CHAPTER 2

Visualization of Molecular Machines by Cryo-Electron Microscopy

Joachim Frank

I. INTRODUCTION

It is difficult nowadays to provide an introduction into cryo-EM within the space of a book chapter, given the current plethora of different methods, and the fact that there are as yet no agreed-on standards in the field. In view of this situation, the best course for the author is to provide the reader with an illustrated introduction into important concepts and strategies. However, at the same time, the focus on the molecular machine invites an expansion of scope in the most relevant section (Section IV), which concerns itself with heterogeneity, and the challenge to obtain an inventory of conformational states of a molecular machine in a single scoop.

I.1. Preliminaries: Cryo-EM as a Technique of Visualization

The transmission electron microscope (TEM) produces images that are projections of a three-dimensional object. To be more precise, the projections are line integrals of the three-dimensional Coulomb potential distribution representing the object. (For all practical purposes, especially in the resolution range down to to ~3 Å, the Coulomb potential distribution is identical to the electron density distribution "seen" by X-rays). Visualizing a molecular machine in three dimensions therefore entails the collection of multiple images showing the molecule in the same processing state (and hence, identical structure) but in different views. Thus the term "3D electron microscopy" is understood as a combination of two-dimensional imaging, following a particular data collection strategy, with three-dimensional reconstruction.

In the electron microscope, a high vacuum must be maintained that allows electrons to travel collision-free over the distances (in the order of meters) required for imaging. This requirement poses difficulties in imaging biological molecules, as they are hydrated. An initial solution to this problem was the negative staining technique, whereby heavy metal salt is added to the aqueous solution in which the molecules are suspended, and the sample is then air-dried on the grid so that each molecule is encased in a layer of heavy-metal stain. Here the contrast is mainly produced by stain exclusion, hence the finer details of the structure are lost even though the molecule is to some extent preserved in its three-dimensional aspects. Another solution to this problem is the use of glucose as an embedding medium (Unwin and Henderson, 1975), which, however, poses difficulties in its application to single molecules because the density of the medium is closely matched to the density of the protein. The third method, which addresses the problem of vacuum incompatibility without having the drawbacks of negative staining or glucose embedment, is the frozen-hydrated specimen preparation technique (Taylor and Glaeser, 1976; Dubochet et al., 1982): Here molecules are embedded in a thin layer of ice that is kept at a temperature where evaporation is negligible. In the following, we reserve the term "cryo-electron microscopy" (cryo-EM) for this particular preparation and imaging method.

I.2. Feasibility of Imaging Structure and Dynamics of Molecular Machines

In discussing the imaging of molecular machines by three-dimensional electron microscopy, we need to examine in which way the technique lends itself to the study of dynamics, or the different conformational states that either coexist or follow one another as the machine performs its work. In what has been termed *four-dimensional imaging* (e.g., Heymann et al., 2004) – the acquisition of a whole sequence of three-dimensional data as the molecule changes dynamically over time – one must keep in mind that, unlike the case of imaging with light microscopy where live processes can be analyzed relatively undisturbed, molecules are instantly incapacitated by the hostile conditions of sample preparation and the impact of the electron beam, making it impossible to collect data from the same dynamically changing molecule more than once. Thus, the "fourth dimension" must be introduced in a virtual manner, by collecting three-dimensional images from *separate* ensembles of molecules "frozen" at different states of processing and linking these states into "time lines" using additional evidence from different kinds of experiments such as single-molecule FRET. This kind of conjecture is by no means unambiguous,

https://doi.org/10.1017/CBO9781139003704.004 Published online by Cambridge University Press

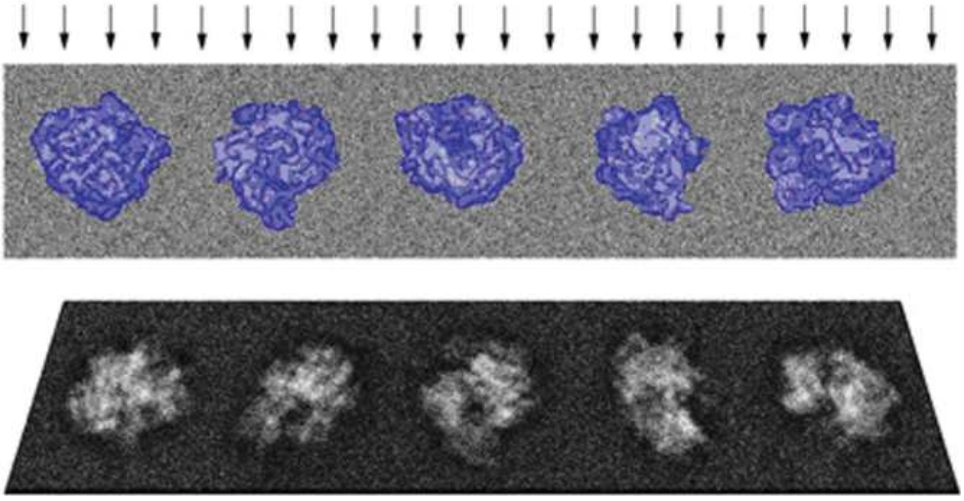


FIGURE 2.1: Data collection for molecules in single-particle form. The molecule exists in multiple copies, all in different orientations in a matrix of vitreous (i.e., virtually amorphous) ice. The electron beam produces a set of projections, from which the molecule can be reconstructed.

because two different states may be connected by multiple pathways, as is well known in enzyme kinetics (Boehr et al., 2006) and protein folding (Hartl and Hayer-Hartl, 2009).

Ideally we would wish to look at a molecular machine in the context of the cell, "frozen" at different time points in the act of performing its work. This type of imaging, however, is difficult to achieve, for several important reasons. One reason is that with a few exceptions, a cell's thickness exceeds the thickness that can be penetrated by electrons in the 100–300 kV range (~0.2 μm) by a large factor, necessitating the use of complicated techniques for high-pressure freezing of the cell and then sectioning the frozen cell to the required thickness (Hsieh et al., 2002; Al-Amoudi et al., 2004; see Koster and Barcena, 2006). Such slices can then be imaged by performing a tilt experiment using a series of small angular increments, collecting a set of projections from which a three-dimensional image can be reconstructed (a technique called "electron tomography" – see Frank, 2006a). The combination of three delicate techniques, especially the sectioning of the frozen specimen, is fraught with difficulties, producing a success rate of just a few percent. Another problem is that the molecule of interest is surrounded by a crowded environment of other structures, making it difficult to discern its boundaries. A third is the accumulation of dose, as many tilt images must be collected from the same area of the sample.

For all the reasons enumerated, imaging molecular machines, with the purpose of obtaining dynamic information, has to date been done almost exclusively using cryo-EM of in vitro samples combined with single-particle reconstruction. In vitro systems have been developed for a variety of fundamental processes in the cell, such as transcription, translation, protein degradation, and protein folding. Such systems, typically dependent in their action on ATP or GTP hydrolysis, may be trapped at key points of their processing path by the use of non-hydrolyzable analogs, or by small molecule inhibitors such as antibiotics – molecules that interfere with the dynamics of the process through targeted binding to specific sites, such as flexible hinge regions, which are required for mobility or for a critical binding interaction.

I.3. Cryo-EM of Molecular Machines in Single-Particle Form

Cryo-EM of molecules in single-particle form (as opposed to molecules occurring in ordered aggregates, such as two-dimensional crystals or helical bundles; Figure 2.1) is a technique of three-dimensional visualization developed over the past three decades (for recent introductory articles or reviews, see van Heel et al., 2000; Frank, 2006b; Wang and Sigworth, 2006; chapter on cryo-EM in Glaeser et al., 2007; Frank, 2009). It is suited for obtaining three-dimensional images of molecular machines in their native state, in vitro, captured in the process of performing their work.

The idea of data collection in such a case is to make use of the fact that the molecule occurs in multiple copies with (essentially) identical structure and that its orientation samples the entire angular range without leaving major gaps. Thus, instead of having to tilt the grid on which the sample is spread into multiple angles (as in electron tomography), it is then possible to take snapshots of multiple fields and, after suitable alignment, combine all projections into a density map depicting the molecule in three dimensions.

In comparison with X-ray crystallography, the absence of intermolecular contacts in the crystal means that the full range of native conformations and binding states can

https://doi.org/10.1017/CBO9781139003704.004 Published online by Cambridge University Press

be observed. On the other hand, lack of a means for uniform fixation of all molecules in a definite conformation, as happens in a crystal, also has the consequence that atomic resolution is much more difficult to achieve, because normally none of the states is populated in sufficient numbers. Failure to reach atomic resolution, as we will see, is at least partially mitigated by methods of approximate atomic interpretation of the reconstructed 3D density map, mostly by flexible fitting of atomic structures obtained by X-ray crystallography.

By now, the combination of the two techniques – cryo-EM and single-particle reconstruction – has yielded spectacular results contributing insights on structure and function of many molecular machines, some of which are featured in separate chapters of this book.

II. THE BASIC TECHNIQUE

II.1. Specimen Preparation: Principle

In the basic experimental technique of cryo-EM, developed at the European Molecular Biology Laboratory (EMBL) in the early 1980s (Dubochet et al., 1982; see Dubochet et al., 1988) following the groundbreaking work by Kenneth Taylor and Robert Glaeser (1976; also see historical perspective by Taylor and Glaeser, 2008), a freeze-plunger (Figure 2.2) is employed to rapidly freeze the liquid aqueous sample, such that vitreous (amorphous) ice is formed. In this simple apparatus, an EM grid (3 mm in diameter, usually made of copper) is suspended at the tip of a pair of tweezers, which is in turn fastened to a vertically mounted, gravity-operated steel rod.

Of particular importance with regard to the imaging of molecular machines is the question to what extent the biological molecule is preserved in its native state. The transition of liquid water to a glass-like "vitreous" state induced by jet-freezing was first described by Brueggeler and Mayer (1980). Subsequently, Dubochet and coworkers demonstrated that this state could also be induced by rapid flash-freezing (or freeze-plunging) of an EM grid on which liquid sample has been deposited in a thin (~100 nm) layer, with the necessary thinning achieved by blotting off excess sample prior to plunging (Dubochet et al., 1982). The method by Dubochet et al. (1988) is now routinely used in the cryo-EM community.

According to studies of the physics of flash-freezing on small (several microns) protein crystals by Halle (2004), however, cryo-structures do not represent the equilibrium structure at the ambient temperature before the freezing, but exhibit extra heterogeneity that affects different parts of the structure differently. The author cites the example of a protein structure determined at 100° K, which might have exposed side chains sampling conformations related to 200° K substates while the same global backbone fold is maintained as at room temperature. These concerns, if

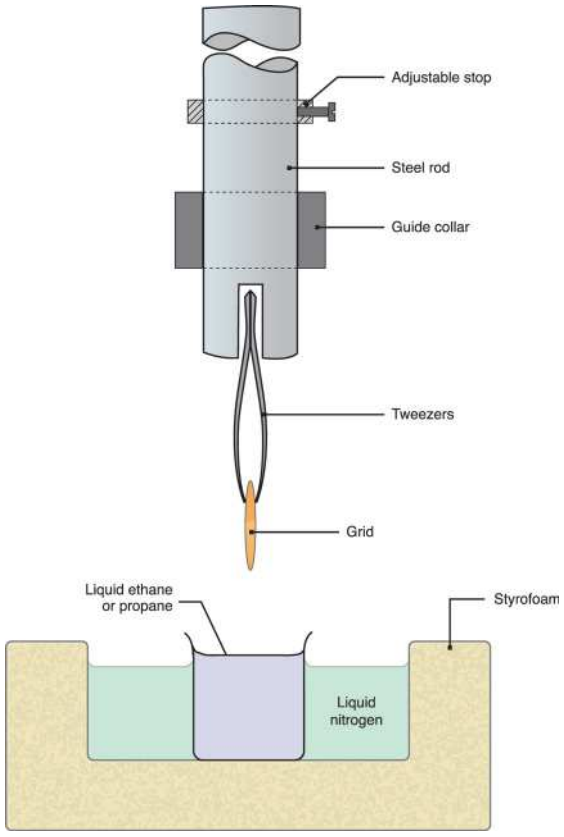


FIGURE 2.2: Preparation of the specimen: Schematic of a freeze-plunger. A droplet of the buffer in which the molecules are suspended is applied to the electron microscope grid, which is held by tweezers. The tweezers are mounted on a gravity-operated rod held by a guide collar. Release of the rod results in the rapid immersion of the grid in the cryogen. The rapid plunge in temperature has the result that the aqueous solvent is converted into vitrified ice.

indeed valid for cryo-EM preparation, may not carry much weight at resolutions at which side chains are not visible, but they will have to be taken into account in the interpretation as the field progresses.

