



ELSEVIER

Contents lists available at ScienceDirect

Data in Brief

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

Data Article

SAXS fingerprints of aldehyde dehydrogenase oligomers

John J. Tanner

Departments of Biochemistry and Chemistry, University of Missouri-Columbia, Columbia, MO 65211, United States

ARTICLE INFO

Article history:

Received 8 September 2015

Received in revised form

10 October 2015

Accepted 15 October 2015

Available online 25 October 2015

Keywords:

Small-angle X-ray scattering

X-ray crystallography

Aldehyde dehydrogenase

Protein oligomeric state

ABSTRACT

Enzymes of the aldehyde dehydrogenase (ALDH) superfamily catalyze the nicotinamide adenine dinucleotide-dependent oxidation of aldehydes to carboxylic acids. ALDHs are important in detoxification of aldehydes, amino acid metabolism, embryogenesis and development, neurotransmission, oxidative stress, and cancer. Mutations in genes encoding ALDHs cause metabolic disorders, including alcohol flush reaction (ALDH2), Sjögren-Larsson syndrome (ALDH3A2), hyperprolinemia type II (ALDH4A1), γ -hydroxybutyric aciduria (ALDH5A1), methylmalonic aciduria (ALDH6A1), pyridoxine dependent epilepsy (ALDH7A1), and hyperammonemia (ALDH18A1). We previously reported crystal structures and small-angle X-ray scattering (SAXS) analyses of ALDHs exhibiting dimeric, tetrameric, and hexameric oligomeric states (Luo et al., *Biochemistry* 54 (2015) 5513–5522; Luo et al., *J. Mol. Biol.* 425 (2013) 3106–3120). Herein I provide the SAXS curves, radii of gyration, and distance distribution functions for the three types of ALDH oligomer. The SAXS curves and associated analysis provide diagnostic fingerprints that allow rapid identification of the type of ALDH oligomer that is present in solution. The data sets provided here serve as a benchmark for characterizing oligomerization of ALDHs.

© 2015 The Author. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Specifications table [please fill in right-hand column of the table below]

Subject area	<i>Chemistry, Biology</i>
--------------	---------------------------

E-mail address: tannerjj@missouri.edu

<http://dx.doi.org/10.1016/j.dib.2015.10.017>

2352-3409/© 2015 The Author. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

More specific subject area	Biochemistry, Structural Biology
Type of data	SAXS data files (.dat) and protein structure coordinate files (.pdb)
How data was acquired	Small-angle X-ray scattering (SAXS) data collected at Advanced Light Source Beamline 12.3.1
Data format	Buffer-subtracted, merged experimental scattering curves (.dat)
Experimental factors	Purified protein samples were subjected to size exclusion chromatography and shipped at 4 °C in 96-well trays to beamline 12.3.1.
Experimental features	Data were collected by the beamline staff as part of the mail-in SAXS program at the SIBYLS beamline. The beamline staff provides the user with buffer-subtracted SAXS curves. The user then performs subsequent data analysis and interpretation.
Data source location	Lawrence Berkeley National Laboratory, Berkeley, CA, USA
Data accessibility	SAXS curves and coordinates of crystal structures are provided as supplementary content

1. Value of the data [describe in 3–5 bulleted points why this data is of value to the scientific community]

- SAXS is a robust method for determining the oligomeric states of proteins in solution.
- When combined with crystal structures, SAXS can also be used to determine quaternary structure.
- The dimeric, tetrameric, and hexameric forms of ALDH have distinctive SAXS curves and SAXS-derived structural parameters.
- SAXS provides a diagnostic fingerprint of ALDH oligomeric state and quaternary structure.
- The data sets provided here serve as a benchmark for characterizing ALDH oligomerization.

