Published in final edited form as:

J Struct Biol. 2016 April ; 194(1): 49–60. doi:10.1016/j.jsb.2016.01.012.

Statistical Modeling and Removal of Lipid Membrane Projections for Cryo-EM Structure Determination of Reconstituted Membrane Proteins

Katrine Hommelhoff Jensen^a, Sami Sebastian Brandt^a, Hideki Shigematsu^{b,c}, and Fred J. Sigworth^b

^aDepartment of Computer Science, University of Copenhagen, Universitetsparken 5, DK-2100 Copenhagen, Denmark

^bDepartment of Cellular and Molecular Physiology, Yale University School of Medicine, 333 Cedar St., New Haven, CT 06520, USA

^cPresent address: RIKEN Center for Life Science Technologies, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Abstract

This paper describes steps in the single-particle cryo-EM 3D structure determination of membrane proteins in their membrane environment. Using images of the Kv1.2 potassium-channel complex reconstituted into lipid vesicles, we describe procedures for the merging of focal-pairs of exposures and the removal of the vesicle-membrane signal from the micrographs. These steps allow 3D reconstruction to be performed from the protein particle images. We construct a 2D statistical model of the vesicle structure based on higher-order singular value decomposition (HOSVD), by taking into account the structural symmetries of the vesicles in polar coordinates. Non-roundness in the vesicle structure is handled with a non-linear shape alignment to a reference, which ensures a compact model representation. The results show that the learned model is an accurate representation of the imaged vesicle structures. Precise removal of the strong membrane signals allows better alignment and classification of images of small membrane-protein particles, and allows higher-resolution 3D reconstruction.

Keywords

liposome; single-particle reconstruction; statistical shape modeling; potassium channel; Kv1.2; higher-order singular value decomposition

1 Introduction

Random spherically-constrained (RSC) reconstruction is a single-particle cryo-EM structure-determination method for membrane proteins reconstituted in spherical lipid vesicles (Wang and Sigworth 2010). Images of individual membrane protein complexes can be extracted and processed as single-particle images, but a very important consideration is the strong signal contributed to particle images by the vesicle membrane. This signal is particularly strong for “side views” of particles which are observed near the limb of the vesicle, where the scattering density of the membrane is maximal (see Fig. 1d).

The 3D reconstruction of a membrane-protein particle embedded in a spherical vesicle is similar to the problem of reconstruction in the presence of mismatched symmetries, for example the reconstruction of a single copy of the exit pore machinery in an otherwise icosahedrally-symmetric virus particle (Agirrezabala, Martin-Benito, Caston, Miranda et al. 2005, Guo and Jiang 2014). In practice however, lipid vesicles have widely variable diameters, which makes reconstruction of an entire vesicle-plus-particle system impractical. On the other hand, structures of large membrane-pore complexes have been obtained (Tilley, Orlova, Gilbert, Andrew and Saibil 2005, Lukoyanova, Kondos, Farabella, Law, Reboul et al. 2015) through the use of side-view images of particles that include membrane density. Because of the large particle size, the signal from the protein in these images predominates over the inconsistent membrane signal so that reliable alignment of images and determination of projection directions is possible.

We consider here the structure determination of relatively small protein particles, about 100 Å in size, embedded in vesicles 300-600 Å in diameter. To allow reliable alignment of the weak particle signals, the membrane contribution to each cryo-EM image is modeled and subtracted. With the assumption of linear superposition of lipid and protein densities – an assumption that appears to be adequately satisfied – the residual protein-particle images can be made sufficiently free of membrane artifacts that they can be aligned and used in single-particle reconstruction (SPR) to obtain a membrane-protein structure. Subsequent re-addition of the membrane density then yields a three-dimensional density map of the membrane protein embedded in the vesicle membrane.

Fig. 1 illustrates the processing of images of the Kv1.2 potassium-channel complex incorporated into lipid vesicles. The channel protein complex is 400 kDa in size, but 40% of that mass lies in the transmembrane region where it is nearly contrast-matched by the membrane lipid and therefore contributes little to particle alignment. We increase the visibility of particles by acquiring focal-pairs of exposures and merge them to form a composite micrograph (Fig. 1c). In this merged image the centers and approximate radii of vesicle images are determined, and particles are selected (Fig. 1d) as described previously (Liu and Sigworth 2014). Meanwhile, a model of the vesicle projections is constructed (Fig. 1e) and used to remove each vesicle's membrane projection from the micrograph (Fig. 1f). The subsequent processing steps are the extraction of individual particle images (Fig. 1g) and 3D reconstruction (Fig. 1h). Replacing the particle images with the unsubtracted ones in the final reconstruction yields the finished density map (Fig. 1i).

