 doi:10.20944/preprints201803.0043.v1

Orientating the future bio-macromolecular electron microscopy

Fei Sun[1,2,3]

[1] Center for biological imaging, Institute of Biophysics, Chinese Academy of Sciences;

[2] Center for Excellent, National Laboratory of Biomacromolecule, Institute of Biophysics, Chinese Academy of Sciences; Datun Road 15, Chaoyang District, Beijing 100101, China.

[3] School of Life Science, University of Chinese Academy of Sciences.

Email: feisun@ibp.ac.cn

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 March 2018 doi:10.20944/preprints201803.0043.v1

ABSTRACT

With forty years of developments, bio-macromolecule cryo-electron microscopy has met its revolution of resolution and is playing a very important role in structural biology study. According to different specimen states, cryo-electron microscopy (cryo-EM) involves three specific techniques, single particle analysis (SPA), electron tomography and sub-tomogram averaging, and electron diffraction. All these three techniques have not realized their full potentials of solving structures of bio-macromolecules and therefore need to be developed in the future. In this review, the current existing bottlenecks of cryo-EM SPA are discussed with theoretical analysis, which includes air-water interface during specimen cryo-vitrification, bio-macromolecular conformational heterogeneity, focus gradient within thick specimen, and electron radiation damage. Besides, potential solutions of these bottlenecks are proposed and discussed, which are worthy of further investigations in the future.

KEYWORDS

Cryo-electron microscopy; Air-water interface; Conformational heterogeneity; Focus gradient; Radiation damage.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 March 2018 doi:10.20944/preprints201803.0043.v1

1. Introduction

Revealing the detailed 3D structure of bio-macromolecules is one of the important steps to understand the life. In recent years, cryo-electron microscopy (cryo-EM) has completed its revolution and been becoming one of the major biophysical techniques to study the 3D structures of bio-macromolecules, especially of membrane protein complexes and supra macromolecular assemblies, giving a big impact on our understanding of the life.

The development of cryo-EM technology started from the last century in 70~80s when the significant electron radiation damage of biological specimen was discovered and the low-dose illumination technique was proposed [18, 34], the cryo-vitrification method to fix the native structure of biological specimen was established [1, 79], and the image analysis theory to process the low signal-to-niose-ratio (SNR) low-dose cryo-EM micrographs of bio-macromolecules was developed [24]. The potential of electron microscopy to determine the high resolution structure of bio-macromolecules had been revealed by the 3D structure of purple membrane in 1975 [37] and seriously discussed with a theoretical consideration in 1995 [35]. In recent years, with the instrucmental advances in electron optics and improved stabilities of electron microscopes, the development of sophisticated image processing software [33, 59, 68] and the automation of data collection [6, 52, 78], the improvement of phase plate technology [43, 44], especially the success of the direct electron detector [3, 48, 49], cryo-EM technology has reached its revolution in resolution [45] with the milestone work that the structure of transient receptor potential cation channel TRPV1 was determined to 3.4 angstrom in 2013 [49] and the structure of the glutamate

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 March 2018 doi:10.20944/preprints201803.0043.v1

dehydrogenase (GDH) was solved to atomic level in 2016 [55] by cryo-EM. More and more supra bio-macromolecular structures, e.g. spliceosome [87], photosynthetic complex [85], and mitochondrial respiratory complex [83], have been solved to near atomic level, which could not be achieved by using traditional structural biology approaches.

According to the states of the biological specimen and the experimental workflow, cryo-EM technologies can be classified into three different techniques, cryo-EM single particle analysis (SPA), cryo-electron tomography (cryo-ET) and sub-tomogram analysis (STA), and cryo-electron diffraction (cryo-ED). Cryo-EM SPA is used to analyze the 3D structure of purified bio-macromolecules in solution, which are cryo-vitrified into a thin ice layer. Ten thousands of cryo-EM images of the bio-macromolecule are needed to increase SNR and reconstruct a high resolution structure [62]. Cryo-ET can reconstruct the native 3D structure of a local region within a cell or tissue and the *in situ* structure of bio-macromolecules can be further analyzed by STA [7]. There is no need to purify bio-macromolecules from cell or tissue while hundreds of tomograms of cryo-lamella of cell or tissue are needed to obtain a high resolution *in situ* structure of bio-macromolecules. Cryo-ED is a kind of technique to use cryo-electron microscope to analyze the crystallized biological specimen. It collects the electron diffraction data from 2D crystals or 3D nano-crystals of bio-macromolecules and then solves the high resolution structures using crystallographic theories, where the technique to study 2D crystals was used to be called electron crystallography [31] and the one to study 3D nano-crystals emerged recently and was called as micro-electron diffraction [74]

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 March 2018 doi:10.20944/preprints201803.0043.v1

Cryo-EM SPA has been developed mature enough in past decades and been becoming the major approach in the current structural biology study. The Nobel prize in chemistry in 2017 was also given to Jacques Dubochet, Joachim Frank and Richard Henderson for their great dedicative work in developing cryo-EM SPA technique. Cryo-ET STA has been also developed quickly in recent years and will become an important also a unique approach to study the *in situ* structure of bio-macromolecules in the future. In this paper, I will discuss the current bottlenecks of cryo-EM SPA and present some potential solutions in my personal view.

2. Current technical bottlenecks in cryo-EM SPA

There have been many good reviews to describe the technique of cryo-EM SPA including the theory, workflow, image processing and applications [14, 57, 80]. In brief, cryo-EM SPA starts from the cryo-vitrification of bio-macromolecular solution and needs to collect thousands of high quality cryo-EM micrographs in a high-throughput way with a limited illumination dose (20 ~ 60 e/Å^2 normally), a defined magnification yielding a proper pixel size (0.8 ~ 1.5 Å/pixel normally), and a proper defocus range (0.8 ~ 3.0 μm). The subsequent image processing includes micrograph correction (motion and distortion correction, dose weighting) and evaluation, contrast transfer function estimation, particle picking and sorting, 2D and 3D classification, orientation refinement and reconstruction, and post-processing (map sharpening). Since the limited illumination dose yields a very noisy raw image of bio-macromolecules that are embedded in vitreous ice, ten thousands of particle images are needed to increase SNR. Thus, the basic assumption behind cryo-

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 March 2018 doi:10.20944/preprints201803.0043.v1

EM SPA is that all the analyzed bio-macromolecules should have an identical structure and conformation, which is actually not always true.

Starting from 2013 when the near atomic structure of TRPV1 was first solved [49], the number of bio-macromolecular structures studied by cryo-EM SPA increased quickly. Till 2017, there are already 5541 cryo-EM maps deposited in electron microscopy database (EMDB, http://www.ebi.ac.uk/pdbe/emdb) [58] while there are only 1566 entries in 2012. Most of the deposited maps were obtained by cryo-EM SPA. However, only a small portion of the map (313/5541) can reach a resolution higher than 4 Å and only a few structures can be solved to the resolution higher than 3 Å (**Figure 1**), which actually forms a barrier for cryo-EM SPA to be widely applied into pharmaceutical industry.

Several bottlenecks are still there in cryo-EM SPA from sample preparation, data collection to image processing, which forms the barrier to reach a better resolution. In cryo-vitrification during sample preparation, the existence of air-water interface increases the possibility of disassociation and denature of bio-macromolecules, which prohibited many fragile macromolecular complexes to be studied by cryo-EM SPA. The existence of intrinsic conformational flexibility of bio-macromolecules rule out the basic assumption upon the identical structure and conformation in cryo-EM SPA, which restricted approaching high resolution. During cryo-EM imaging, the SNR from the current instrument and hardware is still not enough to study the high resolution 3D structures of bio-macromolecules when their molecular weight is smaller than 60 kDa. The phase plate and direct dection detector technologies can be improved better to increase the current cryo-EM SNR. The cryo-

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 March 2018 doi:10.20944/preprints201803.0043.v1

electron microscope can be made more stable, easy to use and dedicative for the cryo-EM SPA workflow with much improved throughput. When the size of bio-macromolecule becomes large or the thickness of biological material increases, the existence of the Ewald sphere effect and focus gradient will limit the approachable resolution of cryo-EM imaging in the current theoretical framework, which needs to be solved in both experimental and image processing procedures. The ultimate bottleneck of cryo-EM SPA is the physical nature of radiation damage of bio-macromolecules during cryo-EM imaging, which can not be avoided and is the key limitation of achieving atomic resolution in cryo-EM SPA.

Glaeser has made a very good discussion of the above bottlenecks and given thoughtful perspectives [27]. Here, I would like to focus on cryo-vitrification, conformational heterogeneity, thick sample, and electron radiation damage, further discuss these limitations in theory, and propose new solutions to solve these bottlenecks.

