Review

Probing Structural Perturbation of Biomolecules by Extracting Cryo-EM Data Heterogeneity

Kira DeVore [ID] and Po-Lin Chiu *[ID]

School of Molecular Sciences, Biodesign Center for Applied Structural Discovery, Arizona State University, Tempe, AZ 85287, USA; kdevore1@asu.edu
* Correspondence: plchiu@asu.edu

Abstract: Single-particle cryogenic electron microscopy (cryo-EM) has become an indispensable tool to probe high-resolution structural detail of biomolecules. It enables direct visualization of the biomolecules and opens a possibility for averaging molecular images to reconstruct a three-dimensional Coulomb potential density map. Newly developed algorithms for data analysis allow for the extraction of structural heterogeneity from a massive and low signal-to-noise-ratio (SNR) cryo-EM dataset, expanding our understanding of multiple conformational states, or further implications in dynamics, of the target biomolecule. This review provides an overview that briefly describes the workflow of single-particle cryo-EM, including imaging and data processing, and new methods developed for analyzing the data heterogeneity to understand the structural variability of biomolecules.

Keywords: single-particle cryo-EM; heterogeneity; molecular dynamics; deep learning; molecular dynamics flexible fitting; image classification

1. Introduction

The revolution in cryogenic electron microscopy (cryo-EM) has dramatically advanced knowledge in the field of structural biology [1,2]. Technical advances in hardware improvement, sample preparation, and algorithms for data analysis have pushed the limit of cryo-EM imaging in biological applications. Thanks to the fast image acquisition of the direct electron detector (DED) camera, the electron images can be recorded in a time series as a molecular movie [3–5]. Massive image processing tasks can be efficiently completed by increasing computing power, making the image reconstruction workflow more efficient. For single-particle cryo-EM, it has now become a routine to collect thousands of electron movies per day on a transmission electron microscope (TEM) and obtain an initial three-dimensional (3D) model from a computer workstation within a few hours or days. The generated high-resolution structural information of biomolecules dramatically expands our understanding of the molecular mechanisms that benefit the fields of biology and medicine.

In a typical single-particle cryo-EM workflow, the protein molecules are rapidly frozen and present multiple states and orientations in thin vitreous ice. The electron scattering by the proteins forms an image by the electromagnetic lenses and is recorded on a detector. Individual molecular images are then used to reconstruct the 3D Coulomb potential density of the target molecule [6]. The reconstruction requires accurate Euler angle assignments on individual two-dimensional (2D) particle images and an appropriate interpolation scheme for 3D back projection [6,7]. High-resolution details of a protein structure, such as residue side chains or ligands, could emerge by averaging a large number of aligned images of the same entity, because the averaging process flattens the noise and increases the signal-to-noise ratio (SNR). However, in some cases, the same averaging procedure does not generate any high-resolution structural details (possibly due to small movements of side chains) or will resolve the protein density at a very low contour level (possibly due to a disordered

structure). It could be that the structural variations recorded in the 2D projections smear the densities in the moving trajectory or flatten out the featureless densities. Thus, a 3D reconstruction can be seen as a back-projected average of 2D particle images with their structural variabilities flattened.

One of the challenging problems in analyzing cryo-EM data is that the image has a low SNR. An efficient and effective algorithm will be needed to extract the data heterogeneity related to the molecule's structural variability. Because biological molecules are sensitive to electron radiation damage, one has to use a lower electron dose to image the specimens to preserve the structural integrity [8–11]. The signal content of the object is then compromised by the low imaging dose, which limits the signal level that can be extracted from the noise. One can increase the particle data size to increase the SNR by averaging. Recently, new strategies of cryo-EM data collection have enabled generating a large particle dataset within a reasonable timeline [12]. Newly developed algorithms, including machine learning-based approaches, open the possibility of analyzing the molecular motions that are captured in cryo-EM images, such as ribosomes [13–15], spliceosomes [13,16–18], and G-protein coupled receptors (GPCRs) [19–22].

This review will first overview the typical processing workflow of single-particle cryo-EM. Next, we will discuss typical sources of heterogeneity within a dataset and list the algorithms used to analyze the image variance by reviewing the principles of conventional multivariate statistical analysis (MSA), optimizations using maximum likelihood (ML), and Bayesian parameter estimation. We will then introduce machine learning approaches to derive a dynamical model from the cryo-EM data. The last part will list some current methods used to analyze the cryo-EM heterogeneity or prepare the cryo-EM sample at different time points, so-called time-resolved cryo-EM. The extracted information from cryo-EM data heterogeneity can provide a knowledge framework about biomolecular dynamics, which may implicate its function and improve the understanding of its working mechanism.

2. Typical Workflow of Single-Particle Cryo-EM

A typical single-particle cryo-EM workflow begins with a biochemically homogeneous protein sample (Figure 1A). To preserve the sample in a hydrated state in the EM column, one can embed the sample in an anti-freezing agent, such as trehalose [23], or vitrify the sample by fast freezing into a cryogen, such as liquid ethane or liquid propane, so-called plunge freezing. Vitrification results in amorphous ice formation below the glass transition temperature and prevents crystalline ice from damaging the proteins [23]. Additionally, because the vitrified sample is in a hydrated state, the preserved structure is close to its native condition. The idea of plunge freezing was initially developed by the 2017 Noble Prize laureate Jacques Dobuchet in the 1980s. It was shown that pure water could be vitrified and used for EM imaging on radiation-sensitive materials [24]. The cryogenic biological sample of adeno-associated viruses was first demonstrated structurally intact and feasible for cryo-EM imaging [25].

A biological specimen is composed of light elements, which can be seen as a weak-phase object that only modulates the phase of the incoming wave [26,27]. Because the intensity variations of a weak-phase object are minimal, the image contrast is minimal in focus. Because one can change the focal plane by adjusting the objective lens strength, this allows for the modulation of the contrast transfer function (CTF). The typical use for imaging a biological specimen is to introduce a defocus to enhance the contrast in the lower spatial frequency region in a non-linear manner [28,29]. However, defocusing causes a high contrast transfer oscillation in the high-frequency region and reduces the spatial coherence, compromising the quality of the phases in the high-frequency region for the subsequent image reconstruction (Figure 1B) [28,29].
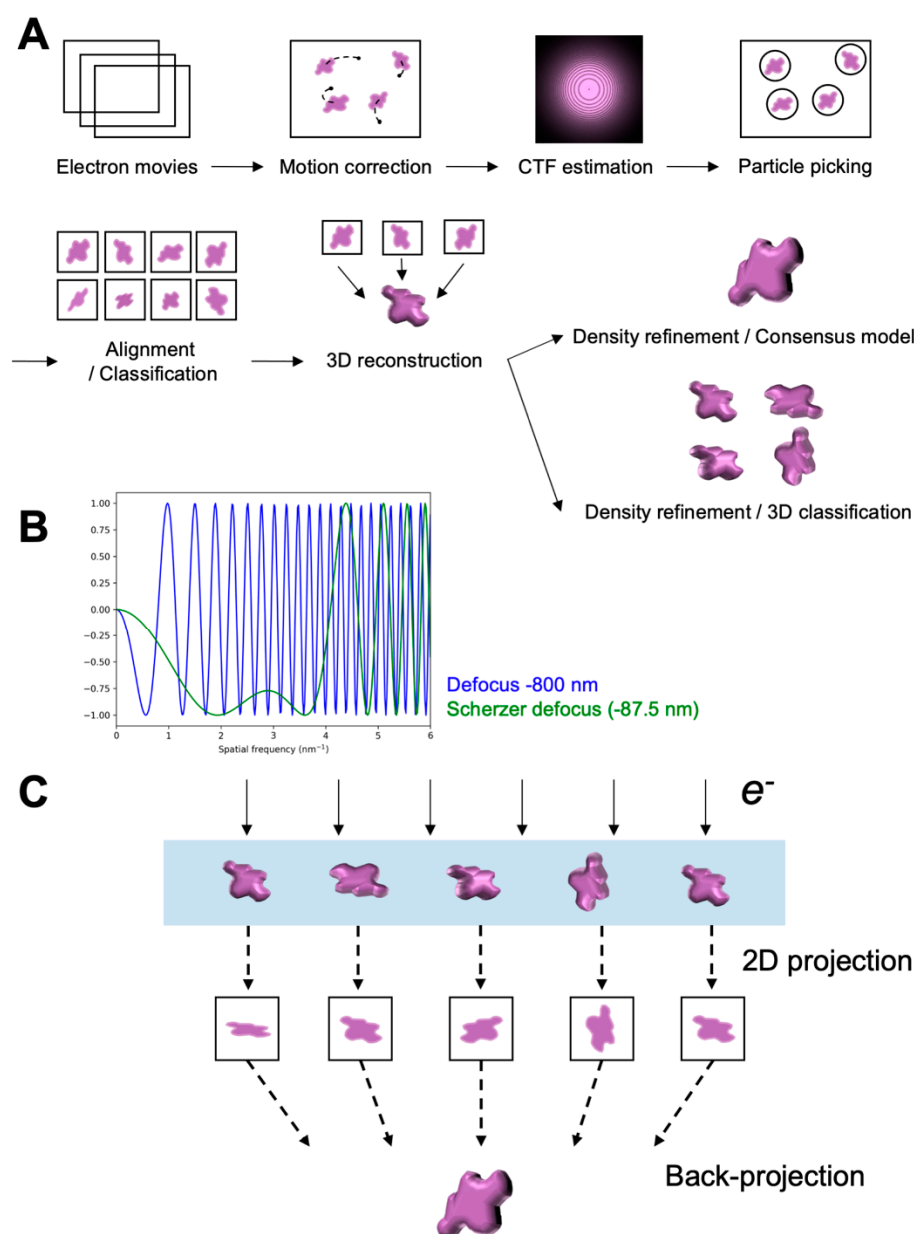
Figure 1. Schematics of the single-particle cryo-EM processing workflow. (**A**) General processing workflow for single-particle reconstruction workflow. (**B**) One-dimensional plot of the contrast transfer function (CTF) versus spatial frequency. Curves are calculated using an accelerated voltage of 300 keV and a spherical aberration coefficient (*Cs*) of 2.7 mm. Blue and green curves are for the defocus at −800 and −87.5 nm (Scherzer defocus), respectively. (**C**) Schematics of 2D projections from an imaged object and a 3D back-projection from the 2D projections. Light blue represents thin ice that embeds the particles in different orientations.