The technique was originally developed for samples in which the molecules are highly ordered, such as in thin "two-dimensional" crystals or helical fibers, for which specific grid preparation methods were designed (see Glaeser et al., 2007). Indeed, ordered aggregates have two advantages: (1) They provide identical environments for the individual molecules, ensuring structural stability and uniformity; and (2) the ordered arrangement lends itself to the application of convenient Fourier processing methods. Most importantly, the structural information in images of two-dimensional crystals is concentrated in spots lying on the reciprocal lattice, making the separation of signal and noise straightforward. Similarly, in helical arrangements, structural information is concentrated in layer lines from where it can be retrieved following well-established rules. In the following, however, we deal exclusively with the sample preparation for single-molecule imaging, predicated

https://doi.org/10.1017/CBO9781139003704.004 Published online by Cambridge University Press

on the assumption that the liquid sample contains large numbers of the same molecule in random orientations.

II.2. Specimen Preparation Protocol

A step-by-step protocol of the cryo-EM grid preparation has been given by Grassucci et al. (2007). At the start, a droplet of the liquid sample containing the molecules is applied to the grid. Excess liquid is blotted off using normal blotting paper, such that only a thin (ideally max. ~100 nm) liquid film remains. This step is critical, as increasing thickness leads to an increasing proportion of deleterious inelastic electron scattering and, eventually, to complete electron-opaqueness. On the other hand, a thickness close to, or smaller than, the size of the molecule will lead to the formation of damaged, partially "freeze-dried" samples. Critical parameters determining the thickness of the vitreous ice layer are temperature and humidity, and the length of time the sample is exposed to these conditions before the fast plunge. Because of the critical importance of these factors, manual operation is now increasingly replaced by use of computer-operated robots with a climate-controlled chamber (see Frederik and Storms, 2005).

Through the release of the rod (Figure 2.1), the EM grid is plunged into liquid ethane that is kept slightly above the temperature of liquid nitrogen (77.2° K) by means of a small heating coil. From that point on, the grid is kept at the liquid-nitrogen temperature, which serves to protect biological molecules from radiation damage (see Glaeser and Taylor, 1978). Essentially, the interpretation of the radiation protection is that the cleavage of bonds and fragmentation of the structure may still occur, but the low temperature "traps" or "cages" the fragments and prevents them from diffusing away. The radiation protection factor afforded by liquid nitrogen compared to room temperature is estimated as being in the range 5–10 (Glaeser et al., 2007). An additional factor may apply when going from the temperature of liquid nitrogen to that of liquid helium, depending on the nature of the specimen (crystalline or single-particle).

After the plunging, the grid is first stored in a storage box, from which it is transferred to a cryo-specimen holder or to a cartridge, depending on the type of electron microscope. In all these transfers, the grid is always kept at the temperature of liquid nitrogen. The specimen holder or cartridge is finally inserted in the EM. A step-by-step protocol is available in Grassucci et al. (2008), albeit specifically formulated for FEI (Portland, Oregon) instruments. This protocol, however, can be easily translated into a protocol for instruments of other manufacturers.

Imaging is done under low-dose conditions, again a measure designed to reduce radiation damage without compromising the image through excessive "shot noise," that is, noise due to the statistical variations of electron flux. "Low dose" is a somewhat diffuse term as used in the literature. Unwin and Henderson (1975), in their pioneering work on two-dimensional crystals of bacteriorhodopsin, used a dose of less than 1 $e^-/Å^2$, but at such a dose, molecules are virtually invisible. However, their sample was embedded in glucose and imaged at room temperature, without the radiation protection afforded by cooling to liquid nitrogen temperature, now routinely used in cryo-EM. Another factor is that higher electron voltages (200 to 300 kV versus 100 kV in Unwin and Henderson's experiment) lead to a reduced ratio of inelastic:elastic scattering (see below), hence reduced radiation damage. For these reasons, electron doses employed in the imaging of molecules are much higher now, in the range of 10 to 20 $e^-/Å^2$, enabling individual particles to be tracked down in the electron micrograph automatically by cross-correlation.

There has been a recent development in specimen preparation techniques that has the potential of revolutionizing cryo-EM imaging of molecular machines. If the objective is to image a macromolecular complex that has a particular ligand bound (for instance EF-G on the ribosome), the traditional way is to employ purification by affinity methods using a column prior to the preparation of the EM grid. The problem with this approach is that the ligand might come off in the process of the transfer, often leading to a heterogeneous mixture on the grid that must then be addressed by image classification (see Section IV.2 in this chapter). Kelly et al. (2008a; 2008b; 2010) developed a method of purification directly on the EM grid ("affinity grid"), ensuring that only molecules with the desired composition are recruited to the EM grid for imaging. The technique uses functionalized nickel-nitrilotriacetic acid (Ni-NTA) lipid monolayers deposited on the grid.

II.3. Imaging in the Electron Microscope

Electron micrographs are recorded either on film, for subsequent scanning in a microdensitometer, or by means of a charge-coupled device (CCD) camera for direct readout. Current CCD cameras that are affordable in price have a size of maximally 4,000 × 4,000 pixels, which gives a useful resolution of merely 2,000 × 2,000 independent pixels because of inter-pixel cross-talk – much less than the number of independent pixels on a film, estimated to be in the range of 6,000 × 10,000. Studies aiming at the highest spatial resolution are therefore still conducted by recording the TEM images on film (see, for instance, Seidelt et al., 2009). Electronic recording with comparable spatial resolution and high dynamic range awaits maturation and implementation of other technologies such as Complementary Metal-Oxide-Semiconductor (CMOS; see Faruqi, 2009).

Voltages being used are normally in the range of 200 to 300 kV. Voltages lower than 200 kV are avoided because of the increase in radiation damage, whereas voltages higher than 300 kV lead to diminishing contrast and inefficient electron recording. In the transmission electron

https://doi.org/10.1017/CBO9781139003704.004 Published online by Cambridge University Press

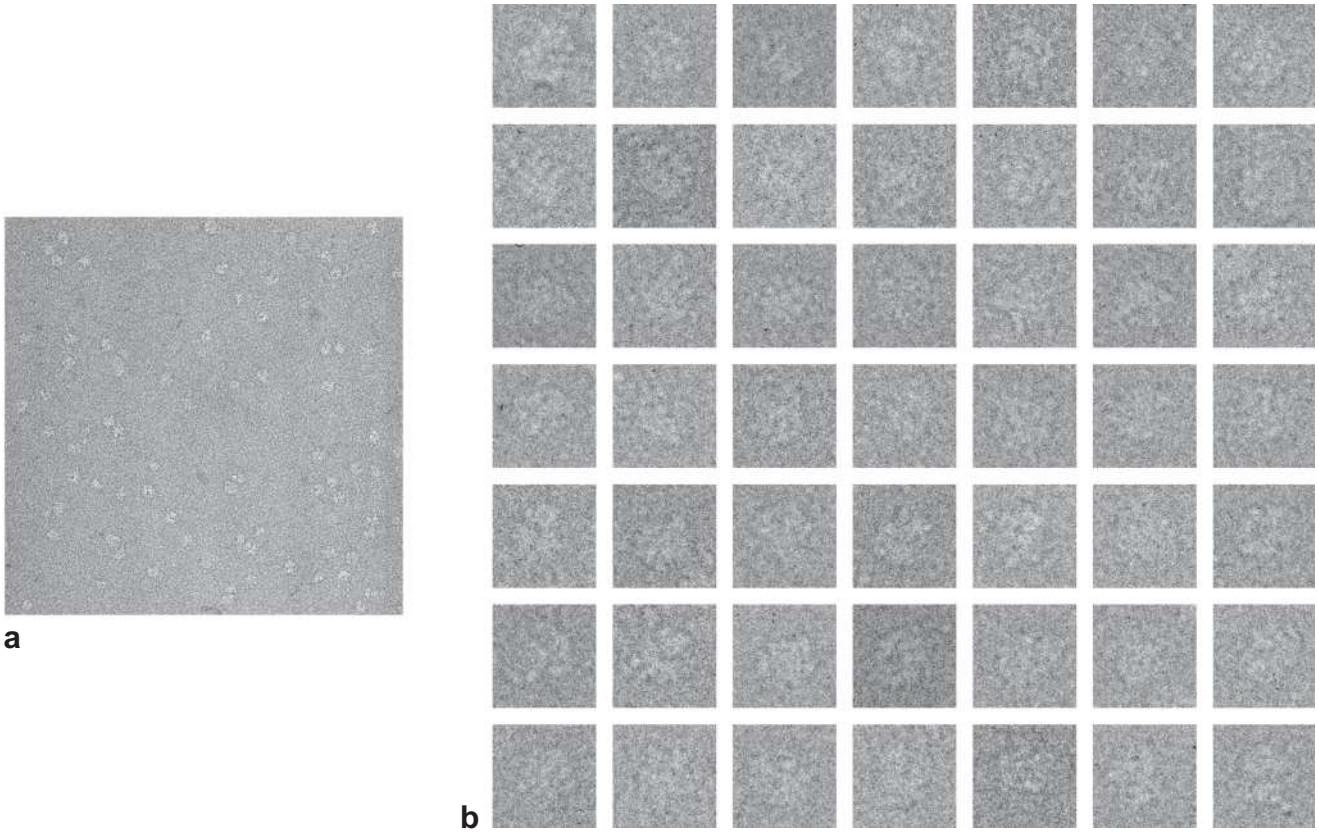


***FIGURE 2.3:** Raw data collected on film with the FEI F30 electron microscope at 300 kV and a magnification of 59,000. (a) Micrograph of a ribosome complex. (b) Gallery of selected images.*

microscope, image formation is based on the scattering interaction between the electron beam and the atoms of the sample. Two types of interaction occur: elastic and inelastic scattering. The former is without dissipation of energy, such that a coherent relationship is maintained between scattered and unscattered beam, which is the origin of productive, high-resolution image formation. The latter, in contrast, is accompanied by energy loss, leading to an incoherent relationship between scattered and unscattered beams and making no productive contribution to the image. Thus the preference for higher voltages – within the limits set by the diminishing contrast – is based on the fact that the ratio between (good) elastic and (bad) inelastic scattering increases with voltage. However, going into details of image formation exceeds the scope of this chapter, and the interested reader is referred to authoritative treatments in the works by Spence (2003), Glaeser et al. (2007), and Reimer and Kohl (2008). A brief introduction is also provided in Frank (2006b).

III. IMAGE PROCESSING

There are numerous steps of image processing that bring us from the raw data to the finished product, the three-dimensional image. Extensive software packages have been written with specialization on electron microscopy, among these EMAN (Ludtke et al., 1999), SPIDER (Frank et al., 1996), FREALIGN (Grigorieff, 2007), IMAGIC (van Heel et al., 1996), MRC suite (Crowther et al., 1996), and XMIPP (Sorzano et al., 2004). (The latest special issue focused on EM-related software tools is the January 2007 issue of the *Journal of Structural Biology*). There is an increasing trend toward the use of multiple platforms, mixing and matching the modules as they have different strengths and weaknesses in the different areas of methodology. For example, Appion (Lander et al., 2009) is a pipeline designed to extend existing software applications and procedures, and provides transparent interconversions between the different file formats and angle conventions used in the various packages.

