Visualizing Flexibility at Molecular Resolution: Analysis of Heterogeneity in Single-Particle Electron Microscopy Reconstructions

Andres E. Leschziner1 and Eva Nogales1,2,3

1Department of Molecular and Cell Biology, University of California, Berkeley, California 94720; email: aeleschziner@berkeley.edu
2Howard Hughes Medical Institute and 3Lawrence Berkeley National Laboratory, Berkeley, California 94720; email: ENogales@lbl.gov

Key Words
cryo-EM, macromolecular assemblies, conformational flexibility, Maximum Likelihood, 3D variance

Abstract
It is becoming increasingly clear that many macromolecules are intrinsically flexible and exist in multiple conformations in solution. Single-particle reconstruction of vitrified samples (cryo-electron microscopy, or cryo-EM) is uniquely positioned to visualize this conformational flexibility in its native state. Although heterogeneity remains a significant challenge in cryo-EM single-particle analysis, recent efforts in the field point to a future where it will be possible to tap into this rich source of biological information on a routine basis. In this article, we review the basic principles behind a few relatively new and generally applicable methods that show particular promise as tools to analyze macromolecular flexibility. We also discuss some of their recent applications to problems of biological interest.
INTRODUCTION

It was not so long ago that cryo-preservation demonstrated the feasibility of imaging biological molecules in the electron microscope at high resolution (44), and even less that the problem of water vitrification was experimentally solved (2). Since then, cryo-electron microscopy (cryo-EM) has firmly established itself as one of the major techniques in structural biology. This rapid growth is due to its unique advantages: (a) the ability to image a sample under physiological or well-characterized in vitro conditions; (b) the capacity to study very large assemblies (several megadaltons); (c) the relatively small amount of sample required to obtain a structure; (d) the speed of vitrification, which allows for samples to be trapped at different stages along a reaction pathway; and (e) finally, only fully appreciated recently, the ability to sort out, computationally, different biochemical or conformational species present in the sample and to extract important biological information from their differences. Much of this review deals with this last point.

Until recently, a large proportion of cryo-EM studies focused mainly on a relatively small number of model systems that were well suited for the methods and tools being developed. A significant emphasis has been placed on collecting large data sets with the overarching goal of improving resolution and eventually obtaining atomic models. An important underlying assumption is that all images represent views originating from identical copies of the same 3D entity. While this may be the case for most studies of 2D crystals or even some helical arrangements of protein filaments (where lattice may constrain/select a given conformation), this assumption breaks down for most studies of single particles, especially when they lack internal symmetry. In fact, the application of single-particle methods to describe deviations from helical symmetry has proven successful in the study of a variety of biological polymers (9).

In recent years, cryo-EM, and in particular the single-particle methodology, has gone from a specialized technique to a general tool to explore structure, and the number of biological systems under study has grown exponentially. With this growth has come the realization that the assumption of sample homogeneity is most often not true; numerous samples are composed of a mixture of different conformers, different biochemical species (varying in their subunit composition, ligand occupancy, etc.), or a combination of both (6, 12, 13, 17, 20, 35, 47, 49). While some of this heterogeneity might be considered a nuisance (compositional differences not resolved through purification—a heterogeneity
that we refer to as biochemical)—much of it should be seen not as a problem to be circumvented, but rather as a rich source of biological information that cryo-EM is particularly well suited to tap into.

Identifying and analyzing heterogeneity is, however, still technically challenging. This article aims to present an overview of the current state of the field in regards to the treatment of heterogeneity as well as a discussion of some of the major challenges in this area for the years to come. Special emphasis is given to methodologies with general applicability rather than to strategies that build on special characteristics of the data.

We begin by briefly outlining the process of obtaining a 3D reconstruction by single-particle methods. We then discuss some of the main tools currently being developed to address heterogeneity. We introduce, in a qualitative way, the principles behind each one of these new methods and then discuss some recent applications.

SINGLE-PARTICLE RECONSTRUCTION: AN OVERVIEW

Throughout the overview of the reconstruction process the reader will find entries in square brackets; these indicate points where one or more of the methods discussed below can be applied to address sample heterogeneity. This is meant to help the reader place each of the methods discussed in the context of a full EM reconstruction.

Sample Preparation and Data Collection

Once a relatively pure specimen is available, there are three general methods of preparing a sample for EM imaging: (a) negative staining, in which the sample is embedded in a solution of a heavy metal salt and subsequently air-dried; (b) vitrification for cryo-EM analysis, in which a very thin layer of the sample solution is rapidly plunged into a potent cryogen (typically liquid ethane) (2); and (c) cryo-negative staining, a combination of the first two methods, in which a sample surrounded by a heavy metal salt solution is vitrified (1, 14).

Negatively stained samples are limited in terms of the resolution that can be achieved owing to potential flattening during air-drying, incomplete coverage by the stain, uneven stain distribution, and the fact that one is imaging the stain rather than the macromolecule itself. Although some variations of the traditional negative-staining technique can reduce these effects (25), vitrified samples are by far preferred because they represent the native species and are minimally affected by sample-preparation artifacts. Nevertheless, they suffer from limited contrast, making image analysis more challenging. They are particularly limiting for ab initio reconstruction of new samples if the methodology involves collection of images from tilted specimens (a technical bottleneck with vitrified samples). Cryo-negative staining, when tolerated by the sample, is a powerful approach that combines the high contrast resulting from the heavy metal salt with the preserving effects.

