Reciprocal space refinement and map calculation for cryo-EM single particle structures

K. Yamashita, R. Warshamanage, G. N. Murshudov

MRC Laboratory of Molecular Biology, Francis Crick Avenue, Cambridge, Biomedical Campus, Cambridge, CB2 0QH, UK

kyamashita@mrc-lmb.cam.ac.uk

The number of high-resolution structure determination by cryo-EM single particle analysis (SPA) is growing rapidly. Focusing on maps having better resolution than 3 Å deposited in the Electron Microscopy Data Bank (EMDB), there were 316 depositions in 2019 while it was 82 in 2018. It increases the importance of method developments for accurate determination of atomic coordinates and thus the model validation, where not only the geometric quality but also the fitness to the map is of great importance.

Here we present a new program *Servalcat* for the refinement and map calculation of cryo-EM SPA structures. *Servalcat* implements a refinement pipeline using *REFMAC5*, which uses a dedicated likelihood function for SPA [1]. It takes as inputs unsharpened and unweighted half maps from independent reconstructions. The variance of noise in Fourier coefficients is estimated using the half maps. A weighted and sharpened F_o-F_c map is calculated after the refinement. The Fourier coefficients for the difference maps are derived as expectation values of unknown Fourier coefficients using their posterior distribution given observations and model parameters. Refinement of atomic displacement parameters is crucial for calculation of a sensible F_o-F_c map. It was shown to be useful for visualization of weak features like hydrogen atoms and model errors as it is done routinely in crystallography. Although hydrogen densities are weaker than heavier atoms (e.g. C, N, O), they are stronger than in the electron density maps produced by X-ray crystallography, and some hydrogen atoms are even visible at ~1.8 Å.

About half of the EMDB-deposited SPA maps have non-*C*1 point group symmetry. If the map has been symmetrised during reconstruction, then all downstream programs should be aware of it and the atomic structure model must follow the symmetry. A user can give an asymmetric unit model and a point group symbol to *Servalcat* for refinement. The NCS constraint function in *REFMAC5* was updated to consider non-bonded interactions and ADP similarity restraints between symmetry copies. The MTRIX records in the PDB format and _struct_ncs_oper in the mmCIF format are used to encode the symmetry information. Currently, there are only few asymmetric unit model depositions to the PDB except viruses. We think that refining and depositing asymmetric unit models with annotations of symmetry will be a common practice in future.

We are also developing a new program *EMDA*, for cryo-EM map and model manipulation with the main focus on validation. EMDA offers several metrics for map validation including FSC combined with mask correction by high resolution noise substitution [2], local correlation using half maps, optimal alignment between maps and magnification scaling using maximum-likelihood method. Also, EMDA includes metrics for map-model validation such as local correlation between map-and-model, which can be used to investigate the quality of the map-to-model fit.

Both *EMDA* [3] and *Servalcat* [4] are freely available under an open-source licence. They are also available within the CCP-EM package.

[1] Murshudov (2016). Methods in Enzymology 579, 277-305.

[2] Scheres and Chen (2012). *Nature Methods* **9**, 853-854

[3] EMDA https://emda.readthedocs.io

[4] Servalcat https://github.com/keitaroyam/servalcat

Keywords: cryoEM, structure refinement, validation, difference map, symmetry

This work was supported by the Medical Research Council, as part of UK Research and Innovation (MC_UP_A025_1012 to K.Y. and G.N.M.) and the Wellcome Trust (grant number 208398/Z/17/Z to R.W.).