2. Data, experimental design, materials and methods

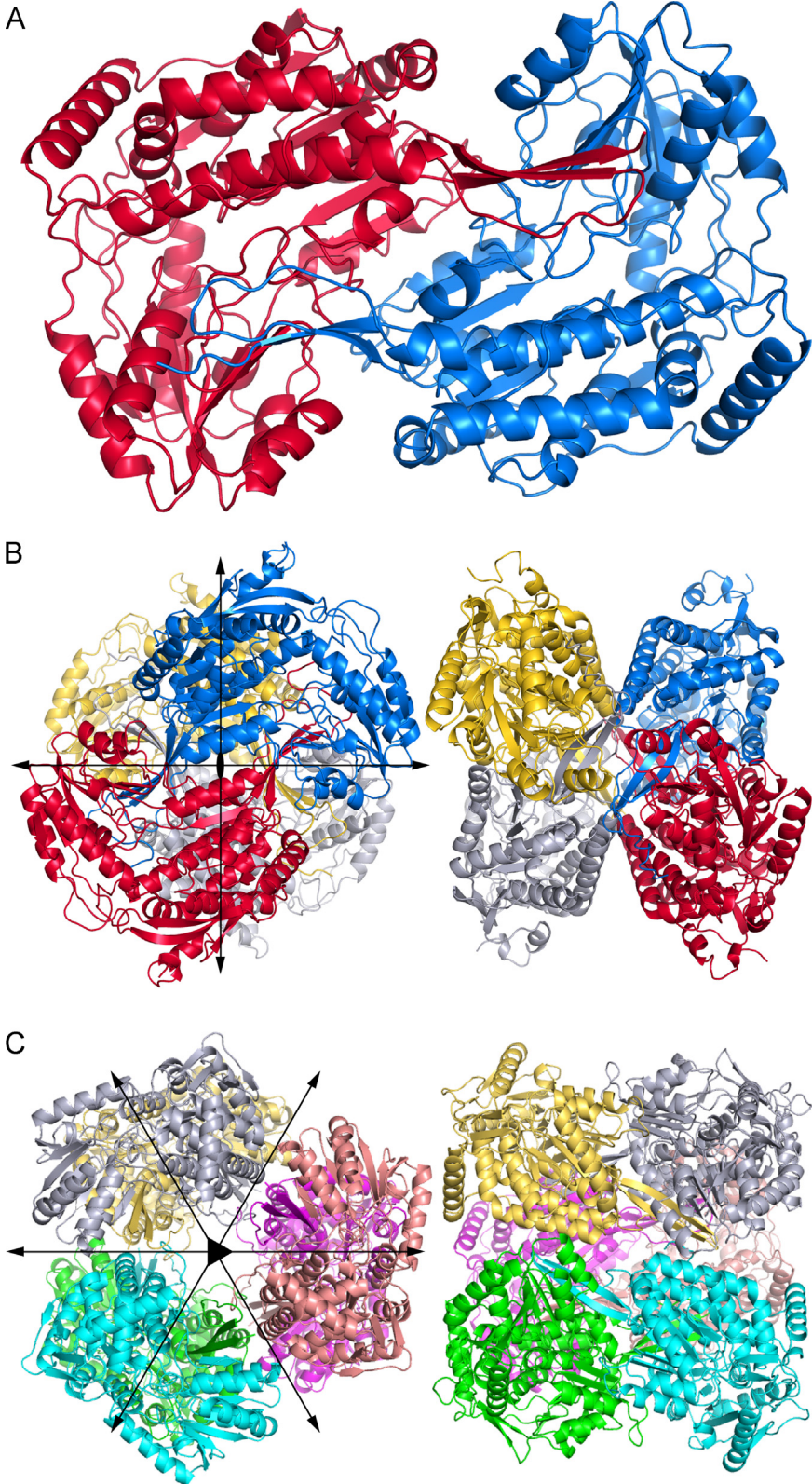
2.1. Representative examples of ALDH oligomers

SAXS fingerprints are provided for prototypical dimeric, tetrameric, and hexameric ALDHs. No other oligomeric forms of ALDH have been described to date.

Bacillus halodurans Δ^1 -pyrroline-5-carboxylate dehydrogenase (BhP5CDH) is presented here as an example of a dimeric ALDH. P5CDHs are part of proline catabolism and catalyze the oxidation of L-glutamate- γ -semialdehyde to L-glutamate [1]. P5CDHs belong to ALDH family 4 (member A1) and are also known as ALDH4A1. The ALDH dimer consists of two domain-swapped protomers and is the fundamental building block of higher order ALDH oligomers (Fig. 1A). The BhP5CDH dimer corresponds to chains A and B of the C2 asymmetric unit of PDB entry 3QAN; the coordinates of this dimer are provided in [Supplementary material](#). Other examples of dimeric ALDHs include human and mouse ALDH4A1 [2,3].