2.1 Cryo-vitrification and air-water interface

The current cryo-vitrification method was originally invented by Dubochet's group in 1984 [1], which is called plunge freezing. The EM grid coated with a carbon supporting film is pre-treated with plasma cleaner and then nipped by a fine tweezer and mounted onto the plunge freezer. 3~5 μl bio-macromolecule solution is dropped onto the supporting film. After a few seconds of incubation, most of the liquid is blotted with filter paper, leaving a thin layer of the solution (30 ~ 50 nm thickness) on the grid. Subsequently, the grid is then quickly plunged into liquid ethane that is pre-cooled by liquid nitrogen, resulting the vitrification of the thin solution layer. Finally, the bio-macromolecules are embedded in the

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 March 2018 doi:10.20944/preprints201803.0043.v1

vitreous ice with their native structure preserved. The procedure of this cryo-vitrification method did not change much since its invention and is still now widely used in the world for cryo-EM SPA. Several venders made commercialized instruments (Thermo Fisher Vitrobot, Leica EM GP and Gatan CP3) for cryo-vitrification, which increase the throughput and reproducibility by controlling experimental parameters (e.g. humidity, temperature, blotting time) accurately.

However, more and more labs found that the homogeneity of the bio-macromolecule became significantly decreased from cryo-vitrified sample when compared with the negatively stained sample. The worst case is that although the specimen shows even distribution with homogenous size and shape from the negative stain electron micrographs, few particles could be identified/recognized from the cryo-EM micrographs (**Figure 2A**). The reason of this phenomena has been fully discussed by Glaeser and colleagues [28, 29] and now widely recognized to be the effect from the air-water interface (**Figure 2B**).

During plunge freezing, the thin solution layer after filter paper blotting results a very large surface-to-volume ratio (~ 20 um^{-1}) in comparison with its original value of (~ 2.0 e-3 um^{-1}). Thus the bio-macromolecule in the thin layer of the solution has large opportunity to reach the surface. The movement of the bio-macromolecule $\sqrt{x^2}$ follows the Brownian motion law according to the following equation,

$$\frac{x^2}{t} = \frac{k_B T}{3\pi\eta r} \qquad (1)$$

where, t is the time of the motion, r is the radius of the bio-macromolecule, η is the

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 March 2018 doi:10.20944/preprints201803.0043.v1

coefficient of viscosity of the solution (for water at 10℃, it is 1.308 mPa.s), k_B is the Boltzmann constant, and T is the temperature in Kalvin. For the thickness of 50 nm of the layer and a 20 nm diameter of the bio-macromolecule, the averaged time (at T = 283 K) for the bio-macromolecule reaching the surface can be estimated as ~ 20 ms. The time can be even shorter (~ 6 ms) for a thinner ice (40 nm thickness) and smaller bio-macromolecule (10 nm diameter). As a result, upon the formation of the thin layer of the solution, the bio-macromolecules can quickly approach the surface called the air-water interface, which has the large possibility to induce denaturation of the bio-macromolecule [28]. The thermodynamics of this procedure can be further described as below (**Figure 2B**),

$$\mathrm{A_V} \underset{k_2}{\overset{k_1}{\rightleftharpoons}} \mathrm{A_S} \xrightarrow{k_3} \mathrm{A_D} \qquad (2)$$

where A_V represents the concentration of bio-macromolecule in the solution, A_S represents the concentration of bio-macromolecule bound to the air water interface, and A_D represents the concentration of denatured bio-macromolecule. Thus, the speed of bio-macromolecule denaturation is determined by the reaction constants k_1, k_2, and k_3. If the order of k_3 can be estimated of $1 \times 10^6 s^{-1}$, the denaturation of the bio-macromolecule can occur within 1 μs.

During plunge freezing, the time normally takes seconds from the completion of blotting to plunging into liquid ethane, which are much longer than the one for bio-macromolecules approaching to the air-water interface and denatured. Therefore, it can be explained why it was difficult to obtain good cryo-EM micrographs for those fragile bio-macromolecules

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 March 2018 doi:10.20944/preprints201803.0043.v1

that are easily denatured (or disassocated) although they can be well captured in negative stain electron micrographs (**Figure 2A**).

The existence of air-water interface has become one of the most important bottlenecks for cryo-EM SPA to obtain high resolution structures of many bio-macromolecules. To overcome this bottleneck, people have developed multiple ways to reduce the air-water interface by utilizing an additional ultrathin (~ 2nm) carbon film [28] or developing affinity grids [42, 90]. Adding surfactant into the solution could be also useful to form a 'cover-slip' at the air-water interface and therefore protect the bio-macromolecules from denaturation [60]. In addition, the new instruments using automatic robotics have been recently developed, e.g. Spotiton [61] and 'grid writer' [2], which avoid paper blotting and can minimize the time between the formation of thin layer and the plunge freezing to ~ 50 ms, which therefore greatly reduce the possibility of the molecule approaching to the surface.

According to the equations (1) and (2), there are also other ways to reduce the effect of air-water interface. Using a high concentration of bio-macromolecules during vitrification may increase the coefficient of viscosity and thus increase the time of approaching to the surface. Meanwhile, the high concentration can also have a chance to saturate the air-water interface with the denatured molecules and thus allow enough number of native molecules left in the solution. To be noted that, the high concentration could increase the difficulty in the following particle picking and image processing. Adding a proper chemical reagent to reduce the interaction between the molecule and the air-water interface would be an alternative way. In such case, the reaction constant k_1 in equation (2) is much

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 March 2018 doi:10.20944/preprints201803.0043.v1

smaller than k_2, and thus large portion of molecules are kept in the solution. Using chemical cross-linker to increase the structural stability of bio-macromolecules would be an additional way that the denaturing reaction constant k_3 is reduced.

2.2 Conformational heterogeneity

Cryo-EM SPA assumes that all the bio-macromolecules analyzed in the electron micrographs have an identical structure and conformation. However, this assumption is not rigorously true in most cases because of the unavoidable thermodynamics of bio-macromolecules and would become much worse when there are heterogeneities existing in the sample.

There are two kinds of heterogeneities of bio-macromolecules, the composition heterogeneity and the conformational heterogeneity. The composition heterogeneity refers to the specimen composed of a mixture of molecules with different ligand-bound states [69], a mixture of macromolecular complexes with different subunits stoichiometry, a mixture of macromolecular assembly with different symmetries, or even in a worst case a mixture of the target molecule and contaminations. The conformational heterogeneity refers to the specimen containing the target bio-macromolecules but in different functional states, and can be further divided into two cases, the heterogeneity with discrete conformations [93] and the one with continuous conformations [21].

The existence of heterogeneity will add difficulty in image processing of SPA and prevent achieving high resolution structure. The composition heterogeneity can be possibly

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 March 2018 doi:10.20944/preprints201803.0043.v1

and efficiently solved by improving biochemical preparation procedure, e.g. more specific affinity chromatography. In addition, recent image processing algorithms have been well developed by applying sophisticated statistics tools, e.g. principal component analysis [81, 82], maximum log likelihood [76] and Bayesian method [67]. Utilizing these advanced image processing tools, we are able to perform efficient image classification to separate bio-macromolecular particles with different compositions and solve the composition heterogeneity [84].

The conformational heterogeneity reflects the functional and thermodynamics nature of bio-macromolecules, which could not be easily improved by conventional biochemical approaches. For the case of discrete conformations, the current image classification algorithms can work very well if SNR of the particles are large enough to discriminate the different conformations. However, for the case of continuous conformations, it would be difficult to improve the reconstruction resolution by carrying out 3D classification approaches unless incredible large number of particles are collected. There have been a few image processing approaches developed to try to solve the heterogeneity of continuous conformations, including local optimization refinement [73], masked refinement [4], normal mode method [40] and manifold-embedding method [25]. When using local optimization or masked refinement approach, we assumed that the bio-macromolecular particle can be divided into a number of rigid parts and the relative orientations and positions of different rigid parts contribute to the flexibility of the molecular complex. This assumption worked in many cases to improve the resolution of bio-macromolecular flexible

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 March 2018 doi:10.20944/preprints201803.0043.v1

modules [87]. However, it is not always true and the internal conformational changes of different modules should be also considered in many cases.

The recent proposed normal mode and manifold-embedding methods would be good solutions to study the intrinsic conformational dynamics of bio-macromolecules directly from the raw cryo-EM images of bio-macromolecular particles. The normal mode method first performs atomization of cryo-EM map and then calculates various normal modes of clustered pseudo atoms. The specific modes are selected to simulate cryo-EM maps with continuous conformations, which are then compared with the raw cryo-EM images[40, 66, 77].

