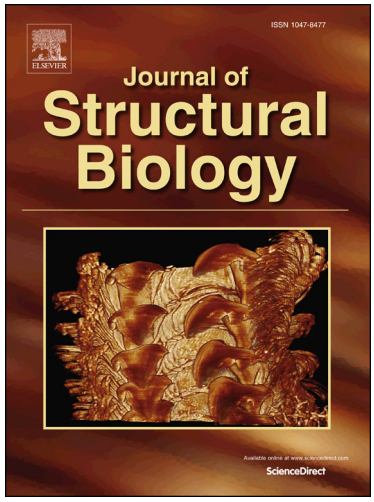

Improved metrics for comparing structures of macromolecular assemblies determined by 3D electron-microscopy

Agnel Praveen Joseph, Ingvar Lagerstedt, Ardan Patwardhan, Maya Topf, Martyn Winn

PII: S1047-8477(17)30087-4
DOI: http://dx.doi.org/10.1016/j.jsb.2017.05.007
Reference: YJSBI 7060

To appear in: *Journal of Structural Biology*

Received Date: 15 October 2016
Revised Date: 19 May 2017
Accepted Date: 23 May 2017

Please cite this article as: Joseph, A.P., Lagerstedt, I., Patwardhan, A., Topf, M., Winn, M., Improved metrics for comparing structures of macromolecular assemblies determined by 3D electron-microscopy, *Journal of Structural Biology* (2017), doi: http://dx.doi.org/10.1016/j.jsb.2017.05.007

Improved metrics for comparing structures of macromolecular assemblies determined by 3D electron-microscopy

Agnel Praveen Joseph[ab], Ingvar Lagerstedt[cd], Ardan Patwardhan[c], Maya Topf[a&], Martyn Winn[b&]

a) Institute of Structural and Molecular Biology, Department of Biological Sciences, Birkbeck College, University of London, Malet Street, London WC1E 7HX, United Kingdom

b) Scientific Computing Department, Science and Technology Facilities Council, Research Complex at Harwell, Didcot OX11 0FA, United Kingdom

c) European Molecular Biology Laboratory, European Bioinformatics Institute, Wellcome Genome Campus, Hinxton, Cambridge CB10 1SD, United Kingdom

d) Current affiliation: Computational Chemistry and Cheminformatics, Lilly UK, Windlesham GU20 6PH, United Kingdom

& Corresponding authors: martyn.winn@stfc.ac.uk, m.topf@cryst.bbk.ac.uk

Abstract

Recent developments in 3-dimensional electron microcopy (3D-EM) techniques and a concomitant drive to look at complex molecular structures, have led to a rapid increase in the amount of volume data available for biomolecules. This creates a demand for better methods to analyse the data, including improved scores for comparison, classification and integration of data at different resolutions. To this end, we developed and evaluated a set of scoring functions that compare 3D-EM volumes. To test our scores we used a benchmark set of volume alignments derived from the Electron Microscopy Data Bank. We find that the performance of different scores vary with the map-type, resolution and the extent of overlap between volumes. Importantly, adding the overlap information to the local scoring functions can significantly improve their precision and accuracy in a range of resolutions. A combined score involving the local mutual information and overlap (LMI_OV) performs best overall, irrespective of the map category, resolution or the extent of overlap, and we recommend this score for general use. The local mutual information score itself is found to be more discriminatory than cross-correlation coefficient for intermediate-to-low resolution maps or when the map size and density distribution differ significantly. For comparing map surfaces, we implemented two filters to detect the surface points, including one based on the 'extent of surface exposure'. We show that scores that compare surfaces are useful at low resolutions and for maps with evident surface features. All the scores discussed are implemented in TEMPy (http://tempy.ismb.lon.ac.uk/).

Keywords: 3D electron cryo-microscopy; integrative modelling; scoring functions; macromolecular assemblies; density fitting

Introduction

A major leap in structure characterization of large bio-molecular machines and cellular components has been brought in by biophysical techniques like electron microscopy (EM) and tomography (ET) (Bai et al., 2013; Kuhlbrandt, 2014; Milne et al., 2013), which result in 3D volume representations of the structure. The Electron Microscopy Data Bank (EMDB) (http://emdb-empiar.org) currently holds over 4000 volume reconstructions from EM and ET, and the number of entries has doubled in the last four years due to increasing interest and development of better image reconstruction