SINGLE PARTICLES ALWAYS FIT THE MOLD

Nikolaus Grigorieff, Howard Hughes Medical Institute, Brandeis University, MS029, 415 South Street, Waltham, MA 02454-9110

Introduction

An explanatory footnote to the Preface of the latest edition of the popular text "Three-Dimensional Electron Microscopy of Macromolecules and Their Assemblies" [1] explains: "X-ray crystallography was a common technique to determine the 3D structure of proteins around the turn of the century. It required extensive amounts of protein and cumbersome crystallization to obtain diffraction patterns, leading to the well-known phase problem. It is still used today to study small molecules at ultrahigh-resolution". The Preface notes that by far, the most common technique to determine 3D protein structures is electron microscopy and single particle averaging because of "its in-house accessibility, its universal scope, reaching from protein molecules to complex assemblies, its simplicity and ease of use and, most importantly, its guarantee for an atomic structure". It concludes that "this technique was instrumental in the completion of one of our most challenging quests, the visualization of the Human Proteome".

Though this book is not yet written, it is useful for this discussion to project into the future and evaluate what will remain science fiction. Using Single Particle Electron Microscopy (SPEM), structures have been obtained at about 7 Å resolution with highly symmetrical viruses [2-4], and at about 10 Å with the asymmetrical ribosome [5, 6]. SPEM has a unique niche in the visualization of large protein complexes, as exemplified over the last 30 years by the ribosome structure. It is gaining scope in image processing of helical structures [7] and 2D crystals [8-10], pointing to a single method for all possible sample geometries. However, with the potential of a universally applicable technique comes also the potential of strongly biasing a structure towards an initial (sometimes wrong) reference used for alignment. SPEM is also being introduced as a "new powerful technique" in many publications and grant applications. However, when compared to X-ray crystallography, the number of publications reveals that the latter is at least 50 times more powerful [1], and when comparing the amount of data contained in a structure on average in each case, X-ray crystallography is about 300 times as powerful. What are the reasons for this stark contrast?

Problems

Automation of data collection and processing: Rapid data collection and robust, semi-automatic data processing has made X-ray crystallography accessible to a large community. Automatic data collection on the EM is being developed [11, 12]. However, many EM labs are developing their own customized software that is often not easy to use by outsiders. In addition, programs are often not well documented and tend to have bugs, which led a prominent crystallographer to ask for the development of "programs and packages in a way that could be used by merely outstanding mortals, rather than demi-gods" [13].

Conformational variability: Conformational inhomogeneity in single particles cannot always be readily detected, especially when the conformational differences are small. Though insignificant at low resolution, conformationally variable parts might not be resolved at the atomic level. This is also a problem in crystallographic approaches though crystal contacts tend to promote a single conformation. Further discussed by Edward Egelman.

Charging and beam-induced movement: Both lead to blurring and degraded high-resolution contrast. They appear to be more severe for single particle work than 2D crystallography as single particles are often imaged while suspended in ice over a hole in the carbon support film [14].

Quantitative understanding of image contrast: In order to determine a 3D structure the experimental data is usually interpreted by a model. Its accuracy therefore depends on a complete and quantitative accounting of the image contrast. This should include not only single elastic scattering events, but also inelastic scattering and, if significant, multiple scattering.

[1] Someone may say there are many more X-ray crystallographers than electron microscopists, but one could reply that there are many more electron microscopes than syncrotrons.

Furthermore, the solvent and support film should be taken into account, including contrast originating from rough ice and carbon surfaces. The theory of image contrast is well developed in materials science, where the quantitative interpretation of electron micrographs and diffraction patterns is done routinely (see for example, [15]). Several studies of inelastic image contrast of frozen-hydrated biological specimens embedded in ice have been performed [16-18], of which the analysis by

Langmore & Smith [17] is the most complete. It has been suggested to use energy filtering to remove inelastically scattered electrons [17-19]. However, energy filters like the Omega filter built by Zeiss, or the Imaging Electron Energy Loss Spectrometer sold by Gatan, are presently not suitable for high resolution since scattering angles are restricted [16], the size of the CCD camera only records a small area of an image, and the filters introduce non-linear distortions in the images. Even more serious is the fact that most inelastically scattered electrons (with energy losses of about 20 eV) retain partial coherence and carry a significant amount of image contrast [20]. The image formed by the inelastically scattered electrons is a replica of the elastic image [15, 20]. To exclude those electrons from the analysis would increase the number of images necessary to obtain a given signal-to-noise ratio in the final structure. A better approach would be to include inelastically scattered electrons in the image analysis by modeling inelastic image contrast.

