

A Primer to Single-Particle Cryo-Electron Microscopy

Yifan Cheng,¹ Nikolaus Grigorieff,² Pawel A. Penczek,³ and Thomas Walz^{4,*}

¹Department of Biochemistry and Biophysics, University of California, San Francisco, CA 94158, USA

²Janelia Research Campus, 19700 Helix Drive, Ashburn, VA 20147, USA

³Department of Biochemistry and Molecular Biology, The University of Texas–Houston Medical School, 6431 Fannin Street, MSB 6.220, Houston, TX 77030, USA

⁴Department of Cell Biology and Howard Hughes Medical Institute, Harvard Medical School, 240 Longwood Avenue, Boston, MA 02115, USA

*Correspondence: twalz@hms.harvard.edu

<http://dx.doi.org/10.1016/j.cell.2015.03.050>

Cryo-electron microscopy (cryo-EM) of single-particle specimens is used to determine the structure of proteins and macromolecular complexes without the need for crystals. Recent advances in detector technology and software algorithms now allow images of unprecedented quality to be recorded and structures to be determined at near-atomic resolution. However, compared with X-ray crystallography, cryo-EM is a young technique with distinct challenges. This primer explains the different steps and considerations involved in structure determination by single-particle cryo-EM to provide an overview for scientists wishing to understand more about this technique and the interpretation of data obtained with it, as well as a starting guide for new practitioners.

Introduction

Cryo-electron microscopy (cryo-EM) has the ability to provide 3D structural information of biological molecules and assemblies by imaging non-crystalline specimens (single particles). Although the development of the cryo-EM technique began in the 1970s, in the last decade the achievement of near-atomic resolution (<4 Å) has attracted wide attention to the approach.

The remarkable progress in single-particle cryo-EM in the last 2 years has primarily been enabled by the development of direct electron detector device (DDD) cameras (Faruqi and McMullan, 2011; Li et al., 2013a; Milazzo et al., 2011). DDD cameras have a superior detective quantum efficiency (DQE), a measure of the combined effects of the signal and noise performance of an imaging system (McMullan et al., 2009), and the underlying complementary metal-oxide semiconductor (CMOS) technology makes it possible to collect dose-fractionated image stacks, referred to as movies, that allow computational correction of specimen movements (Bai et al., 2013; Campbell et al., 2012; Li et al., 2013a). Together, these features produce images of unprecedented quality, which, in turn, improves the results of digital image processing. In parallel, the continually increasing computer power allows the use of increasingly sophisticated image processing algorithms, resulting in greatly improved and more reliable 3D density maps (see also Cheng, 2015, this issue).

Much effort has been invested in simplifying and automating the collection of EM images and the use of image processing software (reviewed in Lyumkis et al., 2010). The problematic issue with single-particle EM, however, is that there is still no objective quality criterion that is simple and easy to use, such as the R-free value in X-ray crystallography, that would allow one to assess whether the determined density map is accurate or not. Even the resolution of a density map remains subject to controversies. The remaining unresolved issues may not always

be fully appreciated by new practitioners and, if overlooked, can lead to questionable results. A recent example is the 6-Å-resolution structure of the HIV-1 envelope glycoprotein (Mao et al., 2013), which prompted a number of commentaries questioning the validity of the structure (Henderson, 2013; Subramaniam, 2013; van Heel, 2013). This primer seeks to inform about the practical nuts and bolts behind determining a structure by single-particle cryo-EM and to guide new practitioners through the workflow (Figure 1) and important caveats and considerations. Also, as these authors' opinions may not always be shared by everybody in the field, the reader is encouraged to consult other texts on single-particle EM, such as Bai et al. (2015), Frank (2006), Lau and Rubinstein (2013), Milne et al. (2013), and Orlova and Saibil (2011).

Protein Purification for Single-Particle Cryo-EM

Single-particle EM depends on the computational averaging of thousands of images of identical particles. If particles exhibit variable conformation or composition (heterogeneity), more homogeneous subsets can be generated using classification procedures (more below). However, whenever possible, structural heterogeneity should be minimized through biochemical means to simplify structure determination. Biochemical analyses by SDS-PAGE and gel-filtration chromatography are not sufficient to assess whether a sample is suitable for EM analysis, as apparently intact complexes can be a mixture of compositionally different sub-complexes, and even compositionally homogeneous complexes can potentially adopt many different conformations. The most informative way to judge the quality of a protein sample is to visualize it by negative-stain EM. In addition to providing high contrast, the negative staining procedure also tends to induce proteins to adsorb to the carbon film in one or only few preferred orientations, making it easier to assess

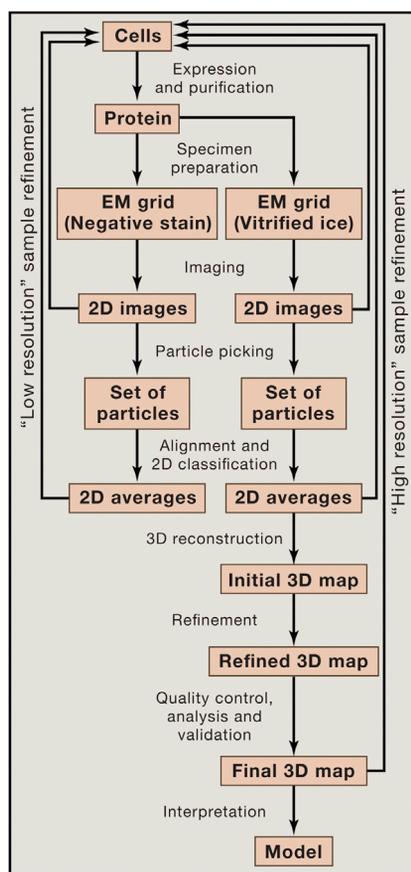


Figure 1. The Steps Involved in Structure Determination by Single-Particle Cryo-EM

A single-particle project should start with a characterization of the specimen in negative stain (left arm of the workflow). Only once the EM images, or potentially 2D class averages, are satisfactory, i.e., the particles are monodisperse and show little aggregation and a manageable degree of heterogeneity (“low-resolution” sample refinement), is the sample ready for analysis by cryo-EM (right arm of the workflow). The images, 2D class averages and 3D maps obtained with vitrified specimens may indicate that the sample requires further improvement to reach near-atomic resolution (“high-resolution” sample refinement).

sample homogeneity (Ohi et al., 2004). The kind of information negative-stain EM provides is described in [Supplemental Information](#).

Structural heterogeneity can be caused by compositional or conformational variability of the target. Compositional heterogeneity, typically the result of sub-stoichiometric components or dissociation of loosely associated subunits, can be addressed in various ways. Ideally, buffer conditions can be found that stabilize the target complex. A promising approach to identify suitable buffer conditions is the Thermofluor-based screening approach (Ericsson et al., 2006). In the case of a sub-stoichiometric subunit, this subunit can be tagged for affinity purification, thus increasing the fraction of complexes containing it in the final preparation. An approach that has proven useful in reducing compositional heterogeneity is mild chemical cross-linking with glutaraldehyde. More control over the cross-linking reaction is obtained with the GraFix technique, in which the sample is

centrifuged into a combined glycerol/glutaraldehyde gradient (Kastner et al., 2008). A variation of this approach is “on column” cross-linking, in which the sample is cross-linked over a size-exclusion column (Shukla et al., 2014). Whichever approach is used, one must keep in mind that cross-linking can introduce artifacts. For example, flexible extensions can become glued together, resulting in a non-physiological structure. Also, if a complex can adopt different conformations, cross-linking can stabilize just one particular state, typically the most compact organization (e.g., Shukla et al., 2014). Hence, native sample always has to be analyzed, too, to understand how cross-linking affects the structure of the target.

Conformational heterogeneity tends to be more difficult to overcome, especially if one or several domains are flexibly tethered to the remainder of a protein. In this case, structural analysis may be restricted to negative-stain EM studies. Alternatively, chemical cross-linking can potentially be used to minimize the conformational heterogeneity, but the physiological relevance of the resulting structures will have to be carefully assessed. Another way to reduce conformational heterogeneity is to lock the target in a defined functional state, which can sometimes be accomplished by adding substrates, inhibitors, ligands, co-factors, or any other molecule affecting the function of the target.

The greatly improved image quality provided by DDD cameras and the availability of ever more sophisticated image-processing software have made structural heterogeneity more manageable. Still, investing time to minimize structural heterogeneity by biochemical tools will always simplify subsequent image processing steps, and it will substantially reduce the risk of obtaining incorrect density maps. Every new project should thus always start with an optimization phase, in which negative-stain EM is used as a tool to optimize protein purification (Figure 1). In rare cases, negative staining will introduce artificial heterogeneity. The only option to exclude this possibility is to look at vitrified specimens by cryo-EM.

Specimen Preparation for Single-Particle Cryo-EM

Before a biological specimen can be imaged, it has to be prepared so it survives the vacuum of the electron microscope, which causes sample dehydration, and the exposure to electrons, which results in radiation damage (the deposition of energy on the specimen by inelastic scattering events that causes breakage of chemical bonds and ultimately structural collapse). The most commonly used preparation techniques, negative staining and vitrification, are briefly discussed in [Supplemental Information](#).

Specimens used for single-particle EM usually consist of purified sample on a carbon film with a support structure. The support structure is most commonly a copper grid, and the carbon film can either be a continuous film, typically used to prepare negatively stained samples, or a holey film, commonly used to prepare vitrified specimens. A problem with EM grids is that thin carbon films are not very stable and are poor conductors at low temperature. This is thought to contribute to the occurrence of beam-induced movement, which can degrade image quality. Therefore, different grid designs have been explored to increase the conductivity of EM grids, such as using doped silicon carbide as the substrate (Cryomesh; Yoshioka et al., 2010),

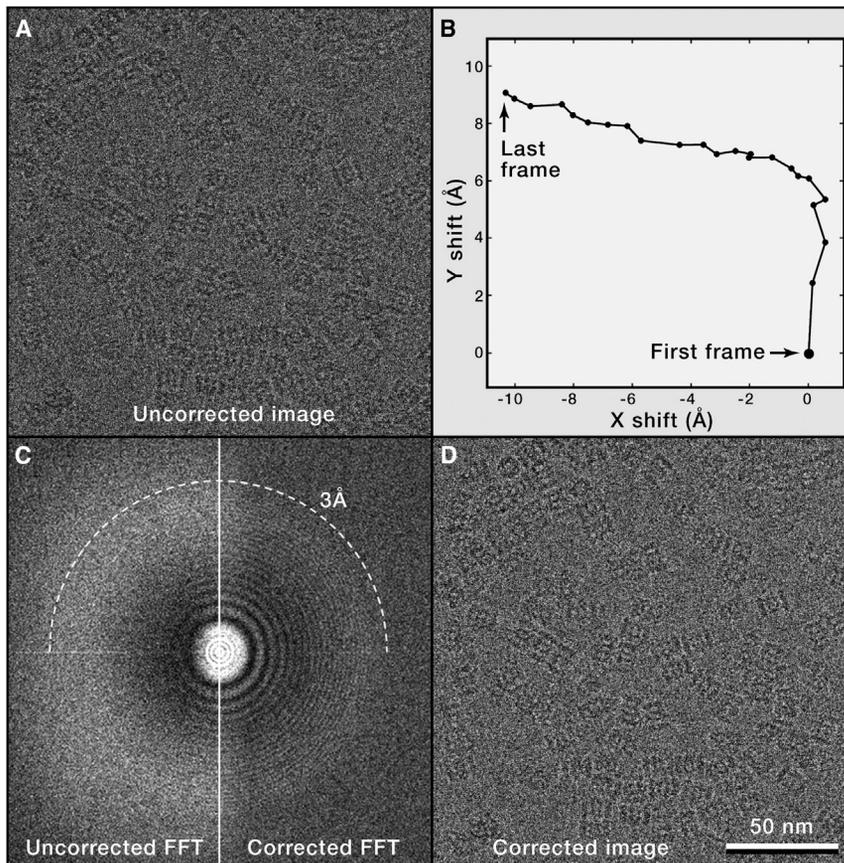


Figure 2. Single-Particle Cryo-EM Images with Motion Correction

Most data recorded with DDD cameras are dose-fractionated image stacks (movies) that can be motion-corrected.

(A) A typical cryo-EM image of vitrified archaeal 20S proteasome particles embedded in a thin layer of vitreous ice. The image is the sum of the raw movie frames without motion correction.

(B) Trace of motion of all movie frames determined using a whole-frame motion-correction algorithm (Li et al., 2013a). Note that the movement between frames is large at the beginning but then slows down.

(C) Left: the power spectrum calculated from the sum of the raw movie frames without motion correction. Right: the power spectrum calculated from the sum of movie frames after motion correction. Motion correction restores Thon rings to close to 3-Å resolution (dashed circle).

(D) Sum of the movie frames that were shifted according to the shifts shown in (B). Note that the images shown in (A) and (D) are indistinguishable by eye, but differ significantly in the quality of the Thon rings seen in their power spectra (C).

and to make them more mechanically stable, such as using gold support (Russo and Passmore, 2014). Before the specimen can be applied, grids have to be rendered hydrophilic, which is typically done with a glow discharger (or, less commonly, with a plasma cleaner).

A perfect vitrified specimen is characterized by an amorphous ice layer of sufficient thickness to accommodate the particles (but ideally not much thicker so that particles are clearly visible), and particles that are well distributed across the field of view and adopt a wide range of orientations. A thin layer of vitrified ice is reasonably transparent and allows particles to be seen clearly (Figures 2 and S1A), while crystalline ice adds a strong texture of dark contrast (bend contours) that usually disguises the embedded particles (Figure S1B).

Semi-automated plungers, such as Vitrobot (FEI) and Cryo-plunge (Gatan), have made it much easier to reproducibly obtain high-quality vitrified specimens. However, care has to be taken to transfer the grids quickly between plunger and cryo-specimen holder and to minimize exposing the liquid nitrogen to air to avoid ice contamination (Figure S1C). An occasional problem is ice that has the appearance of “leopard skin” (Figure S1D). It is unclear what causes this pattern and how it can be avoided, but particles picked from images of such ice areas can still yield reliable 3D maps.