The present paper is concerned with the modeling of the vesicle-membrane images. The goal is to allow quantitative removal of the membrane density from images so that particle-image alignment and orientation assignment is not affected by residual membrane signal. The removal need not be perfect as the final 3D map is generated from particle images with the membrane signal present. Nevertheless the modeling and removal is challenging because, at the limb of a spherical vesicle's projection image, the membrane signal is much stronger than the protein particle's signal and often overlaps substantially with it.

We first introduce the images used in this work (Section 2.1) and the image merging procedure (Section 2.2). Then we describe a physical model of a vesicle image, calculated from a 3D physical model of spherical shells of lipid density (Wang, Bose and Sigworth 2006) (Section 2.3). The 2D projection of this spherical or nearly-spherical membrane model is filtered by the known contrast-transfer function (CTF) to provide a simulation of the vesicle image.

In Section 3 we then describe the main part of the methods developed in this work, the modeling of 2D images of vesicles directly. Based on the approximately scale-invariant properties of the simulated images, and our earlier work (Jensen, Sigworth and Brandt 2016), it is the application of multilinear image analysis in learning and modeling the projections of vesicle membranes. The vesicle-membrane density is modeled as approximately circularly-symmetric rings of constant image intensity, but also included is a non-linear alignment procedure step (in contrast to Jensen et al. 2016) to make highly accurate fits to vesicle-membrane projections that show a certain degree of asymmetric shapes. This approach provides more satisfactory modeling and removal of the membrane signal and allows high-quality 3D reconstruction to be performed with the residual particle images.

2 Materials

2.1 Experimental images

The Kv1.2 “paddle chimera” potassium channel complex (Long, Tao, Campbell and MacKinnon 2007) is 400 kDa in size. It was expressed in *Pichia pastoris* and purified as described by Long, Campbell and MacKinnon (2005); and Long et al. (2007) with some modifications. The construct (a gift of Dr. Yufeng Zhou, University of Pennsylvania) contained a Strep affinity tag at the N-terminus of the Kv1.2 alpha subunit. Membranes were solubilized in dodecylmaltoside detergent, and affinity purification made use of a Strep-Tactin affinity column (IBA GmbH) and was followed by size-exclusion chromatography. Protein was mixed with decylmaltoside-solubilized lipids (POPC:DMPS:Cholesterol 8:1:1) to yield a protein-to-lipid molar ratio about 1:1000 and dialyzed against detergent-free buffer for four days. Vesicles were adsorbed to a thin carbon film, washed with buffer having osmolarity lowered by 20 mOsm to encourage swelling, blotted and plunge-frozen in liquid ethane. Images were taken at 200 keV with an FEI Tecnai F20 microscope equipped with a Gatan K2 electron-counting camera and controlled by the SerialEM program (Mastronarde 2005). Movies were taken at a pixel size of 1.25 Å and a dose rate of 10 e/(pixel · s), giving a total of 30 frames with 1.6 e/Å² of dose per frame, thus 48 e/Å² in total. Initial underfocus was in the range 1.4–2.9 μm but at about frame 18 of each movie the underfocus was

increased by 6 μm . Motion-correction (Shigematsu and Sigworth 2013) was applied separately to the two sections of each movie, and focal-pair images were obtained as the exposure-filtered sums of the sections.

2.2 Image merging

Both for improved visibility of small proteins and for improved fitting of vesicle-membrane density, we combine focal pairs of micrographs. Conway and Steven (1999); and Ludtke and Chiu (2003) have described this process, and ours differs only in details of the image alignment process and in the way in which the weighting factors of images are normalized. A low-defocus image, the first one acquired, provides all of the high-frequency information, while additional low-frequency information is provided, up to the first or second zero of its CTF by a second, high-defocus image (Fig. 2a). The second image is aligned to the first using a general affine transformation whose six coefficients account for translation, change in magnification, rotation and skew. The target function in the alignment is the cross-correlation of the two images, filtered by a function $h(\mathbf{s})$ designed to enhance spatial frequencies of overlap between the CTFs c_1 and c_2 and to give weight to higher spatial frequencies $\mathbf{s}$ according to $h(\mathbf{s}) = c_1 c_2 |\mathbf{s}|$. Once aligned, the images are merged according to weights that are normalized to preserve the spectral density of background noise. In the Fourier domain the merged image $\hat{M}$ is obtained from the exposures $\hat{I}_i$ by a modification of “exposure filtering” (Grant and Grigorieff 2015) to include the effect of differing CTFs,

$$\hat{M}(\mathbf{s}) = \sum_i w_i(\mathbf{s}) \hat{I}_i(\mathbf{s}), \quad (1)$$

where w_i is a frequency dependent weight given by

$$w_i(\mathbf{s}) = \frac{c_i (e^{-d_i/2N_c(\mathbf{s})} - e^{-d_{i+1}/2N_c(\mathbf{s})})}{\sqrt{\sum_j c_j^2 (e^{-d_j/2N_c(\mathbf{s})} - e^{-d_{j+1}/2N_c(\mathbf{s})})^2}} \quad (2)$$

