cage-like structures. This protein complex is formed by 12 or 24 subunits with tetrahedral 23-symmetric or octahedral 432-symmetric arrangements, depending on which sub-family it belongs. There are three sub-families of ferritin: iron-containing ferritin and heme-containing bacterioferritin involved in the iron storage, and DNA-binding protein in starved cells (Dps) that uses iron to protect DNA from free radical-mediated damage [61,62]. Iron is a trace element vital for all living organisms, and although it is the second most common metal on Earth, it is not always bioavailable: for this reason organisms compete for the iron in the environment [63]. The acquisition of iron is considered a key step in the development of any pathogen, and so, host cells have developed mechanisms to sequester the metal from infecting bacteria in an attempt to defend themselves [64]. One of such pathogens is *Mycobacterium tuberculosis*, the causative agent of human tuberculosis.

In this study, a simple high-yield purification protocol is presented for bacterioferritin B (BrfB) of *M. tuberculosis*, together with its high-resolution cryo-EM structure. Previously, only lower resolution crystal structures had been reported. Our structure shows extra residues at both termini. The C-terminal tail of BrfB, being a target peptide for encapsulation [65] and playing a role in ferroxidase activity and iron release in addition to providing stability to the protein [66], loops into the cage region of the complex, prior to folding back via the three-fold channel to the B-pore. Two conformations of the C-terminal residue His175 were observed near the B-pore, suggesting a role of this residue in iron exchange. BfrB could serve the structural community in testing their expanding fleet of equipment as well as aid to a better understanding of iron storage

proteins, which are essential for the survival and progression of important human pathogens like *M. tuberculosis*.

2.2 MATERIALS AND METHODS

2.2.1. Expression and purification of BfrB

Codon-optimised gene of *M. tuberculosis* BfrB was cloned into a modified pRSET backbone (Eurofins Genomics) using *Nde*I and *Hind*III restriction sites to over-express in C41(DE3) *Escherichia coli*. Expression and purification protocol was adapted from Khare, et al. [66] and Parida, et al. [67]. In summary, primary culture of lysogeny broth (LB) medium supplemented with 100 μg/mL ampicillin was prepared from a single colony at 37 °C and 200 rpm overnight, to then inoculate 500 mL fresh LB medium with 1:1000 dilution of the primary culture in the same conditions until reaching an optical density (OD_{600}) of 0.6. Protein over-expression was induced by adding isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 1 mM, for 3 h without changing the culture conditions. Cells were harvested by centrifugation and stored at −20 °C until further use.

Bacterial cells were resuspended in 20 mM Tris-HCl pH 8.0, 300 mM NaCl supplemented with 2 units/mL of benzonase (Sigma-Aldrich), EDTA-free protease inhibitor (Sigma-Aldrich) and 2 mM β-mercaptoethanol (BME). The cell suspension was lysed by sonication. Cellular debris was removed by centrifugation at 30,000 ×g for 30 min. Supernatant was subjected to a saturation of 20% ammonium sulfate and incubated at 5 °C for 1 h with rotation before centrifugation at 15,000 ×g for 20 min. The pellet was resuspended in 20 mM Tris-HCl pH 8.0 and 150 mM NaCl, and centrifuged at 10,000 ×g for 5 min to remove precipitants. Sample was further purified by size exclusion chromatography with a Superdex 200 Increase 10/30 column (GE Healthcare). Fractions were collected and protein purity was evaluated by denaturing polyacrylamide gel electrophoresis. Fractions containing BfrB were pooled together and stored at −80 °C until further use. The final yield was >100 mg for 1 L of culture.

2.2.2 Cryo-electron microscopy sample preparation, data acquisition and image processing

Purified BfrB was first used at a concentration of 11 mg/mL (calculated with Pierce™ BCA Protein Assay Kit). A volume of 2.5 μL was applied on glow discharged UltrAuFoil Au300 R1.2/1.3 grids (Quantifoil). Excess liquid was removed by blotting for 3 s (blot force 5) using filter paper followed by plunge freezing in liquid ethane using a FEI Vitrobot Mark IV operated under 100% humidity at 4 °C. Cryo-EM single particle data were collected on a Titan Krios at 300 kV with a BioQuantum K3 Imaging Filter with a 20eV post-column energy filter. The

detector was utilised in super-resolution counting mode at a nominal magnification of 130,000×. Table 2. 1 shows the statistics of the data set. Data were processed using the RELION pipeline [68]. Movie stacks were corrected for drift (5 × 5 patches) and dose-weighted using MotionCor2 [69]. The local contrast transfer function (CTF) parameters were determined for the drift-corrected micrographs using Gctf [70]. A first set of 2D-references were generated from manually picked particles in RELION [68] and these were then used for subsequent automatic particle picking. Table 2. 1 lists the number of particles in the final data set after particles picking, 2D-classification and 3D classification. The latter was with O symmetry. Beamtilt parameters, anisotropic magnification and local CTF parameters were refined and particles were polished [43]. The resolution of the final full map, listed in Table 2. 1 and shown in Figure 2. 1, was 2.12 Å using the gold-standard FSC= 0.143 criterion [71]. A B-factor plot according to Rosenthal and Henderson [72] was calculated using random subsets of the data with variable number of particles (Supplementary Figure 2. 1).

Table 2. 1 also includes the statistics of a data set collected in-house on a Tecnai Arctica microscope, operating at 200kV, with Falcon3 camera operated in electron counting mode (no energy filter). For this data set, we used a highly concentrated sample of 80 mg/mL, which gave beautiful micrographs with densely packed monolayers of BfrB in the middle of the holes of the UltrAuFoil Au300 R1.2/1.3 grids (Figure 2. 1a). This dataset contained only a third of the number of micrographs compared to the K3 dataset, however, since it was collected on a more concentrated sample with a larger pixel size (0.651 versus 0.935 Å), similar number of particles were obtained for both datasets. The resolution of the 200kV Falcon3 data set was 2.39 Å.

2.2.3   Structure determination and refinement

We used the PDB model 3QD8 [66] as a starting model in Coot [73] for manual docking and building. The final model was refined against a sharpened cryo-EM map obtained by LocSpiral [74]. The model was refined iteratively through rounds of manually adjustment in Coot [75], real space refinement in Phenix [76] and structure validation using MolProbity [77]. The refined model has been deposited in the Protein Data Bank as 7O6E and maps within the Electron Microscopy Data Bank as EMD-12738.

Table 2. 1 Data collection, Refinement and Validation statistics

Data Collection		
Sample	BfrB *M. tuberculosis*	
Concentration (mg/mL)	11	80
Grid type	Quantifoil UltraAuFoil 300 mesh R1.2/1.3	Quantifoil UltraAuFoil 300 mesh R1.2/1.3
Plunge freezer	Vitrobot	Vitrobot
Microscope	Titan Krios	Tecnai Arctica
Voltage (kV)	300	200
Energy filter (eV)	20	None
Camera	K3	Falcon3
Detector Mode	Super-resolution counting	Electron Counting
Nominal magnification (k×)	130	110
Physical pixel size (Å)	0.6514	0.935
Exposure time (s)	1.3	46.33
Fluence (e⁻/Å⁻²)	40	41
Focus range (microns)	-0.4, -0.6,-0.7,-0.8,-0.9,-1.0,-1.2,-1.4	-0.75, -1.0, -1.25
Micrographs	2518	875
Number of fractions	50	50
Particles	163568	186025
Symmetry imposed	O	O
Average resolution (Å)	2.13	2.39
FSC threshold	0.143	0.143
Map sharpening B factor (Å²)	−68	-109
Refinement		
Initial model used (PDB entry)	3qd8	
Model resolution against LocSpiral (Å)	1.89	
FSC threshold	0.5	
Model resolution against Relion map (Å)	2.18	
FSC threshold	0.5	
Model composition monomer		
Atoms	1471	
Hydrogen atoms	0	
Protein residues	177	
Waters	58	
B factors (Å²)		
Protein	28.33	
Water	30.79	
R.M.S. deviations		
Bond lengths (Å)	0.017	

Bond angles (°)	1.324
Correlation coefficients	
Mask	0.88
Box	0.87
Validation	
MolProbity score	1.20
Clashscore	3.49
Poor rotamers (%)	0
Ramachandran plot	
Favoured (%)	97.71
Allowed (%)	2.29
Disallowed (%)	0

2.3   RESULTS AND DISCUSSION

2.3.1   Mycobacterial ferritin structure

We studied the structure of BrfB by cryo-EM and obtained a 2.12 Å-resolution density map based on gold-standard FSC [71] using 2518 micrographs with 163,568 particles. High-quality data allowed observation of secondary structure features already during 2D classification (Figure 2. 1). Clear density of side-chains and holes in aromatic rings illustrates the quality of the EM map (Figure 2. 2). Our final model contains 1392 water molecules, whereas the BfrB model built from the 3 Å X-ray crystallography map has 360 [66] and the 1.15 Å cryo-EM structure of human apo-ferritin has 4622 water molecules [2].

Previous models of *M. tuberculosis* BfrB were determined by X-ray crystallography (pdbid 3QD8 at 3 Å, 3OJ5 at 2.85 Å, & 3UNO at 2.5 Å). We built our model in an enhanced cryo-EM map, a LocSpiral map, calculated using algorithms based on spiral phase transformation (Figure 2. 2) [74]. Supplementary Figure 2.3a shows a comparison between the Relion postprocess map and the LocSpiral sharpened map, revealing some extra features in the latter one. To confirm these, we also calculated the LocOccupancy maps which estimates the density occupancy (Supplementary Figure 2.3b) [74]. The resolution of both maps was also estimated by comparing the FSC between the refined model and the map [76] at a cut-off value of 0.5 (Supplementary Figure 2. 2): this is 1.89 Å and 2.18 Å when comparing to the LocSpiral and Relion postprocess map respectively (Table 2. 1).
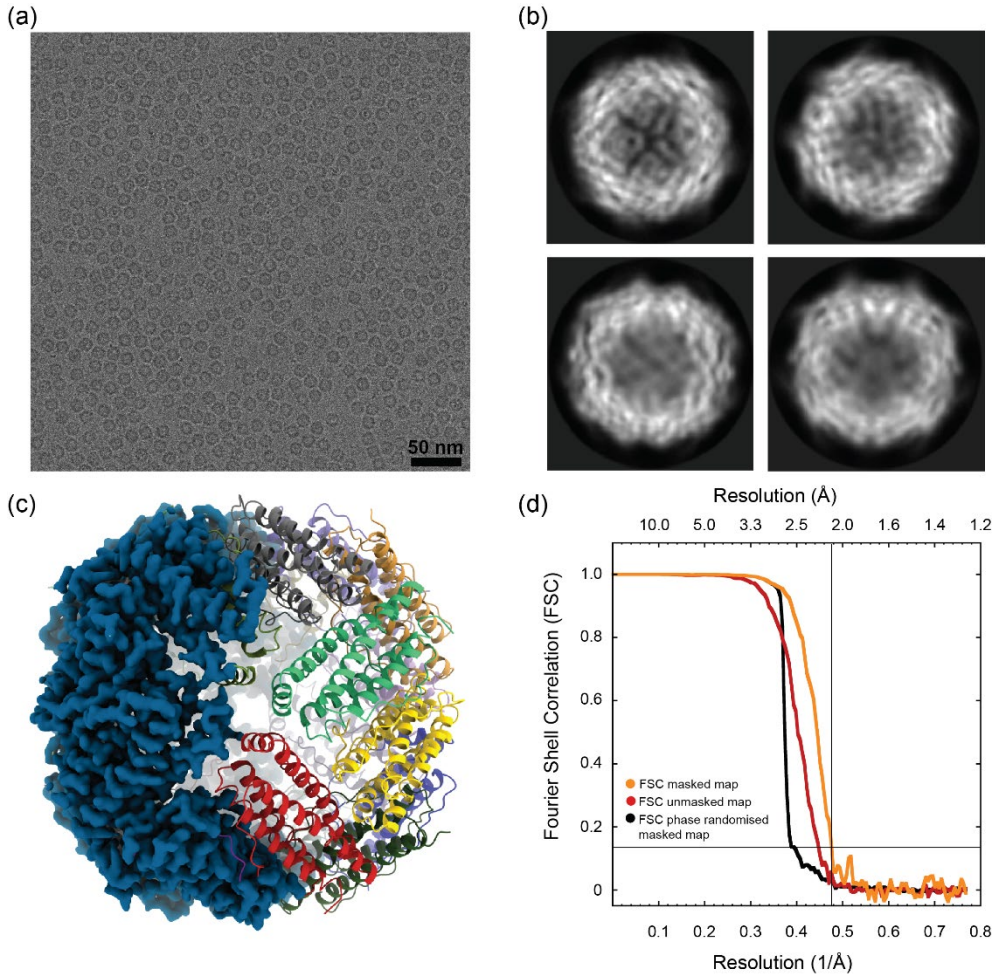
Figure 2. 1 Single-particle analysis of BfrB. a) A micrograph of highly concentrated (80 mg/mL) BfrB sample in vitreous ice, collected on a Falcon3 at 200kV. b) 2D class averages, the size of the shown box is 150 Å. c) 3D reconstruction from 163.568 particles at 2.12 Å resolution, collected on a K3 at 300kV. d) Gold-standard Fourier shell correlation (FSC) before (red line) and after (orange line) masking, and the phase randomised FSC (black line).
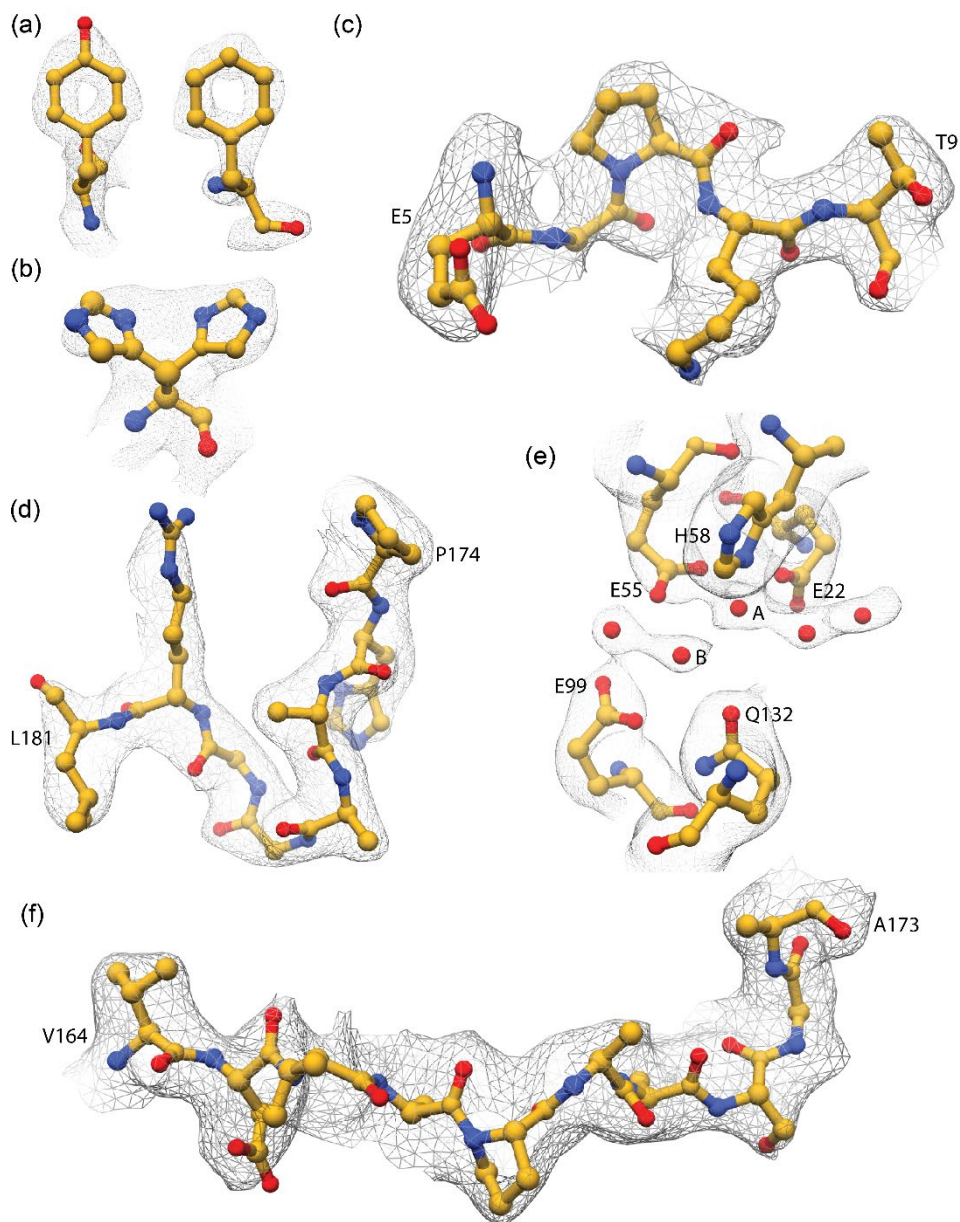
Figure 2. 2 Representative regions of the density. (a) Post-processing map density for Phe23 and Tyr35. Local resolution scaled map of (b) His175, (c) Glu5-Thr9, (d) C_{rigid} (Pro174-Leu181). (e) A string of density blobs near the ferroxidase sites, (f) C_{flex} (Val164-Ala173). The EM density is shown in gray mesh, the residue atoms are represented in ball-stick model.

At the N-terminus, we could add residues Glu5 to Thr9 for which clear density was seen in the enhanced EM map (Figure 2. 2c), but were absent in our starting model 3QD8. The C-terminal region 164-181 consists of a flexible part (C_{flex} 164-173) and a rigid part (C_{rigid} 174-181), as described by Khare, et al. [66]. C_{flex} is ill defined within the X-ray maps. The C-terminal region has been shown to be important for protein stabilisation and iron uptake [66] and so, confident assignment of these residues could aid our understanding of BfrB function. The extension of the C-terminal end in *M. tuberculosis* is unusual for ferritin even compared to heme-containing BfrA from the same organism. This extension has been shown to play an essential role in its function; BfrB exhibits a 3.5-fold reduction in oxidation rate of Fe(II) and a 20% reduction in Fe(III) release rate upon removal of the C-terminal end [66]. We found the extension of the C-terminal end to be located within the interior of the BfrB cage, which is remarkable as Mt-enc has been reported to encapsulate BfrB via this C-terminal extension [65]. Our EM map displays well-defined density for C_{rigid} (Figure 2. 2d), residues 174-181, including a double conformation of residue His175 (Figure 2. 2b). Similar to *S. coelicolor* bacterioferritin, Fe^{2+} ions enters BfrB from the B-pore and converts to Fe^{3+} at the ferroxidase centres [66,78,79]. We hypothesise that the double conformation of His175 might be relevant for the iron exchange of the protein as this residue is located at the interior of the B-pore and in the vicinity of a cavity (Figure 2. 3 and Supplementary Figure 2. 4). This cavity leads towards the ferroxidase sites A (Glu22, Glu55 and His58) and B (Glu55, Glu99 and Glu135) where the ferrous iron is oxidised by molecular oxygen [66,67]. Previously deposited X-ray diffraction electron maps did not place iron ions at these sites. The 2.5 Å map for 3UNO only showed good density for the B-site and weak density for the A-site: at both positions water molecules were placed. The highly conserved Gln132 might favour iron binding at site B. The 3.0 Å map for 3QD8 showed some unmodelled density at the B-site and no density for the A-site. Finally, the 2.85 Å map for 3OJ5 did not show clear density for either the A or the B-site. In our EM map, a string consisting of 5 density blobs could be seen (Figure 2. 2e). We placed water molecules in here, as we lack experimental evidence for these water molecules to be ions. Both the A and B-site have clear densities, where the A-site is coordinated by Glu22 (2.55 Å), Glu55 (2.78 and 2.82 Å), His58 (2.51 Å) and neighbouring water molecules (2.44 and 2.65 Å). The B-site is coordinated by Glu55 (2.77 Å), Glu99 (3.03 Å), Gln132 (2.55 Å) and another water molecule (2.55 Å). Glu135 has different conformation compared to the X-ray map and does not coordinate to the A or B-site. The coordination distances are longer then what is most commonly found for iron ions according to MetalPDB [80]: 2.031-2.236 Å for Fe-N and 2.077-2.414 Å for Fe-O. Another string of multiple water-like densities was found near Asp37, Pro42, Lys46 and Ser50.

Our map reveals density for C_{flex}, which extends into the interior of the cage and is located above the cavity between the B-pore and the ferroxidase centres (Figure 2. 3). The density for C_{flex} is less defined than other areas of the molecule (Figure 2. 2f and Supplementary Figure 2. 3), which would reflect flexibility and a possible functional role for this [81]. LocSpiral [74] helped to further improve this part of the density.
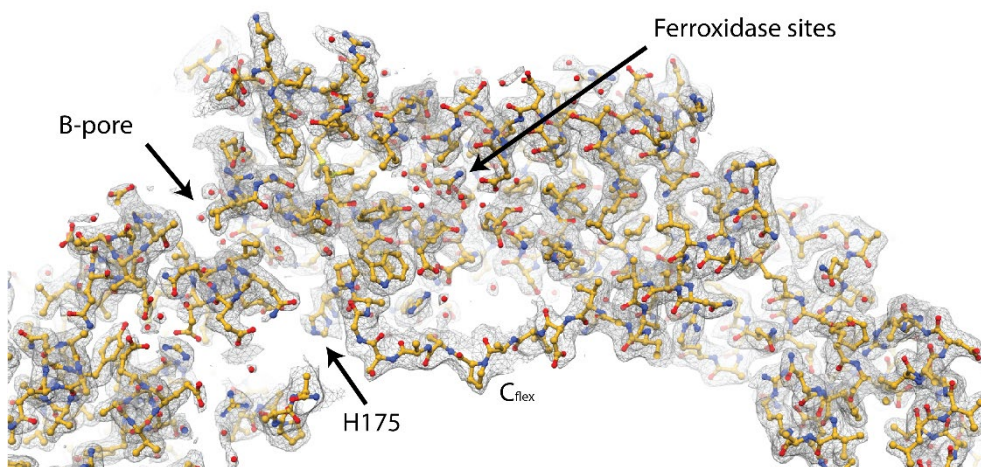
Figure 2. 3 Density map and model for C_{flex}, which extends into the interior of the cage and is located above the cavity between the B-pore and the ferroxidase centres.

2.3.2 Suitable Characteristics for a good standard protein for cryo-EM

Ferritin has become more and more popular to serve as a standard protein for testing and training purposes in microscope labs. The good solubility and stability of most ferritins, its good intermediate size (12-13 nm diameter), combined with a 432-point group symmetry allows for swift characterisation of the performance of the microscope using minimal number of micrographs. To this date, the most commonly used samples are the commercially available horse spleen ferritin, and self-produced mouse ferritin [82] and human ferritin [2]. However, samples and protocols are not always optimal. The most affordable sample is horse spleen ferritin, but, unfortunately, it seldom provides the sample quality needed to push the resolution, possibly due to the presence of broken particles that introduces sample heterogeneity. Khare, et al. [66] compared the stability of ferritin subunits from different organisms and reports that horse ferritin is one of the least stable of all, together with mouse ferritin. The highest resolution achieved by cryo-EM is on human ferritin [2], however, the procedure described is an extensive protocol with more than 10 steps that includes precipitation of nucleic acids, two ammonium sulfate precipitations, two sucrose gradients, two 24-h dialyses and an ion exchange chromatography. Although the final yield was not reported, it is expected that some protein will be lost with every step of the purification, resulting in high purity but low yield. The length and the use of so many different steps is inconvenient if we aim for a protocol that can be reproduced worldwide, as it will require infrastructure and expertise that might not be available everywhere. Previously, we have tried expression and purification protocols of constructs from different

organisms leading to yields that were, in our hands, not sufficient for the extensive and routine use of the proteins with classical grid preparation techniques such as the Vitrobot.
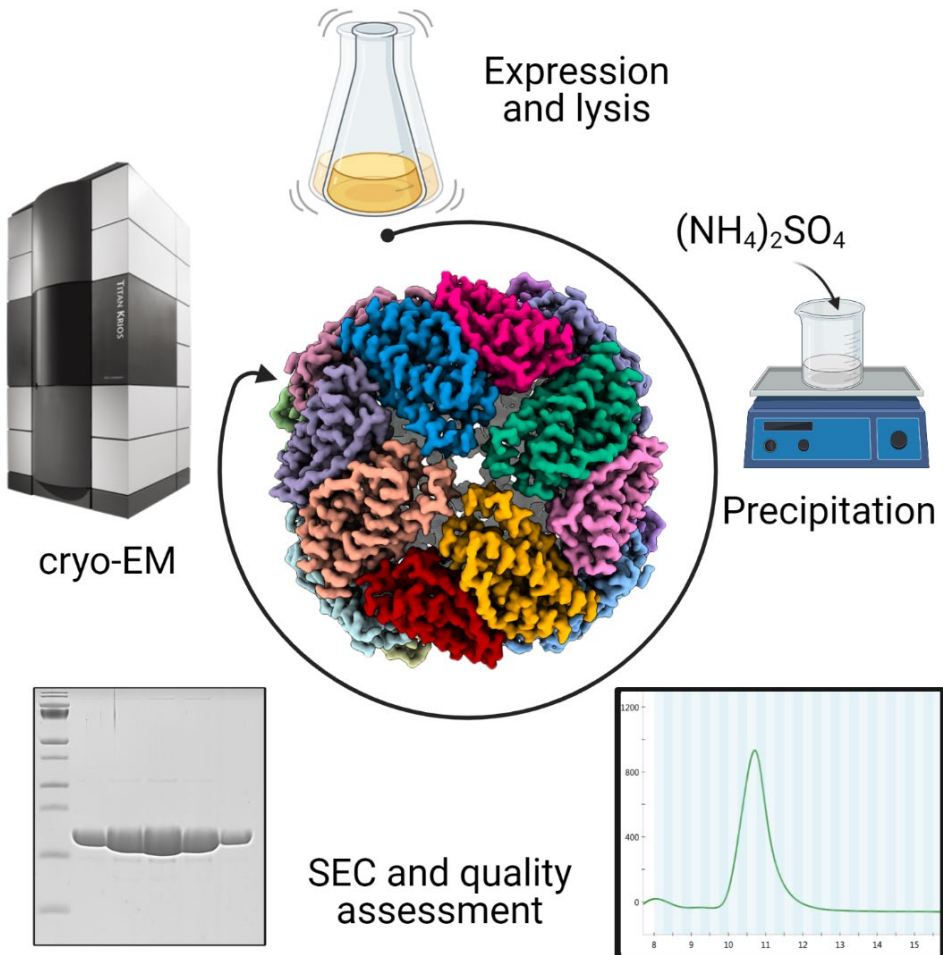
Figure 2. 4 Simple purification workflow of BfrB protein. After protein expression, E. coli cells are lysed by sonication and cellular debris removed by centrifugation. Protein sample is precipitated by ammonium sulfate, and resuspended to be further purified by size exclusion chromatography. High yield of highly pure protein is obtained, suitable for cryo-EM studies.

In here, we have provided an easy and cheap method to purify the mycobacterial BfrB recombinantly expressed in *E. coli* C41 strain. The expression and purification protocol was adapted from Khare, et al. [66], to increase the yield from ~17.5 mg to >100 mg of protein per litre of culture. Given the good expression, we did not design a tag for the BfrB plasmid, thereby overcoming the need for a tag removal step. One ammonium sulfate precipitation step followed

by size exclusion chromatography was sufficient to obtain samples of high-levels of purity, as confirmed by SDS-PAGE gel (Figure 2. 4 and Supplementary Figure 2. 5) as well as cryo-EM (Figure 2. 1). Protein concentration up to 115 mg/mL (calculated from the theoretical extinction coefficient) was achieved without any obvious precipitation, and no monomers observed in the chromatogram, suggesting high solubility and stability of purified BfrB oligomer. The 2.12 Å structure presented here was obtained from one 11 mg/mL aliquot stored at −80 °C, indicating that the protein is well resistant against at least one freeze/thaw cycle. We prepared grids up to a concentration of 115 mg/mL, and observed a dense packing of non-overlapping BfrB particles at 80 mg/ml (Figure 2. 1**a**), which is remarkably high but reproducible for our local grid preparation system. The B-factor plot according to Rosenthal and Henderson [72] shows that we obtained resolutions better then 4, 3 and 2.5 Å for, respectively, 100, 1000 and 10000 particles (Supplementary Figure 2. 1). Although the fitted B-factor of 75 Å² of our reported data set is too high to expect record-breaking resolutions, it is a perfectly adequate sample to obtain 2.5 Å with modest microscopes and settings (Table 2. 1). Whereas the rigid part of BrfB can serve in characterising the performance of the microscope at regular times, the flexible C-terminus could also help to train users in some of the more advanced image data processing SPA steps such as particle subtraction and focused classification. Plasmids are available upon request.

Over the last decade, there has been a fast development in every aspect of cryo-EM. Recent publications have proven that this technique has the potential to visualise macromolecules at atomic level [1,2]. However, every protein has different physicochemical properties that makes them unique and that could challenge its structural elucidation by cryo-EM. For this reason, further improvements in cryo-EM are still necessary and thus a protein that serve as 'workhorse' to test them. BfrB could be one of them.

Data Availability

The refined model has been deposited in the Protein Data Bank as 7O6E and maps within the Electron Microscopy Data Bank as EMD-12738.

Author Contributions

RBGR designed and directed the project; AG designed the studies and produced the constructs; YG and AG purified the protein; AG and RBGR performed the imaging; RBGR processed the data and YZ refined the structures; AG, YZ, PJP & RBGG interpreted the results; PJP and RBGR acquired funding; AG, YZ and RBGR wrote the manuscript with input from all authors.

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

Supplementary Figure 2. 1 B-factor plot according to Rosenthal and Henderson [72], where $1/d^2$, with d being the resolution of each refinement, is plotted against the natural logarithm of the number of particles in each subset. The fitted B-factor is 75 Å².
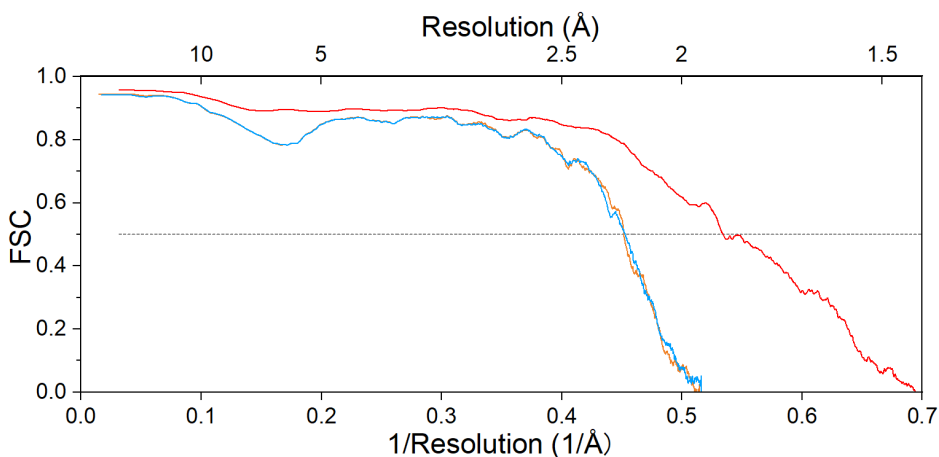
Supplementary Figure 2. 2 Fourier Shell Correlation (FSC) between map and model. The blue and orange lines are the FSC curves between model and both half maps. The red curve reports a resolution for the final map, according to the FSC (model) = 0.5 criterium, of 1.89 Å.
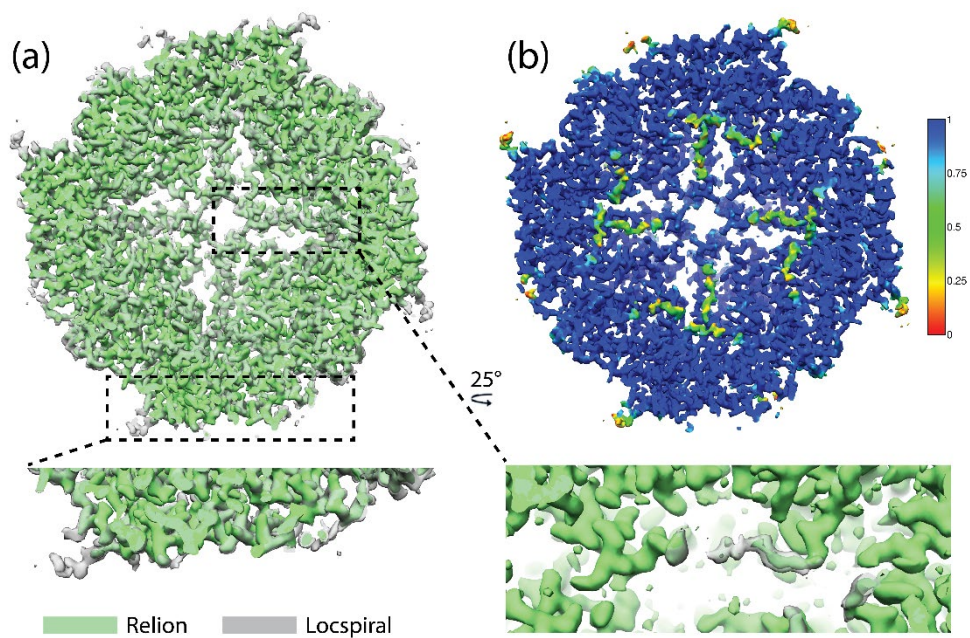
Supplementary Figure 2. 3 The superimposed sharpened maps obtained by LocSpiral (grey colour) and Relion postprocessing (green colour). The figures below show zoomed view of the region indicated with the dashed rectangles. (b) The occupancy map of *Mtb* BfrB model obtained by LocOccupancy, calculated between 8 and 2.1 Å [74]. The occupancy ranges between [0, 1], the occupancy of residues at N-term and C_{flex} have occupancy lower than 1.
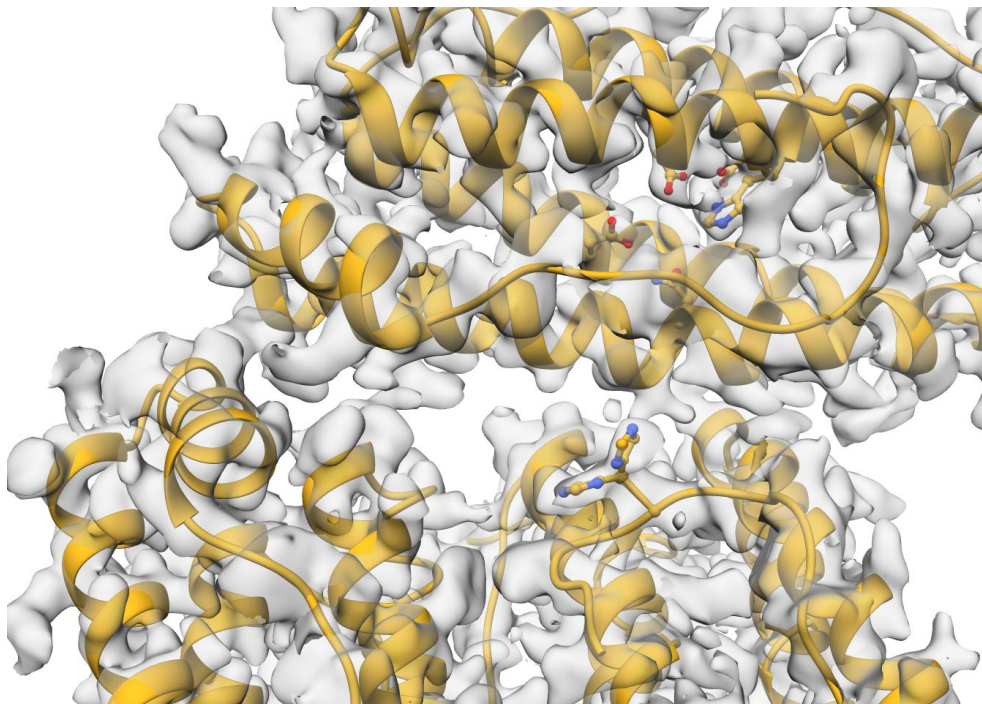
Supplementary Figure 2. 4 Cryo-EM map and Mtb BfrB model with the B-pore channel shown in the middle with a double conformation of His175. His175 as well as the residues at the ferroxidase sites are shown in ball-stick model.
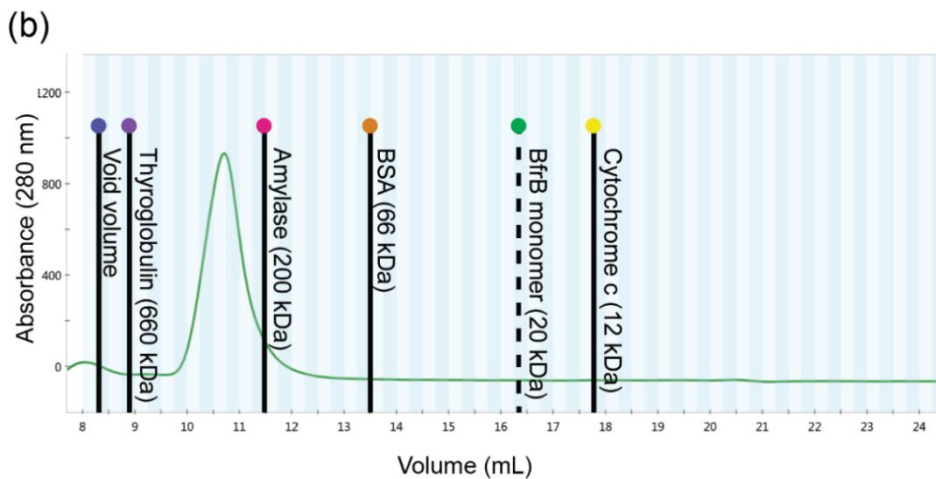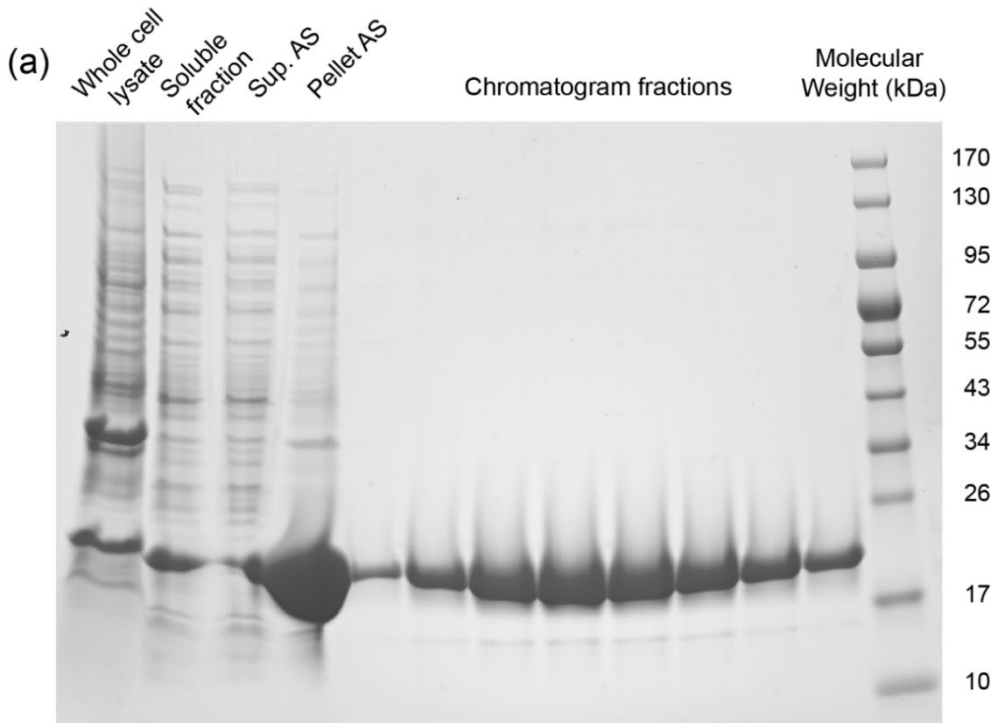
Supplementary Figure 2. 5 Purification of BfrB purification. (a) Denaturing gel electrophoresis of the different purification steps. Lysis of bacterial suspension (Whole cell lysate), removal of cellular debris by centrifugation (Soluble fraction), protein precipitation by ammonium sulfate (Supernatant AS and Pellet AS). Fractions of BfrB oligomer peak obtained from (b) size exclusion chromatography. Reference proteins used for column calibration positioned in their corresponding elution volume.

CHAPTER 3

COULD EGG WHITE LYSOZYME BE SOLVED BY SINGLE PARTICLE CRYO-EM?

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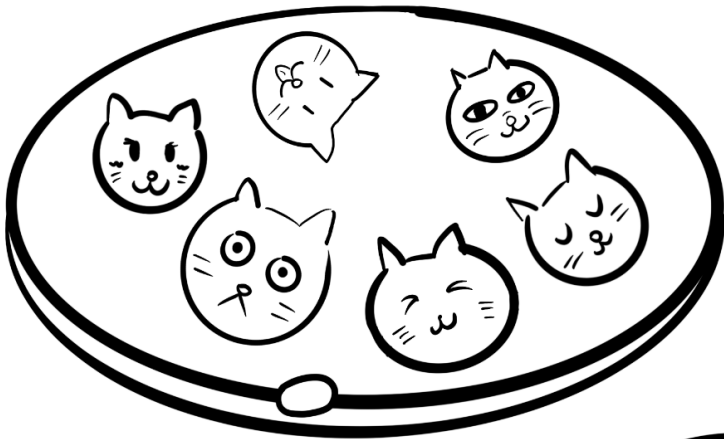
ABSTRACT

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 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 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.

3.1.   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[37,83,84]. Analysis of the size distribution of proteins in plants, animals, fungal and microbial species [50] 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 [48], resulting in high amounts of energy being deposited in the sample for a relative small number of electrons per Å^2 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 [85]. 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 [52,86].

The development of direct electron detectors with more stable microscopes and better data processing software has resulted in an increasing number of published high-resolution cryo-EM structures [43,53,55]. Currently, the highest resolution of a single-particle cryo-EM structure is 1.54 Å for apo-ferritin [54]. To date, the smallest protein solved by single-particle cryo-EM is 43 kDa [83]. Structural determinations of proteins under the 38-kDa theoretical size limit have still not been resolved by single-particle analysis [48,83]. The smaller a particle, the less scattering information that particle can provide. It is difficult to determine the five unknowns (shifts x, y and Euler angles α, β, γ) which limits the alignment of the noisy protein images [48]. The contrast transfer function (CTF), which is inherent with the mechanism of image formation in a microscope, oscillates the contrast of the particles as a function of resolution. It needs to be described accurately in order to be accounted for, which becomes more difficult at higher resolutions and at higher defoci. In conventional electron microscopy, contrast is enhanced by introducing a defocus: smaller particles are normally observed at higher defocus. Spatial and temporal incoherence of the electron source dampens the high-resolution signal at strong underfocus conditions, which impedes reconstructions schemes. Researchers are trying different ways to overcome these limitations, such as high protein concentration, minimal ice thickness [83], and the use of scaffolds [84]. A particularly promising route for small proteins seems to be the use of a phase plate (PP). An ideal PP would offer a $\pi/2$ phase shift without introducing any post-sample scattering. The Volta Phase Plate [87] has made phase plates accessible to a wide community; however, it cannot offer a constant phase shift and generates undesired post-sample scattering. Ongoing research aims to develop next-generation phase plates, such as a laser PP [88,89] or electrostatic PP [90]. Lower energy electron microscopy was proposed by Henderson [91,92], 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 datasets and careful computational analysis would be required to recover high-resolution information.