III.1. Particle Picking

In the TEM micrograph, the molecule projections are visible as very noisy "blobs" with faint contrast (Figure 2.3a). Candidate images (Figure 2.3b) are extracted from the micrographs with the aid of automated search algorithms, which normally, however, still require verification by eye. The search typically involves two-dimensional cross-correlation in some form. The cross-correlation signal is

https://doi.org/10.1017/CBO9781139003704.004 Published online by Cambridge University Press

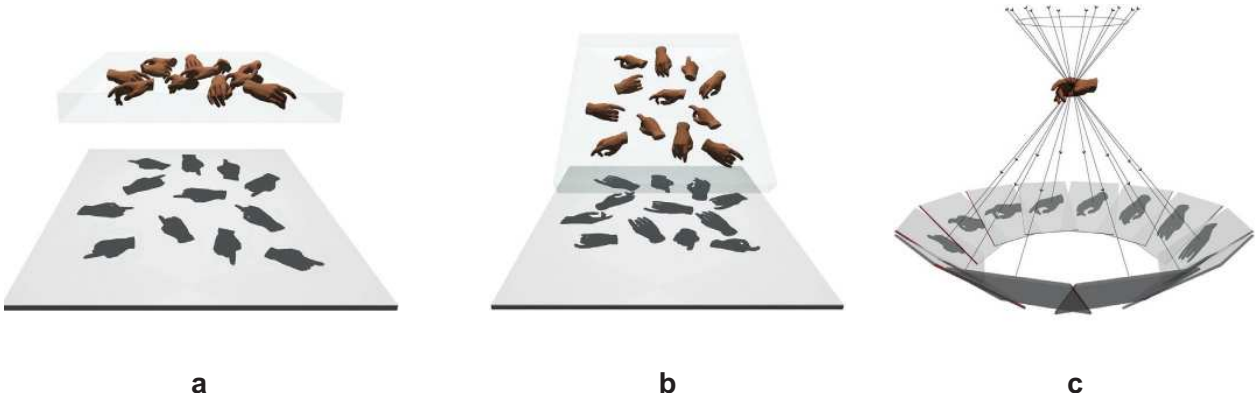


***FIGURE 2.4:** Principle of the random-conical data collection and reconstruction. A given 3D object (here: the hand) is assumed to lie in a defined orientation on the grid, varying only in its azimuthal in-plane position. (a) When the grid is untilted, the projections of the different copies of the object remain the same, except for their rotation within the image. (b) When the grid is tilted, many different 3D projection views are realized. (c) In the coordination system of the object, the 3D projection directions form a cone. The term "random-conical" refers to the fact that the azimuths of the object are normally random, generating random placements in the conical projection geometry. Reproduced from (Frank, 1998) with permission.*

proportional to the contrast and the size of the particle. As a rule of thumb, molecules with molecular mass above 400 kD are easily recognized and processed even in the absence of symmetries.

Increasingly sophisticated methods of automated "particle picking" have been developed to replace the tedious manual selection and verification (recent example: Voss et al., 2009; see also special January 2004 issue of the

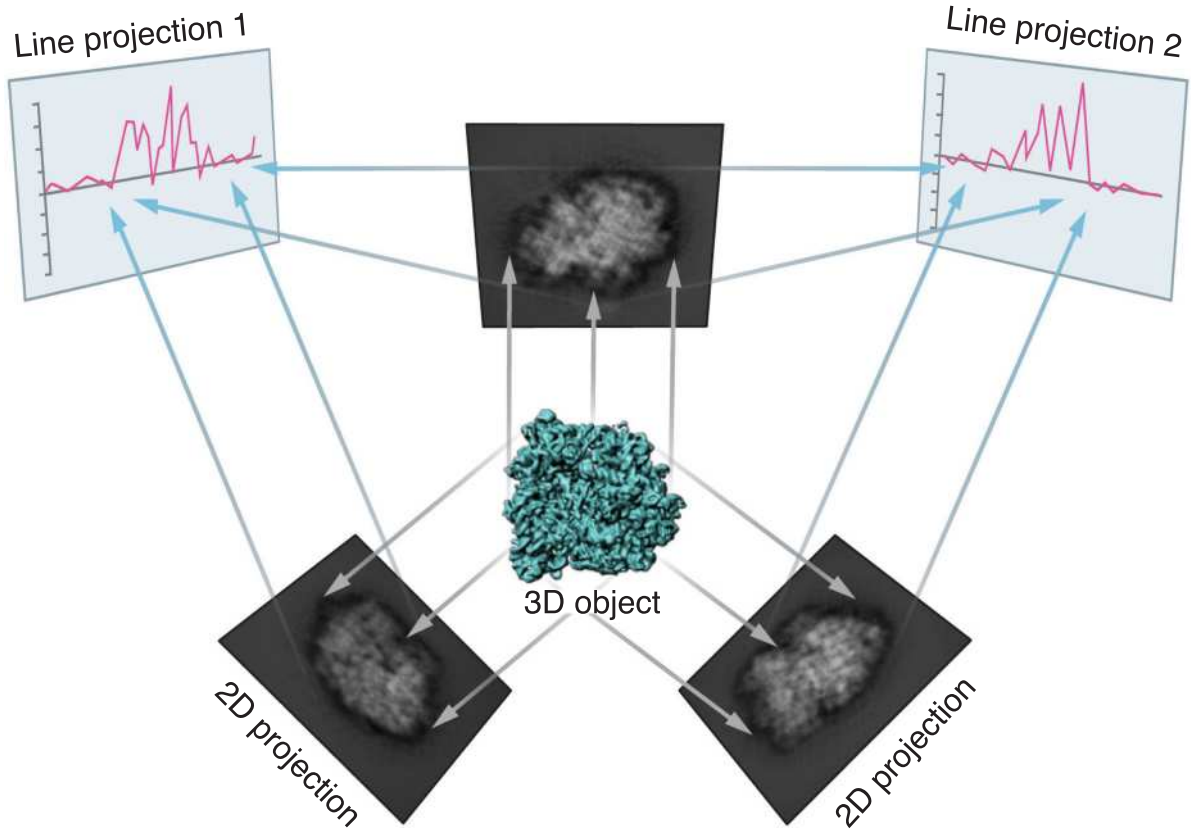
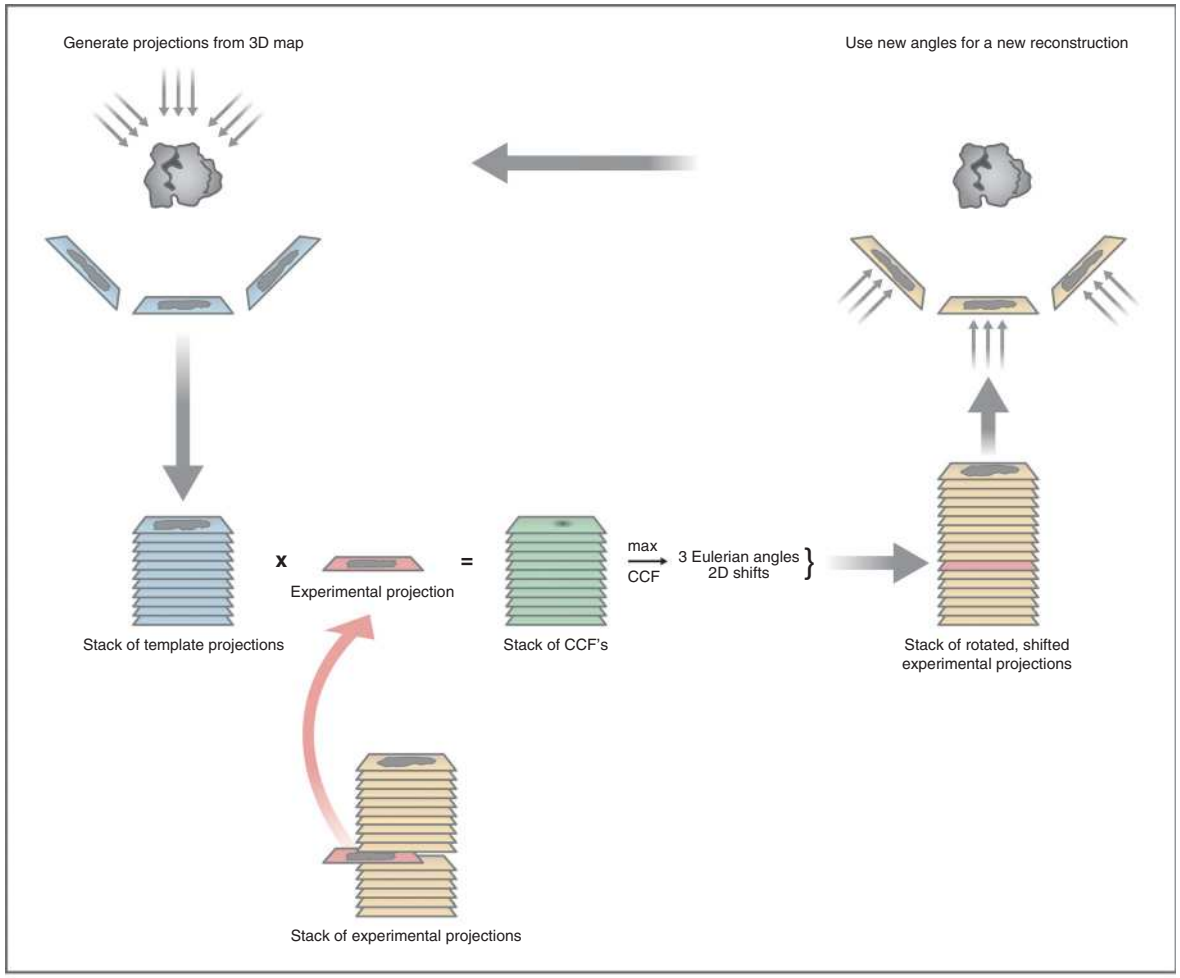


***FIGURE 2.5:** Principle of the angular reconstitution technique. Two different two-dimensional (2D) projections of the same 3D object always have a one-dimensional (1D) line projection in common. By determining these pairwise common 1D line projections for a set of three or more 2D projections, one is able to determine the relative orientations of all 2D projections in a common reference system (van Heel, 1987). Adapted from (van Heel et al., 2000).*

https://doi.org/10.1017/CBO9781139003704.004 Published online by Cambridge University Press



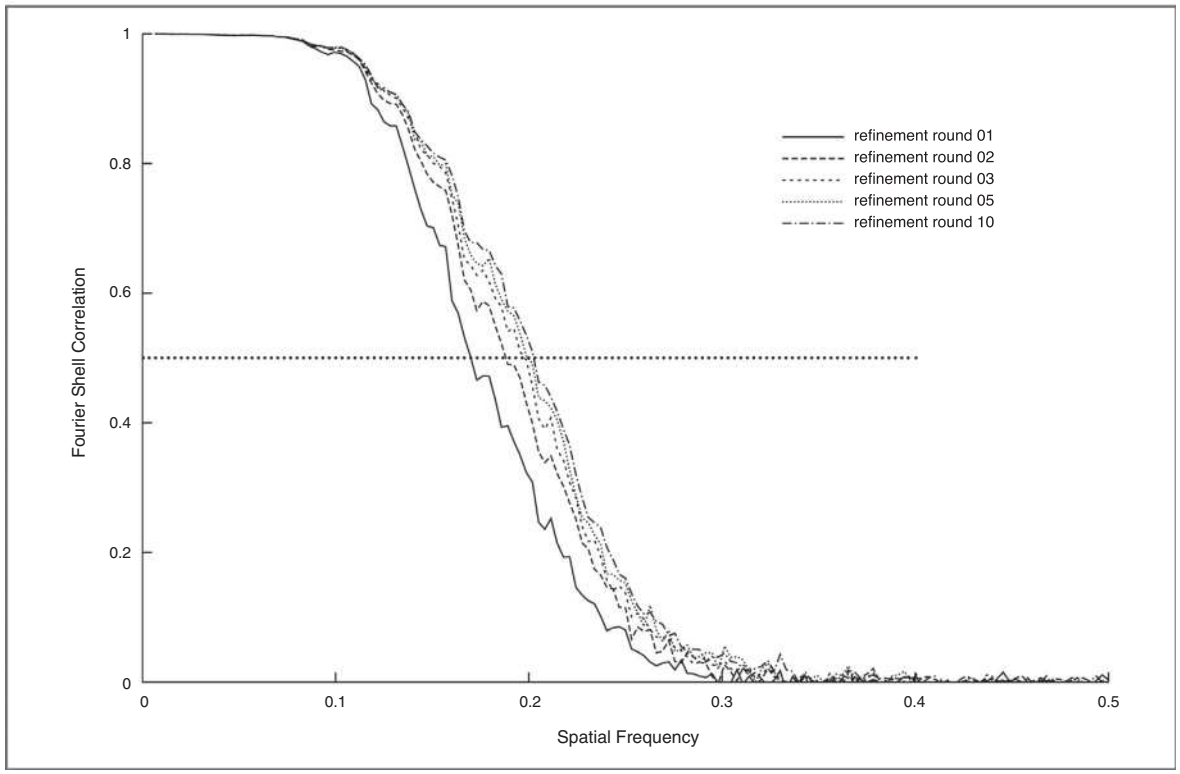


FIGURE 2.6

https://doi.org/10.1017/CBO9781139003704.004 Published online by Cambridge University Press

Journal of Structural Biology 145, Issues 1–2, devoted to this problem). Machine-learning algorithms are the latest addition to the tool set developed over time (Sorzano et al., 2009a; Langlois et al., 2011). The resulting fully verified data set may contain hundreds of thousands particles (i.e., projection images) ready to be processed (Figure 2.3b). There are two main reasons why such large numbers of images are needed: The low signal-to-noise ratio in the images – a consequence of the need to keep the electron dose low; and conformational heterogeneity, given that the minimum statistically required number of particles must be available for each conformation (see Section IV in this chapter).