Human ALDH7A1 (hALDH7A1) forms the classic ALDH tetramer. ALDH7A1 is part of lysine catabolism and is also known as α -aminoadipate semialdehyde dehydrogenase [4–6]. The tetramer is a dimer of dimers having 222 symmetry (Fig. 1B). The crystal structure of hALDH7A1 complexed with α -aminoadipate (PDB entry 4ZUL [5]) has two equivalent tetramers in the C2 asymmetric unit; the coordinates of one of these tetramers are provided in [Supplementary material](#). Other examples of tetrameric ALDHs include ALDH1 and ALDH2 [7].

Thermus thermophilus Δ^1 -pyrroline-5-carboxylate dehydrogenase (TtP5CDH) represents hexameric ALDHs [8]. The hexamer is a trimer of dimers (Fig. 1C). The crystal structure of TtP5CDH (PDB



entry 2BHQ [9]) has a dimer in the H3 asymmetric unit. Application of the crystallographic 3-fold rotation generates the hexamer; the coordinates of this hexamer are provided in [Supplementary material](#). Other hexameric ALDHs include the P5CDHs from yeast [10] and *Deinococcus radiodurans* [8].

2.2. Preparation of protein samples for SAXS data collection

Expression and purification of BhP5CDH, hALDH7A1, and *Thermus thermophilus* P5CDH (TtP5CDH) were described previously [5,8]. Prior to SAXS analysis, each protein was passed through a Superdex 200 size exclusion column to remove any aggregated protein. We note that aggregation must be avoided as it greatly diminishes SAXS data quality, and in some cases, can cause the data to be unusable [11,12]. Effluent from the size exclusion column was reserved for measurement of the background scattering. The protein samples and corresponding buffer samples were pipetted into 96-well PCR plates (Corning Axygen, VWR catalog number 10011-228). Each protein sample was included at three nominal concentrations in the range of 1–10 mg/mL. The total volume in each well was 30 μ L. The trays were sealed with a silicone lid (Corning Axygen, VWR catalog number 10011-130). Each sealed tray was sandwiched between two cold packs that had been incubated at 4 $^{\circ}$ C, and the assembly was stabilized with rubber bands. The assembly was then placed in a Styrofoam box containing additional cold packs (at 4 $^{\circ}$ C) and sent via overnight express courier to beamline 12.3.1 of the Advanced Light Source.

2.3. SAXS data collection and analysis

SAXS data were collected by the beamline staff through the SIBYLS beamline mail-in program (bl1231.als.lbl.gov/htsaxs) [13,14]. For each protein concentration, exposure times of 0.5, 1.0, 3.0, and 6.0 s were used. Scattering curves collected from the protein samples were corrected for background scattering using intensity data collected from the SEC effluent. Composite scattering curves for each protein concentration were generated with PRIMUS [15] by scaling and merging the background-corrected high q region from the 3 s exposure with the low q region from the 0.5 s or 1.0 s exposure. PRIMUS was also used for Guinier analysis. GNOM was used to calculate distance distribution functions [16]. Composite scattering curves for BhP5CDH, ALDH7A1, and TtP5CDH are provided in [Supplementary materials](#).

2.4. SAXS fingerprints of ALDH oligomers

SAXS curves for BhP5CDH (dimer), hALDH7A1 (tetramer), and TtP5CDH (hexamer) are shown in [Fig. 2](#). The dimer curve is distinct from the others in that it is relatively featureless and monotonically decreasing with q in the region of $q < 0.15 \text{ \AA}^{-1}$. In contrast, the tetramer and hexamer curves show peak and valley features in the region $q = 0.075\text{--}0.15 \text{ \AA}^{-1}$. These features are more pronounced in the hexamer curve.