Simulations could help identify and characterise 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 [93], and could be used to evaluate the potential of new instrumentation improvements. Here, we adapted a TEM simulator developed by Vulovic, et al. [86], which is based on physical principles and considers the interaction between solvent, ions, and molecules. The simulator considers electron dose, which is important for TEM of biological samples because biomolecular structures can be altered by radiation damage inflicted by higher electron doses [94,95]. The radiation damage itself can be modeled by introducing a motion-blurring factor to all of the atoms [86]. Other characteristics, such as ice layer thickness, beam parameters, CTF, and detector performance, are also accounted for in the simulator. Simulations were validated by comparing the simulated micrographs with real experimental data [86,93].

First, we started with a subset of an experimental data set of human apo-ferritin (24-mer, 480 kDa) downloaded from EMPIAR-10216 [96]. Based on 29,224 particles extracted from this data set, we were able to obtain a 2 Å density map, faithful to the original experimental 1.64 Å map obtained with the full data set [96]. We then simulated a data set based on the known structure of apo-ferritin (PDB: 2fha) with the exact same parameters as the ones used for the experimental data set and compared the two. We show that our simulations match the existing data. Data processing of the simulated data set resulted in a 1.86 Å density map obtained from 18,062 extracted particles using the program Relion [97], which is in good agreement with a 2 Å density map obtained from the experimental data.

Next, we examined how the simulations of a protein below the 38 kDa theoretical size limit (which still remains to be resolved by single-particle analysis) would look like. We selected the 14 kDa hen egg white lysozyme (HEWL, PDB: 1dpx), a protein standard among X-ray crystallographers. We simulated data sets with and without ideal phase plate and tried to solve the HEWL structure from each dataset. Whereas we could obtain a near-atomic resolution structure from the data set with ideal PP, we were unable to do so for the data set without PP.

3.2. THEORY

Below, a summary is given of the theoretical framework used in this work. Most formulae have been described previously: the theory of image formation [86]; whether one could align the particles [48]; and how many particles are needed to achieve a certain resolution [72]. We recapitulate certain formulae and describe some modifications.

3.2.1. Image formation

The simulation of cryo-EM images of biological samples can be separated into four steps, (i) building of the specimen's interaction potential, (ii) electron propagation through the specimen, (iii) influence of the optics of the electron microscope and (iv) digital direct-electron detector response [86]. Here, we only briefly introduce the theory.

i) Interaction potential

The interaction potential of the specimen can be calculated using the isolated atom superposition approximation (IASA). For biological specimens in cryo-EM, particles are always embedded in ice. The IASA model takes into account the solvent, ions, and molecular interactions. The total interaction potential of the specimen is described as the sum of inelastic scattering potential (v_{ab}), which contributes to amplitude contrast, and elastic scattering potential $\mathcal{V}^{int}(r)$, which contributes to phase contrast.

$$\mathcal{V}_{tot}^{int}(r) = \mathcal{V}^{int}(r) + i\mathcal{V}_{ab} = (\mathcal{V}_{atom}(r) + \mathcal{V}_{bond}(r)) + i\mathcal{V}_{ab} \qquad \text{Eq. 3. 1}$$

where $r = (x, y, z)$ is the position of the electron wave. The combination of $\mathcal{V}_{atom}(r)$ ('atom' contributions) and $\mathcal{V}_{bond}(r)$ ('bonds' contributions) gives the elastic interaction potential $\mathcal{V}^{int}(r)$ of the specimen. The 'bond' contributions mainly come from the influence of solvent, ions and molecular interactions. $\mathcal{V}_{ab}$ is the inelastic part due to the interaction between the incident electrons with the free electrons in the specimen ($\Delta E \sim 20eV$) and atom cores ($\Delta E > 100eV$). It was calculated as the imaginary part of the model.

The solvent interaction potential is calculated from known density of water molecule [20]. The amorphousness of the solvent is modelled as a constant potential. Radiation damage was accounted for by applying a motion factor σ_M, which blurs the interaction potential of the atoms and bonds isotropically [86]:

$$\tilde{V}(q) = \tilde{\mathcal{V}}^{int}(q)\exp(-2\pi^2\sigma_M{}^2q^2) \qquad \text{Eq. 3. 2}$$

where $\tilde{V}(q)$ and $\tilde{\mathcal{V}}^{int}(q)$ are the Fourier transform of the $\tilde{\mathcal{V}}(q)$ and $\tilde{\mathcal{V}}^{int}(q)$, respectively. The relation between the motion blur and B factor is [98,99]:

$$B = 8\pi^2\sigma_M{}^2 \qquad \text{Eq. 3. 3}$$

ii) Electron propagation

A multi-slice algorithm was used to model the interaction between electrons and the specimen [100]. Each slice was 2 nm thick, which is a valid size for weak phase object approximation (WPOA),

as it has multiple scattering events less than 5% at 300 keV [101]. An incident plane wave $\Psi_0(x,y) = 1$ was iteratively propagated though N slices with slice thickness Δz, such that an exit wave $\Psi_{exit}(x,y)$ leaving the sample was calculated.

iii) Optics of TEM

After the wave propagated through the specimen and before passing by the lower piece objective lens, we still have the exit-wave $\Psi_{exit}(x,y)$ in real space. To include the optics of a TEM in the simulation, we multiplied the CTF to the Fourier transformed exit wave function to simulate contrast of the image as a function of spatial frequency (k). Aberrations due to defocus and spherical aberration are included in CTF. The temporal coherence and the spatial coherence of the electron source which dampen the CTF are calculated in envelope functions which was corrected from [86], following [102]:

$$K_c(q) = \exp(-\tfrac{1}{2}\,(\pi\lambda\delta)^2 q^4) \qquad\qquad \text{Eq. 3. 4}$$

$$K_s(q) = \exp(-\left(\tfrac{\pi\alpha}{\lambda}\right)^2 (C_s\lambda^3 q^3 + \Delta f \lambda q)^2) \qquad\qquad \text{Eq. 3. 5}$$

where the c and s represent the chromatic coherence and the spatial coherence of the electron source, respectively. The electron wavelength, calculated at relativistic speed, is given as λ .The spherical aberration is represented by C_s. The defocus value is Δf. The defocus spread δ is calculated as [102]:

$$\delta = C_c \left[4\left(\tfrac{\Delta I_{obj}}{I_{obj}}\right)^2 + \left(\tfrac{\Delta E}{V_{acc}}\right)^2 + \left(\tfrac{\Delta V_{acc}}{V_{acc}}\right)^2 \right]^{1/2} \qquad\qquad \text{Eq. 3. 6}$$

where C_c is the chromatic aberration. The terms of ΔI_{obj} and ΔV_{acc} indicate the instability of the objective lens current and the accelerating voltage. The term $\Delta E / V_{acc}$ is the intrinsic energy spread in the electron gun.

The PP was incorporated by introducing a phase shift $e^{i\Delta\phi}$ to the electron wave in reciprocal space and leaves the central transmitted beam unchanged. The electron intensity on the image plane $I_0(x,y)$ is given by:

$$I_0(x,y) = \left| \mathcal{F}^{-1}[[\mathcal{F}[\Psi_{exit}(x,y)\,]e^{i\Delta\phi}]\text{CTF}] \right|^2 \qquad\qquad \text{Eq. 3. 7}$$

iv) Detector response

Contributions to the detector response in our simulation include conversion factor, modulation transfer function (MTF), and detective quantum efficiency (DQE). The final detected electron intensity, $I(x, y)$, on the detector was calculated according to [86]:

$$I(x,y) = I_{rn} + I_{dc} + CF \cdot \mathcal{F}^{-1}\big[\mathcal{F}\big[P_{oiss}\big(\Phi_e \cdot \mathcal{F}^{-1}\big(\check{I}_0(q)\sqrt{DQE(q)}\big)\big)\big] \cdot NTF(q)\big]$$

Eq. 3. 8

where I_{rn} is the readout current and I_{dc} is the dark current. CF is the conversion factor of the detector, in $[ADU/e^-]$. P_{oiss} describes the Poisson distribution and weights the probability of arrival of an electron for a given dose and expected intensity [86]. NTF is the noise transfer function.

3.2.2   What is the size limitation of particles that can be aligned?

According to the theoretical calculation by Henderson [48], the protein orientation and position can be determined when $X_{sig} > x$. Here, X_{sig} is the signal-to-noise ratio of the image and describes whether the molecule can be detected or not. x is the multiple of sigma expected within the entire volume of five parameter space (shifts x, y and Euler angles α, β, γ) to be examined [48]:

$$X_{sig} = \frac{Signal}{Noise} \times \sqrt{N_{pix}}$$

Eq. 3. 9

$$x = \sqrt{2}erfc^{-1}[\left(\frac{d}{D}\right)^5 /3 \times 10^6]$$

Eq. 3. 10

where $N_{pix} = \frac{D^2}{(d/2)^2}$, is the number of pixels that corresponds to the area of one box containing the particle. 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 the other words, it is the lowest signal-to-noise ratio required for particle alignment. The cross-correlation coefficient for translation (shifts x, y) and rotation (Euler angles α, β, γ) are $(2D/0.2d)^2$ and $(2\pi D/0.2d)^3$, respectively. The value of x will not change much. The key parameter that determines whether a single particle can be aligned is X_{sig}. The derivation of the equation according to Henderson [48] can be found in the supplementary material. Here, the expression of X_{sig} is given by:

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

Eq. 3. 11

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

$$X_{sig} = \sqrt{\frac{\sigma_e \, \pi \, N_e \, D^3}{136}} \qquad\qquad \text{Eq. 3. 12}$$

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 [48] was 5 e$^-$/Å^2, with which he arrived at a 38 kDa theoretical size limit. Assuming a spherical protein with density of 0.8 Da/Å^3, a 38 kDa protein has a diameter of 45 Å. With the incoming 5 e$^-$/Å^2 electron fluence, this protein is not able to be aligned as X_{sig} equals to 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 micro electron diffraction data were obtained from lysozyme crystals at a fluence of 2.6 e$^-$/Å^2 [103]. However, in most modern SPA cryo-EM studies, the fluence is much higher than 5 e$^-$/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 [52]. 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, electron fluence greater than 5 e$^-$/Å^2 may still contribute to the signal up to Nyquist and surely helps with determining particle orientation and translation [72]. In Eq. 3.12, when we use 50 e$^-$/Å^2 for N_e:

$$X_{sig} = 0.068 \times D^{3/2} \qquad\qquad \text{Eq. 3. 13}$$

If we consider HEWL with a diameter of 32 Å, then X_{sig} equals to 12.3. Using a pixel size of 0.5 Å/pixel, d at Nyquist frequency is 1 Å. So, x equals 8 and $X_{sig} > x$. 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 [48], which varies from 0 to 1, the signal in Eq. 3.9 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?

3.2.3 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 [72]:

$$N_{part} = N_{inproj} \left(\frac{\pi D}{N_{asym}d}\right) e^{B/2d^2} \qquad \text{Eq. 3. 14}$$

Where N_{inproj} is the number of images needed per projection. The term $\frac{\pi D}{d}$ is given by the Crowther criterion [104], which describe the minimum number of unique projections needed for reconstructing a particle of diameter D to a resolution of d. N_{asym} 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_{inproj} , we get (supplementary materials and [72]):

$$N_{part} = \frac{1}{N_{asym}} \left[\frac{90\,\pi}{\sigma_e\,N_e\,d_N^2\,d}\right] e^{B/2d^2} \qquad \text{Eq. 3. 15}$$

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. 3.15 and rewrite it:

$$\frac{1}{d^2} = \frac{2}{B} \times \left[\ln(N_{part}) - ln\left(\frac{90\,\pi}{N_{asym}\sigma_e\,N_e\,d_N^2}\right) + \ln(d)\right] \qquad \text{Eq. 3. 16}$$

In Eq. 3.16, $\ln(d) \ll \ln(N_{part})$, therefore, we ignore $\ln(d)$. Thus, the final expression is:

$$\frac{1}{d^2} = \frac{2}{B} \times \left[\ln(N_{part}) - ln\left(\frac{90\,\pi}{N_{asym}\sigma_e\,N_e\,d_N^2}\right)\right] \qquad \text{Eq. 3. 17}$$

3.3   METHODS AND RESULTS

3.3.1   Image formation

Interaction potentials of the specimen were built using IASA. Non-overlapping 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 [105] 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 μm. 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. [105]. We simulated micrographs of human H-chain ferritin (PDB: 2fha) [106] and HEWL (PDB: 1dpx) [107]. For apoferritin, 166 micrographs (~ 20,000 particles) were simulated with an underfocus in the range of 0.2 to 1.3 μm. 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$^-$/Å^2 was used for each of the data sets. Power spectra of the micrographs were calculated with Gctf [70],and CTFs were fitted to these power spectra in the 30 - 2 Å resolution range. Data processing was done using Relion [43]. The motion blur factor σ_M was 0.5 to approximate the beam induced movement of the specimen which corresponding to the B-factor of 19.7 Å^2 according to Eq. 3.3. The quality of the simulated micrographs was compared to experimental ones from EMPIAR [108] (Figure 3. 1).
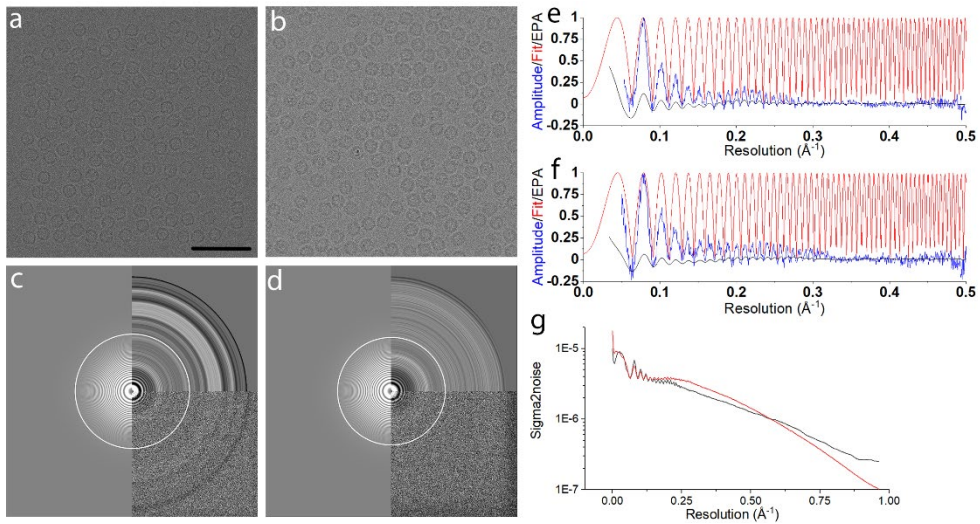
Figure 3. 1 (a) Simulated micrographs for apo-ferritin at 1220 nm underfocus (pixel size 0.5198 Å). (b) Experimental micrographs for apo-ferritin at 1210 nm underfocus (pixel size 0.5198 Å) (c) The Fourier transform of the micrograph shown in a, with equal-phase averaged, and CTF fitted by Gctf. (d) The power spectrum of the real micrograph shown in b, CTF fitted by Gctf. (e) Background-subtracted amplitude spectrum (blue) and the fit (red) from the simulated micrograph, the same as (c). (f) Background-subtracted amplitude spectrum (blue) and the fit (red) from the experimental micrograph, as in (d). (g) Sigma to noise spectra of the particles picked from the simulated micrograph shown in a (black) versus those picked from the experimental micrograph b (red). Scale bar in a represents 50 nm and applies to b.
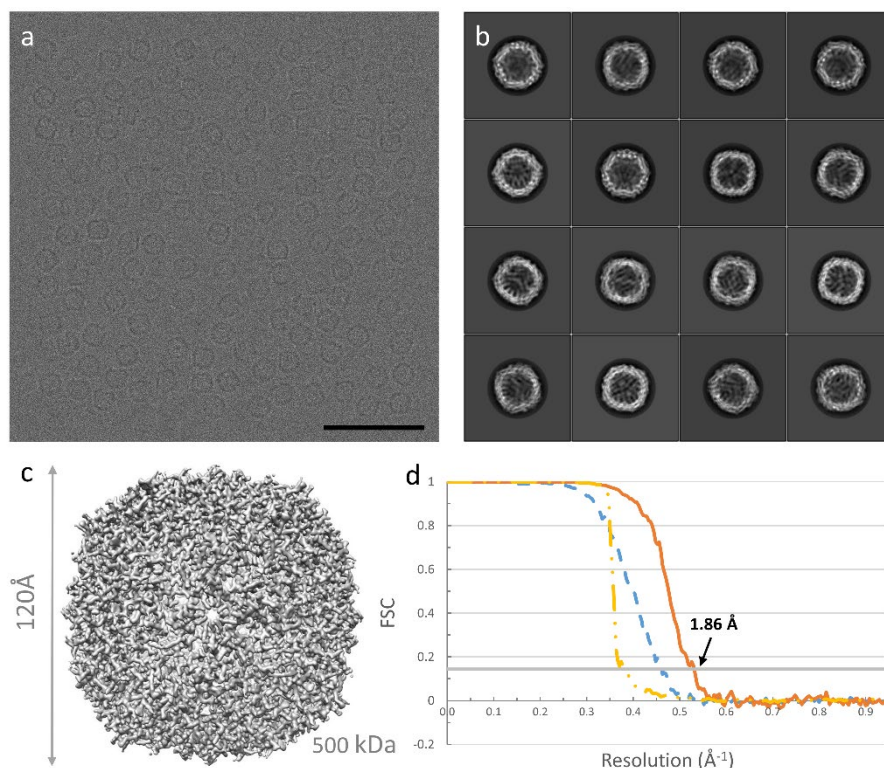
Figure 3. 2 Single-particle analysis of a simulated human apo-ferritin dataset. (a) A typical micrograph of apo-ferritin; the scale bar is 50 nm. (b) 2D class averages. (c) 3D reconstruction from 18,062 particles at 1.86 Å resolution. (d) Gold-standard Fourier Shell correlation (FSC) before (blue line) and after (orange line) masking, and the phase randomised FSC (yellow line).

3.3.2 Data processing

For the human apo-ferritin dataset, 20,214 particles were picked from 166 micrographs (Figure 3. 2a) with a pixel size of 0.5198 Å/pixel. Particles were extracted with a box size of 512 × 512 pixels. For 2D classification, we calculated 100 classes using a regularisation parameter T of 2 (Figure 3. 2b). As the data were simulated homogenously 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 [97]. Finally, Refine3D was run again, the map was post-processed and sharpened (B factor of -50 Å^2), yielding a final 1.86 Å resolution map (Figure 3. 2c).

To compare our simulated micrographs with experimental micrographs, we downloaded 448 experimental micrographs from EMPIAR-10216 [37] (Supplementary Figure 3. 1a). The pixel size was 0.5198 Å/pixel. The reported defocus was in the range of 0.2 um to 1.3 um. From these 448 micrographs, 49,962 particles were picked, and extracted with a box size of 512 × 512 pixels. 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, 5 classes were calculated using a regularisation parameter T equal to 4. After 3D refinement and post-processing, we obtained a 2 Å resolution map from 29,224 selected particles (Supplementary Figure 3. 1c).

For HEWL, we simulated micrographs both with and without PP. The PP used in simulation was an 'ideal' PP, which brought no post-sample scattering, introduced a fixed $\pi/2$ phase shift to the scattered electrons from the specimen, and had infinitely low cut-on frequency. The pixel size was 0.5 Å/pixel for both lysozyme datasets. Without PP, data were simulated with 2.5 to 4 μm underfocus (Figure 3. 3a). From 866 micrographs, 525,053 particles were picked and subjected to 2D classification with T equals to 2, using the 'ignored CTFs until the first peak' option (Figure 3. 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 3. 3c).

For HEWL with PP, the defocus ranged from 0.3 to 0.8 μm underfocus (Figure 3. 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 3. 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 × 256 pixel box and used to build an initial 3D model. This initial model was used as a reference map in 3D refinement with a 15 Å low-pass filter, with selecting 'ignored CTFs until the first peak' and no symmetry. The half maps of Refine3D were combined and sharpened using post-processing, applying a B factor of -116 Å^2, which resulted in a resolution map of 2.78 Å resolution based on gold-standard FSC [71,72] (Figure 3. 4c, 3.4d).
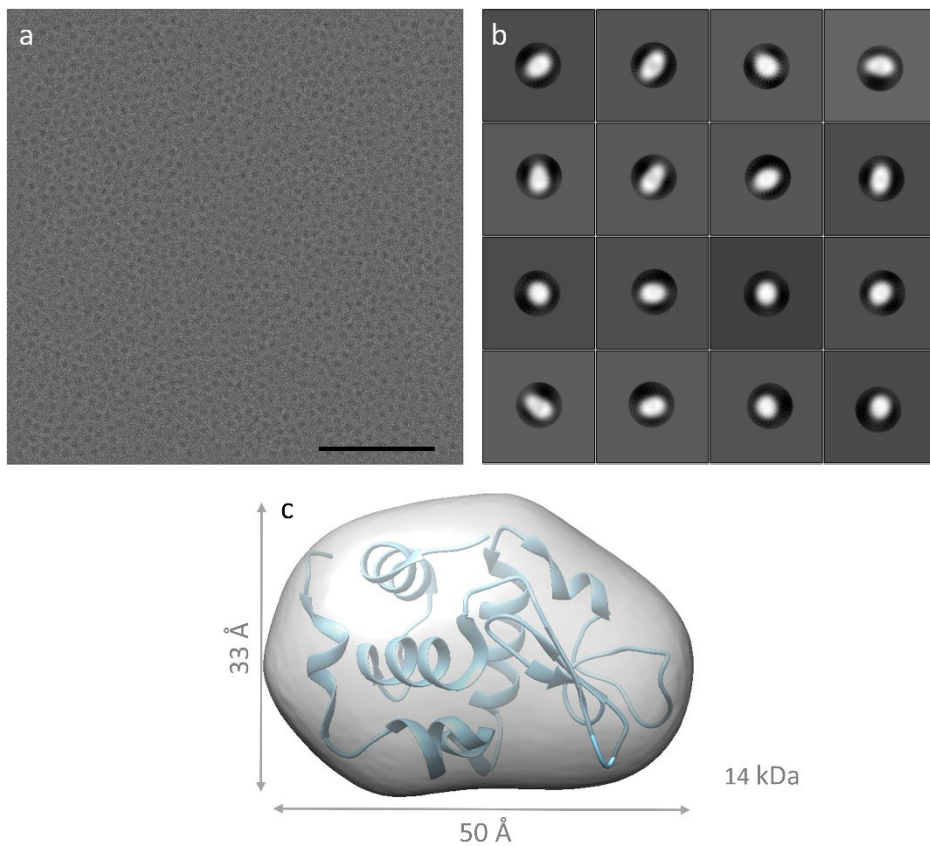
Figure 3. 3 Single-particle analysis of a simulated HEWL dataset without PP. (a) A micrograph of HEWL at 4.0 um underfocus; the scale bar is 50 nm. (b) 2D class averages. (c) The 3D reconstruction of HEWL without PP.
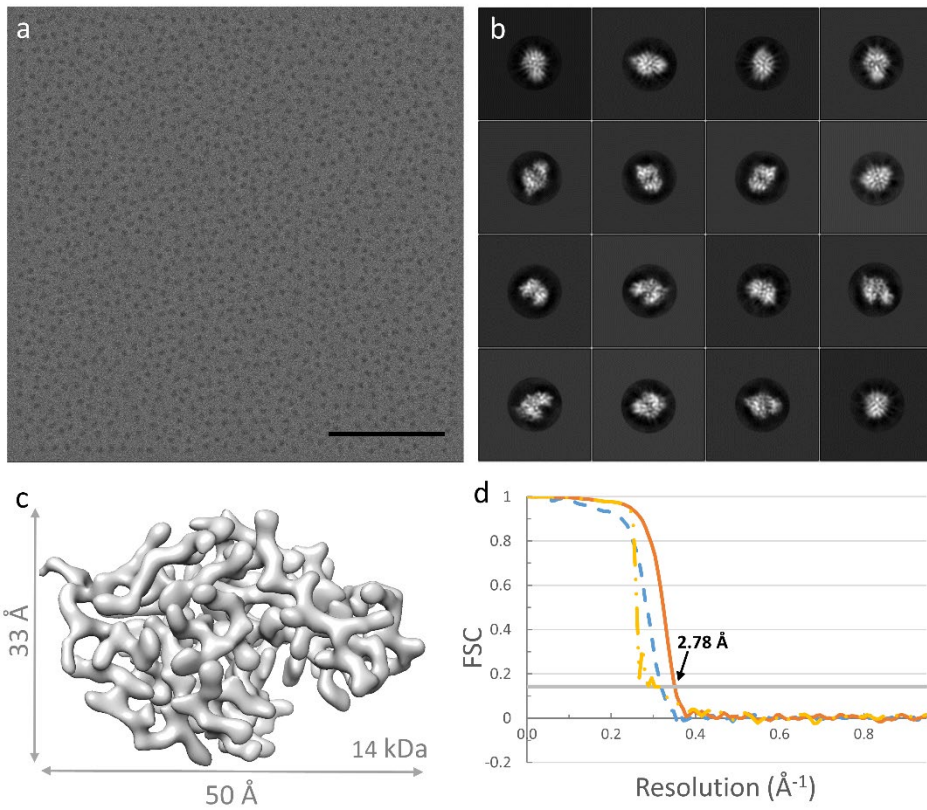
Figure 3. 4 Single-particle analysis of a simulated HEWL dataset with PP. (a) A typical micrograph of HEWL with PP; the scale bar is 50 nm. (b) 2D class averages. (c) The 3D reconstruction of HEWL at 2.78 Å resolution from 225,751 particles. (d) Gold-standard Fourier Shell correlation (FSC) before (blue line) and after (orange line) masking, and the phase randomised FSC (yellow line).

3.4 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 3. 1). The B-factor estimated by Gctf for the simulated micrographs equals to ~40 Å², slightly greater than what was obtained from the experimental micrographs (~30 Å²). The experimental data showed a somewhat stronger signal compared to the simulated one up to 0.5 Å⁻¹ (Figure 3. 1e vs 3.1f, Figure 3. 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 18062 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 inter-particle 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. 3.17 from the Guinier plots (Supplementary Figure 3. 3) was smaller for the simulated data (43 Å²) compared to the experimental dataset (54 Å²). 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 (Supplementary Figure 3. 3). The identical conformation of simulated particles improved the determination of the five parameters (shifts x, y and Euler angles α, β, γ) and particle alignment during 3D refinement in Relion. The estimated accuracy of angles and offsets for the simulated dataset with 4260 particles were 0.13 degrees and 0.2 pixels, respectively, whereas they were 0.32 degrees and 0.39 pixels for the experimental data with 4405 particles.

For HEWL, we observed clear differences between the datasets with and without PP after 2D classification. Without a phase plate, all particles collapsed in one 2D class unless we used the "Ignore CTFs to first peak" option in Relion. However, even then, most 2D classes were extremely blurred, reflecting the fact that the low spatial frequency information, which plays a dominant role in particle alignment, was extremely weak in these data. We were unable to obtain a good initial 3D starting map from the data itself. Starting the 3D refinement with a low-pass filtered map generated from the crystal structure 1dpx, we neither could obtain a higher resolution map (Figure 3. 3c). For the simulated HEWL dataset 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 3. 4b), with high rotational and translational accuracy, and high resolution (~3 Å). Post-processing revealed a B factor -117 Å^2, almost double than the B factor used for apo-ferritin (-50 Å^2). The higher B factor for HEWL compared to ferritin is probably due inaccuracies in orientation and translation alignments which are 2.085 degree and 0.721 pixel in 3D refinement.

The distinct differences between datasets 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 low spatial frequency signal while maximising 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. 3.9 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, preferable $\pi/2$, and introduce no/or small amount of post-sample scattering. It has been shown that the variable phase shift provided by the Volta Phase Plate can be computationally accounted for, however, its post-sample 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 post-sample 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 literature [83], however, these were most likely obtained by proteins attaching to an air-water interface [85], 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 modelled radiation 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 we one we used. An ideal phase plate does not exist yet: it should also be possible to integrate existing phase plate in the simulator. We deposited all our data sets in EMPIAR and distribute the source code of the simulator via GitHub [109], 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 optimised sample preparation [56], 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. Firstly, it could help novel cryo-EM users in data processing training with the unique feature that all parameters are known *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 afterwards 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.

Acknowledgements

Supplementary material

Theory

The expression of signal and noise is given by Henderson [48]:

$$\text{Signal} = 2\sqrt{\frac{\langle I_{obs}\rangle}{I_0}}\sqrt{N_s} \qquad\qquad \text{Eq. S3. 1}$$

$$\text{Noise} = \frac{1}{\sqrt{N_e d^2/4}} \qquad\qquad \text{Eq. S3. 2}$$

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.

$$\frac{\langle I_{obs}\rangle}{I_0} = \frac{f}{2N_s} \qquad\qquad \text{Eq. S3. 3}$$

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.

$$f = \frac{\sigma_e(c)}{3}\left[\frac{N_c}{D^2}\right] \qquad\qquad \text{Eq. S3. 4}$$

$$N_s = \frac{\pi}{2}\left[\frac{D}{d}\right]^2 \qquad\qquad \text{Eq. S3. 5}$$

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 $0.8\ \text{Da/Å}^3$. N_s is the number of unique diffraction spots to resolution d in projection. Combining Eq. S3.3, S3.4 and S3.5, the ratio between mean intensity and total incident beam is:

$$\frac{\langle I_{obs}\rangle}{I_0} = \frac{\sigma_e(c)}{3}\frac{4}{3}\pi\left(\frac{D}{2}\right)^3 \rho\frac{9.1}{110}\left(\frac{1}{D^2}\right)\frac{1}{\pi}\left(\frac{d}{D}\right)^2 \qquad\qquad \text{Eq. S3. 6}$$

As the signal calculated from the mean intensity of the diffracted beam as a fraction of the incident beam, we substitute d in Eq. S3.6 to d_N to indicate the resolution at the Nyquist frequency.

$$\frac{\langle I_{obs}\rangle}{I_0} = \frac{\sigma_e(c)\ d_N^2}{272\ D} \qquad\qquad \text{Eq. S3. 7}$$

By substituting Eq. S3.5 and Eq. S3.7 into Eq. S3.1, the signal from the particle is obtained as:

$$\text{Signal} = \sqrt{\frac{\sigma_e(c)\,\pi\,D\,d_N^{\,2}}{136\,d^2}}$$
Eq. S3. 8

The expression of X_{sig} is given by Eq. S3.9, which is the signal to noise ratio of the image. It describes whether the molecule can be detector not.

$$X_{sig} = \sqrt{\frac{\sigma_e\,\pi\,N_e\,D^3 d_N^{\,2}}{136\,d^2}}$$
Eq. S3. 9

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

$$X_{sig} = \sqrt{\frac{\sigma_e\,\pi\,N_e\,D^3}{136}}$$
Eq. S3. 10

The number of particles that need to build a density map with a certain resolution is given by [72]:

$$N_{part} = N_{inproj}\left(\frac{\pi D}{N_{asym}d}\right)e^{B/2d^2}$$
Eq. S3. 11

Where N_{inproj} is the number of images that need per projection. The term $\frac{\pi D}{d}$ is followed by the Crowther criterion [104] which describes the minimum number of unique projections needed for reconstructing a particle of diameter D to a resolution of d. N_{asym} is the asymmetric unit that a molecule has. B is the temperature factor that describes the effect of contrast loss.

$$N_{inproj} = \frac{\langle S \rangle^2 / \langle N \rangle^2}{\frac{\langle I_{obs}\rangle}{I_0}D^2 N_e}$$
Eq. S3. 12

Based on gold-standard Fourier shell correlation [72,110], $\langle S \rangle/\langle N \rangle = 1/\sqrt{3}$ when $FSC = 0.143$ and the map is interpretable. Thus, the number of particles needed to reach a certain resolution d is:

$$N_{part} = \frac{1}{N_{asym}}\left[\frac{90\,\pi}{\sigma_e\,N_e\,d_N^{\,2}\,d}\right]e^{B/2d^2}$$
Eq. S3. 13

Figures
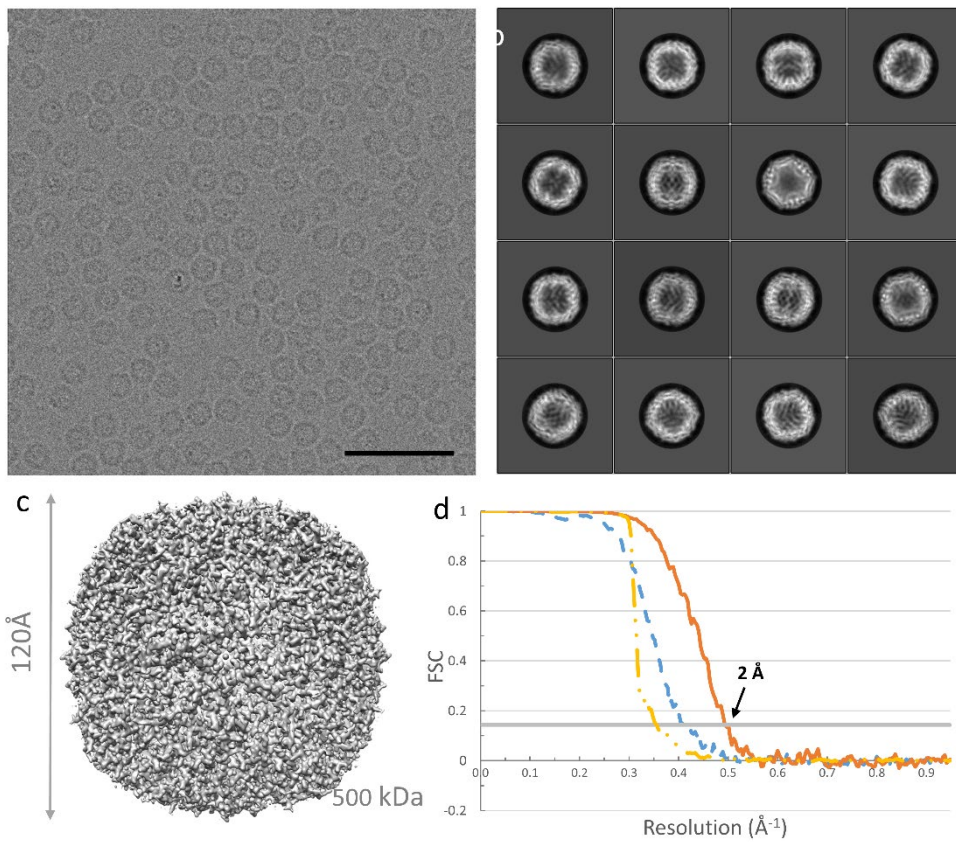
Supplementary Figure 3. 1 Single particle analysis of a subset of the experimental mouse apo-ferritin dataset from EMPIAR-10216. (a) A typical micrograph of apo-ferritin, the scale bar is 50 nm, (b) 2D class averages. (c) 3D reconstruction from 29224 particles at 2 Å resolution. (d) Gold-standard Fourier Shell correlation (FSC) before (blue line) and after (orange line) masking, and the phase randomised FSC (yellow line).
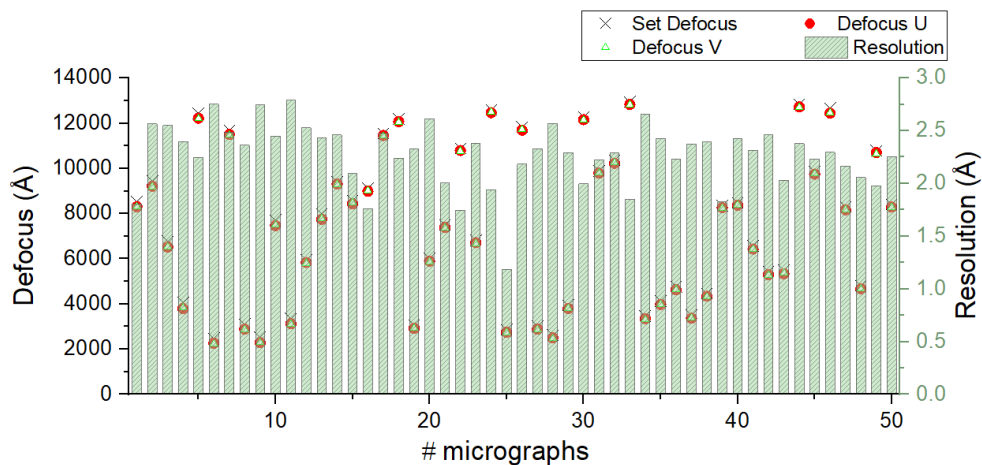
Supplementary Figure 3. 2 CTFs fitted with Gctf. 50 simulated micrographs for apo-ferritin have resolution higher than 3 Å.
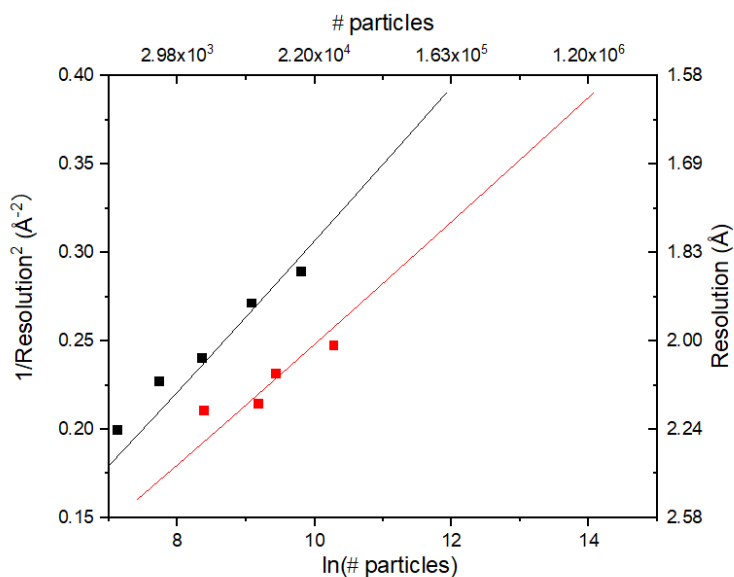
Supplementary Figure 3. 3 Number of particles and resolution plot based on Eq. 3.13. The B-factor for simulated data is 43 Å² (black), according to least squares regression line fitting. The fitted B-factor for the experimental data is 54 Å² (red).

Chapter 4

Charging of vitreous samples in cryo-EM mitigated by graphene

Yue Zhang[*], J. Paul van Schayck[*], Adrián Pedrazo-Tardajos, Nathalie Claes, Willem E.M. Noteborn, Penghan Lu, Hans Duimel, Rafal E. Dunin-Borkowski, Sara Bals, Peter J. Peters, Raimond B.G. Ravelli

[*]both authors contributed equally
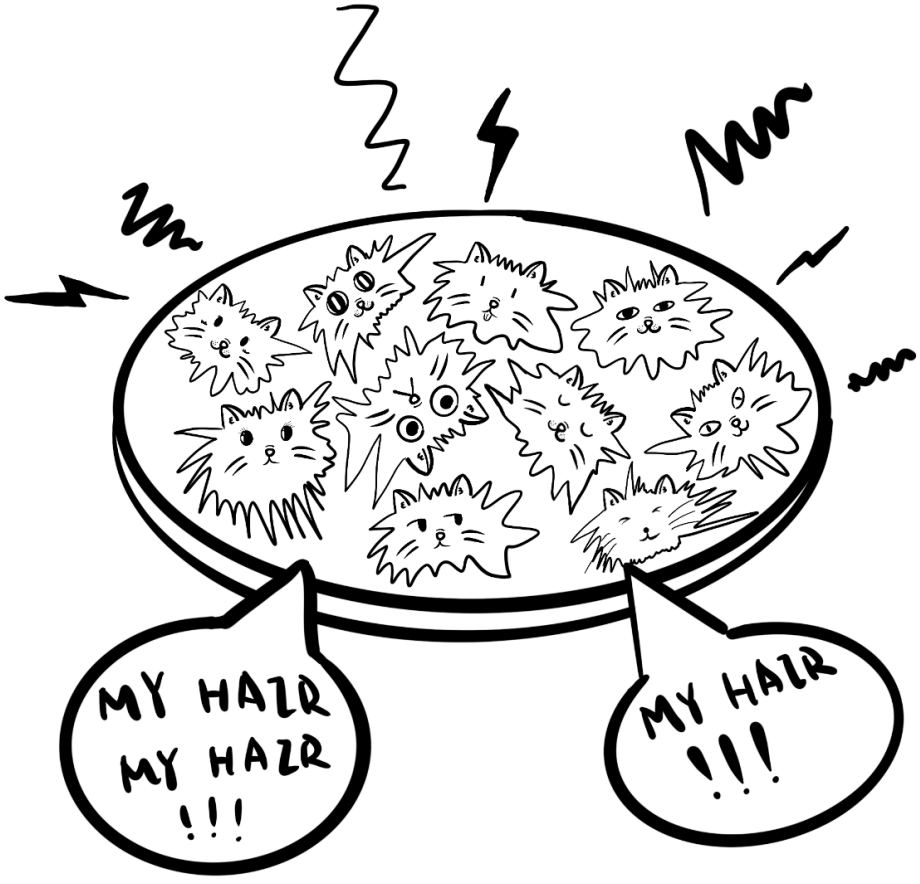
ABSTRACT

Cryogenic electron microscopy (cryo-EM) can provide high-resolution reconstructions of macromolecules embedded in a thin layer of ice, from which atomic models can be built de novo. However, the interaction between the ionising electron beam and the sample results in beam-induced motions and image distortions, which limit the attainable resolutions. Sample charging is one of the contributing factors of beam-induced motions and image distortions, which is normally alleviated by including part of the supporting conducting film within the region exposed to the electron beam. However, routine data collection schemes avoid strategies whereby the beam is not in contact with the supporting film, the rationale of which is not fully understood. Here we characterise electrostatic charging of vitreous samples, both in imaging and diffraction mode. We provide a way to mitigate sample charging by depositing a single layer of conductive graphene on top of regular EM grids. We demonstrate the ability to achieve high-resolution single-particle analysis (SPA) reconstructions at 2 Å when the electron beam only irradiates the middle of the hole on graphene-coated grids, while using data collection schemes that previously failed to produce sub-3 Å reconstructions without the graphene layer. We also observe that the SPA data obtained with the graphene-coated grids exhibit a higher b-factor and reduced particle movement compared to data obtained without the graphene layer. This mitigation of charging could have broad implications for various EM techniques, including SPA and cryo-tomography as well as for the study of the radiation damage and the development of novel sample carriers. Furthermore, it could also pave the way towards more dose-efficient, scanning transmission electron microscopy–based SPA techniques.

4.1 INTRODUCTION

Stimulated by the Resolution Revolution in cryo-EM [111], scientists made spectacular progress in pushing the limits of SPA, both in terms of resolution and particle size. SPA has yielded atomic resolution reconstructions [1,2], as well as reconstructions from particles smaller than the 38-kDa theoretical size limit [48,112]. SPA data collection has become much faster over the last years with the availability of faster detectors, advanced microscope automation, fringe-free imaging, aberration-free image shift and hole clustering [113]. This higher throughput has been combined with schemes to improve signal-to-noise ratios (SNR), such as DQE improvement of detectors, energy filters, lower energy spread of the electron source, enhanced phase contrast, and reduced sample movements.

Practically, it is well-known that upon irradiation by electrons, biomolecules embedded in a thin hole-spanning vitreous ice layer are observed to move [114]. However, the physical mechanisms behind sample movement are not fully understood. Several hypotheses have been proposed to explain this beam-induced motion [34]. It has been argued that the sample is placed under compressive stress upon rapid cryo-cooling. Electron radiation can induce creep in the presence of this stress that results in doming of the sample in the foil openings [32]. Another hypothesis implicates new mechanical stress from the breakage of chemical bonds and generation of hydrogen gas. Alternatively, electrostatic charging generates an attractive force which cause bending and warping of the thin flexible sample layer. Furthermore, it has been shown biomolecules can also appear to shift without physically moving themselves [115]. Non-ideal and dynamically changing lens conditions would result in image distortions in which molecules appear to move [116]. No matter the cause, movement or distortion during imaging leads to a reduced image quality by dampening high-resolution signals. Several schemes have been proposed to reduce beam-induced motions, such as devitrification [117], vitrification at low cooling rate and elevated temperatures [32,118], or the use grids with small holes to have low ice thickness/hole diameter ratio [33].

