

## An objective metric to guide background correction and interpretation of small angle X-ray scattering data

**Yunyun Gao<sup>1,2</sup>, Timothy R. Stachowski<sup>3</sup>, Edward H. Snell<sup>3</sup>, Thomas D. Grant<sup>3</sup>, Arwen R. Pearson<sup>1</sup>**

<sup>1</sup>*Institute of Nanostructure and Solid State Physics, Universität Hamburg, Hamburg, Germany;*

<sup>2</sup>*The Max Planck Institute for the Structure and Dynamics of Matter, Hamburg, Germany;*

<sup>3</sup>*Hauptman-Woodward Medical Research Institute, Buffalo, NY, USA*

*yunyun.gao@cfel.de*

Small-angle X-ray scattering (SAXS) experiments provide an estimation of biological macromolecule geometry on the level of domain structure. The reliability of structural inference drawn from SAXS data is dependent on the accurate measurement as well as the proper post-processing procedure. The methods improving raw data quality and gaining more information are widely explored. Among those innovations, size-exclusion chromatography small-angle X-ray solution scattering (SEC-SAXS) has become a standard method for modern bio-SAXS synchrotron light sources [1-3]. However, the principle of data post-processing for SEC-SAXS remains rather unclear. This includes background correction and averaging of the raw data. Several statistical tools have been developed to assess solution SAXS data quality [4, 5]. These are mostly useful for “rejecting” significantly different data points or data frames, based on the assumption that the rest of the data are close to the “truth”. But this can lead to a situation where mediocre data, for example data contaminated with radiation damage, are not correctable or simply cannot be evaluated before any interpretation is done.

To alleviate this problem, an objective metric, correction-state score (CSS) is proposed. CSS can be used to both verify the data quality and identify the optimal data correction procedure for post-processing of SEC-SAXS data. CSS can be represented as a numerical likelihood with a scale of 0 to 1. Using this objective score it is possible to quantitatively assess the “goodness” or appropriateness of a background correction for SEC-SAXS data. Under the guidance of CSS, the metadata recorded during a SEC-SAXS experiment can be used to maximise the fidelity of the post-processing as well as reduce the ambiguity in further data interpretation.

1. Ryan, T.M., Trehella, J., Murphy, J.M., et al. 2018. An optimized SEC-SAXS system enabling high X-ray dose for rapid SAXS assessment with correlated UV measurements for biomolecular structure analysis. *Journal of Applied Crystallography* 51(1), pp. 97–111.
2. Brennich, M.E., Kieffer, J., Bonamis, G., et al. 2016. Online data analysis at the ESRF bioSAXS beamline, BM29. *Journal of Applied Crystallography* 49(1), pp. 203–212.
3. Blanchet, C.E., Spilotros, A., Schwemmer, F., et al. 2015. Versatile sample environments and automation for biological solution X-ray scattering experiments at the P12 beamline (PETRA III, DESY). *Journal of Applied Crystallography* 48(Pt 2), pp. 431–443.
4. Franke, D., Jeffries, C.M. and Svergun, D.I. 2015. Correlation Map, a goodness-of-fit test for one-dimensional X-ray scattering spectra. *Nature Methods* 12(5), pp. 419–422.
5. Rambo, R.P. and Tainer, J.A. 2013. Accurate assessment of mass, models and resolution by small-angle scattering. *Nature* 496(7446), pp. 477–481.

**Keywords: SAXS, meta-data, objectivity, correction**