The radius of gyration (R_g) is a fundamental solution structural parameter that is quickly determined from SAXS data. R_g can be estimated from Guinier analysis or calculation of the distance distribution function [12,17]. Using either method, one finds that R_g increases with increasing degree of oligomerization ([Table 1](#)). The Guinier R_g values estimated with Primus using the supplied data files are $31.2 \pm 0.1 \text{ \AA}$ for the dimer (using qR_g range of 0.35–1.30), $37.9 \pm 0.5 \text{ \AA}$ for the tetramer ($qR_g = 0.36\text{--}1.30$), and $43.4 \pm 0.3 \text{ \AA}$ for the hexamer ($qR_g = 0.49\text{--}1.28$). The R_g values from calculations of the distance distribution function (real space R_g) are in good agreement with those from Guinier analysis

Fig. 1. The three oligomers of the ALDH superfamily. (A) BhP5CDH is an example of a dimeric ALDH (PDB code 3QAN). (B) Human ALDH7A1 is a dimer-of-dimers tetramer (PDB code 4ZUL). Two orthogonal views are shown. The filled oval and arrows represent the three molecular 2-fold axes of the tetramer. (C) TtP5CDH forms a trimer-of-dimers hexamer (PDB code 2BHQ). Two orthogonal views are shown. The triangle represents the molecular three-fold axis, while the arrows represent the three molecular 2-fold axes. In all three panels, each chain has a different color.

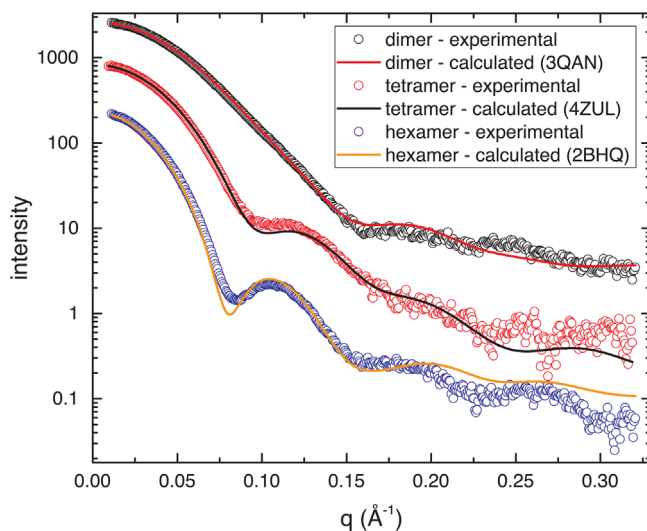


Fig. 2. SAXS data for dimeric (BhP5CDH, black circles), tetrameric (hALDH7A1, red circles), and hexameric (TtP5CDH, blue circles) ALDHs. The smooth curves are theoretical SAXS data calculated from the atomic models provided in the Supplement using FoXS [21]. The fit quality parameters (χ) are 2.2 for the dimer, 1.5 for the tetramer, and 5.5 for the hexamer. Arbitrary scale factors have been applied so that the data sets are offset for ease of comparison.

Table 1
Radii of gyration of ALDH oligomers.

Representative ALDH	Oligomeric state	Guinier R_G (Å)	Real space R_C (Å)	Crystal structure R_C (Å) ^a	M (kDa) ^b
BhP5CDH	Dimer	31	32	31	91 (115)
hALDH7A1	Tetramer	38	37	36	175 (222)
TtP5CDH	Hexamer	43	43	43	309 (343)

^a Calculated from the coordinates provided in [Supplementary materials](#) using MOLEMAN [19].

^b Molecular mass of the oligomer in solution calculated from the SAXS data using the volume of correlation method as implemented in Scatter 1.0 [20]. The theoretical mass calculated from the amino acid sequence is listed in parentheses.

(Table 1). Furthermore, the SAXS R_g values agree well with those calculated from the crystal structures (Table 1).

The molecular mass of proteins in solution can be calculated from SAXS data using the volume of correlation method of Rambo and Tainer [18]. The molecular masses calculated from the ALDH data sets are in good agreement with the theoretical values (Table 1).

Theoretical SAXS data can be calculated from atomic coordinates obtained from crystal structures or homology models. Comparison of the experimental and calculated data allows determination of the quaternary structure of the oligomer in solution. The theoretical SAXS data calculated from the supplied oligomer crystal structure models agree well with the experimental SAXS data (Fig. 2).