The ice layer should be as thin as possible to achieve high contrast between the molecule and the surrounding ice layer

and to minimize defocus spread due to different heights of the molecules in the ice layer, which can hamper high-resolution structure determination. Importantly, if particles cannot be seen reasonably easily by eye, the sample should not be used for data collection. Parameters that affect ice thickness are described in [Supplemental Information](#).

The ice layer usually tends to be thicker around the edge of a hole and thinner in the center. Large molecules, such as viruses and ribosomes, may thus be excluded from the center of a hole. This effect is stronger with specimens containing detergent, which lowers the surface tension, making it also more challenging to produce thin ice. If thin ice is desired, it helps to use holey carbon grids with smaller holes.

A good vitrified specimen shows a high density of molecules in different orientations. Many particles in a hole reduces the number of images that have to be collected, but ideally the molecules should not touch each other. A problem that is often encountered is that only very few molecules are observed in the holes of the carbon film. A large percentage of molecules is removed during blotting with filter paper, and preparation of vitrified specimens thus requires a much higher sample concentration than preparation of negatively stained specimens. It is not unusual, however, that even with very highly concentrated samples, few particles are seen in the holes. Reasons for this problem can be that the molecules preferentially adsorb to the carbon film, diluting them from the holes, or that they denature as they come into contact with the air/water interface due to the surface tension. An effective solution to deal with the preferential adsorption to the carbon film is to apply the sample twice. The first application will saturate the carbon film with protein, and it is therefore more likely that more particles remain in the holes when the sample is applied a second time. Alternatively, the grid can be

covered by a thin carbon or graphene support film or by a lipid monolayer to which the molecules can adsorb. However, with the exception of graphene, additional support films will reduce image contrast, and all substrates have the potential to induce molecules to adopt preferred orientations. Finally, the grid can be decorated with a self-assembled monolayer to pacify the support film and drive the molecules into the holes (Meyerson et al., 2014). Protein denaturation at the air/water interface can be addressed by using thicker ice (which will, however, reduce image contrast), by using a support film that adsorbs the molecules (but also reduces image contrast), or by chemically fixing the sample before vitrification (which has the potential, however, to affect the structure).

Occasionally particles adopt preferred orientations, presumably due to interactions with the air/water interface. This causes a problem for the reconstruction of a 3D density map, which requires multiple views. One can attempt to overcome this problem by using thicker ice, by adding low amounts of detergent (lowering the surface tension of the air/water interface), or by using a thin support film to which the molecules can adsorb and which will thus keep them away from the air/water interface. One can also try to change the glow-discharge parameters or to modify the protein, e.g., by adding/removing affinity tags. If none of these approaches are successful, it is possible (but technically very challenging) to collect images of tilted specimens, but this usually prevents achieving high resolution.

Image Acquisition for Single-Particle Cryo-EM

Structure determination by single-particle cryo-EM, especially if near-atomic resolution is targeted, requires acquisition of high-quality images, i.e., images with high contrast and with sufficient resolution to answer the biological questions being asked. In addition, particularly for high-resolution projects, high efficiency is beneficial to make them economical, i.e., one should be able to collect a large number of micrographs within a reasonable time-frame. Thus, automation of key steps may be called for. While modern electron microscopes are capable of delivering resolutions better than 2 Å, collection of good-contrast, high-resolution images of vitrified specimens remains challenging. It is therefore critical not only to align the electron microscope with great care but also to choose appropriate imaging conditions. Adjustable settings include, but are not limited to, selection of the condenser aperture and spot size, reduction of imaging aberration by coma-free alignment (all briefly discussed in Supplemental Information), as well as issues related to the optimization of image contrast, such as appropriate defocus settings, selection of objective aperture, and the electron dose used. To learn about contrast enhancement by using phase plates, the reader is referred to Glaeser (2013).

The contrast of vitrified biological specimens is very low, and if images were taken in focus, they would contain little, if any, useful information. Images are therefore taken in bright-field mode of the electron microscope while applying underfocus (Frank, 2006). Given a thin object, images are linear projections of the Coulomb potential of the specimen, the fundamental property necessary for subsequent computational reconstruction of its 3D structure. The images are modulated by the contrast transfer function (CTF), a quasi-periodic sine function in reciprocal

space, the periodicity of which depends, among other parameters, on the defocus setting (Wade, 1992; and Supplemental Information). Furthermore, the amplitudes of the high spatial frequencies (high-resolution detail) in an image are attenuated by an envelope function of the CTF. Its rate of decline depends on the spatial coherence of the electron beam, and it increases with increasing image defocus. Therefore, a higher defocus boosts the low-resolution image contrast but weakens the high-resolution contrast, limiting the frequency range of useful information. Thus, it is best to use the smallest possible defocus that still creates sufficient low-resolution image contrast to clearly see the particles. This means that for large molecules, e.g., viruses, a small underfocus can be used. For small particles (molecular mass less than 200 kDa), however, it is often necessary to underfocus by a few micrometers, which will limit the resolution that can be achieved. Importantly, as the CTF has multiple zero crossings, some information within a single image is lost, which is the reason why images have to be collected at different underfocus settings to sample the entire reciprocal space (Penczek, 2010a; Zhu et al., 1997).

The use of an objective aperture increases amplitude contrast by cutting off electrons scattered at high angles. However, as it also sets a cut-off limit for the resolution, a relatively large objective aperture has to be used for high-resolution single-particle cryo-EM imaging (e.g., 70 μm or 100 μm).

Using a higher electron dose also increases image contrast, but higher electron doses will increase radiation damage. Therefore, for single-exposure images and to achieve high resolution, the electron dose is typically kept below $\sim 20 \text{ e}^-/\text{Å}^2$. Much higher electron doses can be used when movies are recorded (see below). The dose rate also needs to be considered and depends on the type of detector being used for imaging. For imaging on film or when a charge-coupled device (CCD) camera is used, a high dose rate (high beam intensity) is typically used to keep the exposure short (~ 1 s or less), which minimizes the extent of specimen drift during exposure. Short exposures are also preferred when integrating DDD cameras are used to collect single-exposure images, but longer exposures can be used when they are operated in movie mode, which reduces or eliminates the problem of specimen drift (see below). The situation is different for electron-counting DDD cameras. To ensure that electrons are counted properly, the dose rate must be kept below $\sim 10 \text{ e}^-/\text{pixel}/\text{sec}$ (based on current technology) on the camera (Li et al., 2013b; Ruskin et al., 2013). Higher dose rates adversely affect electron counting, thus lowering the DQE and image contrast.

A factor contributing to the recent improvement of attainable resolution in cryo-EM is the movie mode available on some DDD cameras. Here, the total electron dose is fractionated into a series of image frames that can be aligned to compensate for specimen drift and beam-induced movement, thus reducing image blurring (Figure 2) (Brilot et al., 2012; Campbell et al., 2012; Li et al., 2013a). After alignment, the frames are averaged, and the resulting image is used for subsequent structure determination. Movies are made possible by the fast readout and the “rolling shutter” mode of CMOS detectors that underlie all DDD cameras and some newer scintillator-based cameras. Some software packages also allow for sub-frame alignment

to account for local motions that occur during beam exposure (Rubinstein and Brubaker, 2014; Scheres, 2014). Movies also offer the possibility to optimize the overall Signal-to-Noise Ratio (SNR) in images of specimens affected by radiation damage. Early frames correspond to a low electron dose and therefore contain high-resolution signal from the least damaged specimen. However, early frames are also often still affected by fast specimen movement (Figure 2B), blurring the high-resolution information. While specimen movement typically slows down and affects later frames less, these correspond to a higher accumulated dose and increasingly lack high-resolution information. When movie frames are averaged, a relative weighting can be applied that optimizes the signal in the final average (Campbell et al., 2012; Scheres, 2014). As an intermediate measure to improve high-resolution image contrast, one can exclude the initial two or three frames (which are often still affected by high initial specimen movement), as well as the later frames that correspond to a total dose of $\sim 20 \text{ e}^-/\text{\AA}^2$ and higher from the frame averages. However, this strategy results in the loss of low-resolution contrast. Therefore, it may currently be best to use images containing all the movie frames in the alignment step during image processing and to use images without the initial and final frames to calculate the final 3D map (Li et al., 2013a).

The attainable resolution depends on the pixel size on the specimen level, which, in turn, depends on the effective magnification. The physical pixel sizes of digital cameras vary as well as the exact position of the cameras in the optical path. Therefore, the image pixel size has to be calibrated not only for each magnification but also for every microscope/camera combination (a protocol for how to calibrate the magnification is described in Supplemental Information). The Nyquist theorem specifies that the theoretically attainable resolution is limited to twice the pixel size, but interpolation errors introduced by image processing operations and low DQE values of the detector near the Nyquist frequency limit the practically attainable resolution further (Penczek, 2010b). As a rule of thumb, the practical resolution limit is closer to three times the pixel size.

Image Processing

A significant part of the workload of a single-particle project is taken up by the processing of the recorded images. The main steps are discussed here, including correction of the microscope CTF, selection of particles and preparation of image stacks, generation of an initial structure and its refinement, treatment of structural heterogeneity, assessment of resolution, and interpretation of the final 3D density maps. A number of software packages exist that have been developed over the last four decades and are still being improved. While the development of software is important for the success of single-particle cryo-EM, the recent groundbreaking results are primarily due to the use of direct detectors and the recording of movies. Prior to their common use, none of the currently employed algorithms and software packages was capable of yielding results comparable to what is now possible. After direct detectors and movies were adopted, near-atomic resolution was achieved with several software packages, including SPIDER (Frank et al., 1981), EMAN2 (Tang et al., 2007), FREALIGN (Grigorieff, 2007), RELION

(Scheres, 2012), and SPARX (Hohn et al., 2007). To date, EMAN/EMAN2 has been, and continues to be, the most popular software, owing to its extensive options, flexibility, and user friendliness. However, users new to cryo-EM may find it easier to start with more specialized software, such as RELION, which offers streamlined processing with fewer options and one main algorithmic approach (maximum likelihood). This primer is not meant to serve as a manual for any specific image processing software package, but instead tries to relate basic concepts, which may be implemented in different ways in different software packages.

Estimation of CTF Parameters and Correction for Its Effects

The accurate estimation of CTF parameters is important for both the initial evaluation of micrograph quality and subsequent structure determination. To calculate the CTF, the parameters that have to be known are acceleration voltage, spherical aberration, defocus, astigmatism, and percentage of amplitude contrast. Voltage and spherical aberration are instrument parameters that are usually used without further refinement (although the value for the spherical aberration provided by the manufacturer may not be completely accurate). The defocus is set during data collection, but the setting is only approximate. More accurate values for defocus and astigmatism are obtained by fitting a calculated CTF pattern (e.g., Mindell and Grigorieff, 2003) to the Thon rings (semi-circular intensity oscillations induced by the CTF seen in the power spectrum of the image [Thon, 1966]). The contribution of the amplitude contrast is typically assumed as 5%–10% for cryo-EM images.

Once the CTF parameters have been determined and as long as a set of particle views that differ by defocus settings is available, correction for the CTF effects is possible and straightforward (Penczek, 2010a). It can be done for both amplitudes and phases (full CTF correction) or only for the phases (phase flipping). For more detail on CTF estimation and correction, see Supplemental Information.

Ultimately, the determined 3D structure should be corrected for the reciprocal space envelope functions that suppress high-frequency information, and thus visual resolvability of map details. These envelope functions describe effects of microscope optics, limitations of digital scanners and cameras, and errors in orientation parameters assigned to particle images (Jensen, 2001 and section on power spectrum adjustment in Supplemental Information).

Particle Picking

Once a dataset has been collected, movies have been aligned and averaged (if applicable), and good micrographs have been selected (e.g., based on Thon rings being visible to high resolution in all directions), a project continues with the labor-intensive process of particle picking. The quality of the selected particles is a major factor in the subsequent analysis, as inclusion of too many poor particles may preclude successful structure determination. Moreover, methods aimed at cleaning up the selected particles are not very robust, and many artifacts pass all tests and adversely affect subsequent data processing efforts. Particles can be selected in a manual, semi-automated, and fully

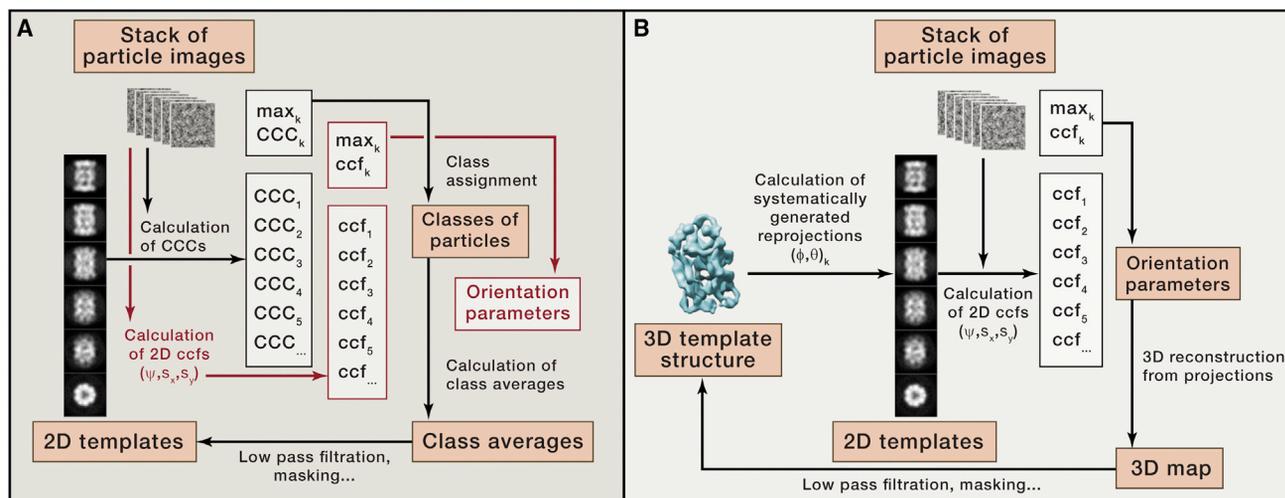


Figure 3. Principle of the K-means Algorithm Used in Single-Particle EM Structure Determination Protocols

(A) In the basic K-means algorithm, the particle images are compared with a set of class averages using a correlation measure that yields the class assignment. Based on the updated class assignments, new class averages are then calculated. Simply by adding 2D alignment of the images to the templates using a correlation function, the algorithm is converted to multi-reference alignment (MRA) (indicated by text in red font).