CTF correction: Although the CTF of an electron microscope is well-understood, its precise parameters have to be determined not only for each image, but for each particle in an image since the vertical particle positions are all different [21]. At high resolution, the curvature of the Ewald sphere also has to be taken into account. Apart from the defocus amount and astigmatism (3 parameters), the beam tilt also has to be determined for each image (2 parameters). This is routinely done in electron crystallography [22] but poses a problem for an image of single particles since the signal at high resolution is much weaker than that from a 2D crystal.

Accurate particle alignment: To determine a high-resolution structure, the 5 parameters determining position and orientation of each particle must be accurately known [23]. Inaccurate determination of particle parameters is likely to be one of the most severe resolution-limiting factors at the moment. Due to the weakening signal towards higher resolution, accurate alignment is difficult at best.

Quality measurement: A common way to evaluate a 3D reconstruction is its resolution. Unfortunately, agreement on measuring resolution is hard to reach [21, 24]. Usually, the resolution is based on the Fourier Shell Correlation (FSC) between reconstructions derived from two halves of the data set, and calculated in resolution zones. One proposal is to set the resolution at the crossover point between the FSC and three standard deviations (3• criterion) of the correlation of random data [21, 25]. Although the 3• criterion indicates a constant probability of presence of data at a particular resolution, it does not say much about the interpretability of the data. Furthermore, the 3• criterion depends on the size of the volume the reconstructed density sits in: Padding the volume will lower the 3• threshold. Another proposal is to set the resolution at the point where the FSC drops below 0.5 [3, 26]. This criterion does not depend on the size of the volume, and it indicates the point where the signal-to-noise ratio (SNR) in each reconstruction is 1 [27]. An additional complication is the bias of the FSC towards higher values as the two reconstructions are not independent but are usually calculated from images which were all aligned against a single reference [28].

Discussion

Developing the technique: When following the development of X-ray crystallography, it is clear that automation of data collection and processing is an important ingredient for a technique to become widely used and accepted. Given the progress already made [11, 12, 29-35] it is reasonable to expect that advances in instrumentation and software will eventually give us high throughput and ease of use: Samples will be applied to engineered grids, allowing for fully automated data collection overnight on a field-emission, liquid-helium electron microscope. As each image is recorded on a CCD, a preliminary evaluation of data quality will be carried out and particles will be selected, aligned and classified without user assistance. Programs will be fully documented, will not crash, and will provide the user with standardized, reliable diagnostic data. Data and a history of operations will be stored in a comprehensive data base. The programs will run on a Linux cluster with 1000 terraflop capacity, and soon after data collection finishes, the user will be presented with a 3D model of his/her protein, together with resolution curves, real-space reliability maps, and a projection of how much more data is needed to reach a specified resolution. The map will be stored in a standard map format, and tools will be provided to fit other models (atomic or otherwise) from a library of known structures into the map.

A fundamental question: A more fundamental problem may be the resolution one could hope to achieve with SPEM. Most of the problems listed above have a direct impact on resolution. Given a sufficiently high SNR, as with 2D crystals, one could obtain an atomic model using EM. Due to their size and symmetry, viruses are leading the way to atomic resolution in the realm of single particles, and this has prompted the use of viruses as scaffolds for other proteins [36]. But is there any

reason to believe that viruses, let alone asymmetric particles like the ribosome, could ever be aligned with sufficient accuracy to obtain an atomic model? It has been predicted that this could be done if the average SNR is sufficiently high to reliably detect the correct correlation peak in a cross-correlation map [37, 38]. This prediction is based on an average signal across the entire resolution range considered and does not account for the rapid signal decay normally observed towards higher resolution [39]. Furthermore, additional degrees of freedom due to parameters such as magnification, defocus/astigmatism, and beam tilt are not considered. *Can an accurate alignment at high resolution be achieved solely based on low-resolution signal?*

This will depend strongly on a detailed understanding of the low-resolution contrast and the noise distribution in the images. New approaches to particle alignment and 3D reconstruction, such as the use of maximum likelihood functions [40], or labeling with heavy-atom clusters as anchor points [41], may help. Furthermore, the signal at higher resolution can be enhanced by reducing charging effects, for example, by using higher electron beam energies, or by preparing the samples with a sandwich of two carbon layers which also reduces beam-induced movement [42, 43].