The distance distribution function is another aspect of the SAXS fingerprints of ALDH oligomers (Fig. 3). In each case, the distribution function exhibits a single maximum; however, the oligomers can be distinguished by the position of the maximum, the width of the distribution, and the maximum particle dimension (D_{max}). The position of the maximum increases with increasing degree of oligomerization, from $r=36$ Å for the dimer, to $r=49$ Å for the tetramer, and $r=58$ Å for the hexamer. Also, the distribution widens with increasing degree of oligomerization. The peak width at half-maximum is 45 Å for the dimer,

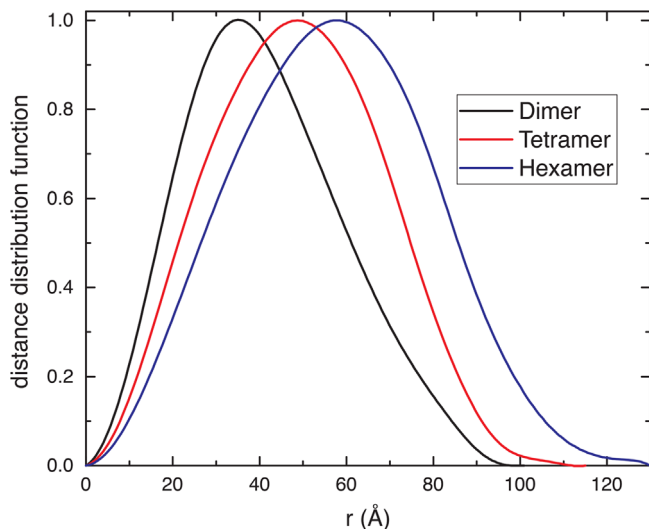


Fig. 3. Distance distribution functions for dimeric (BhP5CDH, black), tetrameric (hALDH7A1, red), and hexameric (TtP5CDH, blue) ALDHs.

53 Å for the tetramer, and 58 Å for the hexamer. D_{\max} is the distance at which the distribution function decays to zero. This value is smallest for the dimer (95–105 Å), intermediate for the tetramer (105–120 Å), and largest for the hexamer (120–125 Å).

In summary, the three oligomeric forms of ALDH are readily distinguishable from SAXS. The data supplied here serves as a benchmark for characterizing ALDH oligomerization.

Acknowledgments

Research reported in this publication was supported by the National Institute of General Medical Sciences of the National Institutes of Health under award numbers R01GM065546 and R01GM093123. Part of the research was performed at the Advanced Light Source, which is supported by the Director, Office of Science, Office of Basic Energy Sciences, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. The Advanced Light Source is a national user facility operated by Lawrence Berkeley National Laboratory on behalf of the Department of Energy, Office of Basic Energy Sciences, through the Integrated Diffraction Analysis Technologies (IDAT) program, supported by DOE Office of Biological and Environmental Research. Additional support comes from the National Institute of Health Project MINOS (R01GM105404).

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at <http://dx.doi.org/10.1016/j.dib.2015.10.017>.

References

- [1] J.J. Tanner, Structural biology of proline catabolism, *Amino Acids* 35 (2008) 719–730.
- [2] T.A. Pemberton, J.J. Tanner, Structural basis of substrate selectivity of Delta(1)-pyrroline-5-carboxylate dehydrogenase (ALDH4A1): semialdehyde chain length, *Arch. Biochem. Biophys.* 538 (2013) 34–40.