Thanks to the fast frame acquisition and high detective quantum efficiency (DQE) of the DED camera, it opens wide possibilities for preserving the structural information of the imaging targets in both spatial and temporal domains [30,31]. Compared to conventional detectors, such as photographic films or charge-coupled device (CCD) cameras, the DED camera directly detects the electrons and allows for a counting mechanism, which has a high-frequency response in the high-resolution region [32]. Additionally, upon electron beam exposure, the specimen moves due to the release of its mechanical strain, resulting in so-called beam-induced movements, which deteriorate the image quality [33–37]. The DED camera is capable of fast image acquisition to enable the recording of a movie, which can

be used to register the translational movements between each frame and obtain an average that improves the SNR of the object [32,38,39]. Furthermore, more improvements in the image quality due to the enhanced capabilities of the DED camera have been pushing the limits of cryo-EM, such as super-resolution [31] and dose weighting [38].

Molecular projections are selected from the electron images and are used to calculate a 3D Coulomb potential reconstruction of the target molecule [6]. Before reconstructing a 3D density volume, the 2D particle projections are aligned to a common origin. In addition to the conventional cross-correlation approach, the maximum-likelihood method is generally used to find the parameters, such as poses, that maximize the maximum likelihood estimates (MLE) [40–42]. In this framework, the expectation-maximization (EM) algorithm postulates the Gaussian probability models and calculates the MLE iteratively [42,43]. One of the disadvantages of the EM algorithm is the occurrence of the convergence to the nearby maximum, rather than the global maximum. Approximate Euler angles are initially assigned to individual 2D particle projections, and along the assigned projection direction, the 2D image intensities are used to integrate and back-project into a 3D volume (Figure 1C) [44]. The back-projection algorithm is based on the central projection theorem, which states that the Fourier transform of a 2D object projection is one slice of the 3D Fourier transform of the object when the Ewald curvature is flat [45]. The assigned Euler angles can be iteratively updated and improved by comparing the newly generated reference density. However, the Euler angle assignment for each particle projection could be challenging given that the SNR of cryo-EM data is extremely low. Therefore, using an unvalidated reference model, or attempting to select particles from a low SNR dataset based on this model, could lead to a biased map that largely resembles the model [46]. Thus, the image processing and the reconstructed model will need rigorous validation to ensure minimal reference bias. The resolution of the final 3D density map is cross-validated using the criteria of the golden Fourier-shell correlation (FSC) [47].

False positives are always generated when identifying particles in an image [48]. To minimize the false positives, the typical approach is to iteratively perform 2D image classification and select the classes with discernible molecular features [48]. However, due to the low SNR cryo-EM image and the error rate of the classification procedure, many particle images could be imprecisely aligned or wrongly assigned, leading to the removal of presentable particle projections. This could down-weight the particle projections in the intermediate states and flatten the final class averages. Therefore, the conventional curation approach may need to be reexamined to prevent the removal of true particle images and preserve the information for extracting the structural variability of the target molecule.

3. Sources of the Cryo-EM Image Data Heterogeneity

Biomolecules perform their function by changing their structures. The structural flexibility of a protein varies across different forms and mostly depends on its physical nature. However, molecules are always mobile because of thermal fluctuations [49]. When triggered by an environmental stimulus or a ligand, a biomolecule can conduct a structural change, globally or locally, and transition to a different energy state to perform its function [50]. Therefore, the proteins in vitreous ice may present different conformational states in an aqueous solution, which can be captured by cryo-EM imaging. The molecular simulation study showed that the fast-cooling process prevents the molecules from crossing the energy barriers, leading to the entrapment of molecules in a conformational ensemble [51]. A consensus 3D density map, a map back-projected using all of the particle images, could be used to estimate the signal certainties or voxel variances of the structural flexibility of the target molecule [52–54]. Local resolution estimation is the approach used to identify the global or local mobility of a protein. However, since the consensus map only presents a static view of the molecule, the estimated map does not contain the magnitude or the direction of motion of the protein local domains.

When mixing protein complex components in vitro for cryo-EM study, the molecules may form homo- or hetero-oligomers. Depending on the stability of the interactions

between proteins, the ensemble may contain different molecular species, leading to compositional heterogeneity in the particle population. Conventional 2D image classification could be used to sort out this compositional heterogeneity, for example, the p97 mutant complex in dodecameric and hexameric forms [55] (Figure 2). In some cases, if the angular distributions of individual molecular species are even for reconstructing their 3D volumes, the particle proportions could help understand the complex formation for the binding reaction.
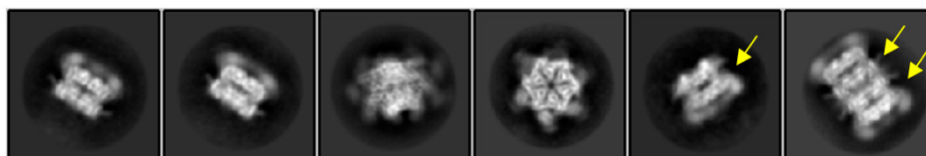

Figure 2. Two-dimensional (2D) class averages of cryo-EM particle images show the compositional heterogeneity. p97 R155H mutant complex [55]. Box side length is 374 Å. Yellow arrows indicate one single protein complex.