The manifold-embedding method maps each projection of bio-macromolecule into a point of hyperspace (N*N dimension, N is the size of projection image). All the points of bio-macromolecules with different orientations and conformations will build a manifold in this hyperspace. The dimension of this manifold is determined by the degree of freedom of bio-macromolecular motion including rotation (3 freedoms), shift (2 freedoms) and conformational changes (various freedoms). Manifold-embedding is a kind of mathematical approach to estimate the degree of freedom of the manifold and decompose these freedoms into different principal coordinates. After decomposition, a specific coordinate can be selected and sorted to reconstruct the conformational changes of bio-macromolecules [15, 25, 71, 72].

The recent published review [41] has discussed various image processing algorithms to solve the conformational heterogeneity, especially the continuous conformation problem.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 March 2018 doi:10.20944/preprints201803.0043.v1

Besides the image processing approaches, here, I would like to propose another biochemical approach to reduce the conformational heterogeneity of bio-macromolecules.

According to the Boltzmann's distribution law, the number of bio-macromolecules at a specific state, N_c, is proportional to $e^{-\frac{E_c}{k_B T}}$, where E_c is the Gibbs energy of the bio-macromolecule in this state c, and k_B is the Boltzmann constant (8.62E-5 eV/K), T is the temperature of the bio-macromolecule solution. Suppose the lowest Gibbs energy of the state is E_L (normally this state is called steady state and it is not degenerative, i.e., only one conformation corresponds to this state), and the highest Gibbs energy of the state is E_H (normally this state is degenerative, i.e., only multiple conformations correspond to this state), the ratio of numbers of molecules between these two states can be determined as,

$$r(T) = \frac{N_L}{N_H} = e^{\frac{E_H - E_L}{k_B T}} = \mathrm{e}^{\frac{\Delta E}{k_B T}} \qquad (3)$$

The Gibbs energy difference ΔE between two states of bio-macromolecules is 40~90 meV [15, 21]. Then the ratio of the numbers can be estimated as, $r(298K) = 4.7 \sim 33.2$ (room temperature), $r(277.6K) = 5.3 \sim 43.0$ (4 degree), $r(253.6K) = 6.2 \sim 61.4$ (-20 degree), $r(193.6K) = 11.0 \sim 219.9$ (-80 degree). Therefore, if we could use chemical cross-linker to fix the steady state at the low temperature (e.g. -20 or -80 degree) before vitrification, there will be more populations of homogenous bio-macromolecular particles of the steady state in the cryo-EM images, which provides an alternative approach to solve the conformational heterogeneity. To utilize this approach, we need to add glycerol or other cryo-protectants into the bio-macromolecular solution in order to keep the solution in liquid

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 March 2018 doi:10.20944/preprints201803.0043.v1

state at low temperature, so that the thermodynamics equilibrium can be reached. Then the chemical cross-linker is added to the cooled solution to allow the cross-linking reaction occur. A good cross-linker needs to be screened and optimized to allow an efficient and fast reaction at low temperature.

2.3 Thick specimen and focus gradient

The current image processing procedure of cryo-EM SPA assumes that the specimen is thin enough so that the dynamic scattering, the Edward sphere and focus gradient effects can be neglected. For 300 kV accelerate voltage and 100 nm thickness of vitrified bio specimen, the dynamic scattering effect can be still neglected because it is smaller than the mean free path (~ 350 nm) of 300 keV electron for the vitrified bio specimen [88]. However, the Edward sphere effect has limited the resolution to 3.8 angstrom according to the formula of $\sqrt{t * \lambda/(2 * 0.7)}$, where t is the specimen thickness and $\lambda\ (= 0.02\ \text{Å})$ is the wavelength of 300 keV electron [16]. The 100 nm focus gradient will induce a phase error of $\frac{\pi}{2}$ at the resolution of 6.3 angstrom according to the formula of $\Delta\chi = \pi\lambda\Delta Z s^2$, where $\Delta\chi$ is the phase error and ΔZ is the focus gradient [92]. Thus, when the size of bio-macromolecule or the vitrified ice is thick, the Edward sphere, especially the focus gradient will take effects and should be corrected to improve the resolution. In cryo-EM, the Edward sphere and focus gradient effects are combined and corrections of these two effects are actually equivalent [16].

Theory of Edward sphere correction has been well developed [16] and various algorithms have been implemented into difference programs and tested with simulated

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 March 2018 doi:10.20944/preprints201803.0043.v1

data [47, 86]. The effect of focus gradient was also carefully discussed recently using simulated data [20]. Here, I will discuss the focus gradient effect in a different way and provide a new approximation to solve and correct this effect.

As shown in **Figure 3**, suppose the thickness of the specimen is D, the underfocus of the proximal side is f_0, and that of the distal side is $f_0 + D$, the averaged underfocus of the specimen is $f_a = f_0 + D/2$. For the thin specimen, the final image can be simply formulated as,

$$I(x,y) = p(x,y) \otimes PSF(f_a, x, y) \qquad (4)$$

where, $p(x,y)$ is the projection of the structural density $f(x,y,z)$ of the specimen,

$$p(x,y) = \int_{f_0}^{f_0+D} f(x,y,z)\, dz \qquad (5)$$

and $PSF(f_a, x, y)$ is the point spread function of the objective lens, and is the Fourier transformation of contrast transfer function,

$$CTF(f_a, s_x, s_y) = Sin\chi(f_a, s) = Sin(\frac{\pi}{2}\lambda {C_s}^3 s^4 - \pi\lambda f_a s^2) \qquad (6)$$

For simplicity, here we do not consider astigmation and amplitude contrast and therefore $s^2 = {s_x}^2 + {s_y}^2$.

However, when the thickness D is large enough, equation (4) has to be corrected by dividing the specimen into a series of thin specimen (n=0, 1, 2, …, N-1), thus the the final image of a thick specimen can be formulated as,

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 March 2018 doi:10.20944/preprints201803.0043.v1

$$I(x,y) = \sum_{n=0}^{N-1} \Delta z * f(x,y,f_0 + n\Delta z) \otimes PSF(f_0 + n\Delta z, x, y) \quad (7)$$

When $N \to +\infty$, we have,

$$I(x,y) = \int_{f_0}^{f_0+D} f(x,y,f_0+z) \otimes PSF(f_0+z,x,y)\, dz$$

$$= \int_{-\frac{D}{2}}^{\frac{D}{2}} f(x,y,f_a+z) \otimes PSF(f_a+z,x,y)\, dz \quad (8)$$

Considering $D\ (\sim 100\ nm) \ll f_a\ (1000 \sim 2500\ nm)$, the point spread function can be expanded at $z = 0$ and approximated as,

$$PSF(f_a+z,x,y) = PSF(f_a,x,y) + z * \frac{\partial PSF}{\partial z}|_{z=0} \quad (9)$$

Combining (8) and (9), we have,

$$I(x,y) = I^0(x,y) + I^1(x,y) \quad (10)$$

$$I^0(x,y) = \int_{-\frac{D}{2}}^{\frac{D}{2}} f(x,y,f_a+z)\, dz \otimes PSF(f_a,x,y) \quad (11)$$

$$I^1(x,y) = \int_{-\frac{D}{2}}^{\frac{D}{2}} f(x,y,f_a+z) z\, dz \otimes \frac{\partial PSF}{\partial z}|_{z=0} \quad (12)$$

Next, we consider Fourier transformation of the structural density $f(x,y,z)$ of the specimen,

$$F(s_x,s_y,s_z) = \iint \int_{z=-\frac{D}{2}}^{\frac{D}{2}} f(x,y,f_a+z) * e^{-i\pi(xs_x+ys_y+zs_z)}\, dzdxdy \quad (13)$$

Thus, we have the projection theorem,

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 March 2018 doi:10.20944/preprints201803.0043.v1

$$F(s_x, s_y, 0) = \iint \int_{z=-\frac{D}{2}}^{\frac{D}{2}} f(x, y, f_a + z)dz * e^{-i\pi(xs_x+ys_y)}\, dxdy$$

$$= \iint p(x, y) * e^{-i\pi(xs_x+ys_y)}\, dxdy \quad (14)$$

and the following relationship,

$$\frac{\partial F}{\partial s_z}|_{s_z=0} = \int f(x, y, f_a + z) * e^{-i\pi(xs_x+ys_y+zs_z)} * (-i\pi z)dxdydz\,|_{s_z=0}$$

$$= \iint \frac{\pi}{i} \int_{z=-\frac{D}{2}}^{\frac{D}{2}} f(x, y, f_a + z)zdz * e^{-i\pi(xs_x+ys_y)}dxdy \quad (15)$$

As a result, combining (14) and (15) and utilizing the convolution theorem, the Fourier transformation of equation (10) becomes,

$$\hat{I}(s_x, s_y) = F(s_x, s_y, 0) * CTF(f_a, s) + \frac{i}{\pi}\frac{\partial F}{\partial s_z}|_{s_z=0} * \frac{\partial CTF(f_a+z,s)}{\partial z}|_{z=0} \quad (16)$$

Combining (6) and (16), we have,

$$\hat{I}(s_x, s_y) = F(s_x, s_y, 0) * Sin\,\chi(f_a, s) - i\lambda\left(\frac{\partial F}{\partial s_z}\right)_{s_z=0} * Cos\chi(f_a, s) * s^2 \quad (17)$$

From (17), it is clear to see, for the thick specimen, the Fourier transformation of cryo-EM image contains an additional term $i\lambda\left(\frac{\partial F}{\partial s_z}\right)_{s_z=0} * Cos\chi(f_a, s) * s^2$. This term has a minor contribution at the low resolution but will interfere significantly with the first term and induce Thron ring fading out at the high resolution.

