ISBN 978-94-007–2529-4

Macromolecular Crystallography

Deciphering the Structure, Function and
Dynamics of Biological Molecules

Edited by
Maria Arménia Carrondo
Paola Spadon

Springer

Chapter 15
Structure Analysis of Biological Macromolecules by Small-Angle X-ray Scattering

Dmitri I. Svergun

Abstract Small-angle X-ray scattering (SAXS) is a low resolution (1–2 nm) structural method, which is applicable to macromolecules in solution providing information about the overall structure and structural transitions. The method covers an extremely broad range of sizes (from a few kDa to hundreds MDa) and experimental conditions (temperature, pH, salinity, ligand addition *etc.*). Recent progress in instrumentation and novel data analysis methods significantly enhanced resolution and reliability of structural models provided by the technique and made SAXS a useful complementary tool to high resolution methods. In particular, SAXS allows for rapid validation of high resolution crystallographic or theoretically predicted models, identification of biologically active oligomers and visualization of missing fragments. Quaternary structure of complexes can be analyzed by rigid body movements/rotations of high resolution models of the individual subunits of domains. The basics of SAXS will be presented and illustrated by advanced applications to macromolecular solutions.

Keywords Solution structure • *Ab initio* methods • Rigid body modeling • Oligomeric mixtures • Flexible proteins

15.1 Introduction

The structural genomics initiatives aiming at large-scale expression and purification for subsequent structure determination using X-ray crystallography (MX) and NMR spectroscopy [10, 12] have already yielded unprecedented numbers of high

D.I. Svergun (✉)
European Molecular Biology Laboratory, Hamburg Outstation, Notkestraße 85,
D-22603 Hamburg, Germany
e-mail: Svergun@EMBL-Hamburg.DE

resolution models for isolated proteins and/or their domains These numbers are expected to grow rapidly in the coming years [24]. The focus of modern structural biology has largely shifted towards the study of macromolecular machines accomplishing most important cellular functions [1]. The macromolecular complexes are usually too large for the structural NMR studies, and they often possess inherent structural flexibility making them difficult to crystallize.

The structural analysis approach to macromolecular complexes includes new crystallographic initiatives complemented by the use of methods yielding structural information in solution at lower resolution. In particular, Cryo-EM allows one to obtain excellent results in many cases [39], but it is usually limited to relatively large macromolecular aggregates (starting from 200–300 kDa).

Small-angle X-ray scattering (SAXS) is a rapid method to characterize low resolution structures of individual macromolecules and complexes in solution and to analyse structural changes in response to variation of external conditions. For establishing the three-dimensional structural models this technique needs monodisperse solutions of purified macromolecules but does not require special sample treatment (growth of crystals, isotopic labelling, cryo-freezing *etc.*). SAXS is applicable to a broad range of conditions and sizes (from a few kDa to hundreds MDa). Unlike most other structural methods, SAXS is able to quantitatively characterize equilibrium and non-equilibrium mixtures and monitor kinetic processes such as (dis)assembly and (un)folding.

Recently, the power of SAXS has been boosted by the significant improvements in instrumentation (most notably, by the high brilliance synchrotron radiation sources) accompanied by the development of novel data analysis methods. These developments made it possible to significantly improve resolution and reliability of the structural models constructed from the SAXS data. Here, the main aspects of SAXS including data processing and interpretation procedures and some applications will be briefly reviewed.

15.2   Basics of a SAXS Experiment

This section will briefly describe the basic theoretical and experimental aspects of SAXS to understand the main principles of the technique as applied to solutions of biological macromolecules. The reader is referred to textbooks [11, 14] or to recent reviews [22, 37, 44] for more detailed description information.