RESOLUTION

In single-particle cryo-EM, resolution refers to a particular cutoff point in a measure of self-consistency of the reconstruction as a function of spatial frequency. Once a final reconstruction is obtained, data are typically split into two halves from which half reconstructions are obtained. The two reconstructions are then compared at different frequency intervals (or shells). One common resolution criterion is based on the Fourier shell correlation and stipulates that the resolution of a structure is the frequency (or its inverse, the real-space distance) at which the two half structures show a correlation of 0.5. Because agreement on the criterion used to select the cutoff point has not yet been reached in the field, it is crucial that the cutoff point used be stated when reporting a resolution value.
of vitrification, and has led to subnanometer resolution (14). Although the ultimate goal of any single-particle EM project is to obtain a structure built exclusively with vitrified data, the use of negatively stained samples can prove useful in the initial characterization of an unknown sample.

Although manual data collection is still widespread, automated procedures are becoming increasingly common (21, 42, 48). The advantages of automation include not only the efficiency of collecting large data sets with limited user supervision but also the archiving of metadata (such as ice thickness, temperature, sample drift) that can prove useful in diagnosing problems and predicting optimal imaging conditions (41).

Two media are available to record EM data: photographic film and charge-coupled device (CCD) cameras. CCDs not only speed up data collection (and are particularly useful when coupled to automation), but also show improved signal-to-noise ratio (SNR) at low to medium resolutions (10–25 Å), an advantage for alignment of single particles (36). Unfortunately, CCDs perform poorly with medium- to high-energy electrons (300–400 keV) (8), requiring additional devices still in the experimental phase (4; K.H. Downing, personal communication).

Two-dimensional Image Analysis
Once molecular images are available in digital form (from CCD images or digitization of film), they must be aligned to each other such that those representing the same view of the particle can be averaged together to improve the intrinsic low SNR of the individual particles.

If a 3D model already exists, the particles can be aligned to references calculated from this model (see below). Otherwise, the most widely used approach for obtaining averaged characteristic views is alignment through classification. In this approach, window-centered particles are subjected to Multivariate Statistical Analysis (MSA) to identify the main features describing the variability in the data. Particles are then grouped into classes (using one of many approaches to partition the data) from which averages are generated (11). These class averages are used as references for further alignment (based on cross-correlation algorithms) and the process is iterated to convergence. It is at this stage that heterogeneity in the data from a new sample can first be detected, either by comparison of class averages or by analysis of the variance within a given class.

[Maximum Likelihood]
[3D Variance-based supervised classification.]

Reconstruction of the Initial Model
The reconstruction of structures from EM images is based on the central section theorem. This states that the Fourier transform (FT) of the 2D projection of a 3D object along a given direction (the EM image corresponding to a certain view of the object) corresponds to a central section (normal to the projection direction) of the FT of the 3D object. If a large-enough number of views were available, we would sample enough of the 3D FT of the object to recover its real space representation. The most essential and often challenging task in single-particle reconstruction is obtaining the relative orientations among the different particle views available.

Two main approaches accomplish this ab initio (where no previous structure of the complex exists): Random Conical Tilt (RCT) and Angular Reconstitution (AR) (see Sidebar). In RCT, each field of particles is imaged twice, first at a relatively high tilt (typically around 50°) and then untilted. The untilted images are used for 2D alignment and classification, and their tilted counterparts are used for reconstruction. The in-plane rotation angles that must be applied to the untilted particles in order to bring them into alignment define the relative angles of the corresponding tilted particles around a cone of views (the angle of the cone corresponding to the tilt)
Because of physical limitations as to how far the sample can be tilted in the microscope, the 3D reconstruction is missing information corresponding to unavailable views (or the equivalent central sections in Fourier space), resulting in anisotropy in the structural information. This effect is known as the missing cone, describing the shape of the missing data in Fourier space. Additional, nontrivial steps must be taken to eliminate or minimize its effect on the final reconstruction.

We recently proposed a new method of obtaining initial reconstructions, Orthogonal Tilt Reconstruction (OTR), that overcomes the missing cone problem (22) (see Sidebar). In this approach, data are collected at $-45^\circ$ and $+45^\circ$ (two orthogonal tilts). One set of views is then used for 2D alignment and classification, and the other set is used for reconstruction. Because in this case the particles used for reconstruction are orthogonal to those used to generate the class averages, the sampling of Fourier space is complete and the missing cone problem is avoided. Whereas the RCT approach is ideally suited for particles with a limited number of available views (where preferential particle orientation exists), the OTR method requires that many such views be experimentally available. Because of the technical challenges involved in collecting good-quality tilted data from vitrified samples, RTC and OTR methodologies are typically used on negatively stained samples.

The AR method utilizes exclusively untitled data and builds on analytical rather than geometrical principles. It is based on the central section theorem and the concept of common lines, although it uses its real space equivalent: Any two given 2D projections have a 1D line projection in common. Because of the noise in the images, class averages rather than individual particle images are used, and the 3D solution is obtained by searching for the orientation between two projections leading to the highest correlation for their 1D projections. This minimization is performed for all common lines among the entire set of available projections (45). As for OTR, this method requires random particle orientations.