III.2. Determination of Projection Angles for *Ab Initio* Reconstruction

Most of the work in single-particle reconstruction goes into the determination of projection angles, which are not *a priori* known. This problem can be compared to the indexing of X-ray diffraction patterns, a task that also places the data collected into a common three-dimensional framework. We distinguish *ab initio* methods from those that require a reference, or a preexisting model. When the structure is unknown, the project has to start with an *ab initio* reconstruction, then normally proceeds with a reference-based angular refinement using the *ab initio* reconstruction as reference. This will become clear in the following.

Two *ab initio* methods of reconstruction were developed in the 1980s: the random-conical reconstruction method (Figure 2.4; Radermacher et al., 1987; Radermacher, 1988) and the method of common lines (or "angular reconstitution") (Figure 2.5; van Heel, 1987; van Heel et al., 2000). The former uses a tilt pair to establish a coordinate system for a sub-population of the molecules that happen to lie in the same orientation on the grid, assigning a set of three Eulerian angles to each molecule. The latter *ab initio* technique uses intrinsic relationships among 2D projections of the same 3D structure – any two such projections have a central line in Fourier space in common, or, in an equivalent real-space formulation, they share a common 1D projection. For details on the mathematical principles and algorithms, the reader is referred to the original literature cited earlier. For a recent step-by-step protocol of the random-conical technique, see Shaikh et al. (2008).

A note on classification is in order here because both *ab initio* reconstruction techniques require classification of the experimental images. In the random-conical reconstruction, particles in the micrograph from the untilted grid have to be sorted into view classes, that is, groups of molecules presenting the same view. Random-conical reconstruction then proceeds separately for each view class. In the angular reconstruction technique, classification is required for a different reason: Here it is needed to boost the signal-to-noise ratio through the formation of class averages, because raw images are much too noisy to allow the comparison of common lines or 1D projections.

Classification of images is routinely done using a recipe developed by van Heel and Frank (1981): given a set of N windowed images, each containing M pixels. The images are first aligned to one another; that is, they are represented in a common two-dimensional reference system, then the whole array of $N \times M$ pixels is subjected to correspondence analysis or principal component analysis (van Heel and Frank, 1981; Lebart et al., 1984; Borland and van Heel, 1990), and some automated classification algorithms such as K-means or hierarchical ascendant classification are applied to the data represented in factor space (Frank, 1990; van Heel et al., 2000).

III.3. Three-Dimensional Reconstruction

Once the angles have been assigned by using either one of the *ab initio* methods, an initial coarse reconstruction is obtained. This is not the place to describe the methods of 3D reconstruction from projections, which has been covered in numerous articles and textbooks because of the wide range of applications; what makes single-particle reconstruction special in some way, affecting the design of algorithms, is that it is a true three-dimensional problem – unlike electron tomography with single-axis tilting – and that the angles are randomly distributed. Still, the best source specific to this subject are Michael

←

FIGURE 2.6: **(continued)** ***Reference-based angle assignment and angular refinement. (a) Flowchart. At the outset, a library of reference or template projections is created from a given 3D reference map (left) on an initially coarse angular grid (see Figure 2.7). Each of the experimental projections is compared with all template projections by cross-correlation (denoted by the symbol X), yielding a stack of cross-correlation functions (middle). Search for the highest cross-correlation peak in the stack yields the three Eulerian angles and x,y shifts that best describe the placement of the projection in the 3D reference system. A reconstruction is performed on the whole stack of experimental data to which these transformations have been applied (right). The resulting reconstruction is the first estimate of the structure. The left-facing arrow at the top shows how this scheme is extended to form a closed loop, termed angular refinement: The previous reconstruction is used as the 3D reference in the next cycle, and so on. In each cycle, the angular spacing, initially 15 degrees, is narrowed so that finer and finer orientation-related features are captured. (b) Resolution, determined by the Fourier shell correlation, as a function of angular refinement iteration. Resolution is defined by the spatial frequency at which the Fourier shell correlation drops below 0.5. It is seen that resolution, so defined, increases steadily with each cycle, but the progress slows down. Cycle 10 is close to convergence.***

https://doi.org/10.1017/CBO9781139003704.004 Published online by Cambridge University Press

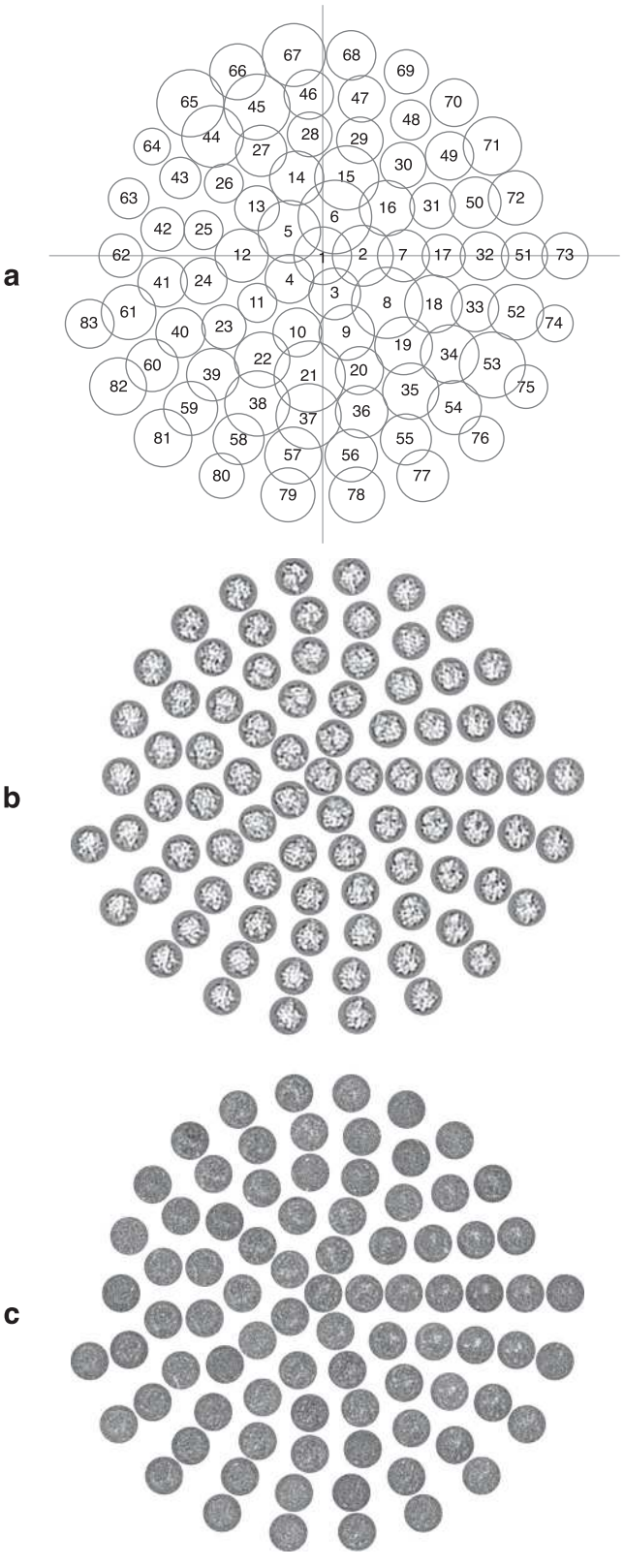


FIGURE 2.7: Reference-based angle assignment in practice, as illustrated by a ribosome project. The Eulerian angles are ordered on a coarse angular grid (83 samples; 15° spacing) on a half-sphere. (a) Orientation statistics: The area of each circle is made proportional to the number of projection images assigned to it. Although there are orientational preferences, reflected by regions with larger circles, there are no real gaps (i.e., regions with circles of vanishing sizes). (b) Orientation class averages. Class averages are pasted into their angular positions in diagram (a). (c) Orientation class variances. Here two-dimensional variance patterns associated with the computations of class averages are pasted into diagram (a).

Radermacher's articles featuring weighted back-projection as the primary approach (Radermacher, 1988; Radermacher, 1991), although new approaches such as a fast gridding-based algorithm for arbitrary geometry (Penczek et al., 2004) and wavelet-based methods (Sorzano et al., 2004) have lately gained popularity.

The initial reconstruction from either *ab initio* method is typically of low resolution and may exhibit artifacts. Angular refinement is a generic term for an iterative procedure that, starting with an initial coarse reconstruction, seeks a solution that is optimally consistent with the experimental data (Penczek et al., 1994; Sorzano et al., 2009a) (Figure 2.6a). To this end, the given reconstruction is used as a 3D reference to generate an indexed library of 2D reference projections. Each experimental image is now matched with each image of the whole library to determine the rotation and shift giving maximum cross-correlation. This search yields an assignment, to the experimental projection, of three Eulerian angles and a 2D shift associated with the best match. After all experimental images have been assigned new angles, a new reconstruction can be computed, as shown on the right in Figure 2.6a. This refinement cycle is now repeated multiple times, with progressively decreasing angular increment, until some criterion of convergence has been fulfilled. A prime criterion for convergence is the progress in *resolution*.

Resolution can be estimated as follows (see recent review by Liao and Frank, 2010). In single-particle reconstruction, the absence of repeats (and symmetries, apart from certain classes of specimens such as viruses) means that the extent of the region in Fourier space where signal is present is not readily apparent – there are no diffraction spots to go by. Instead, one measures the reproducibility, in Fourier space, of reconstructions obtained from randomly drawn half-sets of the data. In other words, one measures the extent of the (usually spherical) region in Fourier space where significant correlation can be detected between the transforms of two reconstructions. The *Fourier shell correlation* curve is a plot of correlation between Fourier coefficients of the two half-set reconstructions averaged on a shell in Fourier space, as a function of shell radius. Typically, the curve falls off from the value 1 at low spatial frequencies to a value close to 0. As the angular refinement proceeds, the falloff region of the curve shifts toward higher spatial frequency (Figure 2.6b) until saturation is achieved.

Another useful criterion in monitoring progress of refinement looks at the number of particles switching angle

https://doi.org/10.1017/CBO9781139003704.004 Published online by Cambridge University Press

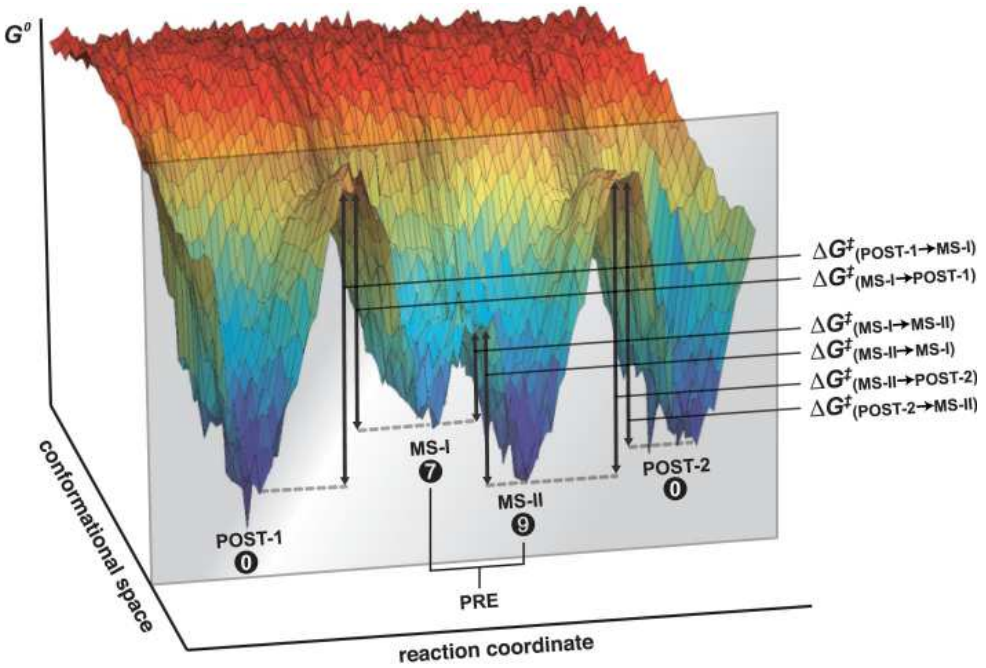


FIGURE 2.8: Schematic diagram of the free-energy landscape for the elongation cycle of translation. The free-energy values correspond to rate constants obtained by smFRET; for details see Frank and Gonzalez (2010). Reproduced from (Frank and Gonzalez) with permission by Annual Reviews.

assignment from one iteration to the next. This number typically stabilizes, indicating that the remaining particles have features that render them "indecisive," and that further progress cannot be expected.