(B) Principle of the projection matching technique used for 3D single-particle EM structure refinement. The best match of an image to a template yields the Euler angles that were used to generate the template, while a 2D alignment step yields the third, in-plane Euler angle and the two in-plane translations, the total of five orientation parameters required for the 3D reconstruction step.

automated manner. In the early stages of analysis, particularly when little is known about the shape of the protein and the distribution of the projection views, the manual approach is preferable. A trained and careful practitioner can obtain much better results than automated approaches, but the risk is that humans tend to focus on more familiar and better visible particle views, thus omitting less frequently appearing orientations that may be needed for successful structure determination. In semi-automated approaches, the computer performs an initial step of detection of putative particles in a micrograph. All candidates are windowed, and the user removes poor particles from the gallery of possible candidates. Fully automated procedures can be divided into three groups: those that rely on ad hoc steps of denoising and contrast enhancement followed by a search for regions of a given size that emerge above the background level (Adiga et al., 2004); those that extract orientation-independent statistical features from regions of micrographs that may contain particles and proceed with classification (Hall and Patwardhan, 2004; Lata et al., 1995); and those that employ templates, i.e., either class averages of particles selected from micrographs or projections from a known 3D structure of the complex (Chen and Grigorieff, 2007; Huang and Penczek, 2004; Sigworth, 2004). The use of fully automated procedures carries even higher risks of introducing bias, as positively correlating noise features are indistinguishable from weak but valid signal. Therefore, one faces the risk of eventually merely reproducing the template structure. The study of the HIV-1 envelope glycoprotein is a prominent example in which template bias likely played a deciding role (Mao et al., 2013). Good practice is therefore to rely on template-based particle picking only if particles are clearly visible in the micrographs.

With particle coordinates identified in the micrographs, the particles are windowed and assembled into a stack. The initial

locations are not very precise. Therefore, the window size should exceed the approximate particle size by at least 30% (more for small particles). For issues relating to aliasing and density normalization, see Supplemental Information.

2D Clustering and Formation of Class Averages

The first step in single-particle EM structure determination is the analysis of the 2D image dataset, particularly the alignment and grouping of the data into homogenous subsets. There are several reasons for why it is best to begin with 2D analysis: (1) 2D datasets contain image artifacts, invalid particles, or simply empty fields that should be removed; (2) the angular distribution of the particle views is unknown and if the set is dominated by just a few views, 3D analysis is unlikely to succeed; and (3) computational ab initio 3D structure determination requires high-SNR input data, as is present in high-quality class averages.

Various strategies have been proposed to deal with the problem of alignment and clustering of large sets of single-particle EM images (Joyeux and Penczek, 2002; Penczek, 2008), but all are fundamentally rooted in the popular K-means clustering algorithm (Figure 3A). As most steps in single-particle EM analysis use a variant of this algorithm, including 2D multi-reference alignment, 3D multi-reference refinement, even 3D structure refinement (projection matching), the principles and properties of K-means clustering are described in Supplemental Information.

A straightforward implementation of the K-means algorithm in single-particle EM analysis is 2D multi-reference alignment (MRA) (van Heel and Stöffler-Meilicke, 1985), a process in which the dataset is presented with K seed templates, and all images are aligned to and compared with all templates and assigned to the one they most resemble. The process is iterative: a new

set of templates is computed by averaging images based on results from the initial grouping (including transformations given by alignment of the data in the previous step), and the whole procedure is repeated until a stable solution is reached. To accelerate the procedure, one can employ an additional step of principal component analysis (PCA) executed so that the clustering is actually performed using factorial coordinates, not the original images (for in-depth reading, see [Frank, 2006](#)). All major single-particle EM software packages contain a version of MRA, often with various heuristics aimed at improved performance, particularly with respect to the problem of “group collapse”: as MRA combines alignment with clustering, the process is unstable in that the more common particle views produce large, high-SNR class averages, which in turn “attract” less common or more noisy images, eventually leading to the disappearance of less populous groups.

In light of the fundamental shortcomings of MRA (see [Supplemental Information](#)), the iterative stable alignment and clustering (ISAC) method has been developed ([Yang et al., 2012](#)). This method uses a dedicated clustering algorithm to counteract group collapse and employs a multi-level validation strategy of the identified groups, thus yielding uniquely homogeneous classes of images (see [Supplemental Information](#) for more information).

Calculation of Initial Structures

Ab initio 3D structure determination is necessary in cases in which no reasonable templates or guesses for the structure exist. Even though new implementations of 3D structure refinement algorithms are increasingly robust, initial templates, when available, can contain significant errors and an attempt to initialize structure refinement with such templates and raw EM particles is likely to fail. When available, 3D templates can be used, e.g., a low-resolution negative-stain EM 3D reconstruction, an appropriately filtered X-ray model or an EM map of a homolog ([Beckmann et al., 1997](#)). If high point-group symmetry is present, particularly icosahedral symmetry, some refinement algorithms will converge properly with random initialization. However, it is always better to execute all steps indicated in this and the previous sections, because extensive validation methodology built into the 2D analysis and ab initio steps significantly increases confidence in the final outcome.

Ab initio structure determination methods can be broadly divided into those that require additional experiments, typically in the form of tilt pairs, and those that use only data of untilted specimens and rely entirely on computational strategies to deliver the structure.

The earliest and still the most commonly used ab initio tilt-based structure determination method is the random conical tilt (RCT) approach ([Radermacher et al., 1987](#)). Because one of the orientation parameters is set experimentally (tilt angle) and others can be computed in a robust manner (in-plane rotation, tilt angle correction), the method will deliver a reliable initial structure. It is, however, difficult to collect high-tilt data of acceptable quality, especially for vitrified specimens, in which case charging and beam-induced movement can be severe. Most RCT work is thus done using negatively stained specimens, but the artifacts associated with staining ([Cheng et al., 2006](#)) and the missing

cone problem ([Frank, 2006](#)) that further degrades the quality of the 3D map limit the utility of the resulting structures. However, RCT is a virtually foolproof method and its outcome will immensely increase the confidence in the final structure.

Computational ab initio structure determination methods seek to determine five orientation parameters (three Euler angles and two translations) for each projection image such that the resulting 3D structure is “best” in the sense of some mathematical criterion. Due to the low quality of EM data and also due to the time needed for the calculations, virtually all ab initio methods in use assume the input to be a relatively small set (<1,000) of class averages that result from 2D analysis. Since the success of the 3D orientation search strongly depends on the data quality, it is particularly important that the used class averages represent homogeneous particle groups.

The earliest computational ab initio structure determination approach is based on the central section theorem: since Fourier transforms of 2D projections of a 3D object are central sections through the 3D Fourier transform of the object, Fourier transforms of any two projections will intersect along a line, called a “common line.” The common-lines approach was first implemented in IMAGIC as “angular reconstitution,” taking advantage of the existence of a mathematical solution for orienting three projections ([van Heel, 1987](#)). Thus, in angular reconstitution, the user selects triplets of class averages, and multiple triplets are then merged into a common framework, yielding the final 3D structure. The procedure depends critically on user choices and one is thus advised to explore various combinations to gain confidence in the ultimate outcome.

A recently introduced approach to ab initio 3D structure determination, which shows great promise in producing reliable initial models, is based on projection matching using the stochastic hill climbing (SHC) algorithm. The SHC strategy was first implemented in the software package SIMPLE ([Elmlund and Elmlund, 2012](#)), and has been expanded to the “validation of individual parameter reproducibility” (VIPER) approach, which incorporates validation steps into the structure determination process, monitoring the orientation parameters ([Penczek, 2014a](#)). See [Supplemental Information](#) for further information on projection matching, SHC and VIPER.

Structure Refinement and Resolution

After obtaining an initial map, the structure has to be refined to obtain the final map ([Figures 4A–4D](#)). All single-particle EM packages use a more or less elaborate version of the 3D projection matching procedure ([Figure 3B](#) and [Supplemental Information](#)) for structure refinement. It modifies the orientation parameters of single-particle images (projections) to achieve a better match with reprojections computed from the current approximation of the structure ([Penczek, 2008](#)). While all implementations share the same principle of projection matching, the details of the methodology and the degree to which the user can control the process vary widely and are discussed in [Supplemental Information](#).

Progress of the refinement is monitored by a number of indicators, in particular the Fourier shell correlation (FSC) curve ([Figure 4E](#)), which provides information on the level of the SNR as a function of the spatial frequency ([Penczek, 2010c](#)), and

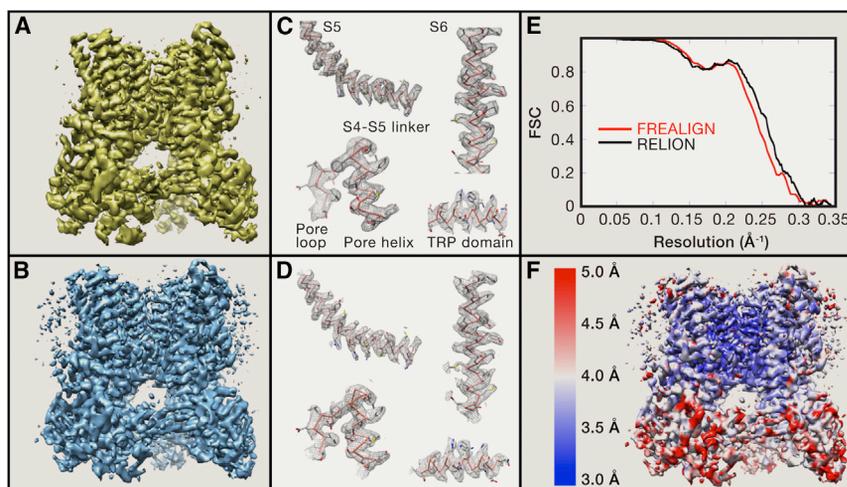


Figure 4. Evaluation and Validation of a 3D EM Structure

Critical evaluation of EM structural results is of utmost importance due to potential model bias and unavoidable noise alignment inherent to the single-particle EM structure determination method. Ultimate confirmation of the map, particularly of the details at the limit of the resolution claimed, is best done by independent structure determination, possibly using different software packages, even if one uses the same dataset. Here, we show the results of two outcomes for the structure determination of the TRPV1 channel.

(A) Originally, the structure was solved using RELION (Scheres, 2012): the refinement was initialized with an RCT model, and the final map represents the best class produced by 3D MRA (Liao et al., 2013).

(B) The structure determination was repeated using the same 2D dataset. 2D MRA was performed using IMAGIC (van Heel et al., 1996), an initial model was generated using EMAN2 (Tang et al., 2007), and refinement and 3D MRA were done in FREALIGN (Grigorieff, 2007; Lyumkis et al., 2013).

For consistency, the rotationally averaged power spectrum of map (B) was set to that of map (A). Interestingly, while the two maps are visually very similar, only ~60% of particles in the best class determined by RELION coincide with those in the best class determined by FREALIGN. This difference likely reflects limitations of *K*-means-based clustering approaches and, possibly, points to the fact that the number of classes used was too low.

(C and D) The side-chain densities in the best parts of the map shown in (A) agree with those of the map shown in (B), validating these details. However, some weaker peripheral density features in the maps shown in (A) and (B) exhibit noticeable differences (see Supplemental Information and Figure S2).

(E) Angular uncertainty and blurring affects the FSC curves, and thus the resolution reported: calculation of 3D reconstructions using multiple, probability-weighted copies of each particle image (“soft matching,” see Supplemental Information) can lead to an apparent improvement in the resolution (RELION, black curve) while hard matching yields more conservative results (FREALIGN, red curve). The difference is, however, too small to affect the interpretation of the maps and also lies within the general uncertainty bounds of the FSC methodology, which also depends on other data-processing steps, as, for example, masking of the map.

(F) The resolution of the map is non-uniform. The local resolution of the map shown in (B) was calculated (Penczek, 2014c) and indicates that densities within the membrane domain, and particularly around the pore, are better resolved than those in the extracellular domains. 3D maps were rendered using UCSF Chimera (Pettersen et al., 2004).

the resolution of the map. The FSC curve is obtained by splitting the dataset into halves, calculating a volume from each half, and computing correlation coefficients within resolution shells extracted from Fourier transforms of the two volumes. Importantly, the definition requires that the noise in the two structures should be independent, a condition difficult to meet in practice and often compromised by refining a single dataset while evaluating the FSC with two structures computed from half-subsets of the entire set. “Resolution” in single-particle EM is then a somewhat arbitrarily chosen cut-off level of the SNR or FSC curve. For example, the resolution can be defined as the spatial frequency at which the FSC curve is 0.5 or as the spatial frequency at which the SNR is 1.0 (corresponding to an FSC of 0.33), the level at which the power of the signal is equal to the power of the noise. Another common choice of threshold is 0.143, the value selected based on relating EM results to those in X-ray crystallography (Rosenthal and Henderson, 2003).