with c_i being the CTF and d_i, d_{i+1} the cumulative electron dose at the beginning and end of exposure i ; N_c is the frequency-dependent critical dose (Hayward and Glaeser 1979). We have used the following function for this quantity, modified from a fit to the data by Baker, Smith, Bueler and Rubinstein (2010), in units of $\text{e}/\text{\AA}^2$ by

$$N_c(\mathbf{s}) = 5 + \frac{80}{(s/.014)^2 + 1} \quad (3)$$

where $\mathbf{s}$ is spatial frequency in $\text{\AA}^{-1}$. A function having larger values of N_c at low frequencies (as in Grant and Grigorieff 2015) would seem to fit our observations better.

In the case of movie data acquired from a direct-electron detector, exposure filtering (Grant and Grigorieff 2015), also known as damage compensation (Wang, Hryc, Bammes, Afonine et al. 2014), is applied to the sets of movie frames obtained at the two defocus values. The net result of the combination of exposure filtering and merging for movie data is that all of

the frames are accumulated according to eqns. (1) and (2) where in this case the $\hat{I}_i$ represent the individual movie frames.

2.3 Physical model of vesicle images

The image of a lipid vesicle can be built up from the projected density of thin spherical shells H of radius ρ and thickness $d\rho$,

$$H(\mathbf{x}, \rho) = \frac{2\rho d\rho}{\sqrt{\rho^2 - |\mathbf{x}|^2}}, |\mathbf{x}| < \rho \quad (4)$$

where $\mathbf{x}$ is the pixel coordinate. Assuming a membrane of total thickness t having a cross-sectional density that does not depend on vesicle radius, the i^{th} shell is weighted by the inner potential v_i (Fig. 2b) which determines the degree of interaction with incident electrons. The result is a two-dimensional projected potential

$$V(\mathbf{x}) = \sum_{i=-t/2}^{t/2} v_i H(\mathbf{x}, \rho_0 + i) \quad (5)$$

which is illustrated in Fig. 2c-e. The modeled image of a vesicle is then obtained by filtering with the CTF operator c and scaling by the interaction factor σ such that the phase-contrast image of a vesicle is obtained as the ratio of the detected and incident intensity

$I_v(\mathbf{x})/I_0 = 1 + 2\sigma c \circ V(\mathbf{x})$ at the electron detector. For 200 keV electrons the true value of the interaction factor σ_0 is 7.3×10^{-4} rad/V for 200 keV electrons (Kirkland 1998). Through the fitting of the physical model to experimental vesicle images, one obtains an effective value of σ that serves as a local measure of image contrast. This contrast is reduced to a variable extent by the ice thickness.

Fig. 2h plots the radial intensity in the simulated images of a “large” and a “small” vesicle, of diameters 300 and 600 Å. Although the underlying radial density profiles of these vesicles are not identical (see Fig. 2e), the CTF filtered images in Fig. 2f and g appear almost like annular rings of density: the main feature is a circumferential dark band of 40 Å width, surrounded on the outside by a bright ring or overshoot of intensity. The form of the radial profile of intensity is nearly independent of vesicle radius (Fig. 2h) and in the merged image set that we used, the features are insensitive to the variations in CTF due to varying defocus values as well. The invariance of these features suggested to us that a simplified approach to modeling vesicle images is possible. Instead of the explicit computation of the projection of a 3D vesicle model, it should be possible to model the 2D image of a vesicle directly from circularly-symmetric features like these. This is the approach underlying the vesicle-removal method described below.

3 Methods

3.1 Overview of the Approach

Jensen et al. (2016) presented a method to accurately separate the vesicle signals from the micrograph that was aimed at improving the protein reconstruction by reducing vesicle artifacts near the protein. The work was built upon the principle of first learning the vesicle subspace from the vesicle structures in the micrographs, and then projecting the experimental micrographs onto the orthogonal complement of the subspace, which is the residual subspace. The residual subspace represents everything which is not true vesicle structure, i.e. protein particles and noise. The vesicle subspace, defined by orthogonal basis shapes, was learned by the higher-order singular value decomposition (HOSVD) augmented with principal component analysis (PCA). The circular symmetry of the vesicles was taken into account by making the decomposition of individual vesicles in the polar coordinate plane, where the data were normalized for vesicle position, size, and the signal magnitude.

Here, we propose an improved vesicle model for cases where the assumption of circular symmetry is not sufficient, as is the case in the dataset considered here. The non-roundness of the vesicles is handled by non-linear shape alignment of the vesicle polar coordinate representation to the mean, which is iteratively updated to an accurate representative of the aligned vesicles. The non-linear shape alignment is principally achieved by fitting a Fourier series to the radial location of the inner wall and warping it onto a constant position. In the construction of the non-linear warping, it is additionally assumed that the vesicle wall has a constant absolute thickness.

