Archives of Biochemistry and Biophysics 581 (2015) 39–48

Contents lists available at ScienceDirect

Archives of Biochemistry and Biophysics

journal homepage: www.elsevier.com/locate/yabbi

Review

Three-dimensional reconstruction methods in Single Particle Analysis from transmission electron microscopy data

J.M. Carazo [a,*], C.O.S. Sorzano [a,c], J. Otón [a], R. Marabini [b], J. Vargas [a]

[a] Biocomputing Unit, National Center for Biotechnology (CSIC), c/Darwin, 3, Campus Universidad Autónoma, 28049 Cantoblanco, Madrid, Spain
[b] Escuela Politécnica Superior, Universidad Autónoma de Madrid, Campus Universidad Autónoma, 28049 Cantoblanco, Madrid, Spain
[c] Bioengineering Lab., Universidad CEU San Pablo, Campus Urb. Montepríncipe s/n, 28668 Boadilla del Monte, Madrid, Spain

ARTICLE INFO

Article history:
Received 29 January 2015
and in revised form 11 May 2015
Available online 22 May 2015

Keywords:
Cryo electron microscopy
Image processing

ABSTRACT

The Transmission Electron Microscope provides two-dimensional (2D) images of the specimens under study. However, the architecture of these specimens is defined in a three-dimensional (3D) coordinate space, in volumetric terms, making the direct microscope output somehow "short" in terms of dimensionality. This situation has prompted the development of methods to quantitatively estimate 3D volumes from sets of 2D images, which are usually referred to as "three-dimensional reconstruction methods". These 3D reconstruction methods build on four considerations: (1) The relationship between the 2D images and the 3D volume must be of a particularly simple type, (2) many 2D images are needed to gain 3D volumetric information, (3) the 2D images and the 3D volume have to be in the same coordinate reference frame and (4), in practical terms, the reconstructed 3D volume will only be an approximation to the original 3D volume which gave raise to the 2D projections. In this work we will adopt a quite general view, trying to address a large community of interested readers, although some sections will be particularly devoted to the 3D analysis of isolated macromolecular complexes in the application area normally referred to as Single Particle Analysis (SPA).

Introduction

Our field of work is the experimental resolution of the three-dimensional structure of macromolecular complexes using the Transmission Electron Microscope (TEM)[1] under cryogenic condition, an area also known as cryo EM. Within this broad topic, we will focus on three-dimensional reconstruction techniques, which is one of the basic steps in the structural resolution process. Note that cryo EM is experiencing a profound "revolution" nowadays thanks to several key technological and methodological advancements, such as the advent of Direct Electron Detectors and new image processing methods. We refer to other contributions in this Special Issue to properly review the state of the art in this field, so that in the following we focus on the crucial step of how to obtain three-dimensional quantitative information from TEM images.

The search towards always richer information is intrinsic to the human being. Indeed, there are many situations in which a certain type of information is experimentally measured, but our real interest goes beyond these measurements and it pertains to another property "related" to them. In other words, we measure "something", but we are interested in "something else". In a very broad sense, these cases are usually referred to as "inverse problems", which can be expressed in a more formal way as

$$g = Hf \tag{1}$$

with g being our measurements, f being our desired property, and H describing the physical process that links our measurement with the desired property. Since we want to obtain f from g, we have to invert, or "reverse", H leading to

$$f = H^{-1}g \tag{2}$$

and thus the name of "inverse problems".

Quite naturally, our ability to obtain f from g will greatly depend on the inversion properties of H. In the case of Transmission Electron Microscopy (TEM), g refers to sets of 2D images collected at the microscope, f to the 3D structure of our specimen, and H conveys the detailed information on how the electron microscope interacts with the specimen under study, producing concrete sets of 2D images. Once H is known, we have to find the conditions under which H^{-1} can indeed be realized, first from a somehow

* Corresponding author. Fax: +34 913720112.

[1] *Abbreviations used:* TEM, Transmission Electron Microscope; SPA, Single Particle Analysis; FT, Fourier Transform; FFT, Fast Fourier Transform; ML, maximum likelihood; CTF, Contrast Transfer Function.

http://dx.doi.org/10.1016/j.abb.2015.05.003
0003-9861/

abstract mathematical perspective, and then as practically implemented in a computer.

The procedure described above is very general, and it applies not only to electron microscopy but to most areas of biomedical imaging. However, the work with isolated macromolecular complexes, normally referred to as Single Particle Analysis (SPA), introduces some crucial differences with respect to other imaging modalities. Indeed, in a typical biomedical imaging application in a clinical context, we have a well-defined and unique f, the patient, from whom a number of images (X-ray radiographies) are going to be collected in order to calculate a 3D map. However, a macromolecular complex is a very dynamic entity, so the probability is large to have in our sample under investigation not only one f, but a whole sets of different f's, corresponding to different conformational states, giving rise to a mixed population of g's. Clearly, the formulation above has to be extended to take into account this situation. Furthermore, a large number of applications in EM are characterized by uncertainties about the way images have been collected, besides always been affected by heavy noise.

