

Contents lists available at [SciVerse ScienceDirect](#)

Journal of Structural Biology

journal homepage: www.elsevier.com/locate/yjsbi

Image formation modeling in cryo-electron microscopy

Miloš Vulović^{a,b}, Raimond B.G. Ravelli^b, Lucas J. van Vliet^a, Abraham J. Koster^b, Ivan Lazić^c, Uwe Lücken^c, Hans Rullgård^d, Ozan Öktem^{e,1}, Bernd Rieger^{a,*,1}

^aQuantitative Imaging Group, Faculty of Applied Sciences, Delft University of Technology, Lorentzweg 1, 2628 CJ Delft, The Netherlands

^bElectron Microscopy Section, Molecular Cell Biology, Leiden University Medical Center, P.O. Box 9600, 2300 RC Leiden, The Netherlands

^cFEI Company, Achtseweg Noord 5, 5651 GG Eindhoven, The Netherlands

^dCOMSOL, Tegnérgatan 23, 11140 Stockholm, Sweden

^eCIAM, Department of Mathematics, KTH – Royal Institute of Technology, 10044 Stockholm, Sweden

ARTICLE INFO

Article history:

Received 23 February 2013

Received in revised form 7 May 2013

Accepted 14 May 2013

Available online xxx

Keywords:

Cryo-electron microscopy

Phase contrast

Amplitude contrast

Image simulation

InSilicoTEM

Interaction potential

ABSTRACT

Accurate modeling of image formation in cryo-electron microscopy is an important requirement for quantitative image interpretation and optimization of the data acquisition strategy. Here we present a forward model that accounts for the specimen's scattering properties, microscope optics, and detector response. The specimen interaction potential is calculated with the isolated atom superposition approximation (IASA) and extended with the influences of solvent's dielectric and ionic properties as well as the molecular electrostatic distribution. We account for an effective charge redistribution via the Poisson–Boltzmann approach and find that the IASA-based potential forms the dominant part of the interaction potential, as the contribution of the redistribution is less than 10%. The electron wave is propagated through the specimen by a multislice approach and the influence of the optics is included via the contrast transfer function. We incorporate the detective quantum efficiency of the camera due to the difference between signal and noise transfer characteristics, instead of using only the modulation transfer function. The full model was validated against experimental images of 20S proteasome, hemoglobin, and GroEL. The simulations adequately predict the effects of phase contrast, changes due to the integrated electron flux, thickness, inelastic scattering, detective quantum efficiency and acceleration voltage. We suggest that beam-induced specimen movements are relevant in the experiments whereas the influence of the solvent amorphousness can be neglected. All simulation parameters are based on physical principles and, when necessary, experimentally determined.

© 2013 Elsevier Inc. All rights reserved.

1. Introduction

The structures of macromolecules, macromolecular complexes and subcellular assemblies provide insight into their functions. Knowledge of the 3D structure of a macromolecule is also the cornerstone for rational drug design (Wang et al., 2005).

Cryo-electron microscopy (cryo-EM) of biological specimens in an unstained, frozen-hydrated state has become an indispensable tool for structural biology (Sali et al., 2003). Advances in cryo-EM single particle analysis (SPA) (Frank, 2006) and cryo-electron tomography (cryo-ET) (Lucić et al., 2005; McIntosh et al., 2005; Leis et al., 2009) provide opportunities to characterize the structures of macromolecular complexes that are either too flexible, heterogeneous or transient to be explored by crystallographic methods (Henderson, 2004; Glaeser et al., 2006). The level of structural detail that can be obtained by cryo-EM is largely limited by specimen

heterogeneity, the effective contrast transfer function (CTF), the detector's detective quantum efficiency (DQE), and radiation damage which limits the integrated electron flux that can be used, resulting in a poor signal-to-noise ratio (SNR) in images.

In addition to hardware developments, computational methods will continue to improve, enabling more information to be extracted from inherently noisy cryo-EM images. Simulations of electron images will be increasingly important in order to optimize the data acquisition strategy, to improve image interpretation and resolution, and to provide insight on ways to improve instrumentation. An accurate forward model of image formation in cryo-EM should rely on all relevant physical properties such as the specimen's elastic and inelastic scattering properties and the effects of the CTF and the detector.

Simulation of transmission electron microscope (TEM) images of biological specimens is implemented in a number of software packages for SPA and ET such as Xmipp (Sorzano et al., 2004; Bilbao-Castro et al., 2004), IMAGIC (van Heel et al., 1996), SPIDER (Frank and Shimkin, 1978; Shaikh et al., 2008), EMAN2 (Tang et al., 2007), Bsoft (Heymann and Belnap, 2007), and TOMToolbox

* Corresponding author. Fax: +31 15 2786740.

E-mail address: b.rieger@tudelft.nl (B. Rieger).

¹ These authors contributed equally to this work.

(Nickell et al., 2005). In most cases, these simulations are used to facilitate Euler angles determination in SPA and to evaluate reconstruction methods for SPA (Marabini et al., 1998; Sorzano et al., 2001) and ET (Marabini et al., 1997). Usually a virtual model of a biological specimen is created using 3D primitives (phantoms) such as spheres, ellipsoids, cubes, and cylinders (Bilbao-Castro et al., 2004). In some cases, the specimen volume is constructed based on information from the RCSB Protein Data Bank (PDB) and TEM images are computed by projecting the 3D specimen; the effects of the solvent and detector are rarely accounted for. In general, projecting the 3D electron density distribution into a 2D image is not correct, since it does not represent the actual physical electron–specimen scattering properties (interaction potential). In addition, the noise is often simplified as being additive Gaussian noise. Below, we discuss two related work that aim to provide more realistic simulations.

In Hall et al. (2011), image simulations were performed to assess the attainable benefits of phase plates. The solvent (water) was treated explicitly via molecular dynamics (MD) simulations generating a box of amorphous water and a multislice approach was used to account for the specimen thickness and multiple scattering. The generated noise was Poisson distributed, but the detector response was not included. Unfortunately, the methods were not validated experimentally.

TEM-simulator (Rullgård et al., 2011) aims to provide accurate simulations based on physical principles. It was the first simulator whose results were compared to experimental data, albeit not in depth. There, the specimen thickness has been neglected, and low-pass filtering to a certain resolution exceedingly damps the interaction potential (IP). Although most simulation parameters described there are based on physical principles, a calibration protocol needs to be employed for some parameters that are phenomenologically introduced, leading to a situation where nuisance parameter tuning is required. Examples of such phenomenological parameters are amorphousness (granularity), absorption potential, as well as camera parameters such as the modulation-transfer function (MTF), detective quantum efficiency (DQE), and conversion factor. Furthermore, none of the aforementioned approaches have considered chemical bonding and/or interaction of the sample with solvent and ions. For completeness, it should be mentioned that recently Shang and Sigworth (2012) parameterized a function that describes the distribution of water molecules around a protein. In previous work the solvent was assumed to be water, instead of less dense vitreous ice, leading to possible artificial damping of the contrast between the protein (which has a higher density than water) and solvent.

For material science applications, numerous TEM simulators have been developed (reviewed by Kirkland, 2010). Many assume that the atoms of a specimen are periodically ordered which is not fulfilled for non-crystalline biological specimens. Some of the simulators, such as YAMS (Dinges and Rose, 1995; Müller et al., 1998) and SimulaTEM (Gómez-Rodríguez et al., 2009), have been used for image simulations of biological specimens. They do not assume that the specimen is periodic and although YAMS propagates the mutual coherence function through the specimen, a method more appropriate for treating the partial incoherence, only elastic scattering was assumed for biological specimens (Sorzano et al., 2001). In both simulators the specimen thickness and multiple scattering events were treated via a multislice approach (Cowley and Moodie, 1957), but inelastic scattering, the detector response, and solvent were ignored. In high resolution electron microscopy (HREM) the contrast in experimental images has been frequently reported to be much less, typically about a factor of three, than predicted by image simulation (Hytch and Stobbs, 1994; Boothroyd et al., 1995). It was suggested in Thust (2009) that this discrepancy,

often called the Stobbs-factor, originates from neglecting the detector's MTF in image simulations.

Here we present, analyze and validate an image formation model in TEM based on physical principles. In addition to computing the 3D potential distribution where atoms are treated in isolation, the interaction redistribution potential due to the solvent, ions and molecular interactions is computed. Beam-induced motion and amorphousness of the vitreous ice are also addressed. For validation, comparisons between experiments and simulations were performed on cryo-embedded specimens. Some of the parameters such as defocus, astigmatism and camera properties are accurately estimated from experiments via available toolboxes (Vulović et al., 2010, 2012), without introducing nuisance parameters. The simulator presented here, InSilicoTEM, has been implemented in DIPimage (www.DIPlib.org), a MATLAB toolbox for scientific image processing and analysis, and is freely available for non-commercial use upon request.

2. Theory

Forward modeling approaches in cryo-EM describe the complex image formation process. Below, we will shortly outline our image formation model whose main ingredients are: the interaction potential, electron wave propagation and intensity detection by the camera. The [Supplementary material \(S.M.\)](#) provides a detailed description of all steps and approximations.

2.1. Interaction potential (IP)

The interaction between the incident electron wave and a macromolecule embedded in the surrounding medium is modeled as a sum of two interaction potential components: (1) “atom” contributions, i.e. the superposition of atomic potentials as if each atom was in isolation; and (2) “bond” contributions, i.e. the influence of the charge redistribution due to the solvent, ions and molecular interactions

$$\mathcal{V}^{\text{int}}(\mathbf{r}) = \mathcal{V}_{\text{atom}}(\mathbf{r}) + \mathcal{V}_{\text{bond}}(\mathbf{r}), \quad (1)$$

where $\mathbf{r} = (x, y, z)$ is the position of the electron wave. Since $\mathcal{V}_{\text{atom}}$ considers the specimen as a set of isolated atoms, we get $\mathcal{V}_{\text{atom}}(\mathbf{r}) = \sum_{j=1}^m \mathcal{V}_{Z_j}(\mathbf{r} - \mathbf{R}_j)$, where \mathcal{V}_{Z_j} is the electrostatic potential of an isolated neutral atom with atomic number Z_j centered at \mathbf{R}_j . With the first Born approximation, such a potential can be written as the inverse Fourier transform of the electron scattering factor of the atom (Peng et al., 2004; Rullgård et al., 2011) (see [Section 1.2 in S.M.](#)).

The isolated atom superposition approximation (IASA) ignores the potential due to the charge redistributions, $\mathcal{V}_{\text{bond}}$, which accounts for the interaction with neighboring atoms, solvent and ions. As $\mathcal{V}_{\text{atom}}$ provides the most significant contribution to the scattering of the incident electron, this computationally convenient approximation provides a good starting point for initial interpretation of high-energy electron diffraction and microscopy experiments (Kirkland, 2010; Peng et al., 1996). Biological specimens are embedded in an amorphous solvent and the potential distribution depends also on the dielectric and ionic properties of the solvent. It seems appropriate to include the contribution of the solvent and ions modeled by $\mathcal{V}_{\text{bond}}$. This potential due to the charge redistribution can be accounted for via a continuum electrostatics approach (see [1.3 in S.M.](#)), described by the solution of the linearized Poisson–Boltzmann equation:

$$-\epsilon_0 \nabla \cdot (\epsilon_r(\mathbf{r}) \nabla \mathcal{V}_{\text{bond}}(\mathbf{r})) = \rho_{\text{mol}}^{\text{bond}}(\mathbf{r}) + \rho_{\text{sol}}^{\text{bond}}(\mathbf{r}) - \alpha(\mathbf{r}) \sum_i \frac{q_i^2 n_i^0 \mathcal{V}_{\text{bond}}(\mathbf{r})}{k_B T}, \quad (2)$$

where ϵ_0 is the permittivity of the vacuum, ϵ_r the relative permittivity (ϵ_{sol} for the sub-volume occupied by solvent and ϵ_{mol} for the molecule), $\rho_{\text{mol}}^{\text{bond}}(\mathbf{r})$ and $\rho_{\text{sol}}^{\text{bond}}(\mathbf{r})$ are the partial (net) charges of the molecule and solvent, respectively; q_i and n_i^0 are, respectively the charge and the concentration of an ion of type i ; k_B the Boltzmann constant, T the temperature, and $\alpha(\mathbf{r}) = 1$ for sub-volume occupied by solvent, and $\alpha(\mathbf{r}) = 0$ otherwise. In order to meaningfully add the two potential contributions (Eq. (1)), the assumptions specified in 1.4 in S.M. must be fulfilled.