Charging of biological samples within the TEM has been discussed for decades [115,119-124]. The ionising electron beam produces secondary electrons that escape from the sample, thus leaving a positive charge on non-conductive specimens [123-128]. Charging can result in unwanted contrast changes, known as the "bee swarm effect", characterised by fluctuating granularity caused by random surface charging at low magnification and high-defocus conditions [125,129,130]. Russo and Henderson described that sample charging is a dynamic process that results from the poor conductivity of the specimen at low electron flux conditions [131]. The positively charged "footprint" from electron irradiation forms a microlens on the sample which deflects incoming electron beams, causing a change in phase contrast. This effect, known as the "Berriman effect", fades when the beam scans nearby regions [126]. The microlens can already be formed within a fluence range of ~10^{-3} to ~1 e$^-$/Å^2, and may contribute to the defocus change observed in the early frames of micrographs [122,123,132].

Sample charging is normally alleviated by including part of the supporting conducting film within the region exposed to the electron beam [34,115,126]. Curtis and Ferrier noticed that the "bee swarm effect" does not happen when part of the beam hits the metal grid bar even when the carbon film and the grid bar within the field of view are not connected [125]. Berriman and Rosenthal designed a special seven-hole C2 aperture, and demonstrated secondary electrons emitted from adjacent areas can reduce the charge on the area of interest [115]. Objective apertures can also reduce specimen charge: secondary electrons emitted by the objective aperture can compensate some of the positive charge of the surface of the specimen [120].

One successful technique to reduce beam-induced motion is "spot-scan imaging", which focuses an electron beam to a diameter of ~1000 Å and scans it over the specimen to capture multiple images [133-135]. However, this technique would cause charging on thin hole-spanning vitreous ice layer. This technique was only successful with specimens supported by a continuous conductive carbon film, which is undesirable for SPA due to the SNR reduction.

A suitable alternative to amorphous carbon is graphene, which has orders of magnitude higher conductivity than a carbon layer [38,136], gives minimum background noise and can be overlaid on top of a support layer. Graphene is used in materials science, and increasingly applied in the life sciences [36,38,137-142]. Ultra-flat graphene can result in uniform thin ice layers, allowing for high-resolution structure determination of sub-100 kDa proteins [36]. Others described that graphene can be functionalised to improve particle density and orientation [39,143,144].

In this paper, we investigate charging effects on vitreous biological specimens and demonstrate that charging can be mitigated by depositing a graphene layer on regular EM grids. We present high-resolution, conventional-TEM SPA reconstructions obtained with such grids, and discuss the importance of understanding charging for future conventional and non-conventional SPA schemes. Being able to mitigate charging by deploying graphene could help to further push the boundaries of resolving high-resolution structures of biomolecules via EM.

4.2 RESULTS

4.2.1 Evaluating the effect of charging in defocused diffraction mode

We used regular SPA samples and grids: *Mycobacterium tuberculosis* ferritin (BfrB) [145] applied to glow-discharged R1.2/1.3 Quantifoil and UltrAuFoil grids. The vitreous ice within the holes of this perforated film acts as an insulator, and a beam size smaller than these holes was used. Similar to Brink, et al. [126], we used defocused diffraction image (DIFF image) to observe the effect of charging (Figure 4.1a). A hybrid pixel Timepix3 detector [146-148] was used to record such images. No objective aperture was used in DIFF images/movies collection. As a function of

fluence, the size of the defocused diffracted beam increased in the overfocused condition (Figure 4.1b and c; Supplementary Movies 1 and 2) and decreased in the underfocused condition (Figure 4.1e and 4.1f; Supplementary Movies 3 and 4). The normalised DIFF beam size in both overfocus (Figure 4.1d) and underfocus (Figure 4.1g) became stable at a fluence around 1.5 e$^-$/Å^2. The change of the DIFF beam size seems to be insensitive to the type of foil material (carbon or Au) given that the beam is inside the hole. By comparison, the normalised DIFF beam sizes from a conductive crossline grating replica sample (Supplementary Movie 5 and 6) remained constant throughout irradiation.

To provide a more in-depth observation of how samples are charged and discharged, we collected DIFF movies with a moving stage, staying at each location for 5 s (fluence at 1.9 e$^-$/Å^2). At the beginning of irradiation, the diffraction lens was set in overfocus condition, and the beam was confined within a hole on vitreous sample (Figure 4.2a) until the DIFF image beam size fully expanded and became stable (Figure 4.2b). After that, the stage was moved so that the carbon foil came close to the beam. The DIFF image beam size started to decrease when the beam and foil were in close proximity (100 to 150 nm), yet not overlapping (Figure 4.2c). The DIFF image beam size decreased to its minimum when the beam was partially on the supporting foil (Figure 4.2d) and remained at this minimum even when the beam was completely on the foil (Figure 4.2e). The normalised DIFF beam size as a function of time (Figure 4.2f, Supplementary Movie 7) highlights the variation in beam size at various beam locations. There is a noticeable bump at around 50 seconds, reflecting a rapid charge-discharge process as the beam is scanned from one side of the foil to the other. DIFF movies collected in the same manner at a diffraction lens in underfocus condition (Figure 4.2g–l, Supplementary Movie 8) showed a similar yet inverse trend, wherein the DIFF image beam reached its maximum (instead of minimum) and remained stable near/on the foil then decreased (instead of increased) in holes. We repeated these experiments with UltrAuFoils and found the same trends (Supplementary Movies 9 and 10).
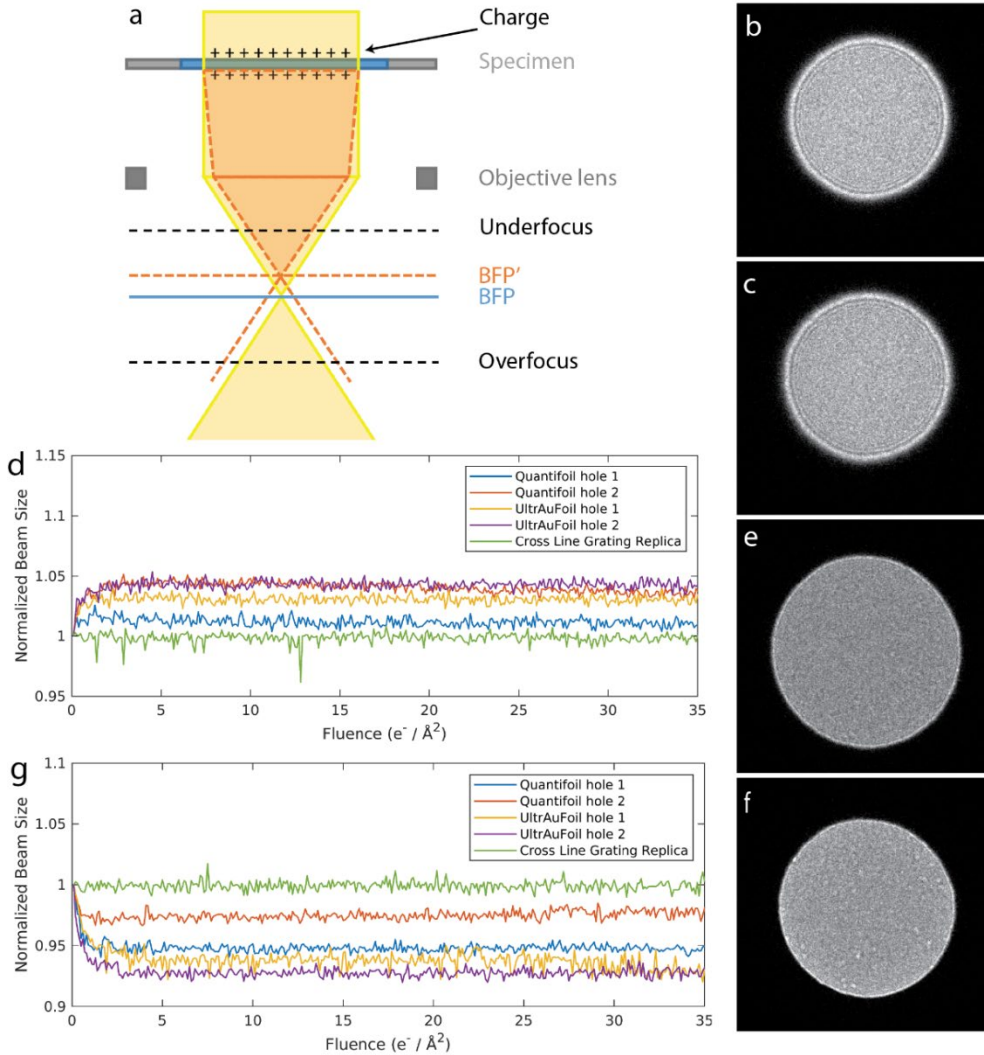
Figure 4. 1 Effect of specimen charging on both Quantifoil and UltrAuFoil shown by DIFF images. (a) Ray diagram of electron-optical effect of charge on the specimen. The electron beam (solid lines), which irradiates a specimen, is focused by the objective lens at the back focal plane (BFP). Charge on sample acts as a lens that converges the beam (dash lines) that induces the expansion of the overfocus pattern (b, c) and the shrinkage of the underfocus pattern (e, f). The DIFF images before (b, e) and after (c, f) 2 e$^-$/Å^2 irradiation of the specimen. The normalised beam radius is plotted as a function of accumulated dose in overfocus (d) and underfocus (g) conditions for both Quantifoil (blue, red curves) and UltrAuFoil (yellow, purple curves). For comparison, the normalised beam radius as a function of accumulated dose from cross line grating replica (Au) samples is shown (green curve). The electron beam flux on the sample was at 0.38 e$^-$/Å^2/s.
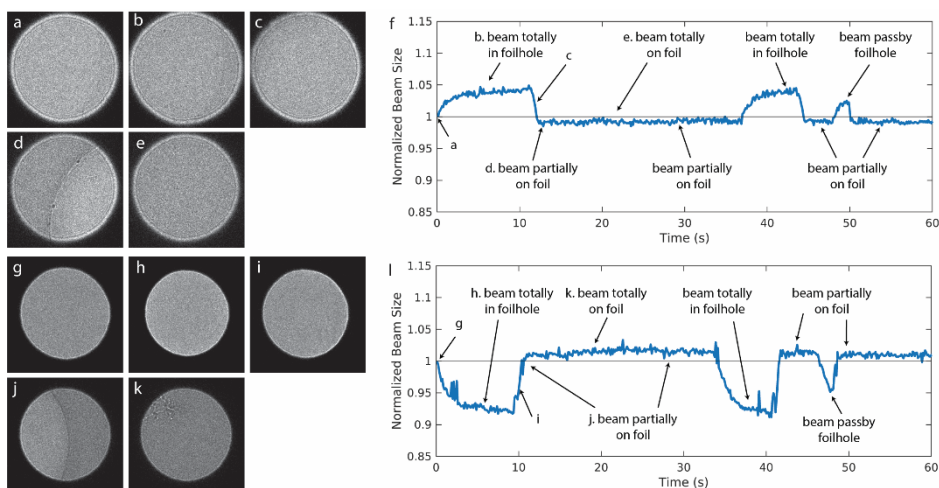
Figure 4. 2 Change of DIFF image beam size upon stage move on Quantifoil grids in overfocus (a–f) and underfocus (g–l) conditions. (f, l) The normalised beam radius is plotted as a function of time, with letters corresponding to the panel images shown to the left. In overfocus condition (a–f), the beam was completely within the foil hole on ice before irradiation (a), increased after a few seconds of irradiation (b), then decreased when the carbon foil moves close but not on the beam edge (c). The size of DIFF image decreased to its minimum when beam partially hits the carbon foil (d), and stayed the same when the beam is completely on the carbon foil (e). (g–h) In the underfocus condition, a similar but inverse trend was observed. The DIFF image size reduced a few seconds after irradiation (h) and increased when the carbon foil moved close to the beam edge (i, j), finally reaching its maximum when the beam hit the carbon foil (j, k). The electron beam flux on the sample was at 0.38 e$^-$/Å^2/s.

4.2.2 The use of graphene-coated grids

Next, we repeated the experiments above with grids (both Quantifoil and UltrAuFoil) with an extra graphene layer applied on top of them, to test whether the conductivity from graphene could alleviate charging. We verified the presence of graphene by collecting electron diffraction patterns from samples with amorphous ice on graphene (Figure 4.3a). These show the hexagonal diffraction pattern of graphene, demonstrating that it withstood the 10 s glow discharge, sample application, vitrification and grid handling. The DIFF beam size remained unchanged from the beginning of irradiation (Figure 4.3b) up to a fluence of 35 e$^-$/Å^2 (Figure 4.3c), in both overfocus (Figure 4.3d) and underfocus (Figure 4.3e) conditions. The DIFF movies with graphene grids are shown as Supplementary Movies 11–14. The DIFF beam size also remained stable for the DIFF image with a moving stage. The normalised DIFF beam size (Supplementary Movies 15–18) as a function of time at both overfocus (Figure 4.3i) and underfocus (Figure 4.3j) conditions showed that it remained unchanged regardless of beam location, whether inside the foil hole

(Figure 4.3f), partially on carbon foil (Figure 4.3g), or totally on carbon foil (Figure 4.3h). Notably, the location of the graphene layer (on top of/underneath the vitreous ice in microscope) did not affect the results (data not shown).
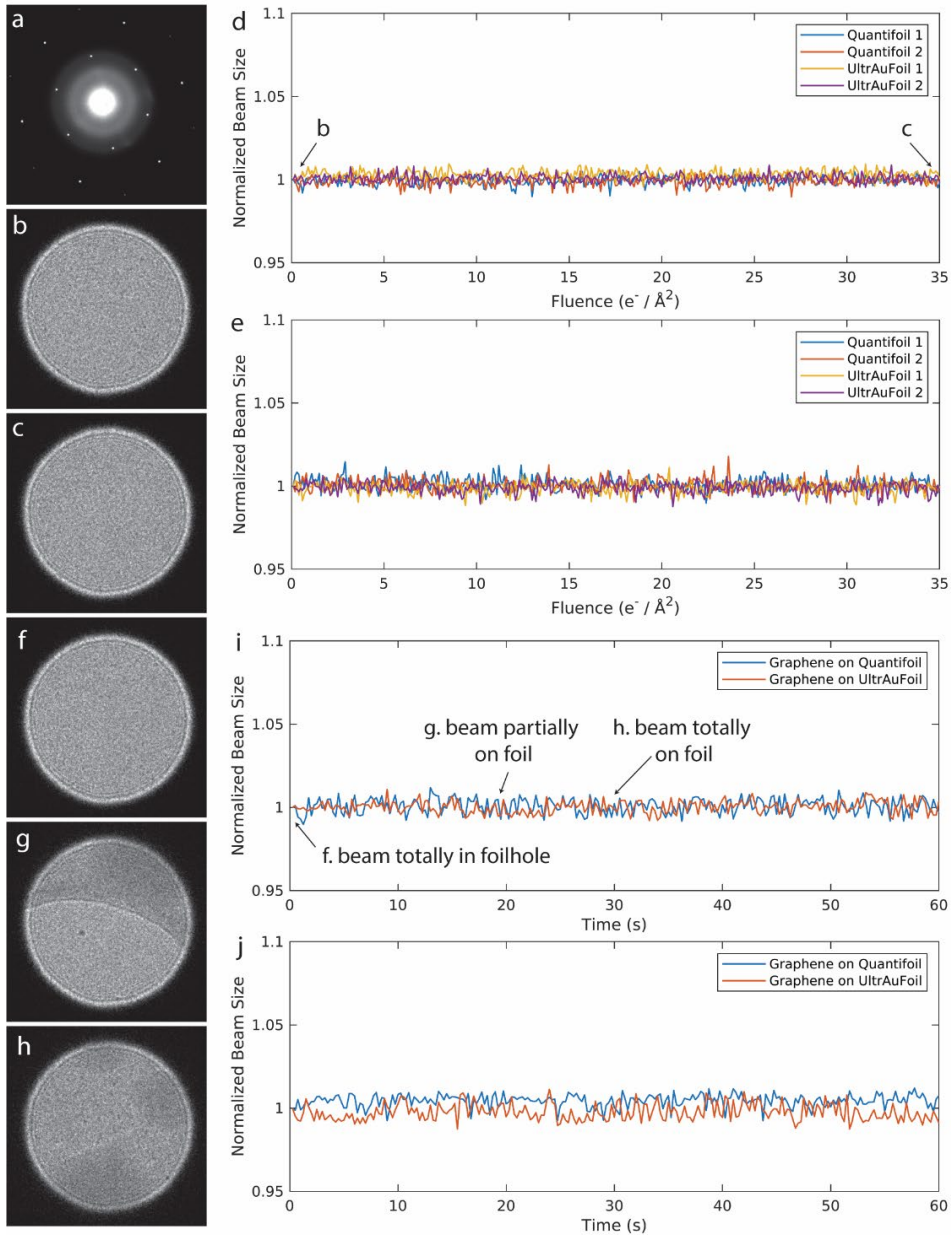
Figure 4. 3 DIFF images of vitreous specimen on Quantifoil with a graphene layer, and the similar results were obtained for UltrAuFoil with a graphene layer as well. (a) The diffraction

pattern of graphene with vitreous ice on it. (b, c) The DIFF images at overfocus conditions are similar before (b) and after (c) 35 e$^-$/Å^2 irradiation. (d, e) The normalised beam radius is plotted as a function of accumulated dose in both overfocus (d) and underfocus (e) conditions for graphene on both Quantifoil (blue, red curves) and UltrAuFoil (yellow, purple curves). When the stage moves, the beam moves from (f) foil hole to (g) partially on the carbon foil to (h) completely on the carbon foil. The DIFF image size was stable as the function of time in (i) overfocus and (j) underfocus conditions for graphene on both Quantifoil (blue curve) and UltrAuFoil (red curve). The electron beam flux on the sample was at 0.38 e$^-$/Å^2/s.

4.2.3 Charging in imaging mode

While the DIFF experiments showed clear charging effects up to a fluence of 1.5 e$^-$/Å^2, after which the beam size remained constant, typical SPA data collection schemes conducted in imaging mode uses fluences of tens of e$^-$/Å^2, which may be more prone to aberrations that are not readily apparent in diffraction mode. Thus, we used imaging mode with parallel illumination (Figure 4.1a) to further investigate charging in these regimes. Images were recorded at the flux of 30 e$^-$/Å^2/s for 1 s with Falcon III detector and no objective aperture at a nominal magnification of 78 000× and averaged without patch-track motion correction. We first used 20 µm C2 aperture to ensure a beam diameter (800 nm) smaller than the hole size (1.2 µm): the beam did not touch the perforated support film. This setup resulted in severely distorted images, as if the particles were moving outwards from a centre (Figure 4.4a). After 10 s, another image was taken at the same spot with 50 µm C2 aperture (beam size 1.9 µm) so that the beam hit the foil. The resulting image was sharp and similar to the initial state (Figure 4.4b), arguing that the sample had not undergone a plastic deformation, but a reversible process only that affected the image. This image distortion (blurring) and restoration (deblurring) is clearly shown in Supplementary Movie 19. To confirm that this effect is not related to the pre-exposure, we collected movies with apertures in the reverse order, first with 50 µm C2 aperture (Figure 4.4c) and then with 20 µm C2 aperture (Figure 4.4d). We found similar results, a sharp image with the larger aperture and a distorted image with the smaller aperture (Supplementary Movie 20). Importantly, when we performed the same experiments using Quantifoil grids with a graphene layer we consistently obtained sharp images, independent of the size of the beam relative to the foil hole size (Figure 4.4e and 4.4f, Supplementary Movie 21; Figure 4.4g and 4.4h, Supplementary Movie 22).

4.2.4 Single particle analysis

Next, we determined whether the use of graphene improved image quality for SPA structure determination. We collected SPA datasets of BfrB samples on Quantifoil grids with and without graphene (Table 4.1). Samples on graphene grids exhibited substantially reduced absolute and collective motion compared to samples on grids without graphene throughout the entire SPA fluence period (Figure 4.5a and 4.5b). We observed overlapping particles when using the

graphene-coated grids and could use a ten-fold dilution of the protein sample to arrive at similar number of particles per micrograph when compared to the grids without graphene. Under conditions where the beam is smaller than the hole and not exposing the supporting foil, we were unable to get a sub-3 Å reconstruction of BfrB using grids without graphene (Supplementary Figure 4.1). However, we could obtain a 2.01 Å reconstruction with graphene grids (Supplementary Figure 4.2). When the beam size was larger than the hole size, and touching the conductive support, we achieved reconstruction of maps at 2.12 Å resolution with Quantifoil grids without graphene (Supplementary Figure 4.3) and 1.90 Å resolution with graphene-coated grids using a similar number of particles (Supplementary Figure 4.4, Table 4.1).
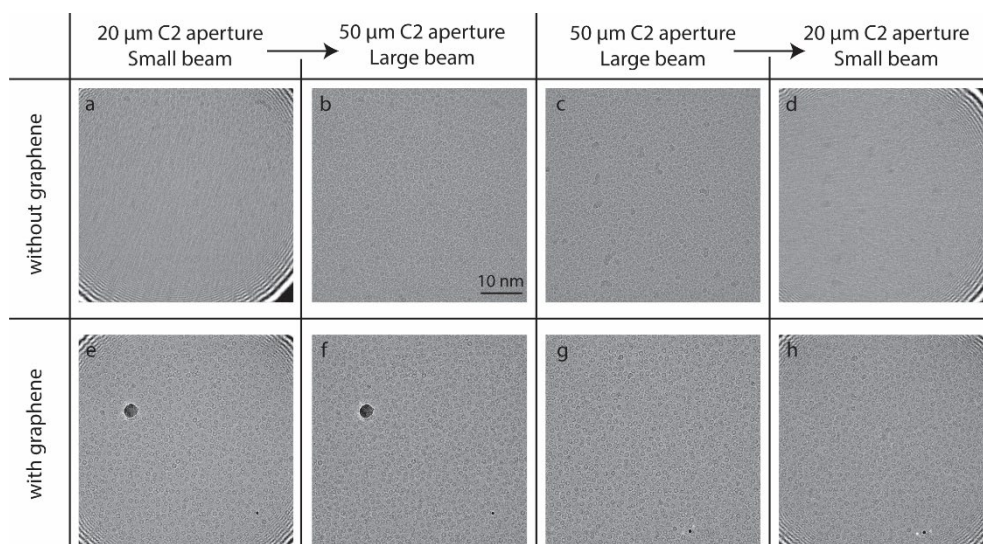
Figure 4. 4 Micrographs of BfrB sample collected at the magnification of 78 000 x on a Falcon III at 200 kV. All micrographs were collected at the flux of 30 e$^-$/Å^2/s for 1 s, fractions averaged without motion correction. No objective aperture was used. Micrographs (a) to (d) have samples at the concentration of 50 mg/ml on normal Quantifoil grids, and micrographs (e) to (h) show samples at the concentration of 5 mg/ml on Quantifoil grids with graphene layer. The grids have the foilhole size of 1.2 µm in diameter. (a, b) Two successive micrographs at the same position on Quantifoil grid, collected with 20 µm C2 aperture, beam size of ~800 nm for 1 s irradiation (a), then with a 50 µm C2 aperture, beam size of ~1.9 µm, for 1 s irradiation (b). (c, d) Two successive micrograph pairs at the same position on Quantifoil grids, but first collected with 50 µm C2 aperture for 1 s irradiation, then with 20 µm C2 aperture for 1 s irradiation. (e, f) Two successive micrographs at the same position on Quantifoil grid with a graphene layer, first collected with 20 µm C2 aperture, beam size of ~800 nm for 1 s irradiation, then with 50 µm C2 aperture, beam size of ~1.9 µm, for 1 s irradiation. (g, h) Two successive micrographs at the

same position on Quantifoil grids with a graphene layer, but first with 50 μm C2 aperture for 1 s, then with 20 μm C2 aperture for 1 s irradiation.
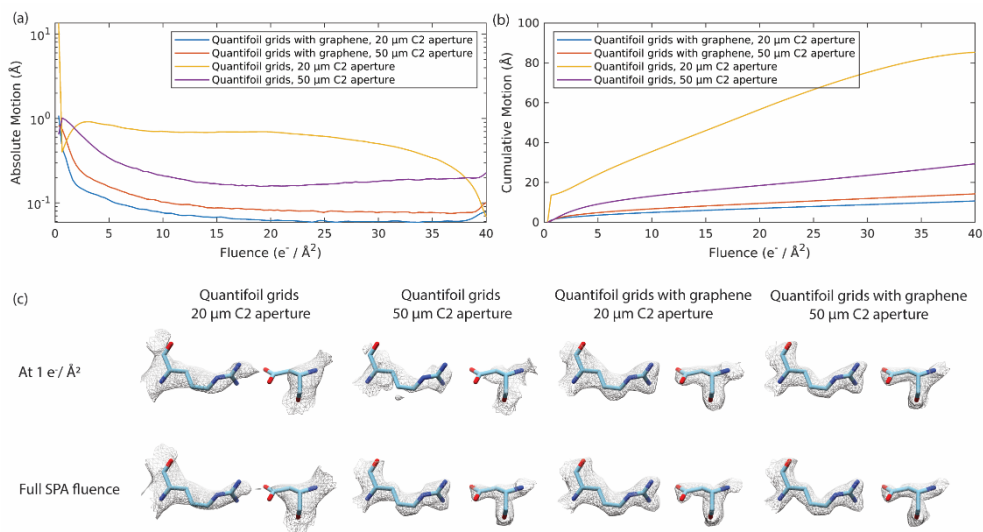
Figure 4. 5 Averaged absolute per-frame motion (a) and averaged accumulated motion (b) as a function of fluence determined by Relion Bayesian polishing of four datasets listed in Table 1. (c) Density maps reconstructed from four data sets each with a fitted BfrB model (PDB: 7O6E). Reconstructions from 1 e-/Å² (top row) and the full SPA fluence up to 40 e-/Å² (bottom row). Density maps are drawn at 1.5 RMSD.

Table 4. 1 Data statistics

Dataset	1	2	3	4
Grid type	Quantifoil 300 mesh R1.2/1.3	Quantifoil 300 mesh with Graphene R1.2/1.3	Quantifoil 300 mesh R1.2/1.3	Quantifoil 300 mesh with Graphene R1.2/1.3
Microscope	Titan Krios			
Voltage (kV)	300			
Objective aperture (μm)	100			
Nominal magnification (1000x)	105			
Physical pixel size (Å)	0.834			
Camera	K3			
Detector mode	Counting			
Focus range (μm)	-0.8 to -2.0		-0.6 to -1.6	
Exposure time (s)	1.7			
Flux (e$^-$/Å²/s)	23.5			
Fluence (e$^-$/Å²)	40			
Fractions (no.)	122			
Beam size (μm)	0.9		1.8	
Micrographs (no.)	633	621	2226	1808
Nr. Particles	85707	146626	596238	494154
Symmetry imposed	O			
FSC threshold	0.143			
Map resolution (Å)	3.50	2.01	2.12	1.90

4.3   DISCUSSION AND CONCLUSIONS

In this study, we examined the effect of charging on cryo-EM SPA samples and demonstrated that the addition of a graphene layer could mitigate this effect, resulting in higher resolution reconstructions allowing for improved SPA data collection schemes.

4.3.1 Charging of sample forms a non-ideal microlens

We set out to observe the effect of charging on cryo-EM SPA sample grids. From Brink, et al. [126], it is known that the effect of charging is particularly noticeable in defocus diffraction mode (DIFF), where it can be observed as a change in the size of the beam (Figure 4.1a). We found that beam size in DIFF mode became stable after a fluence of ~1.5 e⁻/Å² (Figure 4.1d and 4.1g), indicating that the charge saturates at this fluence, in accordance to Schreiber's findings [123]. The beam size change appears to be independent of foil material, here carbon and gold. However, the results shown in Figure 4.2 indicate they do relate to the distance between the beam edge and foil. In the experiment of Figure 4.2, the sample stage is moved while recording and the beam size started to change when foil moved close to the beam edge (100 to 150 nm), but before the beam hit the foil (Figure 4.2c and 4.2i). We speculate that the sample starts to discharge when the beam edge and foil are in close proximity. The electron irradiation induces the positive charge on the non-conductive sample surface with its area broader than the beam size. This charge produces a three-dimensional potential distribution that is extends further in all directions [123,124], with the electric field strength of more than a few MV/m [126,128]. This potential distribution can deflect incoming electrons and cause a drift of the beam (Supplementary Movies 7 – 10). These observations confirm the conclusions of Glaeser [34], who already indicated that charging leads to a deflection of the incident beam and results in the creation of undesirable lenses. Then the sample discharges when the beam hits the foil (Figure 4.2d and 4.2j). Close inspection of the normalised beam size indicates that a non-conducting SPA sample can become charged at an extremely low initial fluence: at the very first frame recorded with 0.047 e⁻/Å² (Figure 4.2f and 4.2l). Overall, the results shown in Figures 4.1 and 4.2 demonstrate that the positive charge induced on the surface of the sample by the electron beam forms a non-ideal microlens that causes the beam size to change in defocused diffraction mode.

4.3.2 Charging of the sample can severely hinder SPA

Surprisingly, whereas the beam size change observed in DIFF mode saturated around 1.5 e⁻/Å², we reproducibly observed continuing distortions in imaging mode, for the full range of fluences normally used in SPA (Figure 4.4a). While distortions were observed in DIFF image as well—features (e.g. ice contamination) were moving outwards even though the DIFF image beam size became stable after 1.5 e⁻/Å² (Supplementary Movie 3), these distortions were reversible, and restored as soon as the beam touched the conductive support film (Figure 4.4b, Supplementary Movie 19), which we attribute to sample discharging. Performing control experiments in reverse

order (Figure 4.4c and 4.4d) showed that the image distortion was not due to pre-exposure. We speculate that, as sample charging is a dynamic process [131], the continuing distortion of the image in imaging mode might be attributed to aberration effects, which occur due to charge redistribution on the sample surface, that further distort the image even when the absolute charge of the sample already reached a maximum. Such distortions have hindered routine SPA data collections at the centre of the hole using beam sizes smaller than the hole size. The SPA dataset we collected in this way displayed severe image distortions (Supplementary Figure 4.1a), and efforts to obtain a proper initial map and reconstruction were unsuccessful. Despite this, we could eventually obtain a correct BfrB reconstruction utilising modern motion-correction techniques combined with Bayesian polishing, albeit only at 3.5 Å and a low b-factor (Supplementary Figure 4.5). This result not only highlights the power of modern image data processing tools, but also provides a warning as the success of these programs might blind the user to the underlying physical phenomena that prevented them from getting better data.

4.3.3 Graphene mitigates effects of charging

Both DIFF image and imaging experiments illustrate that the conductive graphene layer can alleviate charging. DIFF image beam size was unchanged from the beginning of the irradiation to the full dose typically used in SPA (Figure 4.3d and 4.3e). The averaged images presented no blurring independent of the beam location, whether a small beam illuminated middle of a hole (Figure 4.4e) or touched the foil (Figure 4.4g). With a graphene layer, we were able to obtain a good SPA reconstruction of BfrB at 2.01 Å when the beam hits the middle of the hole (Supplementary Figure 4.2b). Additionally, for the BfrB protein that we used to perform the experiment, the graphene helped to concentrate the sample within the holes [39,144], as we could use a 10 times diluted sample compared to the grids without graphene, and still obtained a similar number of particles per micrograph (Table 4.1). The b-factor remained relatively unchanged regardless of the beam size for sample on graphene-coated grids (-82 Å^2 and -84 Å^2), however, it was significantly higher compared to that of Quantifoil grids (-178 Å^2) (Supplementary Figure 4.5). While our results demonstrate that graphene mitigated charging and indeed reduced motion (Figure 4.5a and 4.5b), it cannot be excluded that graphene may also reduce motion by improving sample stiffness.

When using graphene-coated grids, particle motions were smaller when a small beam irradiated only the middle of the hole compared to when a large beam irradiated both the vitrified ice and the foil, which might relate to the fact that smaller beams (with same flux) deposit less energy in the sample. The findings are consistent with the "spot-scan imaging" approach described by Downing [135], where a small beam was utilised to minimize the beam-induced motion. This technique expands the options for reducing motion beyond solely using small hole grids [33].

Nevertheless, early-stage, rapid sample motion [34,149] was observed in Quantifoil grids with and without graphene (Figure 4.5a), indicating that graphene does not prevent stress release at early exposure. Remarkable, our reconstructions of maps obtained from the first fractions of the data

were very similar to the final maps for the graphene-coated grids, whereas these early-exposure reconstructions show a compromised quality compared to the full exposure reconstructions for grids without graphene (Figure 4.5c). Several methods have been suggested to improve on the quality of the maps that can be obtained from the initial frames, including devitrification [117] and vitrification at low cooling rates and elevated temperatures [118]. However, these techniques may result in the formation of crystalline ice. Although the early-stage motion remains high, our data suggest that the use of graphene can still improve the quality of data obtained from these first frames. However, further investigation is necessary to fully characterise this effect.

4.3.4 Conclusion and future perspectives

This study focused on investigating the charging effects on cryo-EM SPA samples, and showed that the incorporation of a graphene layer could effectively alleviate the charging phenomenon. Fabrication of grids with high-quality graphene has been cumbersome due to poor reproducibility, low coverage rate and contamination [41]. We, and others [36], were able to overcome these limitations by using a high quality graphene obtained by an improved transfer method where the reproducibility and coverage rate of graphene grids are increased while minimizing contamination [150]. As a result, these clean graphene grids were able to provide the conductive layer needed to alleviate charging and improve current SPA schemes. While charging is fundamental phenomenon in EM, leading to reduced image contrast and resolution for SPA samples, it also applies for tomography lamellae and non-biological, non-conductive samples. A conductive layer is essential to minimise charging while imaging non-conductive samples, such as biomolecules in vitreous ice, in particular when the beam is not in contact with the conductive supporting layer. A graphene layer provides good conductivity needed to alleviate charging with minimum background noise. Graphene could also help to reduce the radiation damage and provide pristine atomic structures of the inspected samples [151]. Further investigations could be conducted to study the effect of graphene-coated grids on radiation damage in SPA samples, in order to gain a deeper understanding of the electron-induced radiation damage and obtain high-quality structural information of biological molecules. Graphene-coated grids with larger holes could be a promising avenue for future research, allowing for the collection of multiple micrographs with a smaller beam to reduce beam-induced motion and increase data collection throughput.

In summary, the implementation of techniques to mitigate charging is a crucial step in improving current SPA schemes, as well as provide pristine atomic structures of non-biological insulating samples. Charging mitigation could enable low-dose imaging or scanning transmission EM (STEM) techniques such as ptychography and iDPC [152,153]. It is worth noting that the effect of charge may differ in TEM and STEM modes, with local charging and discharging occurring in STEM mode depending on the beam's position relative to the sample. Further research is necessary to explore the impact of charging in STEM mode.

MATERIAL AND METHODS

Production of graphene-coated Grids

Graphene grids were prepared starting from a commercially available graphene-Cu foil, produced by chemical vapour deposition (CVD). First, a thin layer of cellulose acetate butyrate (CAB) was applied via spin coating in a CAB-ethyl acetate solution. The CAB-graphene-Cu stack was placed on top of an ammonium persulfate etching solution, etching away the Cu. The etching solvent was gradually diluted using deionised water to neutralise the solution and remove the Cu residues from the etched foil. Filter paper was placed at the bottom of the petri dish and the grids were placed on top using tweezers. The graphene was transferred onto the grids by removing the water using a micropipette. The filter paper containing the graphene-covered grids was placed on a heating plate for 30 min at 35 °C to dry. Finally, the graphene grids were heated in activated carbon for 15 h at 300 °C, which removed the CAB layer (whose melting temperature is between 170 °C and 240 °C).

Sample preparation and cryo-electron microscopy

M. tuberculosis BfrB was prepared according to Gijsbers, et al. [145]. It was used at a concentration of 50 mg/ml (as calculated with the Pierce BCA Protein Assay Kit) for Quantifoil and UltrAuFoil (Quantifoil Micro Tools, Germany) experiments. We used 300 mesh grids with 1.2 μm diameter holes. A volume of 2.5 ml was applied onto grids which were glow-discharged in vacuum at the current of 7 mA for 30 s. For grids with graphene layer, a volume of 2.5 μl diluted BfrB (1:10, 5 mg/ml) was applied onto mildly glow-discharged grids (current of 7 mA for 10 s). Excess liquid was removed by blotting for 3 s using filter paper followed by plunge freezing in liquid ethane using an FEI Vitrobot Mark IV operated under 95% humidity at 4 °C.

DIFF image collection

Diffraction images were collected with Timepix3 hybrid pixel detector [146-148,154] on a 200 keV Tecnai Arctica (Thermo Fisher Scientific). The beam was blocked with the pre-specimen beam shutter before exposing the sample for recording. The beam was set to be parallel at a flux of 0.38 e$^-$/Å^2/s passing through amorphous ice films. The DIFF image was obtained by defocusing the diffraction lens. No objective aperture was used for DIFF image data collection. Supplementary Movies are available at: https://surfdrive.surf.nl/files/index.php/s/jqHsLgveOLSixxQ

Single particle data acquisition and image processing

Cryo-EM single-particle data were collected on a Titan Krios at 300 kV with a BioQuantum K3 Imaging Filter with a 20 eV post-column energy filter. The detector was utilised in normal counting mode at a nominal magnification of 105 000 ×. Table 1 shows the statistics of the data set. Data were processed using the RELION pipeline [68]. Movie stacks were corrected for drift (7 × 7 patches) and dose-weighted using MotionCor2 [69]. The local contrast transfer function (CTF) parameters were determined for the drift-corrected micrographs using Gctf [70]. A first set of 2D references were generated from manually picked particles in RELION [68] and these were then used for subsequent automatic particle picking. Table 1 lists the number of particles in the final data set after particle picking, 2D classification and 3D classification with O symmetry. Beam-tilt parameters, anisotropic magnification and local CTF parameters were refined and the particles were polished [43].The resolution of the best final map was 2 Å using the gold-standard FSC = 0.143 criterion [71]. The maps have been deposited in the Electron Microscopy Data Bank as entry EMD-xxxxx.

ACKNOWLEDGEMENTS

SUPPLEMENTARY MATERIALS

Supplementary Figure 4. 1 Single-particle analysis of BfrB data set using normal Quantifoil grid, under conditions where the beam (900 nm) is smaller than the hole (1.2 μm) and not exposing the supporting foil. a) A micrograph of highly concentrated (50 mg/mL) BfrB in vitreous ice. b) 3D reconstruction from 85,707 particles at 3.50 Å resolution. c) Gold-standard Fourier shell correlation (FSC) before (red line) and after (blue line) masking, and the phase randomised FSC (yellow line).
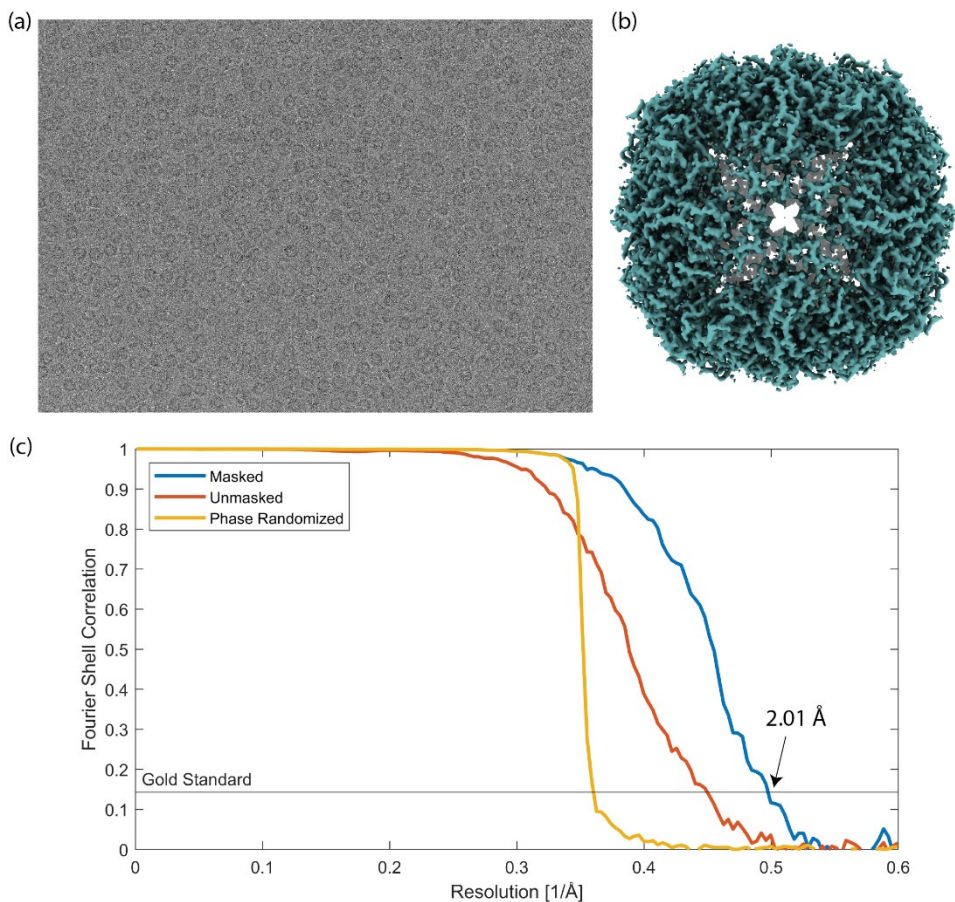
Supplementary Figure 4. 2 Single-particle analysis of BfrB data set using Quantifoil grids with an extra graphene layer, under conditions where the beam (900 nm) is smaller than the hole (1.2 μm) and not exposing the supporting foil. a) A micrograph of diluted (5 mg/mL) BfrB in vitreous ice. b) 3D reconstruction from 146,626 particles at 2.01 Å resolution. c) Gold-standard Fourier shell correlation (FSC) before (red line) and after (blue line) masking, and the phase randomised FSC (yellow line).

