Scipion-ed: a workflow-based and self-documenting program for 3DED data processing

V.E.G. Bengtsson[1], J.M. de la Rosa-Trevin[2], L. Pacoste[1], G. Hofer[1], H. Xu[1], X. Zou[1]

[1]Department of Materials and Environmental Chemistry, Stockholm University, Stockholm, [2]Department of Biochemistry and Biophysics, Science for Life Laboratory, Stockholm University, Stockholm

viktor.bengtsson@mmk.su.se

Three-dimensional electron diffraction (3DED) is an increasingly used complementary technique to X-ray crystallography and cryo-electron microscopy (cryo-EM) [1]. Using methods such as continuous rotation electron diffraction (cRED) [2] and microcrystal electron diffraction (MicroED) [3] with hybrid pixel detectors [4,5] allows fast and increasingly automated 3DED data collection[6]. These data can be processed with software commonly used for X-ray crystallography, such as XDS [7] and DIALS [8]. These programs do not have built-in graphical user interfaces (GUIs), but separate such interfaces do exist[9–12]. However, the interfaces are typically created for processing of individual datasets or data from synchrotron beamlines. As electron diffraction data and experimental geometries are less standardised, they are more likely than X-ray data to require user correction during processing. This is a challenge in automated processing, where relevant parameters can be hard or impossible to change in current interfaces.

We have developed a new GUI for processing 3DED data: scipion-ed. It is optimised for 3DED and intended to be easy for beginners to learn without preventing experienced users of the underlying programs from using specialised or newly implemented functions and parameters. Scipion-ed is an extension of the Scipion framework[13] consisting of python modules and originally designed for processing cryo-EM micrographs by combining multiple underlying programs.

The core module of the Scipion framework is scipion-pyworkflow. It defines the common GUI for all modules, a database for tracking metadata, job execution and the basic workflow engine. The workflow connects multiple steps in a processing pipeline together, and allows the user to split the processing in order to try out different methods or parameters following a certain step without overwriting previous processing. Such workflows can be exported as JSON files and deposited along with the data, imported and reused as the basis for processing of different data, or sent to collaborators.

Layered on top of the core modules are the base plugins scipion-em and scipion-ed. These define basic protocols, metadata structures and other domain-specific information for electron microscopy and electron diffraction respectively. A third layer of plugins is then added with protocols specific for individual software packages. Currently implemented into scipion-ed is scipion-ed-dials. This includes protocols for all typical processing steps: importing diffraction patterns, spot finding, indexing, model refinement, integration and scaling.

The input for each protocol is provided through a window with fields or lists of options for commonly used parameters and default values already filled in. It is thus simple to change individual values. For experienced users it is also possible to add command line parameters directly, enabling the use of additional options that might not yet have been implemented into scipion-ed. In addition to the protocols, scipion offers viewers that quickly show important results of a protocol without having to read the full output. The results summary for scipion-ed-dials protocols also includes buttons for opening the corresponding data in DIALS' image viewer and reciprocal lattice viewer.

It is possible to add plugins for other software, such as XDS or the shelx suite, in the future. Work is also ongoing to implement streaming processing, where a prepared set of protocols are run automatically as soon as the data are available.

[1]   M. Gemmi, E. Mugnaioli, T. E. Gorelik, U. Kolb, L. Palatinus, P. Boullay, S. Hovmöller, J. P. Abrahams, *ACS Central Science* **2019**, DOI 10.1021/acscentsci.9b00394.

[2]   Y. Wang, T. Yang, H. Xu, X. Zou, W. Wan, *J Appl Cryst* **2018**, *51*, 1094–1101.

[3]   B. L. Nannenga, T. Gonen, *Current Opinion in Structural Biology* **2016**, *40*, 128–135.

[4]   E. van Genderen, M. T. B. Clabbers, P. P. Das, A. Stewart, I. Nederlof, K. C. Barentsen, Q. Portillo, N. S. Pannu, S. Nicolopoulos, T. Gruene, J. P. Abrahams, *Acta Cryst A,* **2016**, *72*, 236–242.

[5]   X. Llopart, R. Ballabriga, M. Campbell, L. Tlustos, W. Wong, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **2007**, *581*, 485–494.

[6]   B. Wang, X. Zou, S. Smeets, *IUCrJ* **2019**, *6*, 854–867.

[7]   W. Kabsch, *Acta Cryst D* **2010**, *66*, 125–132.

[8]   G. Winter, D. G. Waterman, J. M. Parkhurst, A. S. Brewster, R. J. Gildea, M. Gerstel, L. Fuentes-Montero, M. Vollmar, T. Michels-Clark, I. D. Young, N. K. Sauter, G. Evans, *Acta Cryst D* **2018**, *74*, 85–97.

[9]   W. Kabsch, "XDS Package," can be found under http://xds.mpimf-heidelberg.mpg.de/, **2016**.

[10]   L. Fuentes-Montero, J. Parkhurst, M. Gerstel, R. Gildea, G. Winter, M. Vollmar, D. Waterman, G. Evans, IUCr, "Introducing DUI, a graphical interface for DIALS," **2016**.

[11]   L. Potterton, J. Agirre, C. Ballard, K. Cowtan, E. Dodson, P. R. Evans, H. T. Jenkins, R. Keegan, E. Krissinel, K. Stevenson, A. Lebedev, S. J. McNicholas, R. A. Nicholls, M. Noble, N. S. Pannu, C. Roth, G. Sheldrick, P. Skubak, J. Turkenburg, V. Uski, F. von Delft, D. Waterman, K. Wilson, M. Winn, M. Wojdyr, *Acta Cryst D* **2018**, *74*, 68–84.

[12]   K. Yamashita, K. Hirata, M. Yamamoto, *Acta Cryst D* **2018**, *74*, 441–449.

[13]   J. M. de la Rosa-Trevín, A. Quintana, L. del Cano, A. Zaldívar, I. Foche, J. Gutiérrez, J. Gómez-Blanco, J. Burguet-Castell, J. Cuenca-Alba, V. Abrishami, J. Vargas, J. Otón, G. Sharov, J. L. Vilas, J. Navas, P. Conesa, M. Kazemi, R. Marabini, C. O. S. Sorzano, J. M. Carazo, *Journal of Structural Biology* **2016**, *195*, 93–99.

Keywords: three-dimensional electron diffraction; 3DED; scipion-ed; software; data processing