3.2 Formal Goal

To separate a protein from vesicle signal, we aim at learning the vesicle structure subspace from the micrographs and remove the vesicle signal in the least squares sense, given the learned vesicle basis elements. In practice, we construct the statistical vesicle model from a stack of windowed vesicle images. The vesicle model is used to construct the projection $\mathbf{P}$ onto the vesicle subspace to compute the model

$$\hat{\mathbf{f}}^{(n)} = \mathbf{P}\mathbf{f}^{(n)}, \quad (6)$$

where $\mathbf{f}^{(n)}$ is the original vectorized, discrete image of the n^{th} vesicle. For the vesicle removal, the vectorized, discrete residual image is computed as

$$\mathbf{r}^{(n)} = \mathbf{f}^{(n)} - \hat{\mathbf{f}}^{(n)}. \quad (7)$$

3.3 Vesicle Representation

The micrographs are first intensity normalized by subtracting the locally estimated background level $b^{(n)}$ for each pixel and dividing the result by the local, pixel-wise vesicle signal level $s^{(n)}$, or,

$$\mathbf{f}_{\text{norm}}^{(n)} = \frac{1}{s^{(n)}} (\mathbf{f}^{(n)} - b^{(n)}). \quad (8)$$

The signal and background levels are estimated as the mean absolute deviation of the windowed vesicle pixel values and the local background pixel values, respectively; see Jensen et al. (2016) for details.

As soon as the intensity normalized vesicle image $\mathbf{f}_{\text{norm}}^{(n)}$ is identified, the geometrically normalized vesicle image is obtained as

$$\mathbf{f}_{\text{geom}}^{(n)} = \mathbf{T}^{(n)} \mathbf{f}_{\text{norm}}^{(n)}, \quad (9)$$

where the matrix $\mathbf{T}^{(n)}$ constitutes (1) transformation into polar coordinates, (2) non-linear shape alignment, and (3) weighted resampling, including soft-windowing, of the aligned polar-coordinate map from the Cartesian grid, as illustrated in Fig. 4.

The non-linear shape alignment refers to the correction of the vesicles for non-circular shapes. This is achieved by modeling the inner radius of the vesicle by a Fourier series in the function of the polar angle, or,

$$R_{\text{in}}(\alpha) = R_{0,\text{in}} + \delta R(\alpha), \quad (10)$$

where $\delta R(\alpha) = \sum_{l=1}^L a_l \cos \alpha + b_l \sin \alpha$. To achieve the non-linear alignment, the normalized polar coordinate plane is mapped to the corrected one where $R'_{\text{in}}(\alpha) = R_{0,\text{in}}$ (Fig. 4). In the mapping, for each α , the radii smaller than the inner radius are linearly stretched while the larger values are only shifted to preserve vesicle wall thickness, see Fig. 5a. The rationale for the mapping is the fact that we intend to map vesicle centre to itself ($\rho = \rho' = 0$) while preserving the absolute vesicle wall thickness. The resampling involves weighting and windowing in the function of the normalized radius, as illustrated in Fig. 5b.

3.4 Vesicle Modeling and Removal

The vesicle model is constructed by learning a set of orthogonal basis elements, which are each separable into an outer product between the radial and angular part. The basis elements are obtained by decomposing the three-dimensional, $I_1 \times I_2 \times I_3$ array $\mathbf{A} - \mathbf{A}_0$ by the higher-order singular value decomposition (HOSVD) (Lathauwer, Moor and Vandewalle 2000), see Jensen et al. (2016), where the image stack $\mathbf{A}$ contains the normalized vesicle images resampled in the polar coordinates, $\mathbf{A}_0$ is the mean, reference array over the vesicle population, and the dimensions I_1 , I_2 and I_3 correspond to the radial variable ρ , angular variable α , and the training vesicle index j , respectively. The basis is further compressed by adding a PCA layer to find the most descriptive paired 2D bases while the least significant new basis elements are truncated. This yields the model

$$\hat{\mathbf{f}}_{\text{geom}} = \mathbf{g}_0 + \mathbf{G}\xi, \quad (11)$$

where $\mathbf{g}_0$ is the mean of the normalized, polar coordinate vesicle images, $\mathbf{G}$ is the truncated, orthogonal basis, and ξ the coordinates of the vesicle in the basis, see (Jensen et al. 2016) for details.

For a novel micrograph, the normalized polar coordinate representation $\mathbf{f}_{\text{geom}}^{(n)}$ of a vesicle is orthogonally projected onto the vesicle subspace that yields the normalized model

$$\hat{\mathbf{f}}_{\text{geom}}^{(n)} = \mathbf{g}_0 + \tilde{\mathbf{P}}(\mathbf{f}_{\text{geom}}^{(n)} - \mathbf{g}_0), \quad (12)$$

where $\tilde{\mathbf{P}} = \mathbf{G}\mathbf{G}^T$ is the projection onto the mean corrected, normalized vesicle subspace. The direct vesicle removal is then achieved by reversing shape and intensity normalization for the model and subtracting the result from the micrograph.