(a)

(b)

(c)
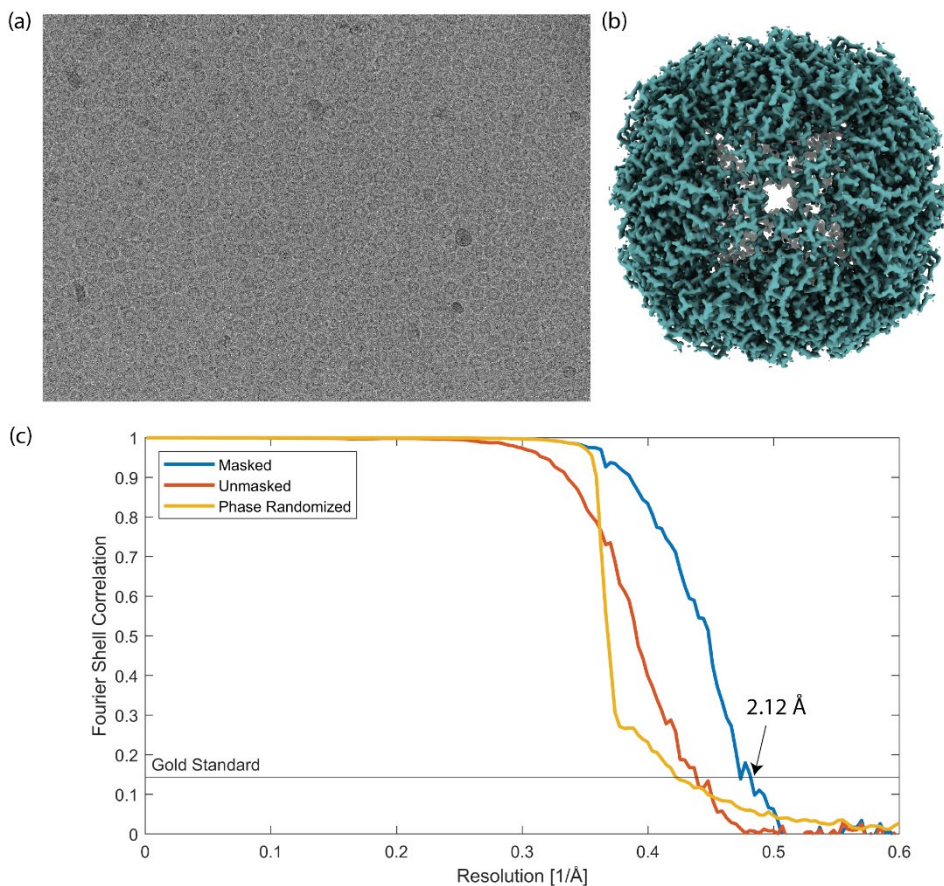
Supplementary Figure 4. 3 Single-particle analysis of BfrB data set using normal Quantifoil grid, under conditions where the beam (1.9 μm) is larger than the hole (1.2 μm). a) A micrograph of highly concentrated (50 mg/mL) BfrB in vitreous ice. b) 3D reconstruction from 596,238 particles at 2.12 Å resolution. c) Gold-standard Fourier shell correlation (FSC) before (red line) and after (blue line) masking, and the phase randomised FSC (yellow line).
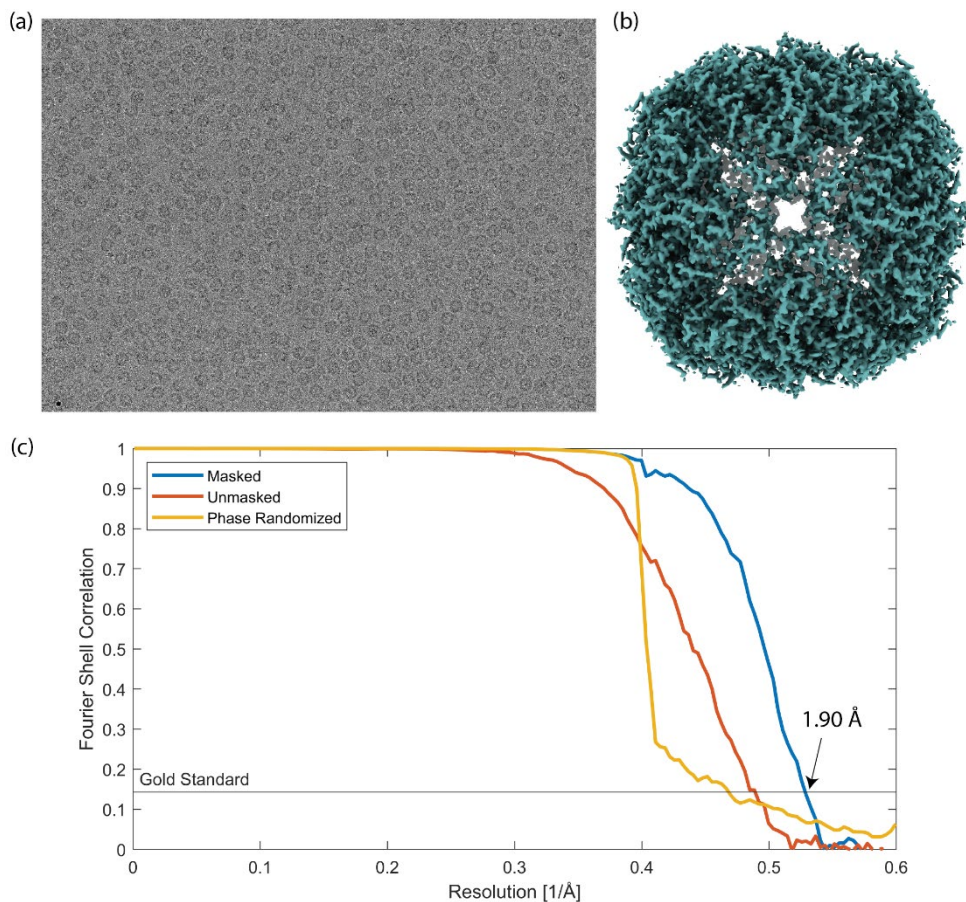
Supplementary Figure 4. 4 Single-particle analysis of BfrB data set using Quantifoil grids with an extra graphene layer, under conditions where the beam (1.9 μm) is larger than the hole (1.2 μm). a) A micrograph of diluted (5 mg/mL) BfrB in vitreous ice. b) 3D reconstruction from 494,154 particles at 1.90 Å resolution. c) Gold-standard Fourier shell correlation (FSC) before (red line) and after (blue line) masking, and the phase randomised FSC (yellow line).
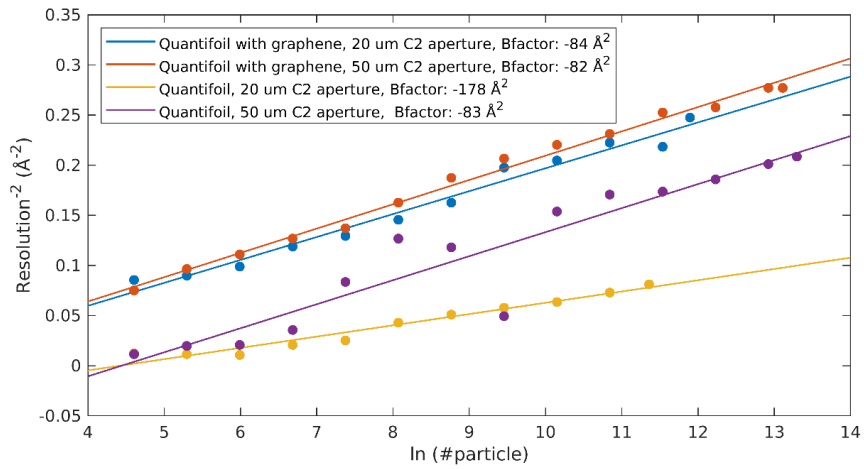
Supplementary Figure 4. 5 The Rosenthal-Henderson B-factor plot showing resolution versus number of particles. B-factors were estimated by fitting a straight line through the inverse of the resolution squared versus the natural logarithm of the number of particles for a range random subsets of full particle list [72].

CHAPTER 5

INTEGRATION OF AN EVENT-DRIVEN TIMEPIX3 HYBRID PIXEL DETECTOR INTO A CRYO-EM WORKFLOW

Yue Zhang[*], J. Paul van Schayck*, Kèvin Knoops, Peter J. Peters and Raimond B. G. Ravelli

[*]both authors contributed equally

A Hybrid
Pixel Box
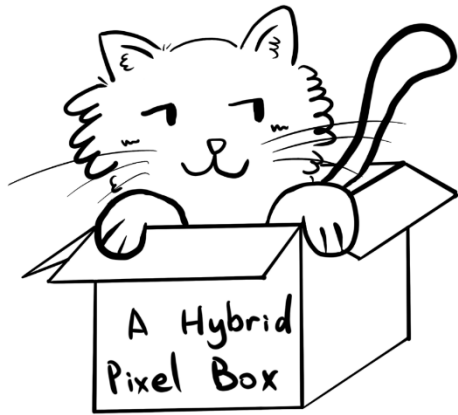
Abstract

The development of direct electron detectors has played a key role in low-dose electron microscopy imaging applications. Monolithic active-pixel sensor (MAPS) detectors are currently widely applied for cryogenic electron microscopy (cryo-EM); however, they have best performance at 300 kV, have relatively low read-out speed and only work in imaging mode. Hybrid pixel detectors (HPDs) can operate at any energy, have a higher DQE at lower voltage, have unprecedented high time resolution, and can operate in both imaging and diffraction modes. This could make them well-suited for novel low-dose life-science applications, such as cryo-ptychography, iDPC, and liquid cell imaging. Timepix3 is not frame-based, but truly event-based, and can record individual hits with 1.56 ns time resolution. Here, we present the integration of such a detector into a cryo-EM workflow and demonstrate that it can be used for automated data collection on biological specimens. The performance of the detector in terms of MTF and DQE has been investigated at 200 kV and we studied the effect of deterministic blur. We describe a single-particle analysis structure of 3 Å resolution and compare it with Falcon3 data collected using the same microscope. These studies could pave the way towards more dose-efficient single-particle techniques.

5.1. Introduction

Cryogenic electron microscopy (cryo-EM) of biological samples depends on the recording of a small number of incident electrons that can be used before the aqueous sample is destroyed by radiation damage [6]. The useful available signal in the recorded image is low and only barely above the shot and background noise. Therefore, subsequent processing of the data is highly dependent on having the best signal-to-noise ratio (SNR) throughout the process. There are many factors that determine the overall SNR in cryo-EM. These can be related to sample preparation, e.g. sample thickness, or related to the microscope, such as spatial and temporal coherence of the electron source and aberration of lenses [48]. The chosen microscope workflow also plays a very important role. The available signal will propagate differently and will have different noise influences, depending on the chosen imaging or diffraction technique [155]. However, no matter the workflow, microscope, or sample, nearly all cryo-EM workflows will use a pixelated detector to record and digitise the signal generated by the microscope.

Ideally, the pixelated detector faithfully records the entire signal generated by the microscope and does not add any additional noise to the signal. The ratio SNR_{out}^2 to SNR_{in}^2 is called the Detective Quantum Efficiency (DQE) of the detector [156]. Because pixelated detectors are inherently limited by their finite pixel size, the DQE is limited to the sinc function $\text{sinc}^2(\frac{\pi\omega}{2})$ as a function of spatial resolution ω. Practically, the DQE can be calculated by measuring the Modulation Transfer Function (MTF) and the Noise Power Spectrum (NPS or Wiener spectrum) and is scaled by the detection efficiency DQE(0) [157-160]. The normalised NPS (NNPS) is obtained after normalising NPS with NPS(0).

$$DQE(\omega) = DQE(0) \frac{MTF(\omega)^2}{NNPS(\omega)^2} \hspace{3cm} \text{Eq. 5. 1}$$

The emergence of direct electron detectors in the form of monolithic active-pixel sensors (MAPS) has been a breakthrough in cryo-EM [26,161]. Three developments surrounding MAPS detectors were pivotal in this breakthrough. First, the MTF of the detectors was significantly improved by back-thinning the sensor layer to reduce the chance that electrons scatter into adjacent pixels [162]. Second, the MTF was further improved by computationally counting single electron events, at the cost of limiting the electron flux below a coincidence level [163,164]. These two developments made MAPS detectors the first cryo-EM detectors to have significantly better SNR performance compared to film-based solutions. The third development was the ability of MAPS detectors to collect multiple frames per second (a movie), to correct for beam-induced motion and sample drift within a single recording [24]. Combined with advances in microscope automation and processing algorithms, MAPS detectors have enabled structural biologists to reveal macromolecular structures at near-atomic resolution using cryo-EM single particle analysis (SPA) [25,121].

However, the use of MAPS detectors is very restricted and is basically limited to this type of cryo-EM workflow. Due to limits in counting algorithms and radiation damage, they cannot be used in experiments with high electron flux or condensed direct beam, such as electron diffraction and ptychography. Although considerable improvements have been made, MAPS detectors have relatively low frame rates (< 1 kHz). Both their radiation fragility and low frame rates can primarily be attributed to their thin back-thinned sensor layers with little room for additional electronics or shielding. Furthermore, current MAPS detectors operate optimally at 300 kV; their MTF and DQE(0) deteriorate at energies below 300 kV [165]. Lower (i.e 100 kV) voltage may provide better SNR due to the more favourable elastic/inelastic scattering rations [48,92]. In addition, lower voltage microscopes might improve access to cryo-EM, due to their lower acquisition and maintenance costs [91]. Therefore, there is interest in developing detectors that can support more versatile workflows and work on a wider range of voltage levels.

Hybrid pixel detectors (HPDs) have many properties that can overcome some of these limitations. By definition, HPDs have separate sensor and electronic layers [154]. This makes them radiation-hard by design and allows additional electronics for much higher readout speeds [166]. On the other hand, this also comes with two important drawbacks, which so far have limited their practical use in many cryo-EM workflows. (1) Hybrid pixel detectors have a relatively large pixel size compared to MAPS detectors (i.e. 55 µm for Medipix HPD family, compared to less than 14 µm for MAPS detectors). This limits, in practise, the maximum field of view of HPDs in any kind of imaging application. (2) Electrons above 80 kV will scatter beyond a single (55 µm) pixel. In effect, the MTF of the HPD at energies above 80 kV deteriorates rapidly [167]. Overcoming these two important HPD drawbacks would make great strides towards HPD applicability in more cryo-EM workflows.

Two different methods have previously been shown to limit the effect of electron charge sharing between pixels in HPDs. The first method is to use a semiconductor material with a higher atomic number. Such a high-Z material reduces the distance from which the electron can scatter and thereby eliminates charge sharing. This is a promising and elegant method, but it comes with additional challenges, such as sensor cooling and crystal imperfections of the high-Z semiconductor material used as the sensor layer [160]. High-Z materials will also lead to a higher number of undetected electrons, thus damping DQE(0) due to increased backscattering compared to a silicon semiconductor. The second method is to electronically assign the sum of the charge of four surrounding pixels to the one pixel that has received the highest charge within that group of pixels. This Charge Sharing Mode (CSM) is present in the Medipix3 HPD and has been used to demonstrate near ideal MTF and DQE up to 80 kV electrons [160,168,169]. However, above 80 kV the charge spreads beyond the four-pixel group, which limits the benefits of CSM. The scattering of electrons through the sensor layer is erratic. They tend to shed most of their energy far away from the impact pixel, which limits the effectiveness of CSM, even if it could operate beyond a four-pixel cluster. Therefore, any sort of computational effort to reduce the effect of charge sharing needs to be more veracious.

In previous work, we have shown how the properties of the Timepix3 HPD can be used to count and localise individual 200 keV or 300 keV electrons to a sub-pixel location [147]. The Timepix3 is capable of simultaneously measuring the time a signal is over the comparator threshold (Time over Threshold; ToT) and the timing of when the signal crosses the comparator threshold (Time of Arrival; ToA) [148]. Previously, we explored these spectroscopic per-pixel properties of the Timepix3 and trained a neural network on simulated Timepix3 data to localise individual electrons from a pixel cluster. This negates the charge-sharing effect and by doubling the pixel matrix dimensions in a super-resolution mode mitigates the downsides of the relatively large physical pixel size. Here, we show how this detector can be incorporated in an actual cryo-EM workflow for a life science sample, and demonstrate its performance at 200 keV, an energy at which we could compare it with the commercial Falcon3 detector.

Cryo-EM workflows require a deep integration of the microscope and detector, both on the hardware and software level. First, the operation of the detector should not negatively affect the microscope itself. A detector comes with cooling and electronics: possible vibrations or electromagnetic effects generated by these should not compromise the quality of the data the microscope could bring. There is also integration at the software level. Often cryo-EM workflows run for hours, up to multiple days, in an automated fashion, where data collection decisions are made based on the output of the detector. These can be adjustments to the lenses to set focus, correct for aberrations, or verify the movement of the sample stage to the next acquisition location. Due to the radiation sensitivity of cryo-EM samples, the blanking and unblanking of the electron beam also needs to be tightly synchronised with the detector.

Turning on the beam, before the detector is enabled, would lead to the loss of precious signal. Different methods can be thought of to reach this level of hard and software integration [170]. Here, we present and explain the choices we made to obtain optimal performance of our Timepix3 setup for a cryo-EM workflow. We have chosen SPA as the cryo-EM workflow for this, being the most widely used by the community right now.

Beyond software and hardware integration, detector data need to be prepared in such a way that downstream processing software can handle it. The Timepix3 is an event-driven detector and outputs a sparse stream of data with, per electron, a cluster of hits containing position, time and energy. After event localisation, this sparse stream is reduced to a single event containing position and time for each electron. In the case of an SPA workflow, the processing software is not yet equipped to deal with this sort of sparse data. With Falcon4 and the EER format, the first steps towards event-driven readout for MAPS detectors have been made [1,171]. However, in its current form, EER is still based on frame-driven detectors, and it is not able to take advantage of the precise nanosecond timing of Timepix3. Therefore, we still had to convert our event stream to a classical movie of fractions (frames). These fractions must be gain-corrected and free from large defects, such as those caused by the edges of the chip. We describe how we have optimised the placement of these events in the image frames and how we have corrected for large defects.

This work describes how we have integrated a Timepix3 HPD into an automated cryo-EM SPA workflow. We measured MTF, NNPS, and DQE, and compared the results obtained with the commercial Falcon3 detector from the same protein and under the same microscope. We discuss how Timepix3 and future HPDs such as Timepix4 could enable the development of new versatile cryo-EM workflows, which could allow one to obtain better and complementary structural information within the limited dose lifetime of aqueous life science samples.

5.2 Experimental Setup

5.2.1. Hardware integration

The detector pod containing the Timepix3 HPD was mounted on a 200 kV Tecnai Arctica (Thermo Fischer Scientific) as previously described [147]. The detector pod (now commercially available from Amsterdam Scientific Instruments as CheeTah T3) provides the vacuum housing, radiation shielding, cooling, mechanical shutter, and houses the readout hardware. The SPIDR readout board developed by the Dutch National Institute for Subatomic Physics (Nikhef) is connected to the detector PC via a 10 Gbit/s fibre optic connection [172].

The detector pod cooling lines were initially connected to a dedicated chiller system (SMC HECR002-A5-FP), which cools the refrigerant to 18 °C. As refrigerant, water mixed with ThermoClean DC (Bioanalytic GmbH) was used to prevent algae growth in the refrigerant. The chiller was located a few metres from the microscope and had a reported, non-adjustable, flow rate of 120 L/h. Both the proximity and the relatively high flow rates of the dedicated chiller were a source of concern, as they may interfere with the operation of the microscope. Fortunately, we also had the possibility of connecting the Timepix3 to the existing central chiller used by the rest of the microscope, enabling us to measure possible interferences of different cooling regimes on the data quality.

We collected a series of high-resolution micrographs of a Crossed Line Grating Replica (EMS diasum, #80051). The micrographs were recorded using Falcon3 in electron counting mode, at a nominal magnification of 215 000 times, 40 s exposure, with a total fluence of 40 e$^-$/Å². Movie stacks were corrected for drift (rigid body) using Relion motion correction [43]. Power spectra were calculated from the resulting sum. An area of the 2.35 Å gold diffraction ring was selected and the mean pixel value in this area was calculated. For each condition, a series of 30 power spectra was recorded with 5-minute intervals. The conditions tested were, in order: dedicated chiller, dedicated chiller when not connected to the detector pod, central chiller at 5 L/h, and central chiller at 25 L/h. As a control, between different conditions, all Timepix3 equipment was turned off. The results are shown in Figure 5. 1. They indicate that there was interference from both the refrigerant that passed through the detector and the chiller itself. This interference could have been electromagnetic or due to vibrations. From these results we concluded that we should

cool our Timepix3 detector using the central microscope chiller at 5 L/h. During such an operation, the temperature of Timepix3 would only increased marginally from 42 °C to 47 °C.
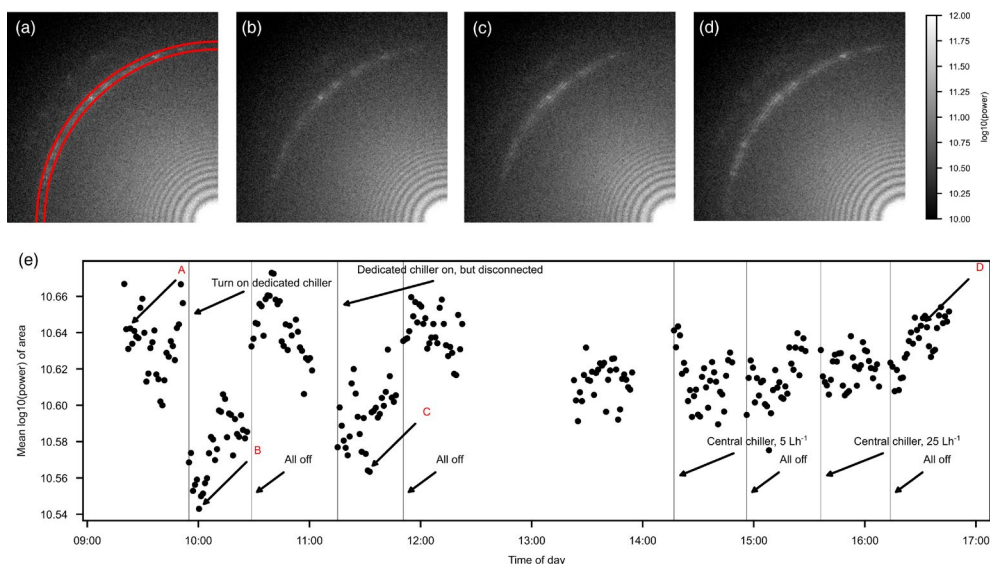
Figure 5. 1 Effect of different cooling and controlling regimes of Timepix3 on the intensity of a 2.35 Å gold diffraction ring (marked red in a). (a - d) show corners of the power spectrum of a motion corrected Falcon3 image recorded in electron counting mode, measured with dedicated chiller off (a), dedicated chiller on (b), dedicated chiller on but disconnected to the microscope (b), and dedicated chiller off (d) like in (a) but 8 hours later. (e) shows the mean intensity of the region marked in red in (a). This demonstrates that the dedicated Timepix3 chiller had a detrimental effect on data quality.

5.2.2. Software integration

The Timepix3 was integrated with SerialEM [173] using a custom camera plugin. This plugin provided communication between SerialEM and Serval, a network service developed by Amsterdam Scientific Instruments (ASI) which facilitates remote procedure calls between the detector PC and the Timepix3 SPIDR readout board. Figure 5. 2a shows the flow diagram schematically. Through the plugin and Serval, SerialEM is able to set the exposure time and can toggle the storage of raw, unprocessed Timepix3 data. In return, the plugin receives unprocessed preview images. These preview images are gain-corrected by SerialEM and used for montage, navigation, or autofocus procedures. Raw unprocessed Timepix3 data can be stored separately on the detector PC and transferred to network storage to be processed on individual workstations.

The synchronisation of the beam blanker and the camera was investigated by unmasking a noisy Timepix3 pixel. This particular noisy pixel blinked at 100 kHz regardless of any actual events. The ToA stamps of this pixel could be used to time the start and end of acquisition. Initially, control of the blanking and unblanking of the beam was performed by software within SerialEM. SerialEM would attempt, after calibration, to time the unblanking of the beam to coincide with the detector starting the acquisition. Because PCs are, without special adaptation, unreliable to do microsecond timings this was later changed to a direct electronic trigger for more accurate control. To do this, the SPIDR readout board had to be configured to output a 3.3V LVTTL signal on its HDMI channel whenever the Timepix3 electronic shutter is open. This signal was inverted (low when the Timepix3 shutter was open, high when it was closed) and connected to the microscope beam-blanking switchboard. Thermo Fisher Scientific had to allow a standalone camera to control the beam-blanker, which a service engineer could do by ticking an option within the Tecnai control software setup (Figure 5. 2c shows the now-enabled option). This direct electronic trigger of the beam blanker reduced the delay between detector acquisition start and beam unblanking to a very reliable 10 μs to 20 μs.

5.2.3. Electron event localisation

The processing of Timepix3 data into micrographs is largely similar to what we described previously [147]. Briefly, raw data are read from its binary Timepix3 format and ToA information is recalculated to one uniform clock. Some known defect pixel columns experience a phase-shifted ToA clock, which is corrected for. Next, non-uniformities in the ToT response are corrected for using previously obtained calibration data. Individual hits are clustered on the basis of being direct neighbours within a 50 ns ToA interval. Clusters are filtered according to their total sum of ToT values and the number of pixels affected. The filtered clusters are then passed to the convolutional neural network, which calculates a subpixel incident position based on the shape of the cluster and the ToT values. An overview of all individual steps is shown in green blocks in Figure 5. 3a.
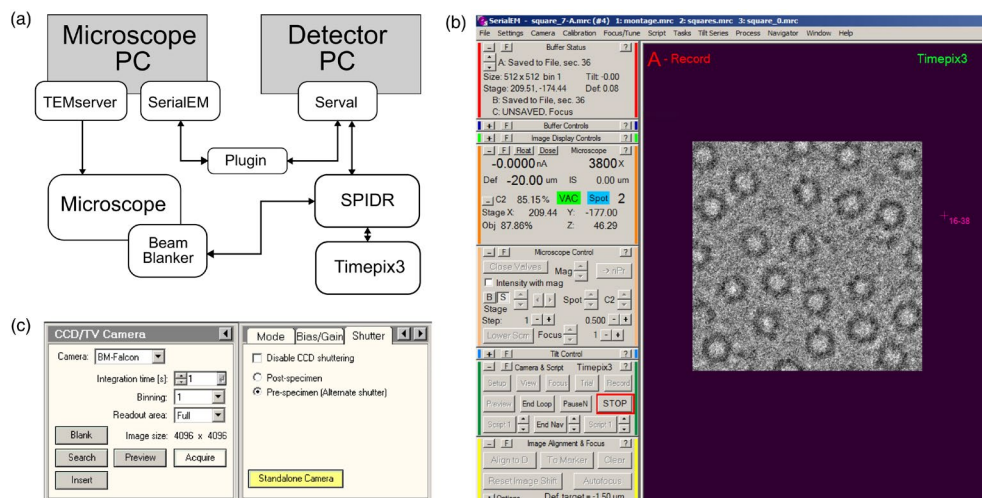
Figure 5. 2 (a) Flow diagram of how the Timepix3 is integrated in the microscope. (b) Computer screen shot of SerialEM controlling the Timepix3 through the custom plugin. (c) Computer screen shot of the Tecnai control software showing the standalone camera control button. Note that the Falcon3 MAPS detector is retracted at the same time.

To improve processing performance, two improvements have been made over the previous work. First, a new custom clustering algorithm was written. This algorithm works by first sorting hits in time, and then searching recursively for neighbouring pixels within a time limit. For extra performance, this algorithm was written in Rust and compiled as a standalone library to be called from Python. Second, a rewrite of the processing code was made to perform all processing steps on chunks of data in parallel. Together, this improved processing performance by more than 10-fold from 80 khit/s to 1 Mhit/s on the same hardware. A typical SPA exposure (120 Mhit, 50 $e^-/Å^2$, 1.25 Å pixel size, 700 MB) is processed in 2 min from raw data to sorted events.

Finally, we adjusted the cluster filter window. In our previous work, we had limited the filtering of the clusters [147](Figure 5. 2) to the main peak, thus excluding events such as those hitting edge pixels. The improved algorithms can make better use of these edge pixels. Therefore, the adjusted clustering window will now accept them, resulting in more events being used.
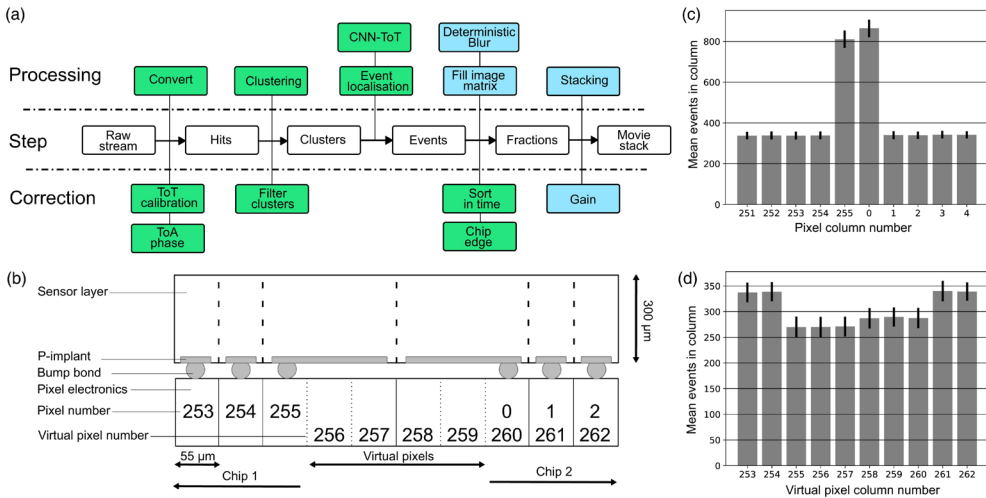
Figure 5. 3 (a) Flow diagram describing how Timepix3 data is processed from raw data to a movie stack. The green blocks show processing steps performed in tpx3HitParser [174]. The blue blocks represent processing steps performed in tpx3EventViewer [175]. (b) Schematic representation of edge pixels. (c and d) Mean values for the number of events in pixels columns between two chip edges, before (c) and after edge correction (d).

5.2.4. Generation and correction of image frames

From the sparse, event-localised data stream consisting of the position and time of each electron, individual image fractions (frames) are reconstructed over several steps. The complete processing workflow, after event localisation, is shown in blue blocks in Figure 5. 3a.

Our Timepix3 setup is in a quad configuration with four individual chips tiled together. Each chip has 256 pixels in each dimension, giving a total of 512 x 512 pixels. The chips are tiled together with a gap of 220μm, the width of four pixels, between them. This means that the pixels at the inner edges of each chip represent an area that is triple the width (165μm) of a normal pixel (Figure 5. 3b). Due to their larger size, they also receive more events (Figure 5. 3c), a bit less than three times more (Figure 5. 3d). To correct for this effect, events in the edge pixels are randomly redistributed to two virtual pixels created next to the original edge pixel. This results in an image frame four pixels larger in both dimensions (516 x 516). The remaining difference in pixel response between the three edge pixels and the rest of the chip can be compensated for via gain correction.

We split the event stream on the basis of their arrival time to generate fractions of arbitrary exposure time. During raw data processing, the fine ToA clock, the ToA, and the SPIDR global timer are combined to produce one uniform clock. All events are sorted on the basis of this

uniform clock, and chunks of events are split out on the basis of the desired per-fraction exposure time.

To form fractions and movies, localised events are placed in an image matrix in several different ways. The convolutional neural network returns, during event localisation, floating precision numbers which can be placed at, practically, any level of sub-pixel accuracy. The first and most naive method is to take the integer of the floating number and place the event in the corresponding pixel (Figure 5. 4b_1). Similarly, by upscaling the image matrix by a factor, it was possible to create super-resolution images at this factor. The second method is to apply a deterministic blur to the event, by upscaling the image matrix N times and then applying a Gaussian convolution of the up-scaled image with variance $N\sigma^2$. Next, the image matrix is Fourier-cropped to the desired final resolution (Figure 5. 4c_1 and d_1). We used $N = 10$ and determined that the optimal σ is 0.5 by calculating all the DQE curves for $\sigma = 0.1, 0.2, \ldots, 1.0$ (Figure 5. 4f).
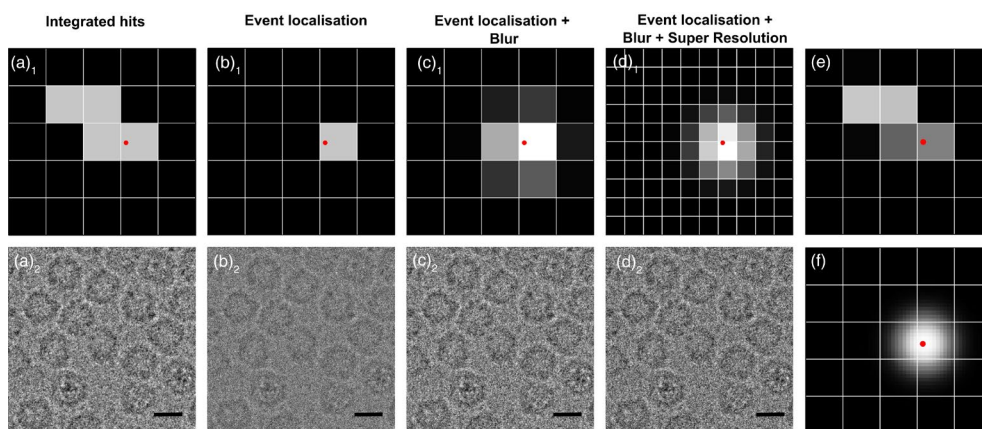
Figure 5. 4 (a1 – d1) Different methods used to generate a pixel cluster representing the same single 200 kV electron event. The red dot denotes the predicted incident location of the electron (CNN-ToT). a1 – c1 are 5 x 5 pixel matrices. d_1 is a 10x10 super resolution pixel matrix. Brighter pixels are assigned more arbitrary counts. (a2 – d2) Different methods used to form an example micrograph of the BfrB protein from the same raw Timepix3 data. Scale bar represents 10 nm. Micrographs have been gain and motion corrected. (e) The input pixel cluster used for the CNN-ToT prediction of the incident location. Brighter pixels received more energy (Timepix3 ToT counter). (f) The applied deterministic blur (Gaussian, $10\sigma^2$, $\sigma = 0.5$) of the same single electron event. The Gaussian is calculated in a 100 x 100 pixel matrix. This is used as precursor step to generate C and D, where the cluster has been Fourier cropped to the desired resolution.

For comparison, the raw Timepix3 data are also processed without clustering and event localisation. Instead, all individual hits are integrated per pixel (Figure 5. 4a$_1$, integrated hits) without further considering their ToT or ToA values. From there, these data are processed in a similar fashion as described above to correct for edge pixels and are sorted and split on the basis of their arrival time.

Finally, the fractions are gain corrected by multiplying with a normalised average image (gain) calculated from a series of flat-field exposures. Gain images are clipped to a maximum of five-fold difference to the median response. The gain was calculated for each method separately. The fractions were individually gain-corrected and placed together in one MRC movie file stack.

5.2.5. Measuring MTF, NNPS and DQE

The MTF was measured using the so-called knife-edge method as previously described [30,160]. Briefly, from an image of the knife edge, the angle and intercept of the edge were precisely determined using thresholding and a Sobel filter. The distance to the edge was then calculated for each pixel. From this an edge spread function could be fitted and the modulation transfer function calculated as Eq. 12 and 13 by McMullan, et al. [30]. In our setup, the edge image was made using a mechanical aluminium shutter placed directly above the Timepix3 detector. The shutter was placed at a 7° angle from the pixel matrix. For Falcon3, an edge image was made using the microscope beam stopper, which was at a 14° angle from the pixel matrix. For both detectors, a sufficiently straight edge was chosen without defects of the shutter or beam stopper.

NNPS was measured as previously described [30,160,176]. Briefly, a series of flat-field images was recorded. For Timepix3, this was done by taking a single exposure and splitting it into multiple fractions. For Falcon3, this was done by taking a single exposure in multiple fractions (a movie). The mean image of the series was calculated and subtracted from each frame in the series. The power spectrum then gives the 2D NPS, and its radial average is the 1D NPS [160](Eq. 2). NPS(0) was estimated by progressively Fourier-cropping the series of images by larger factors b. The variance of the images σ^2, normalised by the square of the binning factor, was evaluated. As b increases, $\sigma^2 b^{-2}$ reaches a plateau that was taken as NPS(0). The plateau was estimated by fitting with a logistic function.

From MTF and normalised NPS (NNPS), DQE can be calculated if the scaling factor DQE(0) is known (Eq. 5.1). DQE(0) can only be calculated from a known gain factor [160](Eq. 4) which, in turn, requires precise knowledge of the beam current. For low-beam currents, this is typically done using a Faraday cup. Our experimental setup did not allow for the positioning of a Faraday cup, and hence we were unable to directly measure DQE(0). Figure 11 from Paton, et al. [160] shows the DQE(0) for a Medipix3 HPD at 200 kV in single-pixel mode (0.8) and in charge-sharing mode (0.9). Taking into account that the minimum operating threshold for Medipix3 is 40% higher than Timepix3 (700e$^-$ vs 500e$^-$), there is no indication Timepix3 DQE(0) at 200 kV

would be lower than 0.9. Therefore, we made the safe assumption that DQE(0) at 200 kV for Timepix3 is at least 0.9.

We are unaware of a published DQE(0) for a Falcon3 operated at 200 kV in electron counting mode. In Kuijper, et al. [105] Falcon3 DQE(0) in integration mode at 200 kV and 300 kV is reported to be, respectively, 0.35 and 0.50. In electron counting mode, DQE(0) at 300 kV is reported to be 0.85. Therefore, we made the safe assumption that DQE(0) at 200 kV Falcon3 in electron counting mode is at most 0.85.

5.2.6. Single-particle analysis workflow

Protein samples of *Mycobacterium tuberculosis* apo-ferritin (bacterioferritin B, BfrB) were prepared as previously described [145]. Purified BfrB was used at a concentration of 40 mg/ml. A volume of 2.5 μL was applied onto glow-discharged UltrAuFoil Au300 R1.2/1.3 grids (Quantifoil). The excess liquid was removed by blotting for 3 s (blot force 5) using filter paper followed by plunge freezing in liquid ethane using a FEI Vitrobot Mark IV operated at 100 % humidity at 4 °C.

Timepix3 single-particle data were collected at 200 kV using SerialEM. SerialEM was directed using scripts adapted from single-particle scripts from the SerialEM Script Repository (https://serialemscripts.nexperion.net/). The Timepix3 required some specific settings, in particular to account for the relatively small field of view of the camera. The tilting angle of the beam used during autofocus had to be tuned and a relatively low magnification had to be used for the hole-centring alignment acquisition such that the entire hole of the grid could be obtained. SerialEM-directed gain correction allowed for correction of the chip edges without accounting for the physical size difference of these pixels. One image per hole was acquired using stage movements to navigate between holes. We waited a minimum of 10 seconds to allow the beam and stage to settle before each data acquisition. Data could be collected autonomously, for up to 72 h in succession, only interrupted by auto-filling of the microscope liquid nitrogen dewars.

Table 5. 1 shows the statistics of the data sets. The same Timepix3 data were processed in four different datasets. Dataset one serves as a control: Timepix3 hits data were integrated without clustering or event localisation. For dataset two, the Timepix3 data were event localised using CNN-ToT. The third Timepix3 dataset was based on event-localisation using CNN-ToT and deterministic blurring. Finally, the fourth dataset is based on event localisation using CNN-ToT, deterministic blurring, at two times super resolution. All four data sets were treated equally from then on. Data were processed using the RELION pipeline [68]. The movie stacks were corrected for drift (single patch) and dose weighted using MotionCor2 [69]. The Timepix3 sample pixel size was measured using a cross-grating grid at low magnification, and initially extrapolated by SerialEM to the recording magnification. Finer pixel size calibration was achieved by aligning the obtained map with an existing model (PDBID:7O6E).

Table 5. 1 Cryo-EM SPA data collection, refinement, and validation statistics of Timepix3 and Falcon3 datasets of BfrB protein.

	Timepix3 Integrated Hits	Timepix3 Localised	Timepix3 Localised + Blur[c]	Timepix3 Localised + SR[a]+ Blur[c]	Falcon3 Electron Counting
Magnification	215k	215k	215k	215k	155k
Voltage (kV)	200	200	200	200	200
Flux (e$^-$/Å^2/s)	29	29	29	29	1.76
Exposure time (s)	1.5	1.5	1.5	1.5	30
Frames (no.)	50	50	50	50	48
Defocus range (μm)	0.5 - 1.4	0.5 - 1.4	0.5 - 1.4	0.5 - 1.4	0.5 - 1.8
Pixel size (Å)	1.2	1.2	1.2	1.2/0.6[b]	0.635
Micrographs (no.)	2970	2970	2970	2970	979
Initial particles (no.)	14658	14535	14878	14911	162975
Final particles (no.)	10420	10525	10513	11422	124577
Map resolution (Å)	3.73	3.56	3.17	3.01	2.54
Sharpening B factor (Å^2)	-183	-132	-132	-119	-93

[a] Two times super resolution. [b] First value in physical pixel size, second value in pixel size at two times super resolution. [c] Deterministic blur.

Falcon3 single-particle data were collected from a different sample grid under the same microscope. Data collection was done using EPU. The detector was used in electron counting mode at a nominal magnification of 155 000 times. Table 5. 1 shows the statistics of this dataset. Data were processed using the RELION pipeline [68]. The movie stacks were corrected for drift (5 x 5 patches) and dose weighted using MotionCor2 [69].

From this point on, the Timepix3 and Falcon3 datasets were treated equally. The contrast transfer function (CTF) parameters were determined for drift-corrected micrographs using Gctf [70]. A first set of 2D references was generated from manually picked particles in RELION [68] and these were then used for subsequent automatic particle picking. Table 5. 1 lists the number of particles in the final dataset initially picked as well as the number of particles that contributed to the final reconstruction. An initial map was generated by RELION with octahedral symmetry (O) and used for 3D refinement. The beam tilt parameters, anisotropic magnification and local CTF parameters were refined and the particles were polished [43]. The final map was calculated and sharpened using the polished particles (Table 5. 1).

For model building, the PDB entry 7O6E [145] was used as a starting model in Coot [73]. The final model was refined against a sharpened cryo-EM map obtained by LocSpiral [74]. The model was

refined iteratively through rounds of manual adjustment in Coot [75], real-space refinement in Phenix [76], redone using PDB-REDO [177], and validated using MolProbity [77] (Supplementary Table 5. 1).

5.3   RESULTS

5.3.1. MTF, NNPS and DQE

Figure 5. 5 shows the MTF, NNPS, and DQE obtained at 200 kV using Timepix3 integrated hits data (Tpx3 hits, blue), event-localised Timepix3 (Tpx3 Loc, orange), event-localised Timepix3 with deterministic blur (Tpx3 Loc+Blur, green), event-localised Timepix3 at two times super-resolution (Tpx3 Loc+SR, red), event-localised Timepix3 with deterministic blur at two times super-resolution (Tpx3 Loc+SR+Blur, purple) and Falcon3 in electron counting mode (Falcon3 EC, brown). Integrating just hits results in an inferior performance of the Timepix3 detector, both in terms of MTF and NNPS, leading to the lowest DQE. Event-localised Timepix3 and event-localised Timepix3 super-resolution have the highest MTF. Their flat NNPS make their DQE appearance intuitive: the improved MTF for event-localisation when comparing original with super-resolution, directly results in a better DQE for event-localisation super-resolution compared to original resolution. The effect of deterministic blur is somewhat less intuitive, as it will dampen both the MTF and NNPS. Overall, the best DQE results for the Timepix3 data were obtained for event-localisation and deterministic blur at super-resolution, closely followed by event-localisation and super-resolution. Falcon3 in electron counting mode performed worst until half Nyquist, due to its lower DQE(0), however, hereafter it performed best.

5.3.2 Single particle analysis

Table 5. 1 lists the statistics and resolutions obtained for the single-particle analysis data. The resolution of the final map of the event localised and deterministic blurred at two-times super resolution Timepix3 data was 3.01 Å using the gold-standard FSC = 0.143 criterion (Figure 5. 6). Its sharpening B-factor is −119 Å². The Timepix3 dataset was recorded over a 72 h period and ≈ 11,000 usable particles were picked from 2,777 micrographs. The Falcon3 dataset was recorded over a 48 h period and from 979 micrographs 162,975 particles (≈ 14 times more than Timepix3 datasets) were picked. The Falcon3 electron counting dataset resulted in a resolution of 2.54 Å with a sharpening B-factor of −93 Å². The Rosenthal-Henderson B-factor plot is shown in Supplementary Figure 5. 1.
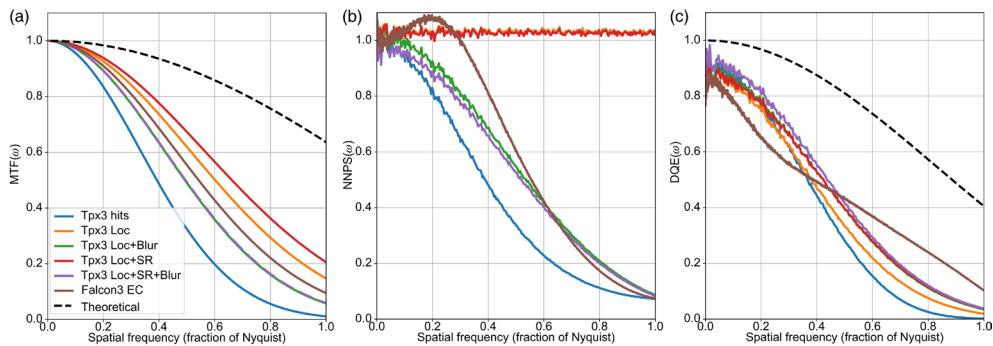
Figure 5. 5 MTF (a), NNPS (b) and DQE (c) as function of spatial resolution (as fraction of the Nyquist frequency) for Timepix3 integrated hits data (Tpx3 hits, blue), event-localised Timepix3 (Tpx3 Loc, orange), event-localised Timepix3 with deterministic blur (Tpx3 Loc+Blur, green), event localised Timepix3 at two times super-resolution (Tpx3 Loc+SR, red), event-localised Timepix3 with deterministic blur at two times super-resolution (Tpx3 Loc+SR+Blur, purple) and Falcon3 in electron counting mode (Falcon3 EC, brown). DQE(0) for Timepix3 is assumed to be 0.9. DQE(0) for Falcon3 electron counting is assumed to be 0.85. To aid in comparison, the spatial frequency of super resolution data has been rescaled to the recorded spatial frequency. The theoretical MTF curve is $\mathrm{sinc}^2(\frac{\pi\omega}{2})$ and the theoretical DQE curve is $\mathrm{sinc}^2(\frac{\pi\omega}{2})$.