2.1.1. Inelastic contributions

The effects of inelastic scattering are modeled as the imaginary part of the interaction potential. The total complex potential is $\mathcal{V}_{\text{tot}}^{\text{int}} = V_{\text{ph}} + iV_{\text{ab}}$ (see 2.3 in S.M.), where V_{ph} is the interaction potential (real value) as described in the previous section. V_{ph} contributes to the phase contrast while V_{ab} influences the amplitude (absorption) contrast. Contributions to the amplitude contrast can be roughly separated into “plasmons”, electrons scattered outside the aperture, and atom core losses. “Plasmons” are not strictly oscillations of free electrons like in metals, but they are producing a similar amount of energy loss (~ 20 eV), hence this commonly used terminology (Egerton, 1996). In a typical electron energy-loss spectroscopy (EELS) spectrum, the intensities due to atom core losses ($\Delta E > 100$ eV) are a couple of order of magnitudes smaller than those of the plasmons ($\Delta E \sim 20$ eV). The influence of the aperture will be taken into account via the optical system. Therefore, the plasmons are considered the most dominant contribution to the inelastic interactions. Since a large part of the specimen consists of embedding medium, the plasmons of vitreous ice damp the useful phase signal. For an amorphous solvent such as vitreous ice and a certain incident electron energy the plasmons can be described via the inelastic mean free path Λ_{in} . We performed Monte-Carlo simulations (Agostinelli et al., 2003; Kieft and Bosch, 2008) to validate the assumption that delocalized processes (represented by a constant absorption potential) are dominant (see 2.3.1 in S.M.). For our purpose the imaginary part of the potential is modeled as

$$V_{\text{ab}}(x, y, z) = 1/(2\sigma\Lambda_{\text{in}}), \quad (3)$$

where $\sigma = 2\pi m |e| \lambda/h^2$ is the interaction constant, λ , e , and m the relativistic wavelength, charge and mass of the incident electron, and h Planck's constant.

2.2. Electron wave propagation

The electron wave propagation through the specimen is based on a multislice method (Cowley and Moodie, 1957) that accounts for the thickness of the specimen and multiple scattering (Kirkland, 2010). An incident electron is described by its wave function and at the top of the $(n + 1)$ th slice of the specimen, the wave function is given by

$$\Psi_{n+1}(x, y) = \mathcal{F}^{-1}[P_n(q_x, q_y, \Delta z_n)\mathcal{F}[\exp(i\sigma V_z(x, y, z))\Psi_n(x, y)]], \quad (4)$$

where $P(q, \Delta z) = \exp(-i\pi\lambda\Delta zq^2)$ is the Fresnel propagator over a slice of thickness Δz , q is the magnitude of the spatial frequency (q_x, q_y) , $\mathcal{F}[\cdot]$ denotes the Fourier transformation, and $V_z(x, y, z) = \int_z^{z+\Delta z} \mathcal{V}^{\text{int}}(x, y, z')dz'$ is the projected potential within the slice. Parallel illumination is modeled as an incident plane wave ($\Psi_0(x, y) = 1$). The propagation of the electron wave through the specimen can be interpreted as recursive transmission and propagation of the wave function through each slice until the wave leaves the specimen ($\Psi_{\text{exit}}(x, y)$).

In cryo-EM the images are mostly generated by phase contrast, as a result of interference between the unscattered and scattered part of the electron exit wave function. The electron wave exiting the specimen $\Psi_{\text{exit}}(x, y)$ is further subject to a frequency dependent

phase shift introduced by the defocus Δf and microscope aberrations such as spherical aberration C_s , and 2-fold astigmatism (A_1, α_1) . The contrast transfer function (CTF) of the lens system in polar coordinates is (Reimer and Kohl, 2008)

$$T(q, \alpha) = KA_p \exp\left(-i\frac{2\pi}{\lambda}\left(\frac{1}{4}C_s\lambda^4q^4 - \frac{1}{2}(\Delta f - A_1 \cos(2(\alpha - \alpha_1)))\lambda^2q^2\right)\right), \quad (5)$$

where A_p is the objective aperture function and K describes spatial and chromatic envelopes. Note that underfocus implies $\Delta f > 0$, as in Reimer and Kohl (2008). The intensity in the image plane is the probability density function given by

$$I_0(x, y) = |\Psi(x, y)|^2 = |\mathcal{F}^{-1}[\mathcal{F}[\Psi_{\text{exit}}(x, y)]T(q, \alpha)]|^2. \quad (6)$$

For details see Section 3 in S.M.

2.3. Detector response

Capturing the final image involves the conversion of the electron wave's intensity distribution into a digital signal via a detector. Electron detectors are characterized by parameters such as conversion factor CF in $[ADU/e^-]$, modulation transfer function (MTF) and detective quantum efficiency (DQE). The measurement process obeys Poisson statistics giving rise to shot noise; the detector adds readout noise I_{rn} and dark current I_{dc} to the final image, and blurs the image with a detector point spread function PSF (x, y) whose Fourier transform is the MTF.

The MTF describes transfer of the signal amplitude for different spatial frequencies. However, the signal and the noise in a TEM detector are not transferred in the same way (Meyer and Kirkland, 1998). The DQE is defined as $DQE(q) = \text{MTF}^2(q)/\text{NTF}^2(q)$, where the NTF is the noise transfer function ($\text{NTF}^2(q) = \text{NPS}_{\text{out}}/(CF^2\Phi_e)$) with NPS being the noise power spectrum, and Φ_e the incident electron flux in $[e^-/\text{area}]$. We model the signal and noise propagation as follows: (1) the Fourier spectrum of the noise-free signal ($\tilde{I}_0(q)$) is damped (multiplied) by the ratio between signal (MTF) and noise (NTF) transfer, (2) this signal is multiplied by the integrated electron flux and shot noise contributions are added, (3) the Fourier spectrum of that (noisy) signal is damped by the NTF, and (4) the number of electrons are scaled with CF to the image gray values in $[ADU]$. Hence, we can write the detected image as

$$I(x, y) = I_{\text{rn}} + I_{\text{dc}} + CF \cdot \mathcal{F}^{-1}\left[\mathcal{F}\left[P_{\text{ois}}\left(\Phi_e \cdot \mathcal{F}^{-1}\left[\tilde{I}_0(q)\sqrt{\text{DQE}(q)}\right]\right)\right] \cdot \text{NTF}(q)\right], \quad (7)$$

where $P_{\text{ois}}(A)$ returns a random number from a Poisson distribution with expected value A . Section 4 in S.M. explains the steps in more detail.

3. Computational methods

The main steps of image formation simulations are (i) construction of the interaction potential (IP) and (ii) electron wave propagation and recording intensity.

Physical parameters of the specimen include pH, dielectric constant, temperature, ion concentration, motion factor and thickness. Microscope parameters involve acceleration voltage and its spread, opening angle, defocus, astigmatism, spherical and chromatic aberrations, objective aperture, magnification, and incident electron flux. Relevant camera parameters are exposure time, binning, conversion factor, MTF, DQE, readout and dark current noise. All parameters influencing the image formation are based on physical principles and when necessary, they were estimated from the experiment, using independent measurements (except beam-induced movements), without introducing nuisance parameters.

In this section, we outline the computational methods for image simulation of biological specimens and parameters estimation.

3.1. Interaction potential (IP)

A forward simulation requires a known model of the specimen. In case of biological specimens, we construct the IP using a hybrid approach combining the isolated atom superposition approximation (IASA) and a Poisson–Boltzmann (PB) description of the interaction between the macromolecule and its solvent and ions (see Section 2.1). The input for computing the IP is a high-resolution X-ray structure as deposited within the PDB which contains a detailed specification of type and position of most atoms in the molecule. Here we used PDB files 1RYP, 1GR5 and 2GTL to model 20S proteasome, GroEL, and earthworm hemoglobin, respectively. The plasmons are accounted for via the inelastic mean free path. The amorphousness of the solvent was generated by an explicit atomic model via MD simulations. An empirical post-blurring can be applied, which results in a similar effect that beam-induced movements could have. The next subsections explain these procedures in more detail.

3.1.1. Isolated atom superposition approximation (IASA)

The dominant part of the interaction potential is the sum of the individual isolated atomic potentials calculated as the Fourier transforms of tabulated electron scattering factors. There are several empirical closed-form approximations of electron scattering factors available (Kirkland, 2010). We use scattering factors that are parameterized as a weighted sum of five Gaussians as given in Table 1 in Peng et al. (1996) and implemented in TEM-simulator (Rullgård et al., 2011). The real potential map calculations are based on a slight modification of their map in such a way that low-pass filtering to a certain resolution does not exceedingly damp the interaction potential (IP), and the solvent is assumed to be vitreous ice instead of water (see Eq. (5) in S.M.). The input PDB file is converted into the electrostatic potential map ν_{atom} . The voxel size of the map in this analysis was set to 1 Å.

3.1.2. The influence of the embedding environment via the Poisson–Boltzmann (PB) approach

We use a continuous electrostatics method to model the influence of the solvent and ions as well as the coarse electrostatic potential redistributions within the macromolecule. There are different implementations for solving the Poisson–Boltzmann (PB) equation. In this study we used APBS (adaptive Poisson–Boltzmann solver) (Baker et al., 2001) that numerically solves the PB equation for solvation energy and potential.

The input for APBS is a modified PDB file (PQR) where the occupancy and temperature fields are substituted with partial charges and the radii fields using PDB2PQR (Dolinsky et al., 2004). Since protein structures deposited in the PDB format usually lack hydrogen atoms, PDB2PQR offers the functionality of adding missing hydrogens atoms and removing steric clashes caused by the newly added hydrogens. Partial (net) charges were calculated with AMBER, one of the forcefields available in PDB2PQR. The pH value

was set to 7. The APBS input file contains both numerical aspects of the computation and physical parameters (c.f. Table 1).

For large molecules such as earthworm hemoglobin we adapted the procedure for calculating PQR files and APBS potentials. For large (constructed) PDB files we adopted a variable column width. The parsers also allowed a more flexible spacing between all fields and larger (unrestricted) field size. In order to calculate the potential map of hemoglobin, the molecule was split into eight parts with an overlap of 10%. A single potential map was assembled from all parts. 20S Proteasome was simulated without ions in the solvent while the ion concentrations for the earthworm hemoglobin sample were 0.05 M, 1 M, and 3 M (mol/l), respectively.

In contrast to typical PB solvers that include two-step solvation energy calculations, here we used a one-step approach. For chemistry and biophysics applications, the reaction fields due to the polarization of the solvent and ions around the molecule are of interest and a two-step approach is needed. In that case, homogeneous dielectric calculations (dielectric constants of the molecule and solvent are equal) are subtracted from heterogeneous calculations (dielectric constants differ). Since we are interested in the electrostatic potential redistribution within the molecule, we did not need to perform homogeneous dielectric calculations (see 1.4 in S.M.), resulting in reduced computation times.

3.1.3. Inelastic scattering

For our purpose the imaginary part of the potential was modeled via the inelastic mean free path (see Eq. (3)). As described in Langmore and Smith (1992) and Wall et al. (1974) the inelastic scattering cross section σ_{in} can be represented via Eq. (26) in S.M. The inelastic mean free path is related to the scattering cross section as

$$\Lambda_{\text{in}} = \frac{M_{\text{W}}}{\rho N_{\text{A}} \sigma_{\text{in}}} = \frac{M_{\text{W}} \beta^2 \cdot 10^{10}}{9.03 \rho Z^2 \ln \frac{\beta^2 (U_0 + mc^2)}{10}} [\text{nm}], \quad (8)$$

where Z is the atomic number, β the ratio between the velocity of electron and light ($\beta^2 = 1 - [mc^2/(U_0 + mc^2)]^2$), U_0 the incident electron energy, mc^2 the rest energy of electron, M_{W} the molar mass, ρ the mass density, and N_{A} Avogadro's number. The dependency of the mean free path on the incident electron energy is given by (8) and plotted in Fig. 1. Experimentally determined values of the

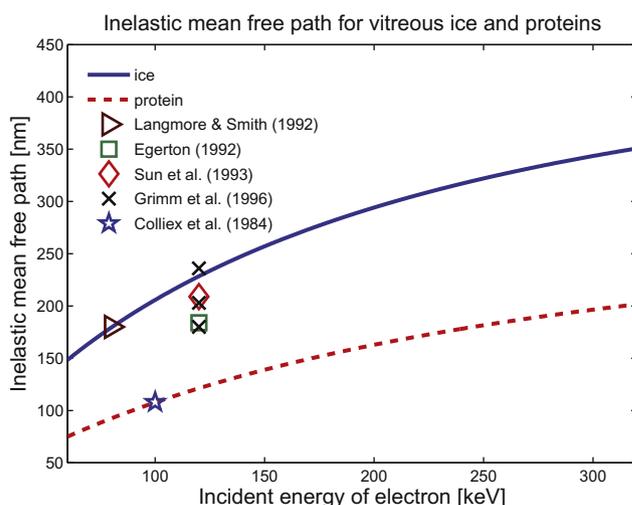


Fig. 1. Inelastic mean free path as a function of the incident electron energy for vitreous ice (blue solid) and protein (red dashed line) using Eq. (8) and data derived from Langmore and Smith (1992) and Colliex et al. (1984). The data points from references Langmore and Smith (1992), Egerton (1992), Sun et al. (1993), Grimm et al. (1996) and Colliex et al. (1984) are included.