4 Results

As the training set, 585 vesicles were extracted from 13 merged micrographs, each binned by 2 into 1920×1920 pixels for a fixed size of $2.5 \text{ \AA}$ per pixel. The vesicle center positions and average radii were estimated as reported by Liu and Sigworth (2014). As the window function (Fig. 5b), we used the Tukey window (Tukey 1967), where its transition region was set to start a few pixels outside the normalized vesicle wall on the bright aura. The nonlinear shape alignment was carried out by optimizing the first $L = 6$ Fourier series coefficients in (10), by minimizing the least squares distance from each vesicle to the mean non-linearly aligned windowed vesicle, while iteratively updating the mean 3 times. The non-linear minimization was performed by Levenberg–Marquardt (Levenberg 1944, Marquardt 1963) optimization in the coarse-to-fine fashion, where the polar coordinate plane was blurred by a Gaussian kernel with standard deviation $\sigma = 3.1$. The optimization was initialized by the result of a rough discrete search for the Fourier coefficients. To avoid undersampling, the normalized vesicle radius was set to 80 pixels which was the radius of the largest relevant vesicle in the data set. The vesicle wall, including the bright aura on both sides (see Fig. 9b, d), was assumed to have the constant thickness of $W = 55$ pixels in the binned micrographs. Some alignment results are illustrated in Fig. 6.

The $I_1 \times I_2 \times I_3 = 103 \times 376 \times 585$ normalized stack was then decomposed by the two layer HOSVD model with $(\tilde{I}_1, \tilde{I}_2, K) = (8, 3, 24)$; that is, the model consisted of 24 basis shapes composed of each of the 8 most significant radial bases paired with each of the 3 most significant angular bases. This choice led to the fewest effective parameters while the residual level matched with the local background level (Jensen et al. 2016) in removal experiments with novel vesicles. The mean vesicle $\mathbf{g}_0$ and components in $\mathbf{G}$ of the model are illustrated in Fig. 7. It shows that the radial components are dominant over the angular ones and represent the basic vesicle shape, whereas the angular components primarily account for simple brightness changes. Further, some qualitative removal results for the vesicles aligned in Fig. 6, are shown in Fig. 8. The projections onto the vesicle subspace (Fig. 8c) and its

orthogonal complement (Fig. 8e) clearly illustrate the separation of vesicle structures and noise. Moreover, the vesicle removal leaves the background and structural noise intact together with the protein signal of interest. One may likewise see that the nonlinear irregularities of the vesicle wall was successfully modeled.

From the original set of 330 merged micrographs containing about 20,000 vesicles, called the unsubtracted (UN) set, two additional sets were created. One came from subtracting vesicles from each micrograph as modeled by the physical model (PH) for spherical vesicles, and the other from subtraction using the statistical vesicle image model (SM). A region of one micrograph from each set is illustrated in Fig. 9. Particle images were extracted using the coordinates of 8019 semi-automatically picked particles, and after visual pruning three stacks of 7042 (UN), 7603 (PH) and 6910 (SM) particle images were obtained. These were first subjected to 2D classification using maximum-likelihood multi-reference alignment with CTF correction, essentially the same as performed in Relion (Scheres 2012) but with a prior probability reflecting the vesicle geometry, and starting with random references. The results are shown in Fig. 10. The membrane density is seen to predominate in the class averages from unsubtracted images, in some cases leaving the protein particle density very indistinct (Fig. 10a). Protein particles are more clearly seen in the PH set, but ring-like artifacts from incomplete vesicle subtraction are prominent (Fig. 10b). Much better removal of membrane artifacts is seen in the SM set (Fig. 10c).

The critical question is whether the vesicle subtraction is sufficient to allow 3D reconstruction of the membrane-embedded protein particles. The published X-ray crystal structure of the Kv1.2 ion channel complex was determined by Long et al. (2007) in the presence of detergent rather than in a lipid membrane, but is thought to represent the native structure. The complex consists of four alpha and four beta subunits (Fig. 11a). Each alpha subunit contributes six transmembrane helices (at top) and part of the T1 domain “stalk”. Beta subunits form the lobes of the “mushroom” at the bottom (intracellular end). A truly native structure of Kv1.2 in a membrane will show membrane density as well.

We performed 3D reconstructions from each of the datasets, using 5744 particle images that were common to all. The 3D reconstructions started with the Kv1.2 crystal structure filtered to 60 Å and were refined using a maximum-likelihood method that employs Fourier reconstruction (Barthel, Tagare and Sigworth 2011). It is similar to that of Relion (Scheres 2012) but the target function includes a prior probability term that provides a constraint to the particle's orientation based on its position relative to the center of its vesicle. Starting with a null model this prior dominates in the first few iterations of the expectation-maximization (EM) algorithm, but has little influence at convergence of the final 3D model.

