Structural bioinformatics

CHROMIXS: automatic and interactive analysis of chromatography-coupled small-angle X-ray scattering data

Alejandro Panjkovich* and Dmitri I. Svergun*

European Molecular Biology Laboratory, Hamburg Outstation, EMBL c/o DESY, 22607 Hamburg, Germany

*To whom correspondence should be addressed.

Received on October 26, 2017; revised on December 21, 2017; editorial decision on December 23, 2017; accepted on December 27, 2017

Abstract

Summary: Size-exclusion chromatography (SEC) coupled to small-angle X-ray scattering (SAXS), also known as inline SEC-SAXS, is being increasingly used for the structural analysis of biological macromolecules, complexes and mixtures in solution. A single SEC-SAXS run generates thousands of individual SAXS profiles from the eluting solute and their analysis requires a correct identification of buffer and sample regions, a rather laborious task. We present CHROMIXS (as in CHROMatography Inline X-ray Scattering), a program for rapid reduction and analysis, both automatically and interactively, of SEC-SAXS data.

Availability and implementation: CHROMIXS is freely available to academic users as part of the ATSAS software suite (www.embl-hamburg.de/biosaxs/download.html).

Contact: apanjkovich@embl-hamburg.de or svergun@embl-hamburg.de

1 Introduction

Small-angle X-ray scattering (SAXS) is becoming a standard tool for the structural characterization of biological macromolecules in solution (Vestergaard, 2016). Using a small (below 1 mg) quantity of purified material, SAXS provides the overall size and overall structure of macromolecules at a resolution of about 10–20 Å. In a typical batch SAXS experiment, scattering from the sample and buffer are measured separately for subsequent subtraction. Cases that are particularly difficult to assess and interpret, such as transient complexes or mixtures, highly benefit from directly coupling the elution on a size-exclusion chromatography column (SEC) to the SAXS sample exposure unit (SEC-SAXS). These experiments can separate the individual components and provide their individual SAXS patterns. During a SEC-SAXS experiment, scattering curves (frames) are recorded continuously every (few) second(s) during the elution of the sample (usually ~30–60 min). The experiment can be coupled to ultra-violet and additional detectors to further characterize the sample biophysically (Jeffries et al., 2016). Instead of a handful of recorded scattering curves as in batch SAXS, a SEC-SAXS experiment produces thousands of un-subtracted curves among which optimal buffer and sample regions need to be identified. Until recently, the reduction and analysis of SEC-SAXS datasets was cumbersome and lengthy, even for experts using a combination of manual processing and ad-hoc scripts.

Here, we present CHROMIXS (as in CHROMatography Inline X-ray Scattering), a user-friendly tool that allows rapid visualization and reduction, both automatically and interactively, of SEC-SAXS datasets. CHROMIXS streamlines the work of expert users and it also allows non-experts to quickly evaluate the results of a SEC-SAXS run. Importantly, CHROMIXS can work fully unsupervised in automatic mode, identifying sample and buffer regions from the elution profile to produce a final subtracted SAXS curve. The graphical interface provides additional functionality, such as automatic estimation of the radius of gyration ($R_g$) and molecular weight values, with options to export both raw data and scalable vector graphics (SVG) for the preparation of publication quality figures (Fig. 1).

2 Description

CHROMIXS is written in C++ and relies on Qt and Qwt for the cross-platform graphical interface, on libsaxsdocument for input/output formats and partially on ATSAS specific library functions for data processing (Franke et al., 2017) (some functions were re-written to optimize SEC-SAXS data processing). The program is cross-platform,
For publication-quality figures, CHROMIXS can export plots in SVG open standard format. The intensity versus frame number curve and $R_g$ values estimated by CHROMIXS can also be exported as raw data in CSV format, for customized plotting or treatment. Alternatively, data can be saved in BioSAXS’ native format, to continue the analysis later. The users can also utilize the ‘open selected frames in Primus’ menu to redirect selected buffer or sample frames to PRIMUS/qt program for the detailed analysis of individual frames, including statistical comparisons (Franke et al., 2017).

CHROMIXS can run completely unsupervised for high-throughput processing without activating the graphical interface. In this fully automatic mode, CHROMIXS loads all frames in the data directory, predicts sample(s) and buffer regions and produces subtracted curve(s). Details of the selection and results are written as metadata in the output file for future reference. An automatic reduction of a bovine serum albumin (BSA) SEC-SAXS dataset consisting of 1800 SAXS one-dimensional frames (Fig. 1) takes a few seconds on a standard desktop computer.

3 Concluding remarks

CHROMIXS allows for a convenient manual or automated reduction of SEC-SAXS runs with well resolved fractions (i.e. baseline separated sample elution peaks). In more intricate cases, CHROMIXS may help in diagnosing and preparing the data for further processing by other programs (as these require buffer frames to be manually defined or subtracted beforehand). For example, overlapping peaks may be decomposed by BioXTAS RAW (Hopkins et al., 2017) or UltraScan-SOMO (Brookes et al., 2016), the latter also providing a function to ease capillary fouling issues.

Currently, CHROMIXS is being routinely used at the EMBL Hamburg P12 BioSAXS beamline, where an increasing percentage of measurements are performed in SEC-SAXS mode. CHROMIXS reads plain ASCII files in and it was tested on the experimental data generated by other SAXS beamlines (e.g. from synchrotrons ESRF, Diamond and Soleil). In addition to a fully automatic SEC-SAXS processing function, CHROMIXS provides a responsive and easy-to-use interface, optimal errors propagation and experiment quality assessment together with thorough data and figure export features.

Acknowledgements

The authors would like to thank Patrice Vachette, Martha Brennich and Grzegorz Chojnowski for advice and for providing data in different formats to test the program. Acknowledgements go as well to the BioSAXS team at EMBL Hamburg P12 BioSAXS beamline, particularly Cj Jeffries, Melissa Graewert, Haydyn Mertens, Alexey Kikhney, Clement Blanchet, Daniel Franke and Nelly Hajizadeh for useful suggestions and discussions.

Funding

Funding for this research was provided by the Horizon 2020 programme of the European Union, iNEXT (H2020 Grant # 653706). A.P. was supported by the EMBL Interdisciplinary Postdoc Programme under Marie Curie COFUND Actions.

Conflict of Interest: none declared.

References