Equation (17) provides at least two ways to solve and correct the focus gradient effect. For the first way, we can take two cryo-EM images of the same specimen with different

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 March 2018 doi:10.20944/preprints201803.0043.v1

underfocus, yielding two equations. The underfocus parameters can be accurately estimated by fitting Thron ring and using the low resolution data. Then the structural factor $F(s_x, s_y, 0)$ can be solved directly. The concern of reduced SNR from additional radiation damage of the second exposure has been discussed carefully and can be eliminated [16].

The second way does not need two experimental images but apply an iterative algorithm to solve the structural factor. Initially, the second term of (17) is neglected and normal cryo-EM SPA procedure is performed to get the first round of structure. Then $\frac{\partial F}{\partial s_z}$ can be computed and will be used to update the structural factor $F(s_x, s_y, 0)$ according to (17). Therefore, the structure of cryo-EM map can be reconstructed again with an improved resolution. This procedure can be iterated several rounds until the convergence reaches.

2.4 Beam induced motion and radiation damage

Cryo-EM of bio-macromolecules embedded in vitreous ice has suffered from beam induced motion (BIM) for many years. When accelerated high-energy electrons interact with the specimen, the electrons from the specimen will be scattered and become secondary electrons coming out of the specimen, leaving positively charged annulus at the illumination area [9]. This is called charging effect ('Berriman effect') from electron beam illumination. Since the thin layer of vitreous ice is an insulator, the positive charges can not be quickly compensated from the environment and thus induce the subsequent physical effects. First, the irregular and metastable structure of the vitrified ice can easily response to the internal electrostatic repulsion stress from the positive charges and thus result a

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 March 2018 doi:10.20944/preprints201803.0043.v1

global mechanical deformation. Such global mechanical deformation is much significant at the beginning of electron illumination [48] and was observed to have a dome-like shape [8]. The mechanical deformation of the ice layer during electron illumination will result a blurred cryo-EM image and weaken the high resolution information significantly. Second, the positive charged annulus will induce additional phase shift of electron beam just like a microscopic electrostatic lens, which will induce an additional contrast loss and blurring of the final image [9], which, however and fortunately, has the minimal effect on the high resolution cryo-EM SPA technique according to the recent study [63].

Besides charging effect that causes BIM, electron beam illumination will induce another more severe effect, which is called radiation damage (or radiolysis). When the secondary electron is kicked out, the chemical covalent bond of molecule will be broken, generating many free-radicals. The free-radicals can migrate quickly and react with the adjacent molecules. As a result, the structure of bio-macromolecule can be damaged effectively from the electron beam illumination and the high resolution structural information will be damped significantly when the illumination dose increased [32, 34]. Thus, electron radiation damage must be carefully considered and the electron dose should be carefully controlled for high resolution cryo-EM SPA work.

The electron radiation damage of vitreous ice embedded bio-macromolecules will further cause the ‘bubbling’ effect that is routinely observed in cryo-EM experiments. Since both water molecules and bio-macromolecules contain abundant hydrogen atoms, electron radiolysis will generate large amount of hydrogen free-radicals and these hydrogen free-

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 March 2018 doi:10.20944/preprints201803.0043.v1

radicals will react subsequently to form hydrogen molecule [46]. At the interfaces between bio-macromolecules and ice or the support carbon film and ice, hydrogen molecules are frequently accumulated into a high concentration and form the hydrogen gas pocket, which was observed as the ‘bubbling’ effect [12]. The generation of hydrogen gas pocket will produce additional mechanical stress within the cryo-vitrified specimen and thus become another factor of BIM. By using a very low electron radiation dose-rate, the accumulation of hydrogen gas can be effectively decreased and thus the ‘bubbling’ effect can be alleviated [12]. To be noted, the ‘bubbling’ effect used to be developed as a kind of technique to study the internal nucleic structure of virus [13].

Recently, another kind of BIM effect, called beam-induced Brownian motion, was proposed and studied [54]. This effect describes a pseudo-Brownian motion of vitreous ice embedded bio-macromolecules, which is generated from the beam-induced movement of water molecules. Fortunately, the experimental data and simulation study by Henderson and coworkers suggested that this beam-induced Brownian motion has the minimal effect on the current cryo-EM SPA work unless the size of bio-macromolecule is small and the target resolution goes beyond 2 angstrom [54].

The existence of BIM and electron radiation damage has been aware of the key bottleneck of high resolution cryo-EM SPA for many years until the emerging of the direct electron detector (DED) [39]. The high detective quantum efficiency (DQE) of the DED camera [53] allows to retain the high resolution weak signal under low dose electron radiation that is important to reduce radiation damage of bio-macromolecules. Equally

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 March 2018 doi:10.20944/preprints201803.0043.v1

importantly, the CMOS architecture of DEM camera enables a high frame rate to record a single exposure into a movie, which can be utilized to correct BIM efficiently by applying appropriate image processing algorithms [8, 48, 94]. In addition, by using dose fractionation and damage compensation algorithms, the movie mode of DED camera can further allow to use a high illumination dose to take cryo-EM images for a better contrast [32].

Besides the direct electron detector that can correct BIM during image processing procedure, additional efforts have been also performed to alleviate BIM effect and these include spot scan imaging [10, 19], paraxial charge compensator [5], and development of various supporting films such Cryo-Mesh grid, graphene film and pure gold grid [64, 65, 89]. The pure gold grid was proved to have sufficient mechanical stiffness and good conductivity, which can therefore reduce BIM [65] and has been successfully applied in many high resolution cryo-EM SPA applications [17].

Overall, the efforts of developing direct electron detector, motion correction algorithms and new types of supporting films in recent years have significantly reduce the effects from BIM and electron radiation damage. More and more bio-macromolecular structures are solved to near atomic resolution (3 ~ 4 angstrom) by cryo-EM SPA approach. However, the electron radiation damage effect is still existing and will be the most important bottleneck of cryo-EM SPA in the future to achieve atomic resolution. Previous studies have showed that the high resolution (~ 3 angstrom) information of vitreous embedded biological specimen already falls off after a low dose ($3\ e\ /\text{Å}^2$) of electron radiation[34]. The severe

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 March 2018 doi:10.20944/preprints201803.0043.v1

mechanical deformation of ice layer at the first dose fractionated frames could not be corrected by image processing algorithms [48]. As a result, while the first couple of frames with less radiation damage contains atomic resolution information, this information can not be restored due to a large BIM and thus these frames have to be discarded in the subsequent image processing.

In the future, there will be two potential ways to further alleviate the electron radiation damage effect. The first possible way is to utilize quantum entanglement effect of electrons to reduce the shot noise of electron beam from the normal scale $\sim 1/N^{1/2}$ to the Heisenberg limit $\sim 1/N$ [27, 56]. As a result, we could utilize a even low electron dose ($1\ e\ /\text{Å}^2$) to capture a good image with enough SNR and less radiation damage [56]. The another possible way is to consider the time scale of electron radiation damage. If we could take cryo-EM image before the specimen damage occurs, we thus could have an opportunity of obtaining the damage-free and high resolution structure of bio-macromolecule. This idea has been proved in the field of serial femtosecond X-ray crystallography with the term of 'diffraction before damage' [51]. In cryo-EM of bio-macromolecules, it is important to estimate the time scale of specimen damage from electron radiation and then verify the possibility of the term of 'imaging before damage'.