Figure 11b and c show reconstructions obtained from the particle stacks using the PH and SM vesicle subtraction. Some details of the protein structure are visible, including some secondary structure such as helix hairpins visible in the beta subunits. As expected, the protein density in the trans-membrane region is reduced due to the membrane subtraction. The two reconstructions are similar but the SM reconstruction shows more detail in the transmembrane region. Parts d-f of the figure show reconstructions with the membrane density present. Figure 11d shows the reconstruction obtained using the UN particle images

with no vesicle subtraction at all. Although the membrane is well defined in the 3D map, the protein density is blurred as expected from the predominance of membrane signal in the alignment process. The application of the Fourier shell correlation (FSC) is not straightforward to reconstructions that include membrane density, so we evaluated local resolution using ResMap (Kucukelbir, Sigworth and Tagare 2014). In that analysis (Figure 11g) only 16% of the protein particle volume showed resolution better than 16 Å.

Figure 11e shows the reconstruction from the PH dataset but with the membrane restored as follows. After the final iteration of the reconstruction with the subtracted dataset (part b of the Figure), the particle alignment and orientation parameters were transferred to the UN dataset for another Fourier reconstruction. To the resulting map, a cylindrical mask was applied to select a small disc of membrane surrounding the particle's transmembrane region, and the masked map is shown in the figure. A local resolution map (Figure 11h) shows that 48% of the particle volume has resolution better than 12 Å.

Figure 11f shows the corresponding membrane-restored result from the SM-subtracted dataset. Its resolution map (Figure 11i) shows more extended regions of high-resolution information, with 58% of the volume having resolution values better than 12 Å. We conclude that the superior SM vesicle subtraction produces a substantial benefit in reconstruction, but even the simple PH vesicle subtraction is vastly superior to no subtraction at all. We further conclude that for a relatively small protein particle like this one (400 kDa molecular weight total, of which 250 kDa is outside the membrane) vesicle subtraction is essential for 3D reconstruction.

5 Discussion

The challenge of reconstructing protein particles embedded in vesicle membranes is that the vesicle size is variable, and the inconsistent membrane density can interfere with the alignment and classification of single-particle images. In cases where the protein complexes are large and their signal predominates in the particle images, no subtraction has been needed (Tilley et al. 2005, Lukoyanova et al. 2015). In cryo-EM images of the relatively small Kv1.2 channel complex, the protein particle's contribution is quite weak compared to the membrane density, and attempts at classification (Fig. 10a) and 3D reconstruction (Fig. 11d) fail because the membrane density interferes with the alignment of the protein images. A simple physical model of the vesicle membrane, which assumes an invariant membrane cross-sectional density and spherical geometry, is successful in reducing the membrane density but leaves small residual artifacts (Fig. 9) which are still substantial in comparison to the weak protein signal and contaminate class-averaged images and reduce the resolution of a final 3D reconstruction (Fig. 11h).

The physical model is inherently three-dimensional, as it is based on computing the projection image of spherical shells of density. In this paper we have considered an alternative approach in which the 2D projected images of vesicles are modeled directly, using a statistical model that is developed from a training set of micrographs. We find that with this model the protein signal is well preserved, while the vesicle structure is accurately

modeled so that membrane-density residuals contribute much less to class-averaging and reconstructions than in the case of the simple physical model.

The statistical vesicle model is very effective because, while it assumes nearly circular symmetry, it includes a nonlinear correction for deviations from the ideal symmetry. In the model the radial components are dominant over the angular components and cover the basic vesicle shape, whereas the angular components primarily account for density gradients in the micrographs. Meanwhile the protein structure is well preserved because the constructed vesicle subspace has low dimension, and little of the protein signal falls into it. The result is membrane-signal artifacts that are of low amplitude and low resolution, and which appear to have negligible effects on the quality of particle alignment.

6 Software

The Matlab source code of the method will become available at authors' web page.

Acknowledgments

This work has been partly supported by the Centre for Stochastic Geometry and Advanced Bioimaging funded by the Villum Foundation, Denmark; JSPS KAKENHI, grant no.15H01656; and National Institutes of Health grants R21 NS085776 and R01 NS021501.