According to the dynamic nature of biomolecules, conformational transitions can be discrete, continuous, or disordered for their function. For instance, ATP synthase is a molecular motor complex that generates or consumes the energy currency, ATP, for sustaining a cell's life, which has been extensively studied using single-particle cryo-EM [56]. The cryo-EM reconstruction of ATP synthase typically generates three distinct states, which are in equilibrium in an aqueous solution [56]. Three different structures present the central stator rotating in various stages, with different proportions of particles in the population. Moreover, chloroplast ATP synthase activity can be modulated by the redox state of the central stator [57,58]. Interestingly, the three distinct rotating states could be resolved in their equilibrium states, but the proportions of the three states were different, implying the structural change in the complex changes the energy profiles of the three states [57] (Figure 3).
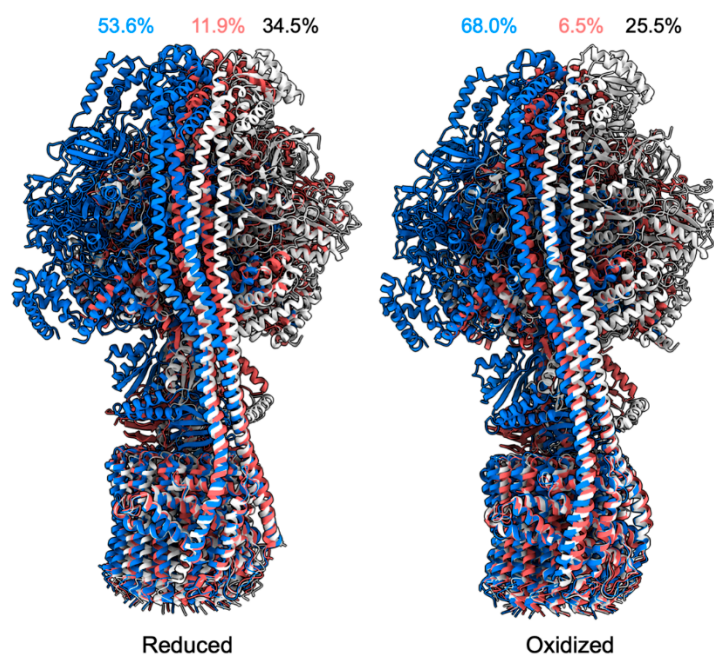

Figure 3. Redox states of chloroplast ATP synthase [57]. Blue, white, and salmon are different rotary states of the ATP synthase. Percentages labeled are the particle proportions within those in individual redox states.

Some biomolecules move continuously and do not exhibit distinct structures in the solution, which is not trivial to be categorized using conventional image classification to identify individual conformations on low-SNR cryo-EM images. In the cases of proteins that are disordered and do not present a structural feature, it is usually challenging to reconstruct high-resolution densities in these flexible regions by the conventional averaging approach.

4. Conventional Approaches to Study Cryo-EM Data Heterogeneity

One of the goals of single-particle reconstruction is to seek an optimal solution for Euler angular assignment to the 2D particle projection for a high-resolution 3D back-projection [7]. To obtain the dynamic information of a target molecule, the imaging will require capturing sufficient 2D projection data to present the structures in different states or different time points. The subsequent image classification scheme will be carried out to detect the structural differences and group the images into different classes in either a supervised or unsupervised manner. The variation between the 3D reconstructions from different groups could provide a dynamic view of the target protein.

4.1. Conventional Multivariate Statistical Analysis (MSA)

Multivariate statistical analysis (MSA) was first introduced to identify different particle views and average the class members to improve the SNR of the images of the negatively stained samples [59]. The fundamental concept is to find the orthogonal axes that can best categorize the image data into several groups and register individual particle images in a hyperspace [59,60]. This will require reducing the image dimensionality into several components, so-called dimension reduction [61]. Principal component analysis (PCA) is commonly used to determine the principal axes and eigenvectors. Correspondence analysis (CA) has also been used to create orthogonal axes and define the data in a relative manner, rather than the absolute Euclidean distance as in PCA [62]. Note that the efficiency of MSA is subjected to the SNR of the data and is highly dependent on the image alignment accuracy, which could be generally improved using an iterative method [59,63]. More EM image analyses using MSA can be found in the following works [59,64–66].

k-means clustering is the most used approach to categorize particle image data, which finds the centroids by starting random seeds and updating them iteratively until the optimal centroids are found [67]. An iterative algorithm could improve the clustering, but it could sometimes converge to a local minimum, complicating the reproducibility [48]. The number of k seeds is critical to the success of the classification, which is usually determined empirically. Additionally, k-means clustering is subjected to the SNR of the data. In the case of low-SNR data, the clustering algorithm usually produces clusters of similar sizes and does not calculate the centroids accurately, biased by the noise level. Moreover, suppose the number of seeds, k, is larger than the number of the representative classes in the data; in that case, some classes will collapse without any members assigned or the particles will be wrongly assigned into dominated classes, lowering the efficiency and accuracy of the clustering.

To mitigate the disadvantage of k-means clustering and prevent the collapse of the classes, some approaches have been developed to improve k-means clustering [68–72]. One of them is the iterative stable alignment and clustering (ISAC), which was proposed to utilize the Equal-Size group k-means (EQK-means) method that equalizes the size of individual classes to prevent the class from collapsing. The classes are then tested with reproducibility to stabilize the classification [68]. The ISAC method is usually applied to 2D image analysis and has the advantage of sorting out the classes that are rarely found in the particle population. Because the algorithm rigorously tests the reproducibility of the particle assignment to the classes, it may take a longer time to complete the classification when the data size is large.

The classification of the 3D volumes could be performed in a supervised manner using projection matching or MSA if the structures of multiple states are known [73,74]. The algorithms iteratively update the assigned Euler angles and classes of the 2D particle

images by comparing the projections of all input 3D volumes. Because prior knowledge of structures is required, the results are strongly dependent on the model and may be subjected to reference bias if the SNR of the data is low or proper cross-validation is not implemented [75].

4.2. Regularized Likelihood Approach

Maximum likelihood (ML) estimation has been long used in crystallographic refinement and single-particle EM in the field of structural biology [41,76]. In single-particle EM, Provencher and Vogel first used the ML algorithm in reconstructing the 3D volumes of biomolecules [77]. Sigworth implemented the ML algorithm in 2D image analysis and proposed the ML optimization for use in single-particle reconstruction [41]. Along with the same concept, Scheres and his coworkers developed a new ML3D (3D maximum likelihood) algorithm to classify single-particle image data for 3D reconstructions [78].

The ML3D algorithm is an expectation-maximization (E-M) algorithm, which seeks structural models that maximize the likelihood of the data by the marginalization over particle orientations and class assignments in an iterative manner [42,79]. Because the approach could converge in a local minimum, it is suggested to begin with a consensus model that is representative of the data when using ML3D [42]. Additionally, the ML3D algorithm uses a multi-reference refinement scheme, which may introduce some extents of reference bias. This will require a validation step by refining independent groups of the data. The current ML approach uses an iterative algorithm to obtain the solution, but this does not ensure convergence and may have a high unnecessary computation cost. On the other hand, the representable number of clusters in the data will need to be determined empirically. It may take a very long time computationally to optimize the class number, k. When k is overdetermined, the generated classes will not be representable for the data, whereas when k is underdetermined, the heterogeneous features may be diminished in the class averages. Additionally, because the class number k is assigned as an integer, presumably describing discrete states in the particle ensemble, it would be trivial to resolve the structural differences within the data or map the trajectory by connecting the resolved 3D class averages if continuous molecular movement is present.

In some cases, the flexible part is only a local portion of the molecule. One can use a smaller mask to limit the density refinement only carried out in the region of interest, improving the accuracy of the alignment and classification [80]. Furthermore, this local refinement could be improved by subtracting the densities that are not of interest, so-called signal subtraction [81]. However, the success of this local refinement procedure is subjected to the size of the selected or masked region, the angular distribution of the particle projections, and the SNR of the data. If the size of the selected region is too small, the accuracy of the image alignment may be compromised. When the target molecule has a concerted movement involving more than one region or domain, one can use a similar strategy to apply to multiple sites of the density and analyze their movements, so-called multi-body refinement [13].

Cossio and Hummer detailed the current methods that adapt the likelihood approach for analyzing cryo-EM data heterogeneity [82].

5. Continuous Structural Heterogeneity Derived from Cryo-EM Data

5.1. Covariance Matrix Estimation

The conventional way to analyze the data heterogeneity is to determine the top principal components or eigenvectors, which are mostly associated with molecular variabilities. However, it becomes challenging when resolving heterogeneity in 3D, but the data are provided in 2D. A recent study proposed a new covariance matrix estimator to analyze the relationships between unknown variables in 2D data and use this estimator to resolve the 3D heterogeneity [83]. The advantage of this approach is that no knowledge about prior distribution or data ranks is required. However, the SNR is still the limiting factor to its

effectiveness. More parameters, such as the CTF, may be required to take into consideration for qualitative comparison with traditional PCA in high dimensions.