Conceptually, a SAXS experiment is rather simple, as illustrated in Fig. 15.1. The samples are exposed to a collimated monochromatic X-ray or neutron beam with the wave vector $k = 2\pi/\lambda$ where λ is the radiation wavelength (Fig. 15.1). The isotropic scattered intensity I is recorded as a function of the momentum transfer $s = 4\pi \sin\theta/\lambda$, where 2θ is the angle between the incident and scattered beam. The scattering from the solvent is measured separately and subtracted to remove the solvent and parasitic background signals.
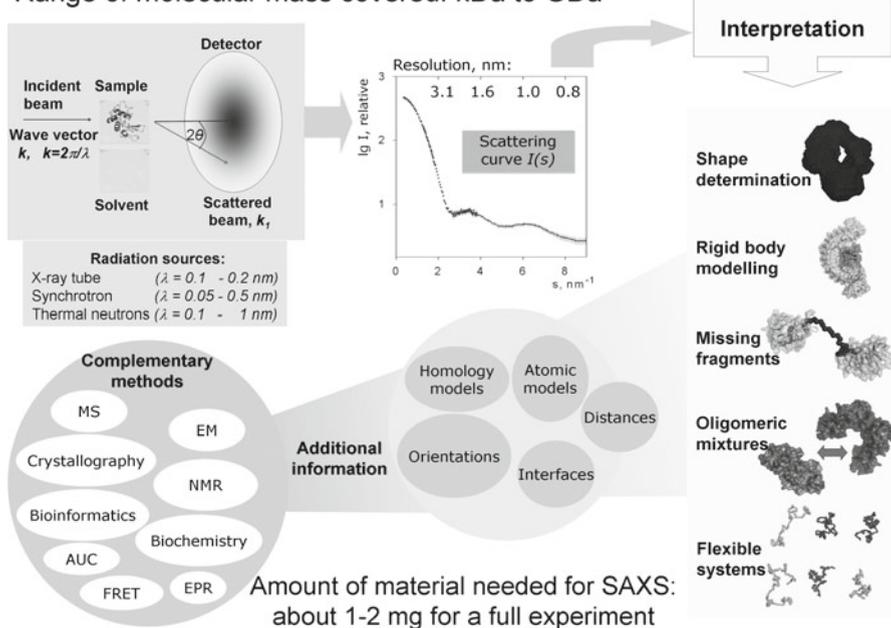
Fig. 15.1 A general scheme of a SAXS/SANS experiment, structural tasks addressed by the technique and its synergistic use with other methods. The nominal resolution of the data in the scattering pattern is indicated as $d = 2\pi/s$. *MS* mass spectroscopy, *AUC* analytical ultracentrifugation, *FRET* fluorescence resonance energy transfer, *EM* electron microscopy, *NMR* nuclear magnetic resonance, *EPR* electron paramagnetic resonance

The SAXS experiments are usually performed at synchrotrons, and the experimental stations offering biological SAXS are available at all synchrotron major sites. Laboratory SAXS cameras, (available from various producers) yield much lower beam brilliance but may be useful at least for preliminary analysis. For structure analysis (shape, quaternary structure), the samples with monodispersity better than 90% are required, which must be verified by other methods (native gel filtration, dynamic light scattering, analytical ultracentrifugation) prior to the synchrotron SAXS experiment. Typical concentrations required are in the range of 0.5–10 mg/ml, and a concentration series is usually measured to get rid of interparticle interference effects. The sample volume per measurement on modern stations is about 10–50 µl so that about 1–2 mg of purified material is usually required for a complete study. The upcoming microfluidic devices [48] will allow one to work on high brilliance sources with nanoliter volumes and µg sample amounts.

One should also mention that neutrons are also employed for small-angle scattering (SANS) analysis of biological macromolecules. SANS (which is performed on research reactors or spallation sources) is sensitive to isotopic H/D exchange. This property is exploited for contrast variation involving measurements in different

H_2O/D_2O mixtures and/or specific perdeuteration of subunits, providing unique information about complex particles [51]. The basic equations and the analysis methods are similar for SAXS and SANS.

15.3   Basics of SAXS Data Analysis

The net SAXS intensity after solvent subtraction contains, generally speaking, two contributions. The so-called form factor *I(s)* emerges from the scattering from individual particles in solution and is employed to extract the structural information. The "structure factor" *S(s)* is due to interference effects between the different particles and yields information about the interparticle interactions (see e.g. [22] for a review).

Purified dilute solutions of macromolecules at concentrations in mM range are usually employed in SAXS to get rid of the interference effects and perform the structural studies assuming that I(s) contains only "form-factor" contrinution. Two important cases are distinguished: (1) monodisperse systems, when all the particles are identical and (2) polydisperse systems, when they are different in size and/or shape.