- [3] D. Srivastava, R.K. Singh, M.A. Moxley, M.T. Henzl, D.F. Becker, J.J. Tanner, The three-dimensional structural basis of type II hyperprolinemia, *J. Mol. Biol.* 420 (2012) 176–189.
- [4] M. Luo, K.S. Gates, M.T. Henzl, J.J. Tanner, Diethylaminobenzaldehyde is a covalent, irreversible inactivator of ALDH7A1, *ACS Chem. Biol.* 10 (2015) 693–697.
- [5] M. Luo, J.J. Tanner, Structural basis of substrate recognition by aldehyde dehydrogenase 7A1, *Biochemistry* 54 (2015) 5513–5522.
- [6] C. Brocker, N. Lassen, T. Estey, A. Pappa, M. Cantore, V.V. Orlova, T. Chavakis, K.L. Kavanagh, U. Oppermann, V. Vasilou, Aldehyde dehydrogenase 7A1 (ALDH7A1) is a novel enzyme involved in cellular defense against hyperosmotic stress, *J. Biol. Chem.* 285 (2010) 18452–18463.
- [7] J.S. Rodriguez-Zavala, H. Weiner, Structural aspects of aldehyde dehydrogenase that influence dimer-tetramer formation, *Biochemistry* 41 (2002) 8229–8237.
- [8] M. Luo, R.K. Singh, J.J. Tanner, Structural determinants of oligomerization of delta(1)-pyrroline-5-carboxylate dehydrogenase: identification of a hexamerization hot spot, *J. Mol. Biol.* 425 (2013) 3106–3120.
- [9] E. Inagaki, N. Ohshima, H. Takahashi, C. Kuroishi, S. Yokoyama, T.H. Tahirov, Crystal structure of *Thermus thermophilus* Delta1-pyrroline-5-carboxylate dehydrogenase, *J. Mol. Biol.* 362 (2006) 490–501.
- [10] T.A. Pemberton, D. Srivastava, N. Sanyal, M.T. Henzl, D.F. Becker, J.J. Tanner, Structural studies of yeast delta(1)-pyrroline-5-carboxylate dehydrogenase (ALDH4A1): active site flexibility and oligomeric state, *Biochemistry* 53 (2014) 1350–1359.
- [11] J.J. Perry, J.A. Tainer, Developing advanced X-ray scattering methods combined with crystallography and computation, *Methods* 59 (2013) 363–371.
- [12] D.A. Jacques, J. Trewthella, Small-angle scattering for structural biology—expanding the frontier while avoiding the pitfalls, *Protein Sci.* 19 (2010) 642–657.
- [13] G.L. Hura, A.L. Menon, M. Hammel, R.P. Rambo, F.L. Poole 2nd, S.E. Tsutakawa, F.E. Jenney Jr., S. Classen, K.A. Frankel, R. C. Hopkins, S.J. Yang, J.W. Scott, B.D. Dillard, M.W. Adams, J.A. Tainer, Robust, high-throughput solution structural analyses by small angle X-ray scattering (SAXS), *Nat. Methods* 6 (2009) 606–612.
- [14] S. Classen, G.L. Hura, J.M. Holton, R.P. Rambo, I. Rodic, P.J. McGuire, K. Dyer, M. Hammel, G. Meigs, K.A. Frankel, J.A. Tainer, Implementation and performance of SIBYLS: a dual endstation small-angle X-ray scattering and macromolecular crystallography beamline at the advanced light source, *J. Appl. Crystallogr.* 46 (2013) 1–13.
- [15] P.V. Konarev, V.V. Volkov, A.V. Sokolova, M.H.J. Koch, D.I. Svergun, PRIMUS: a Windows PC-based system for small-angle scattering data analysis, *J. Appl. Crystallogr.* 36 (2003) 1277–1282.
- [16] D. Svergun, Determination of the regularization parameter in indirect-transform methods using perceptual criteria, *J. Appl. Crystallogr.* 25 (1992) 495–503.
- [17] C.D. Putnam, M. Hammel, G.L. Hura, J.A. Tainer, X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution, *Q. Rev. Biophys.* 40 (2007) 191–285.
- [18] R.P. Rambo, J.A. Tainer, Accurate assessment of mass, models and resolution by small-angle scattering, *Nature* 496 (2013) 477–481.
- [19] G.J. Kleywegt, Validation of protein models from C α coordinates alone, *J. Mol. Biol.* 273 (1997) 371–376.
- [20] R.P. Rambo, Scatter, (<https://bl1231.als.lbl.gov/scatter/>), 2015.
- [21] D. Schneidman-Duhovny, M. Hammel, A. Sali, FoXS: a web server for rapid computation and fitting of SAXS profiles, *Nucleic Acids Res.* 38 (2010) W540–W544.