5.2. Hyper-Molecules

The concept of the "hyper-molecules" is to add high-dimensional variables and map the continuous structural heterogeneity with a Bayesian formulation and use the Markov chain Monte Carlo (MCMC) algorithms to sample the posterior distribution without the need for normalization [84,85]. The advantage of the MCMC approach is to reduce the computational cost and have the capability to expand in combination with other algorithms to improve the performance. The algorithm has been applied on a simulated dataset to demonstrate its feasibility.

5.3. 3DVA (3D Variability Analysis) Approach

A 2D cryo-EM particle projection only presents a partial observation of the 3D structure of the object, and the structural information contained in the image is altered by the imaging. For example, the quality of the image phases depends on the CTF modulation and is also subjected to the noise level. Therefore, when conventional PCA is used to analyze the data for molecular motions, the number of dimensions may need to be increased, which requires numerous amounts of parameterization and complicates the computation. To overcome this drawback, the 3DVA (3D variability analysis) algorithm uses E-M optimization for probabilistic PCA (PPCA) to seek linear subspace models for single-particle image data [86–90]. PPCA generalizes the conventional PCA, when the covariance of the noise is very small. The probabilistic formulation provided by PPCA could improve the efficiency of finding the ML estimates from the principal components. To use the 3DVA algorithm, a precise image alignment for the consensus model is required. It resolves the data heterogeneity better when analyzing the data in a limited resolution range, which could be related to a better SNR level in the low spatial-frequency region. If the size of the flexible moiety is small, 3DVA still has its limitation to extract the variability components to accurately model this moiety that partakes in dramatic distance changes [90].

5.4. CryoDRGN

CryoDRGN (Deep Reconstructing Generative Networks) was designed to probe continuous 3D structural heterogeneity using a generative deep-learning approach [16]. The algorithm encodes the 2D projection data into a latent space using a variational autoencoder (VAE) and decodes slices of a series of 3D reconstructions. A VAE uses a probabilistic formalism to describe the data in a latent space, which can decode the generative models with continuity, or in a smoother way. The early version of the program uses the poses that are optimized for a consensus reconstruction [91], and the latest version includes the optimization of the poses while encoding the data into a latent space [91]. The program provides graphic representations of the projection data via either PCA or UMAP (uniform manifold approximation and projection) [16]. CryoDRGN has been shown to resolve motions of large biomolecular complexes, such as the 80S ribosome and spliceosome complex [16].

Because CryoDRGN adapts a deep neural network framework, the success of resolving continuous motion will be inherently limited by the parameters embedded within its construction. The latent space may need to be optimized empirically [16]. Additionally, overfitting may occur when the data SNR is low or the number of parameters is underdetermined. CryoDRGN utilizes latent space as a form of data reduction, which could be useful for determining the underlying flexibility of molecules but may not implicate any physical properties of the molecules, such as the energy landscape. Additionally, given the diversity of origin regarding heterogeneity within single-particle cryo-EM datasets, this latent space approach may not be representative for all forms of heterogeneity.

6. Mapping Energy Landscape from Cryo-EM Data

The domains within the protein do not move randomly but are constrained by neighboring domains or surrounding environments, and these movements, or conformational changes, usually accompany energy transitions. The proportions of the conformational states follow the Boltzmann distribution, but it is not trivial to reconstruct the energy landscape with individual conformational states from 2D cryo-EM projections, which are mostly noisy with a preferred Euler angle distribution.

Instead of conventional image classification, ManifoldEM was designed to map an energy landscape from the raw image data [92–94]. The method uses particles assigned with the same Euler angles and projects them onto a hyperspace. The particle projections in similar conformational states will more densely populate within a particular manifold in the hyperspace. The advantage of this method is that no preliminary classification is required, and it uses raw image data for the computation. Hence, this method will need a large number of datasets. The program showed promising results for extracting conformational variabilities from single-particle cryo-EM datasets of the 80S ribosome and ryanodine receptor, supported by atomistic simulation results as the ground truth [93].

In a deep-learning framework, AlphaCryo4D uses an energy-based particle-voting algorithm to encode 3D feature densities onto a Boltzmann-weighted embedding landscape [95]. Through searching local energy minima along a minimum-energy path, the 3D densities at transition states could be extracted and reconstructed [95]. This has been used to elucidate the transition of ubiquitin binding on the human 26S proteasome complex at different proposed functional states [95]. However, the method may not accurately analyze a small dataset, which may undersample the particle projections for 3D reconstruction at local minima [95].

The cryo-BIFE (cryo-EM Bayesian inference of free-energy profiles) method extracts the free energy profile using a lower dimensionality with collective variables (CV) involving Bayesian inference [96]. The method investigates the parameters of the collective variable, allowing to uncover the pertinent conformational states of a molecule. The method has been applied to calculate the free energy profile of TMEM16F ion channels and successfully showed two states that are bound with and without calcium ions [96].

7. Hybrid Approaches with Molecular Dynamic Simulations

The image quality resulting from averaging depends on the flexible nature of the target molecule, and it is common that the resolution of the cryo-EM density is not sufficient for accurately modeling the atoms globally or locally. If the target molecule is highly flexible, it may not be trivial to determine a high-resolution structure using the typical single-particle cryo-EM approach since the density of the mobile region will be flattened due to averaging. However, the low-resolution cryo-EM density provides some positional uncertainties for the atoms, which could be useful in understanding the regional flexibility. Low local resolution can be an indicator of a highly flexible region. If the cryo-EM density map has secondary structural features at a medium resolution (4–6 Å), such as α helices or β sheets, one could build the protein backbone to obtain an overall shape of the molecule. Using cryo-EM density as a constraint, molecular dynamics (MD) simulations can be used to understand the movements of domains or residues, which are powerful approaches to studying the structural perturbation and energetics of biomolecules during their functional transitions [97]. Combined with 3D image classification, the MD trajectories derived from the cryo-EM densities could improve the understanding of the continuous motion of the target molecule.

7.1. Detecting Structural Variability Based on the Resolution Anisotropy

Because the conformational change of a biological molecule usually occurs regionally, the local resolution estimation could indicate the level of structural flexibility within the molecule [98,99]. If the atomic coordinates can be obtained from other experiments or the public depositories, such as Protein Data Bank (PDB), one can use the coordinates for

an all-atom molecular simulation for training a model for dynamic information, which is implemented in the DEFMap program [100]. The structural variability of the target molecule can then be directly derived from the experimental cryo-EM image data. The advantage of this method is to reduce the computation cost dramatically compared to conventional all-atom molecular simulations because part of the dynamic information is obtained from the cryo-EM image data.

7.2. Molecular Dynamics Flexible Fitting (MDFF)

In most cases, single-particle reconstruction can only reach a medium resolution range (4–6 Å), which leaves spaces of uncertainty for modeling atomic coordinates. If a template of the atomic model exists, rigid-body fitting could be the first attempt to fit the model into the map, providing knowledge of domain organization in the target protein [101–103]. One can also use a cryo-EM density map as a geometric constraint and perform molecular dynamics flexible fitting (MDFF) to seek the best fit of the atomic model [104,105]. MDFF provides a theoretical framework for model building within a low-resolution cryo-EM map. The MDFF-fitted models built from cryo-EM maps of different states can be used to identify the domains or regions for protein conformational changes. Moreover, another molecular dynamics approach for cryo-EM model building is to optimize the real-space correlation between the model and experimental map based on the theoretical force fields, which results in a model close to the observation map without overfitting [106,107]. The approach is so-called correlation-driven molecular dynamics (CDMD), which has proven to be robust in modeling atomic coordinates in cryo-EM maps at a wide resolution range [106].