A common problem in structure refinement is so-called “over-fitting” of the data—the emergence of features in an EM map due to the alignment of noise. Over-fitting arises due to the fact that the dataset is refined without reference to external standards (at least before the emergence of secondary-structure features whose generic appearance is known), and, therefore, it is not known what constitutes “signal” and what is “noise” (Stewart and Grigorieff, 2004). As a result, artifacts are created by chance and further enhanced by alignment of the noise components in the data, leading to inflated FSC values and an artificially high resolution. It was realized early on that in order to ensure inde-

pendence of noise in the half-dataset maps used to calculate the FSC, the half-datasets must be refined independently (e.g., Grigorieff, 2000). This avoids exaggerated resolution estimates using the FSC, an approach that has recently been reiterated (Scheres, 2012) and is now often referred to as the “gold standard” refinement procedure (Henderson et al., 2012). It has to be noted, however, that even this procedure has limitations, as (1) it is impossible to have true independence between the half datasets; (2) the approach tends to underestimate the resolution potential of the data; and (3) for all existing refinement algorithms, each of the half-structures suffers independently from the described “over-fitting” problem. There are also a number of image-processing steps that result in a nominal improvement of the resolution without actually improving the image alignment parameters (Figure 4E) or map. An obvious example is masking of the structure, as the shape of the mask and the way its edges are attenuated may have a significant impact on the FSC curve. One can also set density values to a constant when they are lower than a certain level, a step that is akin to solvent flattening in X-ray crystallography. Since none of these operations are codified in the field and since the FSC curve is also dependent on other factors beyond the ones mentioned here, it is the opinion of the authors that there is currently no real “gold standard” procedure for structure refinement and resolution estimation of an EM map. An approach equally useful to the “gold standard” procedure to obtain an adequate resolution estimate is simply to limit the refinement frequency to a resolution lower than the one of the reference map.

In conclusion, the quality of an EM map is described by the entire FSC curve, not just the resolution, and there are EM maps with the same nominal resolution that differ significantly in overall quality (Ludtke and Serysheva, 2013). The reverse is also true, namely that the reported nominal resolution reflects the overall resolution of the entire density map but it does not account for local variation. The EM map with the highest nominal resolution is not necessarily the best one, because values at lower frequencies often matter more for connectivity and interpretability of the map. Hence, the resolution reported for an EM map should be treated as a broad guideline rather than a firm number.

3D Multi-Reference Alignment

Many samples will contain structural heterogeneity. When its presence is detected (for example by calculating a variance map, see below), a possibility is to use 3D multi-reference alignment (3D MRA) to extract more homogeneous subsets. Current implementations are natural extensions of the basic projection-matching procedure and employ principles of the *K*-means algorithm: the user has to provide a number of initial 3D templates and the program aligns each single-particle image to all 3D templates to find the best-matching one. When all images are assigned, new 3D reconstructions are calculated and used as new references. The method proved to be successful in many applications (Brink et al., 2004; Heymann et al., 2003; Loerke et al., 2010; Schüler et al., 2006), particularly when “focusing” on a variable sub-region to make the assignments (Penczek et al., 2006). The shortcomings of 3D MRA are those of the *K*-means algorithm: a strong bias toward initial templates, solutions that depend on the number *K* of requested classes, and a lack of validation of the results. In light of these limitations, the applicability of 3D MRA should be guided by the concurrent examination of the local variability of the map (Penczek 2014c). Indeed, if the procedure succeeded in separating the dataset into homogeneous classes, the distribution of 3D variability within each group should be uniform (in practice it tends to be proportional to the density distribution of the map). Any residual local variability that exceeds what is reasonably expected, particularly at locations where map density is low, signals that 3D MRA should be continued with an increased number of classes and possibly in the “focused” mode. The 3D MRA procedure works best for complexes exhibiting substoichiometric ligand binding in which a fragmented appearance of the ligand would indicate failure of the procedure, and results can be validated by the appearance of secondary structure elements in the 3D class averages.

Structure Validation and Interpretation

As explained above, the indication of a certain resolution by the FSC alone does not demonstrate the validity of the refined structure. Independent refinement of two exclusive subsets of the data increases confidence in the resolution but does not necessarily confirm the validity of a structure. This is particularly true for reconstructions that do not resolve secondary structure features. Because refinement is typically initialized with the same 3D template, even if low-pass filtered, this undermines the independence assumption. Furthermore, the FSC may fail entirely to indicate resolution when there is significant misalignment of the particle images. All current refinement software may display

this behavior of the FSC, including software that performs separate refinement of subsets of the data. It is therefore equally important to also apply other plausibility criteria to the results whenever possible (see below).

In case of a heterogeneous dataset, the refinement itself might be correct, but the structure, being a superposition of various states, will have limited biological relevance. Therefore, additional tests are recommended, particularly those that reveal the localized real-space quality of the map. First, it is possible to compute the local resolution of the map using a wavelet-based (Kucukelbir et al., 2014) or an FSC-equivalent approach (Penczek, 2014b) (Figure 4F). Local real-space variability of the map can be assessed using a simplified variance approach (Penczek, 2014c). More information could be obtained from analysis of correlations within the map, as in 3D PCA, by statistical resampling (Penczek et al., 2011), which is computationally demanding and yields only low-resolution information. A local variability analysis can also serve as a means to establish plausible initial templates for 3D MRA (Spahn and Penczek, 2009).

The overall validation of an EM map depends on the resolution reached. We can distinguish three resolution regimes that may help confirm the resolution indicated by the FSC. A *low-resolution* map (>10 Å) reveals the overall shape of a complex and possibly the relative arrangement of major modules. Here, docking of X-ray segments is unreliable, and flexible fitting should be avoided. An *intermediate-resolution* map (4–10 Å) reveals secondary structure details and the relative arrangement of modules. It enables unique fitting of X-ray segments and can be used to detect conformational changes. A *high-resolution* map (<4 Å) clearly resolves secondary structure elements (e.g., α helices) and some individual residues, allowing polypeptide backbone tracing (Figures 4C and 4D) and precise fitting of X-ray segments. It also provides a detailed description of conformational changes. Keeping in mind that the precise resolution number attached to the map can often not be reliably established, one should focus on arguments that give confidence that the overall appearance of the map is correct. Thus, for low-resolution maps, the best evidence is provided by tilt experiments, particularly by initiating the project by RCT reconstruction. While final details of the map might be debatable, at least the possibility of major mistakes is minimized. A map at intermediate resolution can be confirmed if the appearance of subunits agrees with the appearance of segments determined by X-ray crystallography, if available. A measure of confidence can also be provided by a posteriori tilt experiments (Henderson et al., 2011). In these experiments, often referred to as “tilt test,” a small set of image pairs is collected, one untilted and a second with a small sample tilt, for example 10 degree. The test requires projection matching of the particles from the tilt pairs to the EM map that is to be validated. If the difference between the views determined for the tilt pairs corresponds (within error) to the known tilt angle, the EM map is considered valid. High-resolution maps must display known features of secondary structure elements and density for bulky side chains. These features can be further corroborated with a plausible atomic model that can also be used to obtain an independent resolution estimate by converting it to pseudo-electron density and low-pass filtering it to the claimed resolution of the EM map.

The interpretation of EM maps depends mainly on three factors: the resolution of the map, the established presence of multiple conformational states in the sample, and the availability of X-ray crystallographic segments of some components, of the entire complex or of one of its homologs. High-resolution EM maps can be analyzed in the same manner as X-ray maps by performing de novo backbone tracing. Furthermore, because EM experiments yield both amplitudes and phases, it is possible to arrive at reliable atomic models even in cases in which crystallographic efforts were unsuccessful or comparison with an atomic model is difficult. In addition, the availability of local resolution and variability measures is helpful in avoiding over-interpretation of poorly resolved regions of EM maps. At high resolution, docking of X-ray segments can be done with high precision, thus increasing apparent resolvability of the results and making it possible to detect atomic scale conformational changes with respect to the X-ray results. Similarly, availability of high-resolution structures of multiple functional states of the complex makes single-particle EM a unique tool to study protein dynamics.

Intermediate-resolution EM maps offer insights into the arrangement of subunits and localization of functional sites of macromolecular complexes. Structure interpretation can be augmented by docking of X-ray segments, if available, which also improves the precision of feature localization. The docking can be accomplished, for example, using UCSF Chimera (Pettersen et al., 2004). However, as the resolution of EM maps gets worse, so does the precision of docking. While some progress has been made in this area, reliable computational tools to assess docking uncertainty as a function of map quality are lacking, so some caution is needed to avoid over-interpretation of the results.

The main utility of low-resolution EM maps is in revealing the overall architecture of a complex. Results of docking X-ray segments should be interpreted with utmost caution, because determining the best-fitting position of a given segment does not mean that it is its only possible localization, creating the possibility of major mistakes. At the same time, low-resolution EM maps have the added value that they can often provide a stepping stone toward higher resolution, and thus more informative results.

Concluding Remarks

Structure determination by single-particle cryo-EM is an increasingly popular approach, but like most experimental methodologies, it is important not to approach it with “plug and play” assumptions. We hope that the information provided in this Primer will be helpful in guiding the execution of this technique and the interpretation of data obtained with it.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Text and two figures and can be found with this article online at <http://dx.doi.org/10.1016/j.cell.2015.03.050>.

AUTHOR CONTRIBUTIONS

Authors are listed alphabetically and contributed equally to the manuscript. They can be contacted directly: ycheng@ucsf.edu (Y.C.), niko@grigorieff.com

(N.G.), pawel.a.penczek@uth.tmc.edu (P.A.P.), twalz@hms.harvard.edu (T.W.).

ACKNOWLEDGMENTS

The authors thank Rita De Zorzi for providing images of the AMPA receptor and Tim Grant for calculating 2D class averages and an initial reconstruction of the TRPV1 channel using IMAGIC and EMAN2. This work was supported by grants from the NIH (R01 GM098672, R01 GM082893, and P50 GM082250 to Y.C.; R01 GM60635 to P.A.P.). N.G. and T.W. are investigators with the Howard Hughes Medical Institute.

REFERENCES

- Adiga, P.S., Malladi, R., Baxter, W., and Glaeser, R.M. (2004). A binary segmentation approach for boxing ribosome particles in cryo EM micrographs. *J. Struct. Biol.* *145*, 142–151.
- Bai, X.C., Fernandez, I.S., McMullan, G., and Scheres, S.H. (2013). Ribosome structures to near-atomic resolution from thirty thousand cryo-EM particles. *eLife* *2*, e00461.
- Bai, X.C., McMullan, G., and Scheres, S.H. (2015). How cryo-EM is revolutionizing structural biology. *Trends Biochem. Sci.* *40*, 49–57.
- Beckmann, R., Bubeck, D., Grassucci, R., Penczek, P., Verschoor, A., Blobel, G., and Frank, J. (1997). Alignment of conduits for the nascent polypeptide chain in the ribosome-Sec61 complex. *Science* *278*, 2123–2126.
- Brilot, A.F., Chen, J.Z., Cheng, A., Pan, J., Harrison, S.C., Potter, C.S., Carragher, B., Henderson, R., and Grigorieff, N. (2012). Beam-induced motion of vitrified specimen on holey carbon film. *J. Struct. Biol.* *177*, 630–637.
- Brink, J., Ludtke, S.J., Kong, Y., Wakil, S.J., Ma, J., and Chiu, W. (2004). Experimental verification of conformational variation of human fatty acid synthase as predicted by normal mode analysis. *Structure* *12*, 185–191.
- Campbell, M.G., Cheng, A., Brilot, A.F., Moeller, A., Lyumkis, D., Velesler, D., Pan, J., Harrison, S.C., Potter, C.S., Carragher, B., and Grigorieff, N. (2012). Movies of ice-embedded particles enhance resolution in electron cryo-microscopy. *Structure* *20*, 1823–1828.
- Chen, J.Z., and Grigorieff, N. (2007). SIGNATURE: a single-particle selection system for molecular electron microscopy. *J. Struct. Biol.* *157*, 168–173.
- Cheng, Y. (2015). Single-particle cryo-EM at crystallographic resolution. *Cell* *161*, this issue, 450–457.
- Cheng, Y., Wolf, E., Larvie, M., Zak, O., Aisen, P., Grigorieff, N., Harrison, S.C., and Walz, T. (2006). Single particle reconstructions of the transferrin-transferrin receptor complex obtained with different specimen preparation techniques. *J. Mol. Biol.* *355*, 1048–1065.
- Elmlund, D., and Elmlund, H. (2012). SIMPLE: Software for *ab initio* reconstruction of heterogeneous single-particles. *J. Struct. Biol.* *180*, 420–427.
- Ericsson, U.B., Hallberg, B.M., Detitta, G.T., Dekker, N., and Nordlund, P. (2006). Thermofluor-based high-throughput stability optimization of proteins for structural studies. *Anal. Biochem.* *357*, 289–298.
- Faruqi, A.R., and McMullan, G. (2011). Electronic detectors for electron microscopy. *Q. Rev. Biophys.* *44*, 357–390.
- Frank, J. (2006). Three-dimensional electron microscopy of macromolecular assemblies: visualization of biological molecules in their native state (New York: Oxford University Press).
- Frank, J., Shimkin, B., and Dowse, H. (1981). SPIDER—a modular software system for electron image processing. *Ultramicroscopy* *6*, 343–358.
- Glaeser, R.M. (2013). Invited review article: Methods for imaging weak-phase objects in electron microscopy. *Rev. Sci. Instrum.* *84*, 111101.
- Grigorieff, N. (2000). Resolution measurement in structures derived from single particles. *Acta Crystallogr. D Biol. Crystallogr.* *56*, 1270–1277.
- Grigorieff, N. (2007). FREALIGN: high-resolution refinement of single particle structures. *J. Struct. Biol.* *157*, 117–125.