The actual generation of the 3D object (the reconstruction), once the relative orientations of the projections have been assigned, can be accomplished either by Fourier methods (back-transforming the 3D FT) or by a real-space process called back-projection. In back-projection (the reverse of projection) each particle (or class average) is stretched through space in a direction normal to the image plane; the intersection of all these densities gives rise to the reconstruction.

**Refinement**

Initial reconstructions can be improved through refinement. The initial reconstruction is used as a starting reference to determine the orientations of experimental particle images, often from a more extensive data set, and the new orientation parameters (or angular assignments) are used to generate a new, improved 3D structure. The process is iterative, as the new structure is then used as a reference for a new cycle of angular assignments. As the refinement progresses, the fineness in the alignment parameters is increased, leading to higher resolution (see sidebar, Resolution). For a given data set convergence is reached when the angular assignments do not change and the resolution of the structure stops improving. The assessment of resolution for single-particle reconstructions has so far eluded consensus. A variety of criteria exists and the topic has been extensively discussed in the literature (5, 10, 16, 24, 26, 30, 31, 33, 34).

The assignment of orientation parameters is commonly done using either AR (where the reference structure is used as the source of a set of central sections with fixed relative orientations against which the orientations of the input images are determined) or projection matching. In projection matching, projections are calculated from the reference structure...
and correlation methods are used to find the reference projection that best matches each of the input images (28).

[Maximum Likelihood]
[3D variance-based supervised classification.]
[Normal Mode Analysis to generate references]

METHODS FOR ADDRESSING HETEROGENEITY

Normal Mode Analysis: Principles and Applications

Molecular dynamics simulations are a standard way of looking at molecular motion. Yet, because of the dimensions of the macromolecules typically studied by cryo-EM and the timescale of the conformational changes we wish to understand, molecular dynamics are presently beyond our reach. An alternative approach is Normal Mode Analysis (NMA), where the molecular motions are decomposed into a few, typically low-frequency/long-range vibrational modes that can describe the dynamic behavior of the molecule. For a recent review of the technique see Reference 43.

A major application of NMA is in hybrid methods. Here, high-resolution information for a given conformation of a macromolecule (an X-ray crystallography, electron crystallography, or NMR structure) can be used in combination with NMA to obtain pseudoatomic models when low-resolution information (a cryo-EM reconstruction) is available for other conformers (43).

Of more relevance to the problems discussed in this review is the application of NMA to the generation of references for multireference refinement when the source of the heterogeneity is conformational variability. Once an initial reconstruction is available, NMA can be used to predict other possible conformations. References are generated from these predicted structures and used to refine the data by allowing the images to choose simultaneously among them.

A successful application of this approach was published by Brink and colleagues (6). The authors had previously obtained a cryo-EM reconstruction of the human fatty acid synthase (FAS). The enzyme is a large (550 kDa) antiparallel homodimer with a central connector between the two antiparallel regions of density (the overall shape is that of the letter H), and the authors predicted that a certain level of flexibility would be required for its activity (7). In the more recent study, they used NMA to predict a number of vibrational modes for the structure and tested whether the two showing the lowest frequency (that is, the largest conformational changes) were present in their data. The authors used multireference refinement with five references: the original reconstruction with C2 symmetry (7) and both extremes of the two vibrational modes. The refinement gave structures that resembled the predicted conformers. In order to rule out model bias (a common concern in single-particle analysis), Brink et al. performed a single-reference refinement of those particles that had been assigned to one of the NMA-based references but used the C2 symmetric reconstruction as the reference. The refined structure resembled the NMA-predicted conformation and not the starting C2 reference.

3D Multivariate Statistical Analysis and Classification: Principles and Applications

Whereas MSA and classification of images is a standard step during the 2D analysis stage of any new single-particle reconstruction, its application in 3D, to reconstructed volumes, is far less widespread. The technique is more commonly used in cryo-electron tomography, where hundreds of individual 3D representations of a macromolecule can be obtained from a single tomogram. The first application of 3D MSA and classification to tomographic reconstructions was published a decade ago.
For some examples of the most recent applications to tomography, see References 18 and 23.

3D MSA and classification of single-particle reconstructions are less common partly due to the size of the data sets required for this type of analysis. With the growing use of automated data collection and the continuous increase in computing power, this technique should become a powerful tool for the analysis of heterogeneity in single-particle applications as well.

Given that MSA and classification are well-established techniques, the principles behind them will not be covered in detail here. We refer the reader to Joachim Frank’s (11) excellent textbook for a comprehensive introduction. Briefly, MSA is a numerical data reduction tool that identifies patterns in the data and describes the variability present in it using a relatively small number of factors or principal components. Once the images are placed in a multidimensional factor space (based on the contribution each factor makes to a given image), their distribution can be divided into classes (classification) using one of a variety of clustering algorithms. Images in a given group are closely related with a signal that is correlated but noise that is not. Averaging of images within a class results in a significantly improved view of the particle owing to the increased SNR.