For 300kV transmission electron microscope, the accelerated electron gains a high speed $\mathrm{v} = \mathrm{c}\sqrt{1 - \frac{1}{\left(1+{}^{E}/_{E_0}\right)^2}} = 0.78c = 2.3 \times 10^8\, m/s$, where the relativistic effect has to be considered, the static energy of electron is $E_0 = m_0c^2 = 511keV$, the kinetic energy is $\mathrm{E} =$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 March 2018 doi:10.20944/preprints201803.0043.v1

$300\mathrm{keV}$ and the light speed in vacuum is $c = 3 \times 10^8\, m/s$. Then, the time for an electron traveling across a specimen with the thickness of $d = 100\, nm$ is $\Delta t = \frac{d}{v} = 0.33\, fs$. Thus, we could estimate the time scale of the interaction between the specimen and the high energy electron is $\sim 1\, fs$. Then there will be many damage events occurred in the biological specimen, which can be divided into two processes, the primary damage process and the secondary damage process [30, 38]. The primary damage process includes chemical bond breaking, ionization, and production of secondary electrons and free radicals. Previous studies suggested that the time scale of the primary damage process is $1 \sim 10\, ps$ and such damage does not influence the electron microscopic image appreciably [70], because the positions of atoms do not change noticeably at this time scale. The only detectable damage in the electron microscopic image occurs in the secondary process, which initiates from the transition of free radicals and includes subsequent cascade reactions induced by free radicals and productions of new chemical bonds. During the second process, the positions of atoms in the specimen have changed significantly, resulting appreciable damage effect in the final electron microscopic image. The time scale of the second process depends on the rate of free radical transition, which can be estimated as below.

As discussed above, the abundant free radicals generated from vitreous ice embedded bio-macromolecules are hydrogen free radicals H^*. The transition of H^* follows the Fick's laws of diffusion,

$$J = vC = -D \cdot \frac{dC}{dx} \qquad (18)$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 March 2018 doi:10.20944/preprints201803.0043.v1

where, v is the transition rate of H^*, C is the local concentration of H^*, the diffusion coefficient can be calculated according to Stokes-Einstein relationship,

$$D = \frac{k_B T}{6\pi\eta r} \quad (19)$$

where, k_B is the Boltzmann constant (1.38E-23 J/K), T is the temperature of vitrified specimen, η is the viscosity of the vitreous ice, and r $(\sim 10^{-10} m)$ is the radius of H^*. Combining (18) and (19), we can calculate the transition rate of H^* as,

$$v = \frac{1}{C} \cdot \frac{k_B T}{6\pi\eta r} \cdot \frac{dC}{dx} \quad (20)$$

Suppose, initially, the free radicals H^* are concentrated in a small cubic region with the size of 1 nm^3, thus we could have the following estimation,

$$\frac{1}{C} \cdot \frac{dC}{dx} = \frac{1}{C} \cdot \frac{\Delta C}{\Delta x} \sim \frac{1}{\Delta x} \sim 10^9 m^{-1} \quad (21)$$

Thus, at the temperature ($T = 90\ K$) of cryo-EM experiments, the transition rate of H^* can be estimated as,

$$v \sim \frac{1.38\times10^{-23}\times90}{6\times3.14\times10^{-3}\times10^{-10}} \times 10^9 = 0.7\ m/s \quad (22)$$

Then the time for H^* traveling $0.3\ nm$, the averaged distance to reach adjacent groups and then perform radical reaction, can be estimated as,

$$\Delta t \sim \frac{0.3\ nm}{0.7\ m/s} = 0.4\ ns \quad (23)$$

To be noted that, the estimation in (22) utilizes the water viscosity at room temperature,

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 March 2018 doi:10.20944/preprints201803.0043.v1

$\eta_{H_2O} = 10^{-3}\ Pa \cdot s$, where the viscosity of the vitreous ice η_{ice} at $90\ K$ should be much larger (e.g. more than ten times) than the water viscosity at room temperature. Therefore, the time scale for the second damage process is in the nanoseconds $\sim 10\ ns$.

The above estimation suggests that if we could take a cryo-EM exposure within $10\ ns$, the appreciable electron radiation damage during the second process can be almost avoided in the final recorded micrograph. The recent developed ultrafast transmission electron microscopy (UEM) [11, 26, 75, 91] has actually provided an opportunity to test this kind of idea. There are two operation modes of UEM, the stroboscopic mode with picosecond temporal resolution and the single-pulse mode with nanosecond temporal resolution [75]. The stroboscopic mode is useful for ultrafast electron diffraction experiments and suitable to study the reversible process of the material. However, the electron radiation damage of bio-macromolecules is irreversible. Thus, to achieve the concept of 'imaging before damage', it is necessary to develop the cryo-ultrafast transmission electron microscopy (cryo-UEM) that is operated in the single-pulse mode. Although there have been some reports of using UEM to observe the biological specimens [22, 23, 50], all these studies were performed in the stroboscopic mode and utilized the dehydrated specimen, which should not be relevant to biological functions. There is still a big space to develop and improve the single-pulse UEM technology. We are looking forward to the future maturation of cryo-UEM that will bring the bio-macromolecular electron microscopy into a new era.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 March 2018 doi:10.20944/preprints201803.0043.v1

3. Conclusions

Cryo-EM SPA has become the most important technique of bio-macromolecular electron microscopy. The era of studying the structures of bio-macromolecules by using Cryo-EM SPA is just beginning. In the near future, we will see more and more sophisticated bio-macromolecular complexes whose structures are solved into near atomic resolution, making significant implications to their biological functions. However, just as Henderson [36] and Glaeser [27] pointed, cryo-electron microscopy has not yet realized its full potential. In the future, with a better cryo-vitrification technique to avoid air-water interface problem, with a better camera and new type of microscope to further alleviate electron radiation damage effect and with some novel image processing algorithms and experimental techniques to solve the focus gradient problem as well as the conformational heterogeneity issue, Cryo-EM SPA will expand its full ability to solve the atomic resolution of bio-macromolecules.

ACKNOWLEDGEMENTS

I would like to apologize that there are also many other pioneer works of developing bio-macromolecular electron microscopy, which are not mentioned in this review due to limited space. I would like to thank Dr. Xiaojun Huang for her critical comments of the manuscript, Shuangbo Zhang for his assistance of literatures searching, and Ping Shan and Ruigang Su for their assistances in the lab management. This work is supported by grants from Chinese Academy of Sciences (ZDKYYQ20170002 and XDB08030202) and Ministry of Science and Technology of China (2017YFA0504700 and 2014CB910700).

References

[1] Adrian M, Dubochet J, Lepault J and McDowall A W 1984 Cryo-electron microscopy of viruses *Nature* **308** 32-6

[2] Arnold S A, Albiez S, Bieri A, Syntychaki A, Adaixo R, McLeod R A, Goldie K N, Stahlberg H and Braun T 2017 Blotting-free and lossless cryo-electron microscopy grid preparation from nanoliter-sized protein samples and single-cell extracts *Journal of structural biology* **197** 220-6

[3] Bai X C, Fernandez I S, McMullan G and Scheres S H 2013 Ribosome structures to near-atomic resolution from thirty thousand cryo-EM particles *eLife* **2** e00461

[4] Bai X C, Rajendra E, Yang G, Shi Y and Scheres S H 2015 Sampling the conformational space of the catalytic subunit of human gamma-secretase *eLife* **4**

[5] Berriman J A and Rosenthal P B 2012 Paraxial charge compensator for electron cryomicroscopy *Ultramicroscopy* **116** 106-14

[6] Biyani N, Righetto R D, McLeod R, Caujolle-Bert D, Castano-Diez D, Goldie K N and Stahlberg H 2017 Focus: The interface between data collection and data processing in cryo-EM *Journal of structural biology* **198** 124-33

[7] Briggs J A 2013 Structural biology in situ--the potential of subtomogram averaging *Curr Opin Struct Biol* **23** 261-7

[8] Brilot A F, Chen J Z, Cheng A, Pan J, Harrison S C, Potter C S, Carragher B, Henderson R and Grigorieff N 2012 Beam-induced motion of vitrified specimen on holey carbon film *Journal of structural biology* **177** 630-7

[9] Brink J, Sherman M B, Berriman J and Chiu W 1998 Evaluation of charging on macromolecules in electron cryomicroscopy *Ultramicroscopy* **72** 41-52

[10] Bullough P and Henderson R 1987 Use of spot-scan procedure for recording low-dose micrographs of beam-sensitive specimens *Ultramicroscopy* **21** 223-30

[11] Cao G, Sun S, Li Z, Tian H, Yang H and Li J 2015 Clocking the anisotropic lattice dynamics of multi-walled carbon nanotubes by four-dimensional ultrafast transmission electron microscopy *Scientific reports* **5** 8404

