

MS12-03 | PHASER.VOYAGER: DATA-GUIDED MODEL GENERATION AND VISUALIZATION

Sammito, Massimo D (University of Cambridge, Cambridge, GBR); McCoy, Airlie J (University of Cambridge, Cambridge, GBR); Hatti, Kaushik (University of Cambridge, Cambridge, GBR); Oeffner, Robert D (University of Cambridge, Cambridge, GBR); Stockwell, Duncan H (University of Cambridge, Cambridge, GBR); Croll, Tristan I (University of Cambridge, Cambridge, GBR); Read, Randy (Department of Haematology, Cambridge Institute for Medical Research, University of Cambridge, Cambridge, GBR)

Although molecular replacement is a mature technique, new and established users alike can still find it challenging to generate ensemble models and to cope with uncertainty about the contents of the asymmetric unit. Ensembles should include information not only from homologous domains, but also their assembly into macromolecular complexes. Moreover, when multiple homologs are available with varying predicted structural similarity to the target, their optimal clustering into ensembles should depend also on the data resolution. We have developed an automated procedure for building ensembles that accounts for data quality. Ensemble generation is calibrated to the difficulty of the problem, via the expected LLG [1]. Our procedure to rank possibilities for asymmetric unit content takes into account any translational non-crystallographic symmetry and previously observed oligomeric states. We expect to optimize the path to structure solution particularly in challenging cases involving large complexes, low-resolution data, high order non-crystallographic symmetry, low-homology models and combinations thereof.

The automated procedure, *phaser.voyager*, is supported by a graphical interface to

- Visualize data features and their corrections to identify sources of error
- Automate the retrieval, selection and preparation of atomic models, building ensembles optimized to account for limitations in experimental data.
- Guide user decisions including description of the crystal content
- Visualize in real-time the MR tree search, introducing a new user interaction that will allow dynamically branching alternative MR strategies from any intermediate step.
- Optimize the MR solution to output the best model and map.

[1] Airlie J. McCoy et al., (2017) PNAS 114(14):3637-3641