The application of 3D MSA and classification should have unique advantages in those cases in which the heterogeneity in the sample affects a large part of the structure, as might be the case with large conformational changes or biochemical heterogeneity involving a significant portion of the molecule. Under those conditions the assignment of orientation parameters to experimental images necessary to calculate a 3D variance map (see below) may be too unreliable if models describing the heterogeneity are not available. If, on the other hand, a large set of independent initial reconstructions are available, these can be subjected to 3D MSA and classification to obtain a description of the existing heterogeneity and generate references that can be used to sort the data into more homogeneous groups that can be further analyzed.

Critical to the success of MSA and classification (either in 2D or 3D) is the alignment of the images/volumes. Alignment of volumes is often not trivial. If a sample is suspected to be heterogeneous, the initial reconstructions are likely obtained using either tomography or the RCT method. Both techniques use tilting of the sample but suffer from the absence of those views that are physically inaccessible, resulting in areas in Fourier space that contain no information (wedge- or pyramid-shaped for tomography and cone-shaped for RCT). Because the alignment of volumes is performed using cross-correlation methods (i.e., multiplication of complex conjugates in Fourier space), it is affected by these empty areas.

As described above, we have recently proposed OTR as a novel reconstruction method that circumvents the missing data problem while still relying on tilted data (22) (see Reconstruction of the Initial Model, above). Volumes generated with this method can be directly aligned to each other without any biasing imposed by the missing information.

Stark and colleagues have recently applied 3D MSA in a single-particle setting to study the conformational variability of the U5 snRNP spliceosomal component. Whereas U5 snRNPs prepared under standard conditions showed a single, well-defined conformation, complexes prepared under high-salt conditions were found to adopt several conformations (Figure 1). For their analysis, the authors generated large sets of up to 300 individual RCT reconstructions. The initial alignment of the reconstructions was performed using a Maximum Likelihood (ML) approach (see below), with initial references consisting of random Gaussian noise (see Sidebar). In order to avoid the bias imposed by the missing cone in the RCT reconstructions, they carried out the 3D alignment search using projections of the volumes onto
Figure 1

3D MSA and classification of multiple conformations of the 5U snRNP spliceosomal complex. The figure illustrates a few representative reconstructions of the multiple conformations identified for 5U snRNP spliceosomal complexes purified under high-salt conditions. The conformation observed under normal salt conditions is shown on the left. The structures shown on the right were obtained using 3D MSA and classification of a large set of RCT reconstructions. (See text for details.) (Figure courtesy of H. Stark.)

polar coordinates. After a few cycles of this alignment the individual volumes were finally aligned against a consensus model combining all of them (in weighted form) and were then subjected to 3D MSA. Classification and averaging were then used to obtain representations of the U5 snRNP conformers in the sample and to quantitate the extent to which each conformer was represented (Figure 1) (B. Sander, M.M. Golas & H. Stark, personal communication).

3D Variance and Supervised Classification: Principles

Alignment and classification of images are often insufficient to detect and/or characterize the heterogeneity present in a sample. This can be the case if a flexible portion of a molecule becomes blurred out through averaging when its total contribution to the molecular mass (and thus to the alignment) is small, or if small differences become difficult to see in certain views (for example, because they are located in the same line of projection as a major mass in a complex). The detection (and sorting out) of heterogeneity is particularly challenging when initial references that describe it are not available, as is often the case in a new project.

One common sign that a structure might suffer from heterogeneity is a resolution that is significantly lower than expected for a given data set size. While low resolution is a measure of the average error for the entire structure, it tells us nothing about the nature or location of the heterogeneity. This information, however, can be obtained by applying the 3D variance method recently proposed by Penczek and colleagues (29).

In principle we should be able to estimate the variation among all the species present in our sample by calculating the statistical variance from a number of independent reconstructions of the sample. This 3D variance would identify those parts of the structure where the heterogeneity is concentrated (as they should differ most among the reconstructions). The difficulty with this approach lies in the enormous amount of data that would be required. Penczek et al. (29) proposed a bootstrap method that uses sampling with replacements to obtain an estimate of the variance of a given population (the large number of reconstructions we would generate if data were available) from a single sample (our data set). In sampling with replacements each sampling omits some elements and repeats others, with the total number of elements staying constant and equal to the size of the
full data set. Once the bootstrap variance is obtained, one must subtract those contributions made to the variance by (a) the noise and (b) the reconstruction algorithm itself [misalignment of particles also contributes to the variance, but this contribution cannot be estimated independently (29)]. This correction is accomplished by calculating a bootstrap background variance with noise images selected from the same data set and the alignment parameters assigned to the particle images. This background variance, which contains contributions from both the noise and the reconstruction algorithm, is subtracted from the bootstrap variance to obtain the final variance of structures (29) (Figure 2a–f).

In addition to mapping the location of heterogeneity in a reconstruction (areas of high variance), this approach can provide information on the degree of correlation between these areas through the analysis of covariance. For example, if a portion of a macromolecule can oscillate between positions A and B, we would see a region of high variance at both the A and B positions and these variance peaks will be negatively correlated (Figure 2g).

