Archives of Biochemistry and Biophysics 581 (2015) 49–53

Contents lists available at ScienceDirect

Archives of Biochemistry and Biophysics

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

Validation methods for low-resolution fitting of atomic structures to electron microscopy data

Xiao-Ping Xu, Niels Volkmann *

Bioinformatics and Structural Biology Program, Sanford-Burnham Medical Research Institute, 10901 N Torrey Pines Rd, La Jolla, CA 92037, USA

ARTICLE INFO

Article history:
Received 13 April 2015
and in revised form 12 June 2015
Accepted 23 June 2015
Available online 24 June 2015

Keywords:
Fitting
Electron microscopy
Validation

ABSTRACT

Fitting of atomic-resolution structures into reconstructions from electron cryo-microscopy is routinely used to understand the structure and function of macromolecular machines. Despite the fact that a plethora of fitting methods has been developed over recent years, standard protocols for quality assessment and validation of these fits have not been established. Here, we present the general concepts underlying current validation ideas as they relate to fitting of atomic-resolution models into electron cryo-microscopy reconstructions, with an emphasis on reconstructions with resolutions below the sub-nanometer range.

1. Introduction

Due to dramatic improvements in experimental methods and computational techniques, electron cryo-microscopy (cryo-EM) has matured into a powerful collection of methods that allow the high-resolution visualization of the structure and the dynamics of an extraordinary range of biological assemblies in their native aqueous environment. Recent hardware and software developments have revolutionized the field [1]. The increased signal-to-noise ratio of a new generation of cameras that detect electrons directly [2] in combination with their ability to correct for beam-induced movements, have allowed the field to obtain structural information even for particles with low or no symmetry at resolutions around 3 Å [3–8], sufficient to build de novo structural models [9].

However, the majority of reconstructions obtained by cryo-EM are of insufficient resolution for such direct structure determination. In fact, currently over 70% of the reconstructions deposited in the electron microscopy data bank [10] do not reach a resolution of better than 10 Å. While at resolutions between 5 and 10 Å secondary structural elements are often visible as rods (α-helices) and sheets (β-sheets), at resolutions below the 10-Å mark, internal features of the reconstructions are not straightforward to interpret (Fig. 1).

As cryo-EM methodology continues to improve, atomic-resolution reconstructions are likely to become more common. These reconstructions will likely be of highly rigid molecules. At the same time, the advances in cryo-EM technology will also open the door to structure determination of complexes that were previously too small, too heterogeneous, too flexible, or otherwise challenging, albeit at lower resolution. In addition, electron cryo-tomography has become a powerful alternative for structure determination of samples that are not amenable to single-particle approaches. New software developments and careful experimental design [11] enable the determination of structures from cryo tomograms at around 8 Å, but resolutions below 20 Å are more common. As a net-effect, the majority of cryo-EM reconstructions are likely to remain at the resolution range worse than 10 Å for the foreseeable future.

Fitting of atomic components into cryo-EM density maps insufficient for direct de novo model building is routinely used to understand the structure and function of these macromolecular machines. Many fitting methods have been developed, but standard protocols for successful fitting remain to be established. Broadly, fitting methods can be divided into two major groups, rigid-body and flexible fitting methods. In rigid-body fitting approaches the atomic structures of components are fitted as single units. These units can be composed of entire proteins, domains, or even smaller groups of structural element. In flexible fitting approaches the entire atomic structures are allowed to distort in some way to improve the fit with the reconstruction, subject to constraints such as molecular dynamics force-fields or normal modes to counter-balance fitting of spurious noise. Comprehensive reviews of the various fitting approaches that are available were provided in several recent articles [12–14].

Despite the plethora of available fitting techniques, generally accepted criteria for assessing the accuracy and quality of the fitted

* Corresponding author.
E-mail address: niels@sbmri.org (N. Volkmann).