References

- Agirrezabala X, Martin-Benito J, Caston JR, Miranda R, et al. Maturation of phage ϕ 7 involves structural modification of both shell and inner core components. *EMBO J*. 2005; 24:3820–3829. [PubMed: 16211007]
- Baker LA, Smith EA, Bueler SA, Rubinstein JL. The resolution dependence of optimal exposures in liquid nitrogen temperature electron cryomicroscopy of catalase crystals. *J Struct Biol*. 2010; 169:431–437. [PubMed: 19958834]
- Barthel AC, Tagare H, Sigworth FJ. Surface-constrained 3d reconstruction in cryo-em. *Conf Rec Asilomar Conf Signals Syst Comput*. 2011:1026–1030. [PubMed: 24477184]
- Conway JF, Steven AC. Methods for reconstructing density maps of “single” particles from cryoelectron micrographs to subnanometer resolution. *J Struct Biol*. 1999; 128:106–118. [PubMed: 10600565]
- Grant T, Grigorieff N. Measuring the optimal exposure for single particle cryo-em using a 2.6 reconstruction of rotavirus vp6. *Elife*. 2015; 4
- Guo F, Jiang W. Single particle cryo-electron microscopy and 3-d reconstruction of viruses. *Methods Mol Biol*. 2014; 1117:401–443. [PubMed: 24357374]
- Hayward SB, Glaeser RM. Radiation damage of purple membrane at low temperature. *Ultramicroscopy*. 1979; 4:201–210. [PubMed: 473421]
- Jensen KH, Sigworth FJ, Brandt SS. Removal of vesicle structures from transmission electron microscope images. *IEEE Transactions on Image Processing*. 2016; 25(2):540–552.
- Kirkland, EJ. *Advanced Computing in Electron Microscopy*. Plenum Press; New York: 1998.
- Kucukelbir A, Sigworth F, Tagare H. Quantifying the local resolution of cryo-em density maps. *Nat Methods*. 2014; 11(1):63–65. [PubMed: 24213166]
- Lathauwer LD, Moor BD, Vandewalle J. A multilinear singular value decomposition. *SIAM J Matrix Anal Appl*. 2000; 21(4):1253–1278.
- Levenberg K. A method for the solution of certain non-linear problems in least squares. *Quarterly of Applied Mathematics*. 1944; 2:164–168.
- Liu Y, Sigworth FJ. Automatic cryo-em particle selection for membrane proteins in spherical liposomes. *J Struct Biol*. 2014; 185:295–302. [PubMed: 24468290]

- Long SB, Campbell EB, Mackinnon R. Crystal structure of a mammalian voltage-dependent shaker family k⁺ channel. *Science*. 2005; 309(5736):897–903. [PubMed: 16002581]
- Long SB, Tao X, Campbell EB, MacKinnon R. Atomic structure of a voltage-dependent k⁺ channel in a lipid membrane-like environment. *Nat Methods*. 2007; 450(7168):376–382.
- Ludtke SJ, Chiu W. Focal pair merging for contrast enhancement of single particles. *J Struct Biol*. 2003; 144:73–78. [PubMed: 14643210]
- Lukoyanova N, Kondos SC, Farabella I, Law RH, Reboul CF, et al. Conformational changes during pore formation by the perforin-related protein pleurotolysin. *PLoS Biol*. 2015; 13
- Marquardt D. An algorithm for least-squares estimation of nonlinear parameters. *SIAM Journal on Applied Mathematics*. 1963; 11:431–441.
- Mastronarde DN. Automated electron microscope tomography using robust prediction of specimen movements. *J Struct Biol*. 2005; 152:36–51. [PubMed: 16182563]
- Scheres SHW. Relion: Implementation of a bayesian approach to cryo-em structure determination. *J Struct Biol*. 2012; 180(3):519–530. [PubMed: 23000701]
- Shigematsu H, Sigworth FJ. Noise models and cryo-em drift correction with a direct-electron camera. *Ultramicroscopy*. 2013; 131:61–69. [PubMed: 23748163]
- Tilley SJ, Orlova EV, Gilbert RJ, Andrew PW, Saibil HR. Structural basis of pore formation by the bacterial toxin pneumolysin. *Cell*. 2005; 121:247–256. [PubMed: 15851031]
- Tukey JW. An introduction to the calculations of numerical spectrum analysis. *Spectral Analysis of Time Series*. 1967:25–46.
- Wang L, Bose PS, Sigworth FJ. Using cryo-em to measure the dipole potential of a lipid membrane. *National Academy of Sciences of the United States of America*. 2006; 103(4):129–138.
- Wang L, Sigworth FJ. Liposomes on a streptavidin crystal: a system to study membrane proteins by cryo-em. *Methods Enzymol*. 2010; 481:147–164. [PubMed: 20887857]
- Wang Z, Hryc CF, Bammes B, Afonine PV, et al. An atomic model of brome mosaic virus using direct electron detection and real-space optimization. *Nature Communications*. 2014; 5(4808)

Figure 1.
Membrane protein reconstruction workflow.

Figure 2.

Physical model of vesicle images: (a) Effective contrast-transfer function after image merging; The phase-flipped first exposure (defocus $2\mu\text{m}$) has the CTF shown as a dashed curve. With the mixing of Fourier components from the second exposure ($8\mu\text{m}$) the overall CTF has an additional low-frequency peak (solid curve) and visibility of vesicle membranes and small particles is greatly increased. (b) Estimated inner potential of a lipid bilayer, shown in units of volts relative to the inner potential of water. (c-e) Models of vesicles of 600 and 300 $\text{\AA}$ diameter, obtained by weighting and summing shells of density. (f-h) The corresponding simulated images after application of the merged CTF of Fig. 2a and contrast reversal.