IV. CRYO-EM OF HETEROGENEOUS SAMPLES AND DISCOVERY-BASED METHODS

IV.1. The Free-Energy Landscape and Inventory of States

The absence of constraints in the sample preparation of single molecules for cryo-EM implies that molecular machines can adopt all conformations and binding states compatible with the buffer conditions at the moment they are freeze-plunged. To attain the highest resolution, and to maximize the number of particles entering the reconstruction, one is interested in trapping as many molecules as possible in a single state, and indeed this was the strategy that was exclusively used in the beginning as the technique was being developed. However, as powerful classification methods have become available, opportunities for a different strategy have now opened up: to look at a sample of freely equilibrating molecules, with the idea to isolate subpopulations in all the states being present, for an exhaustive characterization of each of these subpopulations by a separate reconstruction (Connell et al., 2007; Fischer et al., 2010; Frank, 2010; Mulder et al., 2010). This approach may be termed discovery-based because the data themselves are rich in latent "scrambled" information, which can be unscrambled by a complex process of analysis.

For further explanation, particularly in view of recent developments, it is helpful to make reference to a schematic diagram of the free-energy landscape of a molecular machine (Frank and Gonzalez, 2010), in this case the Bacterial ribosome, as inferred from smFRET (Figure 2.8). Valleys in this free-energy landscape, which correspond to our colloquial notion of "states," are well-populated, and molecules belonging to these sub-populations can be characterized by separate 3D reconstructions from the various projection classes – provided they can be separated, and provided the number of projections in each class is sufficient in statistical terms to support a reconstruction.

Depending on the ambient temperature – that is, depending on the amount of energy supplied by the thermal environment – molecules may freely interconvert among the different states, but any transient or short-lived intermediate states are too poorly populated to be captured by cryo-EM. Information on these transitions among different states is available through a complementary technique, single-molecule FRET (smFRET). Here fluorophores (a donor-acceptor pair) are placed on components whose distance is expected to change dynamically, and smFRET signals report on their distance changes in real time. smFRET

https://doi.org/10.1017/CBO9781139003704.004 Published online by Cambridge University Press

has been applied to a variety of molecular machines, as described in separate chapters by Xinghua Shi and Takjep Ha (Chapter 1) and MacDougall and coworkers (Chapter 7). These smFRET results show, for instance, that the pretranslocational ribosome is constantly oscillating between two conformational states, termed macrostate I and II (or global state I and II in another terminology; see Fei et al., 2008) (Cornish et al., 2008), and possibly additional states intermediate between these conformations (Munro et al., 2007). As this example of the ribosome shows, the coexistence of multiple states giving rise to multiple conformations is a characteristic of molecular machines that are driven by Brownian motion.

Another example is the general transcription factor IID (TFIID), required for initiation of RNA polymerase II-dependent transcription at many Eukaryotic promoters. Here the binding by a protein (TATA-binding protein) induces a massive conformational change of the factor. Again, both binding states coexist in the sample, calling for a method to sort out the two states by analysis of the projection data (Elmlund et al., 2009).

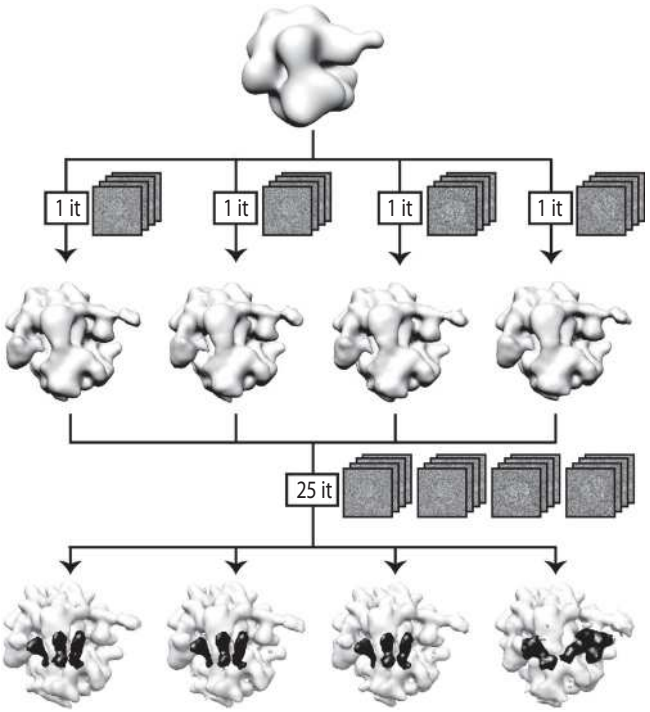


FIGURE 2.9: Maximum likelihood (ML3D) classification of single-particle projections. A low-resolution "consensus" density map (top) is used as initial seed. The low resolution aims to minimize model bias. The data set is split into K subsets if K is the number of classes we are looking for. In the ribosome example (Scheres et al., 2007), K = 4 was chosen. A single iteration of ML3D refinement for each subset yields four density maps with random differences (second row). These maps are then used in a multi-reference ML3D refinement of the entire dataset. Bottom row presents the structures obtained after twenty-five iterations. Three of the maps prove to be virtually identical, showing the tRNA in the A, P, and E positions, whereas the fourth shows a single tRNA bound at the E site, as well as EF-G bound at the A site. Reproduced from (Scheres, 2010) with permission by Elsevier.

IV.2. Classification: Disentangling Orientation and Conformation of the Molecules in the Projection Data

The coexistence of multiple states greatly complicates the reconstruction process in cryo-EM, because in addition to the projection *direction*, the subpopulation or *class* the projection belongs to must be determined. Given that each projection originates from a separate molecule, we are faced with an extremely difficult problem: Variations due to changes in orientation are intermingled with variations due to conformation or binding state.

Initial approaches to this problem of heterogeneity used *supervised classification*, that is, classification based on the similarity to two or more 3D references (see Valle et al., 2002). Sometimes, this scheme is hierarchically employed (Connell et al., 2007; Fischer et al., 2010). Though successful in many cases where the system is well characterized, the approach is unsatisfactory because it not only lacks generality, but also may miss conformations not anticipated, or come to incorrect conclusions if all of the references are dissimilar to the structure the experimental projections originated from. For the past few years, there has been an investment in the development of methods for *unsupervised classification*, which, by contrast, requires little or no prior information.

A position in between supervised and unsupervised methods holds those (Fu et al., 2007; Hall et al., 2007; Elad et al., 2008) that treat the conformational variability as a second-order problem, given that by far the largest variability in the data originates with the changes in orientation of the molecule. Hence, data are first sorted by orientation, then, in a second step, each orientation class is again classified into conformational classes. Finally, a third step is required, which interrelates conformational classes from different orientations.

Entirely unsupervised methods fall in the following categories: maximum-likelihood classification (ML3D; Scheres et al., 2005; 2007 – see Figure 2.9), the bootstrap method (Penczek et al., 1996; Zhang et al., 2008; Hstau and Frank, 2010 – see Figure 2.10), and various methods based on multiple common lines comparisons (Herman and Kalinowski, 2007; Singer et al., 2009; Elmlund et al., 2010; Shatsky et al., 2010).

An example for the recovery of multiple conformational states from a single data set is provided in Figure 2.11. Here maximum-likelihood classification was applied to a set of 216,000 images of a pretranslocational ribosome complex, and four different conformations were extracted, three of which show the ribosome in different states of translocation.

https://doi.org/10.1017/CBO9781139003704.004 Published online by Cambridge University Press

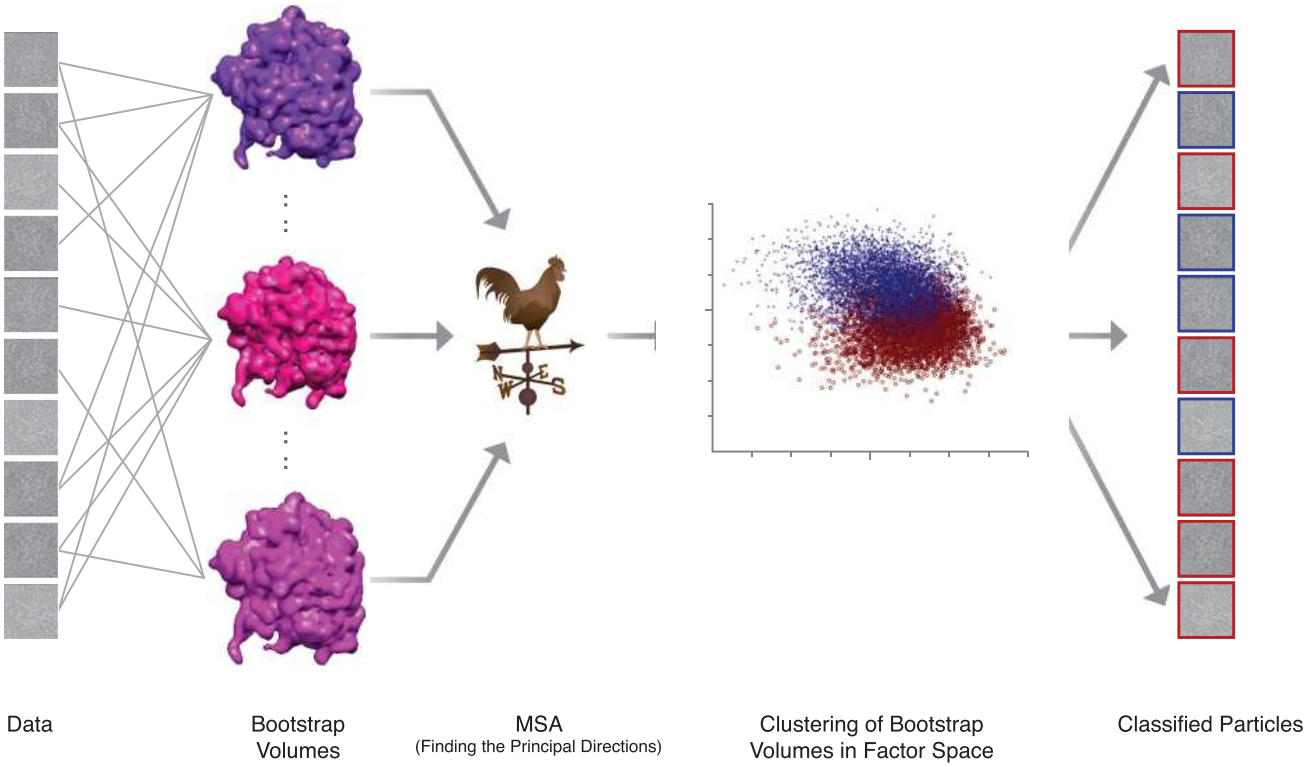


FIGURE 2.10: Flowchart of classification based on bootstrap reconstructions. From the projection set containing N particle images, M < N samples are drawn with replacement. After each draw, a "bootstrap" reconstruction is computed. The resulting M reconstructions are analyzed by multivariate statistical analysis, resulting in a factorial representation of the bootstrap volumes in which K-means classification is performed and classes of bootstrap volumes are identified. For each particle image, its class association is finally determined.

IV.3. Time-Resolved Cryo-EM Imaging

As we have seen, the classification methods described earlier open new avenues of research by which multi-state mixtures can be studied. This technology offers the exciting opportunity to look at systems as they evolve over time; in other words, time-resolved three-dimensional imaging of molecular machines evolving from a non-equilibrium state. For sufficiently slow processes, such as virus capsid maturation (Heymann et al., 2003), ribosome biogenesis (Mulder et al., 2010), and back-translocation in translation (Fischer et al., 2010), the time for blotting and freeze-plunging the grid (in the range of several seconds) is negligible compared with the duration of the entire process, so time resolution is easily achievable by taking aliquots of the sample in regular intervals after the starting reaction for EM imaging. On the other hand, achieving time resolution on the order of a few milliseconds, such as required for studying decoding in translation (see Rodnina et al., 2002; Frank and Gonzalez, 2010), requires a radical departure from the established sample-preparation methods. Pioneering time-resolved techniques developed by Berriman and Unwin (1994) used the principle of spraying one reactant (in this instance, acetylcholine) onto a grid covered with the second reactant (acetylcholine receptor), so that the mixing and reaction occurred on impact of the spray droplets on the target. As discussed in Lu et al. (2009), however, the unsharp definition of reaction time and possible interference of the grid with the target molecule are among the drawbacks of this method when aiming at millisecond resolution.