- Hall, R.J., and Patwardhan, A. (2004). A two step approach for semi-automated particle selection from low contrast cryo-electron micrographs. *J. Struct. Biol.* *145*, 19–28.
- Henderson, R. (2013). Avoiding the pitfalls of single particle cryo-electron microscopy: Einstein from noise. *Proc. Natl. Acad. Sci. USA* *110*, 18037–18041.
- Henderson, R., Chen, S., Chen, J.Z., Grigorieff, N., Passmore, L.A., Ciccarelli, L., Rubinstein, J.L., Crowther, R.A., Stewart, P.L., and Rosenthal, P.B. (2011). Tilt-pair analysis of images from a range of different specimens in single-particle electron cryomicroscopy. *J. Mol. Biol.* *413*, 1028–1046.
- Henderson, R., Sali, A., Baker, M.L., Carragher, B., Devkota, B., Downing, K.H., Egelman, E.H., Feng, Z., Frank, J., Grigorieff, N., et al. (2012). Outcome of the first electron microscopy validation task force meeting. *Structure* *20*, 205–214.
- Heymann, J.B., Cheng, N., Newcomb, W.W., Trus, B.L., Brown, J.C., and Steven, A.C. (2003). Dynamics of herpes simplex virus capsid maturation visualized by time-lapse cryo-electron microscopy. *Nat. Struct. Biol.* *10*, 334–341.
- Hohn, M., Tang, G., Goodyear, G., Baldwin, P.R., Huang, Z., Penczek, P.A., Yang, C., Glaeser, R.M., Adams, P.D., and Ludtke, S.J. (2007). SPARX, a new environment for Cryo-EM image processing. *J. Struct. Biol.* *157*, 47–55.
- Huang, Z., and Penczek, P.A. (2004). Application of template matching technique to particle detection in electron micrographs. *J. Struct. Biol.* *145*, 29–40.
- Jensen, G.J. (2001). Alignment error envelopes for single particle analysis. *J. Struct. Biol.* *133*, 143–155.
- Joyeux, L., and Penczek, P.A. (2002). Efficiency of 2D alignment methods. *Ultramicroscopy* *92*, 33–46.
- Kastner, B., Fischer, N., Golas, M.M., Sander, B., Dube, P., Boehringer, D., Hartmuth, K., Deckert, J., Hauer, F., Wolf, E., et al. (2008). GraFix: sample preparation for single-particle electron cryomicroscopy. *Nat. Methods* *5*, 53–55.
- Kucukelbir, A., Sigworth, F.J., and Tagare, H.D. (2014). Quantifying the local resolution of cryo-EM density maps. *Nat. Methods* *11*, 63–65.
- Lata, K.R., Penczek, P., and Frank, J. (1995). Automatic particle picking from electron micrographs. *Ultramicroscopy* *58*, 381–391.
- Lau, W.C., and Rubinstein, J.L. (2013). Single particle electron microscopy. *Methods Mol. Biol.* *955*, 401–426.
- Li, X., Mooney, P., Zheng, S., Booth, C.R., Braunfeld, M.B., Gubbens, S., Agard, D.A., and Cheng, Y. (2013a). Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. *Nat. Methods* *10*, 584–590.
- Li, X., Zheng, S.Q., Egami, K., Agard, D.A., and Cheng, Y. (2013b). Influence of electron dose rate on electron counting images recorded with the K2 camera. *J. Struct. Biol.* *184*, 251–260.
- Liao, M., Cao, E., Julius, D., and Cheng, Y. (2013). Structure of the TRPV1 ion channel determined by electron cryo-microscopy. *Nature* *504*, 107–112.
- Loerke, J., Giesebrecht, J., and Spahn, C.M. (2010). Multiparticle cryo-EM of ribosomes. *Methods Enzymol.* *483*, 161–177.
- Ludtke, S.J., and Serysheva, I.I. (2013). Single-particle cryo-EM of calcium release channels: structural validation. *Curr. Opin. Struct. Biol.* *23*, 755–762.
- Lyumkis, D., Moeller, A., Cheng, A., Herold, A., Hou, E., Irving, C., Jacovetty, E.L., Lau, P.W., Mulder, A.M., Pulokas, J., et al. (2010). Automation in single-particle electron microscopy connecting the pieces. *Methods Enzymol.* *483*, 291–338.
- Lyumkis, D., Brilot, A.F., Theobald, D.L., and Grigorieff, N. (2013). Likelihood-based classification of cryo-EM images using FREALIGN. *J. Struct. Biol.* *183*, 377–388.
- Mao, Y., Wang, L., Gu, C., Herschhorn, A., Désormeaux, A., Finzi, A., Xiang, S.H., and Sodroski, J.G. (2013). Molecular architecture of the uncleaved HIV-1 envelope glycoprotein trimer. *Proc. Natl. Acad. Sci. USA* *110*, 12438–12443.
- McMullan, G., Chen, S., Henderson, R., and Faruqi, A.R. (2009). Detective quantum efficiency of electron area detectors in electron microscopy. *Ultramicroscopy* *109*, 1126–1143.
- Meyerson, J.R., Rao, P., Kumar, J., Chittori, S., Banerjee, S., Pierson, J., Mayer, M.L., and Subramaniam, S. (2014). Self-assembled monolayers improve protein distribution on holey carbon cryo-EM supports. *Sci. Rep.* *4*, 7084.
- Milazzo, A.C., Cheng, A., Moeller, A., Lyumkis, D., Jacovetty, E., Polukas, J., Ellisman, M.H., Xuong, N.H., Carragher, B., and Potter, C.S. (2011). Initial evaluation of a direct detection device detector for single particle cryo-electron microscopy. *J. Struct. Biol.* *176*, 404–408.
- Milne, J.L., Borgnia, M.J., Bartesaghi, A., Tran, E.E., Earl, L.A., Schauder, D.M., Lengyel, J., Pierson, J., Patwardhan, A., and Subramaniam, S. (2013). Cryo-electron microscopy—a primer for the non-microscopist. *FEBS J.* *280*, 28–45.
- Mindell, J.A., and Grigorieff, N. (2003). Accurate determination of local defocus and specimen tilt in electron microscopy. *J. Struct. Biol.* *142*, 334–347.
- Ohi, M., Li, Y., Cheng, Y., and Walz, T. (2004). Negative staining and image classification – powerful tools in modern electron microscopy. *Biol. Proced. Online* *6*, 23–34.
- Orlova, E.V., and Saibil, H.R. (2011). Structural analysis of macromolecular assemblies by electron microscopy. *Chem. Rev.* *111*, 7710–7748.
- Penczek, P.A. (2008). Single Particle Reconstruction. In *International Tables for Crystallography*, U. Shmueli, ed. (New York: Springer), pp. 375–388.
- Penczek, P.A. (2010a). Image restoration in cryo-electron microscopy. *Methods Enzymol.* *482*, 35–72.
- Penczek, P.A. (2010b). Fundamentals of three-dimensional reconstruction from projections. *Methods Enzymol.* *482*, 1–33.
- Penczek, P.A. (2010c). Resolution measures in molecular electron microscopy. *Methods Enzymol.* *482*, 73–100.
- Penczek, P.A. (2014a). sxvipr. In SPARX Wiki (<http://sparx-em.org/sparxwiki/sxvipr>).
- Penczek, P.A. (2014b). sxlocres. In SPARX Wiki (<http://sparx-em.org/sparxwiki/sxlocres>).
- Penczek, P.A. (2014c). sx3dvariability. In SPARX Wiki (<http://sparx-em.org/sparxwiki/sx3dvariability>).
- Penczek, P.A., Frank, J., and Spahn, C.M.T. (2006). A method of focused classification, based on the bootstrap 3D variance analysis, and its application to EF-G-dependent translocation. *J. Struct. Biol.* *154*, 184–194.
- Penczek, P.A., Kimmel, M., and Spahn, C.M. (2011). Identifying conformational states of macromolecules by eigen-analysis of resampled cryo-EM images. *Structure* *19*, 1582–1590.
- Petersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* *25*, 1605–1612.
- Radermacher, M., Wagenknecht, T., Verschoor, A., and Frank, J. (1987). Three-dimensional reconstruction from a single-exposure, random conical tilt series applied to the 50S ribosomal subunit of *Escherichia coli*. *J. Microsc.* *146*, 113–136.
- Rosenthal, P.B., and Henderson, R. (2003). Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *J. Mol. Biol.* *333*, 721–745.
- Rubinstein, J.L., and Brubaker, M.A. (2014). Alignment of cryo-EM movies of individual particles by global optimization of image translations. *arXiv arXiv:1409.6789*, <http://arxiv.org/abs/1409.6789>.
- Ruskin, R.S., Yu, Z., and Grigorieff, N. (2013). Quantitative characterization of electron detectors for transmission electron microscopy. *J. Struct. Biol.* *184*, 385–393.
- Russo, C.J., and Passmore, L.A. (2014). Electron microscopy: Ultrastable gold substrates for electron cryomicroscopy. *Science* *346*, 1377–1380.
- Scheres, S.H.W. (2012). RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.* *180*, 519–530.
- Scheres, S.H.W. (2014). Beam-induced motion correction for sub-megadalton cryo-EM particles. *eLife* *3*, e03665.

- Schüler, M., Connell, S.R., Lescoute, A., Giesebrecht, J., Dabrowski, M., Schroeer, B., Mielke, T., Penczek, P.A., Westhof, E., and Spahn, C.M.T. (2006). Structure of the ribosome-bound cricket paralysis virus IRES RNA. *Nat. Struct. Mol. Biol.* *13*, 1092–1096.
- Shukla, A.K., Westfield, G.H., Xiao, K., Reis, R.I., Huang, L.Y., Tripathi-Shukla, P., Qian, J., Li, S., Blanc, A., Oleskie, A.N., et al. (2014). Visualization of arrestin recruitment by a G-protein-coupled receptor. *Nature* *512*, 218–222.
- Sigworth, F.J. (2004). Classical detection theory and the cryo-EM particle selection problem. *J. Struct. Biol.* *145*, 111–122.
- Spahn, C.M., and Penczek, P.A. (2009). Exploring conformational modes of macromolecular assemblies by multiparticle cryo-EM. *Curr. Opin. Struct. Biol.* *19*, 623–631.
- Stewart, A., and Grigorieff, N. (2004). Noise bias in the refinement of structures derived from single particles. *Ultramicroscopy* *102*, 67–84.
- Subramaniam, S. (2013). Structure of trimeric HIV-1 envelope glycoproteins. *Proc. Natl. Acad. Sci. USA* *110*, E4172–E4174.
- Tang, G., Peng, L., Baldwin, P.R., Mann, D.S., Jiang, W., Rees, I., and Ludtke, S.J. (2007). EMAN2: an extensible image processing suite for electron microscopy. *J. Struct. Biol.* *157*, 38–46.
- Thon, F. (1966). Zur Defokussierungsabhängigkeit des Phasenkontrastes bei der elektronenmikroskopischen Abbildung. *Z Naturforschung* *21a*, 476–478.
- van Heel, M. (1987). Angular reconstitution: *a posteriori* assignment of projection directions for 3D reconstruction. *Ultramicroscopy* *21*, 111–123.
- van Heel, M. (2013). Finding trimeric HIV-1 envelope glycoproteins in random noise. *Proc. Natl. Acad. Sci. USA* *110*, E4175–E4177.
- van Heel, M., and Stöffler-Meilicke, M. (1985). Characteristic views of *E. coli* and *B. stearothermophilus* 30S ribosomal subunits in the electron microscope. *EMBO J.* *4*, 2389–2395.
- van Heel, M., Harauz, G., Orlova, E.V., Schmidt, R., and Schatz, M. (1996). A new generation of the IMAGIC image processing system. *J. Struct. Biol.* *116*, 17–24.
- Wade, R.H. (1992). A brief look at imaging and contrast transfer. *Ultramicroscopy* *46*, 145–156.
- Yang, Z., Fang, J., Chittuluru, J., Asturias, F.J., and Penczek, P.A. (2012). Iterative stable alignment and clustering of 2D transmission electron microscope images. *Structure* *20*, 237–247.
- Yoshioka, C., Carragher, B., and Potter, C.S. (2010). Cryomesh: a new substrate for cryo-electron microscopy. *Microsc. Microanal.* *16*, 43–53.
- Zhu, J., Penczek, P.A., Schröder, R., and Frank, J. (1997). Three-dimensional reconstruction with contrast transfer function correction from energy-filtered cryoelectron micrographs: procedure and application to the 70S *Escherichia coli* ribosome. *J. Struct. Biol.* *118*, 197–219.

Supplemental Information

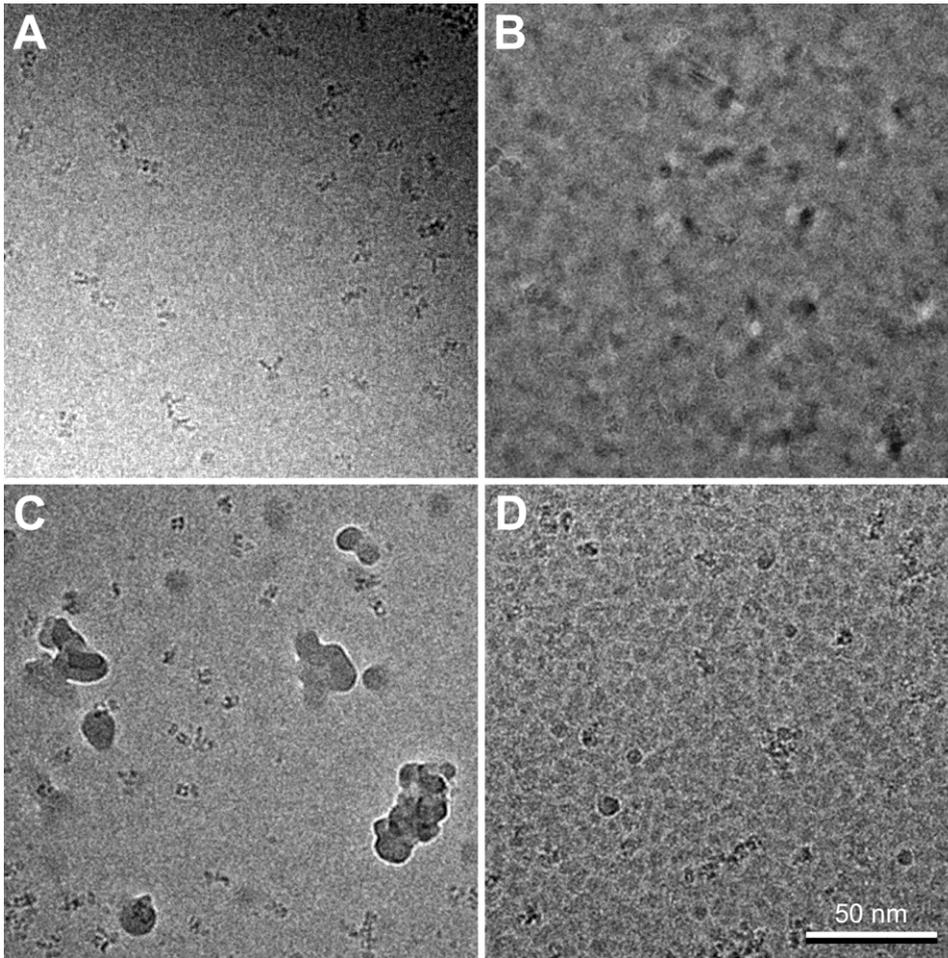


Figure S1. Cryo-EM Images of AMPA Receptors

(A) A good image of AMPA receptor in a hole covered with vitrified ice. The receptors are very distinct and clearly adopt many different orientations. Ideally, the ice layer would be more homogeneous in thickness, and the image would show more particles.