The question of resolution is intimately linked with the problem of resolution measurement. If one assumes for a moment that a reliable measurement exists, it is possible, in principle, to determine all necessary parameters to obtain any desirable resolution (assuming the images contain the necessary data). All one would have to do is to carry out an exhaustive search of all relevant parameters, and to evaluate the corresponding 3D reconstruction for every test parameter set. The problem of conformational flexibility could be handled in a similar manner by testing all possible subsets of a data set to find a homogeneous set of particles. The main problem with this type of approach is the prohibitively high computational cost but, as the past twenty years suggest, this problem will be solved by time.

Measuring resolution: The FSC would be an unbiased resolution measurement if the two reconstructions calculated from two half data sets were truly independent, i.e. if two references were used to refine parameters in each set. The disadvantage is the reduced signal-to-noise ratio of the two references as they would each contain only half the data. An alternative would be the exclusion of a certain percentage of data, say 5 or 10%, from the refinement-reconstruction cycle. A correlation coefficient or phase residual in resolution zones could then be calculated between the excluded images and the reconstruction to measure resolution. This idea, which is already implemented in the program FREALIGN [44], is analogous to the free R-factor used in X-ray crystallography [45].

References

1. Frank, J. and M. van Heel (2010). *Three-dimensional electron microscopy of macromolecules and their assemblies.* 10th ed. San Diego: Academic Press.
2. Conway, J.F., et al. (1997). *Visualization of a 4-helix bundle in the hepatitis B virus capsid by cryo-electron microscopy.* Nature 386, 91-94.
3. Böttcher, B., S.A. Wynne, and R.A. Crowther (1997). *Determination of the fold of the core protein of hepatitis B virus by electron cryomicroscopy.* Nature 386, 88-91.
4. Zhou, Z.H., et al. (2001). *Electron cryomicroscopy and bioinformatics suggest protein fold models for rice dwarf virus.* Nat Struct Biol 8, 868-73.
5. Gabashvili, I.S., et al. (2000). *Solution structure of the E. coli 70S ribosome at 11.5 A resolution.* Cell 100, 537-49.
6. Matadeen, R., et al. (1999). *The Escherichia coli large ribosomal subunit at 7.5 A resolution.* Structure Fold Des 7, 1575-83.
7. Egelman, E.H. (2000). *A robust algorithm for the reconstruction of helical filaments using single-particle methods.* Ultramicroscopy 85, 225-34.
8. Sherman, M.B., et al. (1998). *Multivariate analysis of single unit cells in electron crystallography.* Ultramicroscopy 74, 179-99.
9. Stahlberg, H., et al. (1998). *Are the light-harvesting I complexes from Rhodospirillum rubrum arranged around the reaction centre in a square geometry?* J Mol Biol 282, 819-31.
10. Walz, T. and R. Ghosh (1997). *Two-dimensional crystallization of the light-harvesting I-reaction centre photounit from Rhodospirillum rubrum.* J Mol Biol 265, 107-11.
11. Oostergetel, G.T., W. Keegstra, and A. Brisson (1998). *Automation of specimen selection and data acquisition for protein electron crystallography.* Ultramicroscopy 74, 47–59.
12. Rouiller, I., et al. (2001). *Automated image acquisition for single-particle reconstruction using p97 as the biological sample.* J Struct Biol 133, 102-7.
13. Harrison, S. (2002). Personal communication.
14. Böttcher, B. (1995). *Electron cryo-microscopy of graphite in amorphous ice.* Ultramicroscopy 58, 417-424.