This paper is organized in the following manner. In Section '2D images and 3D volumes: Basic relationships' we will review the basics of the way electrons interact with the specimen in the microscope, producing a 2D image. In practical terms, we will be dealing with the characterization of H. We will also address some of the most common strategies to collect sets of images. Section 'From 2D images to 3D volumes: Reconstruction methods' will then concentrate on ways to invert H, and these will be the different reconstruction methods. At this stage we will present the way the 3D reconstruction process is performed in practice, introducing the notion of a "3D reconstruction workflow", particularized to SPA; this topic will be covered in Section 'A typical 3D reconstruction workflow in SPA'. Quite naturally, any reconstruction process starts with a detailed characterization of the initial experimental images, which will then be addressed in Section 'Ch aracterizing the initial experimental images'. However, we have already indicated that the simple mathematical framework of $g = Hf$ has to be extended to accommodate for the conformational flexibility of macromolecular complexes, besides a large number of experimental uncertainties and noise. This topic will be covered in Section 'From 2D images to 3D volumes: A posteriori projection assignment and classification'. Further elaborating on extensions of the basic reconstruction framework, we will briefly discuss the case of more elaborated H's, typical of certain demanding applications; this will be covered in Section 'On more complicated relationships: When simplification breaks'. Finally, we will present a general discussion in Section 'Discussion and conclusion'.

2D images and 3D volumes: Basic relationships

In this Section we will address three main questions: (1) Which is the relationship between the 3D volume of the specimen under investigation and its associated 2D images?, (2), Is one image enough to obtain a 3D reconstruction, and if this is not the case, how many are needed? and, (3), In practical terms, how 2D images are collected?

A Transmission Electron Microscope works by using highly accelerated electrons as "light source", and focusing these electrons onto an image thanks to electromagnetic lenses. Typical accelerating voltages are in the order of 200 kVolts, producing electrons with associated wavelengths of about 2.5 pm. It is quite clear that, as an instrument, the imaging limitations of the electron microscope are not due to the (very small) wavelength being used (much less than one thousandth of an Å), but to imperfections of the electromagnetic lenses (naturally, the specimen itself may introduce additional limitations, such as those related to dose sensitivity, the material surrounding the sample of interest, or beam induced movement). Electrons interact with the biological specimen under study as negatively charged particles, providing experimental information on the three-dimensional Coulomb (i.e., electrostatic) potential of the specimen. Considering that the typical atomic composition of macromolecular complexes is formed by elements of relatively low atomic number (like carbon, oxygen or hydrogen), and that the specimens themselves are small (a ribosome is in the order of 250Å, as an example), it is normally considered that the interaction between the accelerated electrons and the biological matter is very weak. So weak, in fact, that only some of the electrons going through the specimen interact with some of its atoms, and that the result of this interaction is "only" a change of the associated phase of the electron (they are not absorbed or, in general, loose energy). Under these simplified conditions, it is possible to model electron microscopy images as if the whole three-dimensional structure of the specimen would be "condensed" into an image perpendicular to the electron direction; in other words, as if the whole Coulomb potential would be "summed" (integrated) along the direction of the electron beam into each point of the resulting image. We refer to images formed in this "condensed" manner as "projection images" of the specimen under study (the reader is referred to Hawkes [8], Hawkes and Kasper [9], Frank [7] for further details). We can express the concepts presented above in a simple mathematical way as [2]:

$$\text{EM\_Image} = \text{Projection}(\text{biological specimen}) \tag{3}$$

where "Projection" is an operation performing a summation (line integral) along the electron beam trajectory.

$$g = \textit{line\_integral}(f) \tag{4}$$

Once understood how images are formed, we may start thinking about how the three-dimensional process can take place. Indeed, the field of 3D reconstruction from 2D images may be regarded, at first glimpse, as somehow "magic", and it is not at all obvious that a whole "dimension" can be gained from lower dimensionality data by some mathematical procedure. The question is so fascinating that back in 1917, with no concrete experimental application in mind whatsoever, the Austrian mathematician Johann Radon derived a way to perform this process under a certain set of conditions (a translation in English of this fundamental work can be found in Radon [19]). The first and most critical one was that the lower dimensionality data had to be obtained as line integrals over the higher dimensionality space. Translated into a 2D/3D case, it required that the 2D images had to be projections of the 3D volume, which is exactly the relationship that exists (within approximations) between transmission electron microscopy images and the 3D biological specimen under investigation, as we have presented in previous paragraphs. Radon inversion formula certainly established the feasibility of performing the 3D reconstruction process, but the actual answer was not very practical, since it required an infinite number of noiseless projection images to perform the inversion.

A simple way to have a very practical understanding of the relationship between 2D projection images and its associated 3D volume is to formulate the case in Fourier space. We refer to Fourier space as the range of a very well-known operation known as the

[2] Formally, the projection equation can be written as

$$g(\mathbf{s}) = \text{Proj}\{f(\mathbf{r})\}(\mathbf{s}) = \int_{-\infty}^{\infty} f(H^T\mathbf{s} + z\mathbf{e}_3)dz$$

where $H^T = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$, $\mathbf{s} \in \mathbb{R}^2$ is a 2D coordinate in the image, $\mathbf{r} \in \mathbb{R}^3$ is a coordinate in the 3D volume, and $\mathbf{e}_3 = (0,0,1)^T$ is the z-axis.