Archives of Biochemistry and Biophysics xxx (2016) 1–11

Contents lists available at ScienceDirect

Archives of Biochemistry and Biophysics

journal homepage: www.elsevier.com/locate/yabbi

Analysis of biostructural changes, dynamics, and interactions – Small-angle X-ray scattering to the rescue☆

Bente Vestergaard

Department of Drug Design and Pharmacology, University of Copenhagen, Jagtvej 162, 2100 Copenhagen, Denmark

ARTICLE INFO

Article history:
Received 18 January 2016
Received in revised form
17 February 2016
Accepted 26 February 2016
Available online xxx

Keywords:
Small angle X-ray scattering
Complex structural analysis
Intrinsically disordered proteins
Advanced sample environments
Hybrid biostructural analysis

ABSTRACT

Solution small angle X-ray scattering from biological macromolecules (BioSAXS) plays an increasingly important role in biostructural research. The analysis of complex protein mixtures, dynamic equilibriums, intrinsic disorder and evolving structural processes is facilitated by SAXS data, either in stand-alone applications, or with SAXS taking a prominent role in hybrid biostructural analysis. This is not the least due to the significant advances in both hardware and software that have taken place in particular at the large-scale facilities. Here, recent developments and the future potential of BioSAXS are reviewed, exemplified by numerous examples of elegant applications to challenging systems.

1. Frontiers in biostructural analysis

Prime challenges in biostructural analysis today include the investigation of structural changes, dynamics and polydispersity. The accurate and detailed description of biomacromolecules, derived from high-resolution structures and primarily originating from macromolecular crystallography (MX), has provided the research community with a wealth of insight, greatly enabling our current understanding of biomolecular function. Biomacromolecules, however, are inherently dynamic at several time- and length-scales [1], and this dynamic behavior is crucial for their biological function. The cell is a crowded and ever-changing environment where proteins, lipids, nucleic acids, polysaccharides and other biomolecules interact in a structurally responsive and adaptive manner. A single protein structure should thus not be considered as a three-dimensional rigidly defined entity, but rather be understood as a spatiotemporal distribution of an ill-defined number of conformational states, and this ensemble of conformations defines the biological function of the protein. Likewise, when considering macromolecular interactions (e.g. protein:protein interactions [2]), structural polydispersity plays a significant role: often complexes are only partially and/or transiently formed, and complex formation in addition may induce or require different levels of conformational changes of the individual protomers. One may say that the functionality of macromolecular interactions is defined by a highly refined interplay between two macromolecular entities, each defined as a complex ensemble of structures, and that this interplay introduces an additional level of structural complexity, not exhibited by each individual structure prior to the encounter.

This view on structures challenges biostructural analysis in general, and high-resolution structural investigation in particular. Small angle scattering (SAS), on the other hand, is uniquely suited for these experimental endeavors. In a SAS experiment the sample is in the liquid state, and the experiment can be performed under physiologically or otherwise experimentally relevant conditions, since there are no particular requirements to sample preparation. SAS is a low-resolution method, and hence by no means replaces, but rather uniquely complements high-resolution biostructural analysis. As a consequence, the SAS community is experiencing an almost explosive development in numerous ways: the available software and hardware rapidly advances, the complexity of the scientific questions that are addressed is continuously increasing [3], and there is an ever-

Abbreviations: SAS, small angle (X-ray/neutron) solution scattering; SAXS, small angle X-ray scattering; SANS, small angle neutron scattering; BioSAXS, biological solution SAXS; BioSANS, biological solution SANS; SEC-SAXS, size exclusion chromatography coupled with SAXS data collection; TR, time resolved; ISpyB, Information System for Protein Crystallography Beamlines; ISpyBB, ISpyB for BioSAXS; MW, molecular weight; SASBDB, Small Angle Scattering Biological Data Bank; PDB, protein data bank; RF2, polypeptide release factor 2; IDP, intrinsically disordered protein; NMR, nuclear magnetic resonance; MX, macromolecular crystallography; WAXS, wide angle X-ray scattering.

☆ This article is part of a Special Issue entitled Protein Crystallography, edited by Ana Camara-Artigas and Jose Antonio Gavira.

E-mail address: bente.vestergaard@sund.ku.dk.

http://dx.doi.org/10.1016/j.abb.2016.02.029

growing user community with a significant and growing scientific production [4]. Indeed, the days where SAS was the last resort for the crystallographer *in spe* are over, and the era where SAS greatly empowers biostructural understanding is ongoing.

2. Increasing data quality and decreasing sample quantity

The intensities of X-rays and, to a lesser extent, neutrons at international large-scale facilities are dramatically increasing. This drives impressive hardware advances at the leading established and upcoming facilities within both biological solution small angle X-ray scattering (BioSAXS) [5–11] and neutron scattering (BioSANS) [12] facilities. Not only have the beam parameters (such as brilliance and positional stability) constantly improved, so have the front-end hardware and software developments, the latter dramatically increasing user-friendliness. At several BioSAXS facilities robotics ensure low sample consumption and robust, rapid sample mounting [7,13,14]. Thereby, not only is sample/data-throughput significantly improved, also, the standardized mounting minimizes user mistakes and includes an optimized cleaning procedure, while oscillation of the sample in the capillary diminishes radiation damage, with the net effect of improved data quality. The sensitivity and speed of detectors has also undergone revolutionary development, and one consequence is that now, multiple data-frames from very short exposures are collected for the data collection from a single sample, rather than performing one long exposure. This enables comparison of individual data-frames, and hence detection of potential radiation damage from the high-brilliance X-ray beams, thereby ensuring that the final dataset for a given sample is averaged only from the frames devoid of radiation damage.

