Towards more quantitative results from Three Dimensional Electron Microscopy of Macromolecules

José M. Carazo, Natalia Jiménez-Lozano, C.O.S. Sorzano, Mónica Chagoyen

BioComputing Unit, Centro Nacional de Biotecnología-CSIC, Campus Universidad Autónoma, 28049 Madrid, Spain

Summary: Impressive results on the three-dimensional (3D) architecture of biological macromolecules have been obtained over the last two decades. Also, information integration by a combination of different experimental techniques is also a growing area of interest, especially in the combination of transmission electron microscopy and 3D reconstruction techniques, commonly referred to as 3D-EM, with structural studies by X-ray diffraction. The move towards this data integration is making evident the need to further standardise 3D-EM, as well as the need to obtain more quantitative results. The field of 3D-EM is starting to consider the development of quality standards in areas such as measures of resolution and general quality assessment, as well as methods to produce more quantitative results as far as absolute grey values in the reconstructions is concerned.

1. Introduction

Three-dimensional electron microscopy (3D-EM) has developed over the years into a quite sophisticated approach thanks to the combination of expert knowledge in the fields of biology, physics and data processing (for a review see [1]). Typically, 3D-EM provides structural information on the architecture of biological macromolecules over a resolution range from one to about three nanometers, although a higher resolution has been obtained in certain cases. The number of structures solved over the last three years, as well as the high degree of correlation with data at atomic resolution, obtained mainly by X-ray crystallography, is quite high and raising (for a survey see [2]). As a consequence of this move towards a more central role in the field of structural biology, 3D-EM needs to undertake further developments in a series of fronts, some of which will be covered in this contribution.

2. Resolution and quality assessment of 3D-EM data

In the last years there has been a realisation within the field of 3D-EM that a general criteria should be established to measure the validity and accuracy of the data provided. The current situation in the field is that the "resolution" is the only *de facto* quality measurement supplied by the authors, although different methods can be applied to obtain it. These different criteria range from "reflection-based" resolution estimators, which are typical of electron crystallographic works, to others oriented towards measuring the "reproducibility" degree between two independent reconstructions, commonly used in single particle approaches (namely Fourier ring correlation (FRC) [3], differential phase residual [4], variance map [5], and R-factor$_{AB}$ [6]), as well as other approaches such as the spectral signal-to-noise ratio (SSNR) [7].

Well-established and commonly used procedures and criteria to assess the quality of the 3D-EM data, both in global and local basis, are hence of primary importance. The assessment of the quality of a given reconstruction should be done at three levels:

- quality of the experimental data;
- quality of the reconstruction procedure, i.e. how well the final volume agrees with the experimental data;
- reliability of the final volume independently of the experimental data and reconstruction procedure, i.e. validation of the data *vs. a priori* knowledge.

The first stage should include all the measures and criteria to characterise the experimental technique and goodness of the raw data. Indication of the CTF extrema, defocus range, magnification (both nominal and real) and beam dose are a must, but additionally other parameters should be identified.

Secondly, we need to quantify the quality of the processing steps performed: from the digitization till the reconstruction algorithm, passing through any computational enhancement on the initial images, classification procedures, angular assessment, etc...

Finally, the validation of the reconstruction can also be done in terms of its conformance with any previous knowledge. This is quite straightforward in the case of X-ray structure determination, where stereochemical and energetic restraints are well defined, while it is very difficult to achieve in 3D-EM due to the low resolution of these studies.

An indirect measure of the complexity of the task we are facing may be indicated by the fact that a series of projects addressing this topic as applied to X-ray diffraction and NMR spectroscopy data have been running over the last years [8, 9, 10].

3. Quantitative physical meaning of the 3D results

A common situation as today is that 3D-EM reconstructions provide volumetric data for which the grey values of the individual voxels are not directly (linearly, if possible) related to a physical magnitude. In relative terms, the voxel grey values are related to the Coulomb potential of the sample, but their precise single values do not generally carry a clear meaning. As a consequence, most of the work performed on these data addresses the voxel values in a qualitative manner. However, it would be desirable to work towards better defined voxel values if we were to incorporate precise biochemical knowledge on the direct results of the reconstructions as well as further developing integration tools between data from 3D-EM and other experimental techniques such as X-ray diffraction.