http://dx.doi.org/10.1016/j.abb.2015.06.017
0003-9861/

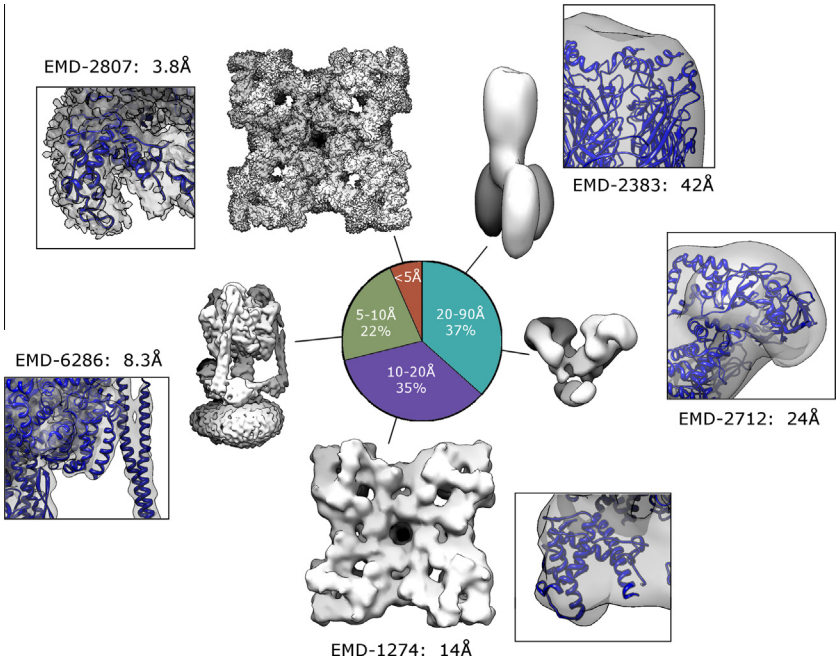

Fig. 1. Spring 2015 snapshot of EM reconstructions deposited in the EM database [10]. Over 70% of the deposited structures do not reach subnanometer resolution and less than 10% reach the 5-Å mark. Some recently deposited example structures are shown to demonstrate the effect of resolution on interpretability in light of atomic resolution structures.

models have not been established yet [15]. In this review, our aim is to present the general concepts underlying current validation ideas as they relate to fitting of atomic models into cryo-EM reconstructions, with an emphasis on reconstructions with resolutions that do not reach the sub-nanometer range.

2. Accuracy versus precision

In the context of fitting, it is important to emphasize the difference between accuracy and precision (Fig. 2). In short, a fit is precise if similar fits are obtained with repeated runs. In contrast, a fit is accurate if it is close to the true structure (or ensemble of structures) underlying the data. Consensus approaches that compare results from different fitting methods [16,17] or from multiple scoring functions [18] using a single data set, for example, can only inform on the precision of the fit. While it has been claimed that precision gives a lower bound for the accuracy, this is not necessarily true (Fig. 2). In fact, in the context of fitting atomic structures into cryo-EM reconstructions, the fitting that appears less precise can actually be more accurate than a fit with the same center position and a narrower spread. The reason is that cryo-EM reconstructions, unlike crystal structures, often represent an ensemble of conformers that co-existed in the sample at the time of freezing, reflecting the structural dynamics of the complex in solution. This can also lead to anisotropy of the resolution in the reconstructions, further complicating the issue.

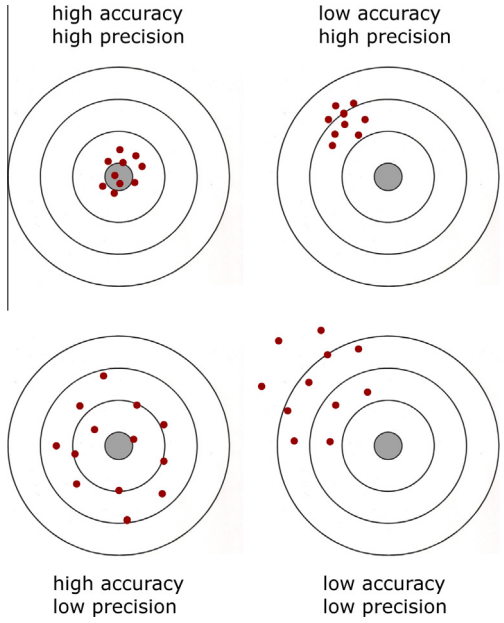

Fig. 2. Difference between accuracy and precision.

It is not immediately clear how precision can be quantified within the context of fitting atomic models. This issue is of major importance especially at low resolution, where ambiguities may arise from the geometry of the reconstruction alone [19] and an objective criterion to allow favoring one solution over another is needed. One approach towards this goal is the use of statistical methods to define confidence intervals (see below), which will then allow to define an objective precision estimate. However, the real quantity that is of interest is the accuracy or how close the obtained fit is to the true structure. Without knowledge of the true structure accuracy cannot be directly assessed.

3. Sources of errors

Every cryo-EM reconstruction has some uncertainty due to the presence of noise and its generally limited resolution. If only such random errors exist, a quantified precision measure may actually be a reasonable estimate for the accuracy of a fit. More troublesome in this context are potential systematic errors such as those originating in the electron microscopy data-collection and reconstruction procedures. These include misestimations of the magnification, incomplete corrections of the microscope's contrast transfer