Specifically, the requirements may be formulated in the following way: First of all, a "zero" time point needs to be established, with little margin of error, denoting the point when the components are brought together. This calls for the design of an efficient mixer. Next, a provision needs to be made for the reaction in the fully mixed sample to take place for a defined time and under controlled conditions, in a suitable reaction chamber. Third, the reacted sample must be rapidly sprayed onto the grid, and the grid be plunged into the cryogen with a minimum of delay. The monolithic microfluidic mixing-spraying device developed by Lu et al. (2009), operated in conjunction with a traditional gravity-operated freeze-plunger, addresses these requirements successfully. Among the first applications of this novel device has been the dynamics of the association of ribosomal subunits into the 70S ribosome (Barnard et al., 2009; Shaikh et al., 2010).

https://doi.org/10.1017/CBO9781139003704.004 Published online by Cambridge University Press

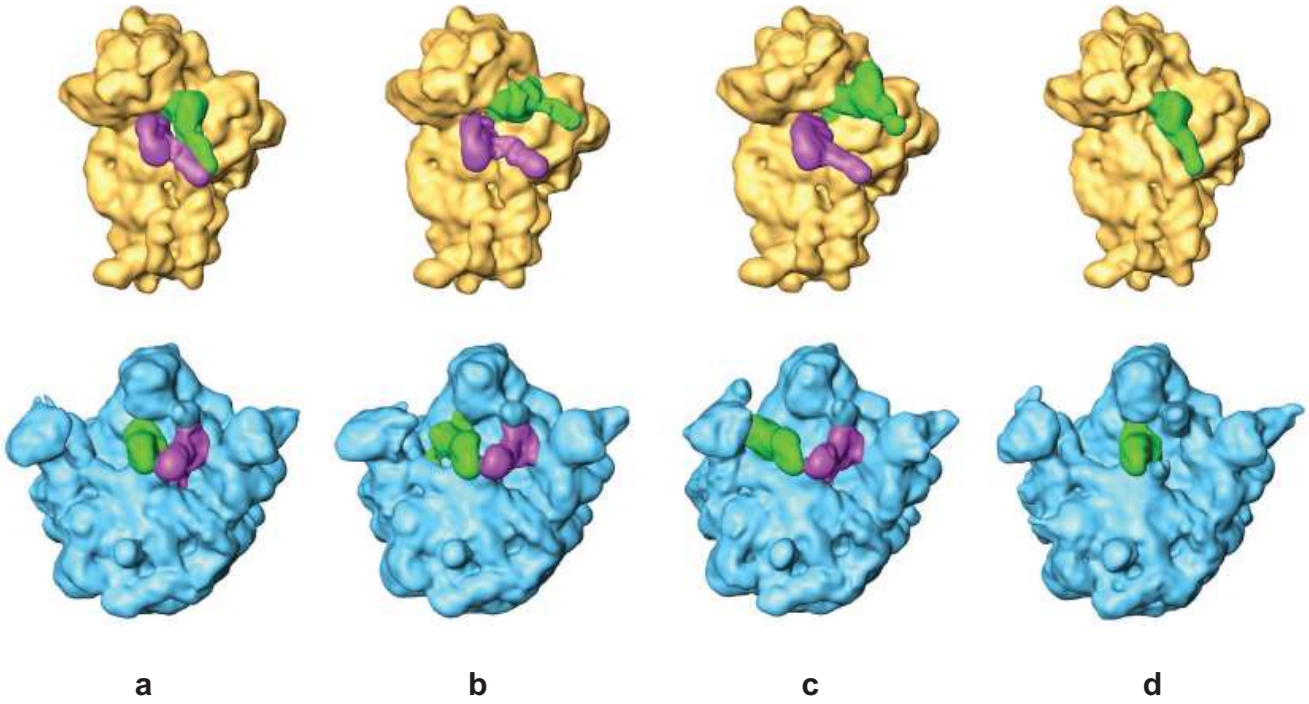


FIGURE 2.11: Example for classification of a heterogeneous data set using maximum likelihood classification (ML3D). Specimen was a pretranslocational ribosome complex in the absence of EF-G described in Agirrezabala et al. (2008). For display of the intersubunit space, the 70S ribosome was computationally separated into its two subunits, 30S (yellow) and 50S (blue). Supervised classification with two references identified two classes, distinguished by tRNA positions and presence versus absence of intersubunit rotation. Number of classes specified was six. Of these, only five were found to be meaningful; the sixth class appeared to have acted as a "garbage bin" for misfits. Of the five remaining classes, one (panel d) lacks the A-site tRNA and hence is not a pretranslocational complex. Of the remaining four, two were found to be closely similar. We have therefore the result that three highly populated subpopulations (panels a, b, and c) exist, in which the positions of the tRNAs and ribosome conformation are well defined. Reproduced from (Frank, 2010, with permission by Wiley-VCH).

IV.4. High-Throughput Data Collection and Processing

Given the fact that tens of thousands of images are required to obtain a sub-nanometer resolution reconstruction of a molecule in a single state, the diversity of coexisting states anticipated in the study of molecular machines means that the data collection and processing has to be streamlined and organized much more efficiently than currently practiced. In the quest for solutions to this problem, pioneering contributions have been made by the groups of Bridget Carragher and Clint Potter at The Scripps Institute, centered around two software tools, Leginon (Suloway et al., 2005) and Appion (Lander et al., 2009), both linked to the same database. Whereas Leginon is designed to collect data on the electron microscope automatically, Appion is a modular Python-based pipeline with interoperability with all major image-processing packages. A demonstration of efficient data collection and processing using these tools has been done by Stagg et al. (2006), who collected more than 200,000 images of GroEl and obtained a sub-nanometer reconstruction of this molecular machine in a span of a day.

Efficient organization of data storage and accelerated means of processing are obvious concerns in this context that are addressed by a variety of recent developments. Graphics Processor (GPU) implementations are now being developed to accelerate the execution of key algorithms such as projection matching significantly. All major EM-specific image-processing packages routinely make use of parallel processing, as many operations on stacks of images are parallel by nature. These new developments, however, are rapid and quite hardware-specific, precluding explicit coverage in this chapter.

V. OBTAINING AN ATOMIC MODEL

The evident goal of cryo-EM applied to a molecular machine is to obtain atomic models that show the machine in different processing states. Only in rare cases, in the presence of symmetries, cryo-EM density maps are of sufficient resolution to allow *de novo* tracing of the chains (Zhou, 2008; Lindert et al., 2009); such atomic models have been obtained for viruses (Yu et al., 2008) and for GroEl (Ludtke et al., 2008). The more typical situation is that atomic structures are available for the molecule and its functional ligands, and that these are used as the "raw material" for the

https://doi.org/10.1017/CBO9781139003704.004 Published online by Cambridge University Press

composition of a comprehensive atomic model showing the conformation of the molecule and the binding interactions of its ligands. Alternatively, if the structure of the whole molecule has not been solved, at least structures of some components may be available.

A number of techniques of increasing sophistication have been developed over the years, parallel to the increase in resolution and demands for increasing fidelity of the fitting. Starting with the most straightforward approach, available structures are fitted as rigid bodies, essentially as an annotation of the (low-resolution) density map that shows putative or confirmed placements of relevant parts, giving an idea which parts of the cryo-EM density map are accounted for (see the example of the spliceosome; Stark and Lührmann, 2006). However, rigid-body docking will not do justice to the observed density map if conformational changes are present, which become noticeable with increasing resolution. Again the simplest solution is piece-wise rigid-body fitting: to break the atomic structure apart at putative flex points and place the individual pieces by eye into the EM map, to serve as an illustration for the kinds of local movements that must be invoked to bring the structure into agreement with the density. Quantitative techniques incorporating this idea, in the next advance of methodology, are SITUS (Wriggers and Chacón, 2001) and real-space refinement (RSREF; Chapman, 1995): The aim of the algorithms incorporated in these packages is to readjust the placement of pieces optimally within the local density by cross-correlation and mend their connections such that the structure is again complete, without violations of the rules of stereochemistry. An example of piecewise flexible fitting using real-space refinement applied to the ribosome is the work by Gao et al. (2003; Gao and Frank, 2005). However, true *flexible fitting methods* go one step further and deform the structure in its entirety. One approach is normal-mode flexible fitting (NMFF; Tama et al. 2004), in which the structure is approximated by an elastic network of pseudoatoms and subjected to a normal mode analysis. Another approach, closer to the atomic nature of the structures to be fitted, is molecular dynamics flexible fitting (MDFF; Trabuco et al., 2008; see application in Villa et al., 2009; Grubisic et al., 2010).

VI. CONCLUSIONS

Cryo-EM, making it possible to visualize molecules under native conditions, is one of the most important tools for the investigation of molecular machines. The field, as I have demonstrated here, currently experiences a widening of perspective, or a shift of paradigm: Instead of focusing on a single state, trapped by some kind of intervention, we now have the capability to make an entire inventory of all those conformational states that are well populated under the conditions of an experiment, and collect data that can be used, in effect, to reconstruct the free-energy landscape of the machine.

This capability invites a comparison with those of other biophysical visualization methods, nuclear magnetic resonance (NMR) and X-ray crystallography. A wide range of techniques in NMR spectroscopy allows characterization of conformational heterogeneity of macromolecules, depending critically on the timescales of inter-conversion between conformations. However, atomic resolution structure determination usually is limited to systems with a small ($\leq$3) number of states (Vallurupalli et al., 2008) or to fitting of ensembles of conformations consistent with experimental data (Lange et al., 2008). Efficient transverse relaxation and spectral congestion normally limit the application of NMR spectroscopy to large molecular machines; however, TROSY techniques have enabled investigations of conformational dynamics in GroEL/GroES (Horst et al., 2005), SecA (Keramisanou et al., 2006), proteosome (Religa et al., 2010), and ribosome (Cabrita et al., 2009).

In X-ray crystallography, the state in which the molecule is visualized is defined by the energetics of crystal formation; under normal circumstances, the molecule exists in a single conformation. In some exceptional cases, two or more copies of the molecules may coexist in the unit cell in different conformations. For instance, Schuwirth et al. (2005) obtained the structure of the *E. coli* ribosome in two conformations, related by a change in small subunit head orientation, which is known to be associated with the translocation process. The same group also found ribosomes paired up in the unit cell, which were in different states of inter-subunit rotation (Zhang et al., 2009).

As a relatively new technique, cryo-EM still lacks uniform standards, data formats, and procedures. As is usually the case, such standards develop in the free marketplace of ideas. In this sense, its present state can be compared with the beginning years of protein X-ray crystallography. However, efforts on both sides of the Atlantic should be mentioned, which strive toward standardization in depositions of maps and an agreement on a common dictionary of terms (Berman et al., 2006; Lawson, 2010; Lawson et al., 2011; Velankar et al., 2010).

ACKNOWLEDGMENTS

This work was funded by HHMI and NIH R01 GM29169. I would like to thank Helen Saibil and Ruben Gonzalez for a critical reading and helpful comments, Bertile Halle for a discussion of possible freeze-plunge artifacts, and Art Palmer for input on the capabilities of NMR in dealing with heterogeneous samples. I further thank Lila Iino-Rubenstein and Melissa Thomas for assistance with the illustrations, and Jesper Pallesen and Hstau Liao for preparing material used in some of the figures.

REFERENCES

Agirrezabala, X. and Frank, J. (2009). Elongation in translation as a dynamic interaction among the ribosome, tRNA, and elongation factors EF-G and EF-Tu. *Q. Rev. Biophys*. 42, 139–158.

Agirrezabala, X. and Frank, J. (2010). From DNA to proteins via the ribosome: structural insights into the workings of the translational machinery. *Human Genomics* 4, 226–237.

Agirrezabala, X., Lei, J., Brunelle, J. L., Ortiz-Meoz, R. F., Green, R., and Frank, J. (2008). Visualization of the hybrid state of tRNA binding promoted by spontaneous ratcheting of the ribosome. *Mol. Cell*. 32, 190–197.