8. Time-Resolved Cryo-EM Studies

Because the protein particles are frozen simultaneously when prepared for cryo-EM imaging, the structures resolved in different states by single-particle cryo-EM do not contain any information in the time domain. Time-resolved cryo-EM aims to use cryo-EM to image or sample a reaction over a period of time until the reaction is in equilibrium. If the reaction is slow, the sampling time interval can be in minutes, allowing for the preparation of a frozen grid of a protein metastate [108,109]. However, if the reaction is fast, the frozen samples at different time points could be obtained using either microfluidic devices that allow the mixing of two reactants to occur or spraying/mixing methods in which a grid is covered with one reactant and the other is sprayed on [110]. The small volumes of samples at different reaction coordinates are 'written' or sprayed onto the grid for imaging, and the subsequent single-particle image classification and reconstruction will be utilized to analyze the molecular interactions or protein conformational changes. The recent development of time-resolved cryo-EM can be found in the following reviews and articles [110–114].

9. Interpretation of the Extracted Information for Biomolecular Perturbation

The molecular motions derived from cryo-EM data could be sampled in different 3D reconstructions from the above-mentioned image classification or data analysis. Because the molecules are all frozen simultaneously, the movements between distinct states may not have a causal connection. Structural variations between the reconstructions can be seen as conformational changes that occur with small energy transitions, and the proportions are weighted according to the Boltzmann distribution. During image processing, discrete motions could be observed to exhibit small subsets of dominant poses. However, continuous motions may arise when a molecule exhibits a high degree of flexibility and can adopt a large number of conformations with nearly equivalent weights in the Boltzmann distribution, which may not be trivial to analyze the data when the SNR is notoriously low.

The interpretation of protein motions learned from cryo-EM data will need to be validated using other biophysical or biochemical methods, such as cross-linking or molecular simulations. In particular, over-interpretation may be encountered when the motion does not relate to any function of the target protein.

10. Summary and Outlook

This review discusses different methods of analyzing the cryo-EM data heterogeneity and extracting structural variability information from the generated models. If a discrete distribution of the conformational states can be determined within the particle population, the validities of individual reconstructions could be verified using a standard approach, such as the gold standard FSC, with their data subsets. However, 3D reconstructions for continuous motion still require a robust validation method to be developed. Additionally, the conventional approach for particle curation may need to be revisited to prevent removing particle images that present the intermediate states. Cryo-EM opens a new route for analyzing structural variabilities of biomolecules and studying molecular dynamics, which provide insight into molecular mechanisms. The development of new data analysis methods will allow cryo-EM to facilitate the discovery of novel dynamic information about biomolecules.

Author Contributions: Conceptualization, P.-L.C.; investigation, K.D. and P.-L.C.; writing—original draft preparation, K.D. and P.-L.C.; writing—review and editing, K.D. and P.-L.C.; visualization, K.D. and P.-L.C.; supervision, P.-L.C.; funding acquisition, P.-L.C. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Callaway, E. Revolutionary Cryo-EM Is Taking over Structural Biology. *Nature* **2020**, *578*, 201. [CrossRef] [PubMed]
2. Bai, X.-C.; McMullan, G.; Scheres, S.H.W. How Cryo-EM Is Revolutionizing Structural Biology. *Trends Biochem. Sci.* **2015**, *40*, 49–57. [CrossRef] [PubMed]
3. Guerrini, N.; Turchetta, R.; Van Hoften, G.; Henderson, R.; McMullan, G.; Faruqi, A.R. A High Frame Rate, 16 Million Pixels, Radiation Hard CMOS Sensor. *J. Instrum.* **2011**, *6*, C03003. [CrossRef]
4. Mooney, P.; Contarato, D.; Denes, P.; Gubbens, A.; Lee, B.; Lent, M.; Agard, D. A High-Speed Electron-Counting Direct Detection Camera for TEM. *Microsc. Microanal.* **2011**, *17*, 1004–1005. [CrossRef]
5. Milazzo, A.-C.; Moldovan, G.; Lanman, J.; Jin, L.; Bouwer, J.C.; Klienfelder, S.; Peltier, S.T.; Ellisman, M.H.; Kirkland, A.I.; Xuong, N.-H. Characterization of a Direct Detection Device Imaging Camera for Transmission Electron Microscopy. *Ultramicroscopy* **2010**, *110*, 744–747. [CrossRef] [PubMed]
6. Frank, J. *Three-Dimensional Electron. Microscopy of Macromolecular Assemblies: Visualization of Biological Molecules in Their Native State*; Oxford University Press: New York, NY, USA, 2006; ISBN 9780198034384.
7. Van Heel, M. Angular Reconstitution: A Posteriori Assignment of Projection Directions for 3D Reconstruction. *Ultramicroscopy* **1987**, *21*, 111–123. [CrossRef]
8. Glaeseral, R.M. Radiation Damage and High Resolution Biological Electron Microscopy. *Proc. Annu. Meet. Electron. Microsc. Soc. Am.* **1973**, *31*, 226–227. [CrossRef]
9. Isaacson, M.; Johnson, D.; Crewe, A.V. Electron Beam Excitation and Damage of Biological Molecules; Its Implications for Specimen Damage in Electron Microscopy. *Radiat. Res.* **1973**, *55*, 205–224. [CrossRef]
10. Glaeser, R.M.; Taylor, K.A. Radiation Damage Relative to Transmission Electron Microscopy of Biological Specimens at Low Temperature: A Review. *J. Microsc.* **1978**, *112*, 127–138. [CrossRef]
11. Hayward, S.B.; Glaeser, R.M. Radiation Damage of Purple Membrane at Low Temperature. *Ultramicroscopy* **1979**, *4*, 201–210. [CrossRef]
12. Cheng, A.; Eng, E.T.; Alink, L.; Rice, W.J.; Jordan, K.D.; Kim, L.Y.; Potter, C.S.; Carragher, B. High Resolution Single Particle Cryo-Electron Microscopy Using Beam-Image Shift. *J. Struct. Biol.* **2018**, *204*, 270–275. [CrossRef] [PubMed]
13. Nakane, T.; Kimanius, D.; Lindahl, E.; Scheres, S.H. Characterisation of Molecular Motions in Cryo-EM Single-Particle Data by Multi-Body Refinement in RELION. *Elife* **2018**, *7*, e36861. [CrossRef] [PubMed]
14. Mitra, K.; Frank, J. Ribosome Dynamics: Insights from Atomic Structure Modeling into Cryo-Electron Microscopy Maps. *Annu. Rev. Biophys. Biomol. Struct.* **2006**, *35*, 299–317. [CrossRef] [PubMed]

15. Tama, F.; Valle, M.; Frank, J.; Brooks, C.L., 3rd. Dynamic Reorganization of the Functionally Active Ribosome Explored by Normal Mode Analysis and Cryo-Electron Microscopy. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 9319–9323. [CrossRef] [PubMed]

16. Zhong, E.D.; Bepler, T.; Berger, B.; Davis, J.H. CryoDRGN: Reconstruction of Heterogeneous Cryo-EM Structures Using Neural Networks. *Nat. Methods* **2021**, *18*, 176–185. [CrossRef] [PubMed]

17. Fica, S.M.; Nagai, K. Cryo-Electron Microscopy Snapshots of the Spliceosome: Structural Insights into a Dynamic Ribonucleoprotein Machine. *Nat. Struct. Mol. Biol.* **2017**, *24*, 791–799. [CrossRef] [PubMed]

18. Haselbach, D.; Komarov, I.; Agafonov, D.E.; Hartmuth, K.; Graf, B.; Dybkov, O.; Urlaub, H.; Kastner, B.; Lührmann, R.; Stark, H. Structure and Conformational Dynamics of the Human Spliceosomal Bact Complex. *Cell* **2018**, *172*, 454–464. [CrossRef]

19. Marino, J.; Schertler, G.F.X. A Set of Common Movements within GPCR-G-Protein Complexes from Variability Analysis of Cryo-EM Datasets. *J. Struct. Biol.* **2021**, *213*, 107699. [CrossRef]

20. Dong, M.; Deganutti, G.; Piper, S.J.; Liang, Y.-L.; Khoshouei, M.; Belousoff, M.J.; Harikumar, K.G.; Reynolds, C.A.; Glukhova, A.; Furness, S.G.B.; et al. Structure and Dynamics of the Active Gs-Coupled Human Secretin Receptor. *Nat. Commun.* **2020**, *11*, 4137. [CrossRef]