(B) An image of a hole that is covered with crystalline ice. Even if particles can occasionally be seen, particles from such ice areas cannot be used.

(C) An image of a hole that has ice contamination. There is no problem when particles can be selected from uncontaminated areas.

(D) An image of a hole covered with ice that shows the “leopard skin” pattern. Although not ideal, such images can be used to select particles.

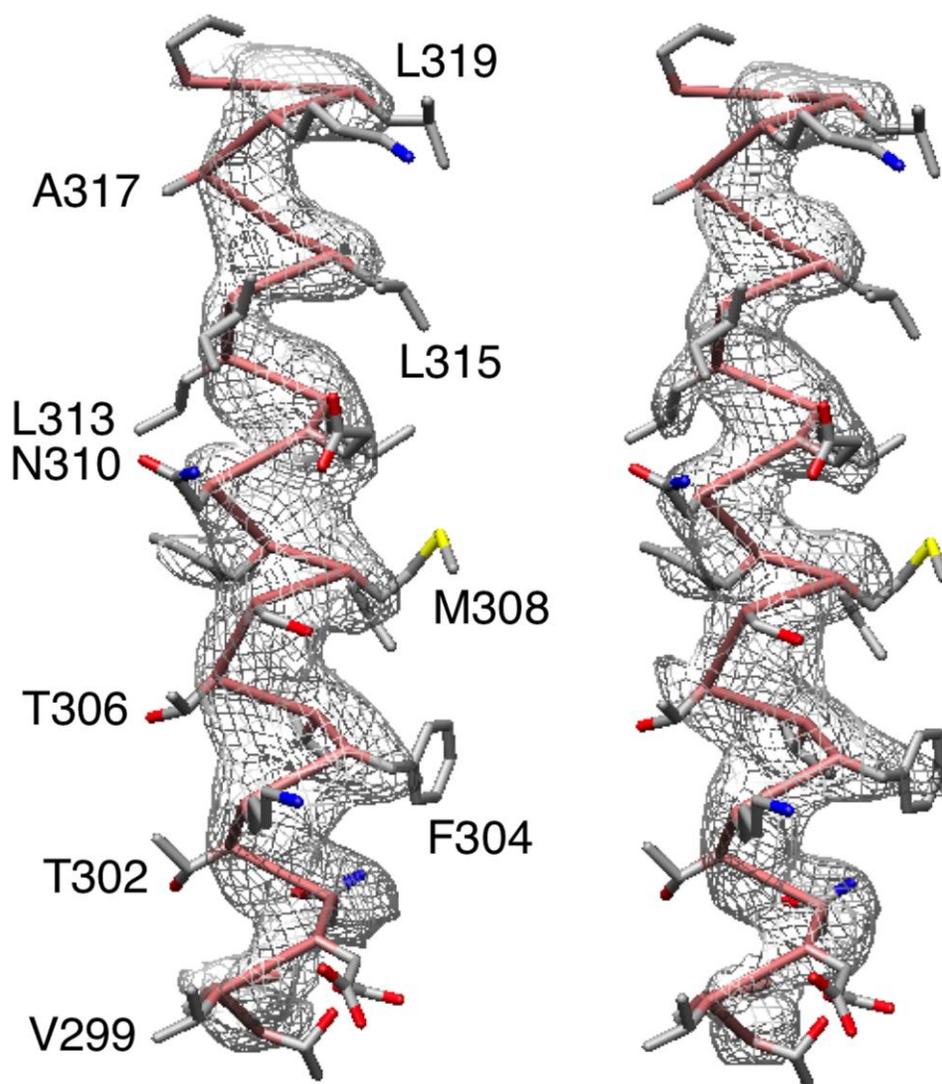


Figure S2. Effects of Soft Matching on Peripheral Density Features

Refinement and reconstruction of the TRPV1 channel structure (Liao et al., 2013) using RELION (Scheres, 2012) and FREALIGN (Grigorieff, 2007; Lyumkis et al., 2013) produces virtually identical features in the well-defined density regions within the transmembrane domain (Figure 4A-D). However, some weaker peripheral features exhibit noticeable differences. The apparent blurring of peripheral features seen in the map refined using RELION (A) is caused by Maximum Likelihood-based “soft matching” of images (see “Projection Matching”), while the map refined with FREALIGN (B) is the result of hard assignments.

**Supplemental Text
Information Provided by Negative-Stain EM**

Negative-stain EM images provide a wealth of information on a sample, such as the presence of contaminants or aggregates, the size, shape and oligomeric state of the target protein or complex, the tendency of a complex to dissociate, creating compositional heterogeneity, as well as potential conformational variability. While impurities are

undesired, they do not need to be problematic. For example, if they differ greatly in size and/or shape from the target, contaminants can simply be excluded during particle picking or removed later during image processing. Knowledge of the size, shape and oligomeric state of the target can be very informative as thin extended domains may be difficult to see in vitrified ice whereas even relatively small but compact molecules may still be discernible. Symmetry will simplify subsequent image processing steps, whereas pseudo-symmetry can cause problems, requiring cautious image processing. If protein aggregation is observed, a low concentration of mild detergent or increasing the ionic strength of the buffer can potentially reduce the problem. Glycerol and sugars are often used to stabilize proteins and complexes in solution, but their use is not optimal for EM studies. In the case of negative staining, they can coat the protein and obscure structural features, and in the case of vitrification, they increase the density of the buffer, thus reducing the contrast created by the protein. These effects are particularly problematic for small molecules.

Negative Staining and Vitrification

In negative staining approaches, such as conventional negative staining and cryo-negative staining, the specimen is embedded in a layer of heavy metal salt crystals (De Carlo and Stark, 2010; Ohi et al., 2004). The advantages of negative staining are that the heavy metals introduce high contrast and that the procedure tends to induce the proteins to adsorb to the carbon support in preferred orientations, which makes it easy to quickly assess the quality and homogeneity of protein preparations. However, negative staining typically introduces artifacts, especially specimen flattening, and limits the achievable resolution to $\sim 20\text{\AA}$. Negative-stain EM is therefore ideal to assess the quality of protein samples, but it can also provide structural information in the cases in which cryo-EM is not feasible, as, for example, when the protein is too small or too heterogeneous.

In vitrification, currently the best specimen preparation procedure, the specimen is applied to a holey carbon grid, blotted with filter paper and then quickly plunged into liquid ethane cooled to liquid nitrogen temperature (Dubochet et al., 1988). Quick freezing solidifies the water without allowing ice crystals to form, which would otherwise damage biological molecules and generate strong contrast that obscures molecular features. Vitrification preserves the specimen in a near-native environment, introduces little or no artifacts, usually allows the proteins to adopt various orientations, and does not limit the resolution that can be achieved. However, vitrified specimens have low inherent contrast, as biological molecules scatter electrons not much more strongly than the surrounding buffer. Therefore, only molecules that scatter electrons sufficiently strongly can be distinguished from the background, imposing a lower mass limit on molecules that can be studied by cryo-EM (theoretically 50 to 100 kDa; Henderson, 1995). In practice the size limit used to be ~ 300 kDa, but, due to the introduction of DDD cameras, it is now possible to see smaller molecules, although this does not guarantee that it will be possible to calculate a 3D map.

Parameters that Affect Ice Thickness

The ice thickness can be modified by adjusting the blotting time and the subsequent time the grid is allowed to dry before it is plunge-frozen. Blotting with filter paper, which can

be done from one or both sides of the grid, physically removes excess fluid. By contrast, waiting before freezing the grid allows water to evaporate. The extent of evaporation depends on the time as well as on the temperature and humidity of the environment, all parameters that can be controlled with commercially available plungers and that need to be optimized empirically. Since evaporation only removes water, the solutes will become more concentrated, and the resulting changes in pH and ionic strength can alter the structure of proteins that are sensitive to these parameters. Other factors that affect ice thickness are the type of sample support film (e.g., homemade carbon, C-flat, Quantifoil), the particular batch of grids (grids can differ substantially between batches and occasionally plastic remnants on holey carbon films have to be removed by washing the grids in ethyl acetate), and the age and thus hydrophobicity of the support film, which can be modified by changing the glow-discharge parameters. A minimum degree of hydrophilicity is required to allow complete wetting of the grid and thus the formation of an ice layer.

Condenser Aperture, Spot Size, and Coma Alignment

The condenser aperture of the electron microscope removes electrons with a large angular spread from the beam and thus improves beam coherence. Hence, smaller apertures are preferable, but these also reduce beam intensity. Therefore, the smallest possible condenser aperture should be used that still provides sufficient beam intensity to record images with the desired dose rate and exposure time. The same rationale applies to the spot size, which is controlled by the first condenser lens. A smaller spot size corresponds to a smaller angular spread of the beam with the same condenser aperture. Thus, the smaller the spot size, the better the spatial coherence of the beam. For an excellent introduction on how to set up an electron microscope for high-resolution imaging, the reader is referred to Lau and Rubinstein (2013) as well as a lecture by John Rubinstein available online (<http://nramm.nysbc.org/2012-workshop-lectures/>).

Coma is an imaging aberration that leads to off-axis phase errors in the recorded images and needs to be minimized when setting up conditions for high-resolution imaging (Glaeser et al., 2011). Coma can be avoided by using parallel illumination, which is complicated by the necessity to have sufficient beam spread for proper illumination of the specimen. In single-particle cryo-EM, the spread is usually adjusted so that the beam diameter is slightly larger than the holes in the carbon film. In this way, the beam symmetrically illuminates the carbon film surrounding the hole, which is a better electrical conductor than vitrified ice and thus helps to reduce beam-induced charge build-up on the illuminated specimen area (Berriman and Rosenthal, 2012). In microscopes that use an electron optical system with three condenser lenses (such as the FEI Titan instruments), parallel illumination can be achieved for a wide range of beam sizes, so that the beam spread can easily be set to the desired size. It is more difficult, however, to achieve fully parallel illumination for microscopes that use a regular two-condenser-lens system, because under parallel illumination conditions the beam size is often too large (illuminating a much larger area than a single hole in the carbon film), resulting in a dose rate that is too low to collect images with a reasonable exposure time. Setting the required beam size thus results in nonparallel illumination, but it is possible to approximate parallel illumination with the use of smaller condenser apertures.

Protocol for Calibrating the Magnification

High-resolution single-particle cryo-EM requires accurate knowledge of the pixel size (equivalent to the effective magnification), which is needed for defocus determination and CTF correction, as well as for map interpretation. Calibration of the pixel size is equivalent to calibration of the effective magnification of a given electron microscope/detector combination. The accuracy of magnification calibration using grating replicas is insufficient for high-resolution single-particle cryo-EM. Typical specimens used for accurate calibration include thin oriented gold (2.04Å spacing), gold particles deposited on carbon film (2.35Å spacing), or graphitized carbon (3.4Å spacing), all of which are commercially available. The pixel size is then determined by measuring the distance of the reflections from the center of the Fourier transform, i.e.: $\text{pixel size} = r * d/D$, where r is the lattice spacing of the specimen in Å, d is the distance of the reflection to the center of the Fourier transform in pixel, and D is the dimension of the image in pixels (assuming the image is square). When using gold particles on carbon film or graphitized carbon, one can also calculate a rotational average and determine the distance of the reflections from the center in a one-dimensional power spectrum. After initial pixel size calibration, one can further refine the value by optimizing the docking of a known crystal structure into a corresponding cryo-EM density map.

The Contrast Transfer Function (CTF) of an Electron Microscope

In traditional light microscopy, optical lenses can be shaped with high precision and are combined to minimize imaging defects, approximating the physical limits imposed by the aperture acceptance angle and wavelength of light. The design and production of ideal electromagnetic lenses is more difficult and imaging aberrations are usually significant. Furthermore, samples prepared for high-resolution cryo-EM are typically only a few tens of nanometers thick and are predominantly made of light elements, including hydrogen, carbon, nitrogen, and oxygen. This means that an in-focus image will not contain much contrast (magnitude of variation between bright and dark features) as the electron transmission across a sample is approximately constant. Contrast in cryo-EM images is primarily generated by setting up the imaging system to obtain interference between unscattered and scattered electrons and, to a lesser extent, by aberrations of the electromagnetic lenses. This type of contrast is called phase contrast because it is generated by the differences in the phase shifts of electrons when they traverse different parts of the sample. Phase contrast can be generated with a phase plate, a technology that is currently being developed (Glaeser, 2013), but the main method to date for generating contrast is to deliberately underfocus the objective to induce phase shifts in the scattered electrons and thus contrast in the images. For weakly scattering samples, the image formation process in the electron microscope is linear, i.e., the densities in the collected image are linearly proportional to the densities in the imaged object, and this is described by the Contrast Transfer Function (CTF) (Frank, 2006; Wade, 1992).