Table 1
Some APBS parameters.

| Symbol | Value | Meaning |
|--------|--------|---|
| pdie | 2.00 | dielectric constant of the solute |
| sdie | 78.54 | dielectric constant of the solvent |
| temp | 298.15 | temperature of the system (K) |
| srad | 1.40 | radius of the solvent molecules (Å) |
| ion | +1 0 2 | ion species, concentration (M) and radius (Å) |

inelastic mean free path reported in the literature vary noticeably (Feja and Aebi, 1999). The reasons for these apparent discrepancies are not always clear. Some of the reported values for a couple of energies are included in Fig. 1.

The fractional composition of a protein was taken to be 0.492, 0.313, 0.094, and 0.101 for elements H, C, N, and O, respectively (Colliex et al., 1984; Reichelt and Engel, 1984). We used the values of Λ_{in} for vitreous ice at 80 kV and protein at 100 kV provided by Langmore and Smith (1992) and Colliex et al. (1984), respectively. The values for any other incident energy of electron U_0 were calculated via Eq. (8).

3.1.4. Amorphousness of the solvent - specimen

As described in Sections 3.1.1 and 3.1.2, the solvent has been modeled as a continuum. Although its influence on the potential redistribution of the macromolecule is accounted for (Section 3.1.2), the solvent potential is calculated from the known density of water molecules using an averaging procedure (see Eq. (5) in S.M.). However, amorphousness of the solvent can influence the appearance of the noise in the image. In order to assess the influence of the amorphousness in cryo-EM under low-flux conditions or to allow one to model it for high fluxes in HREM, we propose two methods for modeling this amorphousness: (i) adding a fixed noise pattern to the specimen's projected potential, and (ii) performing molecular dynamics (MD) simulations via GROMACS (Hess et al., 2008).

(i) Adding a fixed noise pattern to the projections:

This simple method assumes that the overlap of atomic positions in a projection of an amorphous sample is significant and that it is essentially noise with a flat frequency spectrum. This is surely an approximation as every real specimen has limited scattering power. Therefore, we multiply the frequency spectrum by $\exp(-2\pi(qr_d/\Delta_{x,y})^2)$ where q is the spatial frequency, r_d is the average minimum distance between atoms in the amorphous specimen and $\Delta_{x,y}$ is the pixel size in the object plane. The covalent sp^3 radius in carbon is 0.77 Å (Allen et al., 1987), and a model of amorphous carbon should thus have a minimum distance of $r_d = 1.54$ Å. For vitreous ice, the distance between oxygen atoms would be 2.88 Å (Teeter, 1984).

(ii) MD simulations:

In order to produce an explicit description of the solvent (water), we used GROMACS (Hess et al., 2008), a MD simulation package which solves Newton's equations of motion for a system of N interacting atoms. The equations are solved simultaneously in small time steps reaching an equilibrium state of the system. The input was PDB file 1GR5 (GroEL). The missing hydrogen atoms were added and a topology file was generated containing the physical information about all interactions between the atoms of the protein (bonds, angles, torsion angles, Lennard-Jones interactions and electrostatic interactions). Furthermore, the protein was solvated in a $20 \times 20 \times 50$ nm water box with a simple point charge (SPC) 216 model. The specimen box consists of a multitude of small boxes, each containing 216 water molecules. In order to circumvent a crystalline arrangement of small water boxes, energy minimization was performed followed by a short MD simulation (20 ps), effectively randomizing the solvent molecules positions and solvating the protein.

3.1.5. Beam-induced movements

Beam-induced movements can significantly influence the contrast in cryo-EM (Henderson and Glaeser, 1985; Brilot et al., 2012). The whole layer of ice encapsulating the macromolecule

seems to deform upon exposure in a complicated manner. Here, we model these effects empirically via an isotropic motion factor σ_M , which blurs the IP as follows:

$$\tilde{V}(q) = \tilde{V}^{\text{int}}(q) \exp(-2\pi^2 \sigma_M^2 q^2), \quad (9)$$

where $\tilde{V}(q)$ and $\tilde{V}^{\text{int}}(q)$ are the Fourier transforms of the potential $V(r)$ and $V^{\text{int}}(r)$, respectively. This is equivalent to damping of the electron scattering factors in the Fourier domain.

3.2. Electron wave propagation and intensity detection

The incident electron plane wave is propagated through the specimen by a multislice approach inspired by Kirkland (2010). The slice thickness was kept constant at ~ 2 nm. The effective projected potential within this slice thickness in all our simulations $\sigma V_z(\mathbf{r})$ proved to be smaller than 0.36 suggesting that, within a slice, the probability of multiple scattering events is less than 5% and that the weak-phase object approximation and projection assumption are valid (Vulović et al., submitted for publication). As described in Section 2, the CTF accounts for all relevant microscope aberrations, apertures and partial coherence of the electron source. Finally, the image intensity is captured by the detector modeled by the MTF, DQE and various noise sources.

3.3. Parameter estimations

Some imaging parameters vary between acquisitions, while others are stable for a long period of time. To accurately model image formation and validate it with experimental data, we need to know the numerical values of all parameters that influence image formation (see Section 2). The detector parameters are characterized independently via methods described in Vulović et al. (2010). The parameters that must be determined during the data acquisition are magnification, integrated flux, defocus, astigmatism and local ice thickness. The magnification of the microscope was calibrated prior to the acquisitions with a cross grating containing 2160 lines per mm. The integrated electron flux in [$\text{e}^-/\text{Å}$] was estimated from the measured intensities in areas without specimen using the conversion factor of the detector. For each low-flux cryo image, an image of an adjacent carbon support was acquired to accurately measure defocus and astigmatism as well as their uncertainties using the publicly available toolbox described in Vulović et al. (2012). Measurements of the local ice thickness d are based on the Beer-Lambert law and were calculated from the ratio of the integrated intensity of an EELS zero-loss peak I_{z1} relative to the integral of the whole spectrum I . Similar to Eq. (30) in S.M. we have

$$d = \Lambda_{\text{in}} \ln \frac{I}{I_{z1}}. \quad (10)$$

4. Experimental methods

In order to validate our image formation model, cryo-EM experiments were carried out using various test samples and experimental conditions. Numerous defocus and flux series of unfiltered and zero-loss energy filtered images of 20S proteasome and hemoglobin were acquired at 80 kV and 300 kV.

4.1. Sample preparation

Our modeling approach was evaluated with 20S proteasome from *S. cerevisiae*, *Lumbricus terrestris erythrocrurin* (earthworm hemoglobin) and GroEL. Proteasome (Sigma Aldrich, 10 mg/ml) was diluted 10-fold in 50 mM HEPES pH 7.4, 150 mM NaCl, and

1 mM DTT. The hemoglobin sample (Hb) was prepared by diluting the hemoglobin stock solution 25-fold in 50 mM NH₄Ac pH 6.6 (a protocol adapted from Vinogradov and Sharma (1994), Karuppasamy et al. (2011)). The GroEL chaperonin (Sigma Aldrich, 5 mg/ml) was diluted 5-fold in 200 mM MOPS pH 7.4, 100 mM KCl, and 4 mM MgCl₂. Diluted (1:10) protein A (a bacterial surface protein commonly used because of its ability to bind immunoglobins) conjugated with 5 nm colloidal gold particles (CMC-UMC, Utrecht, the Netherlands) was added (~3 μl) as fiducial markers to the samples just before EM grid preparation. Aliquots of 3 μl samples at ~1 mg/ml protein concentration were applied to 200 mesh thick C-flat grids (Protochips Inc., NC, USA) (1.2 mm hole size). All grids were freshly glow discharged for 30 s with a current of 20 mA. Excessive liquid was blotted at room temperature from one side inside a Leica EM GP freezing plunger using 3 s blotting time and 2 s postblotting time with 95% relative humidity. Subsequently, the blotted grid was plunged into liquid ethane for vitrification. The grids were stored in liquid nitrogen pending examination in the electron microscope. In addition to the low-salt hemoglobin sample described above, two more ion concentrations were tested, 1 M and 3 M NH₄Ac, respectively.

4.2. Image acquisitions/data collection

Frozen-hydrated specimens were examined with a Titan Krios electron microscope (FEI Company, The Netherlands), equipped with a field emission gun (FEG) operated at acceleration voltages of 80 and 300 kV. A post-column GIF energy filter (Gatan, USA) equipped with 2 k × 2 k Gatan US1000 camera was used. The energy slit was adjusted to select only electrons with an energy loss less than 5 eV. Other microscope settings were: condenser aperture number 3 (size of 100 μm), objective aperture 4 (100 μm), spot size index 5, and beam diameter of 2 μm. The spherical (C_s) and chromatic (C_c) aberrations for this Titan microscope are both 2.7 mm, while the energy spread (ΔE) and illumination aperture (α_i) are 0.7 eV and 0.03 mrad, respectively. The grids were mounted using the Krios autoloader. A cross-grating was used for magnification calibration. Images of proteasome, hemoglobin and GroEL at 80 kV and hemoglobin at 300 kV were recorded on a 2 k × 2 k Gatan CCD (US1000) camera with a magnification at the detector plane of 44.5 k×. The pixel size of the detector is 14 μm and the final sampling density in the object plane was 3.15 Å/pixel. The requested underfocus ranged from 500 nm to 4000 nm in five steps. The incident flux was derived from the detector analog-to-digital units (ADUs) by taking 1 s exposures without sample and using a conversion factor (in ADU/ e⁻) as calibrated by Vulović et al. (2010). Each defocus series was collected from a previously unexposed sample suspended across one of the holes in the C-flat grid. Electron fluxes of ~2.5 e⁻ Å⁻² s⁻¹ and ~5.5 e⁻ Å⁻² s⁻¹ at respectively 80 keV and 300 keV, were used to record each single frame, while the exposure times used were 0.5 s, 1 s, and 2 s. Images in a defocus series of the same view were taken with and without energy filtering. After each defocus series an image of the adjacent carbon support was acquired using image shift in order to accurately measure defocus and astigmatism on that area (Vulović et al., 2012). These values are then also used for the region of interest.

5. Results

The validation of our image formation model is based on a systematic comparison between simulated and experimental images under various experimental conditions. We present the influence of the solvent including ion concentration, defocus, integrated electron flux, motion factor, amorphousness of the specimen, ice

thickness, MTF and DQE of the camera, and incident electron energy on the image formation of samples embedded in vitreous ice (proteasome and hemoglobin). For an unbiased comparison between experimental and simulated images, the display for each image was stretched between mean value plus/minus 2.2 standard deviations of the corresponding experimental image. Estimated ice thickness *d* and defocus Δ*f* are specified accordingly.