21. Josephs, T.M.; Belousoff, M.J.; Liang, Y.-L.; Piper, S.J.; Cao, J.; Garama, D.J.; Leach, K.; Gregory, K.J.; Christopoulos, A.; Hay, D.L.; et al. Structure and Dynamics of the CGRP Receptor in Apo and Peptide-Bound Forms. *Science* **2021**, *372*, eabf7258. [CrossRef]

22. Hilger, D. The Role of Structural Dynamics in GPCR-Mediated Signaling. *FEBS J.* **2021**, *288*, 2461–2489. [CrossRef] [PubMed]

23. Chiu, P.-L.; Kelly, D.F.; Walz, T. The Use of Trehalose in the Preparation of Specimens for Molecular Electron Microscopy. *Micron* **2011**, *42*, 762–772. [CrossRef] [PubMed]

24. Dubochet, J.; McDowall, A.W. Vitrification of Pure Water for Electron Microscopy. *J. Microsc.* **1981**, *124*, 3–4. [CrossRef]

25. Adrian, M.; Dubochet, J.; Lepault, J.; McDowall, A.W. Cryo-Electron Microscopy of Viruses. *Nature* **1984**, *308*, 32–36. [CrossRef]

26. Cowley, J.M. Image Contrast in a Transmission Scanning Electron Microscope. *Appl. Phys. Lett.* **1969**, *15*, 58–59. [CrossRef]

27. Spence, J.C.H. *High-Resolution Electron Microscopy*; OUP Oxford University Press: Oxford, UK, 2013; ISBN 9780191508400.

28. Erickson, H.; Klug, A. Measurement and Compensation of Defocusing and Aberrations by Fourier Processing of Electron Micrographs. *Philos. Trans. R. Soc. Lond.* **1971**, *261*, 105–118. [CrossRef]

29. Wade, R.H. A Brief Look at Imaging and Contrast Transfer. *Ultramicroscopy* **1992**, *46*, 145–156. [CrossRef]

30. Wu, S.; Armache, J.-P.; Cheng, Y. Single-Particle Cryo-EM Data Acquisition by Using Direct Electron Detection Camera. *Microscopy* **2016**, *65*, 35–41. [CrossRef]

31. Chiu, P.-L.; Li, X.; Li, Z.; Beckett, B.; Brilot, A.F.; Grigorieff, N.; Agard, D.A.; Cheng, Y.; Walz, T. Evaluation of Super-Resolution Performance of the K2 Electron-Counting Camera Using 2D Crystals of Aquaporin-0. *J. Struct. Biol.* **2015**, *192*, 163–173. [CrossRef]

32. Li, X.; Mooney, P.; Zheng, S.; Booth, C.R.; Braunfeld, M.B.; Gubbens, S.; Agard, D.A.; Cheng, Y. Electron Counting and Beam-Induced Motion Correction Enable near-Atomic-Resolution Single-Particle Cryo-EM. *Nat. Methods* **2013**, *10*, 584–590. [CrossRef] [PubMed]

33. D'Imprima, E.; Kühlbrandt, W. Current Limitations to High-Resolution Structure Determination by Single-Particle CryoEM. *Q. Rev. Biophys.* **2021**, *54*, e4. [CrossRef] [PubMed]

34. Vinothkumar, K.R.; Henderson, R. Single Particle Electron Cryomicroscopy: Trends, Issues and Future Perspective. *Q. Rev. Biophys.* **2016**, *49*, e13. [CrossRef]

35. Russo, C.J.; Henderson, R. Charge Accumulation in Electron Cryomicroscopy. *Ultramicroscopy* **2018**, *187*, 43–49. [CrossRef] [PubMed]

36. Russo, C.J.; Henderson, R. Microscopic Charge Fluctuations Cause Minimal Contrast Loss in CryoEM. *Ultramicroscopy* **2018**, *187*, 56–63. [CrossRef]

37. Glaeser, R.M.; Han, B.-G.; Csencsits, R.; Killilea, A.; Pulk, A.; Cate, J.H.D. Factors That Influence the Formation and Stability of Thin, Cryo-EM Specimens. *Biophys. J.* **2016**, *110*, 749–755. [CrossRef]

38. Grant, T.; Grigorieff, N. Measuring the Optimal Exposure for Single Particle Cryo-EM Using a 2.6 Å Reconstruction of Rotavirus VP6. *Elife* **2015**, *4*, e06980. [CrossRef]

39. Zheng, S.Q.; Palovcak, E.; Armache, J.-P.; Verba, K.A.; Cheng, Y.; Agard, D.A. MotionCor2: Anisotropic Correction of Beam-Induced Motion for Improved Cryo-Electron Microscopy. *Nat. Methods* **2017**, *14*, 331–332. [CrossRef]

40. Sigworth, F.J.; Doerschuk, P.C.; Carazo, J.-M.; Scheres, S.H.W. An Introduction to Maximum-Likelihood Methods in Cryo-EM. *Methods Enzymol.* **2010**, *482*, 263–294. [CrossRef] [PubMed]

41. Sigworth, F.J. A Maximum-Likelihood Approach to Single-Particle Image Refinement. *J. Struct. Biol.* **1998**, *122*, 328–339. [CrossRef] [PubMed]

42. Scheres, S.H.W. Classification of Structural Heterogeneity by Maximum-Likelihood Methods. *Methods Enzymol.* **2010**, *482*, 295–320. [CrossRef] [PubMed]

43. Dempster, A.P.; Laird, N.M.; Rubin, D.B. Maximum Likelihood from Incomplete Data via the *EM* Algorithm. *J. R. Stat. Soc.* **1977**, *39*, 1–22. [CrossRef]

44. Radermacher, M. Weighted Back-Projection Methods. In *Electron Tomography*; Springer: New York, NY, USA, 2008; pp. 245–273, ISBN 9780387312347.

45. Nogales, E.; Scheres, S.H.W. Cryo-EM: A Unique Tool for the Visualization of Macromolecular Complexity. *Mol. Cell* **2015**, *58*, 677–689. [CrossRef]

46. Henderson, R. Avoiding the Pitfalls of Single Particle Cryo-Electron Microscopy: Einstein from Noise. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 18037–18041. [CrossRef]

47. Scheres, S.H.W.; Chen, S. Prevention of Overfitting in Cryo-EM Structure Determination. *Nat. Methods* **2012**, *9*, 853–854. [CrossRef] [PubMed]

48. Penczek, P.A. Fundamentals of Three-Dimensional Reconstruction from Projections. In *Methods in Enzymology*; Elsevier: Amsterdam, The Netherlands, 2010; pp. 1–33.

49. Allen, T.W.; Andersen, O.S.; Roux, B. On the Importance of Atomic Fluctuations, Protein Flexibility, and Solvent in Ion Permeation. *J. Gen. Physiol.* **2004**, *124*, 679–690. [CrossRef]

50. Skjaerven, L.; Reuter, N.; Martinez, A. Dynamics, Flexibility and Ligand-Induced Conformational Changes in Biological Macromolecules: A Computational Approach. *Future Med. Chem.* **2011**, *3*, 2079–2100. [CrossRef] [PubMed]

51. Bock, L.V.; Grubmüller, H. Effects of Cryo-EM Cooling on Structural Ensembles. *Nat. Commun.* **2022**, *13*, 1709. [CrossRef]

52. Swint-Kruse, L.; Brown, C.S. Resmap: Automated Representation of Macromolecular Interfaces as Two-Dimensional Networks. *Bioinformatics* **2005**, *21*, 3327–3328. [CrossRef] [PubMed]

53. Aiyer, S.; Zhang, C.; Baldwin, P.R.; Lyumkis, D. Evaluating Local and Directional Resolution of Cryo-EM Density Maps. *Methods Mol. Biol.* **2021**, *2215*, 161–187. [CrossRef]

54. Local Resolution of Cryo-EM Maps with MonoRes. *Nat. Methods* **2018**, *15*, 246. [CrossRef]

55. Nandi, P.; Li, S.; Columbres, R.C.A.; Wang, F.; Williams, D.R.; Poh, Y.-P.; Chou, T.-F.; Chiu, P.-L. Structural and Functional Analysis of Disease-Linked P97 ATPase Mutant Complexes. *Int. J. Mol. Sci.* **2021**, *22*, 8079. [CrossRef] [PubMed]