5.4   DISCUSSION

Our results show a successful integration of Timepix3 as a pixelated detector in a cryo-EM workflow. We were able to integrate this detector both at the hardware and software level. We identified and overcame compromising effects from its original dedicated chiller setup. On the software level, we established a two-way software control of the detector using SerialEM and the remote procedure server Serval from ASI. We arrived at a tightly controlled synchronisation of the electron beam and the Timepix3 shutter event by using SPIDR and a customised trigger box. The successful integration was demonstrated for an SPA workflow, resulting in a 3.01 Å BfrB structure.
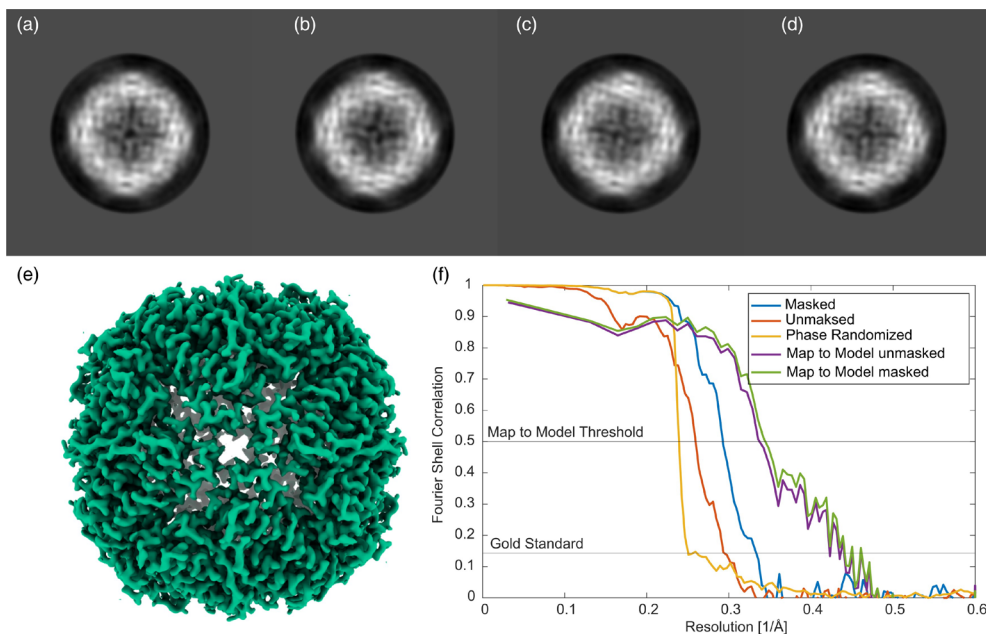
Figure 5. 6 Single-particle analysis of BfrB. (a - d) 2D class averages; the size of the shown box is 150 Å. (e) 3D reconstruction from 11 422 particles at 3.01 Å resolution collected at 200 kV using the Timepix3 with event localisation, deterministic blur and two times super resolution. (f) Gold-standard Fourier shell correlation (FSC) before (orange) and after (blue) masking, the phase randomised FSC (red), and the masked (purple) and unmasked (green) map to model FSC [72].

Integrating a detector into a TEM can be, as with many engineering endeavours, a project full of idiosyncrasies. Some of which were already mentioned, such as the possible interference from the dedicated chiller, but we also experienced others, such as faulty hardware of the detector PC leading to random crashes during workflow execution. The Thermo Fisher Scientific TEMs are rather closed platforms: there is very little open documentation available to make hardware and software integration, which led to a considerable amount of trial-and-error work.

TEMs can produce vast amounts of data, and making these findable, accessible, interoperable, and reusable is hot topic in the field [178,179]. We urge microscope vendors to look beyond the data that TEMs produce and to also practise some of these concepts for their hardware and software to allow for better access and interoperability of their platforms such that advances in the field could accelerate.

Our MTF, NNPS and DQE results (Figure 5. 5) confirm our previously published observations [147]: convolutional neural networks can be used to increase the accuracy at which we can predict the impact location of incident electrons. They also show that placing the event singularly in the predicted pixel is not the best way to use this information. Applying a deterministic blur will

dampen both MTF and NNPS and can result in an increase in DQE, particularly when the NNPS is aliased [160,164]. The deterministic blur effect has been modelled for dose-limited detectors used in medical X-ray devices [180-182]. In the field of cryo-EM, the Falcon3 and Falcon4 implement a deterministic blur; however, we are not aware of work describing its fine details [1]. The NNPS of our Falcon3 in electron counting mode (Figure 5. 5) has a remarkable shape, significantly deviating from the Gaussian blur functions that we have used. Figure 5. 4 shows the comparison of integrated Timepix3 hits data, event localised, and deterministic blurring. Visually, it is hard to see the difference between the integrated hits and the deterministic blurred image: both boost low spatial information, which makes both images visually pleasing. Nonetheless, high spatial-resolution information is much better preserved after event localisation and deterministic blurring (Table 5. 1).

Measuring and calculating the MTF, NNPS and DQE to compare cryo-EM detector performance has clear limitations. The measurement and calculation of the MTF is very sensitive to the way the edge spread function is determined. The estimation of NPS(0) is prone to errors, and the measurement of DQE(0) is practically challenging. This echoes the experiences of Zhu, et al. [183], who reported irreproducible results using the FindDQE programme [159]. Instead of measuring DQE, they examined the quality of their micrographs by using gold-standard Fourier shell coefficient (FSC) curves of their reconstructions. We publish both DQE and FSCs. We have not only openly published our raw data but also the software used to calculate Figure 5. 5, and spur other researchers and detector vendors to do the same whenever reporting an MTF, NNPS and DQE [175,184].

The single-particle analysis of BfrB showed the first cryo-EM workflow integration of a Timepix3 based detector in action. The best Timepix3 results, at 3.01 Å, were obtained using event localisation with deterministic blur at two times super resolution. The quality of the EM map is illustrated by the clear density of side chains, holes in aromatic rings, and the resolution estimate of 2.9 Å obtained by comparing the map with the refined model. For the most part, the SPA results follow the trend seen in the DQE measurements.

Surprisingly, the effect of super-resolution on the final SPA resolution is relatively minor. There is only a minor improvement in map resolution, and the DQE curves are almost the same. This may indicate that, for these data, the event localisation was not much more precise than 1 pixel. Our previously published results showed that, for simulated data, the point of impact of individual electrons can be found, on average for CNN trained on ToT data (CNN-ToT) at 200 kV, 0.50 pixels away from the incident pixel [147]. In this work, we have chosen to use CNN-ToT rather than the CNN trained on ToT and ToA data (CNN-ToT-ToA). The latter was only marginally better compared to CNN-ToT, but gave systematic patterns that needed extra (gain) correction. It could also be possible to adapt our event localisation method for the 80-120 kV energy range. The neural network currently benefits from the higher (>= 200 kV) energy electrons generating larger clusters. However, we believe that it should be possible to further improve the event localisation method, also for the 80-120 kV range, for example by correcting

for more systematic errors in the ToA data or by training directly on experimental data. The latter could be done with a dedicated setup capable of generating a sub-pixel electron beam at the desired energy.

Timepix3 is the first fully event-driven detector to be used in a cryo-EM workflow. For each electron event a separate timestamp is known, with a maximum accuracy of 1.56 ns. Recently, the Electron Event Representation (EER) format was introduced in conjunction with the Falcon4 MAPS detector [171]. Although the EER representation is an improvement to the scheme in which a set of frames is combined into a fraction of e.g. 1 e$^-$/Å^2, it is still frame-based. For each time block, events are grouped together. This makes the format impractical for Timepix3 data, where each event has its own timestamp. We could foresee the development of a new, truly event-based data representation and processing pipeline for event-driven detectors such as Timepix3.

The high maximum hit rate (120 Mhit/s) of the Timepix3 (quad) detector allowed us to record the full useful lifetime of a protein in the electron beam (40 e$^-$/Å^2) as a stream of individual electron events in just 1.5 seconds. To the best of our knowledge, this is the fastest electron counting camera used for SPA so far. Although the data stream has the promise to allow for very fine corrections that occur during the exposure, such as stage drift, lens drift, and beam-induced motions, best results so far were still obtained while allowing for some settling time between hole selection, autofocus and data acquisition. Data throughput was also still limited by processing speed, although we gained a 10-fold increase compared to van Schayck, et al. [147] by improving the clustering algorithm and applying more efficient parallel processing. Further improvements could come from, for example, GPU implementations and the use of software platforms such as LiberTEM [185].

HPDs can be used at much higher electron fluxes in imaging mode or even with direct beams in diffraction mode, e.g. electron diffraction and ptychography [186,187]. Such workflows require full integration of the detector in the TEM as well as automation. In this paper, we demonstrate how we could arrive at such integration and automation for SPA. Next, we will explore diffractive imaging techniques, such as ptychography, starting with near-field electron ptychography with a diffuser [188,189]. Speed is an important practical limitation in cryo-electron ptychography. Our Timepix3 detector should already be much faster than the latest commercially available EMPAD detectors (MAPS detectors cannot be used for ptychography). We are currently building the next generation Timepix4 detector [190], which allows for unprecedented hit-rates (5 Ghit/s) and time resolution (200 ps) and can be tiled on 4-sides without extra gaps. The work presented here provides a solid basis for the use of these innovative detectors for new diffraction and imaging techniques, which should allow one to obtain more and novel structural information on biological samples that have a limited lifetime within the electron beam.

5.5 CONCLUSION

A Timepix3 hybrid pixel detector has been integrated in an automated cryo-EM workflow. Its performance at 200 kV has been demonstrated for different event localisation schemes, both in terms of MTF, NNPS, DQE as well as gold-standard FSCs. High-quality single-particle analysis reconstructions could be obtained with this event-driven HPD, at 200 kV, in electron counting mode, with high hit rates and CNN-based event localisation. HPDs could greatly expand the possibilities of (cryo-)EM for structural biology, since they would allow for, at a wide range of energies, both imaging and diffraction-based experiments.

DATA AVAILABILITY

The refined model has been deposited in the Protein Data Bank as PDB entry 8AEY and the map has been deposited in the Electron Microscopy Data Bank (EMDB) as entry EMD15389. Raw Timepix3 data, processed MRC movie stacks and extracted particles as used for the EMDB map have been deposited in EMPIAR as entry EMPIAR-11113. Raw Timepix3 data and Falcon3 micrographs for calculating MTF, NNPS and DQE have been deposited in Zenodo [184].

SOFTWARE AVAILABILITY

The following software has been deposited at Zenodo: (1) tpx3HitParser to parse raw Timepix3 data in a sparse localised event stream [174], (2) tpx3EventViewer to transform the sparse localised event stream into frames [175], and (3) the scripts to measure MTF, NNPS and calculate DQE [184]. The SerialEM plugin to control Timepix3 through the Serval Remote Procedure Service is available from Amsterdam Scientific Instruments.

ACKNOWLEDGEMENTS

SUPPLEMENTARY MATERIALS

Supplementary Figure 5. 1 The Rosenthal-Henderson B-factor plot showing resolution versus number of particles. B-factors were estimated by fitting a straight line through the inverse of the resolution squared versus the natural logarithm of the number of particles for a range random subsets of full particle list. [72].

Supplementary Table 5. 1 Refinement statistics for the BfrB protein model.

Refinement	
Initial model used (PDB entry)	7O6E
Model resolution against LocSpiral map (Å)	2.9
FSC threshold	0.5
Model composition of monomer	
Atoms	1471
H atoms	0
Protein residues	177
Waters	10
B factors (Å^2)	
Protein	80.3
Water	73.5
R.M.S deviations	
Bond length (Å)	0.008
Bond angle (°)	1.108
Correlation coefficients	
Mask	0.89
Box	0.83
Validation	
MolProbity score	1.15
Clash score	3.6
Poor rotamers (%)	0
Ramachandran plot	
Favoured (%)	98.86
Allowed (%)	1.14
Disallowed (%)	0

SINGLE PARTICLE CRYO-EM: ALTERNATIVE SCHEMES TO IMPROVE THE DOSE EFFICIENCY

Yue Zhang, Penghan Lu, Enzo Rotunno, Filippo Troiani, Paul van Schayck, Amir H. Tavabi, Rafal E. Dunin-Borkowski, Vincenzo Grillo, Peter J. Peters, Raimond B.G. Ravelli
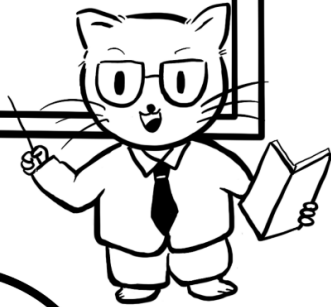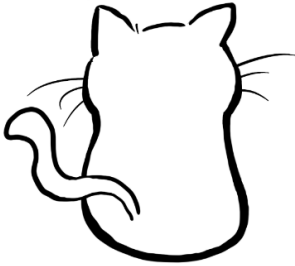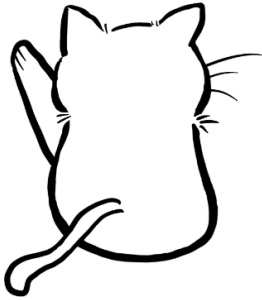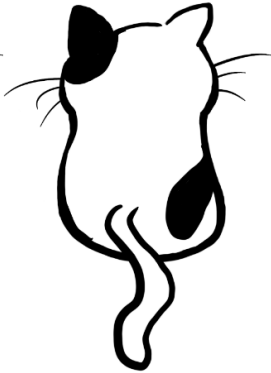
ABSTRACT

Imaging of biomolecules by ionising radiation, such as electrons, causes radiation damage which introduces structural and compositional changes of the specimen. The total number of high-energy electrons per surface area that can be used for imaging in cryo-electron microscopy (cryo-EM) is severely restricted due to radiation damage, resulting in low signal-to-noise (SNR) ratios. High resolution details are dampened by the transfer function of the microscope and detector, and are the first to get lost as radiation damage alters the individual molecules which are presumed to be identical during averaging. As a consequence, radiation damage puts a limit to the particle size and sample heterogeneity with which electron microscopy (EM) can deal. Since a transmission EM (TEM) image is formed from the scattering process of the electron by the specimen interaction potential, radiation damage is inevitable. However, we can aim to maximise the information transfer for a given dose and increase the SNR by finding alternatives to the conventional phase-contrast cryo-EM techniques. Here some alternative transmission electron microscopy techniques are reviewed, including phase plate, multi-pass transmission electron microscopy, off-axis holography, ptychography, and a quantum sorter. We discuss their prospects for providing more or complementary structural information within the limited lifetime of the sample.

6.1. INTRODUCTION

Cryo-electron microscopy (cryo-EM) has become a powerful tool for structural biologists to study the structure-function relationships of their biomolecules of interest. Recent development of transmission electron microscopy (TEM) hardware has led to atomic resolution for single particle cryo-EM [1,2]. For smaller (<100 kDa), more heterogeneous and non-symmetric samples, it is still a challenge to reach near-atomic resolution.

Although transmission electron microscopy has the power to image individual atoms, its performance is limited by the radiation sensitivity of aqueous biological specimens and therefore we need to minimise exposure to the electron beam. Radiation damage results in changes to the specimen due to knock-on damage and radiolysis [48,94,95,102,127,191]. In TEM, the radiation dose is traditionally defined as charge density (C/m^2) or electrons/unit area ($e^-/Å^2$) [102]. Here, we use the absorbed dose in gray (Gy), which is the radiation energy (J) absorbed per unit mass (kg) and relates to the beam parameters (electron energy, electron flux, and shape), sample thickness and composition [192,193]. Henderson proposed a dose limit of 20 MGy at which the diffracted intensity coming from a cryo-cooled crystal would be halved [18], whereas Garman obtained an experimental value of 43 MGy [194]. Glaeser reported a fluence of 6 $e^-/Å^2$ at 80 keV to cause significant damage to biological specimen [195], which can be calculated to correspond to a dose of 44 MGy using Raddose-3D [196]. Micro-electron diffraction (micro-ED) is able to solve the macromolecules to atomic resolution with such a dose [103,197], however, many important biomolecules cannot be crystallised. Ultimately, one wishes to image each particle individually, as each particle itself will define one point within the multidimensional conformation space a macromolecule can adapt. This is done in Single Particle Analysis (SPA), with a fluence of, typically, 40 $e^-/Å^2$ at 300 keV, which would correspond to 149 MGy. Dose-weighting schemes are nowadays routinely applied to account for the loss of high-resolution information at such high doses [52,198].

Because of the large amount of energy being deposited in the sample for a relatively limited fluence, cryo-EM imaging of biological samples results in low signal-to-noise (SNR) ratio images. A large number of images of identical molecules are required to boost the SNR. The SNR of a single image, X_{sig}, is given by [48,59]:

$$X_{sig} = \frac{Signal}{Noise} \times \sqrt{N_{pix}} \propto \sqrt{N_e} \qquad\qquad \text{Eq. 6. 1}$$

Here, N_{pix} is the number of pixels that corresponds to the area of one box containing the particle and N_e is the incident electron fluence. The SNR relates to the detectability of the whole protein, and is proportional to $\sqrt{N_e}$. To increase the SNR, one could increase the signal, decrease the noise, or use a combination of both. The conventional way to increase the signal would be to increase the number of electrons N_e. However, this will not really improve the SNR as the biological specimen will disintegrate upon exposure to the electron beam. For single particle

analysis, a large number of identical protein images is averaged to boost the SNR. The SNR is, therefore, also proportional to $\sqrt{N_i}$, where N_i is the number of individual protein images.

Frozen-hydrated biological specimens consist predominantly of light elements that weakly modify the phase of the electron wave: typical SPA samples of biomolecules in solution are normally thin enough (~20-100nm) to be treated as weak phase objects. The phase shift of the exit-wave is induced by the mean inner potential (MIP) of the sample which provides information about the electrostatic potential of the specimen. For weak phase objects, there is no amplitude contribution and the information transfer function becomes the contrast transfer function (CTF) multiplied by an aperture function and an envelope function [102,199],

$$\text{CTF} = \sin\left(\frac{2\pi}{\lambda}\left(\frac{1}{4}C_s\lambda^4 q^4 - \frac{1}{2}\Delta f \lambda^2 q^2\right)\right)$$ Eq. 6. 2

where C_s is the spherical aberration, λ is the wavelength, Δf is the defocus and $\Delta f > 0$ means underfocus, and q is the spatial frequency. At low defocus, there is very little transfer at low resolution, resulting in very poor contrast. The contrast can be improved by increasing the defocus, which will induce fast oscillations in the CTF. The information transfer is damped by the envelopes function, $E(q)$, which is the product of the envelope functions due to chromatic coherence ($E_c(q)$) and the spatial coherence ($E_s(q)$) of the electron source [102]:

$$E_c(q) = \exp(-\frac{1}{2}(\pi\lambda\delta)^2 q^4)$$ Eq. 6. 3

$$E_s(q) = \exp\left(-\left(\frac{\pi\alpha_i}{\lambda}\right)^2 (C_s\lambda^3 q^3 - \Delta f \lambda q)^2\right)$$ Eq. 6. 4

where δ is the defocus spread due to chromatic aberration. Therefore, increase of the defocus leads via $E_s(q)$ to more damping at high spatial frequency if the illumination semi-angle α_i is not small enough. Recently, Glaeser, et al. [200] pointed out that α can be as low as $1 - 2$ μrad, and that therefore the damping of the CTF is not so much due to the spatial coherence of the electron source, so that defoci values as high as 4 μm could be used. However, a high defocus results in information that becomes delocalised outside the field of view of the detector [200]. To this extent a small defocus value is still useful with the limited camera field that becomes further limited when very small pixel sizes are used for sub-2 Å structure reconstruction. Moreover, to reach atomic resolution reconstruction for biomolecules, aberration correction is essential. The axial coma aberration is the most limiting aberration in the atomic resolution regime, which can be corrected by software [1,44] or by hardware such as by using a spherical aberration corrector [2]. Henderson (1995) represented the contrast of the image with a factor C, which varies from 0 to 1 [48,59]. This multiplicative factor C degrades the value of X_{sig} in Eq. 6. 1, making it difficult to determine structures of smaller proteins (< 100 kDa) to high resolution.

Here, we review alternative cryo-EM techniques for obtaining structural information of biomolecules in solution, mainly for weak phase objects, *e.g.* samples for SPA. SPA is carried out using conventional phase contrast TEM, whereas cryo-STEM has been used for tomography of thicker biological samples [201,202]. Here, we discuss alternative techniques, exotic within the field of structural biology, which have different information transfer functions. We review TEM techniques, such as phase plate, multi-pass TEM, and off-axis holography, as well as STEM techniques, such as ptychography, and a quantum sorter. All these techniques will be discussed qualitatively in the next section. They provide alternative information transfer schemes, which, ultimately, could serve to improve the dose-efficiency in cryo-EM.

6.2. ALTERNATIVE CRYO-EM TECHNIQUES TO OBTAIN STRUCTURAL INFORMATION OF BIOMOLECULES

We review five alternative techniques in this section, namely a laser phase plate (6.2.1), multi-pass TEM (6.2.2), off-axis holography (6.2.3), ptychography (6.2.4), and a quantum sorter (6.2.5). Section 6.2.6 discusses the features, current status, and perspectives of these techniques and compare these to conventional TEM.

6.2.1 Phase plate

One solution to improve information transfer over a wider range of spatial frequencies is to use a phase plate in the back focal plane of the objective lens. A phase plate can introduce a quarter-wave phase shift between transmitted and scattered electrons [87,203-205]. Originating from light optics, Zernike phase plates were proposed and used to increase the phase contrast for weak phase objects by adding a $\pi/2$ phase shift to the diffracted beam [206], altering the CTF into:

$$\text{CTF} = \sin\left(\frac{2\pi}{\lambda}\left(\frac{1}{4}C_s\lambda^4 q^4 - \frac{1}{2}\Delta f \lambda^2 q^2\right) + \frac{\pi}{2}\right) = \cos\left(\frac{2\pi}{\lambda}\left(\frac{1}{4}C_s\lambda^4 q^4 - \frac{1}{2}\Delta f \lambda^2 q^2\right)\right)$$

Eq. 6. 5

As the contrast of images is dominated by the low-frequency components, the forms of the CTF is changed from sine to cosine which is crucial for the information transfer, in particular at low resolution. The cosine term allows one to record images close to focus, which reduces the damping of the envelope function (Eq. 6. 4) as is shown in Figure 6. 1. The practical implementation of a phase plate for TEM, however, has been cumbersome. Most thin film phase plates and even electrostatic phase plates [87,88,90,207-212] undergo electrostatic charging which results in poor reliability and a short working lifetime of the device, and distortion of the final image [213]. The Volta Phase Plate [87], which is commercially available, suffers from the fact that the phase

shift is variable. Furthermore, it introduces inelastic scattering after the sample, resulting in precious electron loss. A phase plate which gives stable $\pi/2$ phase shift, and induces almost no electron loss and post-sample scattering, should be ideal for biological specimens.
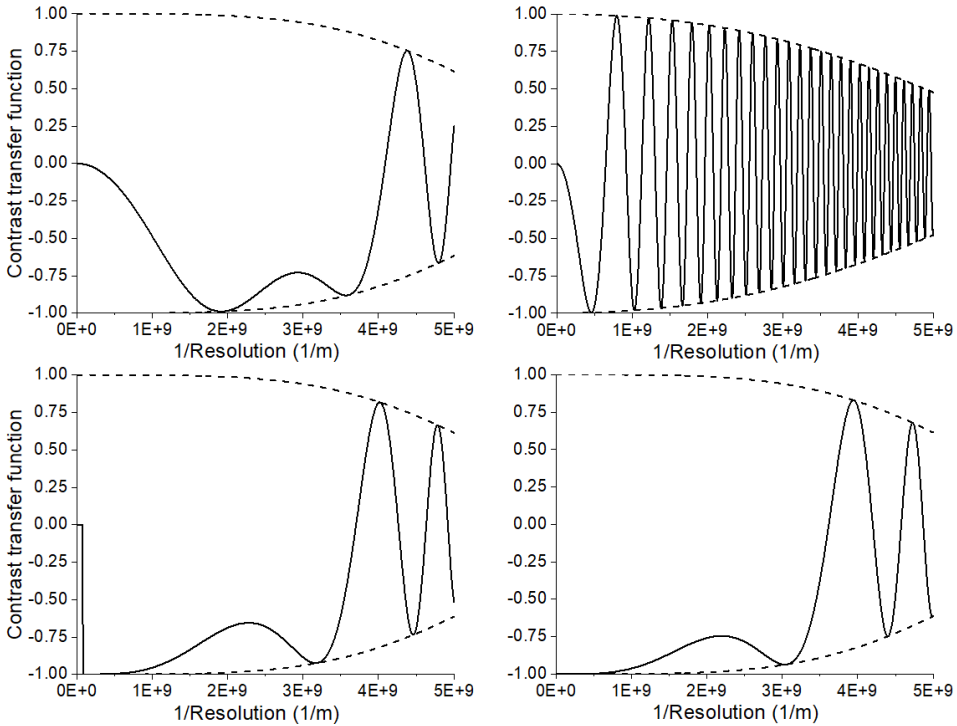
Figure 6. 1 Contrast transfer function of (a) conventional TEM at Scherzer defocus: $\Delta f_{Sch} = 87$ nm, (b) conventional TEM at defocus: $\Delta f = 1.2$ um, (c) with phase plate at defocus of 0.6 x $\Delta f_{Sch} = 52$ nmwith cut-on frequency at $1/25$ nm^{-1}, (d) holography at Gábor defocus: $\Delta f_{Gábor} = 49$ nm. All have the illumination semi-angle set at 15 μrad.

Recently, an alternative approach was devised, based on electron-light interaction: the laser phase plate [214]. This laser-based phase plate still requires the "centring" operation, i.e. shifting the unscattered beam to one of the antinodes of a laser standing wave. However, unlike for most other phase plates, such an alignment procedure should not cause detrimental charging, simply because there is no material along the beam path. This should make long-term operation more reliable. For the same reason, this device is free from inelastic scattering [215] and information cut-off [216], hence allowing for ultimately efficient transfer for almost all spatial frequencies with only reduced contrast along the laser standing wave direction [88].

The incoming charge particles will be affected by a ponderomotive force: a force the charged particles experience in a rapidly oscillating electromagnetic field. The ponderomotive force will induce a phase shift on the electron, as given by [217]:

$$\phi = \sqrt{8\pi} \frac{\alpha}{\beta\gamma} \frac{P}{mc\omega^2 w} \qquad\qquad \text{Eq. 6. 6}$$

where a is the fine structure constant, c is the speed of light, m is the electron mass, β and γ are the electron's relativistic factors, P is the beam power, ω is the laser angular frequency, and w is the size of beam waist. For 300 keV electrons, a $\pi/2$ phase shift requires a laser with an intensity of hundreds of GW/cm^2 [217]. Meanwhile, in order to fit it into the column space of a TEM, the optical cavity needs to be small enough. Several different cavity designs were considered and a near-concentric Fabry-Pérot resonator was chosen to be applied within the TEM column [89,217].

Lots of improvements have been achieved in obtaining a stable and high power laser [88,208]. First, a continuous-wave laser system has been designed and applied. Secondly, the laser intensity was increased from 43 GW/cm^2 to above 120 GW/cm^2 which gives a $\pi/2$ phase shift to 300 keV electrons with longer operating hours. Also, the laser mode waist has been decreased from 13 µm to 8.5 µm. A small mode waist is essential to obtain a smaller cut-on frequency [218], thereby maximising the phase contrast at low spatial frequency. In addition, a transfer lens and a set of deflectors have been used to align the beam to the projection system. With this laser phase plate, a 3.8 Å 20S ribosome reconstruction could be obtained with just 4789 particles at a fluence of 50 e$^-$/Å^2 at 300 keV (185 MGy) [208]. Figure 6. 2 shows a schematic of the laser phase plate within a TEM column.

The mechanical and thermal stability of the laser system was one of the primary concerns, but recent SPA results obtained using the laser phase plate [208] mitigated this doubt. The phase shift produced by this phase plate can in principle be adjusted by varying the laser intensity, which could be advantageous in situations such as the imaging of strong phase objects such as gold fiducial markers used for cryo-electron tomography [218]. However, changing the phase shift will most probably pose further technical challenges to the stability of the laser optics. An additional magnified diffraction plane must be created to match the laser spot size and to accommodate the space for the laser module. The addition of extra transfer lenses substantially increases the chromatic aberration, thereby limiting the spatial resolution that can ultimately be achieved [208]. This issue could possibly be resolved by redesigning the electron optics in future microscopes where the laser phase plate is implemented. All in all, this matter-free phase plate concept is both elegant and technically challenging, and involves serious add-ons to the microscope. We hope to hear much more about it in the years to come.
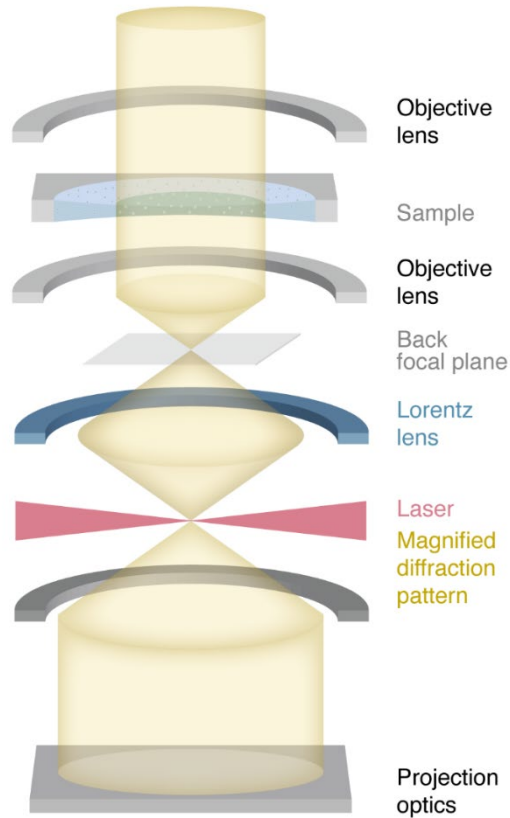
Figure 6. 2 Schematic of a laser phase plate in a TEM. The high-power standing laser wave introduces a phase shift to the unscattered beam. A Lorentz lens after the objective lens is used for magnifying the back focal plane to reduce the constraint of the laser mode waist.

6.2.2 Multi-pass transmission electron microscopy

Biological specimens embedded in a thin layer of vitreous ice are weak phase objects, which means that the phase shift of an electron passing through the sample is small. In order to increase contrast, one could increase the phase shift of the electrons. Let ϕ be the phase shift of the incoming electron beam after it passes through the specimen. If one could pass the same electron m times through the specimen, the accumulative phase shift would become m. This is exactly what is done in multi-pass transmission electron microscopy (MPTEM).
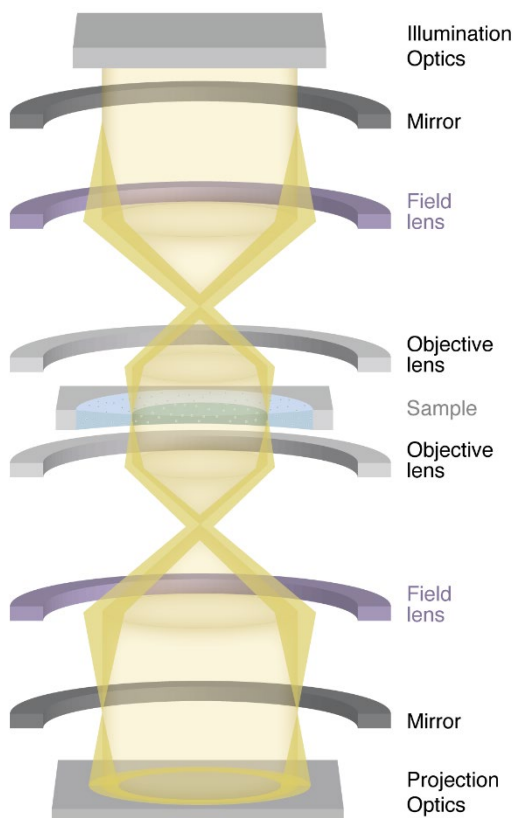
Figure 6. 3 Schematic illustration of the MPTEM setup. The electron beam is generated by the illumination optics, passes through a gated mirror, goes through the sample, and bounces back and forth between the two gated mirrors (shown in purple). After having passed through the sample *m* times, the beam is gated through the second mirror and enters the projection optics and is recorded by the detector.

MPTEM is being developed after the concept was first demonstrated in light microscopy [219]. The sample is placed in a self-imaging optical cavity. A pulse of light enters the self-imaging cavity through an in-coupling mirror and is scattered by the sample. Next, the formed image is reflected on a mirror and re-imaged onto the sample. After passing through the sample *m* times, the light leaves the cavity and is imaged using a microscope objective onto a camera. Analogous to light optic design, multi-pass microscopy was simulated [220] and designed for TEM [221]. Figure 6. 3 shows a schematic view of a MPTEM [221].

In MPTEM, the signal is *m*-fold enhanced whereas the noise is unchanged. Thus, the signal-to-noise ratio X_{sig} **in** equation (1) is *m*-fold enhanced as well, which is directly the main advantage of MPTEM [220]. For a given SNR, MPTEM will cause less radiation damage. For example, conventionally, we will have a SNR proportional to $\sqrt{N}$ if we have N electrons as measurement

probes. In MPTEM, we assume that there are m^2-fold less (N/m^2) incoming electrons passing through the specimen m times. As a result, the signal is proportional to $m \times N/m^2$, which is actually lower than for the conventional way. Fortunately, we care more about the SNR rather than the signal alone. As the noise level is proportional to the square root of the incoming electrons, it is $\sqrt{\frac{N}{m^2}}$. Thus, we have the same SNR as for conventional imaging:

$$SNR = \frac{m \times \frac{N}{m^2}}{\sqrt{\frac{N}{m^2}}} = \sqrt{N}. \qquad\qquad \text{Eq. 6. 7}$$

However the damage to the sample is proportional to the product of electrons and the pass number $m \times N/m^2 = N/m$, which is m-fold less than for a conventional TEM, while the SNR is constant.

The value of m is restricted by several factors. Firstly, once the phase shifts of the elastically scattered electrons have accumulate up to π, additional interaction will lead to a decrease of the phase contrast, thus lowering the signal. Simulations indicated that the SNR decreases at high interaction number, m [220]. Secondly, inelastic scattering restricts m as well. One needs to ensure that the total path length of the electron (m times sample thickness) in the specimen is smaller than the inelastic mean free path (IMFP) of the electron at the specific operating energy. A high interaction number induces high electron losses due to absorption and inelastic scattering, which reduces the efficiency of multi-pass electron microscopy. Simulation illustrates that when $m\sim10$, the MPTEM should have the optimal performance. In this case, the SNR will be three times higher at the same dose limit compared to conventional cryo-EM. This would enhance the resolution of a given sample [221], or allow for SPA to be performed on smaller biomolecules, or enable the processing of highly heterogeneous samples [48,59,222,223].

However, various experimental challenges can be foreseen. Although the optical cavity is available for photons [224,225], a high voltage electron mirror, which is a key component for MPTEM, does not yet exist. Meanwhile, the electron mirrors need to be well synchronised with the electron gun as well as with themselves in order to in-couple and out-couple the beam properly. A laser-triggered electron source fits the experiment best, since a continuous source would result in chromatic aberration [221]. Moreover, the beam needs to interact with the same point of the specimen multiple times to have the correct phase accumulation. This requires very precise optics in the cavity. From all the techniques reviewed in this paper, MPTEM seems to be furthest away from practical applications.

6.2.3 Holography

One of the main issues with conventional TEM is that only the amplitude of the exit wave is recorded, whereas the phase information is lost. We could reconstruct the whole exit

wavefunction if we knew both the amplitude and phase. The phase shift of the exit-wave provides information about the local variations of electrostatic potential in the sample. Full knowledge about the exit wavefunction could also benefit the study of biological objects, as it offers the possibility of correcting the residual aberration digitally by implementing virtually phase plates in software rather than in hardware [226]. One way of obtaining the phase information is by interfering the exit-wave with a reference plane wave. Dennis Gabor was awarded the Nobel Prize in Physics for his invention and development of the in-line holographic method [227], in which the interfering reference and object waves shared the same optical axis. Möllenstedt and Düker proposed off-axis holography [228], in which the exit-wave and reference wave are deflected and overlapped by an electron biprism (Figure 6. 4a). The electron biprism can consist of a positively charged wire. Here, we mainly focus on the off-axis holography technique.

In holography, by superimposing the exit-wave with the reference plane wave, both the phase and amplitude of the exit-wave can be resolved. The intensity of the interference pattern is given by [229-231]:

$$I(x) = |\Psi_O + \Psi_R|^2 = 1 + A^2(x) + 2A\cos(2\pi q_c x + \Phi_o(x)) \qquad \text{Eq. 6. 8}$$

where Ψ_O is the exit-wave, Ψ_R is the reference plane wave, $A(x)$ and $\Phi_o(x)$ are the amplitude and phase of the exit-wave, respectively, and q_c is the spatial frequency of the interference fringes. From **Eq. 6. 8**, we know that the final intensity consists of a reference wave amplitude, an exit-wave amplitude and an exit-wave phase. To reconstruct the complex exit-wave function, one takes the Fourier transform of the hologram:

$$FT[I(x)] = \delta(q) + FT(A^2) + \delta(q - q_c) \otimes FT[A\,exp(i\Phi_o)] + \delta(q + q_c) \otimes FT[A\,exp(-i\Phi_o)] \qquad \text{Eq. 6. 9}$$

This function contains three bands; (i) the central band of the intensity distribution without phase information: $\delta(q) + FT(A^2)$, (ii) the +1 sideband at $q = q_c$ and (iii) the -1 sideband at $q = -q_c$ which both contain the complete exit-wave. Therefore, after the Fourier transform of the hologram, amplitude and phase distribution can be obtained from the inverse Fourier transform of either one of the two sidebands. The spatial resolution of the phase image is normally three-fold better than the fringe spacing [232-234]. It is essential to have fine interference fringes for high spatial resolution. However, this decreases the spatial coherence of the beam, and as a result, the fringe contrast decreases.
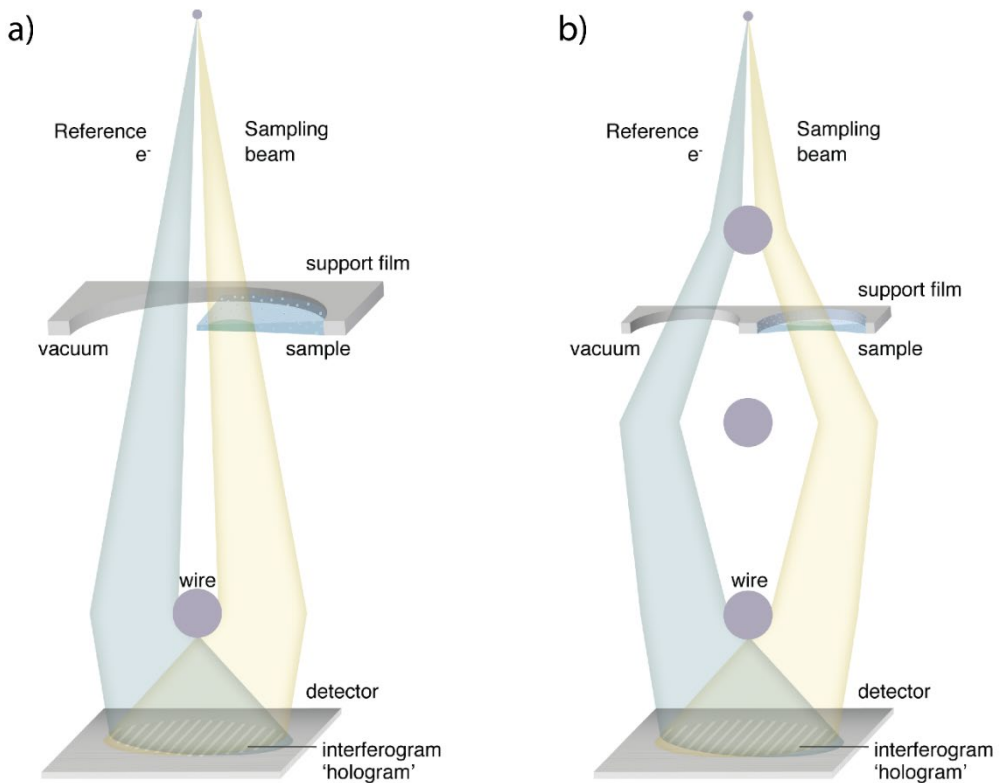
Figure 6. 4 Schematic illustration of the off-axis electron holography setup. a) Conventional off-axis electron holography. Half of the beam passes through the specimen (sampling beam) and the other half goes through the vacuum (reference beam). The two beams are next to each other, and the illuminated area can only be at the edge of the sample. Two beams are superimposed by the electron biprism, a positively charged wire, and interfere at the detector plane. b) Split-illumination holography. The beam is split by condenser biprisms. The reference and sampling beam can go through two neighbouring holes of the sample grid.

Ru, et al. [235] proposed phase-shifting electron holography to overcome this problem. This approach takes a series of holograms with interference fringes shifted every image by a beam-tilt, allowing for the retrieval of both the amplitude and phase information of the objective wave from the periodically changing fringe intensity [236]. As it is unnecessary to filter the signal in Fourier space, since phase-shifting electron holography is able to obtain high-resolution phase images without the loss of fringe contrast [234].

For biological applications, off-axis holography is able to enhance contrast from low spatial frequency to high spatial frequency as the phase contrast transfer function is in the form of a cosine [230]. This is in contrast to conventional TEM, where the contrast is enhanced by applying a high defocus at the cost of fast damping of CTF and loss of high spatial frequency signal. Off-

axis holography has been applied to investigate ferritin [237-239], flagella of bacteria [240], tobacco mosaic virus (TMV) [238,240], T5 bacteriophage virus, the hexagonal packed intermediate layer of bacteria, Semliki Forest virus, as well as collagen fibers and the surface layer of bacteria [230,241,242]. Holography has also been used to study the electrical or magnetic field within a biological sample [226,243].

There are several concerns regarding imaging biological specimens using holography. Firstly, the contrast of the images and the interference fringes can be poor, which has been attributed to the MIP and coherence of the electron beam, respectively. For cryo-EM, the biological specimen is located within a thin layer of vitreous ice and only induces a weak phase shift to the electron beam. To enhance the contrast, one could increase the sample thickness or use a higher dose, but these would induce more inelastic scattering and beam damage, respectively. The inelastically scattered electrons can reduce the contrast of the interference fringes as they are not coherent with the reference beam. High beam coherence requires a field emission gun (FEG) electron source, a small spot size, a small condenser aperture and a low gun extraction voltage [231].