The 3D variance map just described provides the means to carry supervised classification in those areas identified as containing the highest level of heterogeneity (3). The idea is to generate spherical 3D masks centered on one or more areas of high variance (Figure 2b), which are then projected in pseudoevenly spaced directions to generate 2D masks (Figure 2i). Images representing the same view (that is, having the same angular assignment) can now be sorted into subgroups by carrying out MSA and classification within the 2D masks (Figure 2j). User interaction is required during supervised classification in order to estimate, among other things, the number of species present and therefore the number of classes into which the images must be split. A challenging aspect of this method is the assignment of the newly sorted subgroups according to the structure that gave rise to them (Figure 2k). Whereas the presence or absence of density in the case of a ligand may be relatively simple, this assignment might be more difficult in the case of a conformational change that results in high variance in several parts of the structure. Analysis of the covariance map would facilitate this process by indicating those areas where changes in density are correlated (either positively or negatively).

Our lab recently implemented a new method to sort out different views of multiple conformations coexisting in an EM sample using cross-correlation of common lines (40; R.J. Hall, B. Siridechadilok & E. Nogales, manuscript in preparation). During our cryo-EM reconstruction of the human translation initiation factor eIF3 bound to the hepatitis C virus IRES RNA, we found that the RNA adopted several conformations. We used 2D variance-based masks to split the particles in each view (determined by projection matching to the apo eIF3) among three different conformers and calculated a subclass average for each. We then chose a view for which the three conformers looked distinctively different and used these as references to assign other views to one of the three conformations. The method takes advantage of the fact that the spatial relationship among all the views was known from the initial determination of Euler angles against a single eIF3 model and therefore the common lines had been identified. It then assumes that the cross-correlation of common lines between two views will be higher if the conformations are the same. The projection views were therefore sorted on the basis of the highest cross-correlation of common lines with the three references. This procedure was enough to assign sufficient views to the three conformations to generate three distinct references that were further improved by refinement.

Whichever method used, once images have been sorted into groups on the basis of the supervised classification and the new reconstructions generated, multireference refinement (where the new reconstructions are allowed to compete as references for the
3D Variance and Supervised Classification: Applications

Three applications of the 3D variance and its associated supervised classification have been published to date (15, 19, 27). Penczek and colleagues (3) applied the methodology to complexes of the *Escherichia coli* 70S ribosome, tRNA, and the elongation factor G (EF-G).

The data set had been previously analyzed and found to be heterogeneous. Although the complex solved was expected to represent a pretranslocational ribosome containing tRNAs in the A and P sites as well as EF-G, the density of the elongation factor was fragmented, a result of substoichiometric binding of EF-G (3). The 3D variance map confirmed this; the highest level of variability was found in the intersubunit space where the ligands are found.

The variance map was then used as the source for a 3D mask that was projected as 2D masks and used in supervised classification. The experimental images were sorted into two groups (under the assumption that they differed in the presence or absence of EF-G) on the basis of whether they showed high or low density within the 2D mask; two new reconstructions were generated and used as references for a multireference refinement. The final two structures represented a ribosome containing either EF-G and no A or P site tRNAs (but increased density in the E-site) or tRNAs in the A and P sites but no EF-G. Furthermore, a conformational change between the two ribosomal subunits (known as ratchet-like subunit rearrangement), which is believed to be associated with the translocation of tRNAs, was seen in the EF-G structure but not in that containing the tRNAs. The results of the reconstructions after sorting were supported by the analysis of the covariance map. Significant negative correlations were found between the EF-G region and that occupied by the tRNAs in the A and P sites, indicating that their presence is mutually exclusive. The covariance map also indicated that the L1 protuberance oscillated between two distinct conformations (i.e., the covariance showed negative correlation between the two peaks).

Figure 2

3D variance and supervised classification. The sample in this figure contains four species (a). These differ both in conformation and ligand binding (yellow sphere). While the single protrusion at the bottom of the structure oscillates independently of the binding of ligand, the top protrusions are closed in its absence and open in its presence (a). An initial reconstruction results in a heterogeneous structure (b). This reconstruction is used to generate reference projections for the assignment of orientations (d) to a set of experimental images (c). N bootstrap volumes can then be reconstructed (e) using those alignments and sampling with replacements where some images are omitted and others are repeated more than once in a given reconstruction [represented by the relative thickness (or absence) of the arrows leading to the bootstrap volumes]. A variance map (f) can be calculated from the bootstrap volumes (peaks shown in gray superimposed on a semitransparent representation of the initial heterogeneous reconstruction). A covariance map (g) shows whether pairs of peaks of high variance are positively or negatively correlated (or uncorrelated) and thus yields information on the coupling of the observed changes. For example, the negative correlation between the two peaks (blue and yellow) shown at the bottom of the map indicates that occupancy of those locations is mutually exclusive. Similarly, the positive correlation seen between the peak corresponding to the ligand (green) and the peak (orange or red) shown at the top of the map indicates that the presence of density at those locations is coupled. The 3D variance map can be used as the source of a 3D mask (b) that can be projected to generate 2D masks (i), which are then binarized (j). These masks are used in supervised classification (j) to restrict the analysis of the images to those areas that contain the highest variability. Finally, images are sorted into groups representing different species; new reconstructions are obtained and used in multireference refinement (k).
Figure 3