15.4   Monodisperse Systems

For monodisperse solutions, the net intensity *I(s)* is proportional to the scattering from a single particle averaged over all orientations. This allows one to immediately determine the overall geometrical and weight parameters e.g. radius of gyration R_g [16], volume of the hydrated particle V_p [36], and the molecular mass of the particle MM [27]. The Fourier transformation of the scattering intensity provides a characteristic function (averaged Patterson function), which also yields the maximum particle diameter D_{max} [13, 25, 41], Moreover, the low resolution macromolecular shape can be obtained *ab initio* (i.e. without information from other methods). Several approaches have been proposed [2, 7, 8, 18, 43, 46], and *ab initio* shape determination belongs nowadays to routine analysis of the SAXS data. Usually, the shape analysis programs are ran several times and analysed to obtain the most probable and an averaged model [50].

Calculation of the SAXS profiles from atomic models [45] is used to validate theoretically predicted models and verify the structural similarity between macromolecules in crystals and in solution. Moreover, if high resolution models of individual fragments or subunits in a complex are available from crystallography or NMR, rigid body refinement can be employed to model the quaternary structure of the complex. Automated and semi-automated procedures based on screening randomly or systematically generated models were employed by different authors [19, 21, 28]. A comprehensive rigid body modelling program suite is based on the use of spherical harmonics formalism [31, 32, 40, 42].

SAXS is also very useful for the cases when loops or entire domains are missing in high resolution models (e.g. because of flexibility). The missing portions are
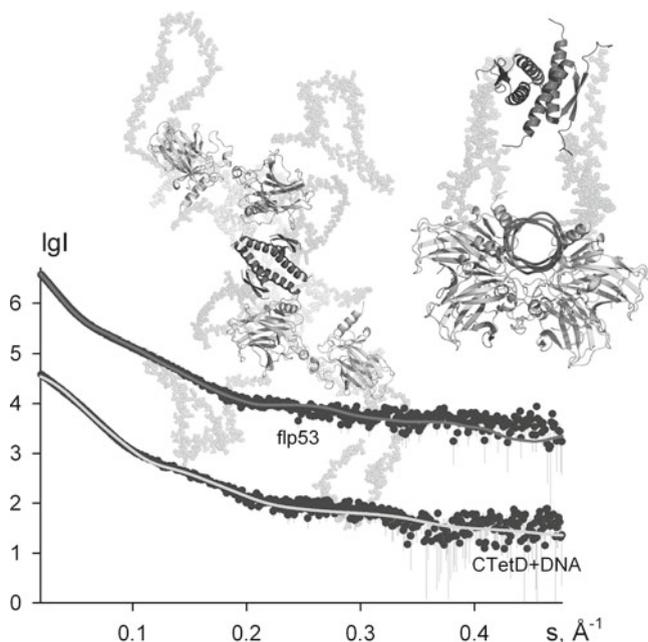
Fig. 15.2 The models of free and DNA-bound tumour suppressor p53 generated by combining SAXS with MX, NMR and EM data. The available high resolution structures of the domains employed in the modelling are displayed as ribbons, the flexible portions of p53 as semi-transparent beads (dummy residues). In the free form (*left panel*) the four domains that recognize DNA (peripheral domains) are arranged in two separate dimers, forming a relatively "flat" structure, with the tetramerization domains in the center. Upon binding to DNA (*right panel*) these core domains wrap around the latter trapping it into a cleft. The models are obtained by rigid body modeling against the SAXS curves displayed as intensity versus momentum transfer in the bottom panel. The data from the free and bound p53 are appropriately displaced along the abscissa axis for better visualization (*dots*, experimental data; *solid lines*, computed patterns from the models)

represented as chains of the so-called dummy residues [34], and the known domains/subunits can be translated and rotated as rigid bodies while simultaneously changing the local conformation of the chains representing the unknown fragments [23]. Numerous applications of rigid body modelling are reported to build structural models of complicated objects in solution (see e.g. references in [33, 37]).

The structural modeling based on the SAXS data is especially effective if one simultaneously incorporates information provided by other methods. An example of such a successful multipronge approach is given by a study of a human tumour suppressor p53 [47], which is a homotetrameric transcription factor (four times 393 residues) playing a central role in the cell cycle. The protein contains a folded core and tetramerization domains, linked and flanked by intrinsically disordered segments. *Ab initio* and rigid body SAXS modelling accounting for NMR-derived interfaces revealed an extended cross-shaped structure with tetrameric contacts and a pair of loosely coupled core domain dimers at the ends (Fig. 15.2). In contrast, the calculated

scattering from a previously published rather compact cryo-EM structure of murine p53 [29] showing dissociated tetramerization domains did not fit the experimental SAXS data. The structure of the complex of p53 with 24 bp DNA independently determined by SAXS and negative stain EM displays a compact complex with the core domains closing around DNA (Fig. 15.2). Interestingly, negatively stained EM analysis of the conformationally mobile, unbound p53 selected a minor compact conformation (less than 20% of the adsorbed particles). The study underlines the significance of the synergistic use of different techniques together with SAXS, in particular, for the structural characterization of a rapidly growing number of proteins with inherent disorder.