[12] Chen J Z, Sachse C, Xu C, Mielke T, Spahn C M and Grigorieff N 2008 A dose-rate effect in single-particle electron microscopy *Journal of structural biology* **161** 92-100

[13] Cheng N, Wu W, Watts N R and Steven A C 2014 Exploiting radiation damage to map proteins in nucleoprotein complexes: the internal structure of bacteriophage T7

Journal of structural biology **185** 250-6

[14] Cheng Y 2015 Single-Particle Cryo-EM at Crystallographic Resolution *Cell* **161** 450-7

[15] Dashti A, Schwander P, Langlois R, Fung R, Li W, Hosseinizadeh A, Liao H Y, Pallesen J, Sharma G, Stupina V A, Simon A E, Dinman J D, Frank J and Ourmazd A 2014 Trajectories of the ribosome as a Brownian nanomachine *Proc Natl Acad Sci U S A* **111** 17492-7

[16] DeRosier D J 2000 Correction of high-resolution data for curvature of the Ewald sphere *Ultramicroscopy* **81** 83-98

[17] des Georges A, Clarke O B, Zalk R, Yuan Q, Condon K J, Grassucci R A, Hendrickson W A, Marks A R and Frank J 2016 Structural Basis for Gating and Activation of RyR1 *Cell* **167** 145-57 e17

[18] Downing K H 1988 Observations of restricted beam-induced specimen motion with small-spot illumination *Ultramicroscopy* **24** 387-97

[19] Downing K H and Glaeser R M 1986 Improvement in high resolution image quality of radiation-sensitive specimens achieved with reduced spot size of the electron beam *Ultramicroscopy* **20** 269-78

[20] Downing K H and Glaeser R M 2017 Estimating the effect of finite depth of field in single-particle cryo-EM *Ultramicroscopy* **184** 94-9

[21] Fischer N, Konevega A L, Wintermeyer W, Rodnina M V and Stark H 2010 Ribosome dynamics and tRNA movement by time-resolved electron cryomicroscopy *Nature* **466** 329-33

[22] Fitzpatrick A W, Lorenz U J, Vanacore G M and Zewail A H 2013 4D cryo-electron microscopy of proteins *Journal of the American Chemical Society* **135** 19123-6

[23] Fitzpatrick A W, Park S T and Zewail A H 2013 Exceptional rigidity and biomechanics of amyloid revealed by 4D electron microscopy *Proc Natl Acad Sci U S A* **110** 10976-81

[24] Frank J 1990 Classification of macromolecular assemblies studied as 'single particles' *Quarterly reviews of biophysics* **23** 281-329

[25] Frank J and Ourmazd A 2016 Continuous changes in structure mapped by manifold embedding of single-particle data in cryo-EM *Methods*

[26] Fu X, Chen B, Tang J, Hassan M T and Zewail A H 2017 Imaging rotational dynamics of nanoparticles in liquid by 4D electron microscopy *Science (New York, N.Y* **355** 494-8

[27] Glaeser R M 2016 How good can cryo-EM become? *Nature methods* **13** 28-32

[28] Glaeser R M and Han B G 2017 Opinion: hazards faced by macromolecules when confined to thin aqueous films *Biophys Rep* **3** 1-7

[29] Glaeser R M, Han B G, Csencsits R, Killilea A, Pulk A and Cate J H 2016 Factors that Influence the Formation and Stability of Thin, Cryo-EM Specimens *Biophys J* **110** 749-55

[30] Glaeser R M and Taylor K A 1978 Radiation damage relative to transmission electron microscopy of biological specimens at low temperature: a review *Journal of microscopy* **112** 127-38

[31] Gonen T, Cheng Y, Sliz P, Hiroaki Y, Fujiyoshi Y, Harrison S C and Walz T 2005 Lipid-protein interactions in double-layered two-dimensional AQP0 crystals *Nature* **438** 633-8

[32] Grant T and Grigorieff N 2015 Measuring the optimal exposure for single particle cryo-EM using a 2.6 A reconstruction of rotavirus VP6 *eLife* **4** e06980

[33] Grigorieff N 2016 Frealign: An Exploratory Tool for Single-Particle Cryo-EM *Methods Enzymol* **579** 191-226

[34] Hayward S B and Glaeser R M 1979 Radiation damage of purple membrane at low temperature *Ultramicroscopy* **04** 201-10

[35] Henderson R 1995 The potential and limitations of neutrons, electrons and X-rays for atomic resolution microscopy of unstained biological molecules *Quarterly reviews of biophysics* **28** 171-93

[36] Henderson R 2015 Overview and future of single particle electron cryomicroscopy *Arch Biochem Biophys*

[37] Henderson R and Unwin P N 1975 Three-dimensional model of purple membrane obtained by electron microscopy *Nature* **257** 28-32

[38] Hendrickson W A 1976 Radiation damage in protein crystallography *Journal of molecular biology* **106** 889-93

[39] Jin L, Milazzo A C, Kleinfelder S, Li S, Leblanc P, Duttweiler F, Bouwer J C, Peltier S T, Ellisman M H and Xuong N H 2008 Applications of direct detection device in transmission electron microscopy *Journal of structural biology* **161** 352-8

[40] Jin Q, Sorzano C O, de la Rosa-Trevin J M, Bilbao-Castro J R, Nunez-Ramirez R, Llorca O, Tama F and Jonic S 2014 Iterative elastic 3D-to-2D alignment method using normal modes for studying structural dynamics of large macromolecular complexes *Structure* **22** 496-506

[41] Jonic S 2017 Computational methods for analyzing conformational variability of macromolecular complexes from cryo-electron microscopy images *Curr Opin Struct Biol* **43** 114-21

[42] Kelly D F, Dukovski D and Walz T 2010 Strategy for the use of affinity grids to prepare non-His-tagged macromolecular complexes for single-particle electron microscopy *Journal of molecular biology* **400** 675-81

[43] Khoshouei M, Pfeffer S, Baumeister W, Forster F and Danev R 2016 Subtomogram analysis using the Volta phase plate *Journal of structural biology*

[44] Khoshouei M, Radjainia M, Phillips A J, Gerrard J A, Mitra A K, Plitzko J M, Baumeister W and Danev R 2016 Volta phase plate cryo-EM of the small protein complex Prx3 *Nat Commun* **7** 10534

[45] Kuhlbrandt W 2014 Biochemistry. The resolution revolution *Science (New York, N.Y* **343** 1443-4

[46] Leapman R D and Sun S 1995 Cryo-electron energy loss spectroscopy: observations on vitrified hydrated specimens and radiation damage *Ultramicroscopy* **59** 71-9

[47] Leong P A, Yu X, Zhou Z H and Jensen G J 2010 *Methods in Enzymology,* ed J J Grant: Academic Press) pp 369-80

[48] Li X, Mooney P, Zheng S, Booth C R, Braunfeld M B, Gubbens S, Agard D A and Cheng Y 2013 Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM *Nature methods*

[49] Liao M, Cao E, Julius D and Cheng Y 2013 Structure of the TRPV1 ion channel determined by electron cryo-microscopy *Nature* **504** 107-12

[50] Lorenz U J and Zewail A H 2013 Biomechanics of DNA structures visualized by 4D electron microscopy *Proc Natl Acad Sci U S A* **110** 2822-7

[51] Martin-Garcia J M, Conrad C E, Coe J, Roy-Chowdhury S and Fromme P 2016 Serial femtosecond crystallography: A revolution in structural biology *Arch Biochem Biophys* **602** 32-47

[52] Mastronarde D 2003 SerialEM A Program for Automated Tilt Series Acquisition on Tecnai Microcopes Using Prediction of Specimen Position *Microscopy and Microanalysis* **9** 1182-3

[53] McMullan G, Faruqi A R, Clare D and Henderson R 2014 Comparison of optimal performance at 300keV of three direct electron detectors for use in low dose electron microscopy *Ultramicroscopy* **147** 156-63

[54] McMullan G, Vinothkumar K R and Henderson R 2015 Thon rings from amorphous ice and implications of beam-induced Brownian motion in single particle electron cryo-microscopy *Ultramicroscopy* **158** 26-32

[55] Merk A, Bartesaghi A, Banerjee S, Falconieri V, Rao P, Davis M I, Pragani R, Boxer M B, Earl L A, Milne J L and Subramaniam S 2016 Breaking Cryo-EM Resolution Barriers to Facilitate Drug Discovery *Cell* **165** 1698-707

[56] Okamoto H 2012 Possible use of a Cooper-pair box for low-dose electron microscopy *Physical Review A* **85** 043810

[57] Orlova E V and Saibil H R 2011 Structural analysis of macromolecular assemblies by electron microscopy *Chem Rev* **111** 7710-48