15. Hirsch, P.B., et al. (1965). *Electron microscopy of thin crystals.* London: Butterworths.
16. Angert, I., et al. (1996). *Elastic and inelastic scattering cross-sections of amorphous layers of carbon and vitrified ice.* Ultramicroscopy 63, 181-192.
17. Langmore, J.P. and M.F. Smith (1992). *Quantitative energy-filtered electron-microscopy of biological molecules in ice.* Ultramicroscopy 46, 349-373.
18. Zhu, J., et al. (1997). *Three-dimensional reconstruction with contrast transfer function correction from energy-filtered cryoelectron micrographs: Procedure and application to the 70S Escherichia coli ribosome.* Journal of Structural Biology 118, 197-219.
19. Grimm, R., et al. (1997). *Energy filtered electron tomography of ice-embedded actin and vesicles.* Biophysical Journal 72, 482-489.
20. Rossouw, C.J. and M.J. Whelan (1981). *Diffraction contrast retained by plasmon and K-loss electrons.* Ultramicroscopy 6, 53-66.
21. van Heel, M., et al. (2000). *Single-particle electron cryo-microscopy: towards atomic resolution.* Q Rev Biophys 33, 307-69.
22. Henderson, R., et al. (1986). *Structure of purple membrane from Halobacterium halobium: recording, measurement and evaluation of electron micrographs at 3.5 Å resolution.* Ultramicroscopy 19, 147-178.
23. Jensen, G.J. (2001). *Alignment error envelopes for single particle analysis.* J Struct Biol 133, 143-55.
24. Ruprecht, J. and J. Nield (2001). *Determining the structure of biological macromolecules by transmission electron microscopy, single particle analysis and 3D reconstruction.* Prog Biophys Mol Biol 75, 121-64.
25. Saxton, W.O. and W. Baumeister (1982). *The correlation averaging of a regularly arranged bacterial cell envelope protein.* J Microsc 127, 127-38.
26. Beckmann, R., et al. (1997). *Alignment of conduits for the nascent polypeptide chain in the ribosome- Sec61 complex.* Science 278, 2123-6.
27. Frank, J. and L. Al-Ali (1975). *Signal-to-noise ratio of electron micrographs obtained by cross correlation.* Nature 256, 376-379.
28. Grigorieff, N. (2000). *Resolution measurement in structures derived from single particles.* Acta Crystallogr D Biol Crystallogr 56, 1270-7.
29. Downing, K.H. and F.M. Hendrickson (1999). *Performance of a 2k CCD camera designed for electron crystallography at 400 kV.* Ultramicroscopy 75, 215-33.
30. Frank, J., et al. (1996). *SPIDER and WEB: processing and visualization of images in 3D electron microscopy and related fields.* Journal of Structural Biology 116, 190-199.
31. Lata, K.R., P. Penczek, and J. Frank (1995). *Automatic particle picking from electron micrographs.* Ultramicroscopy 58, 381–391.
32. Ludtke, S.J., P.R. Baldwin, and W. Chiu (1999). *EMAN: semiautomated software for high-resolution single-particle reconstructions.* J Struct Biol 128, 82-97.
33. Nicholson, W.V. and R.M. Glaeser (2001). *Review: automatic particle detection in electron microscopy.* J Struct Biol 133, 90-101.
34. Stewart, P.L., et al. (2000). *Digitally collected cryo-electron micrographs for single particle reconstruction.* Microsc Res Tech 49, 224-32.
35. van Heel, M., et al. (1996). *A new generation of IMAGIC image processing system.* Journal of Structural Biology 116, 17-24.
36. Kratz, P.A., B. Bottcher, and M. Nassal (1999). *Native display of complete foreign protein domains on the surface of hepatitis B virus capsids.* Proc Natl Acad Sci U S A 96, 1915-20.
37. Glaeser, R.M. (1999). *Review: electron crystallography: present excitement, a nod to the past, anticipating the future.* J Struct Biol 128, 3-14.
38. Henderson, R. (1995). *The potential and limitations of neutrons, electrons and X-rays for atomic resolution microscopy of unstained biological molecules.* Q Rev Biophys 28, 171-93.
39. Saad, A., et al. (2001). *Fourier amplitude decay of electron cryomicroscopic images of single particles and effects on structure determination.* J Struct Biol 133, 32-42.
40. Sigworth, F.J. (1998). *A maximum-likelihood approach to single-particle image refinement.* Journal of Structural Biology 122, 328-339.
41. Jensen, G.J. and R.D. Kornberg (1998). *Single-particle selection and alignment with heavy atom cluster- antibody conjugates.* Proc Natl Acad Sci U S A 95, 9262-7.

42. Brink, J., et al. (1998). *Reduction of charging in protein electron cryomicroscopy*. J Microsc 191, 67-73.
43. Jakubowski, U., W. Baumeister, and R.M. Glaeser (1989). *Evaporated carbon stabilizes thin, frozen-hydrated specimens.* Ultramicroscopy 31, 351–358.
44. Grigorieff, N. (1998). *Three-dimensional structure of bovine NADH:ubiquinone oxidoreductase (complex I) at 22 Å in ice*. Journal of Molecular Biology 277, 1033-1046.
45. Brünger, A.T. (1992). *Free R value: a novel statistical quantity for assessing the accuracy of the structure.* Nature 355, 472-475.