CTF Estimation

Some of the parameters of the CTF are given by the settings of the microscope and some are established experimentally. In practice, the wavelength and spherical aberration

coefficient are considered to be instrument parameters defined by the acceleration voltage used (wavelength) and the lens design (spherical aberration), while the defocus and astigmatism have to be determined for each image. This is usually accomplished by adjusting the parameters so that the computed CTF approximates the amplitude oscillations (Thon rings) observed in the power spectrum of the image (Mindell and Grigorieff, 2003). Once values of all parameters are known, the mathematical form of the CTF is used to correct the image for the CTF effects. As the CTF is an oscillating function that crosses zero multiple times, some information is lost and a single image cannot be fully corrected. Therefore, it is necessary to collect many images at different defocus settings, as these give rise to CTFs with zeros at different spatial frequencies, and averaging them with appropriate weights (Penczek, 2010) will result in a CTF-corrected average.

Most single-particle EM software packages have associated CTF analysis programs, but there are also independent applications (e.g., Mallick et al., 2005; Mindell and Grigorieff, 2003; Sander et al., 2003; Zhou et al., 1996). While all programs generally yield similar results, assessment of errors of the results has so far only been implemented in SXCTER, which adopts a statistical resampling strategy to assess errors of defocus, astigmatism amplitude and angle (Penczek et al., 2014a). Some software evaluates broader image formation characteristics, such as the envelope function (B-factor) (Huang et al., 2003; Saad et al., 2001), or even the distribution of the Spectral-Signal-to-Noise Ratio (SSNR) in the data. Since there is no generally accepted methodology of single-particle EM structure determination that would include the SSNR characteristics of the data, this class of methods is closely associated with particular software packages, such as EMAN2 (Tang et al., 2007). Other software packages only estimate fundamental parameters of the CTF, sometimes only the defocus, but often also astigmatism and amplitude contrast (Mindell and Grigorieff, 2003). These methods tend to be generally usable, especially since a standardized CTF convention is now emerging (Marabini et al., 2013).

Accurate CTF estimation depends on the image size used to window the particles. To have reasonable accuracy, the CTF parameters should be estimated using a relatively large window size. The fact that the fit is good and Thon rings are visible to high frequency does not mean that the CTF will be properly represented in a much smaller particle window as the point spread function associated with the CTF is very broad. Its truncation upon real-space windowing has an adverse impact on the representation of the CTF in reciprocal space. The issue is further complicated by the impact of the pixel size on the representation of the CTF in a discrete image grid (the larger the pixel size the less accurate is the CTF represented). The general relations have been worked out (Penczek et al., 2014a), and it is advisable to consider them prior to data collection as otherwise the attainable resolution might be compromised.

CTF estimation has to be done at the very beginning of a project by computing power spectra using the entire micrograph field. Additional improvement of the power spectrum SNR is typically accomplished by averaging local power spectra computed from small partially overlapping sections of the micrograph (Zhu et al., 1997). However, the estimate can be inaccurate, particularly if the image includes part of the support carbon film. In this case, the power spectrum is dominated by the signal from the carbon,

resulting in an estimated defocus that is different from that of the particles, which can be suspended a few hundred Ångstroms above or below the surface of the carbon film. It is thus advisable, particularly for high-resolution projects, once the structure determination process has yielded an intermediate-resolution CTF-corrected map, to adjust the CTF estimates using Fourier cross-resolution (FCR) (Mouche et al., 2001; Penczek et al., 2014a).

Correction for the CTF

CTF-correction methodologies differ in the exact mechanism of CTF application. Most commonly the CTF is applied to the windowed particles, for which the window size could be relatively small. Current software packages do not assess the accuracy with which the CTF is represented, as described in “CTF estimation”, and this could impose a limit in resolution. Alternatively, assuming that the defocus is constant across the image (Mindell and Grigorieff, 2003), one can apply the CTF to micrographs prior to the windowing of particles (as can be done, for example, in SPARX). This has the advantage that the large micrograph field affords proper representation of the CTF, even at high underfocus.

As long as a set of particle views that differ by defocus settings is available, the full correction for the CTF effects is possible and straightforward (Penczek, 2010), but there are practical problems. For negatively stained specimens, due to their high contrast, it is thus advisable to collect data close to focus, so that CTF considerations are unnecessary. However, cryo-EM projects usually require CTF correction. The exact strategies differ between software packages to the extent that the transfer of intermediate results between them is challenging if not impossible as CTF correction steps might be performed incrementally at different stages of the project. Proper CTF correction calls for reciprocal-space multiplication of the data by the CTF, but for low-quality data and particularly for 2D analysis this procedure can introduce artifacts, such as excessive enhancement of low frequencies. Phase flipping (changing the sign of the CTF in the appropriate regions but leaving the amplitudes unchanged) is the method of choice for such data. However, it is suboptimal in regions of reciprocal space in which the CTF values and thus the Spectral SNR (SSNR) of the data are low, as they have the same weight assigned as high CTF/SSNR regions (Penczek, 2010). Hence, some software packages perform full CTF correction independently for each step of the data processing, i.e., 2D alignment, 3D reconstruction and so on, whereas others perform CTF correction as a more or less independent step. For example, EMAN2 (Tang et al., 2007) estimates the SSNR characteristics of the data, applies phase flipping in the initial stages of the project, and only incorporates information about CTF amplitudes in later stages.

Windowing of Particles: Aliasing and Density Normalization

Very low frequencies in the micrograph originate from unevenness in the illumination, in the ice layer thickness and/or in the supporting carbon film, if used. When particles are windowed into smaller boxes, these low frequencies will be incorrectly mapped into the low frequency range at the box scale. A simple remedy is to apply a high-pass Gaussian Fourier filter with a half-width inversely proportional to the box size to the micrographs

prior to windowing. To eliminate a linear density trend in a particle image, a 2D ramp density gradient is fit and subtracted from the particle image.

The commonly used strategy to normalize density is to subtract from the image the average pixel density computed outside of the particle region and to set the variance such that it equals one within the particle region. This procedure yields acceptable results, but it is not entirely correct as it assumes that the signal in each image contributes the same amount of variance and neglects the influence of defocus on noise distribution. Furthermore, this assumption is not true because the variance will depend on the particle view and will vary strongly if the particle is significantly elongated in one direction. Ideally, the variance of the background around each particle is measured and the image is then scaled to set this variance to a defined value. However, this procedure is marred with difficulties as the high noise level in EM data precludes the reliable estimation of signal variance, so that there is currently no better procedure than the commonly used one.

Principles and Properties of K -Means Clustering

The goal of clustering is to assign (exclusively) n objects to K classes (or groups) such that objects within each class are similar to each other, while the group averages are as dissimilar from each other as possible (Duda et al., 2001). The only approach that guarantees finding an optimum solution to this problem is to try all possible assignments and select the one with the best score, but this procedure by far exceeds what is computationally feasible. Therefore, various more or less *ad hoc* solutions were proposed, and K -means clustering is the one most commonly used in scientific applications.

The algorithm has to be initialized: the number of classes K has to be decided and then the algorithm is either started in a deterministic or random manner. In deterministic initialization, the algorithm is presented with K templates. In 3D MRA, for example, these can be K plausible structures of the complex obtained by some means earlier in the project. Preferable is random initialization, which can be accomplished in one of two ways: either K randomly selected objects from the entire set are used as initial templates or all objects are randomly assigned to K classes such that each class has the same number of objects and within-group averages are used as initial templates.

After initialization the algorithm proceeds in an iterative manner by alternating two steps: (1) for each object its distances (similarity) to all class averages are computed and the object is assigned to the nearest (most similar) average; (2) based on the new assignments, new class averages are computed.

This simple algorithm is guaranteed to terminate (converge) in a finite number of steps, it is very fast and easy to implement. Finally, it yields good results if the data are not exceedingly noisy, if the groups are linearly separable, and if the number of groups K was guessed correctly. However, in applications to EM the shortcomings outweigh the positives: (1) the algorithm is not guaranteed to converge to a global optimum (the best possible solution) and finding the global optimum is not feasible in practice using the strategy described above, (2) outliers (rare objects whose appearance is partially or entirely unrelated to that of the bulk of the data) have a very negative impact on the

outcome, (3) there is nothing in the design of the algorithm that would address the possibility of “group collapsing”, i.e., the possibility of a group losing its members to the point of vanishing, finally and most importantly, (4) if the number of groups is not guessed correctly and the groups are not well separable (which is always the case for very noisy data), the result depends dramatically on the initialization. To expand on the last point, in applications to EM, when the algorithm is initialized with predetermined templates, the outcome will most likely strongly resemble these templates, irrespective of how correct they are, a problem known as “template bias”. For a randomized initialization the results tend to be unstable, i.e., irreproducible in the sense that the outcome will vary significantly upon repeated application to the same dataset. In most applications of the K -means strategy in EM, the problems are further compounded by the need for particle alignment in addition to classification, i.e., concurrently with clustering, the images have to be brought into register. This significantly increases the number of parameters to be determined and weakens the performance, particularly with respect to the problem of “group collapsing”.

Convergence properties of K -means can be improved using the concept of so-called “soft assignment”. In the assignment step of the algorithm, instead of a binary decision (object can belong to one group only) one allows “partial” membership, i.e., an object can be a member of several groups with weighting factors that are derived from the similarities between the object and the class averages. In heuristic approaches, the weights are progressively “sharpened” (an increasingly large weight is used for the most similar assignment) and at convergence weights become binary. The procedure can be formalized using the Maximum Likelihood methodology, in which in addition to membership other statistical properties of the data are estimated (MacKay, 2003). It has to be noted that while such modifications improve convergence by reducing the chance of premature convergence, the fundamental problems remain: the convergence is only local, results depend on initialization, and groups can collapse.

The Iterative Stable Alignment and Clustering (ISAC) Method

In light of the fundamental shortcomings of MRA (see “Principles and properties of K -means clustering”), one has to be judicious when using MRA-like methods and seek validation of the results as, for example, implemented in the Iterative Stable Alignment and Clustering (ISAC) method (Yang et al., 2012). ISAC is based on the observation that a homogeneous group of 2D images can be reproducibly aligned, i.e., multiple and independent randomly initialized reference-free alignments of such a set of images will yield virtually identical 2D averages, and the alignment parameters of individual images will be approximately the same. However, if outliers are present, their alignment parameters tend to differ dramatically between alignment trials. In the ISAC approach, image members that always appear in the same orientation (within a predefined error threshold) are called stable, while the remaining ones are considered outliers and can thus be removed from the group. ISAC also addresses the issue of collapsing groups by using a dedicated clustering algorithm (EQK-means) that delivers equal size groups (Yang et al., 2012). Despite the group validation built into the method, there is no guarantee that all the classes are reliable, i.e., would be reproduced by another, independent run of the program. Therefore, in ISAC, the basic algorithm is executed four times on the same

dataset and the assignments of images to classes are checked for agreement. Groups that are reproducible between independent runs are ultimately presented as the result of ISAC.

Due to the intrinsic multi-level validation strategy of ISAC, it delivers uniquely homogeneous (and thus easy to orient in 3D) classes of images. It almost never assigns all particle images to classes. Many remain “unaccounted for” and can be processed in subsequent “generations” of the ISAC analysis. It is also possible for ISAC to fail to deliver any classes, which indicates poor data quality. The major drawback of ISAC is that it is very computation-intensive, and it is therefore best suited for *ab initio* projects or for exploratory data analysis.

Projection Matching

Projection matching was developed for the purpose of refinement of a low-resolution structure using a set of 2D EM projection images (Penczek et al., 1992). The initial structure is typically generated using one of the *ab initio* methods, for example random conical tilt or a computational approach. The procedure alternates between two steps: (1) Quasi-evenly distributed 2D reprojections of the current 3D template are generated using a resolution-dependent angular step (for structures with point-group symmetry, only the asymmetric subunit is covered) and EM 2D images are computationally matched (aligned) with all 2D templates. The best match yields orientation parameters that are assigned to the EM images. (2) Once all EM images have assigned parameters, a 3D reconstruction is computed, yielding an updated structure, which is then post-processed (low-pass filtered, masked, etc.) and used as the new template in the next iteration of the process. The procedure is iterated until projections stop changing angular assignments or some other criterion is met, for example, the target resolution is reached.

The algorithm can be seen as an extension of the K -means algorithm. The number of classes K is the number of reference projection directions into which the current approximation of the 3D object is reprojected, creating a set of K 2D templates. The assignment step is the same as in the MRA algorithm, i.e., all images are compared with all templates and assigned to the most similar one. The reprojections of the subsequently reconstructed 3D template can be considered to be “directional, properly weighted class averages”. In this description, projection matching is a K -means algorithm with image alignment and enforced consistency among class averages.

At the end of every projection matching iteration, an FSC curve is computed using half-subsets of the projection data and the current resolution is derived from it. Based on the FSC, the structure is low-pass filtered and passed to the next iteration of the procedure as the new template. The procedure shares many features of the K -means algorithm: rapid convergence, ease of implementation, but also bias towards the initial template and inability to overcome local minima. It falls into the category of greedy optimization algorithms as at each step it accepts the best possible matches, resulting in premature convergence (entrapment in a local minimum). As a result, the final structure is unlikely to be the best possible one or, if the initial 3D template was incorrect, the procedure is unlikely to converge to the correct one. In addition, it suffers from a unique problem known as over-fitting or noise alignment, which causes exaggerated resolution estimates when a single dataset is refined and there is no external check (Grigorieff, 2000; Stewart and Grigorieff, 2004).