5.1. Macromolecules embedded in vitreous ice

5.1.1. “Bond” contributions

As described in Section 2, $\mathcal{V}_{\text{atom}}$ is modeled using the isolated atom superposition approximation (IASA), while the redistribution potential as a result of the bond contributions $\mathcal{V}_{\text{bond}}$ is modeled by a Poisson–Boltzmann (PB) method. The ratio between the mean squared “bond” and “atom” potential contributions $R_{\text{bond}} = \langle \mathcal{V}_{\text{bond},0}^2 \rangle / \langle \mathcal{V}_{\text{atom},0}^2 \rangle$ was calculated for each of the simulated interaction potentials \mathcal{V}^{int} (Eq. (1)). $\mathcal{V}_{\text{bond},0}$ and $\mathcal{V}_{\text{atom},0}$ represent mean-subtracted $\mathcal{V}_{\text{bond}}$ and $\mathcal{V}_{\text{atom}}$ potentials, respectively. The values of R_{bond} for proteasome, hemoglobin in 50 mM, 1 M and 3 M NH₄Ac are 5.3%, 9.5%, 7.9%, and 7.7%, respectively. Fig. 2 permits comparisons between (1) experimental images, (2) simulated images which potential is calculated using only $\mathcal{V}_{\text{atom}}$, and (3) using combined potential $\mathcal{V}_{\text{atom}} + \mathcal{V}_{\text{bond}}$. For the experimental conditions used here, the $\mathcal{V}_{\text{bond}}$ contribution to \mathcal{V}^{int} is not significant. In general, the combined potential produces weaker ringing effects on the surface of the molecule and lower peaks inside the proteins (Fig. 2). The SNR in the experimental images was not high enough to notice apparent differences due to the redistribution potential $\mathcal{V}_{\text{bond}}$ within the molecule. We performed simulations with various integrated fluxes, magnifications, defoci and acceleration voltages to assess when it is needed to include $\mathcal{V}_{\text{bond}}$ in the modeling. Fig. 3 compares images from $\mathcal{V}_{\text{atom}}$ and combined potential for some of the parameters. Figs. 3B and 3D suggest that the dark hexagon produced by large defocusing (Δ*f* = 6 μm) is weaker when using the combined potential. The differences inside the protein are more pronounced at higher magnification (Figs. 3A and 3C), and at 300 kV (Figs. 3C and 3D), producing stronger signal for $\mathcal{V}_{\text{atom}}$ than for the combined potential. In general, assuming no beam-induced motion, a higher integrated flux better reveals minute differences inside the molecule. In Figs. 3B and 3D we used an integrated flux of 10 e⁻/Å², which is four times higher than in the actual experiments (at the same magnification). A corresponding SNR (assuming perfect alignment and no beam-induced motion) would be achieved experimentally by averaging 16 equivalent particles.

5.1.2. Defocus series

Various defocus series were acquired with a requested defocus ranging from 500 nm to 4000 nm. From the adjacent carbon area next to each region of interest the defocus values Δ*f* are estimated and provided in the figures captions. For readability we omit to display the astigmatism values as well as uncertainties of the defocus estimations as provided by tools described in Vulović et al. (2012). The astigmatism was always smaller than 6% of the defocus value. The uncertainties of defocus estimation were on average 1.6%.

Fig. 4 shows experimental and simulated defocus series of 20S proteasome, top and side view at 80 kV for 0.5 s and 1 s exposure time, respectively. The simulations correctly predict the changes in the image when the defocus value is altered. For small defocus values the contrast at low frequencies is too small to be distinguished from the noise. However, at larger amounts of underfocus the white fringes and the central channel in the top view (second and third column) are readily recognized and they appear comparable in both experimental and simulated

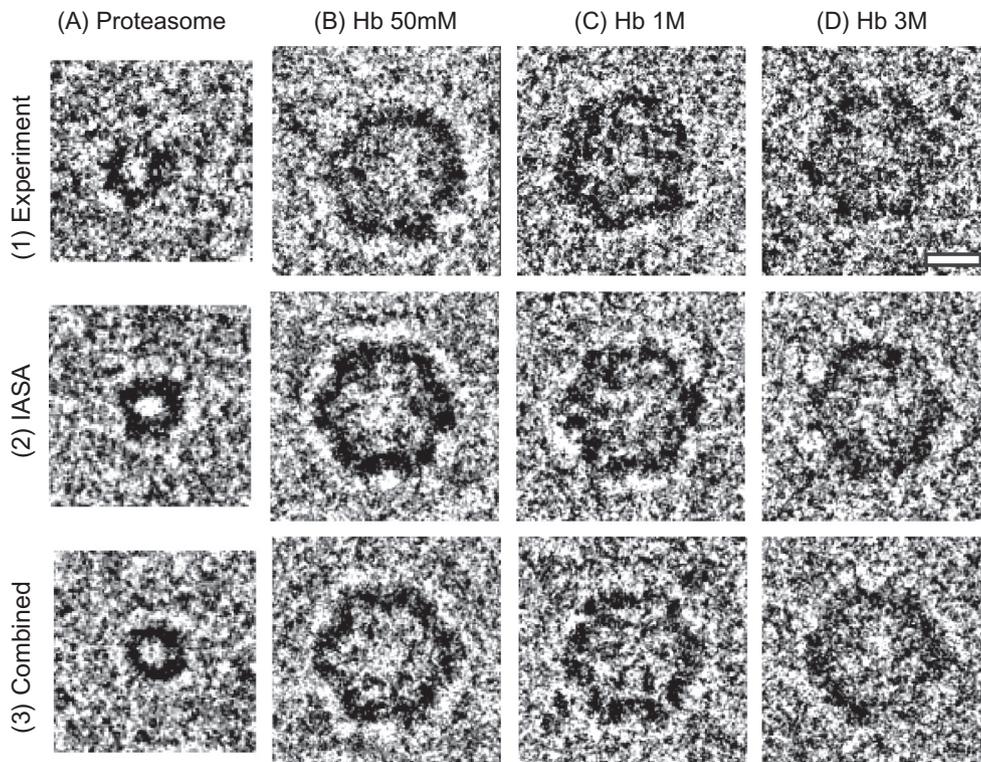


Fig. 2. Examples of (1) experimental images, (2) simulated images where the interaction potential (IP) was constructed from only ν_{atom} , and (3) simulated images with the IP calculated as combined potential $\nu_{\text{atom}} + \nu_{\text{bond}}$. The flux was $\sim 2.5 \text{ e}^{-} \text{ \AA}^{-2} \text{ s}^{-1}$ at 80 kV. From left to right are examples of (A) proteasome ($t_{\text{exp}} = 2 \text{ s}$, $\Delta f = 2509 \text{ nm}$, $d = 69 \text{ nm}$), (B) hemoglobin (Hb) in 50 mM ($t_{\text{exp}} = 2 \text{ s}$, $\Delta f = 4621 \text{ nm}$, $d = 82 \text{ nm}$), (C) Hb in 1 M ($t_{\text{exp}} = 2 \text{ s}$, $\Delta f = 4505 \text{ nm}$, $d = 196 \text{ nm}$), and (D) Hb in 3 M NH_4Ac ($t_{\text{exp}} = 1 \text{ s}$, $\Delta f = 2754 \text{ nm}$, $d = 169 \text{ nm}$). The scale bar corresponds to 10 nm.

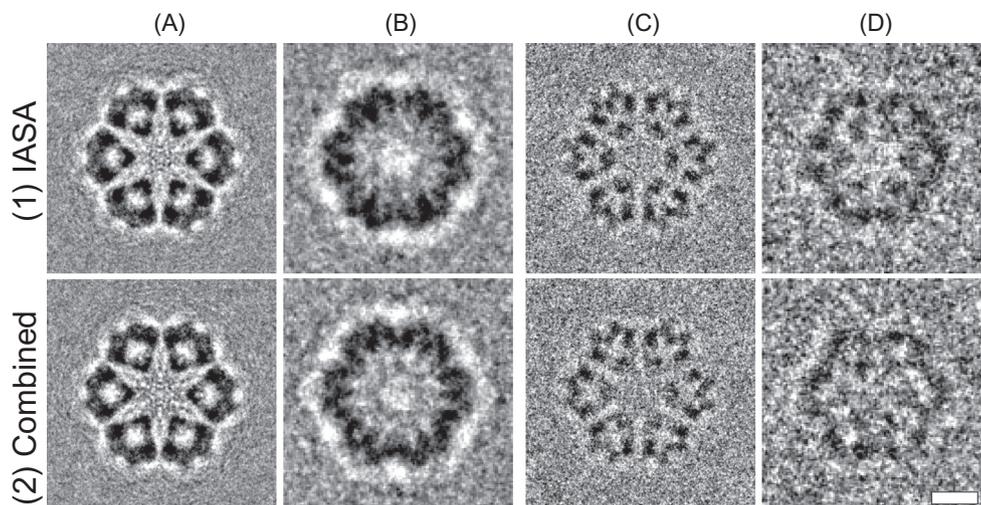


Fig. 3. Examples of (1) simulated images where interaction potential (IP) was constructed from only ν_{atom} , and (2) with the IP calculated as combined potential $\nu_{\text{atom}} + \nu_{\text{bond}}$. (A) Voltage 80 kV, Magnification 100 k \times , $\Delta f = 2000 \text{ nm}$, and integrated flux $100 \text{ e}^{-}/\text{\AA}^2$; (B) Voltage 80 kV, Magnification 42 k \times , $\Delta f = 6000 \text{ nm}$, and integrated flux $10 \text{ e}^{-}/\text{\AA}^2$; (C) and (D) are similar to (A) and (B), respectively but at a voltage of 300 kV. The scale bar corresponds to 10 nm.

images. The experimental images at higher defocus values provide less details as is predicted by the simulations (the fourth and the fifth column (side view) in Fig. 4).

5.1.3. Integrated flux series and motion factor

Subsequently, we tested whether the simulations can predict the effect of different integrated fluxes. After each defocus series, another region of interest was selected and imaged with a different integrated electron flux. The flux was kept constant

($\sim 2.5 \text{ e}^{-} \text{ \AA}^{-2} \text{ s}^{-1}$ at 80 kV), while exposure times were set to 0.5 s, 1 s, and 2 s, producing an integrated flux per single frame of $\sim 1.25 \text{ e}^{-}/\text{\AA}$, $\sim 2.5 \text{ e}^{-}/\text{\AA}$, and $\sim 5 \text{ e}^{-}/\text{\AA}$, respectively. Fig. 5 shows experimental and simulated integrated flux series of 20S proteasome top view (three parts subdivided in quadrants). We expect, based on the experiments shown in Brilot et al. (2012), that the beam-induced motion depends on the integrated flux. The effective motion factors ranged from 4 Å to 10 Å. Modeling smaller motion factors is not needed given our sampling density of 3.15 Å/pixel.

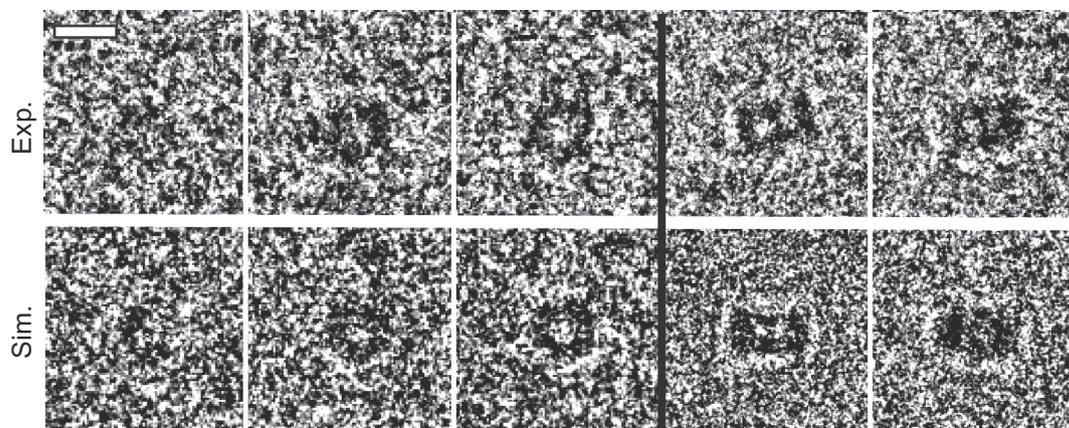


Fig. 4. Experimental and simulated defocus series of 20S proteasome at 80 kV and at a flux of $\sim 2.5 \text{ e}^- \text{ \AA}^{-2} \text{ s}^{-1}$. First three columns (top view): $t_{\text{exp}} = 0.5 \text{ s}$ and defoci from left to right 0.75, 1.3, 1.9 μm , respectively. Last two columns (side view): $t_{\text{exp}} = 1 \text{ s}$ and defoci 4.4, and 6.7 μm , respectively. The scale bar corresponds to 10 nm.