The reasons of this lack of direct meaning for the actual individual values of the voxels are quite diverse. To start with, the Transmission Electron Microscope is not a perfect device, and the images that it provides are modulated by the so-called Contrast Transfer Function (CTF) (for a review see [1]). The actual parametric model of this CTF is based on a detailed physical understanding of the image forming process, and has been worked out over the years for a number of relevant situations that are mostly applicable to the work with macromolecules. However, the models are still far from being perfect, especially in the low frequency region, which is vital to define the general grey level of an image. As for the higher frequencies, special care has to be given to both the contrast inversion as well as the amplitude weighting characteristic of the CTF. However, there are many cases in which the CTF is not considered in the

reconstruction process, basically by low-pass filtering the images in order to avoid encountering the high frequency effects. Still, this does not work out low frequency effects. (Exceptions are certainly found in high resolution electron crystallography as well as in certain viral and single particles studies).

Solutions to the problems referred above may also come from various directions. To start with, the main effect that dominates the low frequency region of the CTF is due to inelastic scattering, and one of the best approaches to its resolution may rest upon the usage of zero-loss energy filters. Once in possession of an accurate model for the CTF over a wide frequency interval, there are approaches that take this CTF directly into account in the reconstruction process [11,12]), while there are others that work *a posteriori* by a process of image/volume restoration ([13, 14]). In general, all the possible solutions sketched before still need substantial further investigation and validation, which are expected to be the theme of a number of research studies over the coming years.

Additionally, two other areas that should be further addressed in this quest for quantitative results are: a precise study of the image acquisition process, including whatever appropriated normalisations, and the characterisation of the 3D reconstruction algorithms in terms of their ability to reproduce the density voxel values.

4. Conclusion

At a time in which data integration plays a central role in modern biology, the need for producing quantitative data with standardised quality control measures becomes essential. The coming years are to witness a broader integration of an increasingly growing flow of data in structural biology into databases, and the issues of standardisation and reliability will then become crucial. The field of 3D-EM is certainly working in this direction, and in this contribution several lines of action have been briefly sketched

Acknowledgements:

This work has been supported, in part, by grant number BIO 98-0761 from the Plan Nacional de Biotecnología of Spain

References

[1] J. Frank. Three-dimensional electron microscopy of macromolecular assemblies. Albany, New York 1996.
[2] J.M. Carazo, G.J. Barton, S.D. Fuller, M. Radermacher, in: Proc. 12th EUREM, Brno 2000.
[3] W.O. Saxton, W. Baumeister J. Microsc 127 (1982) 127.
[4] J. Frank, A. Verschoor, M. Boublik, Science 214 (1981) 1353.
[5] N.A. Farrow, F.P. Ottensmeyer, J. Opt. Soc. Am. A9 (1992) 1749.
[6] D.A. Winkelmann, T.S. Baker, I. Rayment, J Cell Biol 114 (1991) 701.
[7] M. Unser, M.J. Vrhel, J.F. Conway, M. Gross, P. Thévenaz, A.C. Steven, B.L. Trus, in: Proc. 11th EUREM, Dublin 1996, p. 260.
[8] EU 3-D Validation Network, J Mol Biol 276 (1998) 417.
[9] R.A. Laskowski, M.W. MacArthur, J.M. Thornton, Curr Opin Struct Biol 8 (1998) 631.
[10] E.J. Dodson, G.J. Davies, V.S. Lamzin, G.N. Murshudov, K.S. Wilson, Structure 6 (1998) 685.
[11] J. Zhu, P.A. Penczek, R. Schröder, J. Frank, J Struct Biol 118 (1997) 197.
[12] U. Skoglund, L.-G. Öfverstedt, R.M. Burnett, G. Bricogne, J Struct Biol 117 (1996) 173.
[13] J.F. Conway, N. Cheng, A. Zlotnick, P.T. Wingfield, S.J. Stahl, A.C. Steven, Nature 386 (1997) 91.
[14] B. Böttcher, S.A. Wynne, R.A. Crowther, Nature 386 (1997) 88.