MS3-O3 Analysis of multi-crystal datasets with xia2

Richard J. Gildea[1], Graeme Winter[1], Gwyndaf Evans[1]

1. Diamond Light Source

email: richard.gildea@diamond.ac.uk

Continued development of microfocus beamlines has enabled data collection from smaller crystals than ever before, with commensurate decrease in signal, whilst simultaneous advancements in beamline automation and detector technology permit ever faster and lower noise data collections. Consequently, automated data processing plays an essential part both during and after data collection. Furthermore, it is often necessary to merge together many partial datasets in order to obtain a complete dataset with sufficient signal-to-noise for subsequent structure solution and analysis. We present an overview of recent developments in xia2 [1] with an emphasis on automated processing of multi-crystal datasets and the incorporation of BLEND [2]. We will discuss improvements to parallel processing of multiple datasets in xia2, and highlight new tools for analysis of multi-crystal datasets, including clustering on unit cell parameters [2] and intensities [3, 4].

[1] G. Winter, J. Appl. Cryst. (2010). 43, 186-190

[2] J. Foadi, P. Aller, Y. Alguel, A. Cameron, D. Axford, R. L. Owen, W. Armour, D. G. Waterman, S. Iwata and G. Evans, *Acta Cryst.* (2013). D69, 1617-1632.

[3] R. Giordano, R. M. F. Leal, G. P. Bourenkov, S. McSweeney and A. N. Popov, *Acta Cryst.* (2012). D68, 649-658.

[4] Q. Liu, T. Dahmane, Z. Zhang, Z. Assur, J. Brasch, L. Shapiro, F. Mancia, and W. A. Hendrickson, *Science* (2012). 336, 1033-1037.

Keywords: xia2, blend, micro-crystal, multi-crystal, automation

MS3-O4 MoRDa, an automatic molecular replacement pipeline

Alexey Vagin[1], Andrey Lebedev[2]

1. Harwell, Didcot, OX11 0HX, UK
2. STFC Rutherford Appleton Laboratory, Didcot OX11 0FA, UK

email: vagin44@gmail.com

MoRDa is a pipeline for molecular replacement protein structure solution using x-ray data. The software package includes a database and a set of programs for the structure solution. The database (2G) is derived from the PDB and contains a compact description of non-redundant protein chains, domains, homo- and hetero-oligomers. The domain models are pre-processed to exclude flexible loops, which may hinder molecular replacement search. The automatic structure solution involves the search of the homologous chains and domains in the database, further modification of domain and chain models in accordance with the target sequence, composing oligomeric models if possible, and molecular replacement trials (Molrep) with the several search models. Accordingly, trials may include sequential searches for different domains or different chains, dependent on the target sequence and available models. In all cases, ensemble models are generated if there are several homologues of a particular domain or chan. Intermediate and final solutions are refined (Refmac), and selection of the most likely solution is done based on results of the refinements. User input consists of fasta-format file with one or more protein sequences and mtz- or cif-file with experimental structure amplitudes. The complete software package operates off-line. A smaller size package (0.9G) with coordinate files excluded requires Internet connection in order to download necessary files from the PDB.

Keywords: crystallography, pipeline