15.5   Poyldisperse Systems and Mixtures

For polydisperse systems consisting of different types of non-interacting particles, the measured scattering pattern can be written as a linear combination

$$I(s) = \sum_{k=1}^{K} \nu_k I_k(s) \tag{15.1}$$

where $\nu_k > 0$ and $I_k(s)$ are the volume fraction and the scattering intensity from the k-th type of particle (component), respectively, and K is the number of components.

When neither the number nor intensities of the components are known *a priori*, but multiple data sets are recorded from the system with varying volume fractions of the components, the number of components can be determined extracted by model-independent analysis using singular value decomposition (SVD [15]). If the number of components and their scattering intensities are known, the volume fractions can be readily found by a linear least-squares fit to the experimental data. Numerous applications of these approaches encompass e.g. the analysis of intermediates during folding and assembly processes and quantitative description of oligomeric equilibria [9, 17, 49, 52].

SAXS belongs to very few structural methods able to quantitatively characterize flexible macromolecules, and the method was traditionally used to monitor the processes of protein folding/unfolding [30]. For flexible systems, SAXS data reflect conformational average over the entire ensemble and the scattering patterns are to be interpreted accounting for this average instead of searching for a single model. This has recently become possible with a general approach called 'ensemble optimization method' (EOM) allowing for coexistence of multiple conformations [4]. Given a pool of (random) conformers, EOM selects sub-ensembles from them, which, taken as mixtures, fit the experimental profile using Eq. 15.1. The EOM is already actively used to characterize flexible proteins and complexes [5, 26] and it is expected to find broad applications, in particular, in combination with NMR to provide information about both structure and dynamics of the system [3, 6].

15.6   Conclusions

During the last decade, biological SAXS has become increasingly popular in molecular biology revealing low resolution structures of macromolecule in close to native conditions. SAXS can be readily and usefully combined with other computational and experimental techniques to yield comprehensive description of complex objects and processes. The advanced analysis methods are well established by now and are publicly available e.g. in the program package ATSAS (http://www.embl-hamburg.de/biosaxs/). Automated sample changers and pipelines are being developed for high throughput SAXS on synchrotrons [20, 35, 38]. All these developments taken together make the technique readily available for a broad scope of tasks and a broad community of scientists in structural biology.

References

1. Aloy P, Russell RB (2006) Structural systems biology: modelling protein interactions. Nat Rev Mol Cell Biol 7(3):188–197
2. Bada M, Walther D, Arcangioli B, Doniach S, Delarue M (2000) Solution structural studies and low-resolution model of the Schizosaccharomyces pombe sap1 protein. J Mol Biol 300(3):563–574
3. Bernado P, Blanchard L, Timmins P, Marion D, Ruigrok RW, Blackledge M (2005) A structural model for unfolded proteins from residual dipolar couplings and small-angle x-ray scattering. Proc Natl Acad Sci USA 102(47):17002–17007
4. Bernado P, Mylonas E, Petoukhov MV, Blackledge M, Svergun DI (2007) Structural characterization of flexible proteins using small-angle X-ray scattering. J Am Chem Soc 129(17):5656–5664
5. Bertini I, Calderone V, Fragai M, Jaiswal R, Luchinat C, Melikian M, Mylonas E, Svergun DI (2008) Evidence of reciprocal reorientation of the catalytic and hemopexin-like domains of full-length MMP-12. J Am Chem Soc 130(22):7011–7021
6. Blobel J, Bernado P, Svergun DI, Tauler R, Pons M (2009) Low-resolution structures of transient protein-protein complexes using small-angle X-ray scattering. J Am Chem Soc 131(12):4378–4386
7. Chacon P, Moran F, Diaz JF, Pantos E, Andreu JM (1998) Low-resolution structures of proteins in solution retrieved from X-ray scattering with a genetic algorithm. Biophys J 74(6):2760–2775
8. Chacon P, Diaz JF, Moran F, Andreu JM (2000) Reconstruction of protein form with X-ray solution scattering and a genetic algorithm. J Mol Biol 299(5):1289–1302
9. Doniach S (2001) Changes in biomolecular conformation seen by small angle X-ray scattering. Chem Rev 101(6):1763–1778
10. Edwards AM, Arrowsmith CH, Christendat D, Dharamsi A, Friesen JD, Greenblatt JF, Vedadi M (2000) Protein production: feeding the crystallographers and NMR spectroscopists. Nat Struct Biol 7(Suppl):970–972
11. Feigin LA, Svergun DI (1987) Structure analysis by small-angle x-ray and neutron scattering. Plenum Press, New York, pp. xiii, 335
12. Gerstein M, Edwards A, Arrowsmith CH, Montelione GT (2003) Structural genomics: current progress. Science 299(5613):1663
13. Glatter O (1977) A new method for the evaluation of small-angle scattering data. J Appl Crystallogr 10:415–421