[58] Patwardhan A 2017 Trends in the Electron Microscopy Data Bank (EMDB) *Acta Crystallogr D Struct Biol* **73** 503-8

[59] Punjani A, Rubinstein J L, Fleet D J and Brubaker M A 2017 cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination *Nature methods*

[60] Quinn P J and Dawson R M 1970 An analysis of the interaction of protein with lipid monolayers at the air-water interface *Biochem J* **116** 671-80

[61] Razinkov I, Dandey V, Wei H, Zhang Z, Melnekoff D, Rice W J, Wigge C, Potter C S and Carragher B 2016 A new method for vitrifying samples for cryoEM *Journal of structural biology* **195** 190-8

[62] Rosenthal P B and Henderson R 2003 Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy *Journal of molecular biology* **333** 721-45

[63] Russo C J and Henderson R 2018 Microscopic charge fluctuations cause minimal contrast loss in cryoEM *Ultramicroscopy* **187** 56-63

[64] Russo C J and Passmore L A 2014 Controlling protein adsorption on graphene for cryo-EM using low-energy hydrogen plasmas *Nature methods*

[65] Russo C J and Passmore L A 2016 Ultrastable gold substrates: Properties of a support for high-resolution electron cryomicroscopy of biological specimens *Journal of structural biology* **193** 33-44

[66] Sanchez Sorzano C O, Alvarez-Cabrera A L, Kazemi M, Carazo J M and Jonic S 2016 StructMap: Elastic Distance Analysis of Electron Microscopy Maps for Studying Conformational Changes *Biophys J* **110** 1753-65

[67] Scheres S H 2012 A Bayesian view on cryo-EM structure determination *Journal of*

molecular biology **415** 406-18

[68] Scheres S H 2012 RELION: implementation of a Bayesian approach to cryo-EM structure determination *Journal of structural biology* **180** 519-30

[69] Scheres S H, Gao H, Valle M, Herman G T, Eggermont P P, Frank J and Carazo J M 2007 Disentangling conformational states of macromolecules in 3D-EM through likelihood optimization *Nature methods* **4** 27-9

[70] Schnabl H 1980 Does removal of hydrogen change the electron energy-loss spectra of DNA bases? *Ultramicroscopy* **5** 147-51

[71] Schwander P, Fung R, G. N. Phillips J and Ourmazd A 2010 Mapping the conformations of biological assemblies *New Journal of Physics* **12** 035007

[72] Schwander P, Fung R and Ourmazd A 2014 Conformations of macromolecules and their complexes from heterogeneous datasets *Philos Trans R Soc Lond B Biol Sci* **369** 20130567

[73] Shan H, Wang Z, Zhang F, Xiong Y, Yin C C and Sun F 2016 A local-optimization refinement algorithm in single particle analysis for macromolecular complex with multiple rigid modules *Protein & cell* **7** 46-62

[74] Shi D, Nannenga B L, Iadanza M G and Gonen T 2013 Three-dimensional electron crystallography of protein microcrystals *eLife* **2** e01345

[75] Shorokhov D and Zewail A H 2016 Perspective: 4D ultrafast electron microscopy-Evolutions and revolutions *J Chem Phys* **144** 080901

[76] Sigworth F J, Doerschuk P C, Carazo J M and Scheres S H 2010 An introduction to maximum-likelihood methods in cryo-EM *Methods Enzymol* **482** 263-94

[77] Sorzano C O, de la Rosa-Trevin J M, Tama F and Jonic S 2014 Hybrid Electron Microscopy Normal Mode Analysis graphical interface and protocol *Journal of structural biology* **188** 134-41

[78] Suloway C, Pulokas J, Fellmann D, Cheng A, Guerra F, Quispe J, Stagg S, Potter C S and Carragher B 2005 Automated molecular microscopy: the new Leginon system *Journal of structural biology* **151** 41-60

[79] Taylor K A and Glaeser R M 1974 Electron diffraction of frozen, hydrated protein crystals *Science (New York, N.Y* **186** 1036-7

[80] Thompson R F, Walker M, Siebert C A, Muench S P and Ranson N A 2016 An introduction to sample preparation and imaging by cryo-electron microscopy for structural biology *Methods* **100** 3-15

[81] van Heel M 1984 Multivariate statistical classification of noisy images (randomly oriented biological macromolecules) *Ultramicroscopy* **13** 165-83

[82] Van Heel M and Frank J 1981 Use of multivariate statistics in analysing the images of biological macromolecules *Ultramicroscopy* **6** 187-94

[83] Vinothkumar K R, Zhu J and Hirst J 2014 Architecture of mammalian respiratory complex I *Nature* **advance online publication**

[84] Voorhees R M, Fernandez I S, Scheres S H and Hegde R S 2014 Structure of the mammalian ribosome-Sec61 complex to 3.4 A resolution *Cell* **157** 1632-43

[85] Wei X, Su X, Cao P, Liu X, Chang W, Li M, Zhang X and Liu Z 2016 Structure of spinach photosystem II-LHCII supercomplex at 3.2 A resolution *Nature* **534** 69-74

[86] Wolf M, DeRosier D J and Grigorieff N 2006 Ewald sphere correction for single-particle electron microscopy *Ultramicroscopy* **106** 376-82

[87] Yan C, Hang J, Wan R, Huang M, Wong C C and Shi Y 2015 Structure of a yeast spliceosome at 3.6-angstrom resolution *Science (New York, N.Y* **349** 1182-91

[88] Yan R, Edwards T J, Pankratz L M, Kuhn R J, Lanman J K, Liu J and Jiang W 2015 Simultaneous determination of sample thickness, tilt, and electron mean free path using tomographic tilt images based on Beer-Lambert law *Journal of structural biology* **192** 287-96

[89] Yoshioka C, Carragher B and Potter C S 2010 Cryomesh: a new substrate for cryo-electron microscopy *Microsc Microanal* **16** 43-53

[90] Yu G, Li K and Jiang W 2016 Antibody-based affinity cryo-EM grid *Methods*

[91] Zewail A H 2010 Four-dimensional electron microscopy *Science (New York, N.Y* **328** 187-93

[92] Zhang K 2016 Gctf: Real-time CTF determination and correction *Journal of structural biology* **193** 1-12

[93] Zhang K, Foster H E, Rondelet A, Lacey S E, Bahi-Buisson N, Bird A W and Carter A P 2017 Cryo-EM Reveals How Human Cytoplasmic Dynein Is Auto-inhibited and Activated *Cell* **169** 1303-14 e18

[94] Zheng S Q, Palovcak E, Armache J P, Verba K A, Cheng Y and Agard D A 2017 MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy *Nature methods* **14** 331-2

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 March 2018 doi:10.20944/preprints201803.0043.v1

Figures

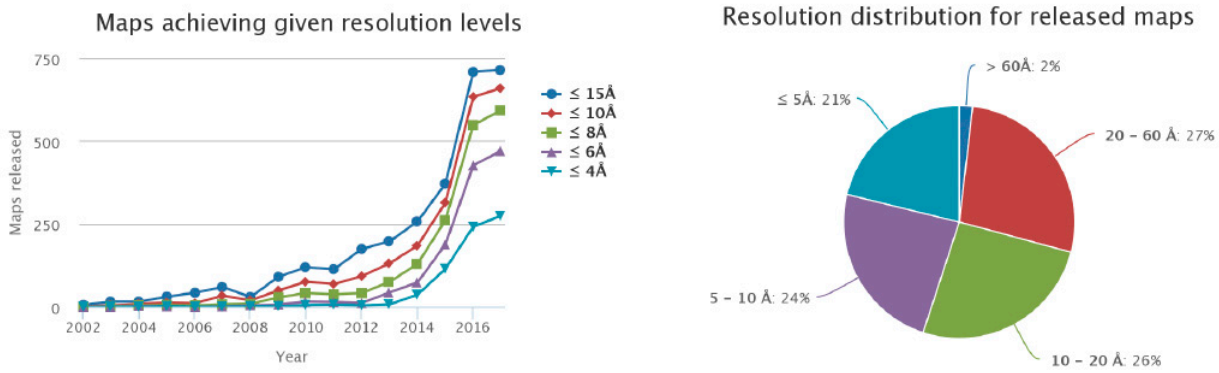

Figure 1. Statistics of the electron density maps deposited in EMDB (Electron Microscopy Data Bank). The annual numbers of released maps with different reported resolutions are plotted from 2002 to 2017 (left). And the distribution of all the released maps till 2017 is statistically plotted vs different resolutions (right). Both panels were generated using the tool of EMDB (http://www.ebi.ac.uk/pdbe/emdb).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 March 2018 doi:10.20944/preprints201803.0043.v1