3D variance analysis of (a) human transcription factor II D (hTFIID) and (b) human RNA polymerase II (hRNAPII). (a) The 3D variance map obtained for the cryo-EM reconstruction of hTFIID (left) was used to generate masks for supervised classification to sort images into groups representing conformation 1 or conformation 2. Two new reconstructions were obtained (center) and the difference map between them (right) was calculated. Yellow and orange areas represent peaks of high variance (left), and arrows indicate regions with high correlation (positive or negative) in the covariance map. The difference map confirms that negative correlations correspond to pairs of densities that are mutually exclusive in a given conformer, whereas positive correlations indicate they coexist. (b) An initial reconstruction of hRNAPII was generated using cryo-negative stain EM data. The 3D variance map was calculated, and the aligned experimental images underwent supervised classification within the 2D projections of the 3D variance-based mask around one of the main variance areas. Images were sorted into two groups on the basis of whether density was (i) present or (ii) absent within the mask. New reconstructions (3D) were generated from the sorted data representing closed and open conformations. (Figure courtesy of P. Grob.)

Our lab has applied the 3D variance and supervised classification approach to the study of the human transcription factor II D (hTFIID) (15) (Figure 3a) and human RNA polymerase II (hRNAPII) (19) (Figure 3b). In both cases, after a 3D variance map showed the presence of heterogeneity, supervised classification was used to sort the images using 2D masks derived from the 3D variance map. For hTFIID most of the variability could be accounted for by using two classes per view. Two new reconstructions were generated from these two sets and used for a multireference refinement of the whole data set. The closed hTFIID conformation showed improved resolution and reduced variance, whereas the open conformation did not. Thus, this open reconstruction may still represent a mixture of a few conformations, while the closed reconstruction has increased stability as a result.
of a larger number of protein contacts. The idea that these two structures represent different conformations of an otherwise similar complex (with limited biochemical heterogeneity) is supported by a number of observations (see Figure 3a): (a) the difference map between the two structures and the 3D variance map resemble each other; (b) a positive peak in the difference map was generally accompanied by a corresponding negative one; and (c) the regions of high variance were strongly correlated. This last observation suggests a cooperative conformational breathing of the complex that is likely an intrinsic property of hTFIID that enables it to interact with a variety of other factors as it directs the assembly of the transcriptional machinery (15).

In our studies of the human RNAPII (Figure 3b) we found that the different conformations reported for the yeast polymerase in crystallographic studies with different ligands were sampled by the human protein in (vitrified) solution in the absence of substrate, suggesting that this flexibility is intrinsic and plays a biological role. Fitting the yeast crystal structures into the cryo-EM densities showed that neither of the two major conformers corresponded exactly to the yeast crystallographic models. These apparent disparities may be due to an intrinsic difference between the yeast and human complexes or, alternatively, to conformations prevalent in solution that may not have yet been captured in a crystal (18).

**Maximum Likelihood: Principles**

The major source of all challenges in cryo-EM stems from the noisy character of the images, which have to be collected using very low electron doses (10–20 e−/Å²) in order to minimize radiation damage. The low SNR makes comparisons involving particles (in any of the steps outlined in the overview) uncertain to some degree. For example, when particles are aligned relative to some set of references, it is typical to use cross-correlation coefficients as indicators of the best match. Yet, the best match is often associated with a cross-correlation coefficient that is only slightly higher than those of several other candidates. Although we would choose that reference as the correct match, it might be more accurate to say that it is likely the best match but others have a certain probability of being correct as well. The ML approach is a formal expression of this uncertainty.

The application of ML to alignment in cryo-EM was initially proposed and demonstrated with synthetic data by Sigworth (39) and later applied to experimental data by Carazo and colleagues (38). An outline of the basic principles of the ML approach follows (readers are referred to the two articles cited above for a rigorous mathematical description of the methodology).

The objective in the ML method is to maximize the probability that a given reference (2D or 3D) would give rise to a given experimental data set. When a set of references are used (as is the case in alignment and refinement), each experimental image has a given probability of having arisen from each of the available references depending on (a) its pixel-by-pixel similarity to the reference and (b) the actual parameters required to transform the latter into the former (extent of the shifts and statistics of the added noise). This probability is turned into a weighting factor that controls the contribution that a particle image will make to each of the new averages (one for each available reference). As the cycles of alignment (or refinement) progress, the likelihood that the current model will give rise to the experimental data is maximized.

Consider a given reference X that matches (in a pixel-by-pixel sense) two experimental images, A and B, equally well. Images are assumed to be roughly centered; therefore, if a much larger shift is required to bring X in alignment with B, then the weight applied to B when generating the new version of X will be significantly smaller than the weight applied to A to reflect the lower probability associated
with the larger shift. Figure 4 illustrates with a synthetic example the steps taken during a cycle of refinement involving two references and a heterogeneous data set containing two similar but distinct structures.

ML methods can be applied to both 2D and 3D problems. The principle is the same; the implementation only differs in that in 3D applications the 2D alignment is used also as the source of the angular information (Euler angles) required to generate the new reconstruction.