Al-Amoudi, A., Norlen, L. P., and Dubochet, J. (2004). Cryo-electron microscopy of vitreous sections of native biological cells and tissues. *J. Struct. Biol*. 148, 131–135.

Barnard, D., Lu, Z., Shaikh, T. R., Yassin, A., Mohamed, H., Agrawal, R., Lu, T.-M., and Wagenknecht, T. (2009). Time resolved cryo-electron microscopy of ribosome assembly using microfluidic mixing. *Microsc. Microan*. 15, 942–943.

Berman, H. M., Burley, S. K., Chiu, W., Sali, A., Adzhubei, A., Bourne, P. E., Bryant, S. H., Dunbrack, R. L., Jr., Fidelis, K., Frank, J., Adam Godzik, A., Henrick, K., Joachimiak, A., Heymann, B., Jones, D., Markley, J. L., Moult, J., Montelione, G. T., Orengo, C., Rossmann, M. G., Rost, B., Saibil, H., Schwede, T., Standley, D. M., and Westbrook, J. D. (2006). Outcome of a workshop on archiving structural models of biological macromolecules. *Structure* 14, 1211–1217.

Berriman, J. and Unwin, N. (1994). Analysis of transient structures by cryo-microscopy combined with rapid mixing of spray droplets. *Ultramicroscopy* 56, 241–252.

Boehr, D. D., McElheny, D., Dyson, H. J., and Wright, P. E. (2006). The dynamic energy landscape of dihydrofolate reductase catalysis. *Science* 313, 1638–1641.

Borland, L. and van Heel, M. (1990). Classification of image data in conjugate representation spaces. *J. Opt. Soc. Am*. A7, 601–610.

Brueggeler, P. and Mayer, E. (1980) Complete vitrification in pure liquid water and dilute aqueous solutions. *Nature* 288, 569–571.

Cabrita, L. D., Hsu, S. T., Launay, H., Dobson, C. M., and Christodoulou, J. (2009). Probing ribosome-nascent chain complexes produced in vivo by NMR spectroscopy. *Proc. Natl. Acad. Sci. USA* 106, 22239–22244.

Chapman, M. S. (1995). Restrained real-space macromolecular atomic refinement using a new resolution-dependent electron density function. *Acta Crystallogr*. A51, 69–80.

Connell, S. R., Takemoto, C., Wilson, D. N., Wang, H., Murayama, K., Terada, T., Shirouzu, M., Rost, M., Schüler, M., Giesebrecht, J., Dabrowski, M., Mielke, T., Fucini, P., Yokoyama, S., and Spahn, C. M. T. (2007). Structural basis for interaction of the ribosome with the switch regions of GTP-bound elongation factors. *Mol. Cell* 25, 751–764.

Cornish, P. V., Ermolenko, D. N., Noller, H. F., and Ha, T. (2008). Spontaneous intersubunit rotation in single ribosomes. *Mol Cell*. 30, 578–588.

Crowther, R. A., Henderson, R., and Smith, J. M. (1996). MRC image processing programs. *J. Struct. Biol*. 116, 9–16.

DoG Picker and Tilt Picker: software tools to facilitate particle selection in single particle electron microscopy. *J. Struct. Biol*. 166, 205–313.

Dubochet, J., Adrian, M., Chang, 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.

Dubochet, J., Lepault, J., Freeman, Z. R., Berriman, J. A., and Homo, J.-C. (1982). Electron microscopy of frozen water and aqueous solutions. *J. Microsc*. 128, 219–237.

Elad, N., Clare, D. K., Saibil H. R., and Orlova, E. V. (2008). Detection and separation of heterogeneity in molecular complexes by statistical analysis of their two-dimensional projections. *J. Struct. Biol*. 162, 108–120.

Elmlund, H., Baraznenok, V., Linder, T., Szilagyi, Z., Rofougaran, R., Hofer, A., Hebert, H., Lindahl, M., and Gustafsson, C. M. (2009). Cryo-EM reveals promoter DNA binding and conformational flexibility of the general transcription factor TFIID. *Structure* 17, 1442–1452.

Elmlund, D., Davis, R., and Elmlund, H. (2010). Ab initio structure determination from electron microscopic images of single molecules coexisting in different functional states. *Structure* 18, 777–786.

Faruqi, A. R. (2009). Principles and prospects of direct high-resolution electron image acquisition with CMOS detectors at low energies. *J. Phys.: Condens. Matter* 21, 314004 (9pp) doi:10.1088/0953–8984/21/31/314004.

Fei, J., Kosuri, P., MacDougall, D. D., and Gonzalez, R. L. Jr. (2008). Coupling of ribosomal L1 stalk and tRNA dynamics during translation elongation. *Mol. Cell* 30, 348–359.

Fischer, N., Konevega, A. L., Wintermeyer, W., Rodnina, M. V., and Stark, H. (2010) Ribosome dynamics and tRNA movement by time-resolved electron cryomicroscopy. *Nature* 466, 329–333.

Frank, J. (1990). Classification of macromolecular assemblies studied as single particles. *Quart. Rev. Biophys*. 23, 281–329.

Frank, J. (1998). How the ribosome works. *American Scientist* 86, 428–439.

Frank, J. (ed.) (2006a). *Electron tomography – methods for three-dimensional visualization of structures in the cell*. New York: Springer Verlag.

Frank, J. (2006b). *Three-dimensional electron microscopy of macromolecular assemblies – visualization of biological molecules in their native state*. New York: Oxford University Press.

Frank, J. (2009). Single-particle reconstruction of biological macromolecules in electron microscopy – 30 years. *Quart. Rev. Biophys*. 42, 139–158.

Frank, J. (2010). The ribosome comes alive. *Israeli J. Chem*. 50, 95–98.

Frank, J. and Gonzalez, R. (2010). Structure and dynamics of a processing Brownian motor: the translating ribosome. *Ann. Rev. Biochem*. 79, 381–412.

Frank, J., Radermacher, M., Penczek, P., Zhu, J., Li, Y., Ladjadj, M., and Leith, A. (1996). SPIDER and WEB: processing and visualization of images in 3D electron microscopy and related fields. *J. Struct. Biol*. 116, 190–199.

Frederik, P. M. and Storms, M. H. (2005). Automated robotic preparation of vitrified samples for 2D and 3D cryo electron microscopy. *Microsc. Today* 13, 32–36.

Fu, J., Gao, H., and Frank, J. (2007). Unsupervised classification of single particles by cluster tracking in multi-dimensional space. *J. Struct. Biol*. 157, 226–239.

Fu, J., Kennedy, D., Munro, J.B., Lei, J., Blanchard, S. C., and Frank, J. (2009). The P-site tRNA reaches the P/E position through intermediate positions. (Abstract) *J. Biomol. Struct. Dyn*. 26, 794–795.

Gao, H. and Frank, J. (2005). Molding atomic structures into intermediate-resolution cryo-EM density maps of ribosomal complexes using real-space refinement. *Structure* 13, 401–406.

Gao, H., Sengupta, J., Valle, M., Korostelev, A., Eswar, N., Stagg, S. M., Van Roey, P., Agrawal, R. K., Harvey, S. C., Sali, A., Chapman, M. S., and Frank, J. (2003). Study of the structural dynamics of the *E. coli* 70S ribosome using real space refinement. *Cell* 113, 789–801.

Glaeser, R. M., Downing, K. H., DeRosier, D., Chiu, W., and Frank, J. (2007). *Electron crystallography of biological macromolecules*. New York: Oxford University Press.

Glaeser, R. M. and Taylor, K. A. (1978). Radiation damage relative to transmission electron microscopy of biological specimens at low temperature: review. *J. Microsc*. 112, 127–138.

Grassucci, R. A., Taylor, D. J., and Frank, J. (2007). Preparation of macromolecular complexes for cryo-electron microscopy. *Nat. Protoc*. 2, 3239–3246.

Grassucci, R.A., Taylor, D., and Frank, J. (2008). Visualization of macromolecular complexes using cryo-electron microscopy with FEI Tecnai transmission electron microscopes. *Nat. Protoc*. 3, 330–339.

Grigorieff, N. (2007). FREALIGN: High-resolution refinement of single particle structures. *J. Struct. Biol*. 157, 117–125.

Grubisic, I., Shokhirev, M. N., Orzechowski, M., Miyashita, O., and Tama, F. (2010). Biased coarse-grained molecular dynamics simulation approach for flexible fitting of X-ray structure into cryo electron microscopy maps. *J. Struct. Biol*. 169, 95–105.

Hall, R. J., Siridechadilok, B., Nogales, E. (2007). Cross-correlation of common lines: a novel approach for single-particle reconstruction of a structure containing a flexible domain. *J. Struct. Biol*. 159, 474–482.

Halle, B. (2004). Biomolecular cryocrystallography: structural changes during flash-cooling. *Proc. Natl. Acad. Sci. USA* 101, 4793–4798.

Hartl, F. U. and Hayer-Hartl, M. (2009). Converging concepts of protein folding *in vitro* and in vivo. Nat. *Struct. Mol. Biol*. 16, 574–581.

Herman, G. T. and Kalinowski, M. (2007). Classification of heterogeneous electron microscopic projections into homogeneous subsets. *Ultramicroscopy* 108, 327–338.

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.

Heymann, J. B., Conway, J. F., and Steven, A. C. (2004). Molecular dynamics of protein complexes from four-dimensional cryo-electron microscopy. *J. Struct. Biol*. 147, 291–301.

Horst, R., Bertelsen, E. B., Fiaux, J., Wider, G., Horwich, A. L., and Wüthrich, K. (2005). Direct NMR observation of a substrate protein bound to the chaperonin GroEL. *Proc. Natl. Acad. Sci. USA*. 102, 12748–12753.

Hsieh, C.-E., Marko, M., Frank, J., and Mannella, C. A. (2002). Electron tomographic analysis of frozen-hydrated tissue sections. *J. Struct. Biol*. 138, 63–73.

Kelly, D. F., Abeyrathne, P. D., Dukovski, D., and Walz, T. (2008a). The affinity grid: a prefabricated EM grid for monolayer purification. *J. Mol. Biol*. 382, 423–433.

Kelly, D. F., Dukovski, D., and Walz, T. (2008b). Monolayer purification: a rapid method for isolating protein complexes for single-particle electron microscopy. *Proc. Natl. Acad. Sci. USA* 105, 4703–4708.

Kelly, D. F., Dukovski, D. and Walz, T. (2010). Strategy for the use of affinity grids to prepare non-His-tagged macromolecular complexes for single-particle electron microscopy. *J. Mol. Biol*. 400, 675–681.

Keramisanou, D., Biris, N., Gelis, I., Sianidis, G., Karamanou, S., Economou, A., and Kalodimos, C. G. (2006). Disorder-order folding transitions underlie catalysis in the helicase motor of SecA. *Nat. Struct. Mol. Biol*. 13, 594–602.

Koster, A. J. and Barcena, M. (2006). Cryotomography: low-dose automated tomography of frozen-hydrated specimens. In: *Electron tomography – methods for three-dimensional visualization of structures in the cell* (pp. 113–161), ed. J. Frank. New York: Springer.

Lander, G. C., Stagg, S. M., Voss, N. R., Cheng, A., Fellmann, D., Pulokas, J., Yoshioka, C., Irving, C., Mulder, A., Lau, P.-W., Lyumkis, D., Potter, C. S., and Carragher, B. (2009). Appion: an integrated, database-driven pipeline to facilitate EM image processing. *J. Struct. Biol*. 166, 95–102.

Lange, O. F., Lakomek, N. A., Farès, C., Schröder, G. F., Walter, K. F., Becker, S., Meiler, J., Grubmüller, H., Griesinger, C., and de Groot, B. L. (2008). Recognition dynamics up to microseconds revealed from an RDC-derived ubiquitin ensemble in solution. *Science* 320, 1471–1475.

Langlois, R., Pallesen, J., and Frank, J. (2011). Reference-free segmentation enhanced with data-driven template matching for particle selection in cryo-electron microscopy. *J. Struct. Biol*., in press.

Lawson, C. L., Baker, M. L., Best, C., Bi, C., Dougherty, M., Feng, P., van Ginkel, G., Devkota, B., Lagerstedt, I., Ludtke, S., Newman, R. H., Oldfield, T. J., Rees, I., Sahni, G., Sala, R., Velankar, S., Warren, J., Westbrook, J. D., Henrisck, K., Kleywegt, G. J., Berman, H. M, and Chiu, W. (2011). EMDataBank.*org: unified data resource for cryoEM*. Nucl. Acid. Res. (Database issue), 34, 287–290.