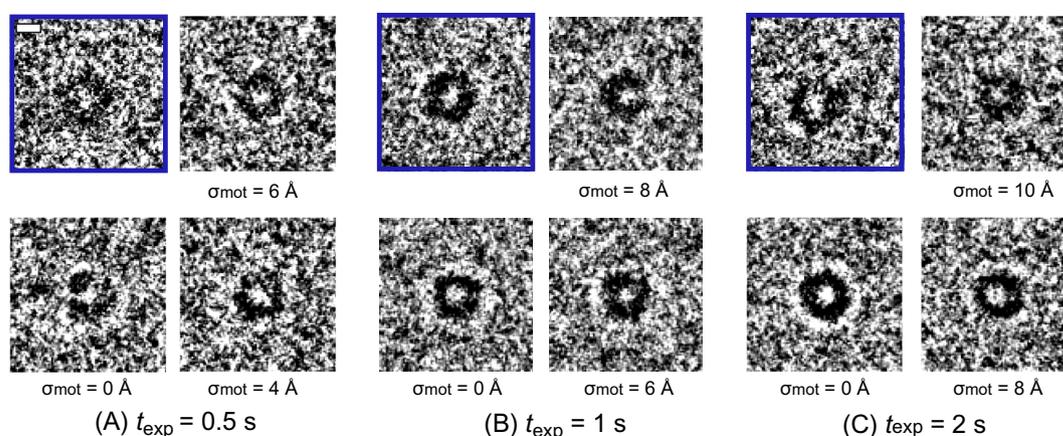


Fig. 5. Integrated flux series and varying motion factors σ_{mot} at 80 kV. The flux was $\sim 2.5 \text{ e}^- \text{ \AA}^{-2} \text{ s}^{-1}$. The experimental images (upper left quadrants) are framed. The simulated images with increasing motion factor are presented in anticlockwise direction. A higher integrated flux requires a larger motion factor. (A) $t_{\text{exp}} = 0.5 \text{ s}$, $\Delta f = 2492 \text{ nm}$, $d = 85 \text{ nm}$ (B) $t_{\text{exp}} = 1 \text{ s}$, $\Delta f = 4392 \text{ nm}$, $d = 92 \text{ nm}$, and (C) $t_{\text{exp}} = 2 \text{ s}$, $\Delta f = 2509 \text{ nm}$, $d = 69 \text{ nm}$. The scale bar corresponds to 5 nm.

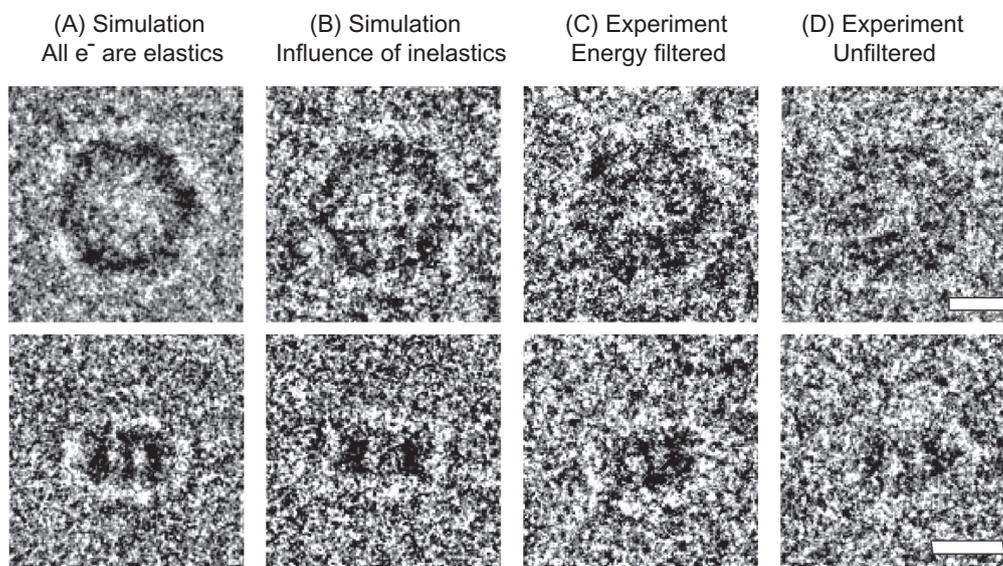


Fig. 6. Influence of inelastic scattering. (A) simulations of pure phase contrast, (B) simulations with inelastic scattering, (C) experimental zero-loss filtered images, and (D) experimental unfiltered images. From top to bottom are presented hemoglobin in 3 M NH_4Ac ($t_{\text{exp}} = 1 \text{ s}$, $\Delta f = 4918 \text{ nm}$, $d = 142 \text{ nm}$, $\sigma_{\text{mot}} = 8 \text{ \AA}$), and side view of 20S proteasome ($t_{\text{exp}} = 1 \text{ s}$, $\Delta f = 6713 \text{ nm}$, $d = 80 \text{ nm}$, $\sigma_{\text{mot}} = 0 \text{ \AA}$). In order to use the same display stretching as in the other examples, the overall higher intensity in (A) was scaled with a thickness dependent constant $\exp(-d/\Lambda_{\text{in}})$, while in (D) we used the ratio between the median value of the filtered and unfiltered images. The scale bar corresponds to 10 nm.

It can be seen that in the absence of motion factor modeling ($\sigma_{\text{mot}} = 0 \text{ \AA}$) the simulated images at higher integrated fluxes display a higher contrast and appear sharper than the experimental data. Incorporating a motion factor of $\sigma_{\text{mot}} \sim 4 - 8 \text{ \AA}$ and $\sigma_{\text{mot}} \sim 6 - 10 \text{ \AA}$ at $\sim 2.5 \text{ e}^-/\text{\AA}$ and $\sim 5 \text{ e}^-/\text{\AA}$, respectively let the simulations be in good agreement with the experiments. Note that the particles were selected from different areas of the specimen, so they differ slightly in defocus and specimen thickness.

5.1.4. Inelastic contributions

Fig. 6A shows simulations where only pure phase contrast is considered for the image formation and electron–specimen interaction. When inelastic events are considered (Fig. 6B), the vitreous ice will damp the amplitude of the propagating wave exponentially with increasing ice thickness. However, the difference between inelastic scattering properties of the protein and that of the vitreous ice (see Fig. 1) produces amplitude contrast. Since the inelastics are modeled as the imaginary part of the interaction potential, they are assumed to be removed from the image. Therefore, the simulated images must be compared with zero-loss energy filtered experimental images (Fig. 6C). The latter excludes most of the electrons with plasmon energy-losses. The objective aperture was large ($100 \mu\text{m}$) allowing us to assume that all elastically scattered electrons reached the detector. Fig. 6D shows unfiltered experimental images where both elastics and inelastics were detected, contributing to a stronger signal. However, the images appear more blurry because the inelastics that reached the detector lost their coherency.

5.1.5. Camera's DQE

Fig. 7 illustrates the necessity of modeling the detector's DQE instead of the commonly used MTF-only approach. The left image (Fig. 7A) was simulated using Eq. (7), assuming that the signal and noise are transferred with the same MTF (NTF = MTF). Fig. 7B shows a simulation which takes into account the DQE and the influence of the conversion factor on the image quality. The experimental image (Fig. 7C) is comparable to Fig. 7B, showing the importance of modeling the DQE.

5.1.6. Acceleration voltage influence

The low-frequency contrast in experimental and simulated images at 300 kV acceleration voltage is smaller than at 80 kV whereas the incident integrated flux was higher (see Figs. 8 and 2). This is in agreement with the energy dependent scattering properties of the incident electrons, interaction constant (see Eq. (4)), and the CTF. Additionally, the MTF and DQE of the CCD camera decrease with increasing acceleration voltage contributing to a reduced low-frequency contrast (Meyer and Kirkland, 1998). However, these combined effects provide an apparent higher level of details in the images (see Fig. 3). At 300 kV the motion factor appears to be smaller (Fig. 8), (data not shown for $\sigma_{\text{mot}} > 0 \text{ \AA}$). In

Fig. 8, it appears that the simulated images at 300 kV using only $\mathcal{V}_{\text{atom}}$ (2) provide a stronger signal compared to the experimental images (1) and to the images that use the combined potential $\mathcal{V}_{\text{atom}} + \mathcal{V}_{\text{bond}}$ (3) (see also Fig. 3).

5.1.7. Amorphousness of the solvent

Fig. 9 shows the influence of the amorphousness of the solvent on the image. The positions of the water molecules were simulated via MD (see Section 3.1.4) in a $20 \times 20 \times 50 \text{ nm}$ box and the interaction potential was generated via IASA. A region 1 modeled with amorphousness is compared to a region 2 where the solvent is modeled as a constant potential and the noise is only due to Poisson statistics. At the integrated flux used in experiments (Fig. 9A), the difference between those two regions is not noticeable (Fig. 9B). Simulated integrated flux series (Figs. 9C–F) suggests that only at high integrated fluxes ($> 100 \text{ e}^-/\text{\AA}^2$), high magnification, and without beam-induced motion, the difference between Poisson noise and solvent amorphousness becomes apparent (Fig. 9E). At 300 kV, the differences are less pronounced, even at a high integrated flux (Fig. 9F).

6. Discussion

Here we highlight and discuss the aspects of our simulation model.

6.1. Forward model

A structure deposited in the PDB contains type and position of atoms in the molecule, although hydrogen atoms are usually lacking. Some of the programs that offer the functionality of adding hydrogen and other missing atoms are described in Hess et al. (2008), Dolinsky et al. (2004), Vriend (1990). In our case, to calculate $\mathcal{V}_{\text{atom}}$, scattering factors for frequencies up to $q = 4 \text{ \AA}^{-1}$ are parameterized as a weighted sum of five Gaussians and provided in Table 1 in Peng et al. (1998). Parameterizations of the scattering factors up to $q = 12 \text{ \AA}^{-1}$ (provided by Table 3 in Peng et al. (1998)) or using a combination of Gaussians and Lorentzians (Kirkland, 2010) would only be beneficial for very high scattering angles and/or heavy atomic elements. Biological specimens mainly consist of lighter elements such as H, C, O, and N, and the deviations of the parameterized curves in Peng et al. (1998) for these elements, from the parameterizations in Kirkland (2010), are less than 0.1%. An advantage of using the parameterization as implemented here is that it avoids singularities at zero distances from the atomic nucleus. Here, calculations of $\mathcal{V}_{\text{atom}}$ are based on a slight modification of Rullgård et al. (2011), in such way that low-pass filtering to a certain resolution does not exceedingly damp the IP, and the solvent is assumed to be vitreous ice instead of water (see Eq. (5) in S.M.). Note that the difference between the inelastic mean free paths of vitreous ice and protein (Fig. 1) contributes to the

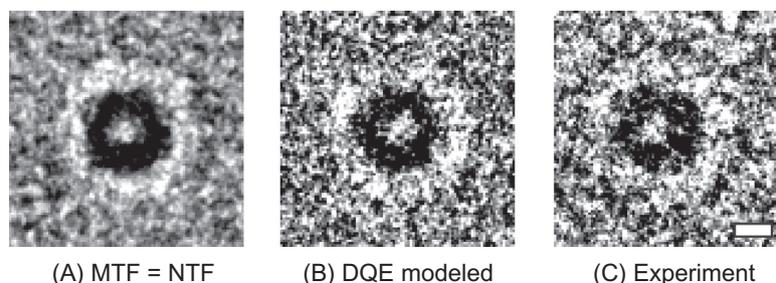


Fig. 7. Influence of the camera's DQE. (A) Simulated image assuming the same MTF for the signal and the noise. (B) Simulated image by taking into account the measured DQE. (C) Experimental image ($t_{\text{exp}} = 1 \text{ s}$, $d = 92 \text{ nm}$, $\Delta f = 6713 \text{ nm}$, and $\sigma_{\text{mot}} = 6 \text{ \AA}$). The scale bar corresponds to 5 nm.