14. Glatter O, Kratky O (1982) Small Angle X-ray Scattering. Academic, London, p 515

15. Golub G, Reinsh C (1970) Singular value decomposition and least squares solution. Numerische mathematik 14(5):403–420

16. Guinier A (1939) La diffraction des rayons X aux tres petits angles; application a l'etude de phenomenes ultramicroscopiques. Ann Phys (Paris) 12:161–237

17. Hamiaux C, Perez J, Prange T, Veesler S, Ries-Kautt M, Vachette P (2000) The BPTI decamer observed in acidic pH crystal forms pre-exists as a stable species in solution. J Mol Biol 297(3):697–712

18. Heller WT, Abusamhadneh E, Finley N, Rosevear PR, Trewhella J (2002) The solution structure of a cardiac troponin C-troponin I-troponin T complex shows a somewhat compact troponin C interacting with an extended troponin I-troponin T component. Biochemistry 41(52):15654–15663

19. Heller WT, Finley NL, Dong WJ, Timmins P, Cheung HC, Rosevear PR, Trewhella J (2003) Small-angle neutron scattering with contrast variation reveals spatial relationships between the three subunits in the ternary cardiac troponin complex and the effects of troponin I phosphorylation. Biochemistry 42(25):7790–7800

20. Hura GL, Menon AL, Hammel M, Rambo RP, Poole FL 2nd, Tsutakawa SE, Jenney FE Jr, Classen S, Frankel KA, Hopkins RC, Yang SJ, Scott JW, Dillard BD, Adams MW, Tainer JA (2009) Robust, high-throughput solution structural analyses by small angle X-ray scattering (SAXS). Nat Methods 6(8):606–612

21. King WA, Stone DB, Timmins PA, Narayanan T, von Brasch AA, Mendelson RA, Curmi PM (2005) Solution structure of the chicken skeletal muscle troponin complex via small-angle neutron and X-ray scattering. J Mol Biol 345(4):797–815

22. Koch MH, Vachette P, Svergun DI (2003) Small-angle scattering: a view on the properties, structures and structural changes of biological macromolecules in solution. Q Rev Biophys 36(2):147–227

23. Konarev PV, Petoukhov MV, Volkov VV, Svergun DI (2006) ATSAS 2.1, a program package for small-angle scattering data analysis. J Appl Crystallogr 39:277–286

24. Levitt M (2007) Growth of novel protein structural data. Proc Natl Acad Sci USA 104(9):3183–3188

25. Moore PB (1980) Small-angle scattering: Information content and error analysis. J Appl Crystallogr 13:168–175

26. Mylonas E, Hascher A, Bernado P, Blackledge M, Mandelkow E, Svergun DI (2008) Domain conformation of tau protein studied by solution small-angle X-ray scattering. Biochemistry 47(39):10345–10353

27. Mylonas E, Svergun DI (2007) Accuracy of molecular mass determination of proteins in solution by small-angle X-ray scattering. J Appl Crystallogr 40:s245–s249

28. Nollmann M, He J, Byron O, Stark WM (2004) Solution structure of the Tn3 resolvase-crossover site synaptic complex. Mol Cell 16(1):127–137

29. Okorokov AL, Sherman MB, Plisson C, Grinkevich V, Sigmundsson K, Selivanova G, Milner J, Orlova EV (2006) The structure of p53 tumour suppressor protein reveals the basis for its functional plasticity. EMBO J 25(21):5191–5200