A highly useful tool which was recently developed, and which, among many other useful applications, can detect the potential onset of radiation damage during multiple exposures, is the correlation map (CorMap) [15]. With this tool, it is possible to perform pair-wise or multiple comparisons of data-curves, independent of the error estimates. The method is highly sensitive, and has a large range of important applications, some of which are mentioned in this review. One example, applied during data collection, is the continuous monitoring of the cleanliness of the sample cell. If collecting a large number of datasets, repeated buffer measurements can be compared, and should correlate perfectly, if the sample cell remains clean. The implementation of such automated procedures at the advanced beamlines greatly enhances the output for users visiting the facilities.

Also, the greatest part of basic data reduction, processing and analysis has been automated. These routines have been included in a major pipeline of data evaluation which call individual advanced software packages, providing the user with on-the-spot information about data quality, biophysical parameters and *ab initio* structures [16]. These and other automated procedures are also included in ISpyBB (Information System for Protein crystallography Beamlines (ISpyB) for BioSAXS) [17]. ISpyBB integrates data-management from the point of preparing the samples (strategy for data collection, calculation of the needed sample quantities etc.), over logging and controlling the data collection to providing the results of the initial automated data analysis in a GUI. In addition IspyBB enables access to and comparison with relevant data from the same or previous data collections [17]. With low sample consumption and rapid data collection BioSAXS users collect tens, hundreds, and sometimes thousands of datasets within a project period, which increases demands for comparative data evaluation, or even data archiving and tracking. It is planned to expand the program, such that the measured data guides additional sample preparation using liquid-handling robotics at high-throughput crystallization facilities. As an example, if the SAXS data reveal that certain conditions promote complex formation and minimize unspecific aggregation, the following data collection will use these or further optimized experimental conditions [17] which ultimately enables a scanning of a multidimensional experimental space, searching for relevant structural states of the investigated macromolecules [18].

Even lower sample consumption can be achieved by the use of microfluidic sample environments [6,18–21]. Some microfluidic systems aim at providing versatile off-the-shelf sample environments for standard screening purposes, however, microfluidics have also in several cases been applied in a customized setup with a particular purpose, such as the screening of structural changes in lipidic mesophases induced by the experimental conditions applied during membrane protein crystallization attempts [22], or the on-chip dialysis setup enabling *in situ* sample concentration and buffer exchange for fragile protein samples [23]. A particularly useful application when combining microfluidics and SAXS is in time-resolved (TR) studies. In fact, microfluidic mixing, both stopped-flow and continuous flow, has been used for more than two decades for the TR study of protein and nucleic acid un- and refolding [24–27]. TR studies are applied today with as low as nl sample consumption [20] and to increasingly complex samples, such as the intermediate nucleosome states during DNA unwinding [28] or intermediate filament formation formed under ionic gradients [29].

3. In-line sample purification and orthogonal data

More slowly evolving mixtures can be efficiently analyzed using SEC-SAXS (size exclusion chromatography coupled with BioSAXS data collection, Fig. 1) [5,30,31]. That is, SAXS data are collected as the sample is purified, by directing the liquid flow directly from the purification column and through the X-ray beam. As simple as this idea may sound, this development is by no means trivial, and has been made possible in part because the increasing intensities at synchrotron beamlines allow for data collection in very short time-intervals. That is, a very large number of data-frames are collected over the elution profile, each frame of a sufficient quality to allow for individual data analysis. Individual data-frames are hence comparatively evaluated, and data-frames from a single eluting species can be isolated in an (semi-) automated manner [5,32–34]. SEC-SAXS can dramatically improve data quality from aggregation prone proteins. If small amounts of un-specific aggregates co-exist with the molecule of interest, this compromises the SAXS data, and in many cases makes any attempt of further analysis futile. With SEC-SAXS, however, the sample of interest is separated from these aggregates and data are collected before the purified sample re-aggregates. SEC-SAXS is also very useful in the case of partial complex formation, where only a fraction of protomers form complexes, and these complexes co-exist with the non-complexed individual proteins (Fig. 1).

Analysis of such mixed states can be very difficult without significant prior information, but SEC-SAXS often allows sufficient separation of the species such that spectra can be obtained from the pure species. Recent examples of such successful analysis are e.g. the studies of prion protein in complex with antibodies [36], co-eluting monomers and dimers of intrinsically disordered alpha-synuclein [37], monomeric, oligomeric or degradation products of fibrinogen [33], MnME:MnmG complexes of different stoichiometric states [38] or co-existing apo- and ligand-bound forms of *Arabidopsis thaliana* acyl acid-amido synthetase [31]. A particularly useful application is in the analysis of membrane proteins [39]. Detergent solubilized membrane protein samples are a mixture of the protein:detergent complexes, soluble detergent and varying amounts of detergent micelles. Since successful analysis of SAXS data crucially relies on correct background subtraction, SAXS data

Please cite this article in press as: B. Vestergaard, Analysis of biostructural changes, dynamics, and interactions – Small-angle X-ray scattering to the rescue, Archives of Biochemistry and Biophysics (2016), http://dx.doi.org/10.1016/j.abb.2016.02.029