**Maximum Likelihood: Applications**

The initial implementation of a ML approach to the problem of alignment of noisy images was carried out by Sigworth (39). In this study, the author used synthetic data derived from a projected view of a bacteriorhodopsin trimer, to which different amounts of Gaussian noise were added. He compared the performance of ML with that of the more traditional cross-correlation alignment (CC) algorithm. The two algorithms produced similar results unless the images had very low SNR. A modified CC approach was used in this study; it included, as ML does, a weighting factor that took into consideration the expected distribution of transformations applied to the images during the alignment. This modified CC performed better than the traditional one. Once the SNR was reduced to 0.005, CC failed. Furthermore, at this SNR, an artificial, incorrect initial reference was reproduced by the CC method regardless of the number of iterations performed in the alignment. ML, on the other hand, converged to a structure similar to the true model with either an initial noisy reference or the incorrect one, the first indication of the robustness of this technique with respect to the initial reference. Importantly, whereas noise images reproduced a reference when a CC approach was used, they diverged from it when treated by ML.

More recently, the group of Carazo (38) applied the ML approach to both synthetic and experimental data with structural heterogeneity. They used a cryo-EM data set of the large SV40 T-antigen bound to an asymmetric DNA. The structure consists of two hexameric rings stacked head-to-head with the DNA running through their centers. Because of the asymmetric nature of the DNA, they expected to visualize a single protruding end.

---

**Figure 4**

The Maximum Likelihood (ML) method. Shown is a sample containing two different species: a narrow and a broad H (a). Initial alignment of the images shows the two underlying structures but fails to fully resolve the heterogeneity (b). In the ML method, the goal is to maximize the likelihood that the references (b) give rise to the experimental data (c). For each experimental image, transformations (shifts, rotation, and noise) are applied (d) to every available reference and a best match is found. Each transformation applied (except for the rotation, which is evenly distributed and has been ignored in this simplified example) has a probability associated with it. Experimental images are assumed to be roughly centered within their windows; therefore, large shifts have lower associated probabilities. In the example shown in (e), both the narrow and broad H’s best matched the experimental image when shifted rightward along X. Yet, the shift that was applied to the broad H is much larger and therefore has a lower probability associated with it (f). Similarly, the noise in the images is assumed to follow a Gaussian distribution. If the noise applied to a reference (that gave a good match to an experimental image) deviates from this distribution, its probability will also be low (g). A cross-correlation is also calculated between each transformed reference and experimental image (h) to score their similarity. When new averages are calculated (i) (one for each original reference), a weighting function is applied that combines the probabilities associated with the different transformations as well as the measure of similarity between images (i). The relative thicknesses of the arrows leading from the experimental images represent their relative weights in the formation of the new averages. These averages show reduced heterogeneity (j). The process is then iterated.
All alignments were started with a reference consisting of an average of a subset of unaligned images (a blob); this removes any initial bias from the alignment. Although both ML and CC reached similar final structures that depict the known flexibility along the long axis of the dodecamer, ML resulted in a significantly higher level of detail. Furthermore, an extension believed to be the DNA is visible in one of the averages obtained with ML but in none of those obtained with CC. Confirming the initial observation of Sigworth, the alignment of a negatively stained sample (with a higher SNR) of the replicative helicase G40P gave similar results with either the CC or ML methods.
An exciting result for an application of ML in 3D was recently reported by Scheres et al. (37). In this study, the authors set out to resolve the heterogeneity that was detected in a sample consisting of 70S *E. coli* ribosomes, tRNAs, and EF-G. The heterogeneity appeared in the initial reconstruction as a blurred density for the small subunit of the ribosome and an incomplete density for EF-G. A single initial reference that was strongly filtered (to eliminate all but the most general features of the complex) (Figure 5a) was used in a single iteration of 3D ML using four random sets of images. This resulted in four similar (but not identical) reconstructions that could be used as seeds for the iterative portion of the 3D ML algorithm. (The need for four seeds was determined empirically.) After 25 iterations against a data set of more than 90,000 cryo-EM images, Scheres et al. obtained two distinct subpopulations. The first subpopulation, represented by three of the four reconstructions, consisted of a ribosome containing three tRNAs (in the A, P, and E sites), and the second subpopulation, represented by a single reconstruction, consisted of a ribosome with EF-G and a single tRNA bound in the hybrid P/E site (Figure 5b, c). These four structures were further refined using standard projection-matching algorithms (i). (Figure courtesy of S. Scheres.)
THE FUTURE

Given that only two decades have elapsed since the initial vitrification of a sample for the electron microscope, we feel justified in our optimism that the next two decades will prove to be an even more exciting time. We feel particularly encouraged by the concerted effort the field is making to address the difficult issue of heterogeneity, developing tools that will enable us to tap into this rich source of biological information. The findings highlighted in this review already demonstrate the importance of these methods when interpreting cryo-EM reconstructions.

The new methods just described (and those we could not cover) must now be tested more extensively in a wide variety of samples to make them as generally applicable as possible. At the same time, we expect to see more and more examples of projects that combine many of these tools to make them even more powerful.

The analysis of heterogeneity is most often a trade-off between the extent to which one wishes to describe the variability in the sample (i.e., the number of discrete structures one sets out to characterize) and the resolution desired for the individual reconstructions. Because of limitations in the size of data sets, an increase in the former is necessarily coupled to a decrease in the latter (and vice versa). With the continuous growth in automation of data acquisition and computer power, it is reasonable to expect that this compromise will decrease over time.