Lawson, C. L. (2010). Unified data resource for Cryo-EM. *Methods Enzymol*. 483, 73–90.

Lebart, L., Maurineau, A., and Warwick, K. M. (1984). *Multivariate descriptive statistical analysis*. New York: John Wiley.

Liao, H. and Frank, J. (2010). Definition and estimation of resolution in single-particle reconstructions. *Structure* 18, 768–775.

Lindert, S., Stewart, P. L., and Meiler, J. (2009). Hybrid approaches: applying computational methods in cryo-electron microscopy. *Curr. Opin. Struct. Biol*. 19, 218–225.

Lu, Z., Shaikh, T. R., Barnard, D., Meng, X., Mohamed, H., Yassin, A., Mannella, C. A., Agrawal, R. K., Lu, T.-M., and Wagenknecht, T. (2009). Monolithic microfluidic mixing-spraying devices for time-resolved cryo-electron microscopy. *J. Struct. Biol*. 168, 388–395.

Ludtke, S. J., Baker, M .L., Chen, D.-H., Song, J.-L., Chuang, D. T., and Wah Chiu, W. (2008). *De novo* backbone trace of GroEL from single particle electron cryomicroscopy. *Structure* 16, 441–448.

Ludtke, S. J., Baldwin, P. R., and Chiu, W. (1999). EMAN: semi-automated software for high-resolution single-particle reconstructions. *J. Struct. Biol*. 128, 82–97.

Mulder, A. M., Yoshioka, C., Beck, A. H., Bunner, A. E., Milligan, R. A., Potter, C. S., Carragher, B., and Williamson, J. R. (2010). Visualizing ribosome assembly: a structural mechanism for 30S subunit assembly. *Science* 330, 673–677.

Munro, J. B., Altman, R. B., O'Connor, N., and Blanchard, S. C. (2007). Identification of two distinct hybrid state intermediates on the ribosome. *Mol. Cell* 25, 505–517.

Munro, J. B., Sanbonmatsu, K. Y., Spahn, C. M., and Blanchard, S. C. (2009). Navigating the ribosome's metastable energy landscape. *Trends Biochem. Sci*. 34, 390–400.

Orzechowski, M., and Tama, F. (2008). Flexible fitting of high-resolution X-Ray structures into cryoelectron microscopy maps using biased molecular dynamics simulations. *Biophys. J*. 95, 5692–5705.

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., Grassucci, R., 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., Renka, R., and Schomberg, H. (2004). Gridding-based direct Fourier inversion of the three-dimensional ray transform. *J. Opt. Soc. Amer*. A21, 499–509.

Penczek, P. A., Zhu, J., and Frank, J. (1996). A common-lines based method for determining orientations for N>3 particle projections simultaneously. *Ultramicroscopy* 63, 205–218.

Radermacher, M. (1988). The three-dimensional reconstruction of single particles from random and non-random tilt series. *J. Electron Microsc. Tech*. 9, 359–394.

Radermacher, M. (1991). Three-dimensional reconstructionof single particles in electron microscopy. In: *Image Analysis in Biology* (D.-P. Haeder, ed.), pp. 219-249. CRC Press, Boca Raton, Fl.

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.

Reimer, L. and Kohl, H. (2008). *Transmission electron microscopy: physics of image formation (5th Edition)*. New York: Springer Verlag.

Religa, T. L., Sprangers, R., and Kay, L. E. (2010). Dynamic regulation of archaeal proteasome gate opening as studied by TROSY NMR. *Science* 328, 98–102.

Rodnina, M. V., Daviter, T., Gromadski, K., and Wintermeyer, W. (2002). Structural dynamics of ribosomal RNA during decoding on the ribosome. *Biochimie* 84, 745–54.

Scheres, S. H. (2010). Visualizing molecular machines in action: single particle analysis with structural variability. In: *Recent advances in electron cryomicroscopy* (ed. Steve Ludtke). Advances in Protein Chemistry and Structural Biology, 81, 89–119.

Scheres, S. H., Gao, H., Valle, M., Herman, G. T., Eggermont, P. P., Frank, J., and Carazo, J. M. (2007). Disentangling conformational states of macromolecules in 3D-EM through likelihood optimization. *Nat. Methods* 4, 27–29.

Scheres, S. H., Valle, M., Nunez, R., Sorzano, C. O., Marabini, R., Herman, G. T., and Carazo, J. M. (2005). Maximum-likelihood multi-reference refinement for electron microscopy images. *J. Mol. Biol*. 348, 139–149.

Schuwirth, B .S., Borovinskaya, M. A., Hau, C. W., Zhang, W., Vila-Sanjurjo, A., Holton, J. M., and Doudna Cate, J. H. (2005). Structures of the bacterial ribosome at 3.5 Å. *Science* 310, 827–834.

Seidelt , B., Innis, C. A., Wilson, D. N., Gartmann, M., Armache, J.-P., Villa, E., Trabuco, L. G., Becker, T., Mielke, T., Schulten, K., Steitz, T. A., and Beckmann, R. (2009). Structural insight into nascent polypeptide chain–mediated translational stalling. *Science* 326, 1412–1415.

Shaikh, T. R., Gao, H., Baxter, W. T., Asturias, F. J., Boisset, N., Leith, A., and Frank, J. (2008). SPIDER image processing for single-particle reconstruction of biological macromolecules from electron micrographs. *Nat. Protoc*. 3, 1941–1974.

Shaikh, T. R., Yassin, A., Lu, Z., Barnard, D. Meng, X., Mohamed, H., Lu, T.-M., Wagenknecht, T., and Agrawal, R. K. (2010). Association of the ribosomal subunits as studied by time-resolved cryo-EM. (Abstract) *Microscopy and Microanalysis* 15 (suppl. 2), pp. 974–975.

Shatsky, M., Hall, R. J., Nogales, E., Malik, J., and Brenner, S. (2010). Automated multi-model reconstruction from single-particle electron microscopy data. *J. Struct. Biol*. 170, 98–108.

Singer, A., Coifman, R. R., Sigworth, F. J., Chester, D. W., and Shkolnisky, Y. (2009). Detecting consistent common lines in cryo-EM by voting. *J. Struct. Biol*. 169, 312–322.

Sorzano, C. O., Marabini, R., Velázquez-Muriel, J., Bilbao-Castro, J. R., Scheres, S. H., Carazo, J. M., and Pascual-Montano, A. (2004). XMIPP: a new generation of an open-source image processing package for electron microscopy. *J. Struct. Biol*. 148, 194–204.

Sorzano, C. O. S., Jonić, S., El-Bez, C., Carazo, J. M., De Carloe, S., Thévenaz P., and Unser, M. (2009a). A multiresolution approach to orientation assignment in 3D electron microscopy of single particles. *J. Struct. Biol*. 146, 381–392.

Sorzano, C. O. S., Recarte, E., Alcorlo, M., Bilbao-Castro, J. R., San-Martín, C., Marabini, R., and Carazo, J. M. (2009b). Automatic particle selection from electron micrographs using machine learning techniques. *J. Struct. Biol*. 167, 252–260.

Spence, J. (2003). *Experimental HREM*. (3rd Edition). New York: Oxford University Press.

Stagg, S. M., Lander, J. C., Pulokas, J., Fellmann, D., Cheng, A., Quispe, J. D., Mallick, S. P., Avila, R. M., Carragher, B., and Potter, C.S. (2006). Automated cryoEM data acquisition and analysis of 284 742 particles of GroEL. *J. Struct. Biol*. 155, 470–481.

Stark, H. and Lührmann, R. (2006). Cryo-Electron microscopy of spliceosomal Components. *Annu. Rev. Biophys. Biomol. Struct*. 35, 435–457.

Suloway, C., Pulokas, J., Fellmann, D., Cheng, A., Guerra, F., Quispe, J., Stagg, S., Potter, C.S., and Carragher, B. (2005). Automated molecular microscopy: the new Leginon system. *J. Struct. Biol*. 151, 41–60.

Tama, T., Miyashita, O., and Brooks III, C. L. (2004). Normal mode based flexible fitting of high-resolution structure into low-resolution experimental data from cryo-EM. *J. Struct. Biol*. 147, 315–326.

Taylor, K. A. and Glaeser, R. M. (1976). Electron microscopy of frozen-hydrated biological specimens. *J. Ultrastruct. Res*. 55, 448–456.

Taylor, K. A. and Glaeser, R. M. (2008). Retrospective on the early development of cryoelectron microscopy of macromolecules and a prospective on the opportunities for the future. *J. Struct. Biol*. 163, 214–223.

Trabucco, L. G., Villa, E., Mitra, K., Frank, J., and Schulten, K. (2008). Flexible fitting of atomic structures into electron microscopy maps using molecular dynamics. *Structure* 16, 673–683.

Unwin, P. N., Henderson, R. (1975). Molecular structure determination by electron microscopy of unstained crystalline specimens. *J. Mol. Biol*. 94, 425–440.

Valle, M., Sengupta, J., Swami, N. K., Grassucci, R. A., Burkhardt, N., Nierhaus, K. H., Agrawal, R. K., and Frank, J. (2002). Cryo-EM reveals an active role for aminoacyl-tRNA in the accommodation process. *EMBO J*. 21, 3557–3567.

Vallurupalli, P., Hansen, D. F., and Kay, L. E. (2008). Structures of invisible, excited protein states by relaxation dispersion NMR spectroscopy. *Proc. Natl. Acad. Sci. USA*. 105, 11766–11771.

van Heel, M. (1987). Angular reconstitution – a posteriori assignment of projection directions for 3-D reconstruction. *Ultramicroscopy* 21, 111–123.

van Heel, M. and Frank, J. (1981). Use of multivariate statistical analysis in analysing the images of biological macromolecules. *Ultramicroscopy* 6, 187–194.

van Heel, M., Gowen, B., Matadeen, R., Orlova, E. L., Finn, R., Pape, T., Cohen, D., Stark, H., Schmidt, R., Schatz, M., and Patwardhan, A. (2000). Single-particle electron cryo-microscopy: towards atomic resolution. *Quart. Rev. Biophys*. 33, 307–369.

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.

Velankar, S., Best, C., Beuth, B., Boutselakis, C. H., Cobley, N., Sousa Da Silva, A. W., Dimitropoulos, D., Golovin, A., Hirshberg, M., John, M., Krissinel, E. B., Newman, R., Oldfield, T., Pajon, A., Penkett, C. J., Pineda-Castillo, J., Sahni, G., Sen, S., Slowley, R., Suarez-Ureuena, A., Swaaminathan, J., van Ginkel, G., Vranken, W. F., Henrick, K., and Kleywegt, G. J. (2010). PDPe: protein data bank in Europe. *Nucl. Acids Res*. (Database issue) 34, 308–317.

Villa, E., Sengupta, J., Trabuco, L. G., LeBarron, J., Baxter, W. T., Shaikh, T. R., Grassucci, R. A., Nissen, P., Ehrenberg, M., Schulten, K., and Frank, J. (2009). Ribosome-induced changes in elongation factor Tu conformation control GTP hydrolysis. *Proc Natl Acad Sci USA* 106, 1063–1068.

Voss, N. R., Yoshioka, C. K., Radermacher, M., Potter, C. S., and Carragher, B. (2009). DoG Picker and TiltPicker: software tools to facilitate particle selection in single particle electron microscopy. *J. Struct. Biol*. 166, 205–313.

Wang, L. and Sigworth, F. J. (2006). Cryo-EM and single particles. *Physiology* 21, 13–18.

Wriggers, W. and Chacón, P. (2001). Modeling tricks and fitting techniques for multiresolution structures. *Structure* 9, 779–788.

Yu, X., Jin, L., and Zhou, Z. H. (2008). *3.88 Å structure of cytoplasmic polyhedrosis virus by cryo-electron microscopy*. Nature 453, 415–419.

Zhang, W., Dunkle, J. A., and Cate, J. H. (2009). Structures of the ribosome in intermediate states of ratcheting. *Science* 325, 1014–1017.

Zhang, W., Kirnmel, M., Spahn, C. M. T., and Penczek, P. A. (2008). Heterogeneity of large macromolecular complexes revealed by 3D cryo-EM variance analysis. *Structure* 16, 1770–1776.

Zhou, Z. H. (2008). Towards atomic resolution structural determination by single-particle cryo-electron microscopy. *Curr. Opin. Struct. Biol*. 18, 218–228.