30. Perez J, Vachette P, Russo D, Desmadril M, Durand D (2001) Heat-induced unfolding of neocarzinostatin, a small all-beta protein investigated by small-angle X-ray scattering. J Mol Biol 308(4):721–743

31. Petoukhov MV, Svergun DI (2005) Global rigid body modelling of macromolecular complexes against small-angle scattering data. Biophys J 89(2):1237–1250

32. Petoukhov MV, Svergun DI (2006) Joint use of small-angle X-ray and neutron scattering to study biological macromolecules in solution. Eur Biophys J 35:567–576

33. Petoukhov MV, Svergun DI (2007) Analysis of X-ray and neutron scattering from biomacromolecular solutions. Curr Opin Struct Biol 17(5):562–571

34. Petoukhov MV, Eady NA, Brown KA, Svergun DI (2002) Addition of missing loops and domains to protein models by x-ray solution scattering. Biophys J 83(6):3113–3125

35. Petoukhov MV, Konarev PV, Kikhney AG, Svergun DI (2007) ATSAS 2.1 – towards automated and web-supported small-angle scattering data analysis. J Appl Crystallogr 40(s1):s223–s228
36. Porod G (1982) General theory. In: Glatter O, Kratky O (eds) Small-angle X-ray scattering. Academic, London, pp 17–51
37. Putnam CD, Hammel M, Hura GL, Tainer JA (2007) X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution. Q Rev Biophys 40(3):191–285
38. Round AR, Franke D, Moritz S, Huchler R, Fritsche M, Malthan D, Klaering R, Svergun DI, Roessle M (2008) Automated sample-changing robot for solution scattering experiments at the EMBL Hamburg SAXS station X33. J Appl Crystallogr 41:913–917
39. Sali A, Glaeser R, Earnest T, Baumeister W (2003) From words to literature in structural proteomics. Nature 422(6928):216–225
40. Svergun DI (1991) Mathematical methods in small-angle scattering data analysis. J Appl Crystallogr 24:485–492
41. Svergun DI (1992) Determination of the regularization parameter in indirect-transform methods using perceptual criteria. J Appl Crystallogr 25:495–503
42. Svergun DI (1994) Solution scattering from biopolymers: advanced contrast variation data analysis. Acta Crystallogr A50:391–402
43. Svergun DI (1999) Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. Biophys J 76(6):2879–2886
44. Svergun DI, Koch MHJ (2003) Small angle scattering studies of biological macromolecules in solution. Rep Prog Phys 66:1735–1782
45. Svergun DI, Barberato C, Koch MHJ (1995) CRYSOL – a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. J Appl Crystallogr 28:768–773
46. Svergun DI, Petoukhov MV, Koch MHJ (2001) Determination of domain structure of proteins from X-ray solution scattering. Biophys J 80(6):2946–2953
47. Tidow H, Melero R, Mylonas E, Freund SM, Grossmann JG, Carazo JM, Svergun DI, Valle M, Fersht AR (2007) From the Cover: Quaternary structures of tumor suppressor p53 and a specific p53 DNA complex. Proc Natl Acad Sci USA 104(30):12324–12329
48. Toft KN, Vestergaard B, Nielsen SS, Snakenborg D, Jeppesen MG, Jacobsen JK, Arleth L, Kutter JP (2008) High-throughput Small Angle X-ray Scattering from proteins in solution using a microfluidic front-end. Anal Chem 80(10):3648–3654
49. Vestergaard B, Groenning M, Roessle M, Kastrup JS, van de Weert M, Flink JM, Frokjaer S, Gajhede M, Svergun DI (2007) A helical structural nucleus is the primary elongating unit of insulin amyloid fibrils. PLoS Biol 5(5):e134
50. Volkov VV, Svergun DI (2003) Uniqueness of ab initio shape determination in small angle scattering. J Appl Crystallogr 36:860–864
51. Whitten AE, Trewhella J (2009) Small-angle scattering and neutron contrast variation for studying bio-molecular complexes. Methods Mol Biol 544:307–323
52. Xu XF, Reinle WG, Hannemann F, Konarev PV, Svergun DI, Bernhardt R, Ubbink M (2008) Dynamics in a pure encounter complex of two proteins studied by solution scattering and paramagnetic NMR spectroscopy. J Am Chem Soc 130(20):6395–6403