One can envision a not-too-distant future where a large number of reconstructions are obtained from a single sample with automated data collection applying either the Random Conical or Orthogonal Tilt geometries and using automated Maximum Likelihood–based alignment and classification of the images. The volumes would be analyzed by 3D MSA and classification to identify and describe any heterogeneity. References would be used to sort out new (untilted) data into more homogeneous groups and new reconstructions would be obtained. 3D variance maps and supervised classification would further increase the homogeneity within groups of images and the final reconstructions would be subjected to Maximum Likelihood–based multireference refinement.

SUMMARY POINTS

1. Macromolecular flexibility is an important source of biological information.
2. Single-particle analysis of vitrified samples is ideally suited for the visualization of conformational flexibility.
3. The noisy character of cryo-EM images creates the necessity for computational means to sort out images originating from different 3D objects.
4. Normal Mode Analysis can be used to predict different conformations for a given structure. The predicted conformations can then be used as references for further refinement.
5. 3D MSA can help in identifying the different conformations present in a sample.
6. 3D variance and covariance maps can localize areas of variability in a reconstruction and provide information about coupling among them.
7. Maximum Likelihood methods take into account the uncertainty involved in aligning (comparing) noisy images.
ACKNOWLEDGMENTS

This work was funded by grants from the Office of Biological and Environmental Research for the U.S. Department of Energy, the National Cancer Institute, and the National Institutes of Health. E.N. is a Howard Hughes Medical Institute Investigator.

LITERATURE CITED


8. Downing KH, Hendrickson FM. 1999. Performance of a 2k CCD camera designed for electron crystallography at 400 kV. *Ultramicroscopy* 75:215–33


17. Time-resolved description of viral capsid maturation by studying the relative contributions from different conformations to samples obtained at different time points.


47. An example of the use of eigen images to sort out size differences in single-particle analysis.

49. Illustrates the degree to which “breathing” of a macromolecular complex can be detected in a single sample.
# Contents

**Frontispiece**  
*Martin Karplus* ................................................................. xii

Spinach on the Ceiling: A Theoretical Chemist’s Return to Biology  
*Martin Karplus* ................................................................. 1

Computer-Based Design of Novel Protein Structures  
*Glenn L. Butterfoss and Brian Kuhlman* ...................................... 49

Lessons from Lactose Permease  
*Lan Guan and H. Ronald Kaback* .............................................. 67

Evolutionary Relationships and Structural Mechanisms of AAA+ Proteins  
*Jan P. Erzberger and James M. Berger* ......................................... 93

Symmetry, Form, and Shape: Guiding Principles for Robustness in Macromolecular Machines  
*Florence Tama and Charles L. Brooks, III* .................................... 115

Fusion Pores and Fusion Machines in Ca\(^{2+}\)-Triggered Exocytosis  
*Meyer B. Jackson and Edwin R. Chapman* .................................... 135

RNA Folding During Transcription  
*Tao Pan and Tobin Sosnick* .................................................. 161

Roles of Bilayer Material Properties in Function and Distribution of Membrane Proteins  
*Thomas J. McIntosh and Sidney A. Simon* ................................... 177

Electron Tomography of Membrane-Bound Cellular Organelles  
*Terrence G. Frey, Guy A. Perkins, and Mark H. Ellisman* ................. 199

Expanding the Genetic Code  
*Lei Wang, Jianming Xie, and Peter G. Schultz* .............................. 225

Radiolytic Protein Footprinting with Mass Spectrometry to Probe the Structure of Macromolecular Complexes  
*Keiji Takamoto and Mark R. Chance* ........................................ 251
The ESCRT Complexes: Structure and Mechanism of a Membrane-Trafficking Network
James H. Hurley and Scott D. Emr .................................................. 277

Ribosome Dynamics: Insights from Atomic Structure Modeling into Cryo-Electron Microscopy Maps
Kakoli Mitra and Joachim Frank .................................................. 299

NMR Techniques for Very Large Proteins and RNAs in Solution
Andreas G. Tzakos, Christy R.R. Grace, Peter J. Lukavsky, and Roland Riek ........ 319

Single-Molecule Analysis of RNA Polymerase Transcription
Lu Bai, Thomas J. Santangelo, and Michelle D. Wang .................................. 343

Quantitative Fluorescent Speckle Microscopy of Cytoskeleton Dynamics
Gaudenz Danuser and Clare M. Waterman-Storer ........................................ 361

Water Mediation in Protein Folding and Molecular Recognition
Yaakov Levy and José N. Onuchic .................................................. 389

Continuous Membrane-Cytoskeleton Adhesion Requires Continuous Accommodation to Lipid and Cytoskeleton Dynamics
Michael P. Sheetz, Julia E. Sable, and Hans-Günter Döbereiner ..................... 417

Cryo-Electron Microscopy of Spliceosomal Components
Holger Stark and Reinhard Lührmann .................................................. 435

Mechanotransduction Involving Multimodular Proteins: Converting Force into Biochemical Signals
Viola Vogel .................................................................................. 459

INDEX

Subject Index .................................................................................. 489
Cumulative Index of Contributing Authors, Volumes 31–35 ...................... 509
Cumulative Index of Chapter Titles, Volumes 31–35 ............................... 512

ERRATA

An online log of corrections to Annual Review of Biophysics and Biomolecular Structure chapters (if any, 1997 to the present) may be found at http://biophys.annualreviews.org/errata.shtml