56. Guo, H.; Rubinstein, J.L. Cryo-EM of ATP Synthases. *Curr. Opin. Struct. Biol.* **2018**, *52*, 71–79. [CrossRef]

57. Yang, J.-H.; Williams, D.; Kandiah, E.; Fromme, P.; Chiu, P.-L. Structural Basis of Redox Modulation on Chloroplast ATP Synthase. *Commun. Biol.* **2020**, *3*, 482. [CrossRef] [PubMed]

58. Hisabori, T.; Sunamura, E.-I.; Kim, Y.; Konno, H. The Chloroplast ATP Synthase Features the Characteristic Redox Regulation Machinery. *Antioxid. Redox Signal.* **2013**, *19*, 1846–1854. [CrossRef] [PubMed]

59. van Heel, M.; Frank, J. Use of Multivariate Statistics in Analysing the Images of Biological Macromolecules. *Ultramicroscopy* **1981**, *6*, 187–194. [CrossRef]

60. van Heel, M. Multivariate Statistical Classification of Noisy Images (Randomly Oriented Biological Macromolecules). *Ultramicroscopy* **1984**, *13*, 165–183. [CrossRef]

61. Van Der Maaten, L.; Postma, E. Dimensionality Reduction: A Comparative Review. *J. Mach. Learn. Res.* **2009**, *10*, 66–71.

62. Van Heel, M.F.J. *Classification of Particles in Noisy Electron. Micrographs Using Correspondence Analysis. Pattern Recognition in Practice*; North-Holland Publishing: Amsterdam, The Netherlands, 1980.

63. Frank, J.; van Heel, M. Correspondence Analysis of Aligned Images of Biological Particles. *J. Mol. Biol.* **1982**, *161*, 134–137. [CrossRef]

64. Frank, J. Differential Averaging of Single Molecule Images Using Multivariate Statistical Classification. *Proc. Annu. Meet. Electron. Microsc. Soc. Am.* **1982**, *40*, 706–709. [CrossRef]

65. van Heel, M.; Portugal, R.V.; Schatz, M. Multivariate Statistical Analysis of Large Datasets: Single Particle Electron Microscopy. *Open J. Stat.* **2016**, *06*, 701–739. [CrossRef]

66. Harauz, G.; Welsh, E.A. Multivariate Statistical Analysis of Electron Micrographs of a Mammalian Transcription Initiation Complex. *J. Electron. Microsc.* **1992**, *41*, 264–266. [CrossRef]

67. Likas, A.; Vlassis, N.J.; Verbeek, J. The Global K-Means Clustering Algorithm. *Pattern Recognit.* **2003**, *36*, 451–461. [CrossRef]

68. Yang, Z.; Fang, J.; Chittuluru, J.; Asturias, F.J.; Penczek, P.A. Iterative Stable Alignment and Clustering of 2D Transmission Electron Microscope Images. *Structure* **2012**, *20*, 237–247. [CrossRef] [PubMed]

69. Sorzano, C.O.S.; Bilbao-Castro, J.R.; Shkolnisky, Y.; Alcorlo, M.; Melero, R.; Caffarena-Fernández, G.; Li, M.; Xu, G.; Marabini, R.; Carazo, J.M. A Clustering Approach to Multireference Alignment of Single-Particle Projections in Electron Microscopy. *J. Struct. Biol.* **2010**, *171*, 197–206. [CrossRef] [PubMed]

70. Wu, J.; Ma, Y.-B.; Congdon, C.; Brett, B.; Chen, S.; Xu, Y.; Ouyang, Q.; Mao, Y. Massively Parallel Unsupervised Single-Particle Cryo-EM Data Clustering via Statistical Manifold Learning. *PLoS ONE* **2017**, *12*, e0182130. [CrossRef]

71. Chen, T.-L.; Hsieh, D.-N.; Hung, H.; Tu, I.-P.; Wu, P.-S.; Wu, Y.-M.; Chang, W.-H.; Huang, S.-Y. γ-SUP: A Clustering Algorithm for Cryo-Electron Microscopy Images of Asymmetric Particles. *Ann. Appl. Stat.* **2014**, *8*, 259–285. [CrossRef]

72. Rao, R.; Moscovich, A.; Singer, A. Wasserstein K-Means for Clustering Tomographic Projections. *arXiv* **2020**, arXiv:2010.09989. [CrossRef]

73. Ludtke, S.J.; Baldwin, P.R.; Chiu, W. EMAN: Semiautomated Software for High-Resolution Single-Particle Reconstructions. *J. Struct. Biol.* **1999**, *128*, 82–97. [CrossRef]

74. Frank, J.; Radermacher, M.; Penczek, P.; Zhu, J.; Li, Y.; Ladjadj, M.; Leith, A. SPIDER and WEB: Processing and Visualization of Images in 3D Electron Microscopy and Related Fields. *J. Struct. Biol.* **1996**, *116*, 190–199. [CrossRef] [PubMed]

75. Rosenthal, P.B.; Henderson, R. Optimal Determination of Particle Orientation, Absolute Hand, and Contrast Loss in Single-Particle Electron Cryomicroscopy. *J. Mol. Biol.* **2003**, *333*, 721–745. [CrossRef]

76. Pannu, N.S.; Read, R.J. Improved Structure Refinement through Maximum Likelihood. *Acta Crystallogr. A* **1996**, *52*, 659–668. [CrossRef]

77. Provencher, S.W.; Vogel, R.H. Three-Dimensional Reconstruction from Electron Micrographs of Disordered Specimens. I. Method. *Ultramicroscopy* **1988**, *25*, 209–221. [CrossRef]

78. Scheres, S.H.W.; Gao, H.; Valle, M.; Herman, G.T.; Eggermont, P.P.B.; Frank, J.; Carazo, J.-M. Disentangling Conformational States of Macromolecules in 3D-EM through Likelihood Optimization. *Nat. Methods* **2007**, *4*, 27–29. [CrossRef] [PubMed]

79. Scheres, S.H.W. RELION: Implementation of a Bayesian Approach to Cryo-EM Structure Determination. *J. Struct. Biol.* **2012**, *180*, 519–530. [CrossRef]

80. Rawson, S.; Iadanza, M.G.; Ranson, N.A.; Muench, S.P. Methods to Account for Movement and Flexibility in Cryo-EM Data Processing. *Methods* **2016**, *100*, 35–41. [CrossRef]

81. Bai, X.-C.; Rajendra, E.; Yang, G.; Shi, Y.; Scheres, S.H.W. Sampling the Conformational Space of the Catalytic Subunit of Human γ-Secretase. *eLife* **2015**, *4*, e11182. [CrossRef]

82. Cossio, P.; Hummer, G. Likelihood-Based Structural Analysis of Electron Microscopy Images. *Curr. Opin. Struct. Biol.* **2018**, *49*, 162–168. [CrossRef]

83. Katsevich, G.; Katsevich, A.; Singer, A. Covariance Matrix Estimation for the Cryo-EM Heterogeneity Problem. *arXiv* **2013**, arXiv:1309.1737. [CrossRef]

84. Lederman, R. Hyper-Molecules: High Dimensional Maps of Molecular Conformations. *Acta Crystallogr. A Found. Adv.* **2020**, *76*, a61. [CrossRef]

85. Lederman, R.R.; Andén, J.; Singer, A. Hyper-Molecules: On the Representation and Recovery of Dynamical Structures for Applications in Flexible Macro-Molecules in Cryo-EM. *Inverse Probl.* **2020**, *36*, 044005. [CrossRef]

86. Tagare, H.D.; Kucukelbir, A.; Sigworth, F.J.; Wang, H.; Rao, M. Directly Reconstructing Principal Components of Heterogeneous Particles from Cryo-EM Images. *J. Struct. Biol.* **2015**, *191*, 245–262. [CrossRef]

87. Neal, R.M.; Hinton, G.E. A View of the Em Algorithm That Justifies Incremental, Sparse, and Other Variants. In *Learning in Graphical Models*; Springer: Dordrecht, The Netherlands, 1998; pp. 355–368, ISBN 9789401061049.