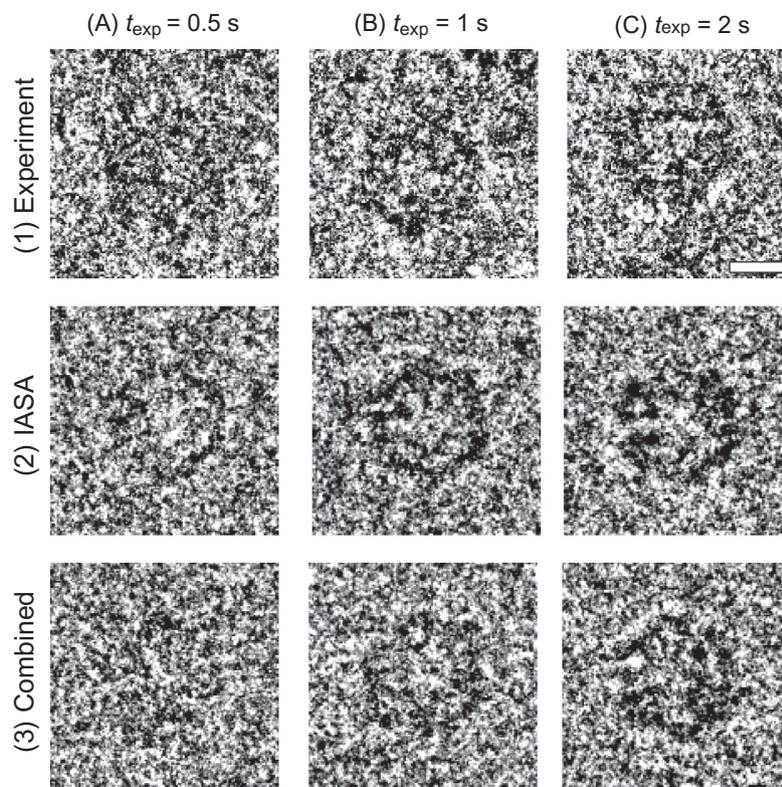


Fig. 8. Integrated flux series and ν_{bond} influence of hemoglobin (Hb) at 300 kV. Flux was $\sim 5.5 \text{ e}^- \text{ \AA}^{-2} \text{ s}^{-1}$. Experimental images (top row (1)), simulated images where the interaction potential (IP) was constructed from only ν_{atom} (middle row (2)), and simulated images with the IP calculated as combined potential $\nu_{\text{atom}} + \nu_{\text{bond}}$ (bottom row (3)). The integrated flux series (A) $t_{\text{exp}} = 0.5 \text{ s}$, $\Delta f = 5607 \text{ nm}$, $d = 176 \text{ nm}$ (B) $t_{\text{exp}} = 1 \text{ s}$, $\Delta f = 5026 \text{ nm}$, $d = 61 \text{ nm}$, and (C) $t_{\text{exp}} = 2 \text{ s}$, $\Delta f = 5750 \text{ nm}$, $d = 180 \text{ nm}$. Under these imaging conditions there appears to be no requirement for inclusion of the motion factor. The scale bar corresponds to 10 nm.

amplitude contrast, but the plasmons of the vitreous ice attenuate the useful phase signal.

To describe electron wave propagation through a specimen with a finite thickness and to account for multiple scattering events, a multislice approach, inspired by Kirkland (2010) is used. Criteria for applicability of the weak-phase object approximation, projection assumption and multislice approach are presented in Vulović et al. (submitted for publication). The criteria indicate that the projected potential map of hemoglobin sampled with a 3 Å pixel size does not, strictly speaking, satisfy the projection assumption, while the weak-phase object approximation holds. This implies that the thickness of the specimen cannot be neglected. Here, we simulated images of a single protein in a field of view smaller than 400×400 pixels for which the multislice approach took only a couple of seconds to compute. However, if one simulates a (tilt) series of e.g. $4 \text{ k} \times 4 \text{ k}$ images, several hours of computational time would be required. If we assume not more than one (weak) elastic scattering event per incoming electron (first Born approximation), the free-space (Fresnel) propagation through a thick and/or tilted sample can be incorporated in the CTF (Philippesen et al., 2007; Winkler and Taylor, 2003; DeRosier, 2000; Wan et al., 2004). We provide the possibility of including such a geometry in the CTF, which speeds up the forward computation (Voortman et al., 2011) as well as the 3DCTF correction (Voortman et al., 2012). This approach assumes the weak-phase object approximation which is in our case satisfied but might not fulfilled for thicker and/or tilted specimens and for higher resolution.

Performing MD simulations on a system consisting of both protein (GroEL) and solvent instead of doing it separately (Hall et al., 2011) should provide a more realistic modeling of the hydration shell of the protein (Shang and Sigworth, 2012). We expect it should reduce the contrast between protein and environment,

thereby further bridging the gap between simulations and experiments. Incorporating such a model might be subject for further studies. It has been reported recently (Shang and Sigworth, 2012) that such MD simulations can be used to derive a continuum model which describes the density of the water molecules surrounding a protein surface. Our current multislice algorithm does not require an explicit atomistic model of the solvent as in Hall et al. (2011).

As TEM image formation usually involves only small angle scattering events, it is possible to ignore off-axis and higher order aberrations and only consider axial aberrations (Peng et al., 2004). As our implementation is modular, there is a possibility of including higher order axial and non-axial aberrations in the future.

The insufficient SNR due to the low-flux imaging conditions and/or due to the beam-induced movements caused that we could not provide experimental evidence of amorphousness due to the solvent in our samples (compare Fig. 9). Consequently, the solvent can be modeled as a continuous medium, simplifying the simulations. Furthermore, by modeling the solvent as vitreous ice, 7% less dense than water at room temperature (Alcorn and Juers, 2010), our predicted contrast would be slightly increased compared to Hall et al. (2011), Rullgård et al. (2011).

Noise in the images is mainly Poisson distributed, and strictly speaking signal-dependent. Here, we introduce a new way of modeling DQE which separates the signal and the noise transfer. An accurate description of the signal/noise transfer may facilitate the regularization in the reconstruction methods.

6.2. Poisson–Boltzmann (PB) approach

We characterize the influence of the solvent dielectric properties, ionic strength and electrostatic distribution within a molecule for TEM simulations and compare these with the isolated atom

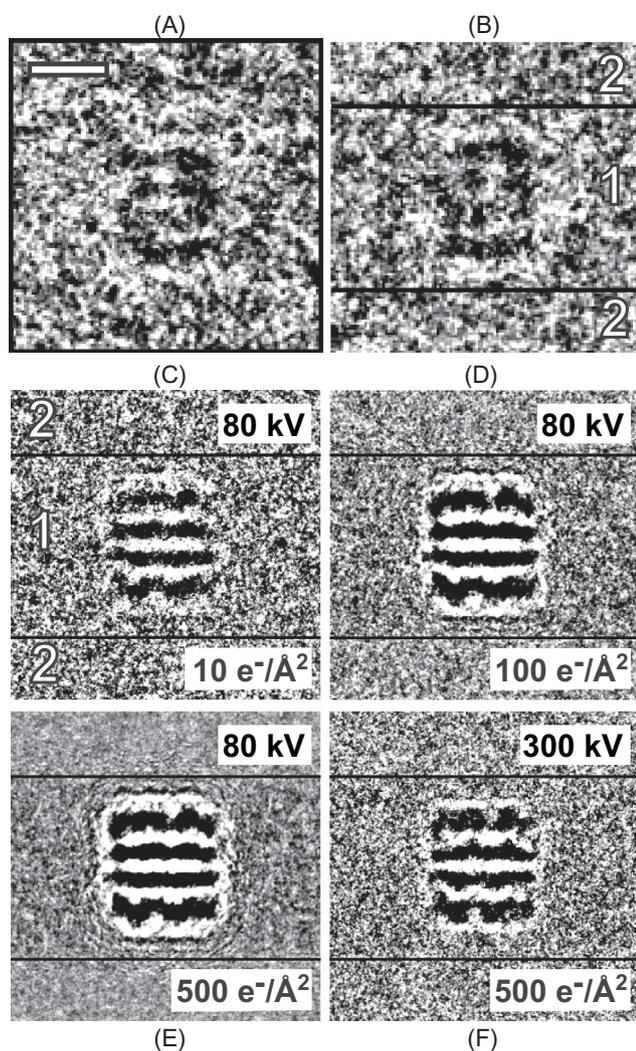


Fig. 9. Influence of the amorphousness of the solvent (region 1 between two horizontal lines) compared to Poisson noise only (region 2 below and above lines). (A) experiment and (B) simulation (voltage 80 kV, magnification 42 \times , integrated flux $\sim 5 \text{ e}^-/\text{\AA}^2$, $\Delta f = 2718 \text{ nm}$, $d = 120 \text{ nm}$, $\sigma_{\text{mot}} = 6 \text{ \AA}$); (C)–(F) Amorphousness dependence in simulations with varying integrated flux and voltage (magnification 100 \times , $\Delta f = 1 \mu\text{m}$, $d = 20 \text{ nm}$); for display purposes a percentile stretch was used (the lower and upper 1% of the gray values were clipped before stretching) (C) integrated flux $10 \text{ e}^-/\text{\AA}^2$ at 80 kV; (D) integrated flux $100 \text{ e}^-/\text{\AA}^2$ at 80 kV; (E) integrated flux $500 \text{ e}^-/\text{\AA}^2$ at 80 kV; (F) integrated flux $500 \text{ e}^-/\text{\AA}^2$ at 300 kV. The scale bar corresponds to 10 nm.

superposition approximation (IASA) where atoms are treated in isolation. The redistribution of the potential due to the interactions is modeled via a PB approach. The ratio R_{bond} between the mean squared “bond” (PB) and “atom” (IASA) potential contributions ranged from 6 to 10%, suggesting that the $\mathcal{V}_{\text{atom}}$ contribution is the dominant part of the interaction potential. The mean value was subtracted from these potentials prior to calculating R_{bond} since the mean value does not influence the phase contrast (Vulović et al., submitted for publication). Comparing simulated images where the IP was constructed only via $\mathcal{V}_{\text{atom}}$ with the ones where the IP was calculated by combining $\mathcal{V}_{\text{atom}}$ and $\mathcal{V}_{\text{bond}}$ did not show significant differences. In general, the images with the combined IP show weaker ringing effects around the protein edges (Fig. 2), better matching the corresponding experimental images. The simulations suggested that the differences would be more pronounced for higher SNR (Fig. 3). Increasing the ion concentration in the hemoglobin solution resulted in a slight decrease of the mean $\mathcal{V}_{\text{bond}}$. A possible explanation is that the electrostatic shielding of

the protein with a negative net charge produces smaller absolute values of $\mathcal{V}_{\text{bond}}$ for higher ion concentrations.

In material science, it has been reported (Kirkland, 2010) that errors up to 10% in calculation of electron scattering factors can occur due to the modeling of atoms as isotropic. The PB approach does not aim to accurately characterize the bonding between the individual atoms but focuses on the macroscopic influence of the solvent and ions on the potential distribution. The accuracy of PB approaches decreases in the region very close to the nucleus since the partial (net) charges are placed at the position of the nucleus (see Section 1.4 in S.M.). The main benefit of the PB method lays therefore in the description of the potential redistribution due to the interaction with the solvent and its ions. Here, we have chosen APBS (Baker et al., 2001), a software package for numerically solving the PB equation based on finite elements. There are, however, other approaches, such as the boundary element solution (Bajaj et al., 2011), which may provide a faster and more accurate description of the potential at the protein boundaries. In this analysis we studied oligomeric macromolecules. The influence of the PB approach might be different for non-oligomeric macromolecules. Furthermore, the PB approach might facilitate the interpretation of transient states.

6.3. Beam-induced movements

Beam-induced specimen movements have long been recognized as one of the main factors attenuating the signal in cryo-EM (Henderson and Glaeser, 1985; Bullough and Henderson, 1987). It has been suggested that the main causes of this local motion are specimen deformation and radiation damage during the exposure (Glaeser and Taylor, 1978; Glaeser, 2008; Brilot et al., 2012; Karimi Nejadasl et al., 2013), and/or charging (Henderson, 1992; Glaeser and Downing, 2004). The inclusion of the motion factor blurs the simulated images to better match the experimental images. This effect is analogous to the damping envelope due to misalignment in SPA (Jensen, 2001). Our approach to include this damping effect is inspired by recent experiments of Brilot et al. (2012) who aimed to quantify the flux-dependent beam-induced movements.

In our analysis, the derived motion factors are similar to the displacement values reported in Brilot et al. (2012), Li et al. (2013). Our observations are consistent with their suggestions that (i) the motion is larger for higher fluxes, and (ii) the motion rate decreases with exposure time showing that the motion is worst at the beginning of the exposure. In our case, the total dose that a specimen received could be up to $10\times$ larger than the dose used to acquire individual images since we acquired numerous exposure series, e.g. at different defoci or with/without energy filter.

Henderson and Glaeser (1985) suggested that some type of beam-induced movements of the specimens (around 5 \AA or more) must occur in approximately equal amounts in all directions. Li et al. (2013) found that this motion is not unidirectional, whereas Brilot et al. (2012) reported directional preference of movements. The simulator accounts for isotropic motion and can be extended to model any particle trajectory during acquisition. However, if such a trajectory is known, it is preferred to correct for it by aligning and averaging the frames captured by a direct electron detector.

Our effective motion factor is smaller at 300 kV than at 80 kV. This could be related to numerous effects including differences in inelastic cross-sections, beam quality, or ice thickness. Note that the integrated electron flux used for 300 kV was about two times larger than at 80 kV in order to have similar deposited energy per mass (dose). The ability to recognize amorphousness of the specimen/solvent decreases due to the beam-induced motion. Due to the large variation in the magnitude of the movements,

some particles can have better contrast than others within the same field of view. The challenge remains how to avoid or automatically correct for beam-induced motion (Brilot et al., 2012).