Secondly, to obtain a reference beam, either vacuum or vitreous water without single particles is needed within the direct vicinity of the specimen for conventional electron holography. The required short distance between the sample and reference beams restricts the region of the sample that can be studied. In Figure 6. 4a, we drew a hole of a typical Quantifoil support which is often used for SPA. These holes of, e.g. 1.2 μm width, are normally filled with biomolecules embedded within a thin layer of vitreous ice. For conventional holography to work, half of this hole should be empty to allow passage of the reference beam. The split-illumination electron holography (Figure 6. 4b) method involves inserting extra biprisms in the condenser system to allow for a much larger distance between sample and reference beam [244,245]. In combination with the new sample preparation methods, such as the pin-printing of sample on specific parts of the grid [56], split-illumination holography might become possible for a wider range of biological applications.

Due to the low contrast and resolution of the biological specimen obtained with holography images, only very few such experiments have been reported. For ice-embedded sample, there are several hurdles: the requirement of a nearby hole, the small differences between the MIP of a protein and its surrounding ice, and destruction of the contrast due to charging of the ice. Recently, low-energy (20 keV) in-line holography experiments have been reported on ice-embedded samples of TMV virions, T4 bacteriophages, and erythrocruorin [246]. They showed that the sample is clearly visible in thin ice (~ 40 nm) and blurred in relatively thick ice (~ 90 nm), and discuss the importance of reduced exposure times to prevent the washing out of interference fringes by mechanical or environmental factors. In addition, compared with the charge-coupled device (CCD) camera used in the previous experiments, direct electron directors have much better detective quantum efficiency (DQE) and modulation transfer function (MTF). This greatly improves the contrast of the interference fringes and reduces the phase error of the reconstructed electron wave function, and thus improves the dose efficiency [247]. Nowadays, with

new sample preparation methods [56], more stable microscopes, direct electron detectors, motion correction [24,52], and additional approaches such as phase-shifting electron holography and split-illumination electron holography, off-axis electron holography might see a revival of interest that could complement SPA cryo-EM.

6.2.4 Ptychography

Ptychography is a method where an image is computationally reconstructed from a series of diffraction patterns. It is based on the measurement of a matrix of correlated diffraction images as a function of probe position. It allows the reconstruction of both the phase and the amplitude of an object by reverse calculating it from the reciprocal space data. The basic principle of this technique was proposed more than 50 years ago [248] and it has been widely used in X-ray [249] and extreme ultraviolet [250-252] studies in the past decades [253]. Only recently, catalysed by the introduction of fast direct electron counting pixelated detectors [147,254], has cryo-electron ptychography started to receive more attention.

In ptychography, a small scanning convergent electron beam is transmitted through the sample (Figure 6. 5). Ptychography has been stated to be the most dose-efficient approach compared to a few other scanning transmission electron microscopy (STEM) imaging modes [255], including differential phase contrast (DPC). Since ptychography is able to directly reconstruct the phase of the specimen, it ought to be beneficial for the imaging of weak phase objects such as biological molecules. One important potential benefit of ptychography is that, depending on the spatial frequency range of interest, the optimal CTF can be chosen by varying the probe convergence semi-angle [255,256]. This allows the retrieval of information at both low spatial frequency, such as morphology, as well as at high frequency, such as macromolecular structures. One can design the experiment in such a way that the single particle data collection is split into a few groups, each of which is taken with significantly different probe convergence semi-angle to enhance the contrast for particular spatial frequencies. Since ptychography is an *in silico* imaging process, the reconstruction algorithm also plays a key role in the information transfer and, in fact, each algorithm possesses its own intrinsic contrast transfer function. For example, it was reported [256] that the extended ptychographic iterative engine (ePIE) [257], the most commonly used iterative ptychography reconstruction algorithm, delivers a broader contrast transfer range compared to the single-side-band (SSB) method [258], a typical non-iterative algorithm for electron ptychography. Therefore, careful selection and optimisation of the reconstruction algorithm seems vital to obtain high contrast ptychographic phase images.
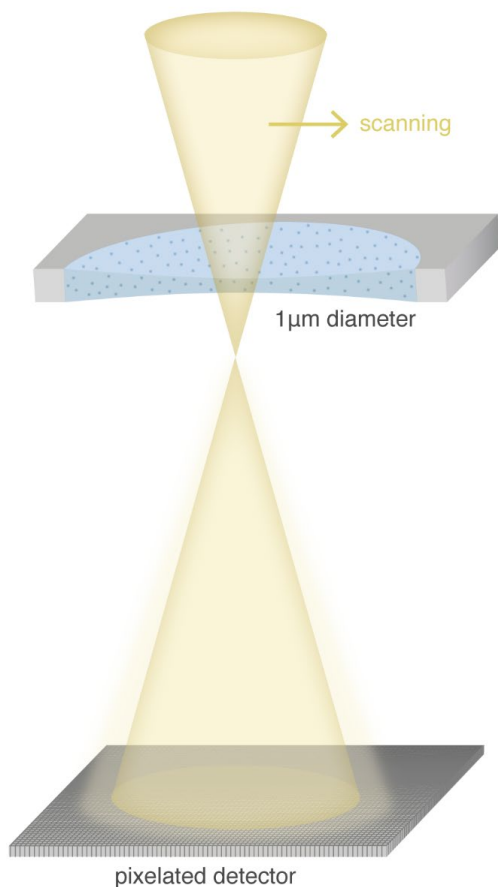
Figure 6. 5 Schematic illustration of cryo-ptychography. A small convergent electron beam scans through the specimen at a certain defocus. A high speed pixelated detector is used for recording the diffractograms that are subsequently collected from partially overlapping sample positions.

Like with any cryo-EM technique, the absolute total information transfer in ptychography will be limited by the total electron dose the sample can withstand. Recently, it has been experimentally demonstrated that electron ptychography can resolve sub-2 Å resolution on 2D materials using less than 400 e-/Å^2 (~1.5 GGy) [259]. Also last year, cryo-ptychographic reconstructed phase images were shown for virus particles that were inspected using a typical cryo-EM dose of ~20 e-/Å^2 (~75 MGy) down to nanometre resolution [153].

Simulations have shown that the use of a structured probe with random speckles can introduce strong fluctuations to the diffraction pattern which are particular helpful for ultra-low-fluence data as is the case in cryo-EM [260]. It was predicted that ptychography would need two orders of magnitude fewer molecules imaged, compared to conventional defocus cryo-EM, to obtain the

same target resolution within the same dose-lifetime of the specimen [260]. Similar probe shaping proposals have been carried out experimentally in STEM ptychography or its variant setups [189,261,262] using thin film-based phase masks. Alternatively, one can use electrostatic electron wave shaping devices [212,263] to avoid profuse scattering from the thin film masks.

There are a number of concerns regarding applying electron ptychography to biological imaging. The applicability of ptychography has been severely limited by sample thickness, and most of the previous experiments were performed on 2D materials. For samples that are too thick to be treated as a projected potential neglecting beam propagation (which is usually only a few nm), multi-slice ptychography approach has been proposed [257] and recently demonstrated for electrons [264]. This uses multiple slices to represent the sample and retrieve the structure of each slice separately. In practice, a monolayer of biomolecules within a well vitrified ice layer that is only slightly thicker than the macromolecules, could work fine for normal ptychography.

Another concern is that the need for sufficient overlap between neighbouring probe illumination areas will lead to different degrees of beam damage across the sample. The biological specimen inside the probe is going to be slightly damaged after the first diffraction pattern, and part of the illuminated area will be irradiated again for the second diffraction pattern. Even when a neighbouring region is not directly illuminated by the probe itself, it might still suffer from beam damage due to electrostatic charging as well as damage spreading [265,266] which will vary dynamically during the scanning of the beam. The gradual degradation and inconsistent quality of the sample may affect the result of the ptychographic reconstruction. Specific dose-weighting schemes, analogous to those developed for SPA, will have to be implemented for ptychography to correctly account for beam damage.

The signification reduction of the number of images to be averaged could improve throughput and resolution and could reduce the minimum particle size that could be studied. However, the current state-of-the-art pixelated detectors are still not fast enough for efficient cryo-ptychography experiments as they it would take milliseconds to record one diffraction pattern. In order to increase the throughput further, a better detector with more pixels (fine sampling in reciprocal space would allow for coarser sampling in real space) and faster speed is keenly desired. An ultra-fast detector with 512*512 pixels and 100 kHz rate is currently under development [267]. Ptychography dataset are four-dimensional and the reconstruction is usually computationally very expensive. The field of Single Particle cryo-EM faced similar challenges less than a decade ago, and managed to deal with them in a spectacular way. The coming decade should reveal the true potential of cryo-ptychography for structural studies of biomolecules in solution.

6.2.5   Quantum Sorter

Cryo-electron microscopy images of thin biological samples contain high-resolution information about the structure of each imaged molecule. This, however, does not directly reveal the identity of each molecule. In SPA, the imaging is performed on purified macromolecular complexes:

biochemical characterisation should provide a clear prior on what one can expect to see by cryo-EM. Often, impurities can be observed within the micrographs as well. Sometimes, these can be computationally sorted and identified *posteriori* by image processing [223]. In cryo-electron tomography, one looks at unique samples, *e.g.* of lamella of infected dendritic cells [268], in which molecules are identified visually [269] or computationally, and used for sub-tomogram-averaging (STA). Both schemes rely on molecule identification *posterio*, after the tomographic data have been collected and analysed. *A priori* localisation of individual molecules might become possible one day by means of correlative light-electron microscopy (CLEM) [270,271], once the resolution and sensitivity of cryo-light microscopy becomes high enough as well as fully compatible with cryo-EM. Below, we will describe how electrons can be used to identify individual molecules *a priori*, by means of a quantum sorter.

The use of a quantum sorter is a novel idea that allows for dose efficient protein discrimination in cryo-EM by the use of custom electron bases. Once electrons impinge on a given protein, the wavefunction of each electron contains all the information on the protein projection. Unfortunately, when their wavefunction collapses on the final detector, only a small part of that information is captured. Therefore, a large number of electrons must be accumulated in order to create an interpretable image of the electron wavefunction. The common concept of image is related to the spatial representation of the electron wavefunction. Mathematically speaking, every representation, namely every set of commuting observables, can be used to analyse a wavefunction, but some representations may be more convenient, depending on the problem one wants to address. While space representation is the one most commonly used, other representations could be more appropriate. One typical example is the angular basis, which includes the radial degree of freedom and orbital angular momentum (OAM) of the electron beam [272-275], and allows one to improve the efficiency of the measurement by increasing the amount of information that can be accessed in the measurement.
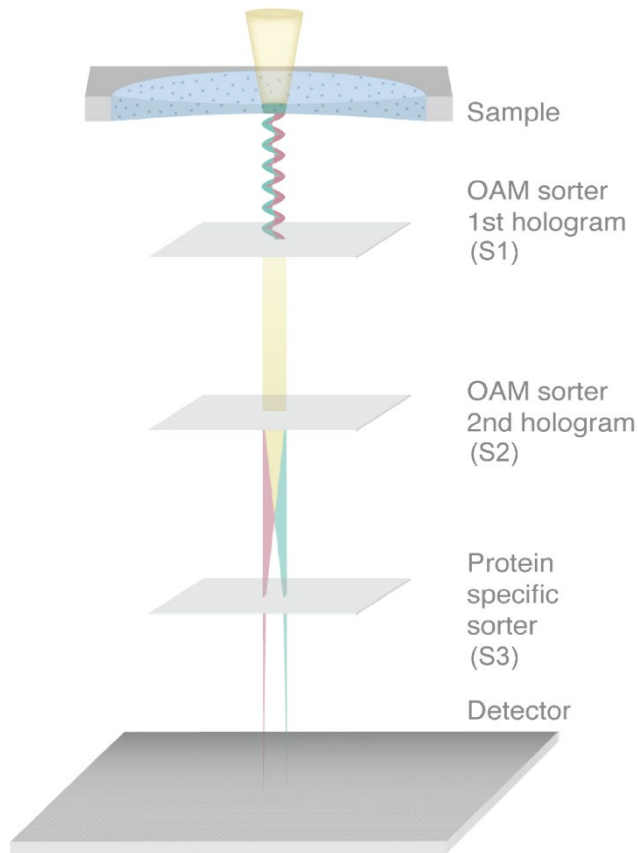
Figure 6. 6 Scheme of the new electro-optical setup allowing optimal discrimination between proteins. An OAM sorter made by three elements performs the Cartesian-polar mapping. Electron vortex beams which contain information hidden in their OAMs are generated after passing through the sample in the specimen plane. The first two holograms (S1, S2) perform the OAM sorting, and the log-polar spectrum is further passed through another hologram (S3) and then through a cylindrical lens (not shown in the figure).

Figure 6. 6 represents the scheme of the electron microscope in the special configuration able to change the representation from a Cartesian to an angular basis. The device has already been demonstrated experimentally [275]. Two coupled phase holograms or two equivalent electrostatic phase elements are placed in the microscope. The sorter 1 (S1) transforms the OAM in exit-wave from azimuthal coordinate to linear transverse coordinate, where the sorter 2 (S2) located in the diffraction plane of S1 corrects the phase distortion caused by S1 [276]. The only additional element described here is the projection element (S3) that permits to implement any measurement scheme. Figure 6. 7 shows the new representation of the protein. Most of the information is located on the OAM axis (ℓ), thus producing a compact representation of the

protein. For example, the presence or absence of rotational symmetry could be directly recognised on this axis. Given prior knowledge about the (low-resolution) structure of the model, the quantum sorter can be used to identify the presence of that molecule when a STEM beam is scanned over it. The capability to discriminate two different protein models using the quantum sorter was recently analysed in [277] within the general framework of quantum state discrimination. For two given proteins, and for the two corresponding wavefunctions of the scattered electrons, one can derive the observable that maximises the discrimination probability. Once the measurement scheme has been implemented, the approach consists of assigning each pixel of the image representation in the OAM space to the model A or B based on a maximum likelihood criterion. The *a priori* probability for an electron to be measured in that pixel was calculated based on the two protein models. A merit factor was established by calculating how many electrons are required in order for the probability of identifying the protein to exceed a given threshold, say 95%. It was found by numerical simulation that an optimal measurement scheme could be constructed by modifying the angular basis by an appropriate projector on a new custom basis [277]. When this strategy is applied, the number of electrons necessary for discrimination would be far less than 1 e$^-$/Å^2 (3.7 MGy). Scanning the sample in advance with such a low fluence would allow identification of the location of specific proteins of interest, while preserving most of the sample lifetime for subsequent imaging by conventional means.
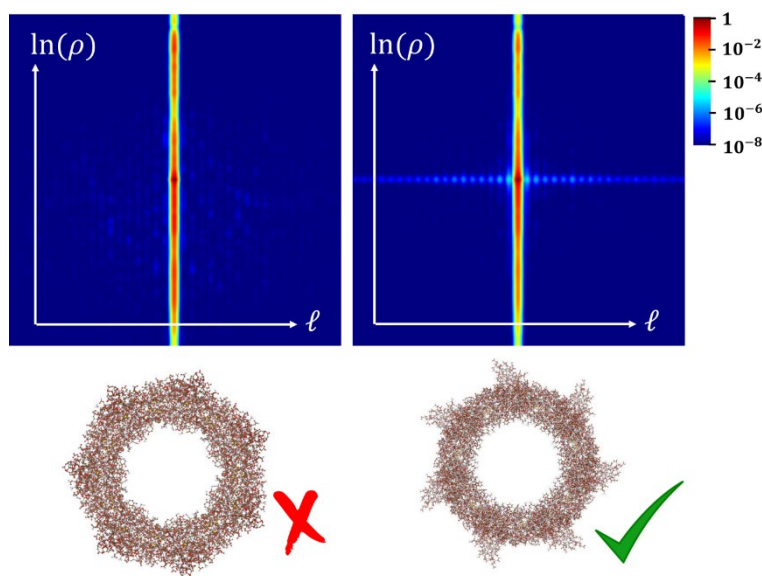
Figure 6. 7 Comparison of the custom basis representation of proteins. The use of the optimised projection permits discrimination between the two models below. It can be appreciated that most of the information is contained in the l=0 axis and in a restricted set of radial coordinates. Reproduced from [277] with permission from the American Physical Society.

6.3   An Overview

We compared the illumination mode, desirable beam energy, state-of-the-art spatial resolution, the required fluence, and tolerance to thicker sample among the techniques listed above (Table 6. 1). Features, current status, and perspectives of these techniques are also included in Table 6. 1. The listed illumination mode refers to the one most often used for that particular technique: for most techniques: other illumination modes are possible as well [189,201,278-280]. While conventional cryo-EM has achieved atomic resolution for ferritin with a fluence 40 e$^-$/ Å^2 at 300 keV (148 MGy) [1,2], smaller macromolecules and heterogeneous sample are still difficult to image. Some of the alternative techniques listed in Table 6. 1 promise prospects to overcome these barriers.

Table 6. 1 Comparison between different Cryo-EM techniques

	Conventional Cryo-EM	Laser Phase Plate	MPTEM	Holography	Ptychography	Q-Sort
Illumination mode	TEM	TEM	TEM	TEM	STEM	STEM
Currently employed beam energies	100 keV – 300 keV	100 keV – 300 keV	10 keV	20 keV – 1.2 MeV	30 keV – 300 keV	100 keV -300 keV
State-of-the-art spatial resolution and required fluence	3D reconstructions at atomic resolution (1.2 Å) ~ 40 e-/ Å² averaged by a few hundred thousands particles	3D reconstructions at near-atomic resolution (3.8 Å) ~ 50 e-/ Å² averaged by a few thousands particles	PSF is 4 - 7 nm at 10 keV according to the simulation ~ 10 to 100 e-/ Å² based on electron fluence and number of passes	2D atomic resolution for structural imaging 2D and 3D nanometre resolution for EM fields Not experimentally optimised for low dose cases	2D deep sub-Å resolution (0.39 Å) with high dose , around 400 e-/ Å² with a single image and no averaging at 2D atomic resolution 2D nanometre resolution with cryo-EM dose, ~ 20 e-/ Å² for virus particle embedded in ice	Nanometre localisation accuracy << 1 e-/Å²
Tolerance to thicker sample*	+	+	- -	-	-	+
Features	Defocus used to enhance contrast	Laser induced electron phase shift Matter free phase plate Adjustable phase shift Extra Lorentz lens High-power laser unit	Contrast enhancement by phase shift accumulation Lower absorbed dose compared to conventional cryo-EM at the same SNR	Direct phase contrast imaging without using a phase plate Exit-wave function retrieval Allows for a wide band of spatial frequency to be transferred while lower spatial frequency is better preserved	Direct phase contrast imaging without using a phase plate Image is computationally reconstructed Phase information retrieval Adjustable contrast transfer band depending on the optical setup and reconstruction algorithm Both sample transmission function and the illumination probe are retrieved which allows for post correction of residual optical aberration	Electron wavefunction is represented in OAM Aiming for specific measurement, *e.g.*, protein identification Extra elements are needed in TEM column

Current status	Atomic resolution achieved with Ferritin 2 Å resolution is becoming more routine Proteins < 50 kDa are still difficult	High power continuous-wave laser available for 300 keV electrons 3.8 Å 20S ribosome has been achieved Laser system stability needs to be ensure Redesign electron optics optimised for laser phase plate	Only simulation and design for 10 keV MPTEM are available No experimental data yet	Low dose cryo-holography has hardly been used so far for biological samples Direct electron detectors start to be used for improving the fringe visibility and therefore the signal to noise ratio.	Advanced reconstruction algorithms are being developed to account for multiple scattering Extremely high spatial resolution beyond the diffraction limit and extremely high precision only limited by thermal noise has been demonstrated for non-biological samples First attempt of ptychographic reconstruction of vitrified biological specimen has been performed	Protein discrimination can be done, according to simulation, with << 1 e$^-$/Å^2
Perspectives	Smaller macromolecules Better sample preparation Better detectors Fewer particles rejected Better treatment of heterogeneous samples	Atomic resolution should be achievable Smaller proteins should become possible Heterogeneous samples should become easier	Reliable and stable electron mirrors are needed, also at high energy Low energy MPTEM has no obvious benefit compares to current cryo-EM	Imaging setups could be designed which carefully optimise spatial resolution, field of view, and fringe contrast Split-illumination holography phase-shifting holography Smart sample preparation optimised for holographic experiments Automated cryo-holographic SPA data collection.	It could become possible to obtain 3D near-atomic resolution reconstruction with much fewer particles compared to what is used by conventional TEM Specific dose-weighting schemes to account for beam damage Automated cryo-ptychography SPA data collection Fast pixelated detector (~100 kHz) Fast data processing	Localisation of points of interest within a sample Use of prior information to guide data collection and processing Structured illumination of the sample provide entirely new data collection schemes

* Tolerance to thicker sample, +: higher tolerance for thicker sample, -: requires very thin sample.

6.4 Conclusion and Outlook

In this work, inspired by the huge interests in SPA from structural biologists, we outlined some alternative TEM and STEM imaging techniques and measurement methods, and discuss their dose efficiency. The radiation damage inflicted by high-energy electrons to the biological specimen is the major factor that limits the image SNR and the resolution. The so-called low-dose technique used in current conventional cryo-EM SPA uses the adjacent area of the sample to focus. This minimises the beam damage to the area of interest before data acquisition, and has achieved good results. However, all SPA routines accept and computationally account for the oscillating character of the CTF, which dampens and even destroys the transfer of information at certain frequencies. Altering the information transfer function could improve the dose-efficiency at which we study biological specimens.

Phase plates make use of the electron-phase information that provides a more dose-efficient approach for radiation-sensitive specimen. By using a phase plate in the back focal plane inside the TEM column, one can enhance the phase contrast of the image by altering the phase difference between the scattered and unscattered beam. Although a perfect Zernike-like phase distribution gives the optimal phase contrast, none of the current phase plates is able to give a delta function-like phase shift to the direct beam while unaffecting the scattered beam. A laser phase plate seems to be an ideal option as it offers a stable phase shift and it does not result in electron loss. With a 3.8 Å 20S ribosome reconstruction obtained with a 300 keV cryo-EM at a dose of 50 e$^-$/Å^2 (185 MGy) [208], we believe that the implementation of the laser phase plate for more cryo-EM SPA can be foreseen in the next few years.

The idea of MPTEM is to apply the transformation to the same electron multiple times to increase the SNR at a certain level of radiation damage. Another operating strategy of MPTEM is to reduce the radiation damage to the specimen while keeping SNR the same as in conventional cryo-EM. This can also be useful, since for SPA the frames above certain dose have much less high resolution information due to beam damage. However, a high voltage electron mirror, the key component for MPTEM, does not yet exist. A low-energy (10 keV) proof-of-concept MPTEM is currently under construction [221], which could be the very first step for the application of MPTEM.

Holography and ptychography have compelling advantages in dose efficient phase retrieval of the exit wave. Instead of imaging the specimen directly, holography records the interference pattern of the sample beam and reference beam. We are not aware of any recent reports of off-axis holography experiments on ice-embedded biological specimens, possibly due to the low contrast. A recent report of low-energy an in-line holography experiment pointed out that the ice thickness plays a key role in the contrast enhancement of the hologram [246]. Together with additional approaches, *e.g.* phase-shifting holography, and split-illumination holography, it could

revitalise the application of holography on ice-embedded biological specimen. Ptychography reconstructs the exit-wave phase information computationally by recording a series of diffraction patterns. Cryo-ptychography has attracted researchers' interest in recent years, and phase images of ice-embedded rotavirus have been reconstructed experimentally with the fluence of 20 e$^-$/Å^2 (75 MGy) [153]. With better illumination schemes and faster pixelated detectors actively being developed, cryo-ptychography for biological applications is likely to produce many more results in the years to come.

The use of a different basis compared to the traditional spatial representation of the electron wavefunction, could open up entirely new concepts. We discussed the use of an angular basis, namely the orbital angular momentum, to identify the presence and location of known proteins within the sample, which here could be done in a highly dose-efficient manner. This might not only have potential to be applied to biomolecules in solution, but also in the crowded cellular environment, such as for cryo-tomography of lamella of (infected) eukaryotic cells.

Conventional TEM has seen tremendous developments over the last 8 years, and the limits of SPA are still being pushed. Here we discussed a number of non-standard techniques, at different pioneering stages. In addition to the techniques discussed here, other techniques, such as aloof beam electron energy-loss spectroscopy (EELS) [281-283], structured illumination with compressed sensing [284-286], and adaptive optics with pixelated phase plates [212] could also have potential applications in imaging or characterisation of biological specimens. TEM at low voltage reduces the radiation damage to the sample [127,287], and ultra-low energy electron microscopy (LEEM) and eV-TEM at 0-30 eV have much less plasmonic and excitonic interaction which results in almost no energy being deposited in the specimen [288,289]. TEM at such low acceleration voltage can, however, only achieve spatial resolution at a few nm range which will not allow for *de novo* structure determination. All in all, there are several alternative cryo-TEM and cryo-STEM schemes for obtaining more information during the limited lifetime of a biomolecule within the electron beam, thereby further pushing the limits of size, structural heterogeneity, and resolution at which one can study the building blocks of life.

ACKNOWLEDGEMENT

CHAPTER 7

DISCUSSION AND SUMMARY
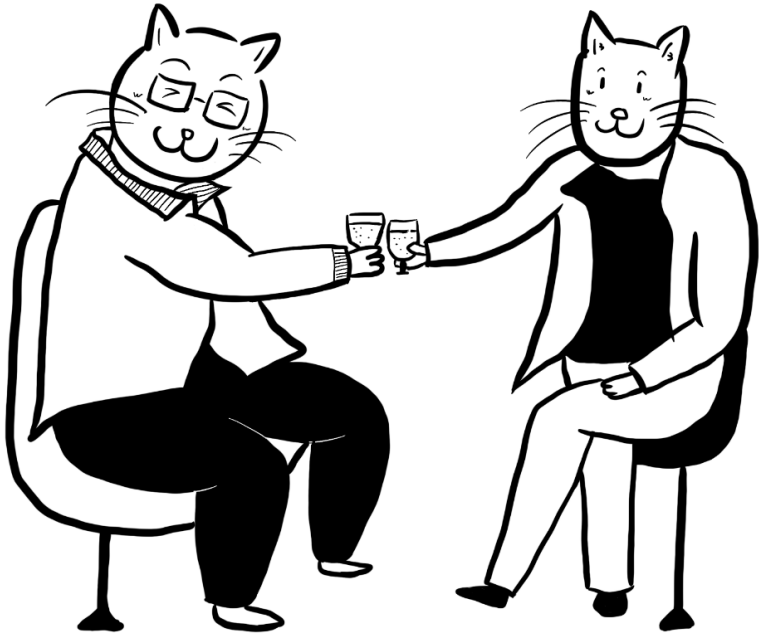
7.1 GENERAL DISCUSSION

Cryo-EM has become a powerful tool in the field of structural biology, enabling the visualisation of macromolecular structures at high resolution. With spectacular advances in hardware and software, cryo-EM has become increasingly accessible, leading to an explosion in its popularity and a surge in the number of structures being solved. Despite progress made thus far, there is still a considerable disparity between current achievements and the theoretically possible outcomes. Further optimisations are required in various areas, beyond those discussed in previous chapters, to bridge this gap and reach the theoretical potential.

Sample preparation

In **chapter 2**, the BfrB protein was introduced as a benchmark sample for characterising cryo-EM hardware and software development, with a simple and high-yield purification protocol. The importance of high yield is highlighted when considering the sample preparation process in SPA. Typically, a droplet of 3 μl of sample is dispensed onto a 3 mm-diameter EM-grid before blotting. Assuming a uniform distribution of a 100 nm-thick liquid on the grid after blotting, only about 0.02% of the sample remains on the grid. Subsequently, even with an optimistic scenario of collecting 5000 micrographs during data collection on one grid, each with 6k x 4k pixels at a pixel size of 1 Å/pix, only 0.02% of the surface area is imaged. Then, during data processing, in another very optimistic scenario, 30% of the particles selected from the micrographs contribute to the final reconstruction. Therefore, obtaining a high yield of the sample is essential as the final high-resolution map is derived from approximately only 12 parts per billion of the original droplet's sample. On the other hand, such calculation illustrates that optimising sample preparation methods is essential if we wish to use less of the sample. For an ideal sample, a single layer of proteins with random orientation and sufficient concentration should be embedded in a vitreous ice layer with a more or less uniform thickness across the entire grid. However, the current blotting and plunge freezing methods are far from optimal and lots of problems remain.

In recent years, various alternative techniques for depositing samples onto grids have been developed aiming to prepare better samples, also in our laboratory in Maastricht. There are two main categories of sample deposition: scribing [56,290] and spraying [57,291-294]. Scribing methods use either pin printing [56] or capillary writing [290] for applying the sample onto the EM grid, which only require (sub)nano-litres of sample. Scribing offers benefits in being fully automated, allowing for good reproducibility of sample thickness by controlling the scribing velocity, while blotting relies on user expertise and experience, which will lead to variability in sample preparation. Spraying requires varying sample volumes per grid, ranging from pico-litres to micro-litres. However, it offers the possibility to capture transient structural states due to the short time interval between deposition and vitrification, which has been exploited for time-

resolved cryo-EM [295]. Meanwhile, samples can be mixed before spraying to the grids, allowing the study of molecular interactions [296,297].

Nevertheless, there are still challenges to be addressed with sample preparation for cryo-EM. One such challenge is that single particles collide with the air-water interface, resulting in preferred orientation, denaturation, and even low concentration of sample within the suspended ice layer [85,298,299]. A rapid process of sample deposition and vitrification could potentially reduce the time for interactions of proteins with the air-water interface [300], but it does not overcome the issue as particles will hit the air–water interface orders of magnitude faster than the time that is used by current sample preparation devices to perform the sample deposition and vitrification process. Innovations in sample carriers, such as functionalised graphene, have been shown to alleviate protein denaturation and preferred orientation at the air-water interface [39,51,85]. Nano-fluidic chips fabricated with Micro-ElectroMechanical Systems (MEMS) technologies can also address this issue by embedding sample in two electron transparent windows (SiN_X), however, such layer can interact with incoming electrons inelastically which introduces noise [301]. Graphene is an ideal material for such windows in terms of electron transparency, but it is also relatively brittle, and imperfections can lead to cracks, preventing it from holding liquid inside two graphene layers.

Another challenge is the stress built up within the specimen during the process of cryo-cooling due to the difference in thermal expansion coefficients between water, support foil and grid [32,33]. The release of stress within the specimen upon electron irradiation results in a "burst" movement at the onset of irradiation [34]. Although several methods have been proposed to address this issue, they have also be shown to introduce new problems such as ice contamination [117,118].

In general, sample preparation is a critical step in cryo-EM, and there are still many challenges that need to be overcome. We look forward to seeing new developments in this field.

More accessible and affordable electron microscopes

The elastic-scattering cross section for carbon at 100 keV is more than twice that at 300 keV [91,302], however, the inelastic-scattering cross-section increases slightly less. As a result, effectively less damage (expressed as energy deposited per sample weight) occurs at lower energies, which should result in better image contrast and higher SNR. Naydenova, et al. [92], demonstrated that good reconstructions could be obtained from data collected on a 100 keV microscope for a number of tested proteins, with resolutions ranging from 3.4 to 8.4 Å. One major advantage of a lower-voltage electron microscope is its relative affordability, which could make it accessible to more researchers and labs. It can also allow for more rapid screening of samples and facilitate high-throughput structural studies. Despite the numerous benefits of lower-voltage cryo-EM, there are still several aspects that require optimisation.

As mentioned in **chapter 5**, MAPS detectors are currently optimised for 300 keV microscopes, and contain a back-thinned sensor layer which helps to improve the MTF at these energies. Lower energy electrons tend to scatter at larger angles and backscatter in the sensor layer, thereby degrading the PSF of the detector. One possible solution to improve the PSF is to further thin the sensor layer. However, the limit of back-thinning has already been reached with the current sensor layer thickness at approximately 30 um [147]. Therefore, alternative approaches are required to optimise detectors for low-energy electron microscopy. One possible solution to limit the impact of a large PSF caused by lower energy electrons is to use larger pixel size. Another promising solution is the use of HPDs that gives energy and time information of each incident particles, making it possible to further boost the MTF with neural network-assisted event localisation methods [146,147].

There are a few fundamental issues regarding imaging with lower energy electrons. One of the issue is that a thin sample is required to avoid multiple scattering, as lower energy electrons have a shorter IMFP. While this approach is effective for SPA, it may not be suitable for tomography, where samples are typically thicker. Meanwhile, lower electron energy increases the curvature of the Ewald sphere, which needs to be computationally accounted for. Moreover, imperfections in the temporal and spatial coherence of the electron beam dampens the CTF, which becomes more severe at lower energies (Eq. 3.4, 3.5, 3.6). A gun monochromator or a cold field emission gun with less energy spread would help to improve the microscope envelope function, in particular at low energies. However, such additions would make the microscope more expensive again.

A 100 keV microscope is expected to be a suitable instrument for most standard laboratories for SPA work. It can serve as a screening tool to evaluate grid quality prior to using a 300 keV microscope for high-resolution data collection. Improved cryo-holders (eg: http://www.hennyz.com) could overcome the need for expensive AutoLoaders. Furthermore, if equipped with suitable detectors, and a monochromator or a cold field emission gun, these microscopes can potentially reduce the time and costs involved in high-resolution data collection and analysis, making cryo-EM more accessible to a wider range of researchers.

How could artificial intelligence help?

In 2018, AlphaFold emerged as the winner of the "Critical Assessment of Structure Prediction" (CASP) competition [303], and its successor, AlphaFold2, achieved an even more remarkable success in the competition in 2020 [304]. This demonstrated the significant potential of AI in predicting 3D protein structures with higher accuracy and speed than traditional experimental methods [305,306]. With more than 200 million proteins predicted by AlphaFold nowadays, almost every known protein on this planet has been covered. It opens up new avenues for drug discovery and protein engineering [307]. Currently, the neural networks of AlphaFold have been trained using all the experimental structures available in the PDB. These structures have been determined using techniques such as X-ray crystallography, cryo-EM, and NMR. However, these

techniques only capture proteins in specific conformations, while proteins are dynamic entities which interact with other proteins at various stages of their functional cycles. Many drugs target dynamic regions, including intrinsically disordered protein regions. Therefore, there is a need for improved accuracy in predicting protein structures, their energy landscape, their possible conformations, their ligand binding sites, and their interactions with other proteins. Future developments in AI-based protein structure prediction should address these challenges to further facilitate drug discovery and protein engineering.

Cryo-EM offers the advantage of capturing macromolecular complexes in their nearly-native state. Particles are embedded in amorphous ice and could retain their structural heterogeneity. However, this heterogeneity poses a significant challenge to image processing as it requires effective determination of different 3D maps from the noisy micrographs. To address these challenges, recent developments in AI have been applied to cryo-EM image processing, including micrograph pre-processing [308,309], particle picking [308-312], 3D reconstruction [313,314], map sharpening and local resolution estimation [315,316], and model building [317-319].

One area where AI has shown particular promise is in particle picking, which is a crucial step in the image analysis pipeline. Historically, particle picking has been a manual and time-consuming task that required a significant amount of expertise. Recent advances in deep learning applications have led to the development of AI-based methods that can (semi-)automate particle picking with high accuracy and efficiency [308,309]. Moreover, AI can also be used to refine 3D maps and to identify different proteins in given micrographs [320]. But addressing heterogeneity of biological molecules is still a main challenge for 3D reconstruction. As is stated in the introduction, a main constraint of cryo-EM is the low SNR of images recorded within the "safe" electron irradiation dose, with ice being one source of noise. Despite the large number of cryo-EM datasets deposited (1248 entries on EMPIAR and 14562 entries on PDB solved by EM as of March 16, 2023), we gain lots of knowledge about proteins but little about noise. Even in simulations, reproducing real noise remains challenging (chapter 3). One promising approach is to train AI with real noise, enabling better accounting for it in data processing. There is potential to learn from noise and improve the SNR of micrographs through AI-based denoising techniques [321].

AI has shown great promise in improving the accuracy and efficiency of cryo-EM image processing, making it possible to solve high-resolution 3D structures of macromolecular complexes that were previously difficult to solve, and reducing the need for human intervention in the structure determination pipeline. As the field of AI continues to advance, we look forward to seeing new developments that further improve the accuracy and efficiency of cryo-EM image processing. The integration of AI with other imaging techniques, such as cryo-electron tomography and correlative light and electron microscopy, could also open up new avenues for studying the structure and dynamics of macromolecular complexes in their native cellular environment.

7.2 FUTURE PERSPECTIVE

Can we routinely achieve atomic resolution for sub 50 kDa macromolecules, map their conformational landscape, and even resolve intrinsically disordered regions with the help of better sample preparation, microscopes, detectors, advanced imaging techniques, and AI?

Furthermore, if we look beyond imaging at cryogenic temperature, can we image biological systems at room temperature to observe dynamic processes in their native state within complex networks at atomic resolution?

Prof. Peter Brzezinski stated in the scientific background on the Nobel Prize in Chemistry 2017: "Perhaps in the future we will be able to obtain high-resolution structural information of molecules, as well as to observe interactions and dynamic processes as they happen, inside cells or organelles." We eagerly anticipate further developments in this area, which will undoubtedly lead to new discoveries and applications in fields, as we continue to strive towards a better understanding of the fundamental workings of life on a molecular level.

7.3 SUMMARY

The five studies presented in this thesis all aim to contribute towards to more dose-efficient imaging and diffraction techniques in cryo-EM such that the SPA limits could be pushed towards smaller or more heterogeneous samples.

For any cryo-EM development, it is crucial to have a suitable workhorse protein. In **Chapter 2**, we introduced *Mycobacterium tuberculosis* ferritin as an excellent model protein due to its good stability and straightforward purification protocol. The protein's octahedral symmetry, good solubility, and appropriate size (~490 kDa) make it possible to determine its structure with just a few micrographs, enabling rapid characterisation of, for example, new microscopes, detectors, or grid preparation techniques. Also, better understanding of mycobacterial proteins in general could help, ultimately, in fighting tuberculosis which still awaits the development of a better vaccine.

Cryo-EM is advancing in the direction of solving smaller proteins, given that 50% of animal and plant proteins have a size of less than 50 kDa, and many of these proteins are relevant drug targets. Recent progress in cryo-EM has resulted in the successful reconstruction of a haemoglobin heterodimer at 32 kDa, surpassing the previously estimated theoretical limitation (38 kDa). What could be the ultimate capabilities of cryo-EM in this field? In **chapter 3**, we simulated single particle data sets using realistic parameters for ice layer, dose, detector performance, and beam characteristics for samples that were ideal in terms of homogeneity, distribution, and stability. Our simulations indicate that, with an ideal phase plate, a protein as small as hen egg white lysozyme (14kDa) could be resolved by cryo-EM.

In **chapter 4**, we discuss the effects of charging and its potential impact on SNR and the ability to achieve high-resolution reconstructions. Charging not only causes image distortions but also motion of the sample due to build-up of electrostatic forces. By deploying a layer of graphene on regular EM grids, we could mitigate charging. These observations were made both in diffraction and imaging mode. Use of graphene could further improve the throughput of routine data collection by inclusion of the middle of the hole. This approach could also reduce beam-induced motions of the sample, thereby further improve the image quality. Mitigating charging might proof to be crucial for STEM based techniques, such as ptychography and iDPC.

In **chapter 5**, a detector with high time resolution and dynamic range, the Timepix3, was integrated in an automated cryo-EM workflow. The Timepix3 setup has been built in Maastricht together with Amsterdam Scientific Instruments. We characterised the Timepix3 in terms of MTF, NNPS, DQE at 200 keV. Using this event-driven HPD in electron counting mode with high hit rates and CNN-based event localisation, high-quality SPA reconstructions of BfrB (chapter 2) at 3 Å have been obtained.

We discuss various techniques that potentially have higher dose efficiency in **chapter 6**. Laser phase plate could become a nearly ideal phase plate for TEM as it is matter free and gives tuneable phase shifts. Such a phase plate could be helpful to further push the size limit of cryo-EM. MPTEM gives possibilities to boost the SNR by having the same electron passing through the specimen multiple times. However, this technique is still far from routine for a multitude of reasons. Primarily, it is the lack of high-voltage electron mirrors. Dose efficient phase retrieval of the exit wave can be achieved through holography, ptychography, and iDPC (not discussed in **chapter 6**), which offer significant advantages. Both ptychography and iDPC are STEM techniques which involves using a small convergent beam scanning through the specimen. The studies in **chapter 4** and **chapter 5** addressed critical challenges when applying these techniques to biological specimens. For instance, charging can limit the achievable resolution, and a high-speed detector capable of operating in diffraction mode is required. The use of a sorter in EM is an entirely novel approach of discriminating particles based on the orbital angular momentum of the exit-wave in a highly dose-efficient manner.

Combined, these approaches could help to expand the limits of cryo-EM in structural biology, thereby enabling the visualisation of a larger range of macromolecular structures, at higher resolutions, in more conformational states. Over the last decade, cryo-EM has revolutionised the field of structural biology, providing insights into the structure-function relationships of many biomolecules which previously failed to be revealed by other techniques. In the decade to come, cryo-EM has the potential to revolutionise structural biology yet again. We hope that our research will contribute a small but valuable contribution towards this goal.

7.4 SAMENVATTING

De vijf in dit proefschrift beschreven studies hebben als doel om een bijdrage te leveren aan de ontwikkeling van efficiëntere beeldvormings- en diffractietechnieken in cryo-EM. Hierdoor wordt het in toekomst mogelijk om kleinere eiwitten en meer heterogene complexen van eiwitten structureel te karakteriseren en begrijpen.

Voor de ontwikkeling van cryo-EM is het belangrijk om een geschikt modeleiwit te hebben. In **hoofdstuk 2** hebben wij het eiwit ferritine van *Mycobacterium tuberculosis* geïntroduceerd als modeleiwit. We laten zien dat dit eiwit eenvoudig te produceren en te purificeren is. Bovendien is het eiwit stabiel, goed te bewaren en kan het prachtige cryo-EM afbeeldingengeven. De octahedrale symmetrie van het eiwit, de goede oplosbaarheid en de geschikte grootte (~490 kDa) maken het mogelijk de structuur te bepalen met slechts enkele opnames, waardoor een snelle karakterisering van bijvoorbeeld nieuwe microscopen, detectoren of technieken voor preparaatvoorbereiding mogelijk wordt. Daarnaast zou een beter begrip van mycobacteriële eiwitten uiteindelijk kunnen helpen bij het ontwikkelen van nieuwe vaccins, die nog steeds hoognodig zijn voor de bestrijding van tuberculose.

Cryo-EM ontwikkelt zich in de richting van het oplossen van steeds kleinere eiwitten. Maar liefst de helft van alle dierlijke en plantaardige eiwitten zijn kleiner dan 50 kDa. Veel van deze eiwitten zijn relevant om medicijnen mee te maken. Recente ontwikkelingen in de cryo-EM hebben ertoe geleid dat de structuur van een heterodimeer van hemoglobine van 32 kDa succesvol bepaald kon worden. Dit is kleiner dan de eerder geschatte theoretische limiet (38 kDa). De vraag is hoe ver cryo-EM uiteindelijk kan komen in het bepalen van steeds kleinere structuren. In **hoofdstuk 3** hebben wij datasets met afzonderlijke deeltjes gesimuleerd met behulp van realistische parameters voor de eigenschappen van de ijslaag, de detector en de elektronenbundel voor preparaten die ideaal waren wat betreft homogeniteit, distributie en stabiliteit. Onze simulaties geven aan dat met een ideale faseplaat, de structuur van een eiwit zo klein als lysozyme (uit een kippenei geïsoleerd) (14kDa) kan worden bepaald met cryo-EM.