88. Tipping, M.E.; Bishop, C.M. Probabilistic Principal Component Analysis. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **1999**, *61*, 611–622. [CrossRef]

89. Roweis, S. EM Algorithms for PCA and SPCA. In *Advances in Neural Information Processing Systems*; Jordan, M., Kearns, M., Solla, S., Eds.; MIT Press: Cambridge, MA, USA, 1997; Volume 10.

90. Punjani, A.; Fleet, D.J. 3D Variability Analysis: Resolving Continuous Flexibility and Discrete Heterogeneity from Single Particle Cryo-EM. *J. Struct. Biol.* **2021**, *213*, 107702. [CrossRef]

91. Zhong, E.D.; Lerer, A.; Davis, J.H.; Berger, B. CryoDRGN2: Ab Initio Neural Reconstruction of 3D Protein Structures from Real Cryo-EM Images. In Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), IEEE Montreal, Montreal, QC, Canada, 10–17 October 2021; pp. 4066–4075.

92. Dashti, A.; Schwander, P.; Langlois, R.; Fung, R.; Li, W.; Hosseinizadeh, A.; Liao, H.Y.; Pallesen, J.; Sharma, G.; Stupina, V.A.; et al. Trajectories of the Ribosome as a Brownian Nanomachine. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 17492–17497. [CrossRef]

93. Dashti, A.; Mashayekhi, G.; Shekhar, M.; Ben Hail, D.; Salah, S.; Schwander, P.; des Georges, A.; Singharoy, A.; Frank, J.; Ourmazd, A. Retrieving Functional Pathways of Biomolecules from Single-Particle Snapshots. *Nat. Commun.* **2020**, *11*, 4734. [CrossRef] [PubMed]

94. Frank, J.; Ourmazd, A. Continuous Changes in Structure Mapped by Manifold Embedding of Single-Particle Data in Cryo-EM. *Methods* **2016**, *100*, 61–67. [CrossRef] [PubMed]

95. Wu, Z.; Zhang, S.; Wang, W.L.; Ma, Y.; Dong, Y.; Mao, Y. Deep Manifold Learning Reveals Hidden Dynamics of Proteasome Autoregulation. *arXiv* **2021**, arXiv:2012.12854. [CrossRef]

96. Giraldo-Barreto, J.; Ortiz, S.; Thiede, E.H.; Palacio-Rodriguez, K.; Carpenter, B.; Barnett, A.H.; Cossio, P. A Bayesian Approach to Extracting Free-Energy Profiles from Cryo-Electron Microscopy Experiments. *Sci. Rep.* **2021**, *11*, 13657. [CrossRef] [PubMed]

97. Karplus, M.; McCammon, J.A. Molecular Dynamics Simulations of Biomolecules. *Nat. Struct. Biol.* **2002**, *9*, 646–652. [CrossRef] [PubMed]

98. Vilas, J.L.; Gómez-Blanco, J.; Conesa, P.; Melero, R.; Miguel de la Rosa-Trevín, J.; Otón, J.; Cuenca, J.; Marabini, R.; Carazo, J.M.; Vargas, J.; et al. MonoRes: Automatic and Accurate Estimation of Local Resolution for Electron Microscopy Maps. *Structure* **2018**, *26*, 337–344. [CrossRef] [PubMed]

99. Kucukelbir, A.; Sigworth, F.J.; Tagare, H.D. Quantifying the Local Resolution of Cryo-EM Density Maps. *Nat. Methods* **2014**, *11*, 63–65. [CrossRef] [PubMed]

100. Matsumoto, S.; Ishida, S.; Araki, M.; Kato, T.; Terayama, K.; Okuno, Y. Extraction of Protein Dynamics Information from Cryo-EM Maps Using Deep Learning. *Nat. Mach. Intell.* **2021**, *3*, 153–160. [CrossRef]

101. Wriggers, W.; Birmanns, S. Using Situs for Flexible and Rigid-Body Fitting of Multiresolution Single-Molecule Data. *J. Struct. Biol.* **2001**, *133*, 193–202. [CrossRef] [PubMed]

102. McGreevy, R.; Teo, I.; Singharoy, A.; Schulten, K. Advances in the Molecular Dynamics Flexible Fitting Method for Cryo-EM Modeling. *Methods* **2016**, *100*, 50–60. [CrossRef]

103. Miyashita, O.; Kobayashi, C.; Mori, T.; Sugita, Y.; Tama, F. Flexible Fitting to Cryo-EM Density Map Using Ensemble Molecular Dynamics Simulations. *J. Comput. Chem.* **2017**, *38*, 1447–1461. [CrossRef] [PubMed]

104. Kulik, M.; Mori, T.; Sugita, Y. Multi-Scale Flexible Fitting of Proteins to Cryo-EM Density Maps at Medium Resolution. *Front. Mol. Biosci.* **2021**, *8*, 631854. [CrossRef] [PubMed]

105. Trabuco, L.G.; Villa, E.; Schreiner, E.; Harrison, C.B.; Schulten, K. Molecular Dynamics Flexible Fitting: A Practical Guide to Combine Cryo-Electron Microscopy and X-ray Crystallography. *Methods* **2009**, *49*, 174–180. [CrossRef] [PubMed]
106. Igaev, M.; Kutzner, C.; Bock, L.V.; Vaiana, A.C.; Grubmüller, H. Automated Cryo-EM Structure Refinement Using Correlation-Driven Molecular Dynamics. *eLife* **2019**, *8*, e43542. [CrossRef]
107. Orzechowski, M.; Tama, F. Flexible Fitting of High-Resolution X-ray Structures into Cryoelectron Microscopy Maps Using Biased Molecular Dynamics Simulations. *Biophys. J.* **2008**, *95*, 5692–5705. [CrossRef] [PubMed]
108. Mulder, A.M.; Yoshioka, C.; Beck, A.H.; Bunner, A.E.; Milligan, R.A.; Potter, C.S.; Carragher, B.; Williamson, J.R. Visualizing Ribosome Biogenesis: Parallel Assembly Pathways for the 30S Subunit. *Science* **2010**, *330*, 673–677. [CrossRef]
109. Fischer, N.; Konevega, A.L.; Wintermeyer, W.; Rodnina, M.V.; Stark, H. Ribosome Dynamics and TRNA Movement by Time-Resolved Electron Cryomicroscopy. *Nature* **2010**, *466*, 329–333. [CrossRef] [PubMed]
110. Frank, J. Time-Resolved Cryo-Electron Microscopy: Recent Progress. *J. Struct. Biol.* **2017**, *200*, 303–306. [CrossRef] [PubMed]
111. Mäeots, M.-E.; Lee, B.; Nans, A.; Jeong, S.-G.; Esfahani, M.M.N.; Ding, S.; Smith, D.J.; Lee, C.-S.; Lee, S.S.; Peter, M.; et al. Modular Microfluidics Enables Kinetic Insight from Time-Resolved Cryo-EM. *Nat. Commun.* **2020**, *11*, 3465. [CrossRef] [PubMed]
112. Dandey, V.P.; Budell, W.C.; Wei, H.; Bobe, D.; Maruthi, K.; Kopylov, M.; Eng, E.T.; Kahn, P.A.; Hinshaw, J.E.; Kundu, N.; et al. Time-Resolved Cryo-EM Using Spotiton. *Nat. Methods* **2020**, *17*, 897–900. [CrossRef] [PubMed]
113. Kontziampasis, D.; Klebl, D.P.; Iadanza, M.G.; Scarff, C.A.; Kopf, F.; Sobott, F.; Monteiro, D.C.F.; Trebbin, M.; Muench, S.P.; White, H.D. A Cryo-EM Grid Preparation Device for Time-Resolved Structural Studies. *IUCrJ* **2019**, *6*, 1024–1031. [CrossRef]
114. Klebl, D.P.; White, H.D.; Sobott, F.; Muench, S.P. On-Grid and in-Flow Mixing for Time-Resolved Cryo-EM. *Acta Crystallogr. D Struct. Biol.* **2021**, *77*, 1233–1240. [CrossRef] [PubMed]