6.4. Validation

In previous work on accurate forward modeling, only the Tobacco Mosaic Virus (TMV) was characterized (Hall et al., 2011) and compared with experiments (Rullgård et al., 2011). The advantage of TMV as a test sample is that the averaged 1D profile across the virus yields a high SNR and can be used for quantitative comparisons. Here, we analyzed three different biological specimens (20S proteasome, hemoglobin, and GroEL) in different embedding materials. Each simulation parameter relates directly to a physical quantity, but a strict quantitative comparison to cryo-EM images is difficult due to the high level of noise and challenging alignment, even for 1D-averaged profiles. For unbiased signal comparison, the display of each image was stretched to match the corresponding experimental image. For visual comparisons we simulated ten different noise realizations (data not shown), confirming that the noise did not change the appearance of the features. A comprehensive quantitative comparison in cryo-EM is mainly compromised by the low SNR. For completeness, it should be mentioned that also in material science, although dealing with much higher SNR, validation of simulations is usually done only visually (Kirkland et al., 2008; Kirkland, 2010), even though there are attempts of using more quantitative approaches (Mobus et al., 1998; Tang et al., 1993; Tang et al., 1994; Saxton, 1997). In Section 5 in S.M., we present simulated and experimental images of carbon edges and carbon nanotubes and their 1D-averaged profiles for a more quantitative comparison. Advantages of using carbon edges and nanotubes for validation include the simplicity of their model and radiation hardness compared to cryo-EM.

Most simulation parameters described in Rullgård et al. (2011) are based on physical principles. They need to employ, however, a calibration protocol for some parameters that are phenomenologically introduced, requiring their tuning. Examples of such parameters are amorphousness, absorption potential, as well as camera parameters such as the MTF, DQE, and conversion factor. The ice thickness in Rullgård et al. (2011) was estimated from one spot although the thickness can vary significantly throughout the field of view. An advantage of the ice thickness measurements via an energy filter as described in this paper is that it is relatively fast experimentally and provides information about the local thickness. We assume that the energy filter was stable during acquisition as characterized in Lücken et al. (2008), without significantly compromising the accuracy of the thickness measurements. However, the experimentally determined values of mean free inelastic path used to estimate the thickness can vary noticeably (Feja and Aebi, 1999). Defocus values normally deviate from the values requested from the microscope. We estimate defocus and astigmatism on the adjacent carbon area which could, in principle, differ from the values at the region of interest due to the non-perpendicular pose of the sample relative to the beam. Although a model for the absorption potential was introduced in Rullgård et al. (2011), simulated data were compared against unfiltered experimental images which also contain inelastically scattered electrons. However, any modeling based on the imaginary part of the IP (analogously to Beer–Lambert law) assumes that those inelastic electrons are removed from the image, requiring a comparison with zero-loss energy filtered images.

Implemented in the C programming language, the user-friendly TEM-simulator (Rullgård et al., 2011) represents a good starting point for image simulations in cryo-EM. The novel aspects described and analyzed in this paper are included in InSilicoTEM, a simulator implemented in MATLAB.

6.5. Outlook

Accurate image simulations help to understand how the recorded image is formed, indicate ways to optimize data acquisition and microscope settings, and provide insight on ways to improve instrumentation. As an integral part of an accurate forward model, the estimation of parameters such as defocus and astigmatism (Vulović et al., 2012) and camera's MTF and DQE (Vulović et al., 2010) is essential and necessary for the CTF correction and/or regularization of the reconstruction approaches.

In addition to improving computational methods, the experimental developments should allow better transfer of the scattered electron wave onto the recorded image intensity. These experimental improvements are being achieved mostly by better sample preparation, by minimizing noise using direct electron detectors and electron counters, by improving the effective CTF via phase plates and more coherent electron sources, and by minimizing the effective beam-induced movements of the specimen.

The magnitude of the beam-induced movements must be reduced in order to increase the contrast in the images. Their effect can be somewhat decreased by lowering the flux, using a smaller carbon hole size, or by pre-irradiation. Postprocessing alignment and averaging the frames captured by a direct electron detector can further reduce blurring in the final images (Brilot et al., 2012; Campbell et al., 2012; Karimi Nejadasl et al., 2013; Bai et al., 2013; Li et al., 2013). It is expected that dose fractionation and superresolution (beyond-Nyquist) EM using electron counting devices can reduce the influence of beam-induced movements, improving the achievable contrast in cryo-EM images. The modularity of InSilicoTEM allows integration of new physical parameters as well as modeling the influence of new hardware components such as the new generation of direct electron detectors.

The simulator could help to assess whether it is possible to resolve a specific macromolecule using a certain set of instrumental and processing parameters. It will be possible to easily and cost-effectively investigate the influence of new data acquisition techniques and advanced instrumentation, and to facilitate the development and evaluation of reconstruction and image processing techniques. In addition to the known PDB file, the input for InSilicoTEM simulator can also be a previously reconstructed 3D potential map of a sample. The simulator could furthermore facilitate the identification of molecular assemblies within the cell, a docking process where atomic models are fitted into cryo-EM obtained maps, and testing whether a proposed 3D model of a macromolecule is in agreement with the features observed in the micrographs. In electron tomography, iterative reconstruction schemes such as simultaneous iterative reconstruction technique attempt to minimize the difference between projections and simulated reprojections of the 3D map. From the differences between observed and simulated images one can often derive information to refine the model. The model parameters are iterated until simulated images best describe those observed. We expect that an accurate forward model based on physical principles will facilitate such iterative scheme and reconstruction resulting in the 3D potential map.

7. Conclusions

We described an approach to simulate image formation in cryo-EM based on physical principles and taking into account the influence of the specimen and its surroundings, the optics of the microscope and the detector. Simulated and experimental images were generated under various settings and visually compared. Generated images adequately predict the effects of the phase contrast introduced by defocusing (Fig. 4), the changes due to the electron

flux (Fig. 5), the influence of inelastic scattering (Fig. 6), camera DQE (Fig. 7), and the acceleration voltage (Fig. 8).

Various buffer compositions have been used to evaluate the influence of charge redistributions for the hemoglobin sample. The contribution of this redistribution to the interaction potential appears to be less than 10% for all these cases and is mostly visible by the slightly less contrast at protein-solvent interfaces compared to the images calculated using only the IASA-based potential.

Increasing the integrated electron flux showed the benefit of introducing a motion factor which could be related to the beam-induced motions. For the 20S proteasome images taken at exposure times of 0.5 s, 1 s, and 2 s, the motion factors were in the range of $\sim 4 \text{ \AA}$, $\sim 6 \text{ \AA}$, $\sim 8 \text{ \AA}$, respectively (see Fig. 5). At 300 kV the motion factor appears to be smaller (Fig. 8). The varying contrast of the particles within a field of view can be explained by the apparently space-variant beam-induced movements.

For typical integrated electron fluxes in cryo-EM ($<100 \text{ e}^-/\text{\AA}^2$), the influence of the amorphousness of the solvent can be neglected since Poisson noise is the dominant noise source in the image (see Fig. 9) and the solvent can be modeled as a continuum.

The theory and methods provided here represent the basis of an expert system that could optimize the data collection strategy and inexpensively and efficiently investigate the influence of the new hardware.

Acknowledgments

The authors thank Paul Mooney (Gatan) for his help related to the MTF and DQE of the GIF camera, Cristina Avramut (LUMC) for assistance in preparation of hemoglobin, Chandra Bajaj and Alex Rand (ICES) for discussions related to the PB methods, Ronald Ligteringen (TU Delft) for help with pdb2pqr parsers, Lennard Voortman (TU Delft) for useful discussions, Willem van Dorp (RUG) for providing graphite flakes, Eric Bosch (FEI) for assistance in using Geant4, Erik Franken (FEI) for critical reading of the manuscript, and acknowledge the use of cryo-EM facility at the Netherlands Center for Electron Nanoscopy (NeCEN) and Gert Oostergetel (RUG). M.V. acknowledges support from the FOM Industrial Partnership Program No. 07.0599.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.jsb.2013.05.008>.

References

- Agostinelli, S., Allison, J., Amako, K., Apostolakis, J., Zschiesche, D., et al., 2003. Geant4 – a simulation toolkit. *Nucl. Instrum. Methods Phys. Res.* 506, 250–303.
- Alcorn, T., Juers, D.H., 2010. Progress in rational methods of cryoprotection in macromolecular crystallography. *Acta Crystallogr. Sect. D* 66, 366–373.
- Allen, F.H., Kennard, O., Watson, D.G., Brammer, L., Orpen, A.G., Taylor, R., 1987. Tables of bond lengths determined by X-ray and neutron diffraction. Part 1. Bond lengths in organic compounds. *J. Chem. Soc., Perkin Trans. 2* (12), 1–19.
- Bai, X., Fernandez, I.S., McMullan, G., Scheres, S.H.W., 2013. Ribosome structures to near-atomic resolution from thirty thousand cryo-EM particles. *eLife Sciences* 2.
- Bajaj, C., Chen, S., Rand, A., 2011. An efficient higher-order fast multipole boundary element solution for Poisson Boltzmann-based molecular electrostatics. *SIAM J. Sci. Comput.* 33, 826–848.
- Baker, N.A., Sept, D., Joseph, S., Holst, M.J., McCammon, J.A., 2001. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. U.S.A.* 98, 10037–10041.
- Bilbao-Castro, J.R., Sorzano, C.O.S., García, I., Fernández, J.J., 2004. Phan3D: Design of biological phantoms in 3D electron microscopy. *Bioinformatics* 20, 3286–3288.
- Boothroyd, C.B., Dunin-Borkowski, R.E., Stobbs, W.M., Humphreys, C.J., 1995. Quantifying the effects of amorphous layers on image contrast using energy filtered transmission electron microscopy. In: *MRS Proceedings*, vol. 354, pp. 495–500.

