

Validation of cryo-EM structures of SARS-CoV-2 and mapping genomic mutations

Sony Malhotra, Martyn Winn, Agnel Praveen Joseph

Scientific Computing Department, Science and Technology Facilities Council, Rutherford Appleton Laboratory, Harwell, United Kingdom.

agnel-praveen.joseph@stfc.ac.uk

Technological developments and growing interests in the study of cellular assemblies have led cryo-EM as a powerful technique to solve the three-dimensional structures of macromolecular complexes. More than 450 structures of molecular complexes from SARS-CoV-2 have been solved using cryo-EM and the structures have been crucial to understand molecular details behind the viral infection and development of drug molecules and vaccines. Given that majority of the cryo-EM reconstructions are solved at resolutions worse than 2.5Å and often the local resolution within the map varies considerably, atomic model validation is crucial to identify errors and less reliable areas of the model.

Here, we present the database CoVal, which is a repository of amino acid replacement mutations identified in the SARS-CoV-2 genome sequences, mapped onto protein structures from cryo-EM and X-ray crystallography. We provide information on the demographic distribution of these mutations, along with details on co-occurring mutations. CoVal gives easy access to mutation sites mapped to known structures with multiple metrics on the quality of the structure and agreement with experimental data. We also provide validation scores for the local quality of mutation site(s) and their structural neighbors. The database is freely accessible at: <https://coval.ccpem.ac.uk>. We also discuss tools for atomic model validation in the CCP-EM software suite [1] and results from the validation analysis of cryo-EM structures from SARS-CoV-2.

[1] Burnley, T., Palmer, M. C., Winn, M. (2017). *Acta Cryst. D.* **73**, 469-477.

Keywords: cryo-EM; SARS-CoV-2; model validation; mutations

This work was supported by the Wellcome Trust research grant WT 208398/Z/17/Z.