

xia2.multiplex*: a multi-crystal data analysis pipeline*R. J. Gildea, J. Beilsten-Edmands, D. Axford, S. Horrell, R. L. Owen, G. Winter***Diamond Light Source Ltd, Harwell Science and Innovation Campus, Didcot, Oxfordshire, OX11 0DE**richard.gildea@diamond.ac.uk*

In macromolecular crystallography, radiation damage limits the amount of data that can be collected from a single crystal. It is often necessary to merge multiple data sets from one or more crystals, for example multiple small-wedge data collections on micro-crystals, in situ room temperature data collections, lipidic mesophase data collections or time-resolved crystallography. Whilst indexing and integration of individual data sets may be relatively straightforward with existing software, additional challenges are commonly encountered when merging multiple data sets. For novel structures, identification of a consensus symmetry can be problematic, particularly in the presence of a potential indexing ambiguity. The presence of non-isomorphous or poor-quality data sets may degrade the overall quality of the merged data set.

To facilitate and help optimise the scaling and merging of multiple data sets, we developed a new program, *xia2.multiplex*, which takes as input the results of data sets individually integrated with *DIALS* [1] and performs symmetry analysis [2], scaling [3] and merging of multi-crystal data sets, as well as analysis of various pathologies that typically affect multi-crystal data sets, including non-isomorphism, radiation damage [4] and preferred crystal orientation.

xia2.multiplex has been deployed as part of the autoprocessing pipeline at Diamond Light Source, including integration with downstream phasing pipelines such as *DIMPLE* [5] and *Big EP* [6].

Using data sets collected as part of in situ room-temperature fragment screening experiments on the SARS-CoV-2 main protease, we demonstrate the use of *xia2.multiplex* within a wider autoprocessing framework to give rapid feedback during a multi-crystal experiment, and how the program can be used to further improve the quality of final merged data set.

[1] Winter, G., Waterman, D. G., Parkhurst, J. M., Brewster, A. S., Gildea, R. J., Gerstel, M., Fuentes-Montero, L., Vollmar, M., Michels-Clark, T., Young, I. D., Sauter, N. K. & Evans, G. (2018). *Acta Crystallographica Section D*.

[2] Gildea, R. J. & Winter, G. (2018). *Acta Crystallographica Section D*, 74(5), 405–410.

[3] Beilsten-Edmands, J., Winter, G., Gildea, R., Parkhurst, J., Waterman, D. & Evans, G. (2020). *Acta Crystallographica Section D*, 76(4), 385–399.

[4] Winter, G., Gildea, R. J., Paterson, N. G., Beale, J., Gerstel, M., Axford, D., Vollmar, M., McAuley, K. E., Owen, R. L., Flaig, R., Ashton, A. W. & Hall, D. R. (2019). *Acta Crystallographica Section D*, 75(3), 242–261.

[5] <http://ccp4.github.io/dimple/>

[6] Sikharulidze, I., Winter, G. & Hall, D. R. (2016). *Acta Crystallographica Section A*, 72(a1), s193.

Keywords: Multi-crystal; data analysis