- Brilot, A.F., Chen, J.Z., Cheng, A., Pan, J., Harrison, S.C., Potter, C., Carragher, B., Henderson, R., Grigorieff, N., 2012. Beam-induced motion of vitrified specimen on holey carbon film. *J. Struct. Biol.* 177, 630–637.
- Bullough, P., Henderson, R., 1987. Use of spot-scan procedure for recording low-dose micrographs of beam-sensitive specimens. *Ultramicroscopy* 21, 223–230.
- Campbell, M.G., Cheng, A., Brilot, A., Moeller, A., Lyumkis, D., Veelsler, D., Pan, J., Harrison, S., Potter, C., Carragher, B., Grigorieff, N., 2012. Movies of ice-embedded particles enhance resolution in electron cryo-microscopy. *Structure* 20, 1823–1828.
- Colliex, C., Jeanguillaume, C., Mory, C., 1984. Unconventional modes for stem imaging of biological structures. *J. Ultrastruct. Res.* 88, 177–206.
- Cowley, J.M., Moodie, A.F., 1957. The scattering of electrons by atoms and crystals. I. A new theoretical approach. *Acta Crystallogr.* 10, 609–613.
- DeRosier, D.J., 2000. Correction of high-resolution data for curvature of the Ewald sphere. *Ultramicroscopy* 81, 83–98.
- Dinges, C., Rose, H., 1995. Simulation of filtered and unfiltered TEM images and diffraction patterns. *Phys. Stat. Sol. A Appl. Mater. Sci.* 150, 23–29.
- Dolinsky, T.J., Nielsen, J.E., McCammon, J.A., Baker, N.A., 2004. Pdb2pqr: an automated pipeline for the setup of poisson boltzmann electrostatics calculations. *Nucleic Acids Res.* 32, 665–667.
- Egerton, R., 1992. A data base for energy-loss cross sections and mean free paths. In: *Proc. 50th Ann. Meet. Elec. Microsc.*, vol. 136, pp. 1248–1249.
- Egerton, R.F., 1996. *Electron Energy-Loss Spectroscopy in the Electron Microscope*, second ed. Springer.
- Feja, B., Aebi, U., 1999. Determination of inelastic mean free path of electrons in vitrified ice layers for on-line thickness measurements by zero-loss imaging. *J. Microsc.* 193, 15–19.
- Frank, J., 2006. *Three-Dimensional Electron Microscopy of Macromolecular Assemblies: Visualization of Biological Molecules in Their Native State*, 2nd ed. Oxford University Press.
- Frank, J., Shimkin, B., 1978. A new image processing software for structural analysis and contrast enhancement. In: Sturgess, J.M. (Ed.), *Electron Microscopy 1978, Papers Presented at the Ninth International Congress on Electron Microscopy*, Toronto, vol. 1. Microscopical Society of Canada, Toronto, pp. 210–211.
- Glaeser, R.M., 2008. Retrospective: Radiation damage and its associated “information limitations”. *J. Struct. Biol.* 163, 271–276.
- Glaeser, R.M., Downing, K.H., 2004. Specimen charging on thin films with one conducting layer: discussion of physical principles. *Microsc. Microanal.* 10, 790–796.
- Glaeser, R.M., Downing, K.H., DeRosier, D., Chu, W., Frank, J., 2006. *Electron Crystallography of Biological Macromolecules*. Oxford University Press.
- Glaeser, R.M., Taylor, K.A., 1978. Radiation damage relative to transmission electron microscopy of biological specimens at low temperature: a review. *J. Microsc.* 112, 127–138.
- Gómez-Rodríguez, A., Beltrán-del Río, L.M., Herrera-Becerra, R., 2009. SimulaTEM: Multislice simulations for general objects. *Ultramicroscopy* 110, 95–104.
- Grimm, R., Typke, D., Bärmann, M., Baumeister, W., 1996. Determination of the inelastic mean free path in ice by examination of tilted vesicles and automated most probable loss imaging. *Ultramicroscopy* 63, 169–179.
- Hall, R.J., Nogales, E., Glaeser, R.M., 2011. Accurate modeling of single-particle cryo-EM images quantitates the benefits expected from using Zernike phase contrast. *J. Struct. Biol.* 174, 468–475.
- van Heel, M., Harauz, G., Orlova, E.V., Schmidt, R., Schatz, M., 1996. A new generation of the IMAGIC image processing system. *J. Struct. Biol.* 116, 17–24.
- Henderson, R., 1992. Image contrast in high-resolution electron microscopy of biological macromolecules: TMV in ice. *Ultramicroscopy* 46, 1–18.
- Henderson, R., 2004. Realizing the potential of electron cryo-microscopy. *Q. Rev. Biophys.* 37, 3–13.
- Henderson, R., Glaeser, R.M., 1985. Quantitative analysis of image contrast in electron micrographs of beam-sensitive crystals. *Ultramicroscopy* 16, 139–150.
- Hess, B., Kutzner, C., van der Spoel, D., Lindahl, E., 2008. GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.* 4, 435–447.
- Heymann, J.B., Belnap, D.M., 2007. Bsoft: image processing and molecular modeling for electron microscopy. *J. Struct. Biol.* 157, 3–18.
- Hytch, M.J., Stobbs, W.M., 1994. Quantitative comparison of high resolution TEM images with image simulations. *Ultramicroscopy* 53, 191–203.
- Jensen, G.J., 2001. Alignment error envelopes for single particle analysis. *J. Struct. Biol.* 133, 143–155.
- Karimi Nejadasl, F., Karuppasamy, M., Newman, E.R., McGeehan, J.E., Ravelli, R.B.G., 2013. Non-rigid image registration to reduce beam-induced blurring of cryo-electron microscopy images. *J. Synchrotron Radiat.* 20, 58–66.
- Karuppasamy, M., Karimi Nejadasl, F., Vulović, M., Koster, A.J., Ravelli, R.B.G., 2011. Radiation damage in single-particle cryo-electron microscopy: effects of dose and dose rate. *J. Synchrotron Radiat.* 18, 398–412.
- Kieft, E., Bosch, E., 2008. Refinement of monte carlo simulations of electron-specimen interaction in low-voltage SEM. *J. Phys. D: Appl. Phys.* 41, 215310.
- Kirkland, A.I., Chang, L.-Y.S., Hutchison, J.L., 2008. Atomic resolution transmission electron microscopy. In: Hawkes, P.W., Spence, J.C.H. (Eds.), *Science of Microscopy*. Springer Verlag, pp. 3–64, chapter 1.
- Kirkland, E.J., 2010. *Advanced Computing in Electron Microscopy*, second ed. Springer Verlag.
- Langmore, J.P., Smith, M.F., 1992. Quantitative energy-filtered electron microscopy of biological molecules in ice. *Ultramicroscopy* 46, 349–373.
- Leis, A., Rockel, B., Andrees, L., Baumeister, W., 2009. Visualizing cells at the nanoscale. *Trends Biochem. Sci.* 34, 60–70.

- Li, X., Mooney, P., Zheng, S., Booth, C.R., Braunfeld, M.B., Gubbens, S., Agard, D.A., Cheng, Y., 2013. Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. *Nat. Methods*. <http://dx.doi.org/10.1038/nmeth.2472>.
- Lucić, V., Förster, F., Baumeister, W., 2005. Structural studies by electron tomography: from cells to molecules. *Ann. Rev. Biochem.* 74, 833–865.
- Lücken, U., Tiemeijer, P., Barfels, M., Mooney, P., Bailey, B., Agard, D., 2008. Image information transfer through a post-column energy filter with detection by a lens-coupled transmission-scintillator CCD camera. *Microsc. Microanal.* 14, 1322–1323.
- Marabini, R., Herman, G.T., Carazo, J.-M., 1998. 3D reconstruction in electron microscopy using ART with smooth spherically symmetric volume elements (blobs). *Ultramicroscopy* 72, 53–65.
- Marabini, R., Rietzel, E., Schröder, R., Herman, G.T., Carazo, J.-M., 1997. Three-dimensional reconstruction from reduced sets of very noisy images acquired following a single-axis tilt scheme: application of a new three-dimensional reconstruction algorithm and objective comparison with weighted backprojection. *J. Struct. Biol.* 120, 363–371.
- McIntosh, R., Nicastro, D., Mastronarde, D., 2005. New views of cells in 3D: an introduction to electron tomography. *Trends Cell Biol.* 15, 43–51.
- Meyer, R.R., Kirkland, A., 1998. The effects of electron and photon scattering on signal and noise transfer properties of scintillators in CCD cameras used for electron detection. *Ultramicroscopy* 75, 23–33.
- Mobus, G., Schweinfest, R., Gemming, T., Wagner, T., Rühle, M., 1998. Iterative structure retrieval techniques in HREM: a comparative study and a modular program package. *J. Microsc.* 190, 109–130.
- Müller, H., Rose, H., Schorsch, P., 1998. A coherence function approach to image simulation. *J. Microsc.* 190, 73–88.
- Nickell, S., Förster, F., Linaroudis, A., Del Net, W., Beck, F., Hegerl, R., Baumeister, W., Plitzko, J.M., 2005. TOM software toolbox: acquisition and analysis for electron tomography. *J. Struct. Biol.* 149, 227–234.
- Peng, L.-M., Dudarev, S.L., Whelan, M.J., 1998. Electron scattering factors of ions and dynamical RHEED from surfaces of ionic crystals. *Phys. Rev. B* 57, 7259–7265.
- Peng, L.-M., Dudarev, S.L., Whelan, M.J., 2004. High-energy electron diffraction and microscopy. In: *Monographs on the Physics and Chemistry of Materials*. vol. 61. Oxford University Press.
- Peng, L.-M., Ren, G., Dudarev, S.L., Whelan, M.J., 1996. Robust parameterization of elastic and absorptive electron atomic scattering factors. *Acta Crystallogr. A* 52, 257–276.
- Philippsen, A., Engel, H., Engel, A., 2007. The contrast-imaging function for tilted specimens. *Ultramicroscopy* 107, 202–212.
- Reichert, R., Engel, A., 1984. Monte carlo calculations of elastic and inelastic electron scattering in biological and plastic materials. *Ultramicroscopy* 13, 279–293.
- Reimer, L., Kohl, H., 2008. *Transmission electron microscopy*. Springer Series in Optical Sciences, fifth ed., vol. 36. Springer-Verlag.
- Rullgård, H., Öfverstedt, L.-G., Masich, S., Daneholt, B., Öktem, O., 2011. Simulation of transmission electron microscope images of biological specimens. *J. Microsc.* 242, 234–256.
- Sali, A., Glaeser, R.M., Earnest, T., Baumeister, W., 2003. From words to literature in structural proteomics. *Nature* 422, 216–225.
- Saxton, O., 1997. Quantitative comparison of images and transforms. *J. Microsc.* 190, 52–60.
- Shaikh, T.R., Gao, H., Baxter, W.T., Asturias, F.J., Boisset, N., Leith, A., Frank, F., 2008. SPIDER image processing for single-particle reconstruction of biological macromolecules from electron micrographs. *Nat. Protoc.* 3, 1941–1974.
- Shang, Z., Sigworth, F.J., 2012. Hydration-layer models for cryo-EM image simulation. *J. Struct. Biol.* 180, 10–16.
- Sorzano, C.O.S., Marabini, R., Boisset, N., Rietzel, E., Schröder, R., Herman, G.T., Carazo, J.-M., 2001. The effect of overabundant projection directions on 3D reconstruction algorithms. *J. Struct. Biol.* 133, 108–118.
- Sorzano, C.O.S., Marabini, R., Velázquez-Muriel, J., Bilbao-Castro, J., Scheres, S.H.W., Carazo, J.M., Pascual-Montano, A., 2004. XMIPP: a new generation of an open-source image processing package for electron microscopy. *J. Struct. Biol.* 148, 194–204.
- Sun, S., Shi, S., Leapman, R., 1993. Water distributions of hydrated biological specimens by valence electron energy loss spectroscopy. *Ultramicroscopy* 50, 127–139.
- Tang, D., Kirkland, A., Jefferson, D., 1993. Optimisation of high-resolution image simulations: I. Image selection in real space. *Ultramicroscopy* 48, 321–331.
- Tang, D., Kirkland, A., Jefferson, D., 1994. Optimisation of high-resolution image simulations: II. Image selection in reciprocal space. *Ultramicroscopy* 53, 137–146.
- Tang, G., Peng, L., Baldwin, P.R., Mann, D.S., Jiang, W., Rees, I., Ludtke, S.J., 2007. EMAN2: an extensible image processing suite for electron microscopy. *J. Struct. Biol.* 157, 38–46.
- Teeter, M.M., 1984. Water structure of a hydrophobic protein at atomic resolution: pentagon rings of water molecules in crystals of crambin. *Proc. Natl. Acad. Sci.* 81, 6014–6018.
- Thust, A., 2009. High-resolution transmission electron microscopy on an absolute contrast scale. *Phys. Rev. Lett.* 102, 2208011–2208014.
- Vinogradov, S.N., Sharma, P.K., 1994. Preparation and characterization of invertebrate globin complexes. In: Everse, J., Vandegriff, K.D., Winslow, R.M. (Eds.), *Hemoglobins Part B: Biochemical and Analytical Methods*. Academic Press, pp. 112–124 (Chapter 8).
- Voortman, L.M., Franken, E.M., van Vliet, L.J., Rieger, B., 2012. Fast, spatially varying CTF correction in TEM. *Ultramicroscopy* 118, 26–34.
- Voortman, L.M., Stallinga, S., Schoenmakers, R.H.M., van Vliet, L.J., Rieger, B., 2011. A fast algorithm for computing and correcting the CTF for tilted, thick specimens in TEM. *Ultramicroscopy* 111, 1029–1036.
- Vriend, G., 1990. WHAT IF: a molecular modeling and drug design program. *J. Mol. Graph.* 8, 52–56.
- Vulović, M., Franken, E., Ravelli, R.B.G., van Vliet, L.J., Rieger, B., 2012. Precise and unbiased estimation of astigmatism and defocus in transmission electron microscopy. *Ultramicroscopy* 116, 115–134.
- Vulović, M., Rieger, B., van Vliet, L.J., Koster, A.J., Ravelli, R.B.G., 2010. A toolkit for the characterization of CCD cameras for transmission electron microscopy. *Acta Crystallogr. D* 66, 97–109.
- Vulović, M., Voortman, L.M., van Vliet, L.J., Rieger, B., 2013. When to use the projection assumption and weak-phase object approximation in phase contrast cryo-EM (submitted for publication).
- Wall, J., Isaacson, M., Langmore, J.P., 1974. Collection of scattered electrons in dark field electron microscopy – 2. Inelastic scattering. *Optik* 39, 359–374.
- Wan, Y., Chiu, W., Zhou, Z.H., 2004. Full contrast transfer function correction in 3D cryo-EM reconstruction. In: *IEEE Proceedings of ICCAS 2004*, vol. 2, pp. 960–964.
- Wang, Y., Chiu, J., He, Q., 2005. Proteomics in computer-aided drug design. *Curr. Computer-Aided Drug Des.* 1, 43–52.
- Winkler, H., Taylor, K.A., 2003. Focus gradient correction applied to tilt series image data used in electron tomography. *J. Struct. Biol.* 143, 24–32.