There are variations of projection matching algorithms. (1) In the basic design of local (or restricted) refinement, the initial comparison with an EM image is exhaustive, i.e., all templates are considered. However, once approximate orientations are known, further searches to identify improved orientations can be restricted to a predefined angular neighborhood of the known orientation of the EM image. This modification significantly accelerates calculations and, once high resolution is reached, prevents images from adopting distant orientations that would be in conflict with what was established at low resolution. (2) In a “soft matching” extension one considers a number of possible angular assignments for a given EM image. Instead of assigning a unique orientation to an image based on the best match, one assigns a number of orientations to the image (in the extreme case, the number would be equal to the number of used 2D template reprojections) with relative weights derived from similarity values computed during matching. This extension significantly improves the convergence properties of projection matching. Variants of soft matching are implemented in EMAN2 and SPARX. A soft matching strategy fundamentally rooted in statistical considerations is the basis of the Maximum Likelihood approach, as implemented in RELION. The inclusion of soft matching carries two dangers: (1) the linearity of map densities is disturbed and the contrast is disproportionally enhanced, which improves the visual appearance of the map, but at the cost of possible suppression of weaker features (Figure S2), (2) the resolution of the map, as evaluated using FSC, is artificially overstated (Figure 4E) and the exact amount of the exaggeration is difficult to quantify as it depends on the noise level in the data and, paradoxically, the shape of the structure. Maximum Likelihood-based projection matching offers an attractive robust alternative to basic projection matching even though, ultimately, it remains a local optimizer and in general cannot overcome initial template bias. It is also time-consuming, with direct implementation prohibitively so, requiring additional *ad hoc* acceleration schemes (Kucukelbir et al., 2012; Scheres, 2012).

Stochastic Hill Climbing (SHC) in Projection Matching

The modification of projection matching from using the greedy algorithm to using the Stochastic Hill Climbing (SHC) optimization (Brownlee, 2012) is straightforward. Instead of finding the best possible match between an EM projection image and all reprojection templates, the templates are scanned in a random order and the first one with a similarity score better than the one obtained at a previous iteration is accepted. It is clear that if started from randomized orientation parameters (or a 3D distribution of random noise as the template), the procedure will initially accept orientations that are entirely incorrect and only slowly drift towards increasingly sensible orientations. This partially incorrect drift allows the procedure to overcome local minima and dramatically increases the possibility of finding the correct solution. While SHC significantly improves the convergence properties of the greedy projection matching method, it does not guarantee convergence to the global minimum, i.e., it may also converge to an incorrect solution and, in the absence of other tests or external information, it is impossible to determine whether the procedure failed or succeeded.

The Validation of Individual Parameter Reproducibility (VIPER) Method

Projection matching is a simple algorithm aimed at improving the quality of an initial 3D template given a set of EM images (Penczek et al., 1994). Recently, the convergence properties of projection matching were dramatically improved by introducing the Stochastic Hill Climbing (SHC) algorithm, which replaces the previous greedy algorithm of the original design and mitigates its shortcomings, particularly the problem of premature convergence (see “Projection matching”). This feature makes SHC-based projection matching a perfect candidate for *ab initio* 3D structure determination, in particular, because in its basic form, it does not require an initial guess of the structure. The SHC strategy was first implemented in the software package SIMPLE (Elmlund and Elmlund, 2012). The method is further enhanced by the inclusion of “soft matching”, i.e., projection images are permitted to simultaneously adopt a predefined number of stochastically determined projection directions. This improves robustness of the algorithm by improving the possibility of finding the correct structure.

While SHC is much more likely to find a plausible solution than basic projection matching, there is no guarantee that it will converge as it can be trapped in a local minimum (i.e., deliver an incorrect structure). Therefore, the Validation of Individual Parameter Reproducibility (VIPER) approach has been introduced, which incorporates validation steps into the structure determination process, monitoring the orientation parameters (Penczek, 2014b). Similarly to ISAC, VIPER uses stability and reproducibility tests. While it cannot be guaranteed to always determine the structure, its failure indicates that the set of input class averages cannot be reproducibly aligned, requiring a critical look at the dataset.

Normalization and Power Spectrum Adjustment of a Refined EM Map

An EM structure that emerges from a multi-staged computational structure determination process has undefined normalization. It is required, though, that the relation between map densities and the original 2D EM data is linear, as only in this case densities in the EM map can be properly interpreted in terms of protein and/or nucleic acid densities. However, there is no consensus as to what the real-space normalization should be, and the density scales of EM maps tend to differ between packages. Moreover, in some packages a utility might be provided to perform final normalization, as in RELION, but most implementations leave this responsibility to the user. Assuming linearity of the data-processing steps, it is reasonable to follow the real-space normalization advocated for the input 2D data: have the average of the background surrounding the structure set to zero and the variance of the structure densities to one.

The issue of correcting for the reciprocal space fall-off of the power spectrum is more complex. Not only is there no adopted standard for such a correction, but it is not even clear what the power spectrum of an electron microscope-derived map should be (for a review and possible solutions see Penczek, 2010). Naïve thinking suggests that the electron microscope is a simple imaging system and thus has an overall envelope function that can be estimated, and the EM map can be adjusted (“sharpened”) to correct for the power spectrum fall-off induced by the microscope. In the common B-factor approach, the EM envelope function is assumed to be a Gaussian function with a half-width defined by B, so one can simply divide the Fourier transform of the map by this function to obtain the “correct” map. The value of B is ideally obtained by examination

of the power spectra of the micrographs (raw data), and a comparison with a “reference” power spectrum of the specimen obtained from an independent SAXS experiment (Saad et al., 2001). This line of thinking neglects the fact that the ultimate EM map is affected by other envelope functions, most notably the alignment error envelope function resulting from the imprecision of alignment programs (Baldwin and Penczek, 2005). Moreover, as SAXS measurements are not routine, it appears that in many cases the value of the B-factor is simply adjusted arbitrarily by the user to achieve a desired “sharpness” of the map, as judged by visual appearance. A sensible alternative is to assume that the reconstructed map should have the same appearance as a map derived from an X-ray model shown at the same resolution. This makes the power spectrum adjustment (sharpening) protocol appealingly simple: the power spectrum of the EM map is adjusted such that its rotational average has the same fall-off as the power spectrum of an X-ray model of the same structure. In the absence of an atomic model of the structure, a homologous model can be used or simply a rotational power spectrum of any X-ray structure that has about the same size and shape. This approximation is justified by the fact that families of similar macromolecules have very similar power spectra (Penczek, 2010).

The power spectrum fall-off adjustment cannot be separated from the final step of the map preparation, which is the low-pass filtration necessary to suppress noise in the map. There are two commonly used ways to “smoothen” the map: (1) truncation of the Fourier transform of the map, i.e., setting Fourier coefficients beyond the nominal resolution to zero or (2) application of a Gaussian function, the half-width of which is derived from the nominal resolution number. However, both methods should be avoided. Fourier truncation (aka top-hat filter) will induce a so-called Gibbs oscillation into the map, resulting in spurious surface features and a somewhat fake high-resolution appearance. A Gaussian low-pass filter has non-zero values in the entire spatial frequency range, so the map, while having a smooth appearance, will always contain a contribution of high-frequency noise, opening up the possibility for spurious artifacts. Ideally, the map is low-pass filtered using a dedicated filter constructed based on the shape of the FSC curve: its values are set to one in low frequencies, then there is a smooth fall-off such that values 0.5 of the FSC curve and the filter appear at the same spatial frequency, and finally the values are set to zero roughly at the point where the FSC drops below 0.1 (Penczek, 2010). The advantage of this method is that it minimizes the arbitrariness of other approaches, and it also makes it possible to directly compare surface representations of EM and X-ray results, which, in turn, provides ultimate validation of the claimed resolution (Penczek et al., 2014).

Supplemental References

Baldwin, P.R., and Penczek, P.A. (2005). Estimating alignment errors in sets of 2-D images. *J Struct Biol* *150*, 211-225.

Berriman, J.A., and Rosenthal, P.B. (2012). Paraxial charge compensator for electron cryomicroscopy. *Ultramicroscopy* *116*, 106-114.

Brownlee, J. (2012). *Clever Algorithms: Nature-Inspired Programming Recipes* (LuLu.com).

- De Carlo, S., and Stark, H. (2010). Cryonegative staining of macromolecular assemblies. *Methods Enzymol* *481*, 127-145.
- Dubochet, J., Adrian, M., Chang, J.J., Homo, J.C., Lepault, J., McDowell, A.W., and Schultz, P. (1988). Cryo-electron microscopy of vitrified specimens. *Q Rev Biophys* *21*, 129-228.
- Duda, R.O., Hart, P.E., and Stork, D.G. (2001). *Pattern Classification* (New York, NY: Wiley).
- Frank, J. (2006). *Three-dimensional electron microscopy of macromolecular assemblies: visualization of biological molecules in their native state* (New York, Oxford University Press).
- Glaeser, R.M., Typke, D., Tiemeijer, P.C., Pulokas, J., and Cheng, A. (2011). Precise beam-tilt alignment and collimation are required to minimize the phase error associated with coma in high-resolution cryo-EM. *J Struct Biol* *174*, 1-10.
- Glaeser, R.M. (2013). Methods for imaging weak-phase objects in electron microscopy. *Rev Sci Instrum* *84*, 111101.
- Grigorieff, N. (2000). Resolution measurement in structures derived from single particles. *Acta Crystallogr D Biol Crystallogr* *56*, 1270-1277.
- Grigorieff, N. (2007). FREALIGN: High-resolution refinement of single particle structures. *J Struct Biol* *157*, 117-125.
- Henderson, R. (1995). The potential and limitations of neutrons, electrons and X-rays for atomic resolution microscopy of unstained biological molecules. *Q Rev Biophys* *28*, 171-193.
- Huang, Z., Baldwin, P.R., Mullapudi, S., and Penczek, P.A. (2003). Automated determination of parameters describing power spectra of micrograph images in electron microscopy. *J Struct Biol* *144*, 79-94.
- Kucukelbir, A., Sigworth, F.J., and Tagare, H.D. (2012). A Bayesian adaptive basis algorithm for single particle reconstruction. *J Struct Biol* *179*, 56-67.
- Lau, W.C., and Rubinstein, J.L. (2013) Single particle electron microscopy. *Methods Mol Biol* *955*, 401-426.
- Liao, M., Cao, E., Julius, D., and Cheng, Y. (2013). Structure of the TRPV1 ion channel determined by electron cryo-microscopy. *Nature* *504*, 107-112.
- Lyumkis, D., Brilot, A.F., Theobald, D.L., and Grigorieff, N. (2013). Likelihood-based classification of cryo-EM images using FREALIGN. *J Struct Biol* *183*, 377-388.
- MacKay, D.J.C. (2003). *Information Theory, Inference and Learning Algorithms* (Cambridge, Cambridge University Press).
- Mallick, S.P., Carragher, B., Potter, C.S., and Kriegman, D.J. (2005). ACE: automated CTF estimation. *Ultramicroscopy* *104*, 8-29.
- Marabini, R., Macias, J.R., Vargas, J., Quintana, A., Sorzano, C.O., and Carazo, J.M. (2013). On the development of three new tools for organizing and sharing information in three-dimensional electron microscopy. *Acta Crystallogr D Biol Crystallogr* *69*, 695-700.

- Mindell, J.A., and Grigorieff, N. (2003). Accurate determination of local defocus and specimen tilt in electron microscopy. *J Struct Biol* *142*, 334-447.
- Mouche, F., Boisset, N., and Penczek, P.A. (2001). *Lumbricus terrestris* hemoglobin – the architecture of linker chains and structural variation of the central toroid. *J Struct Biol* *133*, 176-192.
- Ohi, M., Li, Y., Cheng, Y., and Walz, T. (2004). Negative staining and image classification – powerful tools in modern electron microscopy. *Biol Proced Online* *6*, 23-34.
- Penczek, P.A., Radermacher, M., and Frank, J. (1992). 3-Dimensional reconstruction of single particles embedded in ice. *Ultramicroscopy* *40*, 33-53.
- Penczek, P.A., Grassucci, R.A., and Frank, J. (1994). The ribosome at improved resolution: new techniques for merging and orientation refinement in 3D cryo-electron microscopy of biological particles. *Ultramicroscopy* *53*, 251-270.
- Penczek, P.A. (2010). Image restoration in cryo-electron microscopy. *Methods Enzymol* *482*, 35-72.
- Penczek, P.A., Fang, J., Li, X., Cheng, Y., Loerke, J., and Spahn, C.M. (2014a). CTER – rapid estimation of CTF parameters with error assessment. *Ultramicroscopy* *140*, 9-19.
- Penczek, P.A. (2014b). sxviper. In SPARX Wiki (<http://sparx-em.org/sparxwiki/sxviper>).
- Saad, A., Ludtke, S.J., Jakana, J., Rixon, F.J., Tsuruta, H., and Chiu, W. (2001). Fourier amplitude decay of electron cryomicroscopic images of single particles and effects on structure determination. *J Struct Biol* *133*, 32-42.
- Sander, B., Golas, M.M., and Stark, H. (2003). Automatic CTF correction for single particles based upon multivariate statistical analysis of individual power spectra. *J Struct Biol* *142*, 392-401.
- Scheres, S.H.W. (2012). RELION: Implementation of a Bayesian approach to cryo-EM structure determination. *Journal of Structural Biology* *180*, 519-530.
- Stewart, A., and Grigorieff, N. (2004). Noise bias in the refinement of structures derived from single particles. *Ultramicroscopy* *102*, 67-84.
- Tang, G., Peng, L., Baldwin, P.R., Mann, D.S., Jiang, W., Rees, I., and Ludtke, S.J. (2007). EMAN2: an extensible image processing suite for electron microscopy. *J Struct Biol* *157*, 38-46.
- Wade, R.H. (1992). A brief look at imaging and contrast transfer. *Ultramicroscopy* *46*, 145-156.
- Yang, Z., Fang, J., Chittuluru, J., Asturias, F.J., and Penczek, P.A. (2012). Iterative stable alignment and clustering of 2D transmission electron microscope images. *Structure* *20*, 237-247.
- Zhou, Z.H., Hardt, S., Wang, B., Sherman, M.B., Jakana, J., and Chiu, W. (1996). CTF determination of images of ice-embedded single particles using a graphics interface. *J Struct Biol* *116*, 216-222.

Zhu, J., Penczek, P.A., Schröder, R., and Frank, J. (1997). Three-dimensional reconstruction with contrast transfer function correction from energy-filtered cryoelectron micrographs: procedure and application to the 70S *Escherichia coli* ribosome. *J Struct Biol* 118, 197-219.