In **hoofdstuk 4** bespreken we een fundamenteel effect in cryo-EM: het elektrostatisch opladen van de preparaten en het effect daarvan op de beeldvorming en de signaal-ruisverhouding van de data. Elektrostatisch opladen veroorzaakt niet alleen beeldvervormingen, maar kan ook het preparaat daadwerkelijk laten bewegen door het samenspel van de elektrostatische krachten. We laten in dit hoofdstuk zien dat de effecten van het opladen teniet gedaan kunnen worden door een laag grafeen aan te brengen op het EM-grid. Deze waarnemingen zijn zowel in de diffractie- als in de beeldvormingsmodus gedaan. Het gebruik van grafeen maakt het mogelijk om meer data te verkrijgen van een preparaat vergeleken met de meest gebruikte technieken. We laten zien dat het nu ook mogelijk is om met een kleine elektronenbundel data op te nemen wanneer deze gericht wordt op het midden van een gat in de folie op een EM-grid. Deze aanpak zou ook de bewegingen van het preparaat kunnen verminderen, wat de beeldkwaliteit verder zou

verbeteren. Onze ontdekking zou bovendien van belang kunnen zijn voor STEM technieken, zoals ptychografie en iDPC.

In **hoofdstuk 5** hebben we een detector met een extreem hoge tijdresolutie en dynamisch bereik, de Timepix3, geïntegreerd in een geautomatiseerde cryo-EM workflow. Deze opstelling is in samenwerking met Amsterdam Scientific Instruments in Maastricht gebouwd. We hebben de Timepix3 gekarakteriseerd in termen van MTF, NNPS, DQE bij 200 keV. Met deze gebeurtenis-gedreven HPD in elektronen-telmodus en CNN-gebaseerde gebeurtenislokalisatie konden hoogwaardige SPA-reconstructies van BfrB (**hoofdstuk 2**) van 3 Å worden verkregen.

In **hoofdstuk 6** bespreken we verschillende technieken die mogelijk een hogere dosisefficiëntie hebben. Een laserfaseplaat zou een bijna ideale faseplaat voor TEM kunnen zijn, omdat deze vrij is van materie en een afstembare faseverschuiving geeft. Een dergelijke faseplaat zou nuttig kunnen zijn om de grenzen van de cryo-EM verder te verleggen. MPTEM maakt het mogelijk om de signaal-ruisverhouding te verhogen door hetzelfde elektron meerdere malen door het preparaat te laten gaan. Om tal van redenen is deze techniek echter nog verre van gebruikelijk. Een van de belangrijkste redenen is het ontbreken van een spiegel voor hoogspanningselektronen . Het bepalen van de fase van de uitgangsgolf kan ook bereikt worden door holografie, ptychografie en iDPC (niet besproken in **hoofdstuk 6**). Deze technieken zouden daardoor dosis-efficiënter kunnen zijn. Zowel ptychografie als iDPC zijn STEM-technieken waarbij een kleine convergente bundel punt-voor-punt het preparaat scant. De studies in **hoofdstuk 4** en **hoofdstuk 5** behandelen de belangrijke uitdagingen bij de toepassing van deze technieken op biologische preparaten. Zo kan het statisch opladen de maximaal haalbare resolutie beperken, en is een snelle detector nodig die in diffractiemodus kan werken.

Gecombineerd zouden bovengenoemde technieken kunnen helpen om de grenzen van cryo-EM in de structuurbiologie te verleggen. Hierdoor kan een groter scala van macromoleculaire structuren, bij hogere resoluties en in meer conformatietoestanden worden gevisualiseerd. In de afgelopen tien jaar heeft cryo-EM een revolutie teweeggebracht op het gebied van de structuurbiologie, door inzichten te verschaffen in de structuur-functie relatie van biomoleculen die voorheen niet met andere technieken konden worden bepaald. Cryo-EM heeft de potentie om in het komende decennium de structuurbiologie opnieuw in een stroomversnelling te brengen. Wij hopen dat ons onderzoek een waardevolle bijdrage aan dit doel levert.

Appendices

IMPACT

As of March 2023, the number of cryo-EM labs worldwide has grown exponentially to more than one thousand (1056). The majority of these labs possess state-of-the-art equipment, such as the Titan Krios (355), JEOLs (42), Arctica (94), Glacios (208), Tundra (13), and Talos F200C (39) [322]. As a result, the cryo-EM market has become one of the fastest-growing markets in the scientific industry. A recent report projects that the entire cryo-EM market will drive growth of over 1 billion USD by 2027 [323]. Investing in cryo-EM technology provides companies with the opportunity to develop new products and stay ahead of the competition, leading to increased competitiveness in various industries. Cryo-EM technology has revolutionised the field of structural biology. In particular, cryo-EM SPA has allowed researchers to reconstruct 3D structures of macromolecules at atomic resolution [1,2].

However, it is worth noting that a top-end Titan Krios (> 5 M€) may not always be necessary for SPA, especially when imaging samples with a thickness less than 100 nanometres. In fact, lower-voltage cryo-EMs can produce higher-contrast images due to their more favourable elastic/inelastic cross-section ratio, as mentioned in **chapter 7.** One major advantage of lower-voltage microscopes is their relative affordability, making them more accessible to researchers and labs. The trend on the cryo-EM worldwide map already shows that more labs are obtaining 100 keV Tundra microscopes. The main constraint for obtaining high-resolution structures using a 100 keV electron microscope is the lack of an optimal detector. In **chapter 5**, the integration of a Timepix HPD detector, which works well at all energies, in imaging as well as diffraction mode, has shown promising results. Low-energy microscopes equipped with an HPD offer better affordability, making cryo-EM technology more accessible to researchers and labs in the fields of life and material sciences. This democratisation of the access to cryo-EM technology would allow more researchers to use cryo-EM to study various biological systems, potentially leading to new discoveries and advances in our understanding of complex biological processes.

As it is said by Werner Heisenberg "We have to remember that what we observe is not nature in itself, but nature exposed to our method of questioning." When imaging samples with electrons in the EM, there are multiple processes happening, such as radiation damage, hydrogen gas generation, beam-induced motions, and charging. Charging is an important and fundamental issue, as biological specimens are typically insulators. Charging will affect our observations and the interpretation of the results (**chapter 4**). While charging may not seem to have any significant social or economic impact, studying it is nonetheless crucial. For instance, during the space race between the US and the Soviet Union, the US government invested billions of dollars in its space program between 1957 to 1969, with a significant portion of the funds going towards fundamental research. This research had a substantial impact on many areas of science and technology, foreseen and unforeseen, such as GPS technology, miniaturised electronics, and environmental monitoring. Although the charging issue may not be as significant as exploring

space, it is a fundamental problem when imaging specimens using electrons. It is essential to remember that what we are observing is not the sample in itself, but rather the sample as affected by our method of questioning.

Currently, SPA cryo-EM employs a basic technique that involves irradiating the sample with a parallel flood beam to obtain its projections. This approach, however, comes with several limitations such as low SNR, charging, and radiation damage (although this is inevitable for EM). Quantum mechanics dictates that if we can obtain the wave function of an exit electron, we can potentially retrieve all the information about the sample. Unfortunately, the wave function collapses by the time the electron reaches the detector, limiting the amount of information we can obtain. Nevertheless, recent advances in measurement techniques, as discussed in **chapter 6**, have enabled more dose-efficient methods of data collection, e.g. by measuring all the characteristics of its wave. While some techniques, such as MPTEM and quantum sorters, may not be immediately applicable, others such as laser phase plates, holography, and ptychography hold evident promises for the near future. These methods could allow for more information to be extracted during the limited lifetime of a biomolecule within the electron beam, thus pushing the limits of resolution and enabling the study of the building blocks of life in unprecedented detail. This could lead to a better understanding of biological processes and potentially facilitate the development of new therapeutics to fight disease.

Acknowledgements

Raimond, I would like to dedicate this thesis to you. You have been an exceptional mentor and friend throughout my PhD journey. You never wavered in your commitment to helping me develop as a scientist. You generously shared your vast knowledge, and taught me valuable lessons about the scientific process. Our discussions were always stimulating, and you constantly inspired me with your great ideas. You went above and beyond to support my research, including giving me the freedom to explore different directions and supporting me during our conference presentations. Thank you for your encouragement and guidance, which have been instrumental in shaping my academic and professional trajectory. I will always remember how much you loved vlaai, and how you would bring it to the lab to share with us during our meetings. Your kindness and generosity extended beyond the lab, and I am grateful for the warmth and support you have shown me during my time as your PhD student. May strength, healing, and love surround Maaike, Noe and Seppe, and all the best!

I would like to express my heartfelt thanks to my PhD supervisors Peter and Carmen, for your invaluable guidance and support throughout my research. Your expertise and insights give direction to my work and overcome various challenges. I appreciate for your encouragement, and the opportunities you provided to further my academic and professional growth. I am fortunate to have had such dedicated and inspiring mentors, and I will carry the lessons I learned from you with me throughout my career.

My dear friend Abril, we created wonderful memories during our time working together and the weekend lunch parties we shared. Your enthusiasm and friendly demeanour brought joy, inspiration, and positive energy into my life, and I am lucky to have you as a friend. And of course, let's not forget our plan to travel together!

Paul, I appreciate the time we spent working together. Your expertise and dedication were indispensable to the project's success, and your positive attitude made you a pleasure to work with. I am fortunate to have collaborated with such a knowledgeable and supportive colleague. And best wishes to Pauline, Nolan, and Woud!

Navya, you possess greater strength than you may realise! Developing a new machine is extremely difficult, but your hard work and determination to overcome challenges and achieve success is remarkable. Thank you for your kindness in taking care of Flora when we were not home, and for being an amazing officemate. I have enjoyed all of our conversations and I am confident that you will continue to accomplish great things in your academic and professional pursuits.

Sneha, I appreciate your positive energy and enthusiasm. Our daily coffee breaks are always a highlight of my day! I wish you the best of luck in your PhD journey, and I have no doubt that you will continue to achieve great things!

As I reminisce about my time in Maastricht, the names of Kèvin, Hang, Jianhua, Ye, Casper, Chris, Rene, Hans, Frank, Helma, Willine, Nuria, Anjusha, Penghan, Rafal, and many more come to my mind. These incredible people not only worked with me but also filled my life with joy and unforgettable moments that I cherish to this day. We shared meals, drinks, stories, and laughter, and I learned so much from every one of you. From work-related challenges to personal experiences, you taught me precious lessons. I look forward to continuing our friendship and collaboration in the future.

Hao and Jiali, the years we spent together in Delft will always hold a special place in my heart. We studied and worked on our master's thesis together, and supported each other through thick and thin. We also explored different parts of the world together, from an exchange program in Paris to a trip to Abisko. The countless laughs, conversations, and adventures we shared will stay with me forever, and I am grateful for the lifelong friendship we have built.

Boyao and Chengyu, I want to express my heartfelt gratitude for the amazing times we've shared while working together. The memories we've created together are truly special and I feel fortunate to have you both in my life. Your support, and laughter have made such a positive impact on me, and I look forward to continuing our friendship and making more unforgettable memories together in the future!

Xin and Yucheng, you are the best travel companions, and together, we have explored Spain, Slovenia, Belgium, and Germany. We have made unforgettable memories, and I cannot imagine better travel mates than you both. Looking forward to continuing our travels together in the future!

Jian, it was a pleasure to work with you during our time in Delft. Wishing you all the positivity and happiness in the world as you move forward! I hope we can continue to collaborate in the future and cycle together again!

Yifeng and Yiran, it's hard to believe how quickly time flies. It feels like yesterday that we were together in Delft, learning new things and making unforgettable memories. I will never forget the time we spent together and wish you both all the best in your future. I hope you find success, happiness, and fulfillment in all that you do!

Zidan, Liyin, Guangtao, Zhi, Yize, Kang, Peng, Minglei, Jiaming, Yanqi, Jiayu, Zheng, Tianyi, Sihan and all those who supported me along the way. Your friendship, kindness, and encouragement have been precious to me, and I feel blessed to have you all in my life.

Zheng, I am grateful to have had you as my favourite teacher and friend during my bachelor studies. Your encouragement to explore different paths in life and experience diverse cultures outside of China has been invaluable. Thank you for being an important part of my life and helping me grow.

Lilian Miesen, I am forever grateful to you for saving my life. The ladder you provided rescued me from the water, and the blanket you offered brought warmth to my body. I express my deepest appreciation and send my warmest wishes to you for a life filled with happiness and blessings.

I also want to thank my feline friend, Flora, for the constant comfort and companionship throughout my PhD journey. Her playful nature and soft purrs provided a welcome distraction from the demands of my research, reminding me to take breaks and appreciate life's simple pleasures. Though she may not be able to read this, I thank Flora for her unconditional love and support.

I would like to express my deepest appreciation to my parents! Your love, support, and encouragement throughout my life shaped me into a positive person. Your sacrifices, guidance, and understanding helped me reach where I am today. Thank you for your faith in me, and for instilling the values of perseverance and determination in me from a young age. Without you, I would not have achieved this milestone, and I dedicate this accomplishment to you.

Last but not least, Rong, you have been a constant source of support and encouragement throughout my PhD journey. Your unwavering belief in my abilities and your willingness to listen to my ideas and frustrations have been invaluable. Despite the demands of my research, you have always been there for me, and offered love and understanding when I needed it most. Thank you for your patience, kindness, and the countless sacrifices you have made to help me achieve my goals. Without you, this journey would have been much more difficult. I am fortunate to have you in my life, and I look forward to the future knowing that I will have you by my side.

List of publications

Zhang, Y.*, van Schayck, J. P.*, Pedrazo-Tardajos, A., Claes, N., Noteborn, W. E. M., Lu, P. -H., Duimel, H., Dunin-Borkowski, R. E., Bals, S., Peters, P. J., Ravelli, R. B. G. Charging of vitreous samples in cryo-EM mitigated by graphene. *Manuscript in prep.*

Gao, Y., **Zhang, Y.**, Hakke, S., Peters, P. J., Ravelli, R. B. G. Cryo-EM structure of cytochrome bo3 quinol oxidase assembled in peptidisc reveals an open conformation for ubiquinone-8 release and a Ni-NTA binding motif. *Manuscript in prep.*

van Schayck, J. P.*, **Zhang, Y.***, Knoops, K., Peters, P. J., Raimond, R. B. G. (2022). Integration of an Event-driven Timepix3 Hybrid Pixel Detector into a Cryo-EM Workflow. *Microscopy and Microanalysis* 29, 352-363. doi: 10.1093/micmic/ozac009.

Zhang, Y., Lu, P. -H., Rotunno, E., Troiani, F., van Schayck, J. P., Tavabi, A. H., Dunin-Borkowski, R. E., Grillo, V., Peters, P. J., Ravelli, R. B. G. (2021). Single-particle cryo-EM: Alternative schemes to improve dose efficiency. *Journal of Synchrotron Radiation* 28, 1343–1356. doi: 10.1107/s1600577521007931.

Gijsbers, A.*, **Zhang, Y.***, Gao, Y., Peters, P. J. & Ravelli, R. B. G. (2021). Mycobacterium tuberculosis ferritin: A suitable workhorse protein for cryo-EM development. *Acta Crystallographica Section D* 77, 1077–1083. doi: 10.1107/s2059798321007233.

Tavabi, A. H., Rosi, P., Ravelli, R. B. G., Gijsbers, A., Rotunno, E., Guner, T., **Zhang, Y**., Rocaglia, A., Belsito, L., Pozzi ,G., Tibeau, D., Gazzadi, G., Ghosh, M., Frabboni, S., Peters, P. J., Karimi, E., Tiemeijer, P., Dunin-Borkowski, R. E., Grillo, V. (2021). Symmetry and planar chirality of a protein measured on an angular basis in a transmission electron microscope. Preprint arXiv.

Zhang, Y., Tammaro, R., Peters ,P. J., Ravelli, R. B. G. (2020). Could Egg White Lysozyme be Solved by Single Particle Cryo-EM? *Journal of Chemical Information and Modeling* 60 (5), 2605-2613. doi: 10.1021/acs.jcim.9b01176.

* Shared co-first authorship

REFERENCES

1   Nakane, T. *et al.* Single-particle cryo-EM at atomic resolution. *Nature* **587**, 152-156, doi:10.1038/s41586-020-2829-0 (2020).

2   Yip, K. M., Fischer, N., Paknia, E., Chari, A. & Stark, H. Atomic-resolution protein structure determination by cryo-EM. *Nature* **587**, 157-161, doi:10.1038/s41586-020-2833-4 (2020).

3   Science Museum. *The Microscope*, <https://www.sciencemuseum.org.uk/objects-and-stories/medicine/microscope> (2019).

4   De Rosier, D. J. & Klug, A. Reconstruction of three dimensional structures from electron micrographs. *Nature* **217**, 130-134, doi:10.1038/217130a0 (1968).

5   Fernandez-Moran, H. Low-temperature preparation techniques for electron microscopy of biological specimens based on rapid freezing with liquid helium II. *Ann N Y Acad Sci* **85**, 689-713, doi:10.1111/j.1749-6632.1960.tb49990.x (1960).

6   Taylor, K. A. & Glaeser, R. M. Electron microscopy of frozen hydrated biological specimens. *J Ultrastruct Res* **55**, 448-456, doi:10.1016/s0022-5320(76)80099-8 (1976).

7   Taylor, K. A. & Glaeser, R. M. Electron diffraction of frozen, hydrated protein crystals. *Science* **186**, 1036-1037, doi:10.1126/science.186.4168.1036 (1974).

8   Dubochet, J. & McDowall, A. W. Vitrification of Pure Water for Electron Microscopy. *Journal of Microscopy* **124**, 3-4, doi:10.1111/j.1365-2818.1981.tb02483.x (1981).

9   Dubochet, J., Chang, J. J., Freeman, R., Lepault, J. & McDowall, A. W. Frozen aqueous suspensions. *Ultramicroscopy* **10**, 55-61, doi:10.1016/0304-3991(82)90187-5 (1982).

10   Dubochet, J., Lepault, J., Freeman, R., Berriman, J. A. & Homo, J. C. Electron microscopy of frozen water and aqueous solutions. *Journal of Microscopy* **128**, 219-237, doi:10.1111/j.1365-2818.1982.tb04625.x (1982).

11   Lepault, J., Booy, F. P. & Dubochet, J. Electron microscopy of frozen biological suspensions. *J Microsc* **129**, 89-102, doi:10.1111/j.1365-2818.1983.tb04163.x (1983).

12   Dubochet, J. *et al.* Cryo-electron microscopy of vitrified specimens. *Q Rev Biophys* **21**, 129-228, doi:10.1017/s0033583500004297 (1988).

13   Frederik, P. M., Burger, K. N. J., Stuart, M. C. A. & Verkleij, A. J. Lipid polymorphism as observed by cryo-electron microscopy. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **1062**, 133-141, doi:10.1016/0005-2736(91)90384-K (1991).

14   van Heel, M. & Frank, J. Use of multivariates statistics in analysing the images of biological macromolecules. *Ultramicroscopy* **6**, 187-194, doi:10.1016/s0304-3991(81)80197-0 (1981).

15   Frank, J. & van Heel, M. Correspondence analysis of aligned images of biological particles. *J Mol Biol* **161**, 134-137, doi:10.1016/0022-2836(82)90282-0 (1982).

16   Radermacher, M., Wagenknecht, T., Verschoor, A. & Frank, J. Three-dimensional reconstruction from a single-exposure, random conical tilt series applied to the 50S ribosomal subunit of Escherichia coli. *J Microsc* **146**, 113-136, doi:10.1111/j.1365-2818.1987.tb01333.x (1987).

17   Radermacher, M., Wagenknecht, T., Verschoor, A. & Frank, J. A new 3-D reconstruction scheme applied to the 50S ribosomal subunit of E. coli. *J Microsc* **141**, RP1-2, doi:10.1111/j.1365-2818.1986.tb02693.x (1986).

18   Henderson, R. *et al.* Model for the structure of bacteriorhodopsin based on high-resolution electron cryo-microscopy. *J Mol Biol* **213**, 899-929, doi:10.1016/S0022-2836(05)80271-2 (1990).

19 Kuhlbrandt, W., Wang, D. N. & Fujiyoshi, Y. Atomic model of plant light-harvesting complex by electron crystallography. *Nature* **367**, 614-621, doi:10.1038/367614a0 (1994).

20 Nogales, E., Wolf, S. G. & Downing, K. H. Structure of the alpha beta tubulin dimer by electron crystallography. *Nature* **391**, 199-203, doi:10.1038/34465 (1998).

21 Murata, K. *et al.* Structural determinants of water permeation through aquaporin-1. *Nature* **407**, 599-605, doi:10.1038/35036519 (2000).

22 Liao, M., Cao, E., Julius, D. & Cheng, Y. Structure of the TRPV1 ion channel determined by electron cryo-microscopy. *Nature* **504**, 107-112, doi:10.1038/nature12822 (2013).

23 Cao, E., Liao, M., Cheng, Y. & Julius, D. TRPV1 structures in distinct conformations reveal activation mechanisms. *Nature* **504**, 113-118, doi:10.1038/nature12823 (2013).

24 Li, X. *et al.* Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. *Nat Methods* **10**, 584-590, doi:10.1038/nmeth.2472 (2013).

25 Bai, X. C., Fernandez, I. S., McMullan, G. & Scheres, S. H. Ribosome structures to near-atomic resolution from thirty thousand cryo-EM particles. *Elife* **2**, e00461, doi:10.7554/eLife.00461 (2013).

26 Kuhlbrandt, W. Biochemistry. The resolution revolution. *Science* **343**, 1443-1444, doi:10.1126/science.1251652 (2014).

27 Faruqi, A. R., Henderson, R., Pryddetch, M., Allport, P. & Evans, A. Direct single electron detection with a CMOS detector for electron microscopy. *Nucl Instrum Meth A* **546**, 170-175, doi:10.1016/j.nima.2005.03.023 (2005).

28 Faruqi, A. R., Cattermole, D. M. & Raeburn, C. Direct electron detection methods in electron microscopy. *Nucl Instrum Meth A* **513**, 317-321, doi:10.1016/j.nima.2003.08.055 (2003).

29 Xuong, N. H. *et al.* First use of a high sensitivity active pixel sensor array as a detector for electron microscopy. *Sensors and Camera Systems for Scientific, Industrial, and Digital Photography Applications V* **5301**, 242-249, doi:Doi 10.1117/12.526021 (2004).

30 McMullan, G., Chen, S., Henderson, R. & Faruqi, A. R. Detective quantum efficiency of electron area detectors in electron microscopy. *Ultramicroscopy* **109**, 1126-1143, doi:10.1016/j.ultramic.2009.04.002 (2009).

31 McMullan, G., Faruqi, A. R. & Henderson, R. in *Methods in Enzymology* Vol. 579   1-17 (2016).

32 Thorne, R. E. Hypothesis for a mechanism of beam-induced motion in cryo-electron microscopy. *IUCrJ* **7**, 416-421, doi:10.1107/S2052252520002560 (2020).

33 Naydenova, K., Jia, P. & Russo, C. J. Cryo-EM with sub-1 A specimen movement. *Science* **370**, 223-226, doi:10.1126/science.abb7927 (2020).

34 Glaeser, R. M. Specimen Behavior in the Electron Beam. *Methods Enzymol* **579**, 19-50, doi:10.1016/bs.mie.2016.04.010 (2016).

35 Russo, C. J. & Passmore, L. A. Electron microscopy: Ultrastable gold substrates for electron cryomicroscopy. *Science* **346**, 1377-1380, doi:10.1126/science.1259530 (2014).

36 Zheng, L. *et al.* Uniform thin ice on ultraflat graphene for high-resolution cryo-EM. *Nat Methods* **20**, 123-130, doi:10.1038/s41592-022-01693-y (2023).

37 Fan, X. *et al.* Single particle cryo-EM reconstruction of 52 kDa streptavidin at 3.2 Angstrom resolution. *Nat Commun* **10**, 2386, doi:10.1038/s41467-019-10368-w (2019).

38 Naydenova, K., Peet, M. J. & Russo, C. J. Multifunctional graphene supports for electron cryomicroscopy. *Proc Natl Acad Sci U S A* **116**, 11718-11724, doi:10.1073/pnas.1904766116 (2019).

39   Han, Y. *et al.* High-yield monolayer graphene grids for near-atomic resolution cryoelectron microscopy. *Proc Natl Acad Sci U S A* **117**, 1009-1014, doi:10.1073/pnas.1919114117 (2020).

40   Wang, F. *et al.* General and robust covalently linked graphene oxide affinity grids for high-resolution cryo-EM. *Proc Natl Acad Sci U S A* **117**, 24269-24273, doi:10.1073/pnas.2009707117 (2020).

41   Fan, H. & Sun, F. Developing Graphene Grids for Cryoelectron Microscopy. *Front Mol Biosci* **9**, 937253, doi:10.3389/fmolb.2022.937253 (2022).

42   Scheres, S. H. A Bayesian view on cryo-EM structure determination. *J Mol Biol* **415**, 406-418, doi:10.1016/j.jmb.2011.11.010 (2012).

43   Zivanov, J. *et al.* New tools for automated high-resolution cryo-EM structure determination in RELION-3. *Elife* **7**, doi:10.7554/eLife.42166 (2018).

44   Zivanov, J., Nakane, T. & Scheres, S. H. W. Estimation of high-order aberrations and anisotropic magnification from cryo-EM data sets in RELION-3.1. *IUCrJ* **7**, 253-267, doi:10.1107/S2052252520000081 (2020).

45   Zivanov, J., Nakane, T. & Scheres, S. H. W. A Bayesian approach to beam-induced motion correction in cryo-EM single-particle analysis. *IUCrJ* **6**, 5-17, doi:10.1107/S205225251801463X (2019).

46   Russo, C. J. & Henderson, R. Ewald sphere correction using a single side-band image processing algorithm. *Ultramicroscopy* **187**, 26-33, doi:10.1016/j.ultramic.2017.11.001 (2018).

47   Danev, R. & Baumeister, W. Cryo-EM single particle analysis with the Volta phase plate. *Elife* **5**, doi:10.7554/eLife.13046 (2016).

48   Henderson, R. The potential and limitations of neutrons, electrons and X-rays for atomic resolution microscopy of unstained biological molecules. *Q Rev Biophys* **28**, 171-193, doi:10.1017/s003358350000305x (1995).

49   Glaeser, R. M. Review: electron crystallography: present excitement, a nod to the past, anticipating the future. *J Struct Biol* **128**, 3-14, doi:10.1006/jsbi.1999.4172 (1999).

50   Tiessen, A., Perez-Rodriguez, P. & Delaye-Arredondo, L. J. 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* **5**, 85, doi:10.1186/1756-0500-5-85 (2012).

51   Fan, H. *et al.* A cryo-electron microscopy support film formed by 2D crystals of hydrophobin HFBI. *Nat Commun* **12**, 7257, doi:10.1038/s41467-021-27596-8 (2021).

52   Scheres, S. H. Beam-induced motion correction for sub-megadalton cryo-EM particles. *Elife* **3**, e03665, doi:10.7554/eLife.03665 (2014).

53   Bartesaghi, A. *et al.* Atomic Resolution Cryo-EM Structure of beta-Galactosidase. *Structure* **26**, 848-856 e843, doi:10.1016/j.str.2018.04.004 (2018).

54   Hamaguchi, T. *et al.* A new cryo-EM system for single particle analysis. *J Struct Biol* **207**, 40-48, doi:10.1016/j.jsb.2019.04.011 (2019).

55   Tan, Y. Z. *et al.* Sub-2 A Ewald curvature corrected structure of an AAV2 capsid variant. *Nat Commun* **9**, 3628, doi:10.1038/s41467-018-06076-6 (2018).

56   Ravelli, R. B. G. *et al.* Cryo-EM structures from sub-nl volumes using pin-printing and jet vitrification. *Nat Commun* **11**, 2563, doi:10.1038/s41467-020-16392-5 (2020).

57   Jain, T., Sheehan, P., Crum, J., Carragher, B. & Potter, C. S. Spotiton: a prototype for an integrated inkjet dispense and vitrification system for cryo-TEM. *J Struct Biol* **179**, 68-75, doi:10.1016/j.jsb.2012.04.020 (2012).

58   Carragher, B. *et al.* Current outcomes when optimizing 'standard' sample preparation for single-particle cryo-EM. *J Microsc* **276**, 39-45, doi:10.1111/jmi.12834 (2019).

59   Zhang, Y., Tammaro, R., Peters, P. J. & Ravelli, R. B. G. Could Egg White Lysozyme be Solved by Single Particle Cryo-EM? *J Chem Inf Model* **60**, 2605-2613, doi:10.1021/acs.jcim.9b01176 (2020).

60   Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat Methods* **14**, 290-296, doi:10.1038/nmeth.4169 (2017).

61   Andrews, S. C. The Ferritin-like superfamily: Evolution of the biological iron storeman from a rubrerythrin-like ancestor. *Bba-Gen Subjects* **1800**, 691-705, doi:10.1016/j.bbagen.2010.05.010 (2010).

62   Arosio, P., Ingrassia, R. & Cavadini, P. Ferritins: a family of molecules for iron storage, antioxidation and more. *Biochim Biophys Acta* **1790**, 589-599, doi:10.1016/j.bbagen.2008.09.004 (2009).

63   Expert, D. in *Molecular Aspects of Iron Metabolism in Pathogenic and Symbiotic Plant-Microbe Associations* (eds D. Expert & M. R. O'Brian) (Springer, 2012).

64   Ratledge, C. & Dover, L. G. Iron metabolism in pathogenic bacteria. *Annu Rev Microbiol* **54**, 881-941, doi:10.1146/annurev.micro.54.1.881 (2000).

65   Contreras, H. *et al.* Characterization of a Mycobacterium tuberculosis Nanocompartment and Its Potential Cargo Proteins. *Journal of Biological Chemistry* **289**, 18279-18289, doi:10.1074/jbc.M114.570119 (2014).

66   Khare, G. *et al.* Ferritin structure from Mycobacterium tuberculosis: comparative study with homologues identifies extended C-terminus involved in ferroxidase activity. *PLoS One* **6**, e18570, doi:10.1371/journal.pone.0018570 (2011).

67   Parida, A., Mohanty, A., Kansara, B. T. & Behera, R. K. Impact of Phosphate on Iron Mineralization and Mobilization in Nonheme Bacterioferritin B from Mycobacterium tuberculosis. *Inorg Chem* **59**, 629-641, doi:10.1021/acs.inorgchem.9b02894 (2020).

68   Scheres, S. H. RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J Struct Biol* **180**, 519-530, doi:10.1016/j.jsb.2012.09.006 (2012).

69   Zheng, S. Q. *et al.* MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat Methods* **14**, 331-332, doi:10.1038/nmeth.4193 (2017).

70   Zhang, K. Gctf: Real-time CTF determination and correction. *J Struct Biol* **193**, 1-12, doi:10.1016/j.jsb.2015.11.003 (2016).

71   Scheres, S. H. & Chen, S. Prevention of overfitting in cryo-EM structure determination. *Nat Methods* **9**, 853-854, doi:10.1038/nmeth.2115 (2012).

72   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, doi:10.1016/j.jmb.2003.07.013 (2003).

73   Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr Sect D Biological Crystallogr* **60**, 2126-2132, doi:10.1107/S0907444904019158 (2004).

74   Kaur, S. *et al.* Local computational methods to improve the interpretability and analysis of cryo-EM maps. *Nat Commun* **12**, 1240, doi:10.1038/s41467-021-21509-5 (2021).

75   Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr Sect D Biological Crystallogr* **66**, 486-501, doi:10.1107/S0907444910007493 (2010).

76   Afonine, P. V. *et al.* New tools for the analysis and validation of cryo-EM maps and atomic models. *Acta Crystallogr Sect D* **74**, 814-840, doi:10.1107/S2059798318009324 (2018).

77   Williams, C. J. *et al.* MolProbity: More and better reference data for improved all-atom structure validation. *Protein Sci* **27**, 293-315, doi:10.1002/pro.3330 (2018).

78   Rui, H., Rivera, M. & Im, W. Protein dynamics and ion traffic in bacterioferritin. *Biochemistry* **51**, 9900-9910, doi:10.1021/bi3013388 (2012).

79   Jobichen, C. *et al.* Cryo-EM Structure of Bacterioferritin Nanocages Provides Insight into the Bio-mineralization of Ferritins. *bioRxiv*, 2021.2002.2004.429857-422021.429802.429804.429857, doi:10.1101/2021.02.04.429857 (2021).

80   Putignano, V., Rosato, A., Banci, L. & Andreini, C. MetalPDB in 2018: a database of metal sites in biological macromolecular structures. *Nucleic Acids Res* **46**, D459-D464, doi:10.1093/nar/gkx989 (2018).

81   Teilum, K., Olsen, J. G. & Kragelund, B. B. Functional aspects of protein flexibility. *Cell Mol Life Sci* **66**, 2231-2247, doi:10.1007/s00018-009-0014-6 (2009).

82   Wu, M., Lander, G. C. & Herzik, M. A., Jr. Sub-2 Angstrom resolution structure determination using single-particle cryo-EM at 200 keV. *J Struct Biol X* **4**, 100020, doi:10.1016/j.yjsbx.2020.100020 (2020).

83   Herzik, M. A., Jr., Wu, M. & Lander, G. C. High-resolution structure determination of sub-100 kDa complexes using conventional cryo-EM. *Nat Commun* **10**, 1032, doi:10.1038/s41467-019-08991-8 (2019).

84   Liu, Y., Huynh, D. T. & Yeates, T. O. A 3.8 A resolution cryo-EM structure of a small protein bound to an imaging scaffold. *Nat Commun* **10**, 1864, doi:10.1038/s41467-019-09836-0 (2019).

85   D'Imprima, E. *et al.* Protein denaturation at the air-water interface and how to prevent it. *Elife* **8**, doi:10.7554/eLife.42747 (2019).

86   Vulovic, M. *et al.* Image formation modeling in cryo-electron microscopy. *J Struct Biol* **183**, 19-32, doi:10.1016/j.jsb.2013.05.008 (2013).

87   Danev, R., Buijsse, 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* **111**, 15635-15640, doi:10.1073/pnas.1418377111 (2014).

88   Schwartz, O. *et al.* Laser phase plate for transmission electron microscopy. *Nat Methods* **16**, 1016-1020, doi:10.1038/s41592-019-0552-2 (2019).

89   Muller, H. *et al.* Design of an electron microscope phase plate using a focused continuous-wave laser. *New J Phys* **12**, 073011-073011, doi:10.1088/1367-2630/12/7/073011 (2010).

90   Tavabi, A. H. *et al.* Tunable Ampere phase plate for low dose imaging of biomolecular complexes. *Sci Rep* **8**, 5592, doi:10.1038/s41598-018-23100-3 (2018).

91   Peet, M. J., Henderson, R. & Russo, C. J. The energy dependence of contrast and damage in electron cryomicroscopy of biological molecules. *Ultramicroscopy* **203**, 125-131, doi:10.1016/j.ultramic.2019.02.007 (2019).

92   Naydenova, K. *et al.* CryoEM at 100 keV: a demonstration and prospects. *IUCrJ* **6**, 1086-1098, doi:10.1107/S2052252519012612 (2019).

93   Rullgard, H., Ofverstedt, L. G., Masich, S., Daneholt, B. & Oktem, O. Simulation of transmission electron microscope images of biological specimens. *J Microsc* **243**, 234-256, doi:10.1111/j.1365-2818.2011.03497.x (2011).

94   Weik, M. *et al.* Specific chemical and structural damage to proteins produced by synchrotron radiation. *Proc Natl Acad Sci U S A* **97**, 623-628, doi:10.1073/pnas.97.2.623 (2000).

95   Ravelli, R. B. & McSweeney, S. M. The 'fingerprint' that X-rays can leave on structures. *Structure* **8**, 315-328, doi:10.1016/s0969-2126(00)00109-x (2000).

96   Patwardhan, A. & Lawson, C. L. in *Methods in Enzymology* Vol. 579   393-412 (Academic Press Inc., 2016).

97   Scheres, S. H. W. in *Methods in Enzymology* Vol. 579   125-157 (Academic Press Inc., 2016).

98   Liu, Q., Li, Z. & Li, J. Use B-factor related features for accurate classification between protein binding interfaces and crystal packing contacts. *BMC Bioinformatics* **15 Suppl 16**, S3, doi:10.1186/1471-2105-15-S16-S3 (2014).

99   Carugo, O. How large B-factors can be in protein crystal structures. *BMC Bioinformatics* **19**, 61, doi:10.1186/s12859-018-2083-8 (2018).

100   Kirkland, E. J. *Advanced Computing in Electron Microscopy.* (2010).

101   Vulovic, M., Voortman, L. M., van Vliet, L. J. & Rieger, B. When to use the projection assumption and the weak-phase object approximation in phase contrast cryo-EM. *Ultramicroscopy* **136**, 61-66, doi:10.1016/j.ultramic.2013.08.002 (2014).

102   Williams, D. B. & Carter, C. B. Transmission Electron Microscopy A Textbook for Materials Science. doi:10.1007/978-0-387-76501-3 (2009).

103   Bucker, R. *et al.* Serial protein crystallography in an electron microscope. *Nat Commun* **11**, 996, doi:10.1038/s41467-020-14793-0 (2020).

104   R. A. Crowther, D. J. D. & Klug, A. The Reconstruction of a Three-Dimensional Structure from Projections and its Application to Electron Microscopy. *Proceedings of the Royal Society of London* **317**, 319-340 (1970).

105   Kuijper, M. *et al.* FEI's direct electron detector developments: Embarking on a revolution in cryo-TEM. *J Struct Biol* **192**, 179-187, doi:10.1016/j.jsb.2015.09.014 (2015).

106   Hempstead, P. D. *et al.* Comparison of the three-dimensional structures of recombinant human H and horse L ferritins at high resolution. *J Mol Biol* **268**, 424-448, doi:10.1006/jmbi.1997.0970 (1997).

107   Weiss, M. S., Palm, G. J. & Hilgenfeld, R. Crystallization, structure solution and refinement of hen egg-white lysozyme at pH 8.0 in the presence of MPD. *Acta Crystallogr Sect D Biological Crystallogr* **56**, 952-958, doi:10.1107/s0907444900006685 (2000).

108   Danev, R., Yanagisawa, H. & Kikkawa, M. Cryo-Electron Microscopy Methodology: Current Aspects and Future Directions. *Trends Biochem Sci* **44**, 837-848, doi:10.1016/j.tibs.2019.04.008 (2019).

109   Zhang, Y., Tammaro, R., Peters, P. J. & Ravelli, R. B. G. M4I-nanoscopy/InSilicoTEM: InSilicoTEM_v2.0.0. doi:10.5281/ZENODO.3693834 (2020).

110   Chen, S. *et al.* High-resolution noise substitution to measure overfitting and validate resolution in 3D structure determination by single particle electron cryomicroscopy. *Ultramicroscopy* **135**, 24-35, doi:10.1016/j.ultramic.2013.06.004 (2013).

111   Kuhlbrandt, W. The resolution revolution. *Science* **343**, 1443-1444, doi:10.1126/science.1251652 (2014).

112   Zhang, K. *et al.* Cryo-EM and antisense targeting of the 28-kDa frameshift stimulation element from the SARS-CoV-2 RNA genome. *Nat Struct Mol Biol* **28**, 747-754, doi:10.1038/s41594-021-00653-y (2021).

113   Konings, S. *et al.* Advances in Single Particle Analysis Data Acquisition. *Microscopy and Microanalysis* **25**, 1012-1013, doi:10.1017/s1431927619005798 (2019).

114   Grubb, D. T. Radiation damage and electron microscopy of organic polymers. *Journal of Materials Science* **9**, 1715-1736, doi:10.1007/BF00540772/METRICS (1974).

115   Berriman, J. A. & Rosenthal, P. B. Paraxial charge compensator for electron cryomicroscopy. *Ultramicroscopy* **116**, 106-114, doi:10.1016/j.ultramic.2012.03.006 (2012).

116   Egerton, R. F. Choice of operating voltage for a transmission electron microscope. *Ultramicroscopy* **145**, 85-93, doi:10.1016/j.ultramic.2013.10.019 (2014).

117   Wieferig, J. P., Mills, D. J. & Kuhlbrandt, W. Devitrification reduces beam-induced movement in cryo-EM. *IUCrJ* **8**, 186-194, doi:10.1107/S2052252520016243 (2021).

118   Wu, C., Shi, H., Zhu, D., Fan, K. & Zhang, X. Low-cooling-rate freezing in biomolecular cryo-electron microscopy for recovery of initial frames. *QRB Discovery* **2**, 213-221, doi:10.1017/qrd.2021.8 (2021).

119   Henderson, R. & Glaeser, R. M. Quantitative analysis of image contrast in electron micrographs of beam-sensitive crystals. *Ultramicroscopy* **16**, 139-150, doi:10.1016/0304-3991(85)90069-5 (1985).

120   Henderson, R. Image contrast in high-resolution electron microscopy of biological macromolecules: TMV in ice. *Ultramicroscopy* **46**, 1-18, doi:10.1016/0304-3991(92)90003-3 (1992).

121   Vinothkumar, K. R. & Henderson, R. Single particle electron cryomicroscopy: trends, issues and future perspective. *Q Rev Biophys* **49**, e13, doi:10.1017/S0033583516000068 (2016).

122   Russo, C. J. & Henderson, R. Charge accumulation in electron cryomicroscopy. *Ultramicroscopy* **187**, 43-49, doi:10.1016/j.ultramic.2018.01.009 (2018).

123   Schreiber, M. T., Maigne, A., Beleggia, M., Shibata, S. & Wolf, M. Temporal dynamics of charge buildup in cryo-electron microscopy. *J Struct Biol X* **7**, 100081, doi:10.1016/j.yjsbx.2022.100081 (2023).

124   Cazaux, J. Correlations between Ionization Radiation-Damage and Charging Effects in Transmission Electron-Microscopy. *Ultramicroscopy* **60**, 411-425, doi:Doi 10.1016/0304-3991(95)00077-1 (1995).

125   Curtis, G. H. & Ferrier, R. P. The electric charging of electron-microscope specimens. *Journal of Physics D: Applied Physics* **2**, 1035-1040, doi:10.1088/0022-3727/2/7/312 (1969).

126   Brink, J., Sherman, M. B., Berriman, J. & Chiu, W. Evaluation of charging on macromolecules in electron cryomicroscopy. *Ultramicroscopy* **72**, 41-52, doi:10.1016/s0304-3991(97)00126-5 (1998).

127   Egerton, R. F. Radiation damage to organic and inorganic specimens in the TEM. *Micron* **119**, 72-87, doi:10.1016/j.micron.2019.01.005 (2019).

128   Downing, K. H., McCartney, M. R. & Glaeser, R. M. Experimental characterization and mitigation of specimen charging on thin films with one conducting layer. *Microsc Microanal Official J Microsc Soc Am Microbeam Analysis Soc Microsc Soc Can* **10**, 783-789, doi:10.1017/s143192760404067x (2004).

129   Mahl, H. & Weitsch, W. Nachweis von fluktuierenden Ladungen in d�nnen Lackfilmen bei Elektronendurchstrahlung. *Die Naturwissenschaften* **46**, 487-488, doi:10.1007/bf00626730 (1959).

130   Dove, D. B. Image Contrasts in Thin Carbon Films Observed by Shadow Electron Microscopy. *Journal of Applied Physics* **35**, 1652-1653, doi:10.1063/1.1713709 (1964).

131   Russo, C. J. & Henderson, R. Microscopic charge fluctuations cause minimal contrast loss in cryoEM. *Ultramicroscopy* **187**, 56-63, doi:10.1016/j.ultramic.2018.01.011 (2018).

132   Wang, L. *et al.* Dynamics of the charging-induced imaging instability in transmission electron microscopy. *Nanoscale Adv* **3**, 3035-3040, doi:10.1039/d1na00140j (2021).

133   Downing, K. H. & Glaeser, R. M. Improvement in high resolution image quality of radiation-sensitive specimens achieved with reduced spot size of the electron beam. *Ultramicroscopy* **20**, 269-278, doi:10.1016/0304-3991(86)90191-9 (1986).

134   Bullough, P. & Henderson, R. Use of spot-scan procedure for recording low-dose micrographs of beam-sensitive specimens. *Ultramicroscopy* **21**, 223-230, doi:10.1016/0304-3991(87)90147-1 (1987).

135   Downing, K. H. Spot-scan imaging in transmission electron microscopy. *Science* **251**, 53-59, doi:10.1126/science.1846047 (1991).