Figure 3.
Sketch of vesicle removal workflow.

Figure 4. Geometrical vesicle normalization, with (a) the original vesicle; (b) the polar coordinate transformation; and (c) the alignment of all $R_{in}(\alpha_i)$ into $R'_{in}(\alpha_i) = R_{0,in}$ and resampling of the vesicle pixels into a normalized grid.

Figure 5.

Mapping and windowing of the vesicle pixels for normalization. (a) The radial mapping from the unnormalized to the normalized polar coordinate plane, after aligning the unnormalized, inner vesicle radius R_{in} to the normalized $R_{0,in}$. We assume that the vesicle wall has the constant absolute thickness W , hence, $R_{out} = R_{in} + W$. (b) Window and weight functions illustrated as 1D normalized radial profiles. (i) The window function $w(\rho)$, (ii) the weight function $\sqrt{\rho}$, arising from the polar coordinate transform; (iii) the combined effect of the window and weight functions. The normalized inner and outer vesicle radius are denoted by $R_{0,in}$ and $R_{0,out}$, respectively.

Figure 6.

Non-linear shape alignment of new vesicles: (Top row) intensity normalized vesicles in the polar coordinates; (middle row) the estimated inner vesicle radii; (bottom row) the non-linearly aligned, windowed and weighted vesicle images in the polar coordinates and the windowed, weighted mean vesicle.

Figure 7.

The selected two-layer HOSVD model $(\tilde{I}_1, \tilde{I}_2, \tilde{K}) = (8, 3, 24)$ illustrated in the Cartesian coordinate frame. (Upper left) The mean vesicle $\mathbf{g}_0$ and (upper second left to lower right) the 24 basis images in the order of importance.

Figure 8. Removal of new vesicles (Fig. 6) with the selected model $(\tilde{I}_1, \tilde{I}_2, K) = (8, 3, 24)$: (Column a,b) intensity normalized vesicle images; (c) projection onto the vesicle subspace; (d-e) the removal result. The images (b) and (d) are slightly blurred by a Gaussian filter to aid visual inspection.

Figure 9.

Comparison of spherical-vesicle model and statistical model for subtraction. (a) Portion of a micrograph, about 10% of the total area, (b) Physical model (PM) for spherical vesicles, and (c) the resulting subtracted micrograph. The yellow box indicates an inside-out particle.

Note the circular residual artifacts at the positions of vesicles. (d) Statistical vesicle models (SM) obtained with 6 Fourier terms in the shape normalization and complexity $(\tilde{I}_1, \tilde{I}_2, \tilde{K}) = (8, 3, 24)$. Note the very similar intensity and shape of these images to those in (b), but with improved modeling of non-circularity and shading. (e) Resulting micrograph. The circular artifacts are much less visible. Micrograph images are filtered with a Gaussian filter with half power at 50 Å.

Figure 10.

Class averages. Two-dimensional CTF-corrected class averages were computed by maximum likelihood from stacks of approximately 8000 particle images. Shown are representatives from 50 class averages. (a) The particle images were extracted directly from micrographs with no vesicle removal; the downward and upward membrane curvatures accompany right-side in and inside-out particles, respectively. (b) Classes from a stack obtained using the physical model vesicle removal as in Fig. 9b,c. (c) Classes are shown from particle images after vesicle removal with the statistical model. The statistical model greatly reduces the membrane-artifact rings, allowing better alignment of particle images. Each image is 270 Å wide.

Figure 11.

Three-dimensional reconstructions from particle images. (a) Crystal structure of the Kv1.2 potassium channel (Long et al. 2007). One of the four alpha subunits is colored blue and one of the four beta subunits is colored pink. (b) Side view (projection direction is parallel to the membrane plane) of the 3D map computed from particle images after vesicle subtraction using the physical model PH. The transmembrane region (top) has reduced intensity due to the subtraction process, but secondary-structure features are visible, especially helix-hairpins in the T1 domains and beta subunits. (c) A similar side view as obtained using the statistical model SM shows somewhat more detail in the transmembrane region. (d) Result of single-particle reconstruction on unsubtracted images. The strong membrane signal interfered with protein particle alignment, such that protein density is displaced toward the membrane and no secondary structure is visible. (e) Membrane-restored reconstruction based on the physical model dataset. In the final Fourier reconstruction step the PH-subtracted images were replaced by unsubtracted images, so that membrane density is restored. (f) The similar membrane-restored reconstruction based on the SM dataset. (g-i) ResMap local resolution, rendered on a slice of the UN, PH and SM models, respectively.