

*Automation and inclusion of additional information for biological solution SAXS*Nelly Ros Hajizadeh¹, Daniel Franke¹, Alexey G. Kikhney¹, Dmitri I. Svergun¹¹EMBL Hamburg, Hamburg, Germany

E-mail: nhajizadeh@embl-hamburg.de

Small angle X-ray scattering (SAXS) of biological macromolecules in solution is widely used to assess the size, volume, shape, and flexibility of proteins in near physiological conditions. SAXS is extremely useful as a standalone technique but also in hybrid investigations where information from other methods such as X-ray crystallography, NMR spectroscopy and electron microscopy is combined. Solution SAXS at high-brilliance facilities often allows a large variety of experiments, from high-throughput data collections with robotic sample changers, continuous-flow in-line sample purification and characterization to time resolved applications and even customized sample environments [1]. A dedicated BioSAXS beamline must be able to robustly collect and real-time analyze the output of high-throughput experiments while accommodating the different sample environments and their unique demands on the beamline configuration.

The BECQUEREL user interface at the BioSAXS beamline P12 at PETRA III, DESY, Germany, combines a lean layout with the capabilities of experiment-specific profiles and the power of customizable data collection scripts to allow a user to make the most out of the available beam time. Access levels to hardware devices and their functionality help novice users to avoid accidents while a recommender system guides them through the steps to the next data collection. Power users and beamline scientists are given more complex functionality, including automated alignment operations or manual scans across motor ranges, e.g. to align customized sample environments in the beam. The data collection scripts are designed to minimize dead time during the measurement sequence, e.g. an empty capillary shot is taken while the sample is loading, which allows monitoring of the stability of the instrument, and therefore indirectly the quality of the data. During automated data collection, the user is free to prepare more samples or queue additional measurements.

Following data collection, the acquired data is processed by the SASFLOW data analysis pipeline [2,3]. The pipeline is freely configurable and provides real-time feedback about the sample characteristics, enabling the user to modify the sample conditions if necessary. SASFLOW executes the primary processing steps, including absolute calibration using water, after which it automatically evaluates the overall geometrical parameters of the macromolecules and reconstructs an ab initio model. Latest additions include shape classification, ambiguity assessment as well as determination of the useful data range [3]. Furthermore, the pipeline is able to exchange information with the laboratory-information management system ISPyB. This integration allows the user to conveniently pass over the metadata (sample and buffer names and compositions, concentrations) as well as information from other methods, e.g. available high resolution models. The additional information stored in ISPyB can also be used to evaluate model fits conduct analysis of mixture, and also process the data from membrane proteins (presently in preparation). The modularity of SASFLOW facilitates adaptability and portability as recently demonstrated by its adaption at the Shanghai Synchrotron Radiation Facility.

[1] C. Blanchet et al (2015). J. Appl. Cryst. 48 (2)

[2] D. Franke et al. (2012). Nuc. Inst. Meth. A. 52-59

[3] D. Franke et. al (2017). Submitted

Keywords: [automation](#), [BioSAXS](#), [data analysis](#)