136   Geim, A. K. Graphene: status and prospects. *Science* **324**, 1530-1534, doi:10.1126/science.1158877 (2009).

137   Pantelic, R. S., Meyer, J. C., Kaiser, U., Baumeister, W. & Plitzko, J. M. Graphene oxide: a substrate for optimizing preparations of frozen-hydrated samples. *J Struct Biol* **170**, 152-156, doi:10.1016/j.jsb.2009.12.020 (2010).

138   Pantelic, R. S. *et al.* Graphene: Substrate preparation and introduction. *J Struct Biol* **174**, 234-238, doi:10.1016/j.jsb.2010.10.002 (2011).

139   Russo, C. J. & Passmore, L. A. Controlling protein adsorption on graphene for cryo-EM using low-energy hydrogen plasmas. *Nat Methods* **11**, 649-652, doi:10.1038/nmeth.2931 (2014).

140   Nickl, P. *et al.* A New Support Film for Cryo Electron Microscopy Protein Structure Analysis Based on Covalently Functionalized Graphene. *Small* **19**, e2205932, doi:10.1002/smll.202205932 (2023).

141   Deursen, P. M. G. *et al.* Graphene Liquid Cells Assembled through Loop‐Assisted Transfer Method and Located with Correlated Light‐Electron Microscopy. *Advanced Functional Materials* **30**, 1904468-1904468, doi:10.1002/adfm.201904468 (2020).

142   Mohanty, N., Fahrenholtz, M., Nagaraja, A., Boyle, D. & Berry, V. Impermeable graphenic encasement of bacteria. *Nano Lett* **11**, 1270-1275, doi:10.1021/nl104292k (2011).

143   Fujita, J. *et al.* Epoxidized graphene grid for highly efficient high-resolution cryoEM structural analysis. *Sci Rep* **13**, 2279, doi:10.1038/s41598-023-29396-0 (2023).

144   Xu, J., Cui, X. Y., Liu, N., Chen, Y. A. & Wang, H. W. Structural engineering of graphene for high-resolution cryo-electron microscopy. *Smartmat* **2**, 202-212, doi:10.1002/smm2.1045 (2021).

145   Gijsbers, A., Zhang, Y., Gao, Y., Peters, P. J. & Ravelli, R. B. G. Mycobacterium tuberculosis ferritin: a suitable workhorse protein for cryo-EM development. *Acta Crystallogr Sect D* **77**, 1077-1083, doi:10.1107/S2059798321007233 (2021).

146   van Schayck, J. P., Zhang, Y., Knoops, K., Peters, P. J. & Ravelli, R. B. G. Integration of an Event-driven Timepix3 Hybrid Pixel Detector into a Cryo-EM Workflow. *Microscopy and Microanalysis* **29**, 352-363, doi:10.1093/micmic/ozac009 (2023).

147   van Schayck, J. P. *et al.* Sub-pixel electron detection using a convolutional neural network. *Ultramicroscopy* **218**, 113091, doi:10.1016/j.ultramic.2020.113091 (2020).

148   Poikela, T. *et al.* Timepix3: a 65K channel hybrid pixel readout chip with simultaneous ToA/ToT and sparse readout. *Journal of Instrumentation* **9**, doi:10.1088/1748-0221/9/05/C05013 (2014).

149   Ripstein, Z. A. & Rubinstein, J. L. Processing of Cryo-EM Movie Data. *Methods Enzymol* **579**, 103-124, doi:10.1016/bs.mie.2016.04.009 (2016).

150   Pedrazo-Tardajos, A. & Bals, S. Graphene layer transfer method. (2022).

151   Algara-Siller, G., Kurasch, S., Sedighi, M., Lehtinen, O. & Kaiser, U. The pristine atomic structure of MoS2 monolayer protected from electron radiation damage by graphene. *Appl Phys Lett* **103**, doi:10.1063/1.4830036 (2013).

152   Lazic, I. *et al.* Single-particle cryo-EM structures from iDPC-STEM at near-atomic resolution. *Nat Methods* **19**, 1126-1136, doi:10.1038/s41592-022-01586-0 (2022).

153   Zhou, L. *et al.* Low-dose phase retrieval of biological specimens using cryo-electron ptychography. *Nat Commun* **11**, 2773, doi:10.1038/s41467-020-16391-6 (2020).

154   Heijne, E. H. M. Semiconductor micropattern pixel detectors: a review of the beginnings. *Nucl Instrum Meth A* **465**, 1-26, doi:Doi 10.1016/S0168-9002(01)00340-0 (2001).

155   Zhang, Y. *et al.* Single-particle cryo-EM: alternative schemes to improve dose efficiency. *J Synchrotron Radiat* **28**, 1343-1356, doi:10.1107/S1600577521007931 (2021).

152

156   Dainty, J. C. & Shaw, R. *Image Science: Principles, Analysis and Evaluation of Photographic-type Imaging Processes*. (Academic Press, 1974).

157   De Ruijter, W. J. Imaging properties and applications of slow-scan charge-coupled device cameras suitable for electron microscopy. *Micron* **26**, 247-275, doi:10.1016/0968-4328(95)00054-8 (1995).

158   McMullan, G., Turchetta, R. & Faruqi, A. R. Single event imaging for electron microscopy using MAPS detectors. *Journal of Instrumentation* **6**, C04001-C04001, doi:10.1088/1748-0221/6/04/C04001 (2011).

159   Ruskin, R. S., Yu, Z. & Grigorieff, N. Quantitative characterization of electron detectors for transmission electron microscopy. *J Struct Biol* **184**, 385-393, doi:10.1016/j.jsb.2013.10.016 (2013).

160   Paton, K. A. *et al.* Quantifying the performance of a hybrid pixel detector with GaAs:Cr sensor for transmission electron microscopy. *Ultramicroscopy* **227**, 113298, doi:10.1016/j.ultramic.2021.113298 (2021).

161   Bammes, B. E., Rochat, R. H., Jakana, J., Chen, D. H. & Chiu, W. Direct electron detection yields cryo-EM reconstructions at resolutions beyond 3/4 Nyquist frequency. *J Struct Biol* **177**, 589-601, doi:10.1016/j.jsb.2012.01.008 (2012).

162   McMullan, G. *et al.* Experimental observation of the improvement in MTF from backthinning a CMOS direct electron detector. *Ultramicroscopy* **109**, 1144-1147, doi:10.1016/j.ultramic.2009.05.005 (2009).

163   Battaglia, M., Contarato, D., Denes, P. & Giubilato, P. Cluster imaging with a direct detection CMOS pixel sensor in Transmission Electron Microscopy. *Nucl Instrum Meth A* **608**, 363-365, doi:10.1016/j.nima.2009.07.017 (2009).

164   McMullan, G., Clark, A. T., Turchetta, R. & Faruqi, A. R. Enhanced imaging in low dose electron microscopy using electron counting. *Ultramicroscopy* **109**, 1411-1416, doi:10.1016/j.ultramic.2009.07.004 (2009).

165   Battaglia, M. *et al.* Characterisation of a CMOS active pixel sensor for use in the TEAM microscope. *Nucl Instrum Meth A* **622**, 669-677, doi:10.1016/j.nima.2010.07.066 (2010).

166   Ballabriga, R., Campbell, M. & Llopart, X. Asic developments for radiation imaging applications: The medipix and timepix family. *Nucl Instrum Meth A* **878**, 10-23, doi:10.1016/j.nima.2017.07.029 (2018).

167   McMullan, G. *et al.* Electron imaging with Medipix2 hybrid pixel detector. *Ultramicroscopy* **107**, 401-413, doi:10.1016/j.ultramic.2006.10.005 (2007).

168   Ballabriga, R. *et al.* Medipix3: A 64 k pixel detector readout chip working in single photon counting mode with improved spectrometric performance. *Nucl Instrum Meth A* **633**, S15-S18, doi:10.1016/j.nima.2010.06.108 (2011).

169   Mir, J. A. *et al.* Characterisation of the Medipix3 detector for 60 and 80keV electrons. *Ultramicroscopy* **182**, 44-53, doi:10.1016/j.ultramic.2017.06.010 (2017).

170   Wang, B., Zou, X. & Smeets, S. Automated serial rotation electron diffraction combined with cluster analysis: an efficient multi-crystal workflow for structure determination. *IUCrJ* **6**, 854-867, doi:10.1107/S2052252519007681 (2019).

171   Guo, H. *et al.* Electron-event representation data enable efficient cryoEM file storage with full preservation of spatial and temporal resolution. *IUCrJ* **7**, 860-869, doi:10.1107/S205225252000929X (2020).

172   Visser, J. *et al.* SPIDR: a read-out system for Medipix3 & Timepix3. *Journal of Instrumentation* **10**, C12028-C12028, doi:10.1088/1748-0221/10/12/C12028 (2015).

173   Mastronarde, D. N. SerialEM: A Program for Automated Tilt Series Acquisition on Tecnai Microscopes Using Prediction of Specimen Position. *Microscopy and Microanalysis* **9**, 1182-1183, doi:10.1017/s1431927603445911 (2003).

174   van Schayck, J. P. & Ravelli, R. B. G. *M4i-nanoscopy/tpx3hitparser:v2.2.0*, 2022).

175   van Schayck, J. P. & Ravelli, R. B. G. *M4i-nanoscopy/tpx3event-viewer: v2.0.0*, 2022).

176   Vulovic, M., Rieger, B., van Vliet, L. J., Koster, A. J. & Ravelli, R. B. A toolkit for the characterization of CCD cameras for transmission electron microscopy. *Acta Crystallogr Sect D Biological Crystallogr* **66**, 97-109, doi:10.1107/S0907444909031205 (2010).

177   Joosten, R. P., Joosten, K., Cohen, S. X., Vriend, G. & Perrakis, A. Automatic rebuilding and optimization of crystallographic structures in the Protein Data Bank. *Bioinformatics* **27**, 3392-3398, doi:10.1093/bioinformatics/btr590 (2011).

178   Wilkinson, M. D. *et al.* The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* **3**, 160018, doi:10.1038/sdata.2016.18 (2016).

179   Sarkans, U. *et al.* REMBI: Recommended Metadata for Biological Images-enabling reuse of microscopy data in biology. *Nat Methods* **18**, 1418-1422, doi:10.1038/s41592-021-01166-8 (2021).

180   Rabbani, M., Shaw, R. & Van Metter, R. Detective quantum efficiency of imaging systems with amplifying and scattering mechanisms. *J Opt Soc Am* **4**, 895-901, doi:10.1364/josaa.4.000895 (1987).

181   Cunningham, I. A., Westmore, M. S. & Fenster, A. in *Medical Imaging 1995: Physics of Medical Imaging.* May 1995 edn (eds Richard L. Van Metter & Jacob Beutel) 143-151.

182   Monnin, P., Bosmans, H., Verdun, F. R. & Marshall, N. W. A comprehensive model for quantum noise characterization in digital mammography. *Phys Med Biol* **61**, 2083-2108, doi:10.1088/0031-9155/61/5/2083 (2016).

183   Zhu, D., Shi, H., Wu, C. & Zhang, X. An electron counting algorithm improves imaging of proteins with low-acceleration-voltage cryo-electron microscope. *Commun Biol* **5**, 321, doi:10.1038/s42003-022-03284-1 (2022).

184   van Schayck, J. P., Zhang, Y. & Ravelli, R. B. G. *M4i-nanoscopy/mtf-nps-dqe: v1.0.0*, 2022).

185   Clausen, A. *et al.* LiberTEM: Software platform for scalable multidimensional data processing in transmission electron microscopy. *J Open Source Softw* **5**, 2006, doi:10.21105/joss.02006 (2020).

186   Nederlof, I., van Genderen, E., Li, Y. W. & Abrahams, J. P. A Medipix quantum area detector allows rotation electron diffraction data collection from submicrometre three-dimensional protein crystals. *Acta Crystallogr Sect D Biological Crystallogr* **69**, 1223-1230, doi:10.1107/S0907444913009700 (2013).

187   Jannis, D. *et al.* Event driven 4D STEM acquisition with a Timepix3 detector: Microsecond dwell time and faster scans for high precision and low dose applications. *Ultramicroscopy* **233**, 113423, doi:10.1016/j.ultramic.2021.113423 (2021).

188   Rodenburg, J. M. Ptychography and related diffractive imaging methods. *Adv Imag Elect Phys* **150**, 87-184, doi:10.1016/S1076-5670(07)00003-1 (2008).

189   Allars, F. *et al.* Efficient large field of view electron phase imaging using near-field electron ptychography with a diffuser. *Ultramicroscopy* **231**, 113257, doi:10.1016/j.ultramic.2021.113257 (2021).

190   Llopart, X. *et al.* Timepix4, a large area pixel detector readout chip which can be tiled on 4 sides providing sub-200 ps timestamp binning. *Journal of Instrumentation* **17**, C01044-C01044, doi:10.1088/1748-0221/17/01/C01044 (2022).

191   Karuppasamy, M., Karimi Nejadasl, F., Vulovic, M., Koster, A. J. & Ravelli, R. B. Radiation damage in single-particle cryo-electron microscopy: effects of dose and dose rate. *J Synchrotron Radiat* **18**, 398-412, doi:10.1107/S090904951100820X (2011).

192   Egerton, R. F. Dose measurement in the TEM and STEM. *Ultramicroscopy* **229**, 113363, doi:10.1016/j.ultramic.2021.113363 (2021).

193   Murray, J. W., Garman, E. F. & Ravelli, R. B. G. X-ray absorption by macromolecular crystals: the effects of wavelength and crystal composition on absorbed dose. *Journal of Applied Crystallography* **37**, 513-522, doi:10.1107/S0021889804010660 (2004).

154

194   Owen, R. L., Rudino-Pinera, E. & Garman, E. F. Experimental determination of the radiation dose limit for cryocooled protein crystals. *Proc Natl Acad Sci U S A* **103**, 4912-4917, doi:10.1073/pnas.0600973103 (2006).

195   Glaeser, R. M. Limitations to significant information in biological electron microscopy as a result of radiation damage. *J Ultrastruct Res* **36**, 466-482, doi:10.1016/s0022-5320(71)80118-1 (1971).

196   Bury, C. S., Brooks-Bartlett, J. C., Walsh, S. P. & Garman, E. F. Estimate your dose: RADDOSE-3D. *Protein Sci* **27**, 217-228, doi:10.1002/pro.3302 (2018).

197   Unwin, P. N. & Henderson, R. Molecular structure determination by electron microscopy of unstained crystalline specimens. *J Mol Biol* **94**, 425-440, doi:10.1016/0022-2836(75)90212-0 (1975).

198   Grant, T. & Grigorieff, N. Measuring the optimal exposure for single particle cryo-EM using a 2.6 A reconstruction of rotavirus VP6. *Elife* **4**, e06980, doi:10.7554/eLife.06980 (2015).

199   Tomonaga, K., Takada, M., Ichihara, T. & Kuroda, Y. Apoptosis in Helicobacter pylori gastritis and residual gastritis after distal gastrectomy. *Hepatogastroenterology* **50**, 902-905 (2003).

200   Glaeser, R. M. *et al.* Defocus-dependent Thon-ring fading. *Ultramicroscopy* **222**, 113213, doi:10.1016/j.ultramic.2021.113213 (2021).

201   Wolf, S. G. & Elbaum, M. CryoSTEM tomography in biology. *Methods Cell Biol* **152**, 197-215, doi:10.1016/bs.mcb.2019.04.001 (2019).

202   Waugh, B. *et al.* Three-dimensional deconvolution processing for STEM cryotomography. *Proc Natl Acad Sci U S A* **117**, 27374-27380, doi:10.1073/pnas.2000700117 (2020).

203   Badde, H. G. & Beimer, L. Der Einfluß einer streuenden Phasenplatte auf das elektronenmikroskopische Bild. *Zeitschrift für Naturforschung A* **25**, 760-765, doi:10.1515/zna-1970-0532 (1970).

204   Danev, R., Glaeser, R. M. & Nagayama, K. Practical factors affecting the performance of a thin-film phase plate for transmission electron microscopy. *Ultramicroscopy* **109**, 312-325, doi:10.1016/j.ultramic.2008.12.006 (2009).

205   Murata, K. *et al.* Zernike phase contrast cryo-electron microscopy and tomography for structure determination at nanometer and subnanometer resolutions. *Structure* **18**, 903-912, doi:10.1016/j.str.2010.06.006 (2010).

206   Zernike, F. Phase contrast, a new method for the microscopic observation of transparent objects. *Physica* **9**, 686-698, doi:10.1016/s0031-8914(42)80035-x (1942).

207   Schultheiß, K., Pérez-Willard, F., Barton, B., Gerthsen, D. & Schröder, R. R. Fabrication of a Boersch phase plate for phase contrast imaging in a transmission electron microscope. *Review of Scientific Instruments* **77**, 033701-033701, doi:10.1063/1.2179411 (2006).

208   Turnbaugh, C. *et al.* High-power near-concentric Fabry-Perot cavity for phase contrast electron microscopy. *Rev Sci Instrum* **92**, 053005, doi:10.1063/5.0045496 (2021).

209   Danev, R. & Nagayama, K. Single particle analysis based on Zernike phase contrast transmission electron microscopy. *J Struct Biol* **161**, 211-218, doi:10.1016/j.jsb.2007.10.015 (2008).

210   Walter, A. *et al.* Practical aspects of Boersch phase contrast electron microscopy of biological specimens. *Ultramicroscopy* **116**, 62-72, doi:10.1016/j.ultramic.2012.03.009 (2012).

211   Edgcombe, C. J. *et al.* Characterisation of ferromagnetic rings for Zernike phase plates using the Aharonov-Bohm effect. *Ultramicroscopy* **120**, 78-85, doi:10.1016/j.ultramic.2012.06.011 (2012).

212 Verbeeck, J. *et al.* Demonstration of a 2 x 2 programmable phase plate for electrons. *Ultramicroscopy* **190**, 58-65, doi:10.1016/j.ultramic.2018.03.017 (2018).

213 Glaeser, R. M. Invited review article: Methods for imaging weak-phase objects in electron microscopy. *Rev Sci Instrum* **84**, 111101, doi:10.1063/1.4830355 (2013).

214 Barwick, B. & Batelaan, H. Aharonov-Bohm phase shifts induced by laser pulses. *New Journal of Physics* **10**, 083036-083036, doi:10.1088/1367-2630/10/8/083036 (2008).

215 Marko, M., Meng, X., Hsieh, C., Roussie, J. & Striemer, C. Methods for testing Zernike phase plates and a report on silicon-based phase plates with reduced charging and improved ageing characteristics. *J Struct Biol* **184**, 237-244, doi:10.1016/j.jsb.2013.08.008 (2013).

216 Hettler, S. *et al.* High-resolution transmission electron microscopy with an electrostatic Zach phase plate. *New Journal of Physics* **18**, 053005-053005, doi:10.1088/1367-2630/18/5/053005 (2016).

217 Schwartz, O. *et al.* Near-concentric Fabry-Perot cavity for continuous-wave laser control of electron waves. *Opt Express* **25**, 14453-14462, doi:10.1364/OE.25.014453 (2017).

218 Danev, R. & Nagayama, K. Optimizing the phase shift and the cut-on periodicity of phase plates for TEM. *Ultramicroscopy* **111**, 1305-1315, doi:10.1016/j.ultramic.2011.04.004 (2011).

219 Juffmann, T., Klopfer, B. B., Frankort, T. L. I., Haslinger, P. & Kasevich, M. A. Multi-pass microscopy. *Nat Commun* **7**, 12858, doi:10.1038/ncomms12858 (2016).

220 Juffmann, T. *et al.* Multi-pass transmission electron microscopy. *Sci Rep* **7**, 1699, doi:10.1038/s41598-017-01841-x (2017).

221 Koppell, S. A. *et al.* Design for a 10 keV multi-pass transmission electron microscope. *Ultramicroscopy* **207**, 112834, doi:10.1016/j.ultramic.2019.112834 (2019).

222 Hall, R. J., Nogales, E. & Glaeser, R. M. Accurate modeling of single-particle cryo-EM images quantitates the benefits expected from using Zernike phase contrast. *J Struct Biol* **174**, 468-475, doi:10.1016/j.jsb.2011.03.020 (2011).

223 Su, C. C. *et al.* A 'Build and Retrieve' methodology to simultaneously solve cryo-EM structures of membrane proteins. *Nat Methods* **18**, 69-75, doi:10.1038/s41592-020-01021-2 (2021).

224 Arnaud, J. A. Degenerate optical cavities. *Appl Opt* **8**, 189-195, doi:10.1364/AO.8.000189 (1969).

225 Kolobov, M. I. *Quantum Imaging.* (Springer US, 2007).

226 Prozorov, T., Almeida, T. P., Kovacs, A. & Dunin-Borkowski, R. E. Off-axis electron holography of bacterial cells and magnetic nanoparticles in liquid. *J R Soc Interface* **14**, 20170464-20170464, doi:10.1098/rsif.2017.0464 (2017).

227 Gabor, D. Microscopy by reconstructed wave-fronts. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* **197**, 454-487, doi:10.1098/rspa.1949.0075 (1997).

228 Möllenstedt, G. & Düker, H. Beobachtungen und Messungen an Biprisma-Interferenzen mit Elektronenwellen. *Zeitschrift für Physik* **145**, 377-397, doi:10.1007/BF01326780 (1956).

229 Simon, P. *et al.* Electron holography of organic and biological materials. *Advanced Materials* **15**, 1475-1481, doi:10.1002/adma.200301645 (2003).

230 Simon, P. *et al.* Electron holography of biological samples. *Micron* **39**, 229-256, doi:10.1016/j.micron.2006.11.012 (2008).

231 Dunin-Borkowski, R. E., Kovács, A., Kasama, T., McCartney, M. R. & Smith, D. J. in *Springer Handbook of Microscopy Springer Handbooks* Ch. Chapter 16, 767-818 (Springer, 2019).

232   Völkl, E. & Lichte, H. Electron holograms for subångström point resolution. *Ultramicroscopy* **32**, 177-179, doi:10.1016/0304-3991(90)90035-k (1990).

233   Yamamoto, K., Sugawara, Y., McCartney, M. R. & Smith, D. J. Phase-shifting electron holography for atomic image reconstruction. *J Electron Microsc (Tokyo)* **59 Suppl 1**, S81-88, doi:10.1093/jmicro/dfq033 (2010).

234   Yamamoto, K., Anada, S., Sato, T., Yoshimoto, N. & Hirayama, T. Phase-shifting electron holography for accurate measurement of potential distributions in organic and inorganic semiconductors. *Microscopy (Oxf)* **70**, 24-38, doi:10.1093/jmicro/dfaa061 (2021).

235   Ru, Q., Endo, J., Tanji, T. & Tonomura, A. Phase-Shifting Electron Holography by Beam Tilting. *Applied Physics Letters* **59**, 2372-2374, doi:Doi 10.1063/1.106019 (1991).

236   Ru, Q., Lai, G., Aoyama, K., Endo, J. & Tonomura, A. Principle and Application of Phase-Shifting Electron Holography. *Ultramicroscopy* **55**, 209-220, doi:Doi 10.1016/0304-3991(94)90171-6 (1994).

237   Kawasaki, T., Endo, J., Matsuda, T., Osakabe, N. & Tonomura, A. Applications of holographic interference electron microscopy to the observation of biological specimens. *J Electron Microsc (Tokyo)* **35**, 211-214, doi:10.1093/oxfordjournals.jmicro.a050569 (1986).

238   Weierstall, U. & Lichte, H. Electron holography with a superconducting objective lens. *Ultramicroscopy* **65**, 13-22, doi:Doi 10.1016/S0304-3991(96)00050-2 (1996).

239   Harscher, A. & Lichte, H. Experimental study of amplitude and phase detection limits in electron holography. *Ultramicroscopy* **64**, 57-66, doi:Doi 10.1016/0304-3991(96)00019-8 (1996).

240   Aoyama, K. & Ru, Q. Electron holographic observation for biological specimens: electron holography of bio-specimens. *J Microsc* **182**, 177-185, doi:10.1046/j.1365-2818.1996.133413.x (1996).

241   Lichte, H. *et al.* Electron holography: Applications to materials questions. *Annual Review of Materials Research* **37**, 539-588, doi:10.1146/annurev.matsci.37.052506.084232 (2007).

242   Harscher, A. Elektronenholographie biologischer Objekte: Grundlagen und Anwendungsbeispiele. (1999).

243   Dunin-Borkowski, R. E. *et al.* Magnetic microstructure of magnetotactic bacteria by electron holography. *Science* **282**, 1868-1870, doi:10.1126/science.282.5395.1868 (1998).

244   Tanigaki, T. *et al.* Split-illumination electron holography. *Applied Physics Letters* **101**, 43101-43101, doi:10.1063/1.4737152 (2012).

245   Tanigaki, T. *et al.* Advanced split-illumination electron holography without Fresnel fringes. *Ultramicroscopy* **137**, 7-11, doi:10.1016/j.ultramic.2013.11.002 (2014).

246   Cheung, M., Adaniya, H., Cassidy, C., Yamashita, M. & Shintake, T. Low-energy in-line electron holographic imaging of vitreous ice-embedded small biomolecules using a modified scanning electron microscope. *Ultramicroscopy* **209**, 112883, doi:10.1016/j.ultramic.2019.112883 (2020).

247   Chang, S. L. Y., Dwyer, C., Barthel, J., Boothroyd, C. B. & Dunin-Borkowski, R. E. Performance of a direct detection camera for off-axis electron holography. *Ultramicroscopy* **161**, 90-97, doi:10.1016/j.ultramic.2015.09.004 (2016).

248   Hoppe, W. Beugung im inhomogenen Primärstrahlwellenfeld. I. Prinzip einer Phasenmessung von Elektronenbeugungsinterferenzen. *Acta Crystallographica Section A* **25**, 495-501, doi:10.1107/s0567739469001045 (1969).

249   Pfeiffer, F. X-ray ptychography. *Nature Photonics* **12**, 9-17, doi:10.1038/s41566-017-0072-5 (2018).

250   Baksh, P. D. *et al.* Quantitative and correlative extreme ultraviolet coherent imaging of mouse hippocampal neurons at high resolution. *Sci Adv* **6**, eaaz3025, doi:10.1126/sciadv.aaz3025 (2020).

251   Seaberg, M. D. *et al.* Tabletop nanometer extreme ultraviolet imaging in an extended reflection mode using coherent Fresnel ptychography. *Optica* **1**, 39-44, doi:10.1364/Optica.1.000039 (2014).

252   Odstrcil, M. *et al.* Ptychographic coherent diffractive imaging with orthogonal probe relaxation. *Opt Express* **24**, 8360-8369, doi:10.1364/OE.24.008360 (2016).

253   Rodenburg, J. & Maiden, A. in *Springer Handbooks*   819-904 (Springer, 2019).

254   MacLaren, I., Macgregor, T. A., Allen, C. S. & Kirkland, A. I. Detectors-The ongoing revolution in scanning transmission electron microscopy and why this important to material characterization. *Apl Mater* **8**, 110901-110901, doi:10.1063/5.0026992 (2020).

255   Yang, H., Pennycook, T. J. & Nellist, P. D. Efficient phase contrast imaging in STEM using a pixelated detector. Part II: optimisation of imaging conditions. *Ultramicroscopy* **151**, 232-239, doi:10.1016/j.ultramic.2014.10.013 (2015).

256   O'Leary, C. M. *et al.* Contrast transfer and noise considerations in focused-probe electron ptychography. *Ultramicroscopy* **221**, 113189, doi:10.1016/j.ultramic.2020.113189 (2021).

257   Maiden, A. M. & Rodenburg, J. M. An improved ptychographical phase retrieval algorithm for diffractive imaging. *Ultramicroscopy* **109**, 1256-1262, doi:10.1016/j.ultramic.2009.05.012 (2009).

258   Rodenburg, J. M., Mccallum, B. C. & Nellist, P. D. Experimental Tests on Double-Resolution Coherent Imaging Via Stem. *Ultramicroscopy* **48**, 304-314, doi:Doi 10.1016/0304-3991(93)90105-7 (1993).

259   Chen, Z. *et al.* Mixed-state electron ptychography enables sub-angstrom resolution imaging with picometer precision at low dose. *Nat Commun* **11**, 2994, doi:10.1038/s41467-020-16688-6 (2020).

260   Pelz, P. M., Qiu, W. X., Bucker, R., Kassier, G. & Miller, R. J. D. Low-dose cryo electron ptychography via non-convex Bayesian optimization. *Sci Rep* **7**, 9883, doi:10.1038/s41598-017-07488-y (2017).

261   Van den Broek, W. *et al.* Towards Ptychography with Structured Illumination, and a Derivative-Based Reconstruction Algorithm. *Microscopy and Microanalysis* **25**, 58-59, doi:10.1017/s1431927619001028 (2019).

262   Yang, H., Ercius, P., Nellist, P. D. & Ophus, C. Enhanced phase contrast transfer using ptychography combined with a pre-specimen phase plate in a scanning transmission electron microscope. *Ultramicroscopy* **171**, 117-125, doi:10.1016/j.ultramic.2016.09.002 (2016).

263   Tavabi, A. H. *et al.* Generation of electron vortices using nonexact electric fields. *Physical Review Research* **2**, 013185-013185, doi:10.1103/PhysRevResearch.2.013185 (2020).

264   Chen, Z. *et al.* Electron ptychography achieves atomic-resolution limits set by lattice vibrations. *Science* **372**, 826-831, doi:10.1126/science.abg2533 (2021).

265   Muller, D. A. & Silcox, J. Delocalization in Inelastic-Scattering. *Ultramicroscopy* **59**, 195-213, doi:Doi 10.1016/0304-3991(95)00029-Z (1995).

266   Riekel, C. *et al.* Raster microdiffraction with synchrotron radiation of hydrated biopolymers with nanometre step-resolution: case study of starch granules. *Journal of Synchrotron Radiation* **17**, 743-750, doi:10.1107/S0909049510028335 (2010).

267   Ciston, J. *et al.* The 4D Camera: Very High Speed Electron Counting for 4D-STEM. *Microscopy and Microanalysis* **25**, 1930-1931, doi:10.1017/s1431927619010389 (2019).

268   Berger, C. *et al.* Structure of the Yersinia injectisome in intracellular host cell phagosomes revealed by cryo FIB electron tomography. *J Struct Biol* **213**, 107701, doi:10.1016/j.jsb.2021.107701 (2021).

269   Nickell, S., Kofler, C., Leis, A. P. & Baumeister, W. A visual approach to proteomics. *Nat Rev Mol Cell Biol* **7**, 225-230, doi:10.1038/nrm1861 (2006).

270   Morgan, C., Godman, G. C., Breitenfeld, P. M. & Rose, H. M. A correlative study by electron and light microscopy of the development of type 5 adenovirus. I. Electron microscopy. *J Exp Med* **112**, 373-382, doi:10.1084/jem.112.2.373 (1960).

271   Hanein, D. & Volkmann, N. in *Advances in Protein Chemistry and Structural Biology* Vol. 82 91-99 (Academic Press Inc., 2011).

272   Verbeeck, J., Tian, H. & Schattschneider, P. Production and application of electron vortex beams. *Nature* **467**, 301-304, doi:10.1038/nature09366 (2010).

273   McMorran, B. J. *et al.* Electron vortex beams with high quanta of orbital angular momentum. *Science* **331**, 192-195, doi:10.1126/science.1198804 (2011).

274   Grillo, V. *et al.* Measuring the orbital angular momentum spectrum of an electron beam. *Nat Commun* **8**, 15536, doi:10.1038/ncomms15536 (2017).

275   Tavabi, A. H. *et al.* Experimental Demonstration of an Electrostatic Orbital Angular Momentum Sorter for Electron Beams. *Phys Rev Lett* **126**, 094802, doi:10.1103/PhysRevLett.126.094802 (2021).

276   Pozzi, G. *et al.* Design of electrostatic phase elements for sorting the orbital angular momentum of electrons. *Ultramicroscopy* **208**, 112861, doi:10.1016/j.ultramic.2019.112861 (2020).

277   Troiani, F. *et al.* Efficient molecule discrimination in electron microscopy through an optimized orbital angular momentum sorter. *Physical Review A* **102**, 043510-043510, doi:10.1103/PhysRevA.102.043510 (2020).

278   Yasin, F. S. *et al.* Probing Light Atoms at Subnanometer Resolution: Realization of Scanning Transmission Electron Microscope Holography. *Nano Lett* **18**, 7118-7123, doi:10.1021/acs.nanolett.8b03166 (2018).

279   Winkler, F. *et al.* Absolute Scale Quantitative Off-Axis Electron Holography at Atomic Resolution. *Phys Rev Lett* **120**, 156101, doi:10.1103/PhysRevLett.120.156101 (2018).

280   Wolf, S. G., Houben, L. & Elbaum, M. Cryo-scanning transmission electron tomography of vitrified cells. *Nat Methods* **11**, 423-428, doi:10.1038/nmeth.2842 (2014).

281   Krivanek, O. L. *et al.* Vibrational spectroscopy in the electron microscope. *Nature* **514**, 209-212, doi:10.1038/nature13870 (2014).

282   Egerton, R. F. Vibrational-loss EELS and the avoidance of radiation damage. *Ultramicroscopy* **159 Pt 1**, 95-100, doi:10.1016/j.ultramic.2015.08.003 (2015).

283   Rez, P. *et al.* Damage-free vibrational spectroscopy of biological materials in the electron microscope. *Nat Commun* **7**, 10945, doi:10.1038/ncomms10945 (2016).

284   Padgett, M. J. & Boyd, R. W. An introduction to ghost imaging: quantum and classical. *Philos Trans A Math Phys Eng Sci* **375**, 20160233-20160233, doi:10.1098/rsta.2016.0233 (2017).

285   Li, S. *et al.* Electron Ghost Imaging. *Phys Rev Lett* **121**, 114801, doi:10.1103/PhysRevLett.121.114801 (2018).

286   Leary, R. K. & Midgley, P. A. in *Springer Handbook of Microscopy Springer Handbooks* Ch. Chapter 26, 1279-1329 (Springer, 2019).

287   Kaiser, U. *et al.* Transmission electron microscopy at 20 kV for imaging and spectroscopy. *Ultramicroscopy* **111**, 1239-1246, doi:10.1016/j.ultramic.2011.03.012 (2011).

288   Geelen, D. *et al.* eV-TEM: Transmission electron microscopy in a low energy cathode lens instrument. *Ultramicroscopy* **159 Pt 3**, 482-487, doi:10.1016/j.ultramic.2015.06.014 (2015).

289   Neu, P. S., Geelen, D., Thete, A., Tromp, R. M. & van der Molen, S. J. Complementary LEEM and eV-TEM for imaging and spectroscopy. *Ultramicroscopy* **222**, 113199, doi:10.1016/j.ultramic.2020.113199 (2021).

290   Arnold, S. A. *et al.* Blotting-free and lossless cryo-electron microscopy grid preparation from nanoliter-sized protein samples and single-cell extracts. *J Struct Biol* **197**, 220-226, doi:10.1016/j.jsb.2016.11.002 (2017).

291   Ashtiani, D. *et al.* Delivery of femtolitre droplets using surface acoustic wave based atomisation for cryo-EM grid preparation. *J Struct Biol* **203**, 94-101, doi:10.1016/j.jsb.2018.03.012 (2018).

292   Rubinstein, J. L. *et al.* Shake-it-off: a simple ultrasonic cryo-EM specimen-preparation device. *Acta Crystallogr Sect D* **75**, 1063-1070, doi:10.1107/S2059798319014372 (2019).

293   Feng, X. *et al.* A Fast and Effective Microfluidic Spraying-Plunging Method for High-Resolution Single-Particle Cryo-EM. *Structure* **25**, 663-670 e663, doi:10.1016/j.str.2017.02.005 (2017).

294   White, H. D., Thirumurugan, K., Walker, M. L. & Trinick, J. A second generation apparatus for time-resolved electron cryo-microscopy using stepper motors and electrospray. *J Struct Biol* **144**, 246-252, doi:10.1016/j.jsb.2003.09.027 (2003).

295   Frank, J. Time-resolved cryo-electron microscopy: Recent progress. *J Struct Biol* **200**, 303-306, doi:10.1016/j.jsb.2017.06.005 (2017).

296   Dandey, V. P. *et al.* Time-resolved cryo-EM using Spotiton. *Nat Methods* **17**, 897-900, doi:10.1038/s41592-020-0925-6 (2020).

297   Chen, B. *et al.* Structural dynamics of ribosome subunit association studied by mixing-spraying time-resolved cryogenic electron microscopy. *Structure* **23**, 1097-1105, doi:10.1016/j.str.2015.04.007 (2015).

298   Weissenberger, G., Henderikx, R. J. M. & Peters, P. J. Understanding the invisible hands of sample preparation for cryo-EM. *Nat Methods* **18**, 463-471, doi:10.1038/s41592-021-01130-6 (2021).

299   Lyumkis, D. Challenges and opportunities in cryo-EM single-particle analysis. *J Biol Chem* **294**, 5181-5197, doi:10.1074/jbc.REV118.005602 (2019).

300   Noble, A. J. *et al.* Reducing effects of particle adsorption to the air-water interface in cryo-EM. *Nat Methods* **15**, 793-795, doi:10.1038/s41592-018-0139-3 (2018).

301   Huber, S. T. *et al.* Nanofluidic chips for cryo-EM structure determination from picoliter sample volumes. *Elife* **11**, 1-25, doi:10.7554/eLife.72629 (2022).

302   Reimer, L. & Kohl, H. *Transmission Electron Microscopy Physics of Image Formation.* Vol. 50 (2003).

303   Kryshtafovych, A., Schwede, T., Topf, M., Fidelis, K. & Moult, J. Critical assessment of methods of protein structure prediction (CASP)—Round XIII. *Proteins: Structure, Function, and Bioinformatics* **87**, 1011-1020, doi:10.1002/PROT.25823 (2019).

304   Kryshtafovych, A., Schwede, T., Topf, M., Fidelis, K. & Moult, J. Critical assessment of methods of protein structure prediction (CASP)—Round XIV. *Proteins: Structure, Function, and Bioinformatics* **89**, 1607-1617, doi:10.1002/PROT.26237 (2021).

305   Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583-589, doi:10.1038/s41586-021-03819-2 (2021).

306   Baek, M. *et al.* Accurate prediction of protein structures and interactions using a three-track neural network. *Science* **373**, 871-876, doi:10.1126/science.abj8754 (2021).

307   Varadi, M. *et al.* AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Research* **50**, D439-D444, doi:10.1093/NAR/GKAB1061 (2022).

308   Tegunov, D. & Cramer, P. Real-time cryo-electron microscopy data preprocessing with Warp. *Nat Methods* **16**, 1146-1152, doi:10.1038/s41592-019-0580-y (2019).

309   Bepler, T., Kelley, K., Noble, A. J. & Berger, B. Topaz-Denoise: general deep denoising models for cryoEM and cryoET. *Nat Commun* **11**, 5208, doi:10.1038/s41467-020-18952-1 (2020).

310   Zhu, Y., Ouyang, Q. & Mao, Y. A deep convolutional neural network approach to single-particle recognition in cryo-electron microscopy. *BMC Bioinformatics* **18**, 1-10, doi:10.1186/S12859-017-1757-Y/FIGURES/6 (2017).

311   Wagner, T. *et al.* SPHIRE-crYOLO is a fast and accurate fully automated particle picker for cryo-EM. *Communications Biology 2019 2:1* **2**, 1-13, doi:10.1038/s42003-019-0437-z (2019).

312   Zhang, J. *et al.* PIXER: An automated particle-selection method based on segmentation using a deep neural network. *BMC Bioinformatics* **20**, 1-14, doi:10.1186/S12859-019-2614-Y/TABLES/3 (2019).

313   Zhong, E. D., Bepler, T., Berger, B. & Davis, J. H. CryoDRGN: reconstruction of heterogeneous cryo-EM structures using neural networks. *Nature Methods 2021 18:2* **18**, 176-185, doi:10.1038/s41592-020-01049-4 (2021).

314   Punjani, A. & Fleet, D. 3D Flexible Refinement: Structure and Motion of Flexible Proteins from Cryo-EM. *Microscopy and Microanalysis* **28**, 1218-1218, doi:10.1017/S1431927622005074 (2022).

315   Sanchez-Garcia, R. *et al.* DeepEMhancer: a deep learning solution for cryo-EM volume post-processing. *Communications Biology 2021 4:1* **4**, 1-8, doi:10.1038/s42003-021-02399-1 (2021).

316   Ramírez-Aportela, E., Mota, J., Conesa, P., Carazo, J. M. & Sorzano, C. O. S. DeepRes: A new deep-learning- and aspect-based local resolution method for electron-microscopy maps. *IUCrJ* **6**, 1054-1063, doi:10.1107/S2052252519011692/FQ5008SUP1.PDF (2019).

317   Maddhuri Venkata Subramaniya, S. R., Terashi, G. & Kihara, D. Protein secondary structure detection in intermediate-resolution cryo-EM maps using deep learning. *Nature Methods 2019 16:9* **16**, 911-917, doi:10.1038/s41592-019-0500-1 (2019).

318   He, J., Lin, P., Chen, J., Cao, H. & Huang, S. Y. Model building of protein complexes from intermediate-resolution cryo-EM maps with deep learning-guided automatic assembly. *Nature Communications 2022 13:1* **13**, 1-16, doi:10.1038/s41467-022-31748-9 (2022).

319   Mostosi, P., Schindelin, H., Kollmannsberger, P. & Thorn, A. Haruspex: A Neural Network for the Automatic Identification of Oligonucleotides and Protein Secondary Structure in Cryo-Electron Microscopy Maps. *Angewandte Chemie International Edition* **59**, 14788-14795, doi:10.1002/ANIE.202000421 (2020).

320   Skalidis, I. *et al.* Cryo-EM and artificial intelligence visualize endogenous protein community members. *Structure* **30**, 575-589 e576, doi:10.1016/j.str.2022.01.001 (2022).

321   Gambini, L., Mullarkey, T., Jones, L. & Sanvito, S. Machine-learning approach for quantified resolvability enhancement of low-dose STEM data. *Machine Learning: Science and Technology* **4**, 015025-015025, doi:10.1088/2632-2153/ACBB52 (2023).

322   Noble, A. *High-End Cryo-EM Worldwide*, <https://tiny.cc/cryoem_map> (2023).

323   Fior Markets. *Cryo-electron Microscopy Market is Globally Expected to Drive Growth of USD 1,010.66 million by 2027*, <https://www.globenewswire.com/news-release/2022/05/16/2444160/0/en/Cryo-electron-Microscopy-Market-is-Globally-

Expected-to-Drive-Growth-of-USD-1-010-66-million-by-2027-Fior-Markets.html>
(2023).

CURRICULUM VITAE

Born on June 19, 1993, in Xiaogan, China, Yue Zhang spent most of his life there until he moved to Jinan in 2011 to pursue his bachelor's degree. In 2015, he graduated with a diploma in Energy and Environmental Engineering from Shandong University and subsequently moved to Delft, the Netherlands to continue his graduate studies at the Delft University of Technology. He completed his Master's degree in Sustainable Energy Engineering in 2017 and then began working as a researcher in the university, designing and fabricating in-situ TEM nanochips under the supervision of Prof. Henny Zandbergen.

In 2019, Yue made his way to Maastricht to join the Nanoscopy group at the Maastricht MultiModal Molecular Imaging (M4I) Institute, supervised by Prof. Raimond B. G. Ravelli, Prof. Peter J. Peters, and Dr. Carmen López-Iglesias. During his PhD, he contributed to two projects - Q-sort and TKI/4DEM. In the Q-sort project, Yue worked alongside a team of experts to explore the use of a quantum sorter for low-dose information retrieval from biological specimens. In the 4DEM project, he played a critical role in integrating a Timepix3 Hybrid pixel detector into a 200 keV Thermo Fisher Tecnai Arctica for an automated cryo-EM workflow. He will continue to apply his expertise by building and integrating a Timepix4 Hybrid pixel detector for a cryo-EM workflow. Inspired by the charging effect when using the TPX3 detector, he designed and performed an experiment to characterize the charging effect and deploy graphene to mitigate charging. As Yue continues to work alongside his peers, he is eager to make contribution to the world of cryo-EM.