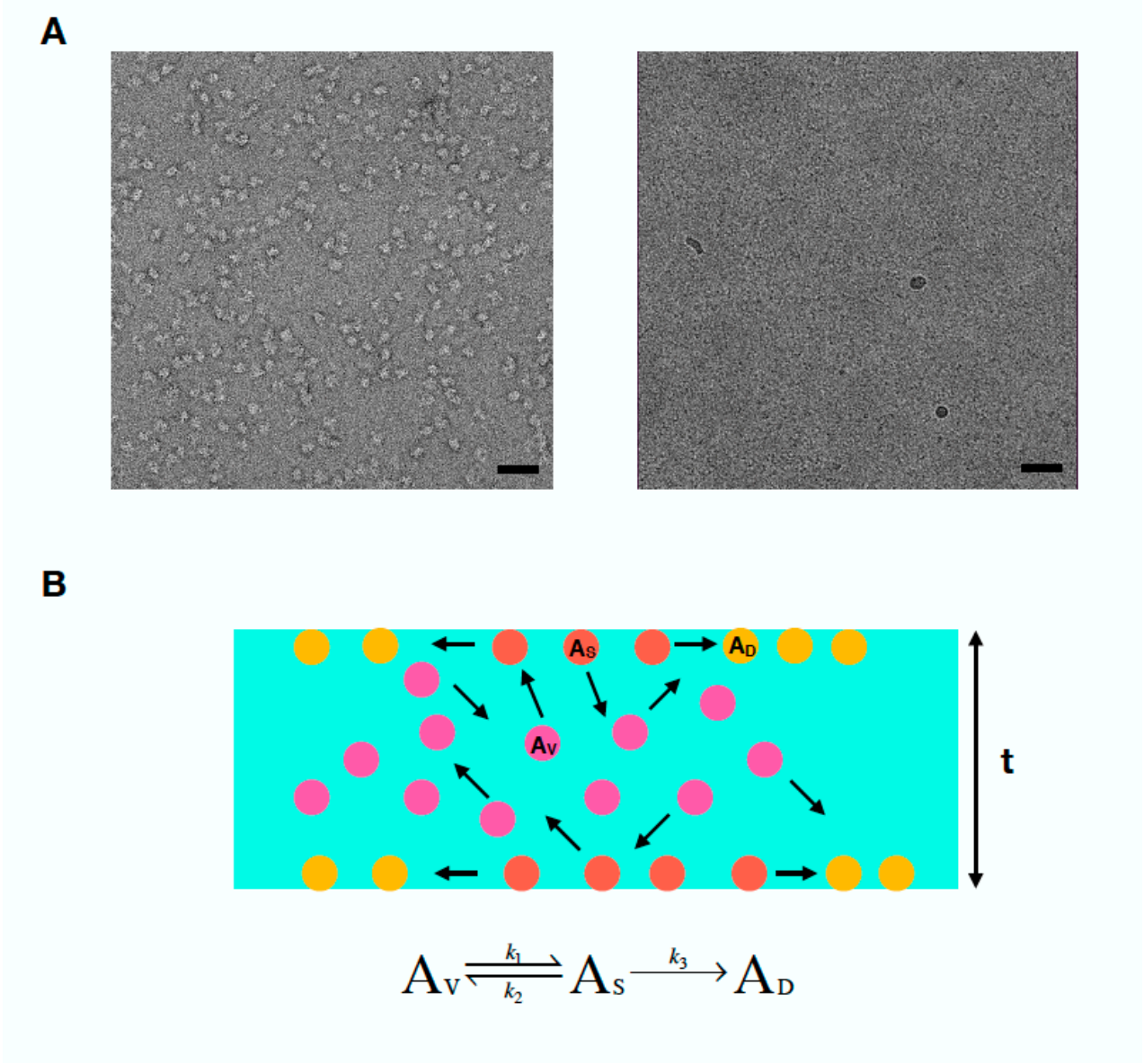

Figure 2. Air-water interface effects during specimen cryo-vitrification. **A**. The bio-macromolecule specimen exhibits homogenous distribution and good shape in negative electron microscopy (left), but is prone to degrade and hard to observe in cryo-electron microscopy (right). Scale bar, 50 nm. **B**. A diagram showing the physical process during specimen cryo-vitrification. Bio-macromolecules in native state (A_V) are colored in pink, those (A_S) absorbed to air-water interface are colored in orange, and that (A_D) denatured in yellow. The thickness (t) of the solution layer after blotting is between 30 ~ 100 nm.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 March 2018 doi:10.20944/preprints201803.0043.v1

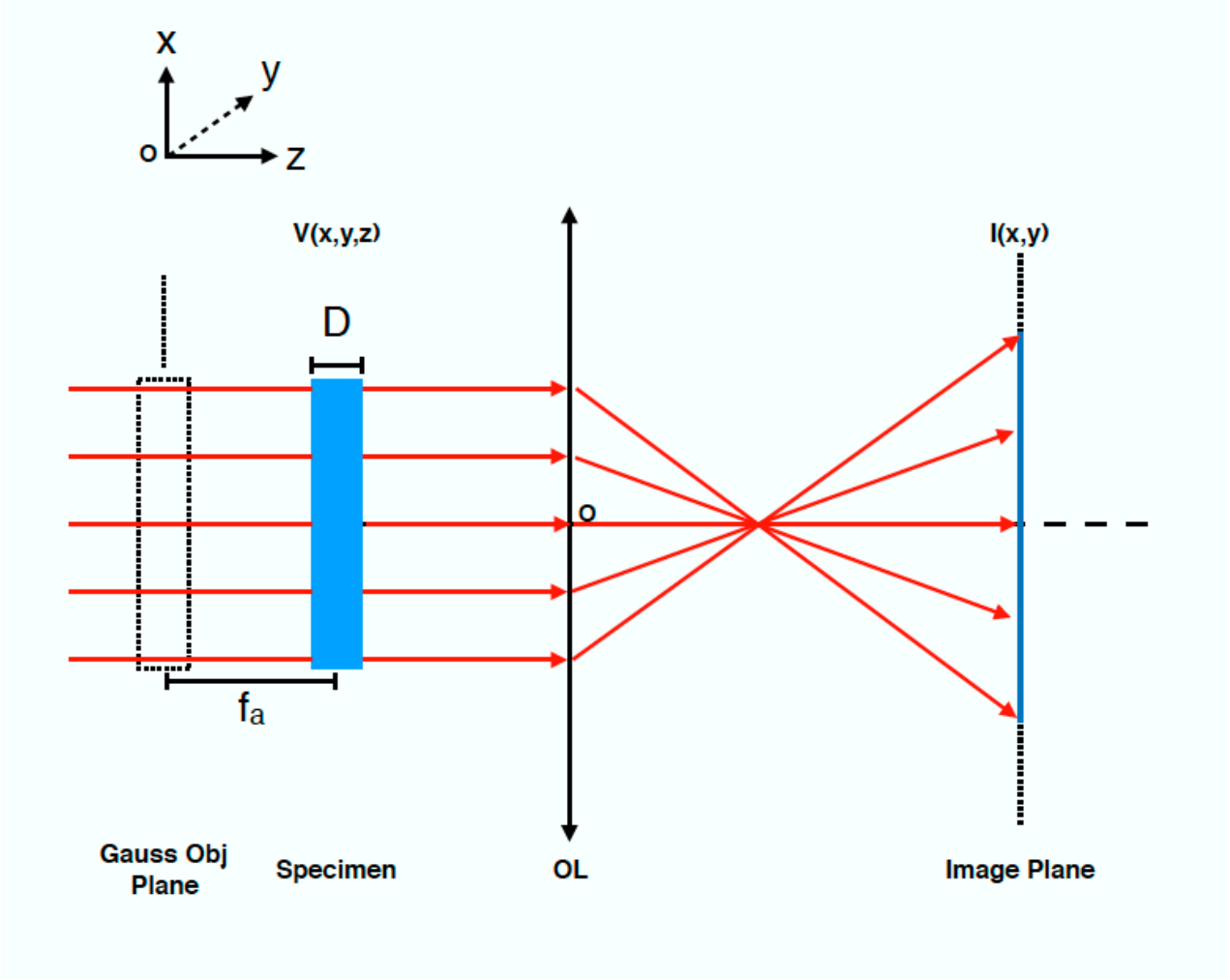

Figure 3. A diagram showing the image formation of a think specimen with the averaged underfocus f_a. Gauss Obj Plane, the plane where the idea objective locates according to Gauss image formation formula. D, the thickness of the specimen. OL, objective lens. V(x,y,z), the density function of the specimen. I(x,y), the image function at the image plane.